Testing cluster analysis on combined petrophysical and geochemical data for rock mass classification

Maria C. Kitzig[1,3] Anton Kepic[1,2] Duy T. Kieu[1]

[1]Department of Exploration Geophysics, Curtin University, GPO Box U 1987, Perth, WA 6845, Australia.
[2]Deep Exploration Technologies Cooperative Research Centre, Curtin University, GPO Box U 1987, Perth, WA 6845, Australia.
[3]Corresponding author. Email: m.kitzig@postgrad.curtin.edu.au

Abstract. New drilling, measurement-while-drilling and top-of-hole sensing technologies are being developed to overcome the challenges of exploration for new mineral deposits under deep cover. These methods will provide continuous, near-real time data collection from every drillhole in the future. Consequently, there will be a need for efficient methods of analysing and interpreting this data stream to complement the exploration strategy. We demonstrate the usefulness of cluster analysis for rapid, automated rock mass classification, and the impact of selecting different subsets of the available data on the classification results. Our study shows that only a few measurements are needed to broadly domain the intersected rock mass and highlights the importance of selecting correct input data depending on the purpose of the classification. Our analysis also indicates the potential of identifying textural and rock mechanical properties from petrophysical measurements via cluster analysis.

Introduction

A variety of data is collected in today's minerals exploration and evaluation campaigns. The type of data collected depends on the deposit or commodity plus the current practice of the exploration company. In most cases, samples are first logged visually by trained geologists and then sent to a laboratory for chemical analysis. Currently, there is a lot of variation in what type of analytical and logging data is collected after visual inspection because of associated costs of acquiring such data. What elements are assayed may vary even for different sections of a drillhole depending on the recommendations of the logging geologist. In addition, geophysical downhole logs are still not commonplace and often available only for a few drillholes throughout a target, despite their potential to considerably aid geological interpretation. Looking towards the future of targeting new deposits deep under cover, this current practice will become increasingly impractical due to long turnaround times and the opportunity cost of making uninformed or poor decisions. Fortunately, new logging-while-drilling and top-of-hole sensing technologies are being developed and will become commercially available over the next few years (Hillis et al., 2014). When implemented, these technologies will provide near-real time data, and more importantly, continuous and consistent measurements from every drillhole. This new wealth of readily available data will both enable and require an automated and timely method of interpretation to aid a quick decision making process.

In this study, we present the outcome of an experiment to automatically domain the intersected rock mass from drillhole RD01 at the DET CRC Brukunga Drilling Research and Training Facility. The drill site is located at the historic Brukunga sulphide mine in the Adelaide Hills in South Australia. Our research objectives were:

(a) testing of a distance-based, unsupervised clustering method for classification of geological data;
(b) determining what and how many of the available datasets are necessary to broadly domain the geological units; and
(c) studying the influence of prior data manipulation on the accuracy of the classification.

We used the fuzzy *c*-means (FCM) clustering algorithm mainly because of its ability to deal with imprecise or mixed data (Bezdek et al., 1984). As opposed to 'hard clustering' where a sample either belongs to a specific cluster or does not, fuzzy clustering allows a sample to be a member or part of several 'fuzzy' clusters simultaneously. The degree to which a sample belongs to a particular cluster is defined by its membership degree, a real number between 0 and 1, summing to unity over all clusters. The usefulness and successful application of fuzzy clustering to geoscientific data has been demonstrated in several studies, e.g. Bosch et al. (2013), Dekkers et al. (2014), Hanesch et al. (2001) and Templ et al. (2008).

Methodology

The first objective was to test a distance-based, unsupervised clustering method for rock mass classification. The FCM clustering method groups a dataset into subsets based on their similarities by minimising the following least-squares objective function:

$$J_{fcm}(U, V) = \sum_{j=1}^{n} \sum_{k=1}^{c} u_{jk}^{m} \|z_j - v_k\|_2^2$$

$$\text{subject to} \sum_{k=1}^{c} u_{jk} = 1,$$

(1)

where n is the total number of sample points $z = \{z_1, z_2,..., z_n\}$, c is the number of clusters, m is the weighting exponent ($m \geq 1$), and $V = \{v_1, v_2,..., v_C\}$ are the centre values. $U = \{u_{jk} \in [0, 1]\}$ is the membership matrix whose elements u_{jk} represent the membership degree of the jth data point to the kth cluster. $\| \|_2$ is the Euclidian norm. We solve for u_{jk} to determine the membership degree and use a cut-off value (α) above which a sample is uniquely assigned to one particular cluster.

The FCM algorithm requires the user to specify the number of clusters before analysis. For most applications, where information is available about the local, deposit scale geology, this prior knowledge does not pose a significant problem. In our case, the number of litho-groups that should be identified was based on the visual core log and a quick grouping based on certain elemental ratios that are useful indicators for different rock classes. The lithologies at Brukunga are mainly psammitic to pelitic metasedimentary units that host variable amounts of sulphide mineralisation (Skinner, 1958). These units are intersected by dolerite dykes up to a few metres thick, which can be significantly altered. The dykes are characterised by elevated Ti, Ca and Mg contents as well as low natural gamma counts. The mineralised units are readily separable from the host rock due to elevated Fe and S contents, higher density and lower resistivity. These units can be further subdivided based on their 'grade' (i.e. sulphide mineral content). High grade units are characterised by a higher S/S+Fe+Al ratio and an increase in P-wave velocity due to the abundance of pyrite. The manually defined rock classes are:

Class1 – psammite (host rock), S/S+Fe+Al ≤ 0.13, Ca/Fe+Si < 0.08, Si > 60%;
Class2 – lower grade mineralised, S/S+Fe+Al < 0.20, Ca/Fe+Si < 0.08, Si < 60%;
Class3 – higher grade mineralised, S/S+Fe+Al > 0.20, Ca/Fe+Si < 0.08, Si > 60%; and
Class4 – dolerite, Ca/Fe+Si ≥ 0.08, Ca ≥ 6%.

