Mapping Rocks, Soils, Vegetation Communities and Vegetation Density with the GERIS using Linear Mixture Modelling and Post-Processing Techniques

N.A. Drake

Department of Geography, University of Reading,
Whiteknights, Reading, U.K.

Abstract

Linear Mixture Modelling is applied to GERIS data using twenty four bands from 1.97 to 2.366 microns that have been processed to suppress the noise and provide an image of relative reflectance. The mixture modelling procedure defines five image components: Sericitic alteration, dolomite, soil, green vegetation and dry vegetation. The resultant mixture maps provide new information on the geology of the area by defining previously unrecognized areas of sericitic alteration; however, much of the geology is obscured by soil and vegetation. Post processing methods are applied to these mixture maps to derive an image of vegetation density and to map the dominant plant communities of the region.

1. INTRODUCTION

GER imaging spectrometer (GERIS) data were acquired over an area of gold mineralization 10km south of Galconda, Northern Nevada in october 1987 by BP Minerals (now RTZ). Though the data were acquired primarily for mapping hydrothermal alteration that might be associated with gold mineralization, the goals of this study were broader. The aim was to map the rock types, soils and vegetation communities of the area as much of the region is covered by varying amounts of soil and vegetation that obscure the underlying geology. An understanding of the controls on the distribution of the soil types and vegetation communities is central to assessing the effectiveness of imaging spectrometry in mapping the geology of semi-arid areas. Also the distribution of these vegetation communities is of interest to ecologists and soils to geomorphologists. Only the 24 bands in the infrared from 1.97 to 2.366 microns were used as in this wavelength range many minerals have diagnostic absorption features, as do many components of the vegetation canopy (Elvidge 1988). The seven wavelengths longword of 2.366 microns were not used because of their extremely low signal to noise ratio. This imagery has a nominal spatial resolution of 20 m.

2. GEOLOGY, SOILS AND VEGETATION

The study area is semi-arid receiving about 200mm of precipitation per year. Being at an altitude of between 5000 and 7000 feet, some of this precipitation occurs as snow during winter months. The geology of the area is complex. It has suffered four periods of intense compressive deformation between middle to late Paleozoic and late Cretaceous to early Territory. No detailed mapping of the area has been undertaken though a regional survey has been conducted (Willden 1964). Two outcropping units are recognised. The Cambrian Preble Formation, which consists of limestones and phyllite/quartzite units covers the vast majority of the study area. A minor outcrop of the Ordovician Valmy formation is found in the north east of the region. This unit is a dolomitic limestone that includes some shale beds. During the last phase of compressive deformation a large granodiorite body was intruded that underlies the study area. Copper and gold mineralization that is thought to be related to this phase of plutonism is found in many areas and is loosely associated with rocks that have been hydrothermally altered to sericite with minor kaolinite.

Outcrops of the above mentioned rocks are mainly found outcropping on ridges. The valley sides and floor being covered by soil which is commonly underlain by alluvium. The composition of this soil (a fine clay) is similar over all rock types. Its origin is thought to be due to aeolian dust fallout derived from the surrounding playas of paleo-lake Lahontan that dried up during the early Holocene (Chadwick and Davies 1989).

The vegetation cover of the area varies between 0 and 100%. This large variation in density is due to the presence of three plant communities. In the major river valleys that drain the Sanoma Range and dissect the study area from east to west is found the annual shrub community. This community occurs in well watered areas, usually in the immediate vicinity of rivers. The dominant species of this community are Dwarf Willow and Rosaceae. In the summer months this community provides a dense canopy (90 to 100%) of active green leaves (from now on termed active vegetation). The chaemophytic shrub community is found in highly variable density throughout the rest of the region. High densities occur in areas with high moisture availability and a well developed soil, with lower densities found in less favorable areas. The dominant plant of this community is Sagebrush though species such as Indian Paint Brush, Mormon Tea and Snakeweed are commonly found. As well as exhibiting a highly variable density these plants have a much lower pecentage of green leaves in their canopy; many branches, dead leaves and dead branches are visible and in most cases the canopy is sparse enough so as to expose the underlying rocks and soils. Finally the graminaceous community is found intermingled with the chaemophytic shrub community at low densities. These grasses also occur as dense patches (100%) in fertile areas where other plant communities are absent. The absence of other species in these fertile areas is probably due to human clearance and/or the effects of previous fires. During the summer these grasses quickly die off and present a canopy of 100% or less of dry vegetation.

It is clear from the description above that in this area the majority of pixels will be mixtures of vegetation, soil and the rock types of the region. Also in some cases the vegetation canopy itself is a mixture of different plant materials. A method of data analysis that takes these factors into account is needed. To do this linear mixture modelling has been implemented (Smith et al. 1985; Drake and Settle 1989; Settle and Drake 1990). Before applying this technique however it is necessary to calibrate the data and reduce the effects of noise.

