Optimal Sampling Schemes for Vegetation and Geological Field Visits

Pravesh Debba

CSIR, Spatial Planning Systems, Decision Support and Systems Analysis, CSIR Built Environment

Outline

BASICS OF SAMPLING SCHEMES

- Sample - small subset of the population of interest.
- Sample should represent the characteristics of the population (parameters / distribution).
- Draw inferences about a population based on incomplete knowledge.
- Distinguish between two general approaches
  - Design-based Methods
    - 'Ignore' the spatial structure
    - Use some form of random sampling
    - Use feature space to design sample
  - Model-based Methods
    - Explicitly model the spatial structure
    - Selection of sample based on optimisation criterion
    - Use geographic space to design sample

IMPORTANCE OF OPTIMAL SAMPLING SCHEMES

- Environmental studies:
  - where to sample?
  - what to sample?
  - and how many samples to obtain?
- Remote sensing as ancillary information in the design of optimal sampling schemes.
- Advantages of using remote sensing images:
  - Provides a synoptic overview of a large area
  - Wealth of information over the entire area
  - In these methods sampling avoids subjective judgement
  - Reduces costs and saves time on the field (fewer samples)

Optimal Sampling Schemes for Vegetation and Geological Field Visits

Debba

Introduction

Classification

Optimized sampling schemes case studies

Optimized field sampling for improved estimates of vegetation indices

Optimized field sampling representing the overall distribution of a particular mineral

Hyperspectral sensors

- record the reflectance in many narrow contiguous bands
- various parts of the electromagnetic spectrum (visible - near infrared - short wave infrared)
- at each part of the electromagnetic spectrum results in an image

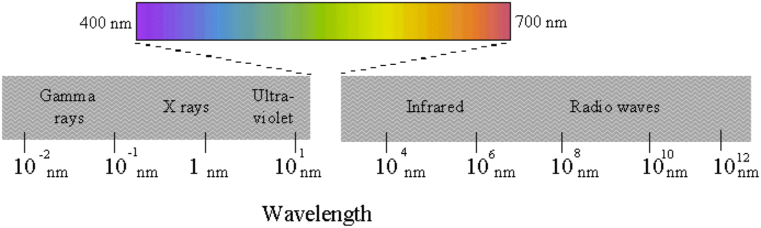


Figure: Spectral Range

OVERVIEW OF HYPERSPECTRAL REMOTE SENSING (cont. . .)

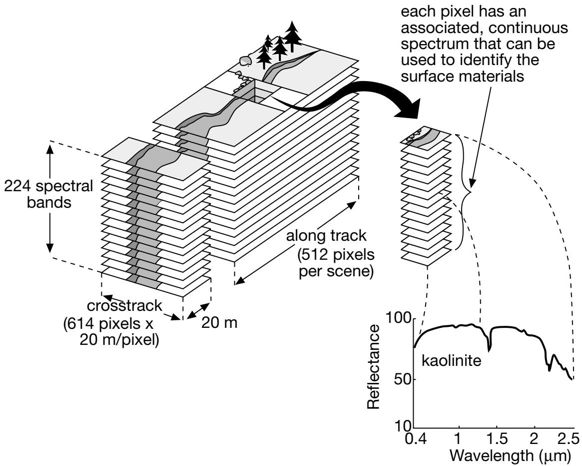


Figure: Hyperspectral cube

OVERVIEW OF HYPERSPECTRAL REMOTE SENSING (cont. . .)

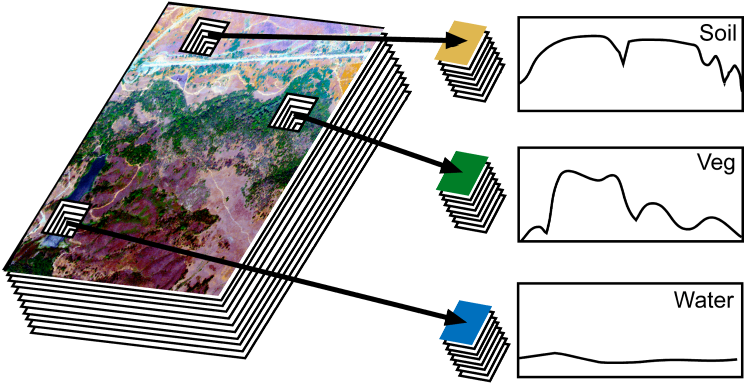


Figure: Pixels in hyperspectral image

OVERVIEW OF HYPERSPECTRAL REMOTE SENSING (cont. . .)