Since the subdivision into unmineralised host rock, high and low grade mineralised rock and dolerite dykes, separates the rock mass at Brukunga into major geologically and economically useful groups, the number of clusters for FCM analysis was set to four. The weighting exponent m (see Equation 1) controls the 'hardness' of the cluster boundaries and can be varied between 1 (hard) and 30 (very soft, blurred). For most datasets, a value between 1.5 and 3 gives good results (empirical advice from Bezdek et al., 1984). We tested values of 2, 1.8 and 1.6 and found

little difference in performance. When implementing the FCM algorithm in MATLAB, a third input parameter, the cut-off value α, can be specified. Since a data point can belong to more than one cluster simultaneously, α is the degree of membership above which a point is assigned to a particular cluster. If a data point belongs to more than one cluster with a membership value below α, it will not be assigned to any one particular cluster, but is then defined as belonging to more than one cluster to a certain degree. This might be useful in detecting intervals that are a mix of different rock types, or intervals that are somewhat different from the main lithologies, such as highly altered rock. For the purpose of broadly classifying the rocks into major groups, this can add noise or small-scale variations, which are undesirable in this context. We tested values of 0.6 and 0.4, where 0.4 forced most data to belong to one particular cluster, which was preferred.

The second objective was to test different subsets of the available log, and assay data to determine the combinations that are best suited to classify the rock types at Brukunga. Subsets of petrophysical (GP) and geochemical (GC) assay data were first analysed separately and then combined to compare their respective performances. The individual subsets are summarised in Table 1. To show the effect of data preconditioning on the cluster outcome, the FCM algorithm was first run on the raw datasets and repeated on standardised data. The individual datasets were standardised with the Z-score, calculated by subtracting the mean from each data point and dividing each mean result by the standard deviation. This works well with normally distributed data; however, some data has a natural log-normal distribution. In this case, we used a log-transformation first.

To compare the results of the cluster analysis with the manual classification, three different indicators of success were calculated: classification complexity, uncertainty in classification and the match with the manual interpretation. A simple measure of complexity is the number of class changes along the drillhole. However, the manual classification has 40 class changes over an interval of 274 samples (metres), so a value considerably below 40 changes would provide a poor representation of the geology, which is undesirable. In practice, simple models are less of a problem than overly complex solutions with many changes due to flipping back-and-forth between cluster values. The second value is the number of non-uniquely classified samples, or more specifically, those intervals with membership degrees below the cut-off value. The degree of match with the manual interpretation is calculated by counting the number of correctly identified intervals divided by the total number of intervals (or samples). There is no consideration to the degree of matching, whether picking one class over another is a better mismatch; either it is the same class (1) or not (0). Thus, the resulting matching (correlation) coefficient is a robust measure of a successful classification.

