ELSEVIER

International Journal of Applied Earth Observation and Geoinformation 10 (2008) 305–311

INTERNATIONAL JOURNAL OF APPLIED EARTH OBSERVATION AND GEOINFORMATION

www.elsevier.com/locate/jag

Classification of radio elements using mutual information: A tool for geological mapping

P.T. Krishna Kumar [a,*], V.V. Phoha [c], S.S. Iyengar [b,1]

[a] *Room No: 114, Central Design Office, Indira Gandhi Centre for Atomic Research, Kalpakkam 603102, India*
[b] *Department of Computer Science, Louisiana State University, Baton Rouge, LA 70803, USA*
[c] *Department of Computer Science, Louisiana Tech University, Arizona Avenue, Ruston, La 71270, USA*

Received 16 May 2007; accepted 5 December 2007

Abstract

A broad range of current airborne gamma ray spectrometry (AGRS) applications involve environmental mapping and mineral exploration. One common goal for such applications is the development of an algorithm for reliable on line classification of radio elements. In this paper, we propose the concept of maximization of correlated information as the similarity measure for classification. In order to achieve this similarity measure, we have developed an algorithm using the concept of minimization of mutual information, which is computationally faster, and requiring less memory than the hierarchical agglomerative clustering (HAC) method. The minimization of mutual information is achieved by maximizing the correlated information of the correlation matrix. The correlated information is maximized by the determination of its lower bound using the technique of determinant inequalities developed by us. We demonstrate the robustness of our results using mutual information and its superiority over that of Ward's method of minimum variance for the aerial survey carried out in central India.

Keywords: Airborne gamma ray spectrometry; Mutual information correlation matrix; Upper and lower bounds; Technique of determinant inequalities; Classification of radioactive ores; Ward's method

1. Introduction

Airborne gamma ray spectrometry (AGRS) is widely being used for environmental mapping, geological mapping and mineral exploration. Large areas of the world have been covered by ground and airborne gamma ray surveys and many national and regional radiometric maps have been compiled and published (IAEA, 2003). AGRS uses NaI(Tl) detectors mounted in small aircraft as sensors to measure gamma radiation, emitted from naturally occurring elements, potassium (K), uranium (U) and thorium (Th) that produce gamma rays of sufficient energy and intensity to be measured at airborne survey heights. The concentrations of U, Th and K are inferred from the abundance of gamma-emitting daughter products ^{214}Bi (energy of γ being 1.76 MeV), ^{208}Tl (energy of γ being 2.62 MeV) and ^{40}K (energy of γ being 1.46 MeV), respectively. Different workers have studied the concentrations among the three elements in the same rock types and they have reported a strong correlation in them (Kline et al., 1989; Neuschel et al., 1971).

Among the multivariate analysis used for classification of radioactive ores, several aggregation criteria like distance to mean (Harris, 1989), partitioned algorithms

* Corresponding author. Tel.: +91 44 27480088; fax: +91 44 27480096.
E-mail addresses: gstptk@yahoo.co.in, gstptk@gmail.com (P.T.K. Kumar), phoha@latech.edu (V.V. Phoha), iyengar@csc.lsu.edu (S.S. Iyengar).
[1] Tel.: +1 225 578 1252; fax: +1 225 578 1465.

0303-2434/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.jag.2007.12.001

like K-means (Graham and Bonham-Carter, 1993) have been used to group the individual samples in AGRS. Some of the statistical routines like, mean differencing (Kovarch et al., 1994), regression analysis (Wellman, 1998) and principal component analysis (PCA) have also been used in several AGRS measurements (Gillespie et al., 1986; Chavez and Kwarteng, 1989). But, the sensitivity of the clustering was not sufficient to identify the geologically significant radioactive elements (Harris, 1989) leading to the conclusion that clustering should be used to complement rather than replace visual interpretation (IAEA, 2003). Hierarchical agglomerative clustering (HAC) used in airborne survey (Martelet et al., 2006), consume enormous CPU time and memory complexity limits its practical use (Dash et al., 2007). In this paper, we propose a new on line technique requiring less CPU time and memory to classify and group radioactive ores containing U, Th and K. Our technique based on the theory of information and the principle of entropy preservation is extremely useful in discriminating low noise level data prevalent in AGRS (Dickson, 2004). The information theoretic concept of mutual information has been known as a criterion for feature separation (Petridis and Perantonis, 2004). Due to its computational complexity (Darbellay, 1999) it has mostly been used in an approximate way for feature selection (Battiti, 1994; Kwak and Chong-Ho, 2002). To the best of our knowledge, we have not come across any signal extraction technique used in AGRS using information theory. In this paper we describe the algorithm we have developed based on the technique of determinant inequalities and the concept of minimization of mutual information to classify and group homogeneous zones of radioactive ores. We compare our method of minimization of mutual information with the Ward's method of minimum variance and bring out the superiority of our method.

