6

Field sampling scheme optimization using simulated annealing

Pravesh Debba
CSIR Built Environment, Logistics and Quantitative Methods LQM, POBox 395, Pretoria, 0001 South Africa

1. Land characterization: problems in deriving optimal sampling schemes

Land has many components. The various components, such as vegetation, and in the absence of vegetation the rocks and sands with all their minerals make up land cover. To adequately characterize the vegetation components or the mineral components of land, detailed maps describing the spatial distributions of, for example, certain crops or certain minerals are required. The spatial distributions of crops or minerals, however, vary from one place to another according to factors at local settings. Therefore, thorough sampling of land is required to generate detailed maps accurately depicting spatial variability of either crops or minerals and associated metals. Such an undertaking would require money, time, and manpower in order to achieve spatial information of interest at the desired level of accuracy. Therefore, planning *where* and *how many* samples should be collected, in order to map accurately the spatial distributions of either crops or minerals and associated metals, is a non-trivial task.

A sampling plan or scheme refers to positions of samples on the ground. There are two types of sampling schemes, (a) a retrospective scheme, whereby sample locations are either removed from or added to an existing sampling scheme, and (b) a prospective scheme, whereby sample locations are pre-determined before actual sampling in the field. A sampling scheme design is considered optimal if there is (i) a reduction in the number of samples but resulting in estimates of population parameters of interest with the same or similar uncertainty, (ii) a reduction in the variability or mean squared error in estimates of population parameters of interest, (iii) a more correct distribution of samples representing the distribution of the population of interest, or a combination of these criteria. Development of optimal sampling requires *a priori* spatial information about a study area.

Around the mid-20th century and a few decades thereafter, those who studied crops (Driscoll & Coleman, 1974; Everitt et al., 1980; Johnson, 1969) and those who searched for minerals (Allum, 1966; Eardley, 1942; Gilbertson et al., 1976; Laylender, 1956; Longshaw & Gilbertson, 1976) developed their sampling schemes by using geographical information from topographic maps and/or stereoscopic aerial photographs and from visual observations during field reconnaissance surveys. From the 1970s, technological developments in remote sensing resulted in the collection of spaceborne multispectral data, which were to a larger extent useful to derive *a priori* spatial information required in sampling campaigns to study agricultural crops (Everitt et al., 1979; McGraw & Tueller, 1983) but were to a lesser extent useful to derive *a priori* spatial information required in searching for minerals (Houston, 1973; Iranpanah,

1977; Lowman, 1976; Siegal & Abrams, 1976; Siegal & Gillespie, 1980). The reasons for the relative contrast of usefulness spaceborne multispectral data to crop vegetation studies and to search for minerals are that multispectral sensors collect broad wavelength data (a) mostly in the visible to near infrared range of the electromagnetic spectrum, where vegetation has diagnostic spectral features, but (b) partly in the shortwave infrared range of the electromagnetic spectrum, where most minerals have diagnostic spectral features. Multispectral data allow mapping of individual crop species quite accurately (Bouman & Uenk, 1992; Brisco et al., 1989; Richardson et al., 1985), but allow mapping of groups and not individual minerals such as in hydrothermally altered rocks (Abrams, 1984; Carranza & Hale, 2002; Kowalik et al., 1983; Rowan et al., 1977).

From the 1990s, however, advanced technological developments in remote sensing resulted in acquiring airborne hyperspectral data, which are better sources of *a priori* information for those who optimize their respective sampling schemes to study crop vegetation or search for minerals and associated metals. The advantage of hyperspectral data over multispectral data can be attributed to their high spatial resolution and much higher spectral resolutions in the visible to the shortwave infrared regions (Clark, 1999; Polder & van der Heijden, 2001), which allow distinction between plant species (Chang, 2006; Okina et al., 2001; Thenkabail, 2002; Thenkabail et al., 2002) or minerals and associated metals (Cudahy et al., 2000; Martini, 2003; Martini et al., 2003; Papp & Cudahy, 2002). Nevertheless, the ability to process and analyze multi-dimensional hyperspectral data promptly requires improved or novel techniques in order to extract and then further process vital information to derive optimal sampling schemes. The availability of airborne hyperspectral data, therefore, raises two problems in deriving optimal sampling schemes to study crops and to search for minerals and associated metals: (1) how to extract accurate *a priori* information of interest; and (2) how to further process *a priori* information of interest to derive an optimal sampling scheme. The first problem is related to the fact that processing and analysis of hyperspectral data results in only estimates of certain parameters such as (a) vegetation indices, which could reflect crop health (Ausmus & Hilty, 1972; Carter, 1994; Knipling, 1970), and (b) mineral indices, which are estimates of relative abundance of minerals (Chabrillat et al., 1999; Crósta et al., 1998; Resmini et al., 1997; Smith et al., 1985). Accurate estimation of these parameters is undermined by several factors that, for example, distort the spectral signal from materials of interest on the ground to the hyperspectral sensor in the air (Gupta, 2003; Lillesand et al., 1994; Richards, 1993; Sabins, 1996). The second problem is related to the statistical correlation or spatial association between parameters estimated from hyperspectral data and the primary variables of interest, which in this chapter are crops or minerals and associated metals. To investigate potential solutions to these two problems in deriving optimal sampling schemes given hyperspectral data, it is important to first understand hyperspectral remote sensing and optimization of schemes separately and to then merge the disparate knowledge gained. The following two sections provide brief literature reviews on hyperspectral remote sensing and optimization of sampling schemes, respectively.

In this chapter, estimates of parameters of interest derived from hyperspectral data or statistical correlation between parameters estimated from hyperspectral data and the primary variables of interest are here referred to as a model. It is hypothesized that model-based optimal sampling schemes can be derived by (a) improving the precision/accuracy of a model, (b) improving the parameter estimates of a model, (c) reducing the variability of a model, (d) reducing the error of model; or by a combination of any of these aspects. Accordingly, to investigate the hypothesis, the main purpose of this chapter is to use airborne hyperspectral

data to obtain models for input into simulated annealing in order to derive optimal sampling schemes.

2. Hyperspectral remote sensing

In the study of electro-magnetic physics, when energy in the form of light interacts with a material, part of the energy at certain wavelength is absorbed, transmitted, emitted, scattered, or reflected due to the property or characteristics of the material (Sabins, 1996). The three most common ways of measuring the reflectance of a material are by (a) using a hand-held spectrometer over the material in the field or laboratory, (b) using a sensor mounted on an aircraft over a land terrain, or (c) using a sensor mounted on a spacecraft over the earth's surface.

Available hyperspectral data are mostly obtained by aircrafts. Hyperspectral data are reflectance measurements at very narrow wavelengths, approximately 10 nm or less, and are acquired simultaneously over a large spectral range, usually between 0.4 μm and 2.5 μm (Chang, 2006). This spectral range includes the visible, near infrared and short wave infrared regions of the electro-magnetic spectrum, resulting in a large number (often > 100) of contiguous spectral bands or channels. Reflectance data in each spectral channel can be pictorially represented as an image, which is composed of discrete picture elements or pixels. The brightness of a pixel represents the reflective value of materials at specific wavelengths of the electro-magnetic spectrum. Every material has unique spectral features (Hapke, 1993), which are distinct arrays of spectral values at certain regions of the electro-magnetic spectrum. Because hyperspectral sensors acquire spectral data from narrow and contiguous bands of the electro-magnetic spectrum, they provide much better capability to identify materials than broad-band sensors (Sabins, 1999). For example, analysis of changes in narrow absorption features (Van der Meer, 2004), which are usually not recorded by broadband sensors, is a powerful tool in remote identification and estimation of individual materials instead of groups of materials.

A vast amount of scientific knowledge has been and is currently being developed in the field of hyperspectral remote sensing of the environment (Chang, 2006; Gupta, 2003; Sabins, 1996). There are several international peer reviewed journals specifically publishing innovative procedures and advancements on hyperspectral remote sensing of the environment. Integration of hyperspectral data or information derived from hyperspectral data into optimization of sampling schemes has been relatively neglected (Stein et al., 1999).

3. Optimization of sampling schemes

Spatial sampling has been addressed by statisticians for many years. In comparing traditional sampling schemes Burgess et al. (1981) found that a regular grid results in only slightly less precise estimates than a triangular grid, for the same sampling density. They concluded that a small loss of precision or small increase in sampling density to achieve a given precision corresponds with a small increase in price to pay for the practical convenience of regular grids. Christakos & Olea (1992) present a case-specific methodology for choosing between different grid designs.