Table 1.   **Different subsets used for cluster analysis.**

	Geophysical data – GP	Assay data – GC
GPGC combined	natural gamma (total count), density, P-wave velocity, S-wave velocity, acoustic impedance (Z), Poisson's ratio (v), shear modulus (G), bulk modulus (K), Young's modulus (E), magnetic susceptibility, 16'' resistivity, resistivity, apparent chargeability	SiO_2, Al_2O_3, Fe_2O_3, CaO, MgO, Na_2O, K_2O, TiO_2, Cr_2O_3, LOI, TOT/S, Ti/Al, Al/Si, Hf, Zr, Cu and Tl.
	GP1	**GC1**
GPGC1 combined	natural gamma (total count), density, P-wave velocity, magnetic susceptibility, 16'' resistivity	SiO_2, Al_2O_3, Fe_2O_3, CaO, MgO, Na_2O, K_2O, TiO_2
	GP2	**GC2**
GPGC2 combined	natural gamma (total count), density, magnetic susceptibility	Fe_2O_3, CaO, Ti/Al
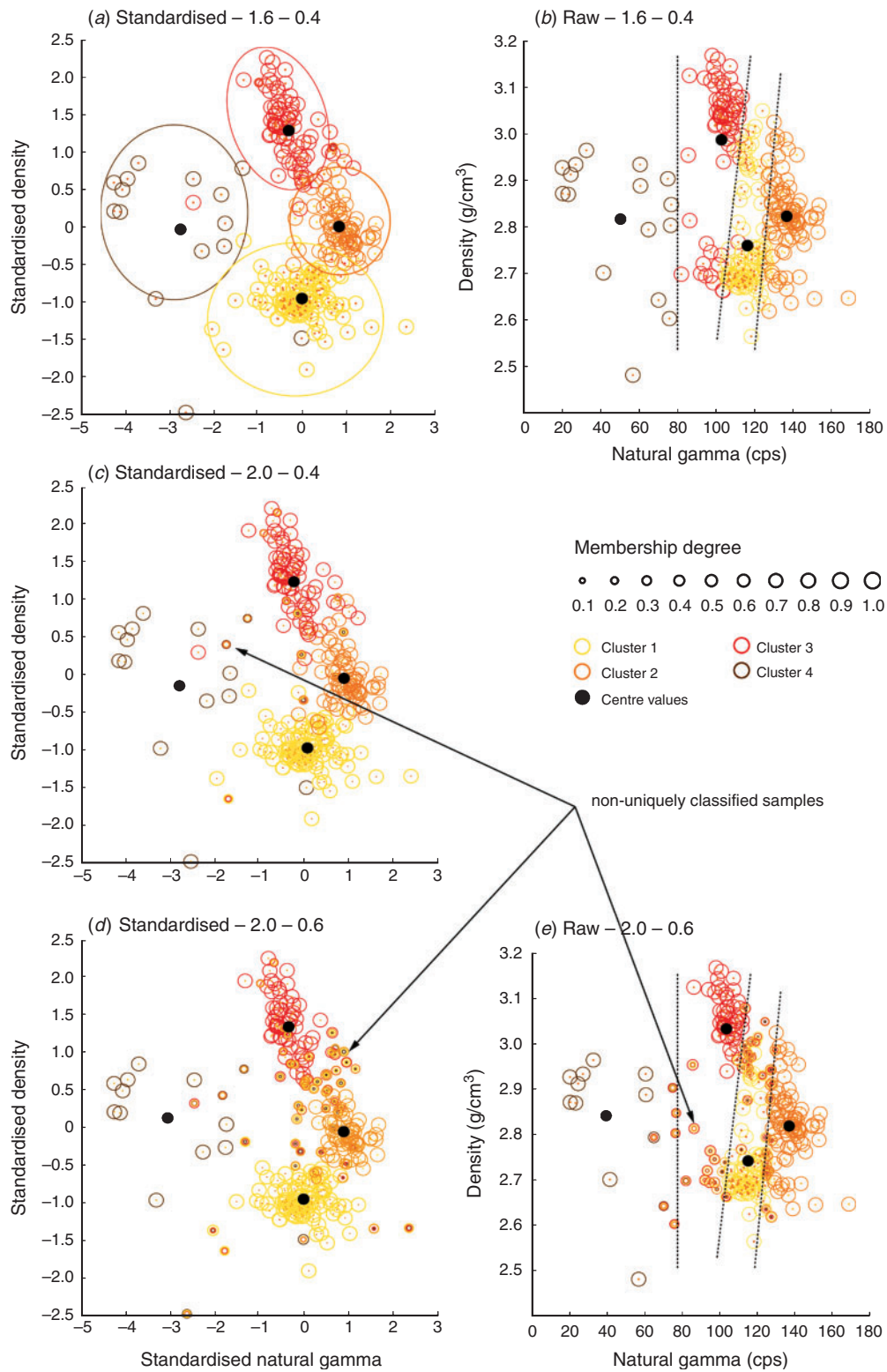
Fig. 1. Results of clustering subset GPGC2 with different FCM parameters. (*a, c, d*) Results for clustering the standardised data: (*a*) α = 0.4, *m* = 1.6 – nearly all data points are uniquely and correctly classified (see Table 2*a*); (*c*) α = 0.4, *m* = 2.0 – a higher value for *m* leads to a larger number of non-uniquely classified samples and more class changes due to 'fuzzier' cluster boundaries (see Table 2*b*); (*d*) α = 0.6, *m* = 2.0 – clustering with a higher value for α reveals samples with membership degrees between 0.4 and 0.6 (smaller multi-coloured circles), predominantly along cluster boundaries where the membership of a sample to one or the other cluster becomes more ambiguous. Those points belong to several clusters with membership degrees indicated by the size and colour of the circle. (*b, e*) Results for clustering the raw data. The trends for changing α and *m* are apparent for the raw data as well (see Table 3*a* versus Table 3*b*). By decreasing α and *m*, all data points could be uniquely classified, but some misclassification is evident when clustering the raw data compared to the standardised data (see Tables 2 and 3). The cluster boundaries seem to follow certain ranges of the natural gamma count (*x*-axis) leading to elongated (banded) cluster shapes.

Results

Initially, we tested a few different parameters of the FCM algorithm to determine the values that are best suited to classify this type of data. A combination of $\alpha = 0.4$ and $m = 1.6$ appears to provide the most robust result (Figure 1*a*). By lowering the cut-off value α, the number of non-uniquely classified intervals was decreased considerably (Figure 1*c* versus Figure 1*d* and Figure 1*b* versus Figure 1*e*). Decreasing the weighting exponent m affects the 'hardness' of the cluster boundaries, and using $m = 1.6$ resulted in a general decrease in the number of unnecessary class changes (Table 2). Figure 1 shows an example of how the different FCM algorithm parameters affect the clustering results.

After determining the best parameters, the algorithm was run both on the standardised and raw data of all subsets. The outcomes are summarised in Tables 2 and 3 respectively. A comparison between the clustering results of the standardised and the raw data of subset GPGC2 with the same parameters is shown

in Figure 1*a* versus Figure 1*b*, Figure 1*d* versus Figure 1*e* and Figure 2*a* versus Figure 2*b*. Our results demonstrate the importance of standardising data before FCM analysis especially when the variables have different units or scales. The clusters of the raw data (Figure 1*b, e*) have elongated shapes, following certain ranges of the natural gamma data. This is due to higher absolute values of the natural gamma compared to the other measurements included in subset GPGC2. The effect is eliminated after standardising, which represent the individual data variables in terms of their mean and standard deviation. Standardising reduced the number of class changes considerably, except for subset GPGC1. An overall good-to-excellent correlation between the standardised data and the manual classification is achieved for all subsets, where the combined subset GPGC2 yielded the highest matching coefficient of 0.94.