2. Details of aerial survey

We present here the analysis of AGRS carried out over central India covering around 3000 km. Heterogeneous groups of rocks are locally distributed in the surveyed area as per geo-chemical maps and hence sample interval was chosen to be around 1000 m. The rocks belong to Sausar group and its members are represented by biotite gneisses, feldspathic gneisses, quartz biotite schist, biotite muscovite granite and muscovite quartz schist and are subjected to metamorphism and granitization (Sarkar, 1972).

The gamma ray spectrometer system installed in the aircraft comprised of twelve NaI(Tl) crystals, each of cubical shape. The dimension of each crystal is 10.2 cm × 10.2 cm × 40.6 cm of 50 l capacity of Na(Tl). The energy range of the total gamma count rate window was kept from 0.41 to 2.81 MeV to accommodate all the γ energies emanating from ^{214}Bi, ^{208}Tl and ^{40}K. The survey was done at a nominal height of 122 m and the sampling time ranged from 1 to 3 s. Since, most data are acquired digitally, the following preprocessing have been applied to the observed count rates (IAEA, 2003):

- subtraction of cosmic, aircraft and radon backgrounds (CF_b)
- stripping corrections to remove effects of Compton scattering (CF_s)
- attenuation corrections to remove variations from nominal survey flying height (CF_a)
- After applying the above correction factors, the corrected count rate is converted into equivalent ground concentrations using the sensitivity constant (S_{Th}).

For example in the case of thorium,

$$N_{Th}(\text{corr}) = N_{Th}(\text{obs}) - (CF_b + CF_s + CF_a), \quad \text{and}$$

$$\text{eTh} = S_{Th} N_{Th}(\text{corr})$$

where N_{Th}(corr) is the corrected count rate, N_{Th}(obs) is the observed count, CFs are the various correction factors mentioned above and S_{Th} is the sensitivity constant which relates the count rate to isotope abundance in parts per million. The term equivalent is used for the parent isotopes of uranium and thorium concentrations (eU, eTh) as they are determined indirectly from their daughter products Bi^{214} and Tl^{208}, respectively, under the assumption that the daughter products are in equilibrium with the parent isotopes. The concentration of potassium is determined directly from K^{40}. Thus in our AGRS, the four measured variables are

- total counts,
- eTh (ppm),
- K (%),
- eU (ppm).

The method to generate the correlation matrix from the above four variables has been described in detail in Davis (1986) and will not be repeated here. The interelement correlation matrix for the above four variables has been depicted in Table 1 for the three sets of data, each set covering a distance of 1000 m.

Table 1
Correlation matrix of the radio-elemental data

Set.	Radiometic parameters	Total	eTh	K	eU
1.	Total	1.0	0.9417	0.8761	0.7087
	eTh		1.0	0.8272	0.6267
	K			1.0	0.5336
	eU				1.0
2.	Total	1.0	0.9002	0.7356	0.6527
	eTh		1.0	0.5870	0.4441
	K			1.0	0.1731
	eU				1.0
3.	Total	1.0	0.7751	0.5412	0.7166
	eTh		1.0	0.2210	0.6188
	K			1.0	0.1315
	eU				1.0

3. Entropy preservation in AGRS

Let us designate the radioactive concentration of the rock to be detected as A and let C be the corresponding gamma count rate from the NaI(Tl) detector. The parameter to be optimized is the average mutual information between the input vector A and the output vector C in the presence of noise. The optimization can be formulated as the principle of entropy preservation under certain circumstances dependant on the type of activation function and the noise model. For simplicity, we consider a linear activation function $C = f(A)$ with an additive noise source n.