In optimization of model-based sampling schemes, Spruill & Candela (1990) considered the prediction accuracy of chloride concentration in groundwater by removing or adding locations to an existing sampling network. In a similar way, Royle & Nychka (1998) used a geo-

metrical criterion in order to optimize spatial prediction. Brus & de Gruijter (1997) compared design-based and model-based sampling schemes.
With applications of geostatistical methods, it has been previously shown that for spatially correlated data a triangular configuration of sampling points is most efficient and for isotropic variations the grid should be equilateral (Burgess et al., 1981). McBratney et al. (1981) and McBratney & Webster (1981) presented procedures for optimizing the spacing grid of a regular rectangular or triangular lattice design by maximizing the prediction variance, given an *a priori* variogram. If a variogram, however, shows a relatively high nugget and sampling density is relatively scarce, then a hexagonal grid can be most efficient (Yfantis et al., 1987). By removing or adding locations to an existing sampling network, Ben-Jemaa et al. (1995) used ordinary co-kriging between sediment concentration of mercury and a sediment grain size index to maximize the prediction accuracy. Lloyd & Atkinson (1999) used ordinary kriging and ordinary indicator kriging to optimize a sampling scheme. Diggle & Lophaven (2006) used a Bayesian criterion to optimize geo-spatial prediction by (a) deleting locations from an existing sampling design and (b) choosing positions for a new set of sampling locations. Other studies of variogram application to optimize sampling schemes include Russo (1984), Warrick & Myers (1987), Zimmerman & Homer (1991) and Müller & Zimmerman (1999).
With applications of simulated annealing, Sacks & Schiller (1988) presented several algorithms for optimizing a sampling scheme out of a small grid of possible locations. McGwire et al. (1993) investigated the impact of sampling strategies on the stability of linear calibrations by enforcing various sample distance constraints in a Monte Carlo approach. Van Groenigen & Stein (1998) extended this design by presenting the optimal sampling scheme using spatial simulated annealing that could handle earlier data points and complex barriers. Van Groenigen & Stein (1998) also developed further the Warrick & Myers (1987) criterion to optimize sampling schemes. Van Groenigen et al. (1999) used spatial simulated annealing to construct sampling schemes with minimal kriging variance. They found that anisotropy of the variogram had considerable influence on the optimized sampling scheme, with the highest sampling density in the direction of the highest variability. Van Groenigen et al. (1999) used spatial simulated annealing and the criterion for minimizing the maximum kriging variance in obtaining the optimal sampling scheme. Van Groenigen, Pieters & Stein (2000) showed how conditional probabilities of exceeding environmental threshold values of several contaminants could be pooled into one variable, indicating health risk and thereby used simulated annealing to optimize the sampling scheme. Van Groenigen, Gandah & Bouma (2000) used yield maps to optimize, via spatial simulated annealing, soil sampling for precision agriculture in a low-tech environment. Lark (2002) maximized the likelihood estimation for the Gaussian linear model, which results in designs consisting of fairly regular array supplemented by groups of closely spaced locations.
In sampling for field spectral measurements to support remote sensing, Curran & Atkinson (1998) used co-kriging to define the optimal 'multiple' sampling design, which could be used to simultaneously sample ground and remote sensing data. Tapia et al. (2005) applied a multivariate *k*-means classifier to delineate vegetation patterns from remote sensing data together with the Van Groenigen & Stein (1998) criterion in order to prioritize the survey to areas with high uncertainty. In the current chapter, sampling schemes are optimized based on remote sensing data or remotely sensed information and the application of simulated annealing.

4. Simulated Annealing in context of sampling scheme optimization

Simulated annealing is a general optimization method that has been widely applied to find the global optimum of an objective function when several local optima exist. Details on simulated annealing can be found in Kirkpatrick et al. (1983), Bohachevsky et al. (1986) and Aarts & Korst (1989).

In application of simulated annealing to sampling scheme optimization, a fitness function $\phi(\mathbf{S})$ has to be minimized, depending on the sampling configuration $\mathbf{S}$. Starting with a random sampling scheme $\mathbf{S}_0$, let $\mathbf{S}_i$ and $\mathbf{S}_{i+1}$ represent two solutions with fitness $\phi(\mathbf{S}_i)$ and $\phi(\mathbf{S}_{i+1})$, respectively. Sampling scheme $\mathbf{S}_{i+1}$ is derived from $\mathbf{S}_i$ by randomly replacing one of the points of $\mathbf{S}_i$ by a new point not in $\mathbf{S}_i$. A probabilistic acceptance criterion decides whether $\mathbf{S}_{i+1}$ is accepted or not. This probability $P_c(\mathbf{S}_i \rightarrow \mathbf{S}_{i+1})$ of $\mathbf{S}_{i+1}$ being accepted can be described as:

$$P_{\mathbf{c}}(\mathbf{S}_i \rightarrow \mathbf{S}_{i+1}) = \begin{cases} 1\,, & \text{if } \phi(\mathbf{S}_{i+1}) \leq \phi(\mathbf{S}_i) \\ \exp\left(\dfrac{\phi(\mathbf{S}_i) - \phi(\mathbf{S}_{i+1})}{\mathbf{c}}\right)\,, & \text{if } \phi(\mathbf{S}_{i+1}) > \phi(\mathbf{S}_i) \end{cases} \tag{1}$$

where $\mathbf{c}$ denotes a positive control parameter (usually called the temperature in simulated annealing problems). Several cooling schedules are possible to reduce the temperature. At each value of $\mathbf{c}$, several transitions have to be made before the annealing can proceed, and $\mathbf{c}$ can take its next value. A transition takes place if $\mathbf{S}_{i+1}$ is accepted. Next, a solution $\mathbf{S}_{i+2}$ is derived from $\mathbf{S}_{i+1}$, and the probability $P_c(\mathbf{S}_{i+1} \rightarrow \mathbf{S}_{i+2})$ is calculated according to an acceptance criterion (Equation 1).

Through this and related studies it has been observed that when several local optima exist, as in the case of designing optimal sampling schemes, simulated annealing is superior to gradient based methods.

5. A prospective sampling scheme

Although for this particular case study, a prospective sampling scheme is designed to target a specific mineral, the method is not restrictive to either prospective sampling or in the field of geology (Debba, Carranza, Stein & van der Meer, 2008; Debba et al., 2005). Retrospective sampling schemes can similarly be designed (Debba et al., 2009; Diggle & Lophaven, 2006). Also, these sampling schemes can be applied to various other fields of study, for example, to better estimate certain vegetation parameters (Debba, Stein, van der Meer & Lucieer, 2008).

Hyperspectral imaging systems are useful in identifying individual iron and clay minerals, which can provide details of hydrothermal alteration zoning (Sabins, 1999) based on specific absorption features of these minerals. Thorough discussions on absorption features of hydrothermal alteration minerals can be found in Clark (1999); Hapke (1993); Salisbury et al. (1991); Van der Meer (2004). Various mapping of minerals using hyperspectral data can be found in Crósta et al. (1998); Kruse & Boardman (1997); Rowan et al. (2000); Sabins (1999); Vaughan et al. (2003).

Surface sampling in the field is often advantageous for starting surveys. Identification of hydrothermal alteration minerals like alunite, kaolinite and pyrophyllite, from hyperspectral images leads to a better understanding of the geology and alteration patterns in a region. As such, the analysis of airborne hyperspectral imagery can aid in selecting follow-up targets on the ground before fieldwork is performed. In this study, focus is on the mineral alunite as it is characteristic of hydrothermal alteration zones in the Rodalquilar area in Spain (Arribas et al.,

1995). Alunite has a distinct spectral signature and is often, although not always, related to high sulphidation epithermal gold (Hedenquist et al., 2000). The purpose was to guide field sampling collection to those pixels with the highest likelihood for occurrence of alunite, while representing the overall distribution of alunite. The method offers an objective approach to selecting sampling points in order to, for example, create a mineral alteration map. However, this method can be easily extended to other hydrothermal alteration minerals that have diagnostic absorption features. Combination of several mineral images can then be used in classification of the image to create an alteration map.
The present study aims to use the spectral angle mapper (SAM) and spectral feature fitting (SFF) to classify alunite and obtain rule images. Each pixel in a rule image represents the similarity between the corresponding pixel in the hyperspectral image to a reference spectrum. These rule images are then used to govern sampling to areas with a high probability of alunite occurring and to intensively sample in areas with an abundance of alunite. This effectively delineates favorable areas from unfavorable ones and provides an objective sampling scheme as an initial guideline. The design of the optimal sampling scheme to target these areas of a particular intense hydrothermal alteration mineral is the objective of this study. Such an optimal sampling scheme defies the conventional methods of mineral exploration, which can be time-consuming, cost-prohibitive and involve a high degree of risk in terms of accurate target selection (Srivastav et al., 2000). The study is illustrated with hyperspectral data acquired over the Rodalquilar area.

5.1 Study area

5.1.1 Geology and hydrothermal alteration of the Rodalquilar area

The area of study is located in the Cabo de Gata volcanic field, in the south-eastern corner of Spain (Fig. 1), and consists of calc-alkaline volcanic rocks of the late Tertiary. Volcanic rocks range in composition from pyroxene-bearing andesites to rhyolites. Extensive hydrothermal alteration of the volcanic host rocks has resulted in formation of hydrothermal mineral zones from high to low alteration intensity in the sequence: silica (quartz, chalcedony, and opal) $\rightarrow$ alunite $\rightarrow$ kaolinite $\rightarrow$ illite $\rightarrow$ smectite $\rightarrow$ chlorite. Associated with this mineral alteration are high sulphidation gold deposits and low sulphidation base metal deposits. Gold mineralization is located in the central part of the volcanic field within the Rodalquilar caldera. Arribas et al. (1995) distinguish five hydrothermal alteration zones: silicic, advanced argillic, intermediate argillic and propylitic.
The silicic zone is dominated by the presence of vuggy (porous) quartz, opal and gray and black chalcedony veins. Vuggy quartz (porous quartz) is formed from extreme leaching of the host rock. It hosts high sulphidation gold mineralization and is evidence for a hypogene event. Alteration in the advanced argillic zone is of two types: hypogene and supergene. Alunite, often associated with quartz, kaolinite, illite, jarosite and very rarely pyrophyllite, is the dominant mineral characterizing this zone. The intermediate argillic zone is composed of quartz, kaolinite group minerals, illite, illite-smectite, and minor alunite, diaspore and pyrophyllite. Near the advanced argillic zone, kaolinite is abundant, whereas in the outer zone closer to the propylitic halo illite-smectite becomes the predominant minerals. The propylitic type of alteration is characterized by the presence of chlorite, illite, and smectite. Table 1 presents an overview of the alteration zones and associated alteration minerals. Detection of such minerals is facilitated in the field by hand-held spectrometers.