3. CALIBRATION AND NOISE SUPPRESSION

Imaging spectrometry data in the 2.0 to 2.4 micron region is inherently noisy due to the low output of energy from the sun and the narrowness of the spectral bands, which for the GERIS is 16.5nm. Mixture modelling techniques are susceptible to noise and a method of noise suppression is needed. To determine the influence of noise on individual spectra, however, it is first necessary to calibrate to reflectance by removing the atmospheric continuum. Atmospheric modelling could not be used due to the lack of any ancillary information. Also, the empirical line method (Elvidge 1988) could not be used due to a lack of bright and dark homogeneous objects. Consequently the Log-residual method was employed (Mackin *et al.* 1988).

Log-residuals is essentially a method of statistical continuum removed, it is assumed that the only constant component in each pixel is the atmosphere. As outlined by Mackin *et al.* (1988) the method has advantages and disadvantages. The advantages are that no ancillary data is required, it is computationally simple and produces an image of relative reflectance. The method, however, can introduce spectral 'artifacts' that could be interpreted as absorption features but are in fact spurious. Also if the scene is relatively homogeneous (ie. it consists of largely one cover type), the spectral response of this cover type will be removed because it will contribute to the perceived continuum. Though these problems can occur they were not noticed in this scene due to the large amount of spectral variation.

The spectral noise was suppressed by use of a five point weighted moving average filter on each image spectra (Mackin *et al.* 1988). It was found that to provide a single pixel spectra that is largely free of noise three passes of the filter were needed. Fig. 1 shows a single pixel sericite spectra before and after this filtering procedure, a large improvement in data quality is evident. Though the method is effective in suppressing spectral noise, visual analysis of single bands of this filtered imagery showed that little or no spatial noise had been suppressed. If the image is thought of as a series of spectroradiometers for each pixel the effect of filtering is such that each spectra is largely free from noise but each spectroradiometer is poorly calibrated to its neighbor so noise in the spatial domain is still evident. This spatial noise was found to affect the mixture modelling procedure so a medium filter was ap-

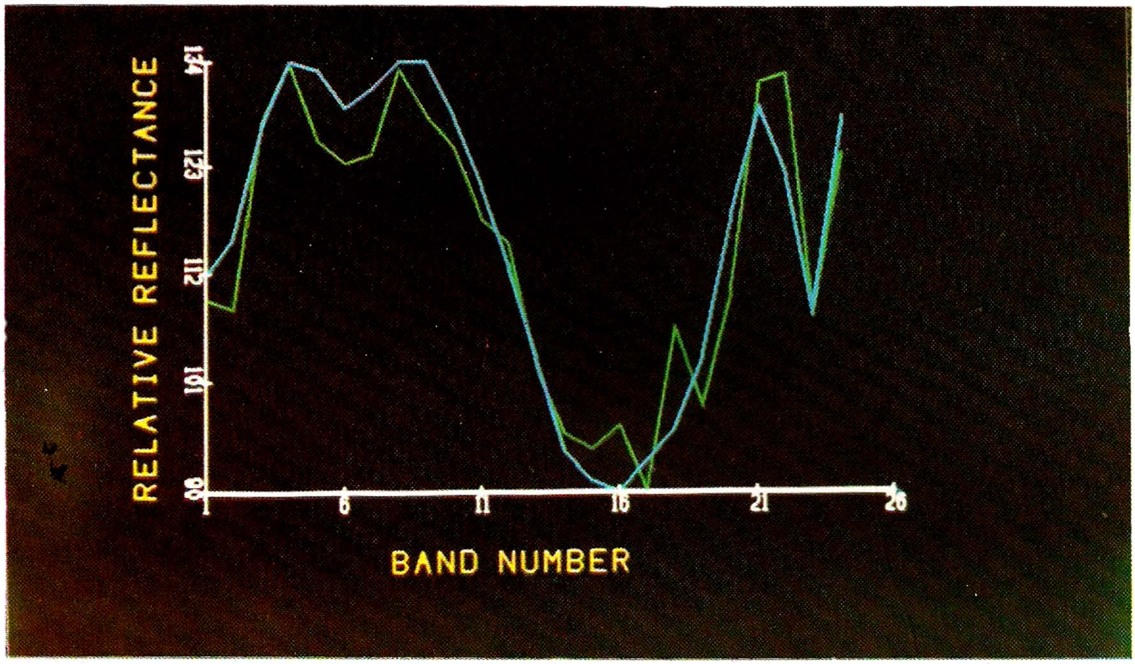

Fig. 1. The effect of a five point weighted moving average filter on a single pixel spectrum of sericite. Green is the original spectrum, while the blue spectrum has had three passes of the filter.

plied to each band of the GERIS imagery in order to reduce it. It was found that a single pass of the median filter was sufficient to reduce the noise to a level where the mixture modelling was largely unaffected by it.