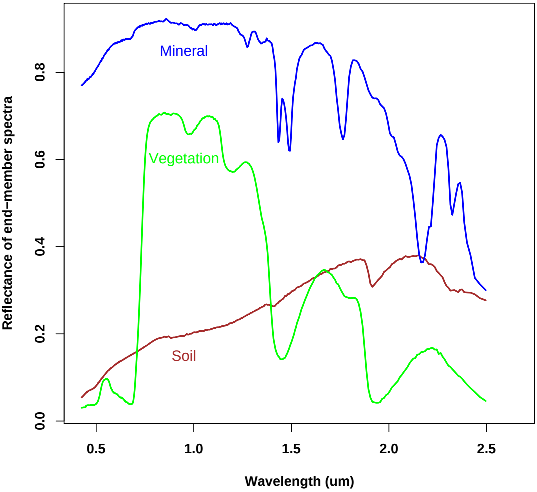


Figure: Example of 3 different spectral signatures

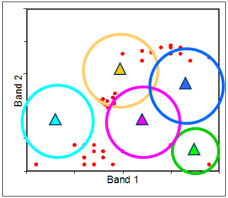
UNSUPERVISED CLASSIFICATION

- No previous knowledge assumed about data.
- Tries to spectrally separate the pixels.
- User has controls over:
  - Number of classes
  - Number of iterations
  - Convergence thresholds
- Two main algorithms: Isodata and k-means

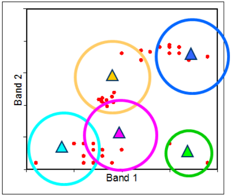
K-MEANS CLUSTERING

- A set number of cluster centres are positioned randomly through the spectral space.
- Pixels are assigned to their nearest cluster.
- The mean location is re-calculated for each cluster.
- Repeat 2 and 3 until movement of cluster centres is below threshold.
- Assign class types to spectral clusters.

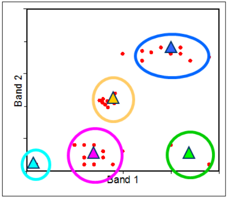
K-MEANS CLUSTERING (cont. . .)



(a) 1st iteration. Cluster centres are set at random. Pixels assigned to the nearest centre.



(b) 2nd iteration. Centres move to the mean-centre of all pixels in this cluster.

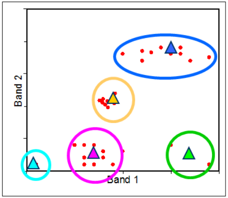


(c) N-th iteration. Centres have stabilised.

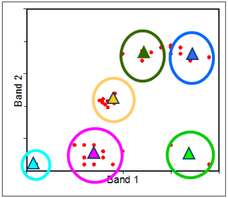
ISODATA

- Extends k-means. Also calculate standard deviation for clusters.
- After stage 3 we can either:
  - Combine clusters if centres are close.
  - Split clusters with large standard deviation in any dimension.
  - Delete clusters that are to small.
- Then reclassify each pixel and repeat.
- Stop on max iterations or convergence limit.
- Assign class types to spectral clusters.

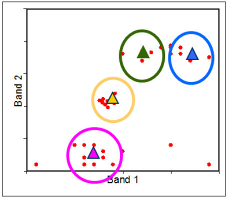
ISODATA (cont. . .)



(d) Data is clustered but blue cluster is very stretched in band 1.



(e) Cyan and green clusters only have 2 or less pixels. So they will be removed.



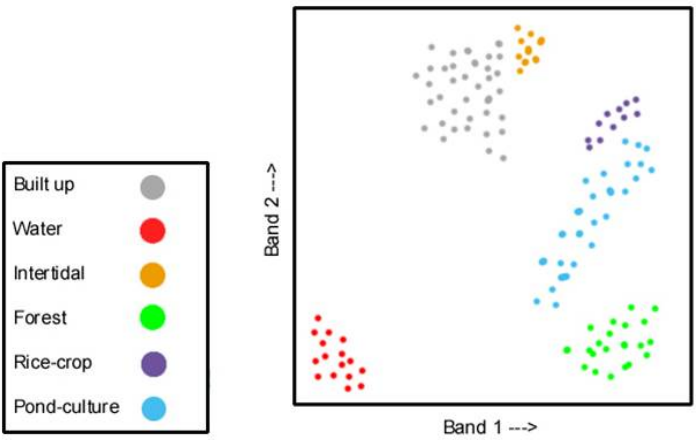
(f) Either assign outliers to nearest cluster, or mark as unclassified.