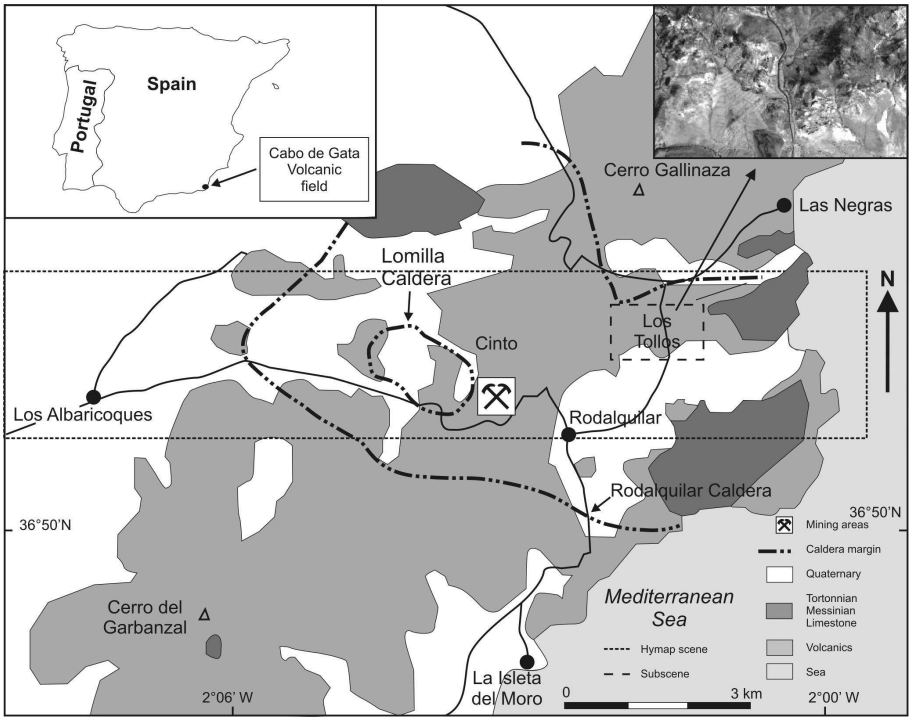


Fig. 1. A generalized geological map (modified after Cunningham et al. (1990)) of the Rodalquilar study area showing the flight line (dotted box) and the hyperspectral data (top right corner and dashed box) used in the present manuscript.

In the Rodalquilar area alunite is associated both with areas of intense hydrothermal alteration that are host to gold mineralization and with barren supergene altered rocks (Arribas et al., 1995; Hedenquist et al., 2000).

5.1.2 Data

We use a sub-scene (350 $\times$ 225 pixels) of the airborne imaging spectrometer data acquired by the Hyperspectral Mapper (HyMAP) in July 2003 during the HyEUROPE 2003 campaign (Fig. 1). HyMAP is a 126-channel instrument that collects data in a cross-track direction by mechanical scanning and in an along-track direction by movement of the airborne platform. The instrument acts as an imaging spectrometer in the reflected solar region of the electromagnetic spectrum (0.4–2.5 μm). Spectral coverage is nearly continuous in the SWIR and VNIR regions with small gaps in the middle of the 1.4 and 1.9 μm atmospheric water bands. The spatial configuration of the instrument accounts for an IFOV of 2.5 mrad along track and 2.0 mrad across track resulting in a pixel size on the order of 3–5 m for the data presented in this chapter. Due to instrument failure the SWIR 1 detector did not function during acquisition, thus no data were acquired in the 1.50–1.76 μm window. The HyMAP data were atmospherically and geometrically corrected using the Atmospheric and Topographic Correction (ATCOR 4) model (Richter, 1996).

Alteration Zone	Alteration Minerals
Silicic	quartz; chalcedony; opal
Advanced Argillic	quartz; alunite; kaolinite; pyrophyllite; illite; illite-smectite
Intermediate Argillic	quartz; kaolinite; illite; illite-smectite
Sericitic	quartz; illite
Propylitic	quartz; illite; montmorillonite
Stage 2 Alunite	alunite; kaolinite; jarosite

Table 1. Summary of alteration zones and dominant minerals in the Rodalquilar area (Arribas et al., 1995).

In support of the imaging spectrometer data, field spectra was collected for some of the prospective targets during the over-flight using the Analytical Spectral Device (ASD) fieldspec-pro spectrometer. This spectrometer covers the 0.35–2.50 μm wavelength range with a spectral resolution of 3 nm at 0.7 μm and 10 nm at 1.4 and 2.1 μm. The spectral sampling interval is 1.4 nm in the 0.35–1.05 μm wavelength range and 2 nm in the 1.0–2.5 μm wavelength range.
The SWIR 2, with a spectral range 1.95–2.48 μm (bandwidth 16 nm), is potentially useful for mapping alteration assemblages as well as regolith characterization (Abrams et al., 1977; Cudahy et al., 2000; Goetz & Srivastava, 1985; Kruse, 2002; Papp & Cudahy, 2002). HyMAP has been used successfully to map minerals (Cudahy et al., 2000; Martini, 2003; Martini et al., 2003; Papp & Cudahy, 2002) and detect faults and fractures (Martini et al., 2003). We reduced dimensionality of the data by considering all channels in the spectral range 1.970–2.468 μm. This spectral range covers the most prominent spectral absorption features of hydroxyl-bearing minerals, sulfates and carbonates, which are common to many geologic units and hydrothermal alteration assemblages (Kruse, 2002). These minerals also exhibit distinctive absorption features at wavelengths in the partly missing range of 1.4–1.7 μm, a range also affected by the water absorption features in the atmosphere.
Fig. 2 shows spectral plots of seven of the most prominent alteration minerals in the study area (Arribas et al., 1995), at a spectral resolution coinciding with HyMAP after continuum removal was applied. Continuum removal normalizes the respective spectra to enable comparison of absorption features from a common baseline. The continuum is a function of the wavelength that is fitted over the top of the spectrum between two local spectra maxima. A straight line segment joins the first and last spectral data values taken as the local maxima (Clark & Roush, 1984; Clark et al., 1991). This figure shows differences in absorption features of the different minerals, in terms of shape, size, symmetry, depth and wavelength position. These distinct characteristics enable researchers to identify individual minerals from hyperspectral data. The spectrum of quartz has no distinctive absorption feature (in this spectral range), but the remaining spectra have distinctive absorption features at wavelengths near 2.2 μm, each differing slightly in position and geometry.
Alunite was chosen among the seven most prominent alteration minerals in the area (Hedenquist et al., 2000) because it has distinct absorption characteristics (Clark, 1999; Hapke, 1993; Salisbury et al., 1991; Van der Meer, 2004), which are recognizable from hyperspectral images (Crósta et al., 1998; Kruse & Boardman, 1997; Rowan et al., 2000; Sabins, 1999; Vaughan et al., 2003). Although this study concentrates on one hydrothermal mineral, namely alunite, the method demonstrated can easily be extended to other minerals of interest. The test image

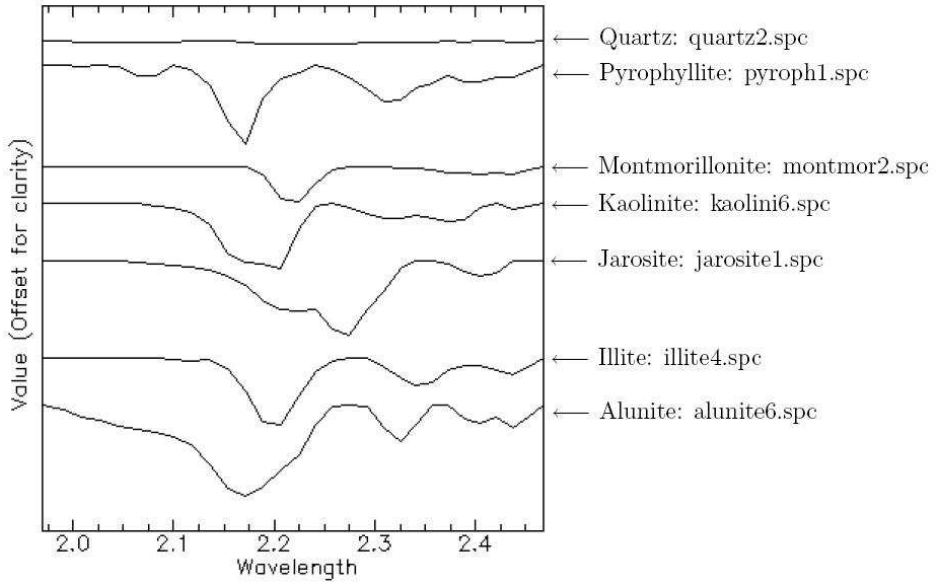


Fig. 2. Plot of 7 endmembers from USGS spectral library (Clark et al., 1993) for the 30 selected channels, enhanced by continuum removal.

selected was in an area that was relatively undisturbed through excavation, hence between 2–3 km from the nearest gold mining area as indicated in Fig. 1.