Subset GPGC2 includes only a small subset of all available data, and its FCM analysis result can be seen as a 'more supervised' form of clustering for the following reasons.

**Table 2. Cluster correlation result of the standardised data.**

Standardised data – $c = 4$, $m = 1.6$, $\alpha = 0.4$								
Dataset	**GPGC**	**GP**	**GC**	**GPGC1**	**GP1**	**GC1**	**GPGC2**	**GP2**	**GC2**
Class matching	**0.93**	0.74	0.91	0.88	0.89	0.86	**0.94**	0.88	**0.92**
Number of class changes	**35**	45	60	53	**41**	73	**44**	50	54
Non-uniquely classified intervals	1	4	1	2	5	4	2	6	1
Standardised data – $c = 4$, $m = 2.0$, $\alpha = 0.4$								
Class matching	0.92	0.73	0.68	0.86	0.86	0.66	0.90	0.81	0.92
Number of class changes	41	50	76	60	45	86	55	65	55
Non-uniquely classified intervals	6	11	25	7	16	29	12	14	7

Table 3. Cluster correlation results of the raw data.

(*a*) Raw data – $c = 4$, $m = 1.6$, $\alpha = 0.4$								
Dataset	**GPGC**	**GP**	**GC**	**GPGC1**	**GP1**	**GC1**	**GPGC2**	**GP2**	**GC2**
Class matching	0.39	0.45	0.36	0.70	0.70	0.72	0.74	0.64	0.92
Number of class changes	86	86	86	43	43	91	96	119	60
Non-uniquely classified intervals	0	4	2	0	0	0	0	0	0
(*b*) Raw data – $c = 4$, $m = 2.0$, $\alpha = 0.6$								
Class matching	0.43	0.27	0.22	0.66	0.66	0.68	0.66	0.59	0.73
Number of class changes	97	97	90	60	60	95	118	141	87
Non-uniquely classified intervals	36	36	32	29	29	37	54	28	35
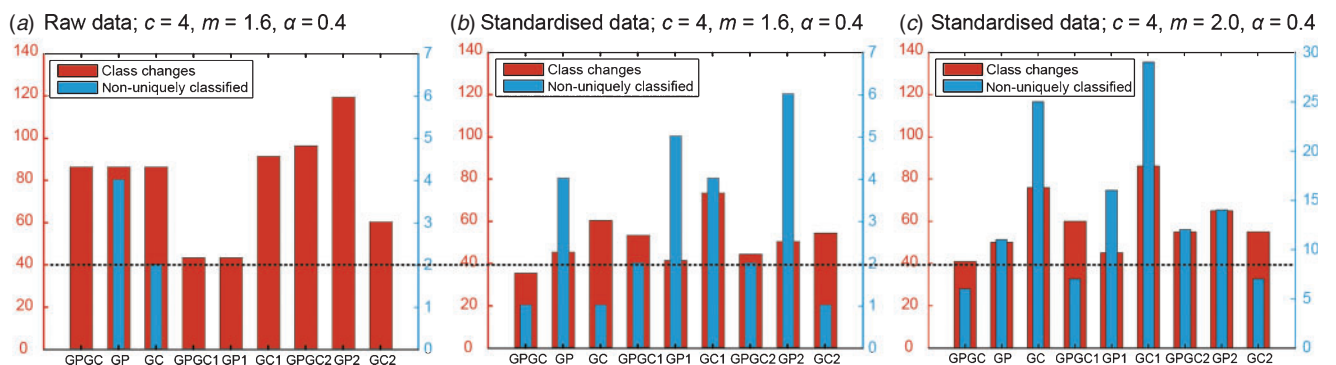