SUPERVISED CLASSIFICATION

- Start with knowledge of class types.
- Classes are chosen at start.
- Training regions are created for each class.
- Ground truth used to verify the training regions.
- Quite a few algorithms. Here we will look at:
  - Parallelepiped
  - Maximum likelihood

DATA

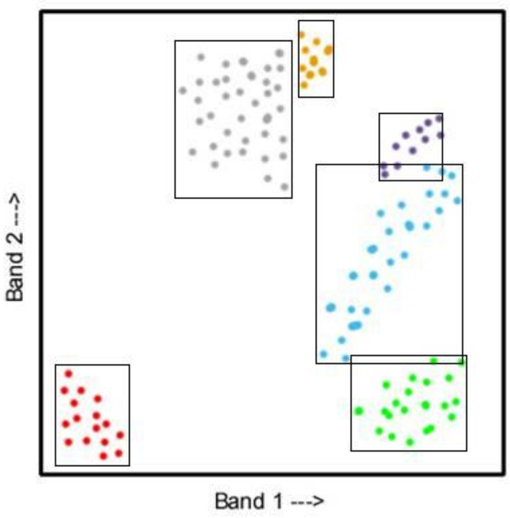
Training classes plotted in spectral space. In this example using 2 bands.



PARALLELEPIPED

- For each training region determine the range of values observed in each band.
- These ranges form a spectral box (or parallelepiped) which is used to classify this class type.
- Assign new image pixels to the parallelepiped which it fits into best.
- Pixels outside all boxes can be unclassified or assigned to the closest one.
- Problems with classes that exhibit high correlation between bands. This creates long ‘diagonal’ data-sets that do not fit well into a box.

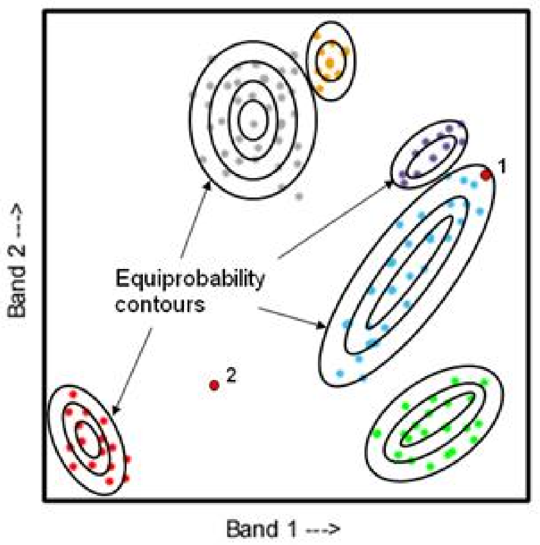
PARALLELEPIPED (cont. . .)



MAXIMUM LIKELIHOOD

- For each training class the spectral variance and covariance is calculated.
- The class can then be statistically modelled with a mean vector and covariance matrix.
- This assumes the class is normally distributed. Which is generally okay for natural surfaces.
- Unidentified pixels can then be given a probability of being in any one class.
- Assign the new pixel to the class with the highest probability — or unclassified if all probabilities low.

MAXIMUM LIKELIHOOD (cont. . .)



Outline

1. Introduction
2. Classification
3. Optimized sampling schemes case studies
   - Optimized field sampling for improved estimates of vegetation indices
   - Optimized field sampling representing the overall distribution of a particular mineral

OBJECTIVE OF STUDY

- The design of the optimal prospective sampling scheme for field visits in an agricultural study, using a segmented hyperspectral image.
- The optimal prospective sampling scheme will be representative of the whole study area for various parameters embedded by the segmentation and bands selected for the segmentation.

STUDY SITE

- Study site – Tedej – Hungary.
- Crops: barely, maize, sugar beet, sunflower, alfalfa.
- Digital Imaging Spectrometer – DAIS-7915 – 79 channel hyperspectral image.
- Spectral range from visible (0.4 μm) to thermal infrared (12.3 μm).
- Spatial resolution 3–20 m depending on the carrier aircraft altitude.

STUDY SITE (cont. . .)

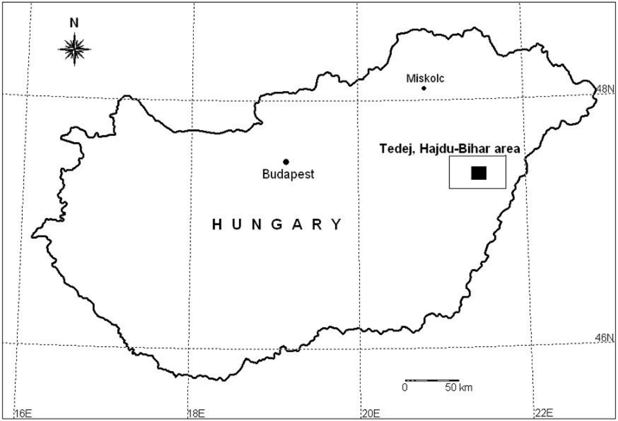


Figure: Study area in Tedej, Hajdu-Bihar area, Hungary.

STUDY SITE (cont. . .)