5.2 Methods

The method for obtaining the optimal sampling scheme commences with application of two classification techniques used, namely, spectral angle mapper (SAM) (Kruse et al., 1993) and spectral feature fitting (SFF) (Clark et al., 1991) to obtain rule images. The digital number (DN) values in a rule image represent similarity between each corresponding pixel's spectrum to a reference mineral spectrum, resulting in one rule image for each mineral considered. Scaled weights are then derived from the rule images. These weights are used in a mathematical objective function (defined in equation 8, see also Van Groenigen, Pieters & Stein (2000)), which is optimized in relation to the spatial distribution of the georeferenced image pixels representing a collection of alunite samples in the field. The aim of optimizing the objective function is to spread the location of the alunite sampling points over the region while targeting pixels that have a high probability of being alunite. In effect, the location of these samples in the field will be dense if distributed in areas with an abundance of alunite and where pixels have a high probability of being alunite. Optimization of the objective function is an exhaustive combinatorial problem. The complexity of the objective function and the iterative process of randomly selecting a pixel in the image as a new sampling point replacing an old one from the collection give rise to many local optima, which is solved through simulated annealing.

5.2.1 Spectral Angle Mapper (SAM) Classifier

SAM is a pixel based supervised classification technique that measures the similarity of an image pixel reflectance spectrum to a reference spectrum from either a spectral library or field spectrum (Kruse et al., 1993). This measure of similarity is the spectral angle (in radians) between the two spectra, where each is an m-dimensional feature vector, with m being the number of spectral channels. Small angles indicate a high similarity between pixel and reference spectra. For an image $\mathbf{I}$, the spectral angle $\theta(\overrightarrow{\mathbf{x}})$, for $\overrightarrow{\mathbf{x}} \in \mathbf{I}$, is given by

$$\theta(\overrightarrow{\mathbf{x}}) = \cos^{-1}\left(\frac{f(\lambda)\cdot e(\lambda)}{||f(\lambda)||\cdot||e(\lambda)||}\right), \tag{2}$$

where λ is the wavelength range of the m spectral channels, $f(\lambda)$ is an unclassified m-dimensional image reflectance spectrum under observation and $e(\lambda)$ is an m-dimensional reference spectrum. SAM is directionally dependent, but independent of the length of the spectral vector, thus insensitive to illumination or albedo effects (Crósta et al., 1998). It is also dependent on the user-specified threshold and wavelength range. The result of using equation 2 are grayscale images (SAM's Rule Images), one for each reference mineral, with DN value representing the angular distance in radians between each pixel spectrum and the reference mineral spectrum (see Fig. 3a). Darker pixels in the rule image indicate greater similarity to the reference mineral spectra. Further, if this angular distance is smaller than a user specified threshold, the pixel is assigned to the category of the respective reference mineral, leading to image classification. This algorithm has been implemented in ENVI^{TM} image analysis commercial software.

5.2.2 Spectral Feature Fitting (SFF)

SFF matches the image pixel reflectance spectrum to reference spectrum from either a spectral library or a field spectrum by examining specific absorption features in the spectrum after continuum removal has been applied to both the image and reference spectrum (Clark et al., 1991). Denote the continuum for the image reflectance spectrum as $c_f(\lambda)$ and for the reference spectrum as $c_e(\lambda)$. The continuum is removed (Clark & Roush, 1984) using

$$\begin{array}{rcl} e_c(\lambda) & = & e(\lambda)/c_e(\lambda) \\ f_c(\lambda) & = & f(\lambda)/c_f(\lambda) \end{array} \tag{3}$$

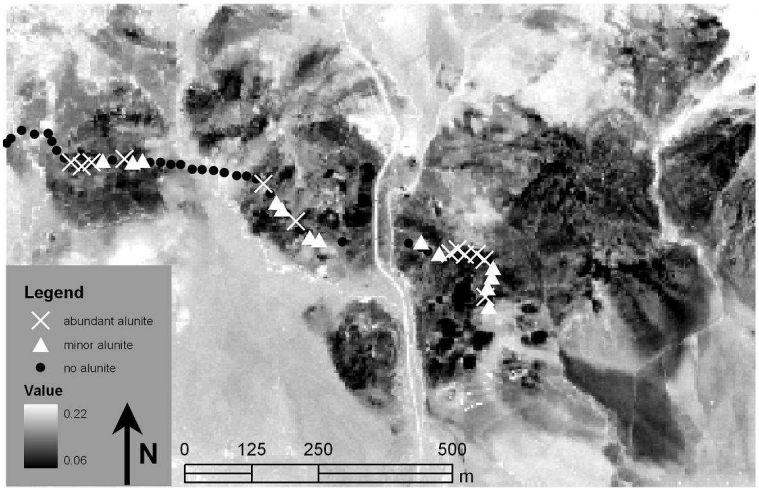
where $e_c(\lambda)$ is the continuum removed reference spectrum and $f_c(\lambda)$ is the continuum removed image reflectance spectrum, respectively. The resulting normalized spectra reflect levels equal to 1.0 if continuum and the spectrum match and less than 1.0 in the case of absorption.

Similarly, the absorption feature depth is defined as

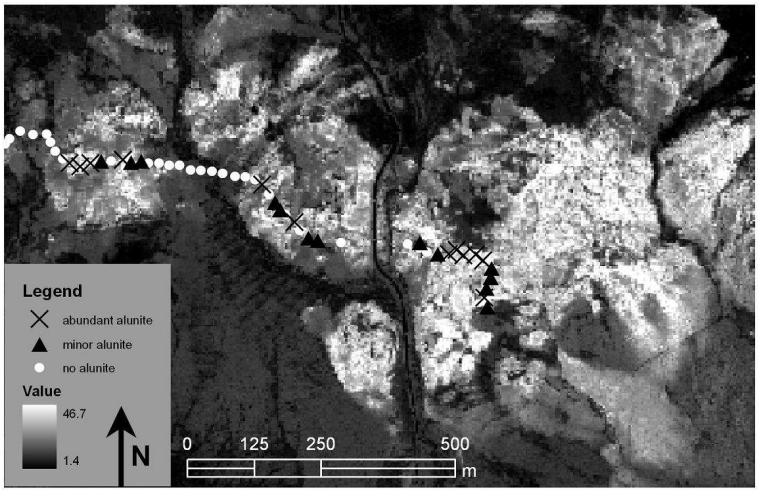
$$\begin{array}{rcl} D[e_c(\lambda)] & = & 1 - e_c(\lambda) = 1 - e(\lambda)/c_e(\lambda) \\ D[f_c(\lambda)] & = & 1 - f_c(\lambda) = 1 - f(\lambda)/c_f(\lambda) \end{array} \tag{4}$$

for each spectrum. The absorption feature depth has a unique magnitude and location, both depending on the mineral and its chemical composition.

Scaling is usually necessary for reference spectra because absorption features in library data typically have greater depth than image reflectance spectra. A simple scaling function of the form $e_c^s(\lambda) = a_0 + a_1 e_c(\lambda)$, where $e_c^s(\lambda)$ is the modified continuum removed reference



(a) SAM classification rule image for alunite. Dark areas indicate smaller angles, hence, greater similarity to alunite. This figure also shows the location of the field data.



(b) SFF fit image for alunite. Lighter areas indicate better fit values between pixel reflectance spectra and the alunite reference spectrum. This figure also shows the location of the field data.

Fig. 3. SAM and SFF (fit) Rule Images.

spectrum that best matches the image spectra, is useful. For an image $\mathbf{I}$, the scale $\tau_S(\overrightarrow{\mathbf{x}})$, for $\overrightarrow{\mathbf{x}} \in \mathbf{I}$, is determined using least squares that gives the best fit to the image spectrum $f_c(\lambda)$

$$D[f_c(\lambda)] = a + \tau_S(\overrightarrow{\mathbf{x}}) D[e_c(\lambda)] \,. \tag{5}$$

Hence the scale image, produced for each reference mineral, is the image of scaling factors used to fit the unknown image spectra to the reference spectra. The result is a grayscale scale image, whose DN value corresponds to $\tau_S(\overrightarrow{\mathbf{x}})$.
The total root-mean-squares (RMS) errors, $\tau_E(\overrightarrow{\mathbf{x}})$, was defined as

$$\tau_E(\overrightarrow{\mathbf{x}}) = \sqrt{\frac{1}{m}\sum_b (D[f_c(\lambda_b)] - D[e_c^s(\lambda_b)])^2} \tag{6}$$

where λ_b denotes the wavelength of channel b, $b = 1, \ldots, m$. The result is a grayscale RMS error image, with DN value corresponding to $\tau_E(\overrightarrow{\mathbf{x}})$.
The fit image equals

$$\tau_F(\overrightarrow{\mathbf{x}}) = \tau_S(\overrightarrow{\mathbf{x}})/\tau_E(\overrightarrow{\mathbf{x}}) \tag{7}$$

providing a measure of how well image pixel reflectance spectra match reference spectra. A large value of $\tau_F(\overrightarrow{\mathbf{x}})$ corresponds to a good match between the image spectrum and the reference spectrum. The fit values are used as a rule image to weigh each pixel to a reference mineral, namely alunite (see Fig. 3b). This algorithm has been implemented in ENVI^{TM} image analysis commercial software. Further details on SFF can be found in (Clark & Swayze, 1995; Clark et al., 1992; 1991; 2003).

5.2.3 Sampling

Sampling by simulation annealing requires definition of a mathematical objective function, called the fitness function.