According to the information theory (Cover and Thomas, 1991) the mutual information I between A and C for the system is expressed as

$$I(A;C) = H(A) + H(C) - H(A,C) \tag{1}$$

where H (A) and H (C) are the entropies of A and C, respectively, and H (A, C) is the joint entropy.

When an airborne spectrometer flies over point source of radiation, the recorded data do not show a single spike, but degrades into a Gaussian curve (Bristow, 1983). Accordingly, we consider, C having Gaussian distribution where its entropy with a fixed variance is maximum (Cover and Thomas, 1991).

The entropy terms in Eq. (1) can be computed as follows (Cover and Thomas, 1991). The entropy of the output vector C is

$$H(C) = \left(\frac{1}{2}\right)\log[(2\pi e)^N \det(M)] \tag{2}$$

where det (M) is the determinant of the covariance matrix $<C\ C^{\mathrm{T}}>$. The joint entropy for a Gaussian distributed noise signal with variance σ_n, for

$$N = 2 \text{ is } H(n) = \log(2\pi e\sigma_n^2), \tag{3}$$

where we have assumed the noise variance to be the same for both the N. The mutual information from Eq. (1), Eq. (2) and Eq. (3) is then

$$I(A;C) = \log\left(\frac{\det(M)}{\sigma_n^2}\right), \tag{4}$$

Thus, the mutual information between the vectors A and C in Eq. (4) depends on the noise variance σ_n and det (M). The mutual information is always positive and cannot be negative (Cover and Thomas, 1991). Minimizing mutual information with a fixed noise variance, thus depends upon the determinant of the covariance matrix M. Minimization of mutual information between different components of the output vector minimizes the redundancy and this phenomenon of redundancy reduction refers to the transformation such that the output components are statistically independent (Barlow et al., 1989; Deco and Brauer, 1995). When A and C are independent, the mutual information is zero. Thus mutual information is a measure of statistical correlation between the variables A and C (Deco and Obradovic, 1996). As an illustration, let us consider a simple case of just two outputs,

$$M = <C\ C^{\mathrm{T}}> = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}$$

The elements of the covariance matrix can be written as the variances of C_1 and C_2,

$$c_{11} = \sigma_1^2 + \sigma_n^2, \quad c_{22} = \sigma_2^2 + \sigma_n^2,$$
$$c_{12} = c_{21} = \sigma_1\sigma_2\rho_{12}$$

where ρ_{12} is the correlation coefficient between C_1 and C_2.

In general, if ρ is the correlation matrix, the elements of ρ are,

$$\rho_{ij} = \frac{\text{covariance}(C_i, C_j)}{\{\text{Var}(C_i)\text{Var}(C_j)\}^{1/2}}.$$

The determinant of the covariance matrix is then

$$\det\ M = \sigma_n^4 + \sigma_n^2(\sigma_1^2 + \sigma_2^2) + \sigma_1^2\sigma_2^2(1 - \rho_{12}^2). \tag{5}$$

In terms of the correlation matrix ρ, Eq. (5) can be expressed as,

$$det\ \rho = \sigma_n^4 + 2\sigma_n^2 + (1 - \rho_{12}^2) \tag{6}$$

Let $G = \det \rho$

All the attempts in the past (Haykin, 1994) for optimizing G, completely neglected the correlated non-diagonal elements of ρ or made the gross assumption that there is no correlation, thereby loosing correlated information. But, for the low noise variance case existing for uranium exploration, third term of Eq. (5) or Eq. (6) is relatively important. An optimal trade off between two options then minimizes the mutual information. In the first option, output variances σ_1^2 and σ_2^2 are kept as small as possible. In the second option, the outputs C_1 and C_2 are made correlated. The value of G in the second option depends upon ρ_{12}. As mutual information is a measure of statistical correlation between the variables A and C, minimizing G by the knowledge of the bounds for the correlated elements of ρ leads to maximization of the linear association or the correlated information. The higher, the correlated information, the more similar are the systems (Massart et al., 1978). Hence, for classification and grouping, we have used in this paper, high value of correlated information as a similarity measure like similarity measure of distances like Euclidean and Mahalonobis (Massart and Kaufman, 1983). Further, maximization of the correlated information leads to Ward's criteria of minimization of variance (Leach and Gillet, 2003; Ward, 1963). Hence, an index of classification and grouping radioactive sources having a low noise variance is the minimization of the mutual information by the estimation of upper and lower bounds for the elements of the correlation matrix ρ.