Fig. 2. Histogram plots comparing the results of clustering raw data (*a*) versus standardised data (*b*) in terms of number of class changes and non-uniquely classified intervals. Although there are a few more non-uniquely classified intervals for the standardised data (right hand side *y*-axis in all plots), the number of class changes (left *y*-axis) is reduced considerably for most subsets, leading to a less complex result that better matches the manual classification (dotted line and see Tables 2 and 3). (*b*) and (*c*) show the effect of decreasing the weighting exponent *m* on the number of non-uniquely classified samples (note the different scales of the right hand side *y*-axes; 0–7 in (*b*) and 0–30 in (*c*)).
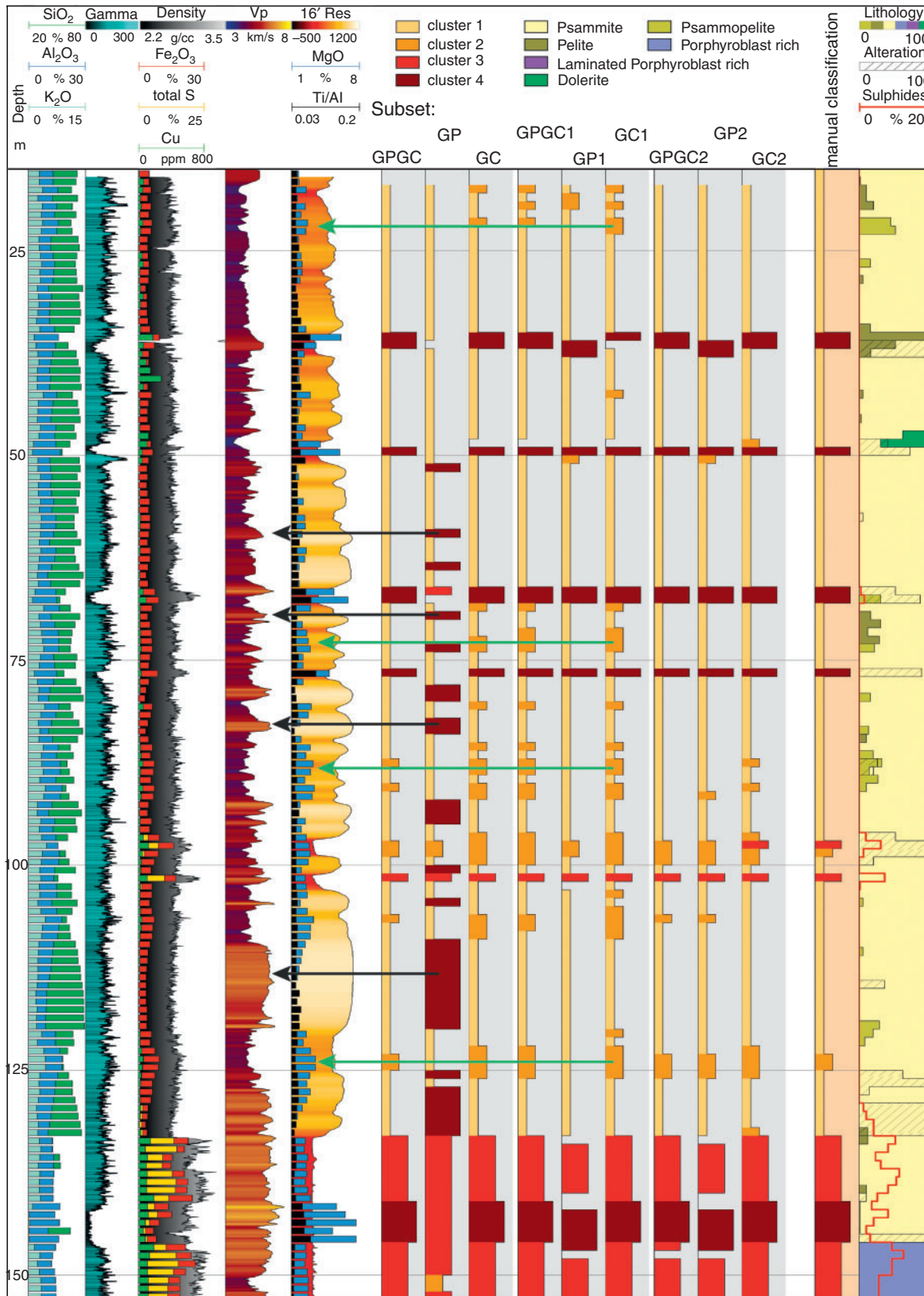
Fig. 3. A log plot (from WellCAD) of the FCM classification results and selected data inputs: geophysical log and chemical assay data (left part), cluster results of the respective subsets of the standardised data and manual logs further to the right. The manual classification log is based on elemental ratios calculated from assay results. 'Lithology', plotted as a percentage log, represents the visual core log of the geologist. 'Alteration' is the visual logging of the thickness of alteration zones in cm per m interval. 'Sulphides' is the visual log of the total amount of sulphides present per interval. Black arrows indicate quarzitic, coarse grained intervals identified by clustering the GP dataset (Cluster 4).These intervals coincide with spikes in P-wave velocity, higher resistivity and higher SiO$_2$ content. Green arrows indicate pelitic intervals identified in subsets dominated by geochemical data (GC, GC1 and GPGC1). These intervals are characterised by lower SiO$_2$ and higher Al, Mg and K contents. The smallest combined subset GPGC2 best matches the manual classification.
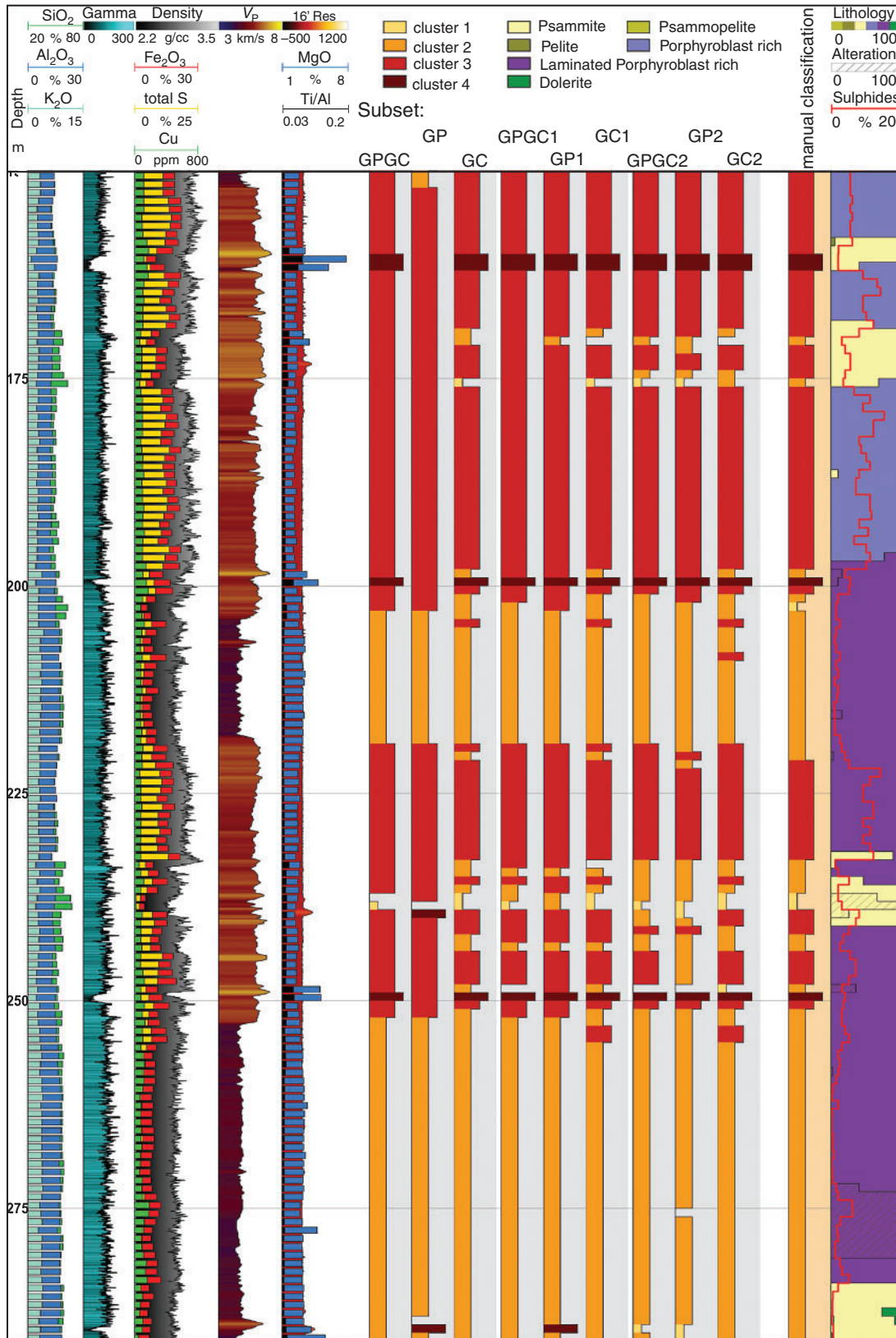
Fig. 3.   (*continued*)

