USING REMOTELY-SENSED DATA FOR OPTIMAL FIELD SAMPLING

BY
DR PRAVESH DEBBA

STATISTICS IS THE SCIENCE pertaining to the collection, summary, analysis, interpretation and presentation of data. It is often impractical – or even impossible – to collect data from the total population of interest. Instead, one uses data from a subset of the population, called a sample. A sample, however, gives rise to uncertainty in terms of the process or population being studied. Subsequently, the design of sampling schemes plays an important role in statistics because a sample should represent the characteristics of the population, in terms of its distribution and associated parameters.

Some of the key questions in environmental studies are: where to sample, what to sample and how many samples to obtain. Conventional sampling techniques are not always suitable in environmental studies and scientists have explored the use of remotely-sensed data as ancillary information to aid in the design of sampling schemes. Remote sensing images provide a synoptic overview of a large area, thus giving a wealth of information over the entire area.

This can be far more informative than a few selected points on the ground for fieldwork. However, fieldwork is still an important exercise because one can measure certain features that are unobservable from the image, or one can estimate more accurately certain features at that specific point. For an environmentalist, knowing the exact location where to carry out a fieldwork sample is an important issue as it avoids subjective judgement and can save on time and costs in the field.

STATISTICAL SAMPLING, USING DATA OBTAINED FROM REMOTE SENSING, FINDS APPLICATION IN A VARIETY OF FIELDS. THESE INCLUDE AGRICULTURE FOR ESTIMATING THE VEGETATION INDICES FOR DIFFERENT TYPES OF CROPS ON FIELDS; GEOLOGY FOR IDENTIFYING PARTICULAR MINERALS; AND SURVEYING FOR ISOLATING AREAS WITH POSSIBLE UNDISCOVERED MINERAL DEPOSITS.