4. Estimation of upper and lower bounds for correlated elements by the technique of determinant inequalities

In this paper, we describe the technique developed by us (Krishna Kumar, 1990) for obtaining upper and lower bounds for the correlated elements of the correlation matrix thereby minimizing G without loosing correlated information.

Consider a quantity q, which is unknown or is difficult to estimate. A rigorous estimate of it, is provided by the upper and lower bounds, say U and L, respectively, such that $\mathrm{U} \geq q \geq \mathrm{L}$. These rigorous bounds are useful in practice only if they (1) closely bracket q and (2) can be readily computed.

Bulk of the world production of uranium is found in quartz pebble conglomerate and paleo surface type of deposits in equilibrium with its decay products. Thorium is frequently associated with uranium in such deposits (Bristow, 1983). As the principal γ emissions range between 1.46 and 2.81 MeV from all the three elements, the unknown quantity q in our case is the constant radioactivity due to ^{214}Bi of uranium series, ^{208}Tl of thorium series and ^{40}K of potassium recorded in the window of the counting system. This constant activity causes correlation or bias in the measurement of A. This constant bias appears in one or several elements of the determinant G. Let us suppose the sign of the determinant G can be determined. Then G can be considered as a polynomial in q, i.e. $G = G(q)$ and the roots of the determinant function $G(q) = 0$ enable us to estimate the permissible values of q and hence the upper and lower bounds can be determined. Thus to determine the bounds

(a) the sign of the determinant G has to be known and
(b) the roots of the polynomial $G(q) = 0$ should be determined.

The determinant G is positive when $\rho_{ij} = 0$. In this case, only the uncorrelated diagonal elements of ρ exist. Similarly the determinant G is zero when ρ_{ij} is either +1 or −1. Such a determinant is called a Gram determinant or Gramian and its positivity is expressed as an inequality

$$G \geq 0. \tag{7}$$

The upper and lower bounds are determined by solving the polynomial equation $G(q) = 0$.

For the purpose of illustration, let us consider a (3×3) correlation matrix with elements of ρ as follows:

$$\begin{pmatrix} 1 & \rho_{12} & \rho_{13} \\ \rho_{21} & 1 & \rho_{23} \\ \rho_{31} & \rho_{32} & 1 \end{pmatrix}$$

Then $G = \det \rho \geq 0$ requires that

$$1 - \rho_{23}^2 - \rho_{12}^2 + 2\rho_{12}\rho_{13}\rho_{23} - \rho_{13}^2 \geq 0$$

From the above equation, it is clear that ρ_{12} must lie between two roots of the quadratic equation, which constitute the upper, and lower bounds of ρ_{12}.

The upper bound is

$$\rho_{13}\rho_{23} + [(1 - \rho_{13}^2)(1 - \rho_{23}^2)]^{0.5}$$

and the lower bound is

$$\rho_{13}\rho_{23} - [(1 - \rho_{13}^2)(1 - \rho_{23}^2)]^{0.5}$$

In order to process the correlation matrix of larger order so that our technique is competitive to the PCA, we developed an efficient generalized algorithm for

Table 2
Upper and lower bounds for the non-diagonal element ρ_{24} of the correlation matrix of Table 1 and the value of the respective determinant, $G = \det \rho$

Set.	Value of ρ_{ij} from Table 1.	Lower bound of ρ_{ij}	Upper bound of ρ_{ij}	G with ρ_{ij} from Table 1.	G with upper bound of ρ_{ij}.	G with lower bound of ρ_{ij}.	Ward's method
1.	$\rho_{24} = 0.6267$	0.4388	0.84	0.0089	8.7E-05	8.2E-05	0.0415
2.	$\rho_{24} = 0.4441$	0.3824	0.81	0.0077	1.56E-05	7.25E-06	0.0159
3.	$\rho_{24} = 0.6188$	0.2684	0.80	0.0294	5.52E-05	1.23E-06	0.0828

The value of G with the lower bound values in each set is compared with Ward's method.

obtaining the bounds (Krishna Kumar, 1991). Let us designate the determinant as

G_i: with ith row and column deleted,
G_{ij}: with ith and jth row and column deleted,
(note that when G has only two rows and columns then $G_{12} = 1$)
g_{ii}: with $\rho_{ii} = 0$,
g_{ij}: with $\rho_{ij} = 0$, and row j and column i deleted.