4. LINEAR MIXTURE MODELLING

4.1 *Assumptions*

Mixture modelling aims to map the amount of each component in each pixel of the image. To apply a mixture model it is necessary to make a number of assumptions. The most important assumption involves the method in which mixing of sub-pixel components occurs. Additive (linear) mixing, is assumed here. This occurs when the components are distributed as patches that are large enough to allow the light to interact only with a specific patch of each component. More complex mixture models attempt to explain first and second order multiple scattering between components (Smith *et al.* 1985). The second assumption is that all the components in the image have sufficient spectral contrast to allow their separation (ie. they all have a diagnostic spectral response in the wavelength region being investigated). It is also assumed that the number and identity of these components can be defined in some way. Finally the pure examples of each component in the mixture must be known in order to fit these pixels to the imagery and estimate their proportions in each pixel. These pure pixels are known as endmembers, a common term in compositional data analysis. Though these endmembers are the purest example of each component in the image they may in fact be complex mixtures. Dry grass for example is an endmember in this image, however, dry grass is a complex mixture of lignin, cellulose and other minor components. The validity of the first two assumption images can only be speculated at unless two models are compared. However, it is possible to determine the factors lying behind the last two assumptions (Drake and Settle 1989; Smith *et al.* 1985).

4.2 *Finding and identifying components and their endmembers*

Smith *et al.* (1985) showed that the principal components analysis can be used to determine the number of spectrally distinct components in the mixture and their respective endmembers. The method involves plotting scattergrams of the significant principal components (those that are not simply noise). The number of components in the image equals the number of extremities in the scattergram and the endmembers lie at the end of these extremities. The method is empirical but relies on the fact that endmembers are pure and will thus lie at the extremes of any distribution in feature space, while mixture pixels will

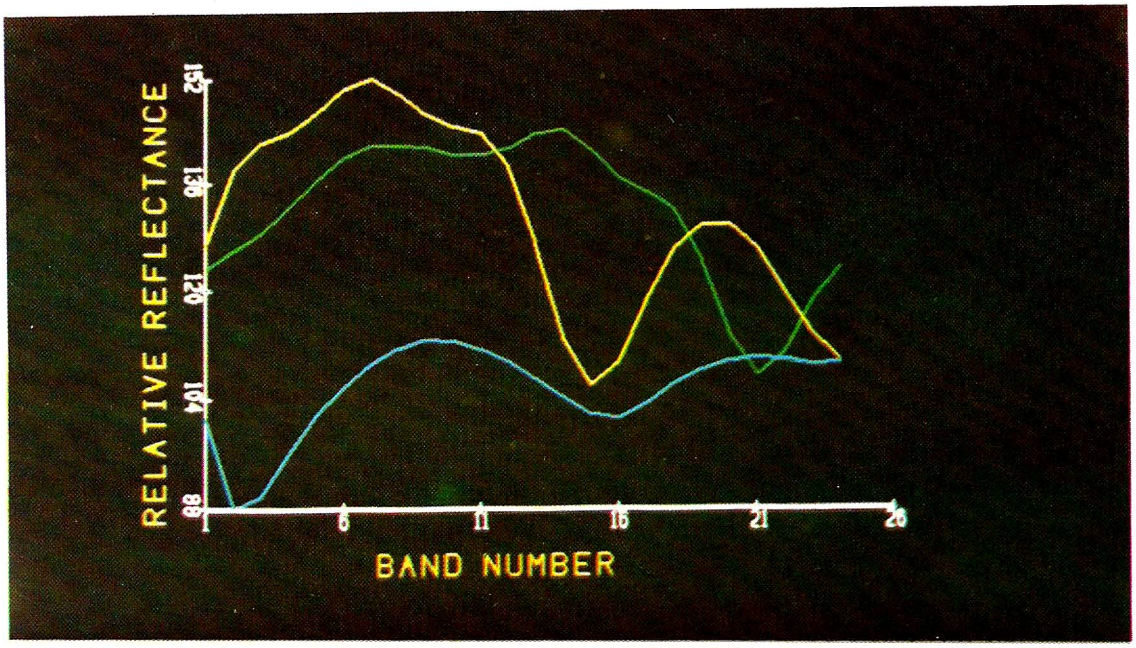

Fig. 2. Geological endmember spectra. The yellow spectrum is sericitic alteration, the green spectrum dolomitic limestone and the blue spectrum soil.