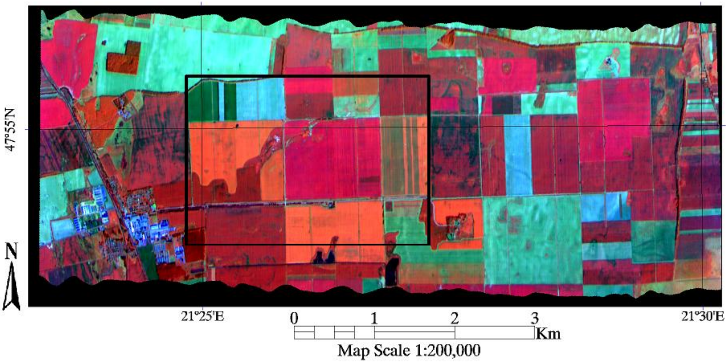


Figure: Hyperspectral image of study area in Tedej, Hajdu-Bihar area, Hungary. Reflectance values for bands 29 (0.988 μm), 39 (1.727 μm) and 1 (0.496 μm).

METHODS: ITERATED CONDITIONAL MODES (ICM) ALGORITHM

- Adequate image segmentation takes into account both spectral features and spatial information.
- Markov Random Fields (MRF) have been useful in this respect.
- For each category $k = 1, 2, \ldots, K$, let
  - $\mathbf{C}_k^{(\alpha)}$ denote the set of pixels which belongs to the kth category and $\mathbf{C}^{(\alpha)} = \bigcup_{k=1}^{K} \mathbf{C}_k^{(\alpha)}$ the segmented image at the αth iteration, $\alpha = 0, 1, 2, \ldots,$
  - $N_k^{(\alpha)}$ denote the number of elements in $\mathbf{C}_k^{(\alpha)}$, i.e. the number of pixels in the kth category at the αth iteration,
  - $\mu_k^{(\alpha)} = \sum_{(i,j) \in \mathbf{C}_k^{(\alpha)}} f_{ij} / N_k^{(\alpha)}$ be the m-dimensional mean vector of the kth category at the αth iteration.

METHODS: ITERATED CONDITIONAL MODES (ICM) ALGORITHM (cont. . .)

$$\arg\min_k \left\{ \left(f_{ij} - \mu_k^{(\alpha)} \right)^T \left(f_{ij} - \mu_k^{(\alpha)} \right) - \beta \nu^{(\alpha)} N_{ij}^{(\alpha)}(k) \right\} \qquad (1)$$

$$\nu^{(\alpha)} = \frac{1}{N} \sum_{k=1}^{K} \sum_{(i,j) \in \mathbf{C}_k^{(\alpha)}} \left(f_{ij} - \mu_k^{(\alpha)} \right)^T \left(f_{ij} - \mu_k^{(\alpha)} \right). \qquad (2)$$

METHODS (cont. . .): SECOND ORDER MRF FOR ICM

A second order MRF was applied in which the neighbors of each pixel consists of its eight adjacencies, with border pixels adjusted appropriately.

1	2	2
1	(i,j)	2
3	3	5

$$N_{ij}^{(\alpha)}(1) = 2$$
$$N_{ij}^{(\alpha)}(2) = 3$$
$$N_{ij}^{(\alpha)}(3) = 2$$
$$N_{ij}^{(\alpha)}(4) = 0$$
$$N_{ij}^{(\alpha)}(5) = 1$$

Figure: Calculation of $N_{ij}^{(\alpha)}(k)$ for an arbitrary interior pixel (i,j) belonging to category k.

METHODS (cont. . .): SAMPLE SIZE PER CATEGORY

For a pre-specified number of n samples, the sample size for category k equals

$$n_k = n_{(0)} + \left(n - K \cdot n_{(0)}\right) \cdot \frac{N_k^{(r)} \sqrt{\nu_k^{(r)}}}{\sum_{t=1}^{K} N_t^{(r)} \sqrt{\nu_t^{(r)}}}, \qquad (3)$$

where $\nu_k^{(r)} = \frac{1}{N_k^{(r)}} \sum_{(i,j) \in \mathbf{C}_k^{(r)}} \left(f_{ij} - \mu_k^{(r)}\right)^T \left(f_{ij} - \mu_k^{(r)}\right)$.

METHODS (cont. . .): SIMULATED ANNEALING

Simulated annealing — optimization method to find the global optimum of an objective function in the presence of local optima. A fitness function $\phi(\mathbf{S})$ has to be minimized. A probabilistic acceptance criterion decides whether $\mathbf{S}_{i+1}$ is accepted or not:

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

METHODS (cont. . .): FITNESS FUNCTION PER CATEGORY

The initial sampling scheme for the kth category $\mathbf{S}_k^{(0)}$ is a random selection of n_k [see Equation 3] points from category k. For $\mathbf{S}_k$, the fitness function equals

$$\phi_{\text{MMSD}}(\mathbf{S}_k) = \frac{1}{N_k^{(r)}} \sum_{(i,j)\in\mathbf{C}_k^{(r)}} \left|\left|c_{k(ij)} - W_{\mathbf{S}_k}(c_{k(ij)})\right|\right| \,, \qquad (5)$$

where $c_{k(ij)} \in \mathbf{C}_k^{(r)}$ is a location vector denoting the (i,j)th pixel belonging to category k and $W_{\mathbf{S}_k}(c_{k(ij)})$ denotes the location vector of the nearest sampling point in $\mathbf{S}_k$.