The individual petrophysical measurements and elements for analysis were chosen based on their expected ability to correctly aid classifying the rocks at Brukunga. From the available petrophysical measurements, the natural gamma and magnetic susceptibility logs were also chosen because they are routinely gathered during most exploration and mining logging campaigns. The density log is expected to be useful to separate mineralised units from the host rock and to identify the dolerite dykes to some degree. From the available elemental analysis, Fe was also chosen to separate mineralised units from host rock as

well as separate high and lower grade. The Ca analysis and Ti/Al ratio are indicative of dolerite dykes.

Discussion

Our results above demonstrate that standardising data comprised of variables with different units and scales assist the FCM algorithm greatly. When running the clustering algorithm with identical parameters on the raw and standardised data, the results for the latter achieved better correlation and less class changes in almost all instances (Figures 1 and 2, Tables 2 and 3). Furthermore, our results indicate that a similar, and usually better, classification result was obtained by merging the separate subsets (e.g. GP, GC) into a larger combined dataset (GPGC combined). Less harm and greater good is done for classification by combining petrophysics with elemental analysis than just including more elements from the geochemistry. The overall best matching classification result was obtained from the smallest combined subset (GPGC2) comprised of only six independent measurements. As mentioned above, these measurements were chosen to reflect a selection that a geologist/geophysicist, with prior knowledge of the local geological setting, might choose. This also highlights the fact that not all measurements and analysis that can be obtained are needed to successfully classify the intersected rock mass. However, the 'right' combination of datasets will vary from deposit to deposit and will differ depending on the purpose of the classification (e.g. broad versus fine, lithology versus mineralisation, alteration, etc.). Figure 3 illustrates the results in the form of well logs. The manual classification, the visual core logs ('Lithology', 'Alteration', 'Sulphides'), as well as selected petrophysical measurements and elemental analysis are included for comparison. Only the results from clustering the standardised data are shown. Figure 3 shows that the cluster results correlate well with the manual classification in terms of identifying major boundaries and rock classes. In most cases, the results of clustering the datasets that include elemental analysis only (GC, GC1), are more complex with more variation (class changes), especially in the unmineralised host rock (upper part of the drillhole (15–134 m)). The GP data of the first subset did not identify the dolerites (non-uniquely classified intervals), but added a subdivision within the unmineralised host rock, apparently based on variations of P-wave velocity.

The 'Lithology', 'Alteration' and 'Sulphides' logs are based on visual core logging (Figure 3). The main features logged by the geologist were: percentage of psammite, pelite and psammopelite (metasedimentary host rock), a laminated, porphyroblast rich variant of psammitic host rock associated with sulphide mineralisation (lower grade), a porphyroblast rich variety of psammitic host rock, also associated with sulphide mineralisation (higher grade) and the dolerite dykes. These respective litho-types were logged in terms of percent of their relative occurrence per metre interval. The 'Alteration' log records the thickness of alteration zones in cm per meter interval. The total amount of visually identified sulphides logged (in %) per interval is shown in the 'Sulphides' log. As described above, the 'Lithology' log is mainly based on rock texture, a feature that is not directly reflected in any of the available measurements. Thus, there are some differences between the 'Lithology' log and the manual classification, as well as the cluster results. The visual logging of the abundance of sulphide minerals on the other hand tends to better match the manual classification and our cluster outcome.