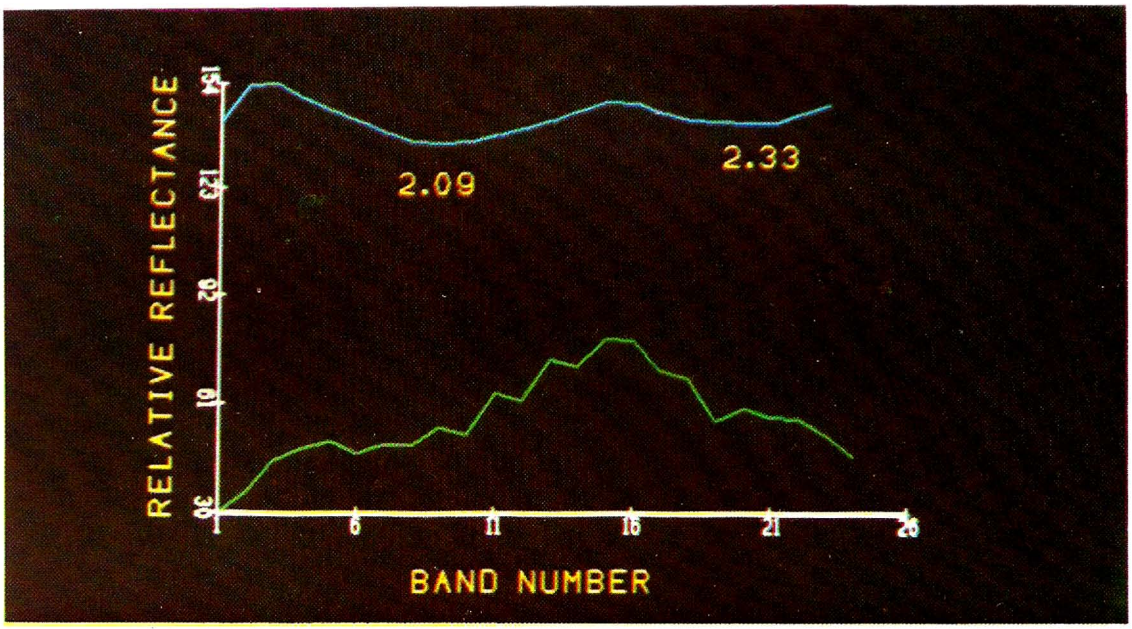

Fig. 3. Vegetation endmembers. The green spectrum is active vegetation and the blue spectrum dry vegetation.

lie in the region defined by these endmembers. The method has been validated on both linear and non-linear mixtures (Smith *et al.* 1985). It was applied to the GERIS imagery and though it was found to be affected by systematic noise it was otherwise successful. Five endmembers were defined, their spectra are shown in Figs. 2 and 3.

Using imaging spectrometry it is possible to remotely identify the components in the mixture by interpreting their spectra. The yellow spectrum in Fig. 2 has an absorption feature at 2.201 microns and is due to hydrothermal alteration of the phyllites to sericite. The green spectrum has an absorption feature at 2.316 due to dolomite and calcite in the dolomitic limestone of the Valmy formation.

The blue spectrum corresponds to soil, the small absorption feature at 2.235 is due to a complex mixture of clays currently being identified by XRD. The final two endmembers shown in Fig. 3 come from areas of dense vegetation. The green spectrum in Fig. 3 is from a patch of dense active vegetation in an area of the annual vegetation community, and the blue spectral is from an area of dry grass. The spectrum response of vegetation is poorly understood compared to those of rocks and minerals, however, work by Elvidge (1988) suggests that the absorption features at 2.09 and 2.33 are due to ligno-cellulose absorptions and the curve of the green vegetation is due to water in the green leaf cells. The reason why there are three plant communities but only two vegetation endmembers is discussed and evaluated in sections 6 and 7.

4.3 *The Linear Mixture Model*

The linear mixture model can be applied in a number of ways (Settle and Drake 1990). The classical or least squares estimator is used here where the following function is minimised:

$$(X - Mf)^T (X - Mf) \tag{1}$$

the symbol X is a pixel spectra, M a matrix whose columns are endmember spectra and f a vector of unknown proportions. This least squares approach is subjected to two constraints, the first being:

$$f_i > 0 \quad i = 1 \ldots c \tag{2}$$

where c is the number of components. This simply says that negative proportions are not allowed. The second constraint is:

$$f_1 + f_2 \ldots + f_c = 100\% \tag{3}$$

This constraint states that the pixels cannot add to more or less than a hundred percent. It implies that we have knowledge of all the components in the pixel. The advantages of applying these constraints are outlined in Settle and Drake (in press).

5. RESULTS

Figs. 4 and 5 show the results of the mixture modelling. Fig. 4a is the sericite mixture map. This map depicts the various intensities of hydrothermal alteration, where it is not obscured by soil or vegetation. The majority of these outcrops, and certainly the highest proportions estimates are found on hill or ridge tops where soil and vegetation is sparse. Notwithstanding this a number of new areas of alteration have been defined that may be associated with mineralization. The dolomite mixture map outlines outcrops of the Valmy formation where it is not obscured by soil or vegetation however, this mixture map has a number of small areas of misclassification associated with systematic scanner noise. The soil mixture map outlines its distribution to a large extent, the large bright area in the middle of the image for example is an alluvial gold working where the vegetation cover has been removed and 100% soil exposed. This map, however, also outlines those rock types of the Preble formation that like soil, have a small absorption feature near 2.235 and also rocks that have no absorption features but a similar reflectance, such as phylliles and the limestone (the lack of absorption features in this limestone could be due to organic matter (Crowley 1986). The soil mixture map therefore is a composite mixture map of a number of materials that have a similar spectral response. They can not be separated by mixture modelling because the assumption that all components have a unique spectral response is violated. The high proportions estimates for green vegetation mixture map in Fig. 5a depict the areas of the annual vegetation community in the valley floors, with much lower proportions in areas of the chaemophytic shrub community. The dry vegetation mixture map is shown in Fig. 5b. The highest proportions of dry vegetation appear to be related to those areas were the graminaceuos community dominates though relatively high proportions are also found in areas where the chaemophytic community dominates. These high proportions seem to be due to both the understory of dry grass and the dry matter that is exposed in the plant canopy. To determine this it is necessary to look at the spectral response of the different parts of the plant canopy of the species that are found in this community.