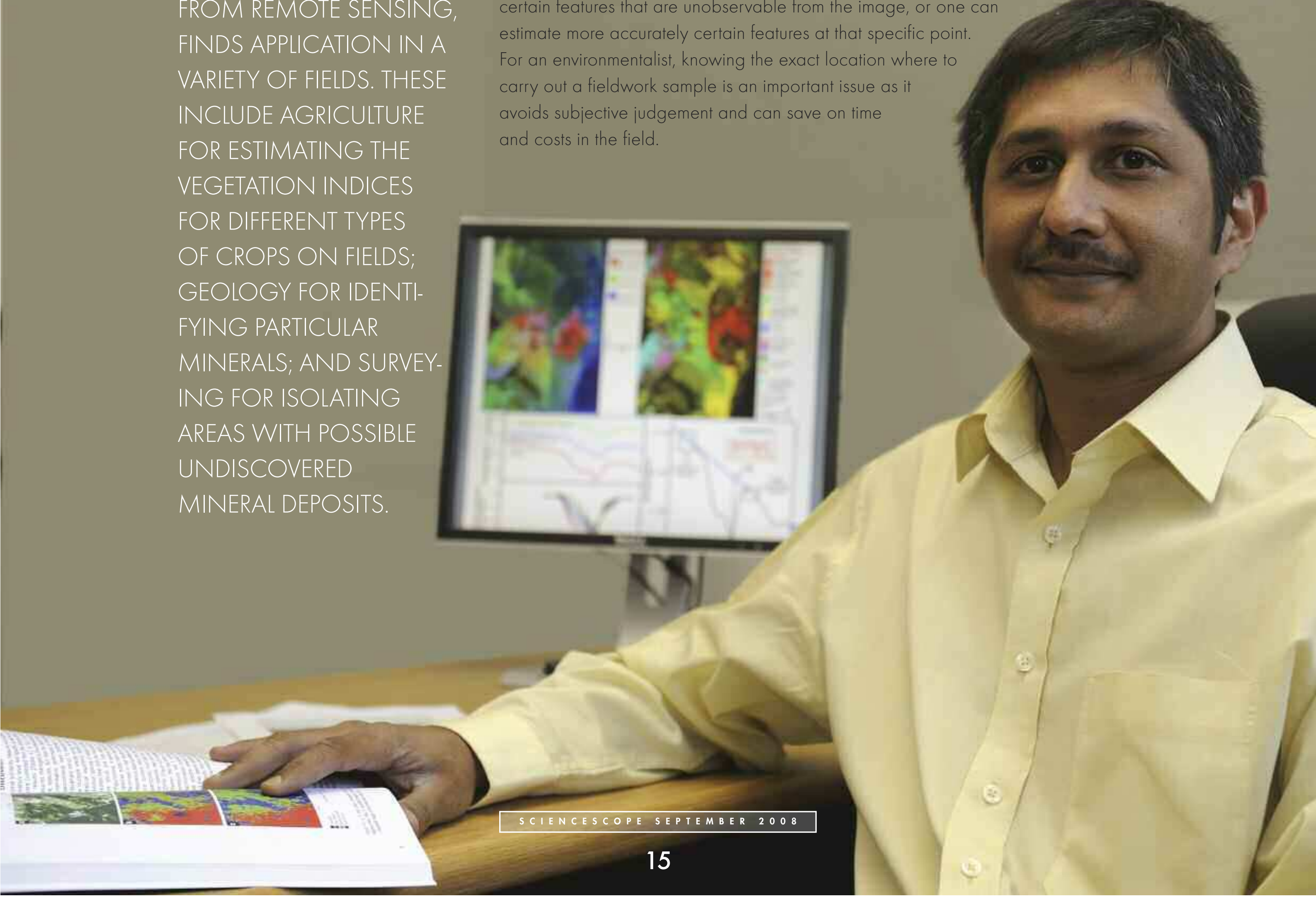
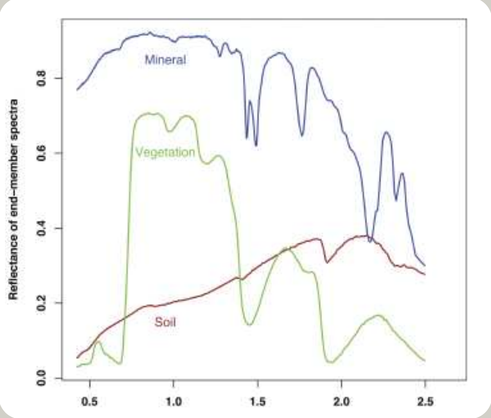


Figure 1: Examples of different spectral signatures



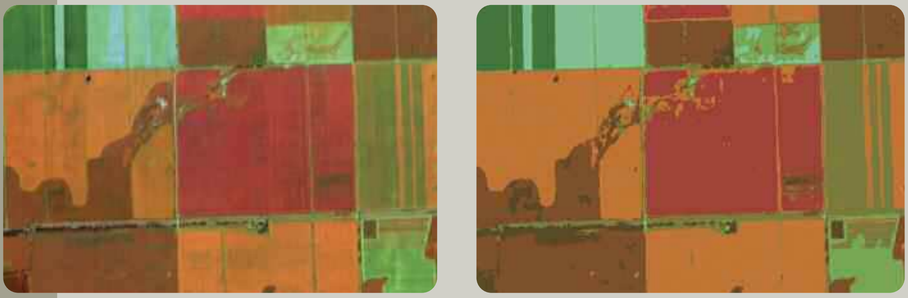


Figure 2: (a) Original image

(b) Classified image

Figure 3: Optimised sampling scheme

Remote sensing is the acquisition of data in the form of images obtained from a sensor mounted on an aircraft, spacecraft, satellite, ship or even hand-held over the earth's surface. Our study in designing optimal sampling schemes used hyperspectral images, obtained from sensors mounted on an aircraft. These sensors typically record the reflectance in many narrow contiguous bands at various parts of the electromagnetic spectrum, ranging from visible through near infrared to short-wave infrared. The recording at a specific part of the electromagnetic spectrum results in a grayscale image.