RESULTS: GENERATED SEGMENTED IMAGE

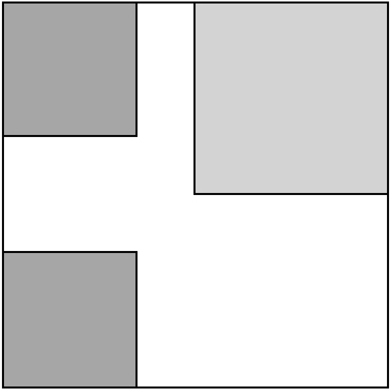


Figure: Generated segmented image.

RESULTS (cont. . .): OPTIMIZED SAMPLING SCHEME

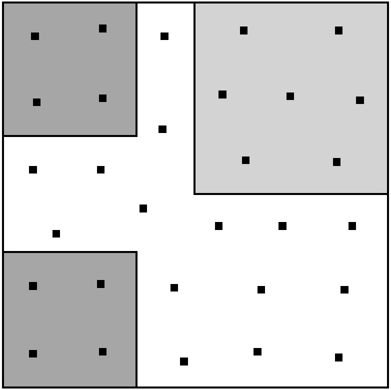


Figure: Optimized sampling scheme.

RESULTS (cont. . .): ORIGINAL HYPERSPECTRAL IMAGE

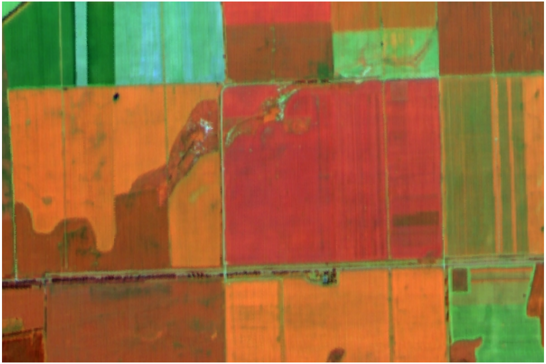


Figure: Original hyperspectral image. Reflectance values for bands 29 (0.988 μm), 39 (1.727 μm) and 1 (0.496 μm).

RESULTS (cont. . .): SEGMENTED IMAGE – 8 CATEGORIES

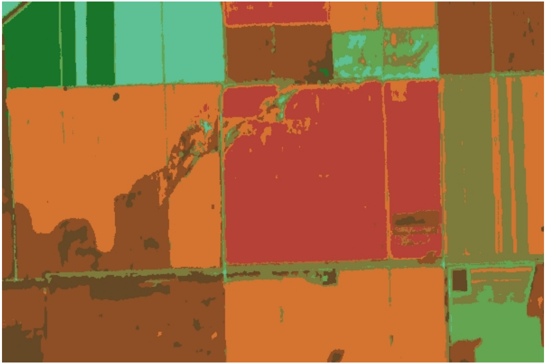


Figure: ICM Segmented image with eight categories.

RESULTS (cont. . .): SEGMENTED IMAGE – 4 ROI CATEGORIES

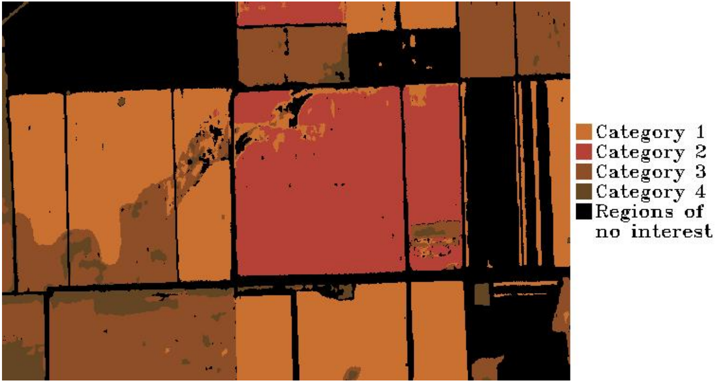


Figure: Segmented image confining sampling regions to the four categories.