6.THE SPECTRAL RESPONSE OF PLANT MATERIALS

The spectral response of many of the species found in this region and of some of the common chemical components of plants has been studied by Elvidge (1988) and Elvidge (unpublished data). A summary of this work is shown in Fig. 6. The spectral response of green leaves for five species of

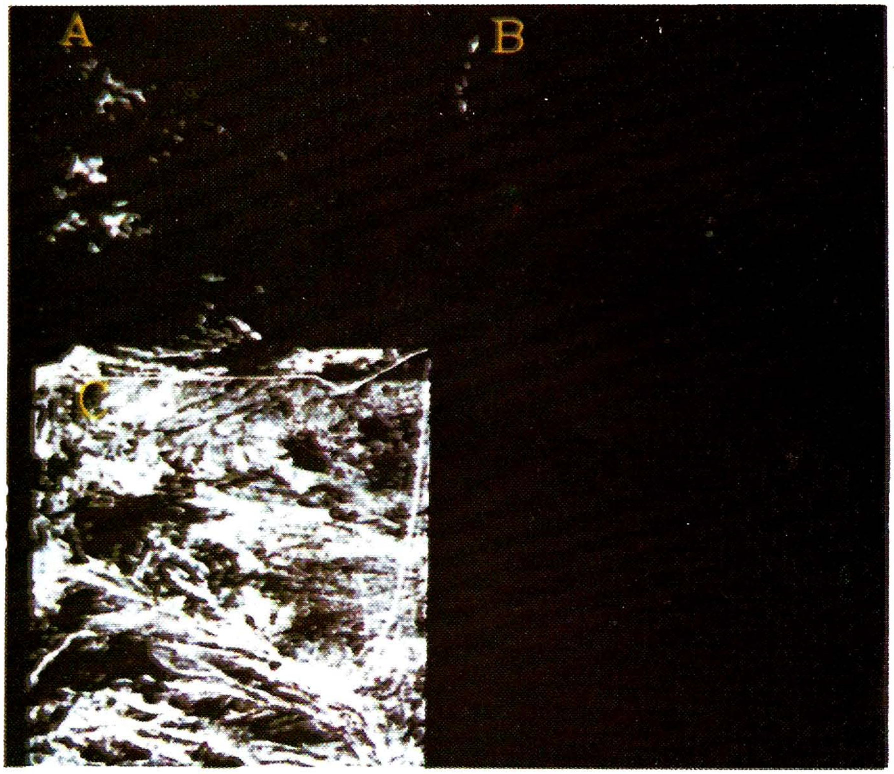

Fig. 4. Geological mixture maps. A is sericitic alteration, B is dolomitic limestone and C soil.

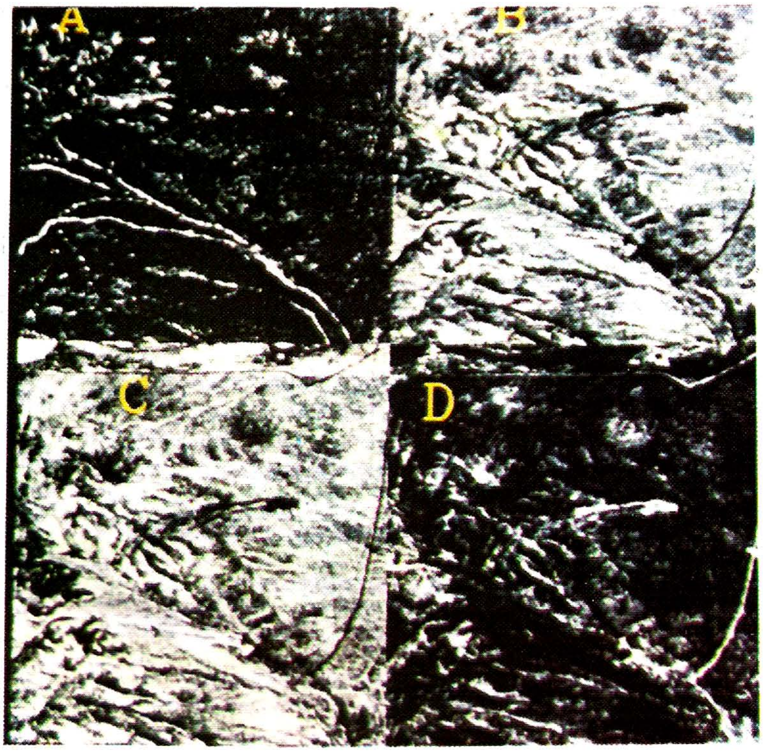

Fig. 5. A is the active vegetation mixture map, B the dry vegetation mixure map, C the vegetation density and D a map of bare ground.