Fitness function

The Weighted Means Shortest Distance (WMSD)-criterion is a weighted version of the Minimization of the Mean Shortest Distances (MMSD)-criterion (Van Groenigen, Pieters & Stein, 2000). The fitness function is extended with a location dependent weight function that is scaled to $[0, 1]$, namely, $w(\overrightarrow{\mathbf{x}}) : \mathbf{I} \rightarrow [0, 1]$ by

$$\phi_{\mathrm{WMSD}}(\mathbf{S}^n) = \frac{1}{N}\sum_{\overrightarrow{\mathbf{x}}\in\mathbf{I}} w(\overrightarrow{\mathbf{x}}) \left|\left| \overrightarrow{\mathbf{x}} - W_{\mathbf{S}^n}(\overrightarrow{\mathbf{x}}) \right|\right| , \tag{8}$$

where $W_{\mathbf{S}^n}(\overrightarrow{\mathbf{x}})$ is the location vector of the sampling point in $\mathbf{S}^n$ nearest to $\overrightarrow{\mathbf{x}}$, N is the number of pixels in the image and $w(\overrightarrow{\mathbf{x}})$ is a weight for the pixel with location vector $\overrightarrow{\mathbf{x}}$. The weights express knowledge or assumptions about the occurrence of alunite in some parts of the region by controlling the sampling density in these areas. Larger weights result in a higher likelihood of a pixel being selected in the final sampling scheme.
This fitness function also spreads the location of the sampling points over the region classified as alunite. Since these points on the image are georeferenced, they will appropriately serve as target points to be sampled in the field. There will be a high probability that the field sample points suggested are alunite and these points will be spread according to the distribution of alunite as in the classified image. This achieves the purpose of the study of obtaining a collection of sampling points in the field that appropriately represent the distribution of the mineral of interest. A weight function is defined below to meet this objective.
For the weight function, scaled weights are used based on several rule images to guide sampling to areas with a high probability of being alunite and to sample more intensely where

an abundance of alunite occurs. Using SAM's rule image and SFF's rule image, derived by applying equations 2 & 7, thresholds θ^t and τ_F^t are selected for SAM and SFF respectively. Pixels exceeding either of these threshold angles receive zero weight, otherwise the weight is a function of the spectral angle and the fit value. Higher weights will emerge from smaller spectral angle between the image pixel reflectance spectrum and reference alunite spectrum, and a larger fit value between these two spectra. The weight $w(\overrightarrow{\mathbf{x}})$, for each pixel $\overrightarrow{\mathbf{x}}$, scaled to $[0,1]$ is defined as

$$w(\theta(\overrightarrow{\mathbf{x}}),\tau_F(\overrightarrow{\mathbf{x}})) = \begin{cases} \kappa_1 w_1(\theta(\overrightarrow{\mathbf{x}})) + \kappa_2 w_2(\tau_F(\overrightarrow{\mathbf{x}})), \\ \qquad \text{if } \theta(\overrightarrow{\mathbf{x}}) \leq \theta^t \text{ and } \tau_F(\overrightarrow{\mathbf{x}}) \geq \tau_F^t \\ 0, \quad \text{if otherwise} \end{cases} \tag{9}$$

where $0 \leq \kappa_1, \kappa_2 \leq 1$ and $\kappa_1 + \kappa_2 = 1$. The weight for SAM: $w_1(\overrightarrow{\mathbf{x}})$, for each pixel $\overrightarrow{\mathbf{x}}$, scaled to $[0,1]$ is defined as

$$w_1(\theta(\overrightarrow{\mathbf{x}})) = \begin{cases} 0, & \text{if} \quad \theta(\overrightarrow{\mathbf{x}}) > \theta^t \\ \dfrac{\theta^t - \theta(\overrightarrow{\mathbf{x}})}{\theta^t - \theta_{\min}}, & \text{if} \quad \theta(\overrightarrow{\mathbf{x}}) \leq \theta^t \end{cases} \tag{10}$$

and the weight for SFF: $w_2(\overrightarrow{\mathbf{x}})$, for each pixel $\overrightarrow{\mathbf{x}}$, scaled to $[0,1]$ is defined as

$$w_2(\tau_F(\overrightarrow{\mathbf{x}})) = \begin{cases} 0, & \text{if} \quad \tau_F(\overrightarrow{\mathbf{x}}) < \tau_F^t \\ \dfrac{\tau_F(\overrightarrow{\mathbf{x}}) - \tau_F^t}{\tau_{F,\max} - \tau_F^t}, & \text{if} \quad \tau_F(\overrightarrow{\mathbf{x}}) \geq \tau_F^t \end{cases} \tag{11}$$

where θ^t is the maximum angle threshold value chosen, $\theta_{\min}$ the minimum spectral angle occurring, τ_F^t is the minimum fit threshold value chosen and $\tau_{F,\max}$ the maximum value.
The weight function if used in the fitness function will be restricted to those pixels with a spectral angle smaller than the threshold chosen and with a fit larger than the chosen threshold. The probability is largest to select a pixel that is most similar to the alunite reference spectrum, in terms of both the angle between these spectra and absorption feature fit. The georeferenced location of each pixel chosen by the algorithm in the final sampling scheme will be a point to be sampled on the ground.
This weight function (equation 9), is based on two rule images. This can easily be extended to more than two rule images, by using different proportions κ_i for each rule image i conditional on $\sum \kappa_i = 1$. Also, in terms of the method of SFF, several absorption features could be considered for a particular mineral, producing a fit image for each feature. These images could be combined in the same way, thereby increasing the weights of image pixels having a spectrum similar to the mineral. This in effect increases the probability of the mineral being selected in the sampling scheme.

5.3 Results

Forty samples were arbitrarily chosen to illustrate the distribution of these points for the proposed sampling scheme. Prior to sampling, isolated segments (< 10 pixels) were removed. This was performed as there was a high chance that they were a result of noise in the image and it seemed impractical to sample in such small areas. However, if these are meaningful targets, with very high probability of alunite, the above procedure can be performed without removal of these pixels.

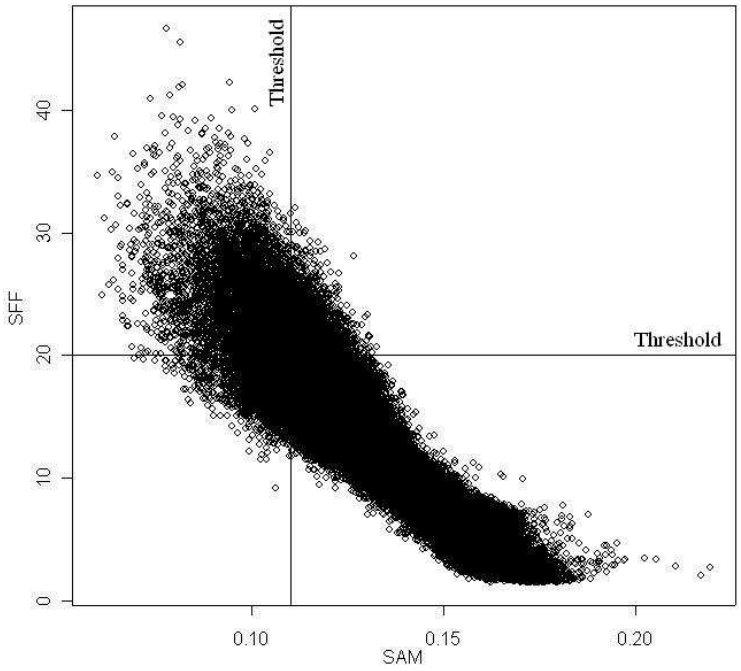


Fig. 4. Scatter plot of values in rule images obtained through SAM and SFF and the respective thresholds chosen to represent similarity or fit to alunite.

The DN values, $\theta(\overrightarrow{\mathbf{x}})$, from SAM's rule image in Fig. 3a were used in equation 10 to obtain scaled weights. We used a threshold, $\theta^t = 0.11$ radians. Pixels lying left of the 0.11 threshold (Fig. 4) correspond to positive weights. The resulting scaled weights correspond to a greater similarity to alunite reference spectrum.
SFF was applied to the alunite reference spectrum, resulting in a scale image and an RMS error image. The ratio of these images, produces a fit image (Fig. 3b). The bright pixels represent the best fit to the alunite reference spectrum. The DN values from the fit image, $\tau_F(\overrightarrow{\mathbf{x}})$, was used in equation 11 to obtain the weights for SFF using a threshold value of 20 for τ_F^t. This threshold was chosen after individual spectral analysis of some pixels and selecting several thresholds. The values of the rule images of SAM and SFF can be seen in Fig. 4. Pixels in the upper left quadrant correspond to positive weights. In equation 9 we have chosen $\kappa_1 = \kappa_2 = \frac{1}{2}$.
Table 2 summaries the weights derived by SAM and the weights derived by SFF. From the first row and first column, 6.5% of the pixels receive zero weight from one classification but weights larger than zero from the other classification. This can also be seen in Fig. 4 corresponding to the pixels in the upper right and lower left quadrants. These weights were then combined using equation 9 and are displayed in Fig. 6. Darker areas have higher weights and hence greater similarity to alunite reference spectrum in terms of both SAM and SFF. The sampling result using this weight function is also displayed in Fig. 6. The sample points are distributed over the alunite region and most of the points are found in the darker areas of the image.

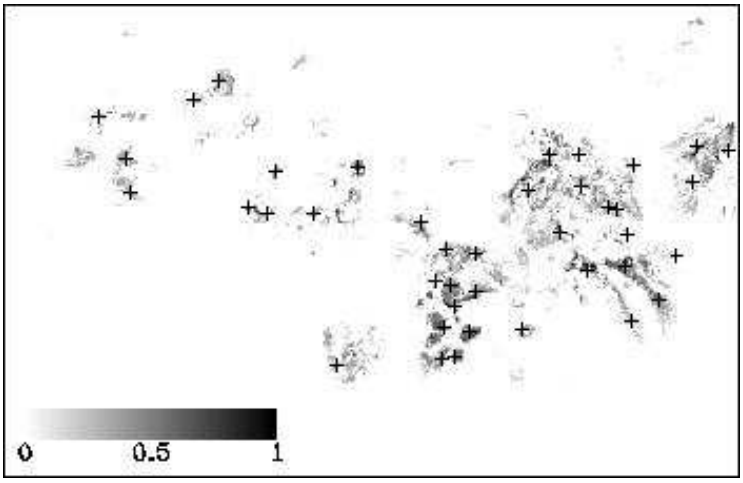


Fig. 5. Weight function: Scaled weights derived using SAM and SFF rule mages for alunite using their respective thresholds. Distribution of 40 sampling points using the weight function. Darker areas indicate greater similarity to alunite.