According to Eq. (7) $G \geq 0$ and hence G_i and G_{ij} are also Gram determinants of lower order. Thus

$G \geq 0,\ \ G_i \geq 0,\ \ G_{ij} > 0$ and we can establish the following inequalities

$$\rho_{ij} + \left(\frac{g_{ii}}{G_i}\right) \geq 0,$$

$$(g_+ - \rho_{ij})(\rho_{ij} - g_-) \geq 0$$

where $g_\pm = \{(-1)^{i+j}\ g_{ij}{}^{\pm}(G_iG_j)^{0.5}\}/G_{ij}$.

Thus for the uncorrelated component, the lower bound is

$$\rho_{ii} \geq \frac{-g_{ii}}{G_i} \tag{8}$$

while, for correlated component the upper and lower bounds are

$$g_+ \geq \rho_{ij} \geq g_-. \tag{9}$$

5. Robustness of the results obtained by lower bound values

As thorium is always associated with uranium, and since we are interested in uraniferous region, we focused on the correlation element ρ_{24} between U and Th in each of the three sets for which upper and lower bounds were obtained using Eq. (9). As per Table 1, the three correlated elements whose bounds have to be determined are 0.6267, 0.4441 and 0.6188 in the sets 1–3, respectively. The values of the upper and lower bounds for these elements and the value of the determinant with the bounds in each of the above three cases are depicted in Table 2 along with the values obtained by Ward's method.

According to Hadamard's inequality

$$G = \det\ \rho \leq \prod\ \rho_{ij}$$

The equality is achieved if and only if $\rho_{ij} = 0$. The maximum value of the determinant is the product of the diagonal elements which is unity and the minimum value is zero, when ρ_{ij} is either +1 or −1. Since, the mutual information cannot be negative, the least positive value of the determinant is the robust value that minimizes the mutual information.

In our analysis, the least value of the determinant in each of the three sets have been obtained with the lower bound value and these values are very much less than the values obtained by the Ward's method. As these least values signify high correlated information, we have thus provided a unique approach where by we can classify similar type of ore formation by the minimization of the mutual information by the estimation of lower bounds for the correlated elements of the correlation matrix ρ.

6. Critical analysis of usefulness of mutual information

As mutual information depends on the determinant of the correlation matrix, the original correlation matrix, does not yield least minimum value of mutual information and hence the necessity to determine the bounds for the correlated elements.

The similarity measure of classifying the rocks according to the correlation coefficient is apparent from the close scrutiny of the value of the correlation coefficient of the respective bounds. The value of the correlation coefficient of the upper bound for all the three sets is around 0.8 and the determinant of the upper bound values does not yield minimum value of mutual information. Thus we cannot cluster and classify the rocks according to the similar upper bound values of the correlation coefficient.

The correlation coefficient of the lower bound values for all the three sets are different and yield different value of least mutual information, signaling three different radioactive ore formation. Thus, we can classify the rocks according to the correlation coefficient of the lower bound values. The difference between the correlation coefficient of the existing value and the lower bound values is least in the case of set 2, signifying that the region under set 2 is rich in uranium as compared to regions under set 1 and set 3.

Minimizing the mutual information, not only aids in classifying the rocks on line, but also helps in future aerial and detailed ground surveys. Future aerial and ground surveys can be undertaken over those areas whose correlation coefficient lie around the lower bound values. This will not only aid for mapping similar zones of ore formation, but also help for exploratory drilling and mining operations.