When examining the individual logs in Figure 3 further, some interesting links between the cluster outcome and the influences

from the individual petrophysical logs and elemental analysis become more evident. The largest subset (GPGC), which includes almost all available data, yielded the least complex result. The number of class changes is 35, five less than for the manual classification. All major units and boundaries were correctly identified whilst omitting most small-scale variations. This result is therefore most valuable when looking for a broad, first pass interpretation of the intersected rock mass. Subset GP, comprised of all petrophysical measurements and derived parameters, shows more variation within the unmineralised host unit. The dolerites (usually cluster number 4) were not identified as a separate cluster. Instead, cluster number 4 was assigned to a different identified 'subclass' of rock that was not separated out in the other results. This class (Cluster 4, brown intervals in the GP log in Figure 3, indicated by black arrows) seems to be characterised by spikes in P-wave velocity and higher resistivity and coincide with intervals of higher Si content (Figure 4). These intervals are likely to represent quartz rich (quarzitic) beds in the metasedimentary host rock unit. The textural features of these intervals are described as medium to coarse grained and massive with little to no fractures. There is no correlation between increasing P-wave velocity (V_P) and density for these intervals but a correlation between V_P and SiO_2 content and V_P and grain size is evident (Figure 4). Coarser grained textures and higher quartz content can be positively correlated with unconfined compressive strength (UCS) of rocks, which in turn might be correlated with P-wave velocity (Tandon and Gupta, 2013). Subsets GC, GPGC1 and GC1 show quite similar results to each other, displaying a comparatively high frequency of class changes in the upper section. These changes occur at different intervals than the coarse grained intervals described above and seem to correlate with slightly higher Al, Mg and K contents and lower Si content (green arrows in Figure 3). These regions coincide with intervals of lower P-wave velocity and small-scale dips in the resistivity log, and are likely to represent pelite rich intervals of the host rock with opposing trends in petrophysical and geochemical characteristics compared to the quarzitic units. Since the number of clusters was set to equal 4, these pelite rich intervals were grouped within class 2 (laminated variety, low grade sulphides) during the clustering of those subsets. The laminated variety is in fact a meta-pelitic rock that is characterised by a fine grained texture and abundant mica. These textural features and the varying abundance of sulphide minerals throughout the rock facies at Brukunga are believed to be of primary depositional origin (Skinner, 1958). Grouping the smaller pelitic intervals with the laminated variety is therefore geologically reasonable, since both represent a very similar litho-type within the larger stratigraphic unit. The clustering results from the separate analysis of the GP and GC/GC1 subsets can be useful for separating the broader class of psammitic host rock into quarzitic, pelitic and intermediate subclasses.

Subsets GP2, GC2 and combined GPGC2 show less variation within the psammitic unit and a reasonable (GP2) or very good (GPGC2, GC2) correlation with the manual classification. Note that the manual grouping isn't necessarily the 'correct' classification. Correctness is a matter of definition and purpose. However, for our purposes the manual grouping is considered the best outcome for the cluster analysis. As mentioned before, the manual classes were largely based on elemental ratios that reflect the differences in rock composition at the Brukunga site. The same reasoning was applied to choose the input data (measurements) for subset GPGC2. It is therefore not surprising that the cluster results of this subset best match the manual grouping. What our results best demonstrate is that it is not
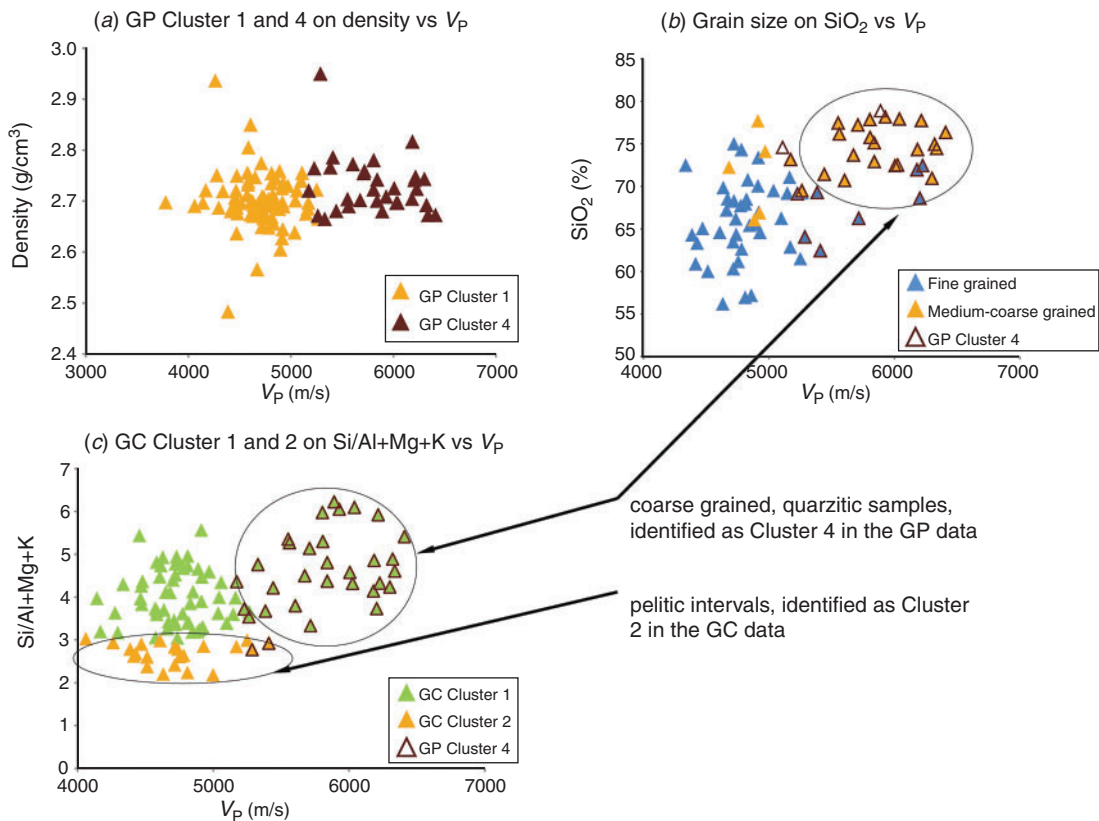
Fig. 4. Intervals from the unmineralised upper part of the drillhole are shown (15–134 m). (*a*) Scatter plot of Clusters 1 and 4 of the GP subset. There is no evident correlation between density and P-wave velocity for these samples. (*b*) Fine and coarse grained samples plotted on V_P versus SiO_2 scatterplot show a correlation between SiO_2 content and V_P. Most coarse grained intervals are correlating with higher velocities and those intervals were identified as Cluster 4 in the GP dataset. (*c*) Clusters 1 and 2 from clustering the GC dataset. Cluster 2 shows lower values for Si/Al+Mg+K and lower velocities indicative for more pelitic intervals. The quarzitic intervals identified by the GP Cluster 4 are plotted for reference.