SFF \ SAM	0.0	(0.0, 0.2]	(0,2, 0.4]	(0.4, 0.6]	(0.6, 0.8]	(0.8, 1.0]
0.0	70603	615	238	84	40	8
(0.0, 0.2]	2546	739	304	126	78	8
(0,2, 0.4]	1183	710	332	111	37	7
(0.4, 0.6]	304	333	156	49	10	1
(0.6, 0.8]	33	42	31	6	1	0
(0.8, 1.0]	4	3	4	4	0	0

Table 2. Weights derived from SAM (column) and SFF (row).Values in the table represent the number (frequency) of pixels that match in a certain range.

Validation

Ground data collected using an ASD fieldspec-pro spectrometer were used to support the proposed sampling schemes by validating the SAM classified image and the images of the weights used in this chapter. Reflectance spectra of 51 ground measurements (see Fig. 3) were analyzed individually for their alunite content and classified into one of three classes, namely, "no alunite", "minor alunite" and "abundant alunite". Using the ground data of those pixels classified as alunite or not, the accuracy of SAM is 78% and for SFF is 79%.

5.4 Discussion and conclusion of the study

Designing sampling schemes that target areas with high probability and in greater abundance of alunite occurring was demonstrated by using a weight function for the WMSD-criterion as an objective function in simulated annealing. Predefined weights allow us to distinguish areas with different priorities. Hence sampling can be focused in areas with a high potential

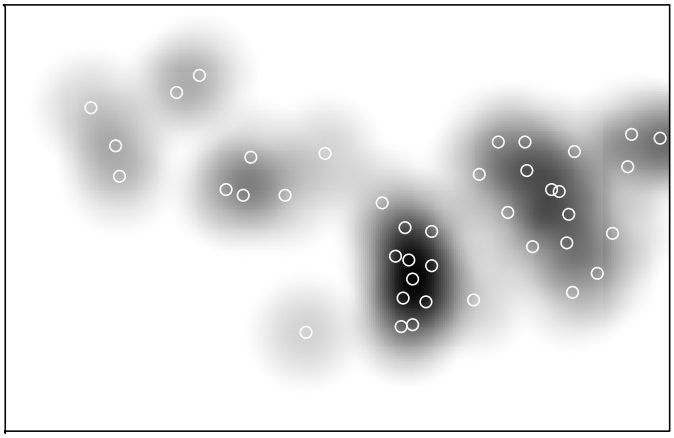


Fig. 6. Distribution of 40 optimized samples with $\theta^t = 0.11$ radians and $\tau_F^t = 20$. Darker patches in the images indicate sampling points are near to each other. This effectively implies greater abundance of alunite.

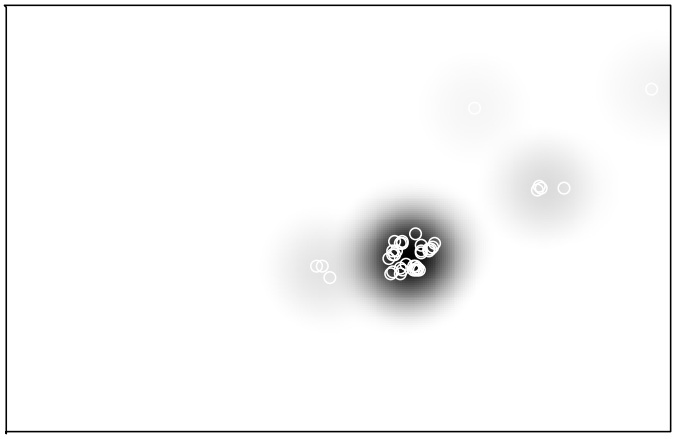


Fig. 7. Distribution of 40 highest weight samples. Darker patches in the images indicate sampling points are near to each other. This effectively implies greater abundance of alunite.

for the occurrence of a mineral of interest and reduces sampling in areas with low potential. This effectively reduces time and costs in the field. Randomly selecting points in the image, as potential sites to sample on the ground, could result in the location of these samples clustered and/or having a low probability of being alunite. Selecting a collection of sampling locations that have the highest probability of being alunite could result in the location of most sampling points clustered in the image (Fig. 7). This implies sampling in a limited area on the ground, and effectively these samples will not represent the overall distribution of alunite over the entire study area. In the proposed sampling schemes there is a balance between selecting samples that have a high probable alunite and the location of samples not to be clustered in

the field. A good sampling scheme will target areas with high probability of alunite and the distribution of sample points will correspond closely to the distribution of alunite. This means intensive sampling in the area with an abundance of alunite.
We used the threshold of 0.11 radians for SAM and a threshold of 20 for SFF. The threshold chosen for SAM in this case can be set higher (similarly the threshold for SFF can be set lower) to include some pixels with a reflectance spectrum similar to that of other minerals, example kaolinite and pyrophyllite. This is not considered to be a major problem, as the scaled weights used by the optimal sampling scheme will be low, thereby reducing the probability of selecting that pixel's location as a point to be sampled on the ground.
The weight function uses two rule images, one derived from SAM and another from SFF. A comparison of the scaled weights derived from SAM and SFF (Table 2), indicates that the methods for SAM and SFF do not always agree. Only the purest pixels classified as alunite have positive weights. The advantage of combining SAM and SFF classification methods in the weights function results in a classified image that is robust for the thresholds and selected channels. The weights derived from SAM and from SFF were then combined into a single weight image, which was used for the design of the optimal sampling scheme. A suitable range for the thresholds has to be known. This can be obtained by observing individual spectra and the purest of these can be selected to train the thresholds. Using the combined weights from SAM and SFF, sample points can be concentrated in the region with a high probability of alunite, which are robust against the thresholds selected. The distribution of sample points corresponds closely to the distribution of alunite (Fig. 6).
The sampling scheme proposed is of interest to (a) exploration geologists for specified target locations of hydrothermally altered minerals (e.g. alunite) with distinct absorption features, (b) researchers trying to understand the geothermal system and hydrothermal zones in a specific region and (c) engineers to better collect field data in relation to flights by improving on ground truthing and calibration measurements. With the aid of new spaceborne launched hyperspectral sensors, e.g. Hyperion and ARIES-1, data are available for most regions and hence will be helpful to geologist's planning phase of selecting important mineral targets in the field. The method presented here could result in reduction of time and effort in the field, but by no means replace the field geologist. It is merely an aid for target selection of minerals as an initial survey, followed by denser surface sampling of interesting anomalies.
Combination of SAM and SFF rule images thus obtained resulted in robust weights to focus sampling in areas of high probability of alunite. Sample points are arranged more intensely in areas with an abundance of alunite. SAM and SFF both lead to a relevant classification of the study area with respect to alunite, as observed from the rule images and validation of the rule images using ground measurements.

6. References

Aarts, E. & Korst, J. (1989). *Simulated Annealing and Boltzmann Machines*, New York: John Wiley.

Abrams, M., Ashley, R., Rowan, L., Goetz, A. F. H. & Kahle, A. (1977). Use of imaging in the .46–2.36 μm spectral region for alteration mapping in the Cuprite mining district, Nevada: USGS OFR-77-585.

Abrams, M. J. (1984). Landsat-4 thematic mapper and thematic mapper simulator data for a porphyry copper, *Photogrammetric Engineering and Remote Sensing* **50**: 1171–1173.

Allum, J. A. E. (1966). *Photogeology and regional mapping*, Oxford : Pergamon Press.

Arribas, Jr., A., Cunningham, C. G., Rytuba, J. J., Rye, R. O., Kelley, W. C., Podwysocki, M. H., McKee, E. H. & Tosdal, R. M. (1995). Geology, geochronology, fluid inclusions, and isotope geochemistry of the Rodalquilar gold alunite deposit, Spain, *Economic Geology* **90**: 795–822.

Ausmus, B. S. & Hilty, J. W. (1972). Reflectance studies of healthy, maize dwarf mosaic virus-infected, and Helminthosporium maydis-infected corn leaves, *Remote Sensing of Environment* **2**: 77–81.

Ben-Jemaa, F., Mariño, M. & Loaiciga, H. (1995). Sampling design for contaminant distribution in lake sediments, *Journal of Water Resource Planning Management* **121**: 71–79.

Bohachevsky, I. O., Johnson, M. E. & Stein, M. L. (1986). Generalized simulated annealing for function optimization, *Technometrics* **28**(3): 209–217.

Bouman, B. A. M. & Uenk, D. (1992). Crop classification possibilities with radar in ERS-1 and JERS-1 configuration, *Remote Sensing of Environment* **40**: 1–13.

Brisco, B., Brown, R. J. & Manore, M. J. (1989). Early season crop discrimination with combined SAR and TM data, *Canadian Journal of Remote Sensing* **15**(1): 44–54.

Brus, D. J. & de Gruijter, J. J. (1997). Random sampling or geostatistical modelling? Choosing between design-based and model-based sampling strategies for soil (with discussion), *Geoderma* **80**: 1–44.