plant in the 0.4 to 2.5 micron region is shown in Fig. 6a, their spectra are very similar at all wavelengths, in the 2.0 to 2.4 region all show a peak at 2.2 microns. This is due to liquid water in the cells of the leaf, as is illustrated by the spectra of wet cotton cellulose in Fig. 6d. The spectral response of the different dry plant materials of big sagebrush are shown in

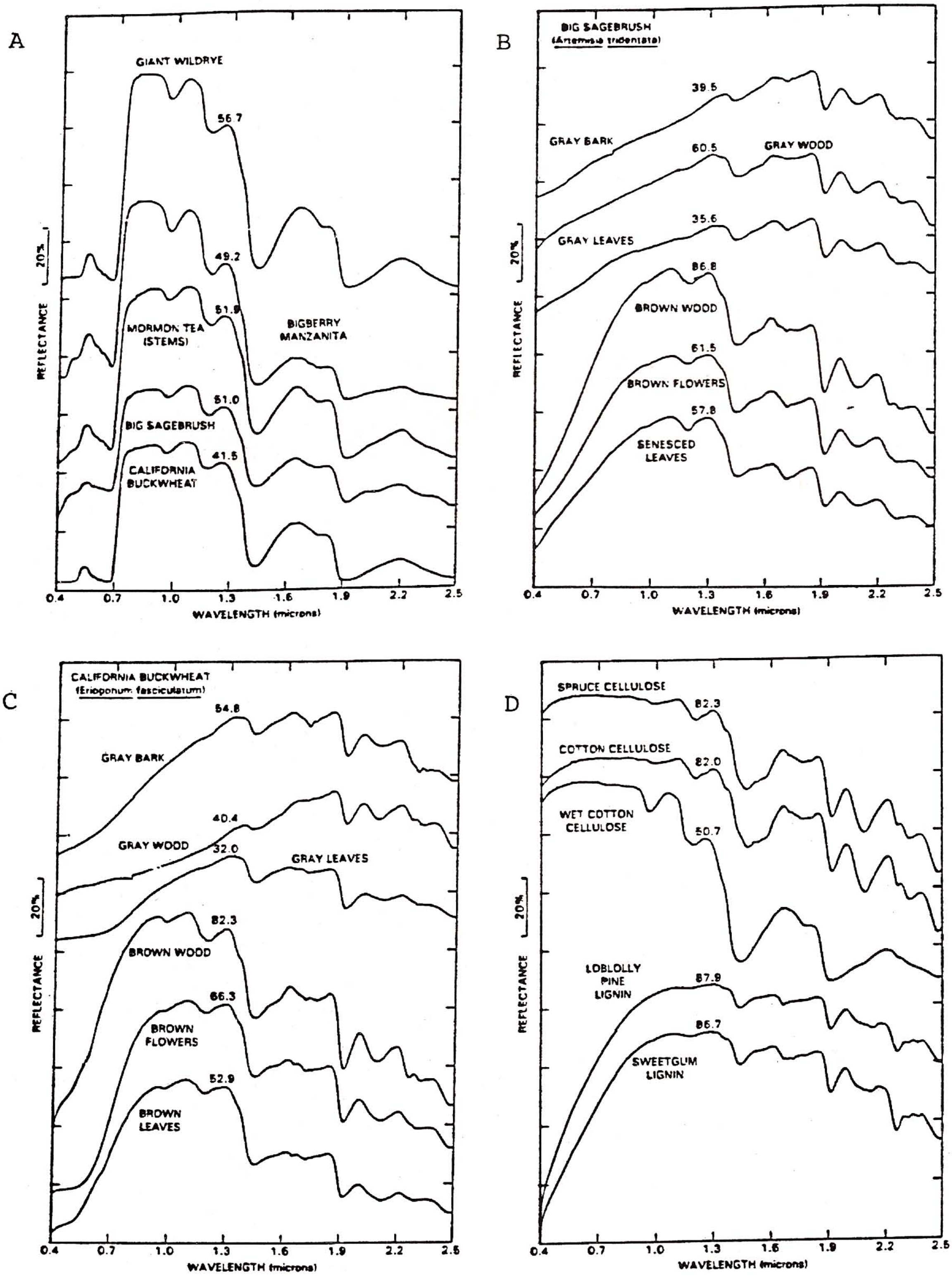

Fig. 6. The spectral response of active and dry plant materials for some of the species found in the region. Also shown are the spectra of some plant constituent materials. After Elvidge (1989) and Elvidge (unpublished data) .