RESULTS (cont. . .): OPTIMIZED SAMPLING SCHEME

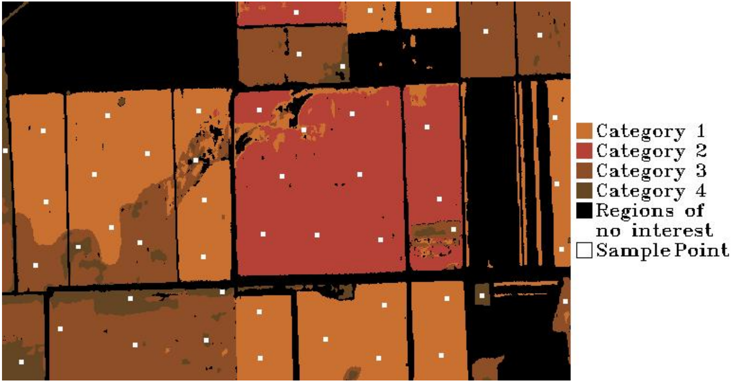


Figure: Optimized sampling locations of 50 points distributed over 4 categories.

RESULTS (cont. . .): DIFFERENT VEGETATION INDICES

- Normalized Difference Vegetation Index (NDVI)

$$\mathrm{NDVI} = \frac{R_{0.886} - R_{0.675}}{R_{0.886} + R_{0.675}} \tag{6}$$

- Renormalized Difference Vegetation Index (RDVI)

$$\mathrm{RDVI} = \frac{R_{0.886} - R_{0.675}}{\sqrt{R_{0.886} + R_{0.675}}} \tag{7}$$

- Modified Simple Ratio (MSR)

$$\mathrm{MSR} = \left(\frac{R_{0.886}}{R_{0.675}} - 1\right) \Big/ \sqrt{\frac{R_{0.886}}{R_{0.675}} + 1} \tag{8}$$

- Soil-Adjusted Vegetation Index (MSAVI)

$$\mathrm{MSAVI} = \frac{1}{2}\left[(2R_{0.886} + 1)^2 - 8(R_{0.886} - R_{0.675})\right] \tag{9}$$

RESULTS (cont. . .): COMPARISON OF SAMPLING SCHEMES

		Mean			
		NDVI	RDVI	MSR	MSAVI
Image		0.59	8.8	1.34	1.24
Optimized sampling scheme		0.58	8.6	1.32	1.22
Random sampling scheme	1	0.49	7.9	1.18	1.09
	2	0.38	6.1	0.94	0.89
	3	0.45	7.0	1.11	1.06
Grid sampling scheme	1	0.49	7.8	1.14	1.13
	2	0.53	8.2	1.25	1.13
	3	0.53	8.3	1.26	1.15

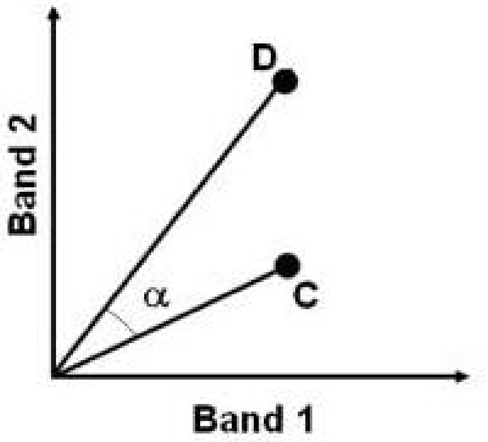
Outline

OBJECTIVE OF STUDY

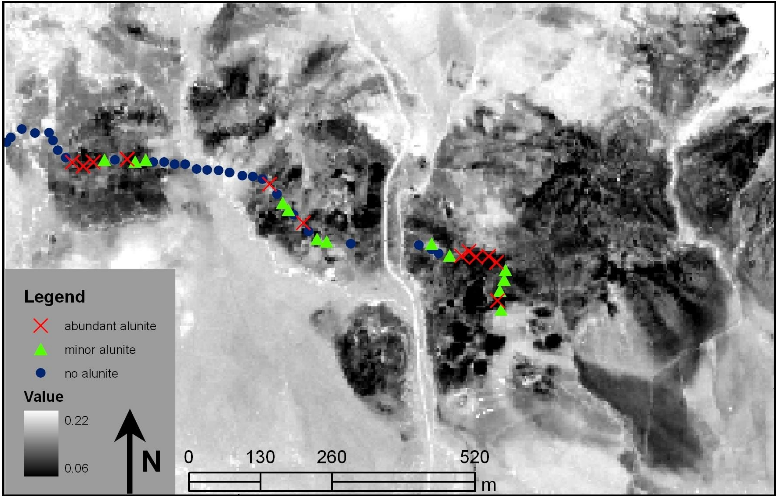
Using a hyperspectral image, to guide field sampling collection to those pixels with the highest likelihood for occurrence of a particular mineral, for example alunite, while representing the overall distribution of alunite.

Usefulness: To create a mineral alteration map

SPECTRAL ANGLE MAPPER



SPECTRAL ANGLE MAPPER (cont. . .)