7. Reduction of noise in our technique

As, the gamma count rate is used to estimate the abundances of uranium, thorium and potassium in the ground, factors affecting the counting statistics and the associated uncertainty are of paramount importance. The dominant source of noise in an individual spectrum is related to the counting uncertainty. The standard error of measurement per unit distance on the ground is approximately inversely proportional to the square root of the count rate. In order to reduce the standard error we must increase the number of counts, which means increasing the volume of the crystal detector to the maximum possible. But, there are practical limitations on the size to which crystals can be grown and hence on increasing the volume by adding more crystals. As, adding more crystals mean adding more weight, larger and more expensive aircraft are needed to carry them. Thus, given the load capacity of the aircraft that is to be used in the survey, there is clearly a limit on the total crystal volume that can be carried. Hence, these limited number of NaI(Tl) crystals are primarily mobile counters. We have mentioned in Section 2, that to minimize the mutual information, we have to obtain the bounds for the correlated elements of the covariance matrix and also keep the output variances as small as possible. Hence, for a fixed noise variance, minimizing det (M) is achieved by minimizing σ_1^2 and σ_2^2 (second term of Eq. (5)). The technique of determinant inequalities enables us to estimate the lower bound for the variances of the covariance matrix using Eq. (8). Using, this lower bound value as the guide, the counting statistics can be improved upon, so as to minimize the variances of the count rate and hence the mutual information.

Apart from counting uncertainty there are many other sources of noise in the aerial data like, high frequency variations in the K, U and Th concentrations in the ground, airborne radon, variable vegetation and soil moisture, energy drift in radiation detectors, changing the angle of the detector due to aircraft movement. These together with the counting noise constitute total measured noise and it is very difficult to get a measure of the noise distribution (Dickson, 2004). In this context, the entropy-based information theoretic approach, is content free and does not make assumptions about the distribution of the noise data. As we have modeled in Section 2, our information theoretic approach has an additive noise and if we perform gradient ascent on the average mutual information $I(A; C)$, it leads to difference between two entropy terms, i.e.

$$I(A;C) = H(C) - H(n) = \left(\frac{1}{2}\right)[G^{\mathrm{L}} - G^{\mathrm{U}}],$$

where the superscripts refer to two phases of our algorithm, called learning (L) and unlearning (U) (Linsker, 2005). During the L phase, G^{L} is the determinant of the covariance matrix with input signal, input noise and output noise and in the U phase, G^{U} is the determinant of the covariance matrix having only the input and output noise. Thus, in finding the difference between two determinants, the noise contribution is substantially reduced.

8. Conclusions

In AGRS, one has to handle, increasingly massive data sets collected over the span of 1000 m for on line ore classification. The algorithm to handle such massive data should require less CPU time and memory. Our technique of determinant inequalities requires less memory and is conceptually elegant and computationally faster to classify radioactive ores in the aerial survey by the similarity measure of maximizing the correlated information of the correlation matrix. Thus the concept of mutual information considers element wise, the entire structure of the correlation matrix where as PCA and other related methods, the prime objective is only decorrelation of the matrix. Hence, element wise processing is not feasible by HCA, PCA or the factor analysis method. Further our algorithm can process correlation matrices of large dimension and hence errors due to collapsing of these matrices do not arise.

References

Barlow, H., Kaushal, T., Mitchison, G., 1989. Finding minimum entropy codes. Neural Comput. 1, 412–423.

Battiti, R., 1994. Using mutual information for selecting features in supervised neural networking. IEEE Trans. Neural Netw. 5, 537–550.

Bristow, Q., 1983. Airborne γ-ray spectrometry in uranium exploration, principles and current practice. Int. J. Appl. Radiat. Isot. 1, 199–229.

Chavez, P.S., Kwarteng, A.Y., 1989. Extracting spectral contrast in Landsat thematic mapper image data using selective principal component analysis. Photogramm. Eng. Remote Sensing 55, 339–348.

Cover, T., Thomas, J., 1991. Elements of Information Theory. John Wiley and Sons, New York, p. 542.

Darbellay, G.A., 1999. An estimator of the mutual information based on a criterion for independence. Comput. Stat. Data Anal. 32, 1–17.

Dash, M., Petrutiu, S., Scheuermann, P., 2007. pPop: fast yet accurate parallel hierarchical clustering using partitioning. Data Knowl. Eng. 61, 563–578.