Observing a single pixel over the range of images, namely over the electromagnetic spectrum, results in a spectral signature (see Figure 1). Each object has a distinct spectral signature and it is this signature that can be used to classify an image into different classes, for example, barley, maize, sugar beet and sunflower.

Three examples follow to illustrate the way in which remotely-sensed data can be used to derive optimal sampling schemes for fieldwork.

ESTIMATES OF VEGETATION INDICES

CSIR researchers and their peers in the Netherlands designed an optimal sampling scheme for field visits, on an agricultural field in Hungary, which would be able to estimate more accurately the various vegetation indices for each of the different types of crops on the field. A vegetation index is a quantitative value used to measure biomass or vegetative vigour. First, the hyperspectral image (Figure 2a) was classified into eight classes (Figure 2b) using an automated process. Four classes were grouped to form a region of no interest (Figure 3) where there was no vegetation. Through a mathematical objective function and an optimisation procedure, 50 sample points were spread over the four classes (Figure 3).

The optimised sampling scheme resulted in more accurate estimates for various vegetation indices compared to simple random sampling, grid sampling or stratified random sampling. The optimised sampling scheme could potentially have an impact on providing improved estimates on the health status of the various types of crops in an agricultural field.

MINERAL DISTRIBUTION

In a geological study, applied to an area in Spain, we designed a sampling scheme that has the highest likelihood for identifying the occurrence of a particular mineral, namely alunite, while sampling from occurrences across the whole area. Geologists are often interested in creating a mineral alteration map. To achieve this, they need to identify hydrothermal alteration minerals (e.g. alunite), through field sampling.

First, each pixel in the hyperspectral image was matched to the alunite reflectance spectrum (Figure 1: mineral). We then created a soft classification of the image. Each pixel ranges between zero and one and the value was used as weights in a mathematical objective function. When this function was optimised, the sampling points were distributed over the alunite region and most of the points are found in the areas where