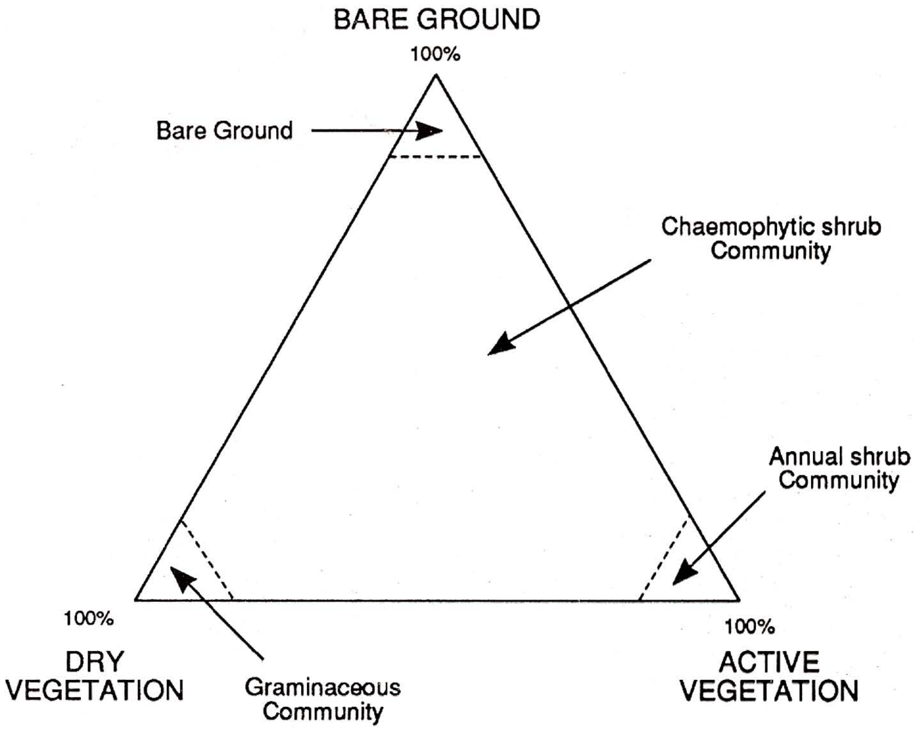

Fig. 7. Schematic ternary mixture space distribution of the vegetation communities found in the region.

Fig. 6b. Their spectral response varies in the visible and near infrared where plant decay causes the breakdown of the lignin absorption that originates from an intense absorption feature in the ultraviolet (Elvidge 1988). Their spectral response in the 2.0 to 2.4 micron region, however, is similar. All dry plant materials exhibit absorption features at 2.09 and 2.33 which are identical to those of the dry vegetation endmembers. Some of these materials also have a small absorption at 2.27 microns which are not evident in the imagery. This is a bit puzzling but it may be that a nominal spectral resolution of 16.5 nm is not sufficient to resolve this absorption feature. These absorption features are termed ligno-cellulose absoption by Elvidge (1988) as lignin and cellulose exhibit these absorptions (see Fig. 6d) and these materials comprise 80 to 90% of the dry weight of most plant materials.

The spectral response of the different dry plant materials from California Buckwheat, one of the grasses found in the study area is shown in Fig. 6c. These exhibit the same absorptions as the dry sagebrush materials. It appears therefore that the spectral response of plants is not controlled by species but by the phenological state of the vegetation and the amount of leaves and dry plant materials exposed in the canopy. This helps to explain why though there are only two vegetation endmembers there are three vegetation communities.

7. POST-PROCESSING OF MIXTURE MAPS FOR THE DERIVATION OF ECOLOGICAL PARAMETERS

The distribution of dry and active plant materials in the vegetation canopies of the region, as mapped by the mixture modelling, may be of use to ecologists, however, there are more important parameters that can be estimated from their distribution. The most simple one is to compute the vegetation density which can be derived from adding together the green and dry vegetation mixture maps. This vegetation density map is shown in Fig. 5c. In this relatively simple situation where only three vegetation communities exist it is also possible to map where the different plant communities are dominant. This classification relies on the fact that during the dry season each plant community has a variable but known amount of dry or active (green) material in its canopy and a variable but known density, and thus a known amount of rock or soil exposure. If the rock types and soil mixture maps are added together a map is derived that depicts the amount of bare ground, this map is shown in Fig. 5d. These vegetation communities can now be thought of in terms of their distribution in the ternary mixture space of bare ground; green vegetation and dry vegetation shown in Fig. 7. Where each community is dominant it occupies a specific region of this mixture space. A simple decision rule classifier can be imple-