Burgess, T. M., Webster, R. & McBratney, A. B. (1981). Optimal interpolation and isarithmic mapping of soil properties. IV Sampling strategy, *Journal of Soil Science* **32**(4): 643–660.

Carranza, E. J. M. & Hale, M. (2002). Mineral imaging with landsat thematic mapper data for hydrothermal alteration mapping in heavily vegetated terrane, *International Journal of Remote Sensing* **23**(22): 4827–4852.

Carter, G. A. (1994). Ratios of leaf reflectances in narrow wavebands as indicators of plant stress, *International Journal of Remote Sensing* **15**: 697–703.

Chabrillat, S., Goetz, A. F. H., Olsen, H. W., Krosley, L. & Noe, D. C. (1999). Use of AVIRIS hyperspectral data to identify and map expansive clay soils in the front range urban corridor in Colorado, *Proceedings of the 13th International Conference on Applied Geologic Remote Sensing*, I, Vancouver, British Columbia, Canada, pp. 390–397.

Chang, C.-I. (2006). *Hyperspectral Imaging: Techniques for Spectral Detection and Classification*, Springer.

Christakos, G. & Olea, R. A. (1992). Sampling design for spatially distributed hydrogeologic and environmental processes, *Advances in Water Resources* **15**(4): 219–237.

Clark, R. N. (1999). Spectroscopy of rocks and minerals, and principles of spectroscopy, *in* A. Rencz (ed.), *Remote Sensing for the Earth Sciences: Manual of Remote Sensing*, Vol. 3, John Wiley and Sons, New York, chapter 1, pp. 3–58.

Clark, R. N. & Roush, T. L. (1984). Reflectance spectroscopy: Quantitative analysis techniques for remote sensing applications, *Journal of Geophysical Research* **89**: 6329–6340.

Clark, R. N. & Swayze, G. A. (1995). Mapping minerals, amorphous materials, environmental materials, vegetation, water, ice, and snow, and other materials: The USGS Ticorder Algorithm, *Summaries of the Fifth Annual JPL Airborne Earth Science Workshop*, Vol. 1, JPL Publication 95-1, pp. 39–40.

Clark, R. N., Swayze, G. A. & Gallagher, A. (1992). Mapping the mineralogy and lithology of Canyonlands. Utah with imaging spectrometer data and the multiple spectral feature mapping algorithm, *Summaries of the Third Annual JPL Airborne Geoscience Workshop*, Vol. 1, JPL Publication 92-14, pp. 11–13.

Clark, R. N., Swayze, G. A., Gallagher, A. J., King, T. V. V. & Calvin, W. M. (1993). The U. S. Geological survey, digital spectral library: Version 1: 0.2 to 3.0 microns, U.S. Geological Survey Open File Report 93-592.

Clark, R. N., Swayze, G. A., Gorelick, N. & Kruse, F. A. (1991). Mapping with imaging spectrometer data using the complete band shape least-squares algorthm simultaneously fit to multiple spectral features from multiple materials, *Proceedings of the Third Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) workshop*, JPL Publication 91-28, pp. 2–3.

Clark, R. N., Swayze, G. A., Livo, K. E., Kokaly, R. F., Sutley, S. J., Dalton, J. B., McDougal, R. R. & Gent, C. A. (2003). Imaging spectroscopy: Earth and planetary remote sensing with the USGS Tetracorder and expert systems, *Journal of Geophysical Research* **108**(E12): 5–1–5–44.

Crósta, A. P., Sabine, C. & Taranik, J. V. (1998). Hydrothermal alteration mapping at Bodie, California, using AVIRIS hyperspectral data, *Remote Sensing of Environment* **65**(3): 309–319.

Cudahy, T., Okada, K. & Brauhart, C. (2000). Targeting VMS-style Zn mineralisation at Panorama, Australia, using airborne hyperspectral VNIR-SWIR HyMap data, *ERIM Proceedings of the 14th International Conference on Applied Geologic Remote Sensing*, Las Vegas, pp. 395–402.

Cunningham, C. G., Arribas, Jr., A., Rytuba, J. J. & Arribas, A. (1990). Mineralized and unmineralized calderas in Spain; Part I, evolution of the Los Frailes Caldera, *Mineralium Deposita* **25 [Suppl.]**: S21–S28.

Curran, P. J. & Atkinson, P. M. (1998). Geostatistics and remote sensing, *Progress in Physical Geography* **22**(1): 61–78.

Debba, P., Carranza, E. J. M., Stein, A. & van der Meer, F. D. (2008). Deriving optimal exploration target zones on mineral prospectivity maps, *Mathematical Geosciences* **41**(4): 421–446.

Debba, P., Carranza, E. J. M., Stein, A. & van der Meer, F. D. (2009). An optimal spatial sampling scheme to characterize mine tailings, Presented at the 57th Session of the International Statistical Institute (ISI) conference for a special topics contributed paper meeting on Statistical methods applied in GIS and remote sensing, 16-24 August 2009, International Conventional Centre (ICC), Durban, KwaZulu Natal, South Africa.

Debba, P., Stein, A., van der Meer, F. & Lucieer, A. (2008). Field sampling from a segmented image, *in* O. Gervasi, B. Murgante, A. Laganá, D. Taniar, Y. Mun & M. Gavrilova (eds), *Computational Science and Its Applications - ICCSA 2008.*, Vol. 5072 of *LNCS*, Springer, Heidelberg, pp. 756–768.

Debba, P., van Ruitenbeek, F. J. A., van der Meer, F. D., Carranza, E. J. M. & Stein, A. (2005). Optimal field sampling for targeting minerals using hyperspectral data, *Remote Sensing of Environment* **99**: 373–386.

Diggle, P. & Lophaven, S. (2006). Bayesian geostatistical design, *Scandinavian Journal of Statistics* **33**: 53–64.

Driscoll, R. S. & Coleman, M. D. (1974). Color for shrubs, *Photogrammetric Engineering and Remote Sensing* **40**: 451–459.

Eardley, A. J. (1942). *Aerial photographs: their use and interpretation*, New York: Harper.

Everitt, J. H., Gerbermann, A. H., Alaniz, M. A. & Bowen, R. L. (1980). Using 70 mm aerial photography to identify rangeland sites, *Photogrammetric Engineering and Remote Sensing* **46**: 1339–1348.

Everitt, J. H., Richardson, A. J., Gerbermann, A. H., Wiegand, C. L. & Alaniz, M. A. (1979). Landsat-2 data for inventorying rangelands in south Texas, *Proceedings of the 5th Symposium Machine Processing of Remotely Sensed Data. Purdue University, West Lafayette, Ind.*, pp. 132–141.

Gilbertson, B., Longshaw, T. G. & Viljoen, R. P. (1976). Multispectral aerial photography as exploration tool. IV-V - an applications in the Khomas Trough region, South Africa; and cost effective analysis and conclusions (for mineral exploration), *Remote Sensing of Environment* **5**(2): 93–107.

Goetz, A. F. H. & Srivastava, V. (1985). Mineralogical mapping Cuprite mining district, Nevada, *in* G. Vane & A. Goetz (eds), *Proc Airborne imaging spectrometer data analysis workshop*, Jet Propulsion Laboratory Publication 85-41, pp. 22–31.

Gupta, R. P. (2003). *Remote Sensing Geology*, second edn, Springer-Verlag New York, LLC.

Hapke, B. (1993). Combined theory of reflectance and emittance spectroscopy, *in* C. Pieters & P. A. J. Englert (eds), *Remote Geochemical Analysis: Elemental and Mineralogical Composition*, Cambridge University Press, Cambridge, UK, pp. 31–42.

Hedenquist, J. S., Arribas, A. R. & Gonzalez-Urien, E. (2000). Exploration for epithermal gold deposits, *Reviews in Economic Geology* **13**: 245–277.

Houston, R. S. (1973). Geologic mapping using space images, *Contributions to geology* **12**(2): 77–79.

Iranpanah, A. (1977). Geologic applications of Landsat imagery, *Photogrammetric Engineering and Remote Sensing* **43**: 1037–1040.

Johnson, P. L. (1969). *Remote sensing in ecology*, University of Georgia Press, Athens, GA.

Kirkpatrick, S., Gelatt, J. C. D. & Vecchi, M. P. (1983). Optimization by simulated annealing, *Science* **220**(4598): 671–680.

Knipling, E. B. (1970). Physical and physiological basis for the reflectance of visible and near-infrared radiation from vegetation, *Remote Sensing of Environment* **1**: 155–159.

Kowalik, W. S., Lyon, R. J. P. & Switzwe, P. (1983). The effect of additive reaidance terms on ratios of Landsat data (for mineral exploration), *Photogrammetric Engineering and Remote Sensing* **49**: 659–669.

Kruse, F. A. (2002). Comparison of AVIRIS and Hyperion for hyperspectral mineral mapping, *SPIE Aerospace Conference, 9-16 March 2002, Big Sky, Montana, published on CD-ROM, IEEE Catalog Number 02TH8593C, Paper 6.0102*, pp. 1–12.

Kruse, F. A. & Boardman, J. W. (1997). Characterization and mapping of Kimberlites and related diatremes in Utah, Colorodo, and Wyoming, USA, using the airborne visible/infrared imaging spectrometer (AVIRIS), *ERIM Proceedings of the 12th International Conference on Applied Geologic Remote Sensing*, Colorado, pp. 21–28.