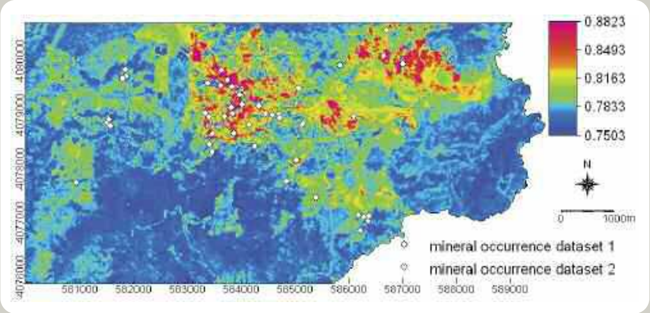


Figure 4(a): Band ratio 1

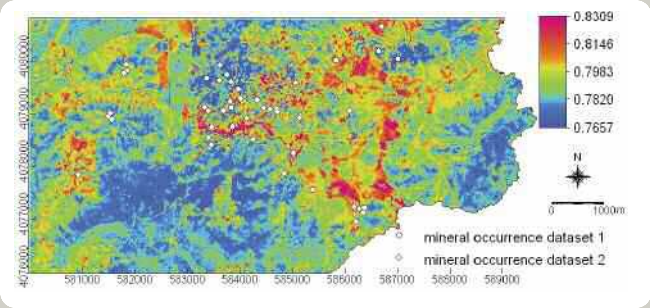


Figure 4(b): Band ratio 2

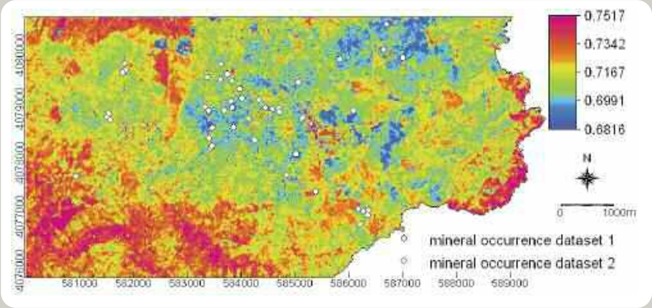


Figure 4(c): Band ratio 3

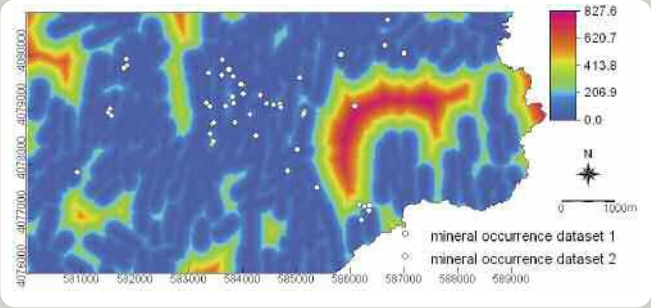


Figure 4(d): Distance to faults and fractures

there is a higher likelihood of alunite. This could potentially save time and costs because with this sampling scheme the geologist is almost certain to sample only alunite on the ground, while accurately reflecting the distribution of alunite in the region.

EXPLORATION TARGET ZONES

Another area in which sampling schemes play an important role is in deciding where to demark exploration target zones for further surveying as potential mining areas. Typically, these areas should have a high likelihood of undiscovered mineral deposit occurrences. This study was also conducted in Spain.

To test our methodology we obtained two independent datasets of discovered mineral deposits. One set was used to arrive at a mineral potential map and an optimised sampling scheme in terms of exploration target zones. The second dataset was used to determine if the predicted exploration target zones from our sampling scheme contains one of the 'now assumed' undiscovered mineral deposits.

To determine the mineral potential map, we created band ratio images from the hyperspectral images and a map of the distance to faults and fractures (Figures 4a-d). These were used as input layers in our modelling.

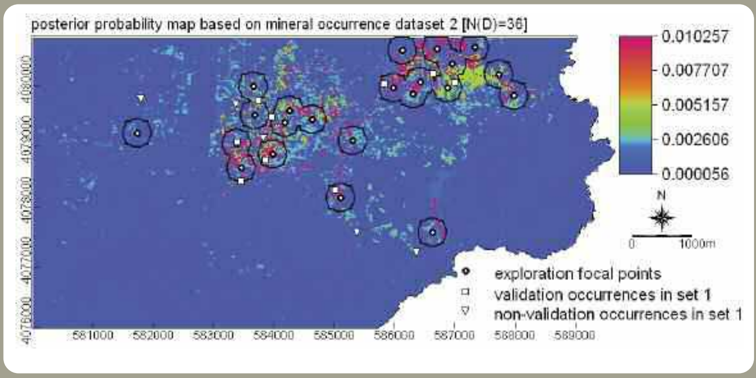


Figure 5: Optimal target zones on a mineral potential map

With the aid of an appropriately defined objective function and through mathematical optimisation, we arrived at the optimal target zones (black circles) on the derived mineral potential map (Figure 5). These target zones contain nine of the 14 assumed undiscovered mineral deposits. The remaining five that were undiscovered are not too far from a target zone. Furthermore, the optimised target zones suggest new potential areas of undiscovered mineral deposits for further surveying.

The advantage of this research is that it enables the geologist to concentrate on specific areas with a high likelihood of undiscovered mineral deposits for further surveying, thereby saving time and costs on surveying areas less likely to be commercially viable.

The research referred to in this article was conducted with collaborators and supervisors from the International Institute for GeoInformation Sciences and Earth Observations (ITC) in the Netherlands.

Enquiries:
Dr Pravesh Debba
pdebba@csir.co.za