CONTINUUM REMOVAL

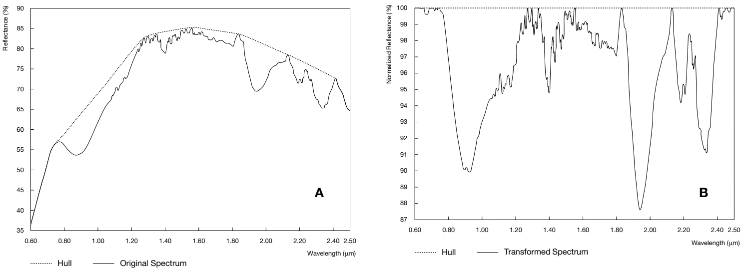


Figure: Concept of the convex hull transform; (A) a hull fitted over the original spectrum; (B) the transformed spectrum.

METHODS (cont. . .): SFF Rule Image for Alunite

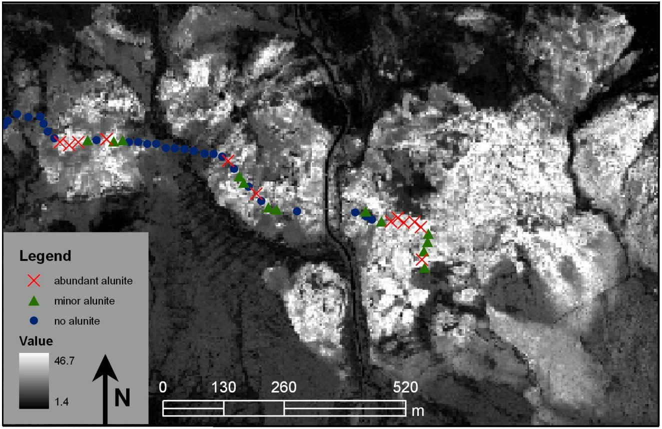


Figure: SFF fit image for alunite. Lighter areas indicate better fit values between pixel reflectance spectra and the alunite reference spectrum.

Combination of SAM and SFF scaled to $[0,1]$ is defined as

$$w(\theta(\overrightarrow{\mathbf{x}}), \tau_F(\overrightarrow{\mathbf{x}})) = \begin{cases} \kappa_1 w_1(\theta(\overrightarrow{\mathbf{x}})) + \kappa_2 w_2(\tau_F(\overrightarrow{\mathbf{x}})), & \\ \quad \text{if } \theta(\overrightarrow{\mathbf{x}}) \leq \theta^t \text{ and } \tau_F(\overrightarrow{\mathbf{x}}) \geq \tau_F^t & \\ 0, & \text{if otherwise} \end{cases} \quad (10)$$

$$\phi_{\text{WMSD}}(\mathbf{S}^n) = \frac{1}{N} \sum_{\overrightarrow{\mathbf{x}} \in \mathbf{I}} w(\overrightarrow{\mathbf{x}}) \left|\left| \overrightarrow{\mathbf{x}} - W_{\mathbf{S}^n}(\overrightarrow{\mathbf{x}}) \right|\right| , \quad (11)$$

METHODS (cont. . .): Fitness Function

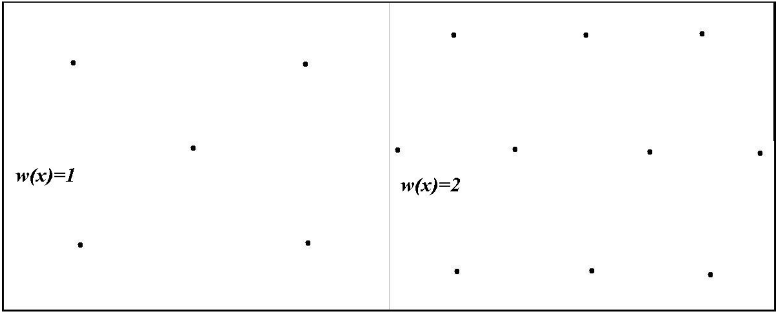


Figure: Fitness function with different weights for $N = 15$.

RESULTS (cont. . .): OPTIMIZED SAMPLING SCHEME

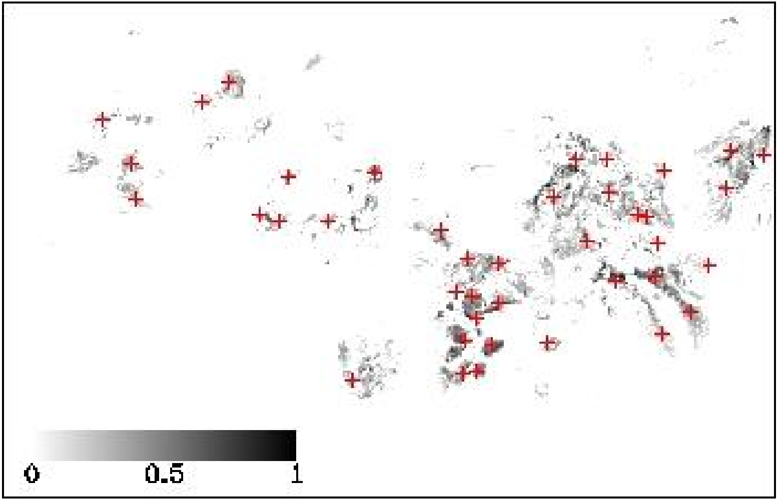


Figure: Optimized sampling scheme.