Davis, J.C., 1986. Statistics and Data Analysis in Geology. John Wiley & Sons, New York, p. 646.

Deco, G., Brauer, W., 1995. Nonlinear higher-order statistical decorrelation by volume conserving neural architectures. Neural Netw. 8, 525–535.

Deco, G., Obradovic, D., 1996. An Information Theoretic Approach to Neural Computing. Springer, New York, NY, p. 261.

Dickson, B.L., 2004. Recent advances in aerial gamma-ray surveying. J. Environ. Radioact. 76, 225–236.

Gillespie, A.R., Kahle, A.B., Walker, R.E., 1986. Colour enhancement of highly correlated images: decorrelation and HIS contrast stretches. Remote Sensing Environ. 20, 209–235.

Graham, D.F., Bonham-Carter, G.F., 1993. Airborne radiometric data: a tool for reconnaissance geological mapping using GIS. Photogramm. Eng. Remote Sensing 58, 1243–1249.

Harris, J.R., 1989. Clustering of gamma ray spectrometer data using computer image analysis system. In: Agterberg, Bonham Carter, G.F. (Eds.), Statistical Applications in the Earth Science. Geological Survey of Canada, Canada, pp. 19–31.

Haykin, S., 1994. Neural Networks: A Comprehensive Foundation. Prentice Hall, New Jersey, NJ, p. 842.

IAEA, 2003. Guidelines for Radio Element Mapping using Gamma Ray Spectrometry Data. IAEA-TECDOC-1363, Vienna, 173pp.

Kline, S.W., Mose, D.G., Zietz, I., 1989. The use of gamma aeroradioactivity in regional indoor radon prediction: a case study from Fairfax County, Virginia, USA. J. Radiol. Prot. 9, 121–128.

Kovarch, A., Flood, P.G., Tyne, E., 1994. Geographical information system for regional scale geological analysis: the Manila case study. In: Proceedings of the 7th Australian Remote Sensing Conference. pp. 1076–1083.

Krishna Kumar, P.T., 1990. Estimation of bounds for Monte Carlo uncertainty analysis by the technique of determinantal inequalities. Ann. Nucl. Energy 17, 483–486.

Krishna Kumar, P.T., 1991. Algorithm for evaluation of bounds for covariance-based uncertainty analysis. Ann. Nucl. Energy 18, 479–481.

Kwak, N., Chong-Ho, 2002. Input feature selection for classification problems. IEEE Trans. Neural Netw. 13, 143–159.

Leach, A.R., Gillet, V.J., 2003. An Introduction to Chemoinformatics. Kluwer, Netherlands, p. 259.

Linsker, R., 2005. Improved local learning rule for information maximization and related applications. Neural Netw. 18, 261–265.

Martelet, G., Truffert, C., Tourliere, B., Ledru, P., Perrin, J., 2006. Classifying airborne radiometry data with agglomerative hierarchical clustering: a tool for geological mapping in context of rainforest (French Guiana). Int. J. Appl. Earth Obs. 8, 208–223.

Massart, D.L., Kaufman, L., 1983. The Interpretation of Analytical Chemical Data by the Use of Cluster Analysis. John Wiley & Sons, New York, NY, p. 237.

Massart, D.L., Dijkstra, A., Kaufman, L., 1978. Evaluation and Optimization of Laboratory Methods and Analytical Procedures. Elsevier, The Netherlands, p. 596.

Neuschel, S.K., Bunker, C.M., Bush, C.A., 1971. Correlation of uranium, thorium and potassium with aero radioactivity in the Berea area, Virginia. Econ. Geol. 66, 302–308.

Petridis, S., Perantonis, S.J., 2004. On the relation between discriminant analysis and mutual information for supervised liner feature extraction. Pattern Recogn. 37, 857–874.

Sarkar, S.N., 1972. Present status of precambrian geochronology of Peninsular India. In: Proceeding of the 24th International Geology Conference, Montreal, Canada, pp. 260–272.

Ward, J.H., 1963. Hierarchical grouping to optimize an objective function. J. Am. Stat. Assoc. 58, 236–244.

Wellman, P., 1998. Mapping of a granite batholith using geological and remotely sensed data. Explor. Geophys. 29, 643–648.