necessary to collect the largest amount of data to be able to robustly classify the intersected rock mass. The most extensive datasets performed well, but not as well as a trimmed set of petrophysical and geochemical data (GPGC2). For such an approach to succeed, it is important to understand how the measurements and analysis will relate to the rock characteristics. This does not pose a difficulty as there is much information about these relationships. However, an appreciation of both the physical and elemental data influences is needed to make the correct input data selections for FCM analysis to work well.

Conclusions

Our results confirm: (i) the usefulness of FCM clustering as a tool for grouping geological data, and (ii) the importance of data preconditioning. The most important factor to consider when using unsupervised cluster methods is the selection of input data. This selection should be based on the specific purpose of the classification. If little prior knowledge about the rock mass exists, a large input dataset comprised of both petrophysical logs and geochemical assay data would appear to be the best choice to get a first pass, broad grouping of the different domains. Including petrophysical measurements reduces the complexity of the classification, which is a desirable result in the early stages of exploration work where the main objective is to broadly classify and identify prospective intervals. However, if the main characteristics of the rock mass to be found are more specific, then a selection of input data that reflect those characteristics will be sufficient to identify the desired rock classes.

There would appear to be a tendency to neglect petrophysical data if an abundance of elemental data is provided, given the paucity of studies integrating both types of data. Our study shows that this approach excludes a very valuable input to understanding the rock mass. For example, other rock characteristics, like mechanical and textural properties, can be distinguished when petrophysical data is included. A dataset dominated by variables derived from velocity measurements, identified intervals with higher grain sizes and different mineralogy in our study area, due to the effect of these characteristics on rock mechanical properties. Our findings not only highlight the variety of applications of this methodology, but also the potential value of obtaining continuous petrophysical measurements from drillholes.

Acknowledgements

References

Bezdek, J. C., Ehrlich, R., and Full, W., 1984, FCM: the fuzzy c-means clustering algorithm: *Computers & Geosciences*, **10**, 191–203. doi:10.1016/0098-3004(84)90020-7

Bosch, D., Ledo, J., and Queralt, P., 2013, Fuzzy logic determination of lithologies from well log data: application to the KTB project data set

(Germany): *Surveys in Geophysics*, **34**, 413–439. doi:10.1007/s10712-013-9242-2

Dekkers, M. J., Heslop, D., Herrero-Bervera, E., Acton, G., and Krasa, D., 2014, Insights into magmatic processes and hydrothermal alteration of in situ superfast spreading ocean crust at ODP/IODP site 1256 from a cluster analysis of rock magnetic properties: *Geochemistry, Geophysics, Geosystems*, **15**, 3430–3447. doi:10.1002/2014GC005343

Hanesch, M., Scholger, R., and Dekkers, M., 2001, The application of fuzzy c-means cluster analysis and non-linear mapping to a soil data set for the detection of polluted sites: *Physics and Chemistry of the Earth Part A: Solid Earth and Geodesy*, **26**, 885–891. doi:10.1016/S1464-1895(01)00137-5

Hillis, R. R., Giles, D., van der Wielen, S. E., Baensch, A., Cleverly, J. C., Fabris, A., Halley, S. W., Harris, B. D., Hill, S. M., Kanck, P. A., Kepic, A., Soe, S. P., Stewart, G., and Uvarova, Y., 2014, Coiled tube drilling and real-time sensing – enabling prospective drilling in the 21st century: Society of Economic Geologists Special Publication, **18**, 243–259.

Skinner, B. J., 1958, The geology and metamorphism of the Nairne pyritic formation, a sedimentary sulfide deposit in South Australia: *Economic Geology and the Bulletin of the Society of Economic Geologists*, **53**, 546–562. doi:10.2113/gsecongeo.53.5.546

Tandon, R. S., and Gupta, V., 2013, The control of mineral constituents and textural characteristics on the petrophysical & mechanical (PM) properties of different rocks of the Himalaya: *Engineering Geology*, **153**, 125–143. doi:10.1016/j.enggeo.2012.11.005

Templ, M., Filzmoser, P., and Reimann, C., 2008, Cluster analysis applied to regional geochemical data: problems and possibilities: *Applied Geochemistry*, **23**, 2198–2213. doi:10.1016/j.apgeochem.2008.03.004