RESULTS (cont. . .): Distribution of 40 highest values

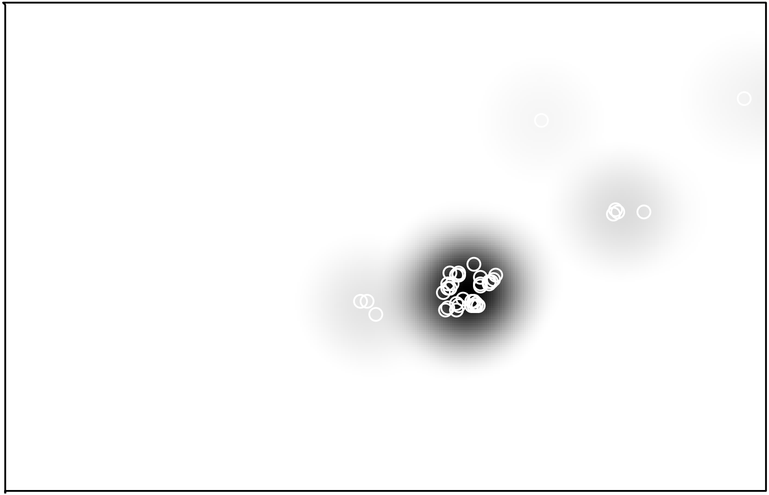


Figure: Sampling scheme: 40 highest values

RESULTS (cont. . .): Distribution of 40 optimized sampling scheme

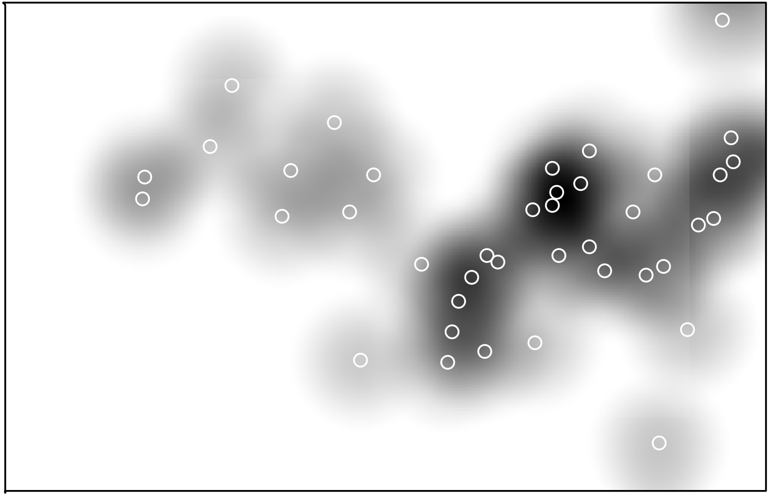


Figure: Distribution of 40 optimized sampling scheme

This is a numerical measure of the quality of the sampling design. The most common are:

- Minimise the maximum kriging variance
- Minimise the average kriging variance
- Maximise the information in a sample variogram

Kriging variance does not depend on the observed values, but only on the spatial structure and the location of the sample points i.e. the only factors influencing the kriging variance are therefore the variogram, the number of observations and the location of the prediction point. This means that it is possible to calculate the kriging variance before actual sampling takes place, provided the variogram is known or can be assumed. This feature is used to optimise spatial sampling schemes for minimal kriging variance.

Example

$$\phi_{\mathrm{OK}}(\mathbf{S}) = \frac{1}{N}\sum_{j=1}^{N} \sigma^2_{\mathrm{OK}}\left(x_j|\mathbf{S}\right) \ , \tag{12}$$

or

$$\phi_{\mathrm{MAX}}(\mathbf{S}) = \max\left(\sigma^2_{\mathrm{OK}}\left(x_j|\mathbf{S}\right)\right) \ , \tag{13}$$

where

$$\sigma^2_{\mathrm{OK}}\left(x_0\right) = \sum_{i=1}^{N} \lambda_i \cdot \gamma\left(x_i - x_0\right) + \Phi \ , \tag{14}$$

where λ_i denotes the weight of the ith observation and Φ a Lagrange multiplier.