Kruse, F. A., Lefkoff, A. B., Boardman, J. W., Heidebrecht, K. B., Shapiro, A. T., Barloon, P. J. & Goetz, F. H. (1993). The spectral image processing system (SIPS) - interactive visualization and analysis of imaging spectrometer data, *Remote Sensing Environment* **44**: 145–163.

Lark, R. M. (2002). Optimized spatial sampling of soil for estimation of the variogram by maximum likelihood, *Geoderma* **105**: 49–80.

Laylender, P. A. (1956). A performance estimate comparing conventional geologic mapping with that accomplished with the aid of color photographs, *Photogrammetric engineering* p. 953.

Lillesand, T. M., Kiefer, R. W. & Chipman, J. W. (1994). *Remote Sensing and Image Interpretation*, New York, John Wiley & Sons.

Lloyd, C. D. & Atkinson, P. M. (1999). Designing optimal sampling configurations with ordinary and indicator kriging, in proceedings of the 4th international conference on geocomputation, Virginia, USA, *GeoComputation 99* .

Longshaw, T. G. & Gilbertson, B. (1976). Multispectral aerial photography as exploration tool - III two applications in the North-Western Cape Province, South Africa (for mineral exploration), *Remote Sensing of Environment* **5**(2): 79–92.

Lowman, P. D. (1976). Geologic structure in California: three studies with Landsat-1 imagery, *California Geology* **29**: 75–81.

Martini, B. A. (2003). Assessing hydrothermal system dynamics and character by coupling hyperspectral imaging with historical drilling data: Long Valley Caldera, CA, USA, *Proceedings 25th New Zealand Geothermal Workshop*, Vol. 25, pp. 101–106.

Martini, B. A., Silver, E. A., Pickles, W. L. & Cocks, P. A. (2003). Hyperspectral mineral mapping in support of geothermal exploration: Examples from Long Valley Caldera, CA and Dixie Valley, NV, USA, *Geothermal Resources Council Transactions*, Vol. 27, pp. 657–662.

McBratney, A. B. & Webster, R. (1981). The design of optimal sampling schemes for local estimation and mapping of regionalized variables - II: Program and examples, *Computers & Geosciences* **7**(4): 335–365.

McBratney, A. B., Webster, R. & Burgess, T. M. (1981). The design of optimal sampling schemes for local estimation and mapping of regionalized variables - I: Theory and method, *Computers & Geosciences* **7**(4): 331–334.

McGraw, J. F. & Tueller, P. T. (1983). Landsat computer-aided analysis techniques for range vegetation mapping, *Journal of Range Management* **36**: 627–631.

McGwire, K., Friedl, M. & Estes, J. E. (1993). Spatial structure, sampling design and scale in remotely-sensed imagery of a California Savanna Woodlands, *International Journal of Remote Sensing* **14**(11): 2137–2164.

Müller, W. G. & Zimmerman, D. L. (1999). Optimal designs for variogram estimation, *Environmetrics* **10**(23–37).

Okina, G. S., Roberts, D. A., Murraya, B. & Okin, W. J. (2001). Practical limits on hyperspectral vegetation discrimination in arid and semiarid environments, *Remote Sensing of Environment* **77**: 212–225.

Papp, É. & Cudahy, T. (2002). Geophysical and remote sensing methods for regolith exploration, *in* É. Papp (ed.), *Hyperspectral remote sensing*, CRCLEME Open File Report 144, pp. 13–21.

Polder, G. & van der Heijden, G. W. A. M. (2001). Multispectral and hyperspectral image acquisition and processing, *in* Q. Tong, Y. Zhu & Z. Zhu (eds), *Proceedings of SPIE*, Vol. 4548.

Resmini, R. G., Kappus, M. E., Aldrich, W. S., Harsanyi, J. C. & Anderson, M. (1997). Mineral mapping with hyperspectral digital imagery collection experiment (HYDICE) sensor at Cuprite, Nevada, U.S.A., *International Journal of Remote Sensing* **18**(7): 1553–1570.

Richards, J. A. (1993). *Remote Sensing Digital Image Analysis: An Introduction*, second edn, Springer-Verlag, Berlin.

Richardson, A. J., Menges, R. M. & Nixon, P. R. (1985). Distinguishing weed from crop plants using video remote-sensing, *Photogrammetric Engineering & Remote Sensing* **51**(11): 1785–1790.

Richter, R. (1996). Atmospheric correction of DAIS hyperspectral image data, *SPIE Proceedings*, Vol. 2756, Orlando, pp. 390–399.

Rowan, L. C., Crowley, J. K., Schmidt, R. G., Ager, C. M. & Mars, J. C. (2000). Mapping hydrothermally altered rocks by analyzing hyperspectral image (AVIRIS) data of forested areas in the Southeastern United States, *Journal of Geochemical Exploration* **68**(3): 145–166.

Rowan, L. C., Goetz, A. F. H. & Ashley, R. P. (1977). Discrimination of hydrothermally altered and unaltered rocks in visible and near infrared multispectral images, *Geophysics* **42**(3): 522–535.

Royle, J. A. & Nychka, D. (1998). An algorithm for the construction of spatial coverage designs with implementation in S-PLUS, *Computational Geoscience* **24**: 479–488.

Russo, D. (1984). Design of an optimal sampling network for estimating the variogram, *Soil Science Society American Journal* **52**: 708–716.

Sabins, F. F. (1996). *Remote Sensing: Principles and Interpretation*, third edn, W.H. Freeman and Company, New York.

Sabins, F. F. (1999). Remote sensing for mineral exploration, *Ore Geology Reviews* **14**(Issues 3–4): 157–183.

Sacks, J. & Schiller, S. (1988). Spatial designs, *in* S. Gupta & J. Berger (eds), *Statistical Decision Theory and Related Topics*, Vol. 2 of *Papers from the forth Purdue symposium*, Springer-Verlag, New York, pp. 385–399.

Salisbury, J. W., Walter, L. S., Vergo, N. & D'Aria, D. M. (1991). *Infrared (2.1-2.5 μm) spectra of minerals*, Johns Hopkins University Press, Baltimore, MD.

Siegal, B. S. & Abrams, M. J. (1976). Geologic mapping using Landsat data, *Photogrammetric Engineering and Remote Sensing* **42**: 325–337.

Siegal, B. S. & Gillespie, A. R. (1980). *Remote sensing in geology*, New York : Wiley.

Smith, M. O., Johnston, P. E. & Adams, J. B. (1985). Quantitative determination of mineral types and abundances from reflectance spectra using principal component analysis, *Journal of Geophysical Research* **90**: 797–804.

Spruill, T. B. & Candela, L. (1990). Two approaches to design of monitoring networks, *Ground Water* **28**: 430–442.

Srivastav, S. K., Bhattacharya, A., Kamaraju, M. V. V., Reddy, G. S., Shrimal, A. K., Mehta, D. S., List, F. K. & Burger, H. (2000). Remote sensing and GIS for locating favourable zones of lead-zinc-copper mineralization in Rajpura-Dariba area, Rajasthan, India, *International Journal Remote Sensing* **21**(17): 3253–3267.

Stein, A., van der Meer, F. & Gorte, B. (eds) (1999). *Spatial Statistics for Remote Sensing*, Vol. 1 of *Remote Sensing and Digital Image Processing*, Kluwer Academic Publishers.

Tapia, R., Stein, A. & Bijker, W. (2005). Optimization of sampling schemes for vegetation mapping using fuzzy classification, *Remote Sensing of Environment* pp. 425–433.

Thenkabail, P. S. (2002). Optimal hyperspectral narrowbands for discriminating agricultural crops, *Remote Sensing Reviews* **20**(4): 257–291.
URL: *http://www.yale.edu/ceo/Projects/swap/pubs/ optimal_bands_text.pdf*

Thenkabail, P. S., Smith, R. B. & De-Pauw, E. (2002). Evaluation of narrowband and broadband vegetation indices for determining optimal hyperspectral wavebands for agricultural crop characterization, *Photogrammetric Engineering and Remote Sensing* **68**(6): 607–621.

Van der Meer, F. D. (2004). Analysis of spectral absorption features in hyperspactral imagery, *JAG: International Journal of Applied Earth Observation and Geoinformation* **5**(1): 55–68.

Van Groenigen, J. W., Gandah, M. & Bouma, J. (2000). Soil sampling strategies for precision agriculture research under Sahelian conditions, *Soil Science Society American Journal* **64**: 1674–1680.

Van Groenigen, J. W., Pieters, G. & Stein, A. (2000). Optimizing spatial sampling for multivariate contamination in urban areas, *Environmetrics* **11**: 227–244.

Van Groenigen, J. W., Siderius, W. & Stein, A. (1999). Constrained optimisation of soil sampling for minimisation of the kriging variance, *Geoderma* **87**: 239–259.

Van Groenigen, J. W. & Stein, A. (1998). Constrained optimization of spatial sampling using continuous simulated annealing, *Journal Environmental Quality* **27**: 1078–1086.

Vaughan, R. G., Calvin, W. M. & Taranik, J. V. (2003). SEBASS hyperspectral thermal infrared data: surface emissivity measurement and mineral mapping, *Remote Sensing of Environment* **85**(1): 48–63.

Warrick, A. W. & Myers, D. E. (1987). Optimization of sampling locations for variogram calculations, *Water Resources Research* **23**(3): 496–500.

Yfantis, E. A., Flatman, G. T. & Behar, J. V. (1987). Efficiency of kriging estimation for square, triangular, and hexagonal grids, *Mathematical Geology* **19**(3): 183–205.

Zimmerman, D. L. & Homer, K. E. (1991). A network design criterion for estimating selected attributes of the semivariogram, *Environmetrics* **4**: 425–441.