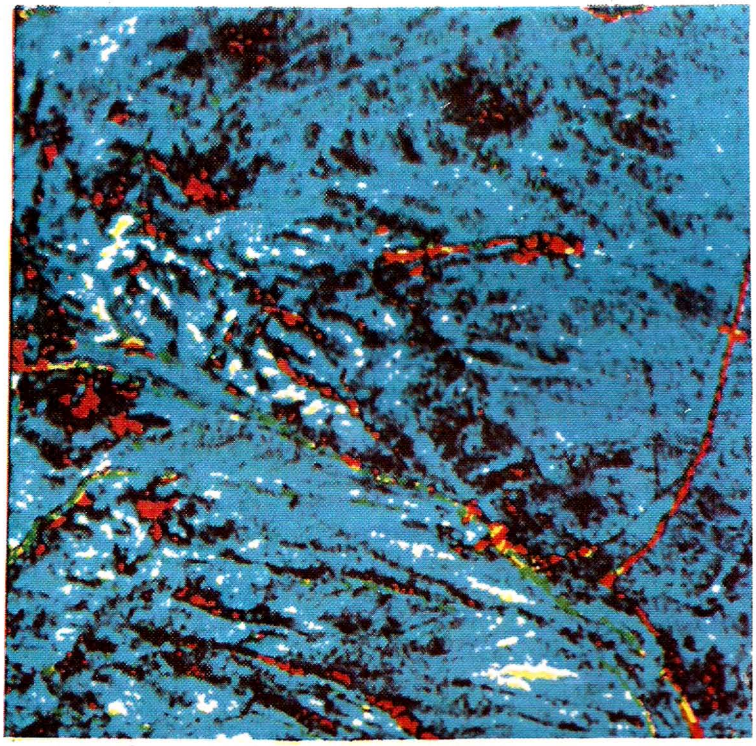

Fig. 8. Vegetation classification map. Brown is bare ground, yellow corresponds to regions were the graminaceous community is dominant, green to regions were the annual shrub community is dominant and the various shades of cyan to grey depict the distribution and density of the chaemophytic shrub community.

mented to classify each community. If an area has 90% green vegetation for example the annual shrub community will be dominant. A similar threshold on the dry vegetation will outline where graminaceous community is dominant. A rather arbitrary threshold of 90% has been used to outline areas of dominantly bare ground.

The chaemophytic shrub community is highly variable in its density and also exhibits variations in the amount of dry and green material in its canopy, it lies in the large region of this ternary mixture space not defined by the other decision rules. The classified map resulting from this mixture space interpretation is shown in Fig. 8. In producing this classified map the classes that are relatively pure as the annual and graminaceous communities and the bare ground have solid colouring of green, yellow and red respectively. As the region where the chaemophytic shrub community dominates exhibits large variations in density, the vegetation density map has been used as a backdrop. This community is displayed in various shades of cyan where grey indicates a low density and bright cyan a high density.

It is clear that linenar mixture modelling combined with field work to determine the vegetation phenology and the various components in the plant canopy is an effective method of deriving ecological parameters from imaging spectroscopy data. This method of data analysis however is still experimental and needs to be valided with ground reference data of field proportions estimates, this validation is currently being undertaken.

8. CONCLUSIONS

Geological applications of imaging spectroscopy in this semi-arid region are restricted by the development of a soil that blankets the valley floor and shallower slopes. The major areas of rock outcrop are found on steeper slopes, hill and ridge tops. The soil is a fine material that originates from the playas of the region. Vegetation grows on the soil even where it is poorly developed and contributes in a large part to the obscuring of the underlying geology. Notwithstanding this many of the major formations of the region have been identified and mapped using linear mixture modelling techniques. The method, however, only provides limited geological discrimination as some rock types have a similar response to the soil and thus are confused with it. Hydrothermal alteration of these rocks to sericite is readily identified and mapped and some new areas of sericitic alteration have been discovered.

There is much information on the vegetation cover and type to be gained by studying the 2.0 to 2.4 micron region. Green and dry plant materials have distinctive spectral response due to water and ligno-cellulose absoptions. Using mixture modelling and post-processing techniques it seems possible to map not only the distribution of these green and dry plant materials but also to estimate the vegetation density and the distribution of the three vegetation communities of the region.

ACKNOWLEDGMENTS

Thanks to BP Minerals (now RTZ) particularly E. Lockhart for providing the imagery and valuable field information, S. Mackin for his assistance in noise removal and calibration, J.J. Settle for helpful discussion on mixture modelling and T.J. Munday whose interest and help made this project possible. The author acknowledges a Reading University research endowment fellowship. I also wish to thank C. Elvidge for helpful discussion on the spectral response of vegetation and for providing copies of the spectra of plant materials.

REFERENCES

Chadwick, O.A. and Davis, J.O., 1990, "Soil-forming intervals caused by eolian sediment pulses in Lahotan basin, northwestern Nevada", *Geology*, 18, 243-246.

Crowley, J.K., 1986, "Visible and near-infrared spectra of carbonate rocks: reflectance variations related to petrographic texture and impurities", *J.Geophys. Res.*, 91 b5, 5001-5012.

Drake, N.A., Settle, J.J., 1989, "Linear mixture modelling of Thematic Mapper data of the Peruvian Andes", *Proc. 9th EARSEL, Symposium,* Espoo, Finland, 27 June - 1 July, 490-495. EUR12827EN.

Elvidge, C.D., 1989, "Vegetation reflectance features in AVIRIS data", *Proc. 6th Thematic Conf. on Remote Sensing for Exploration Geology,* Houston, Texas, May 16-19.

Mackin, S., Munday, T.J. and Hook., S., 1987, "Preliminary results from an investigation of AIS-1 data over an area of epithermal alteration: Plateau, northern Queensland, Australia", Vane, C., (ed) *Proc. 3rd Airborne Imaging Spectrometer Data Analysis Workshop,* June 2-4. JPL publication 87-30.

Settle, J.J. and Drake, N.A., 1990, "Linear mixing and the estimation of ground cover proportions", *Int. J. Remote Sensing* (in press).

Smith, M.O., Jonson, P.E. and Adams, J.B., 1985, "Quantitative determination of mineral types and abundances from reflectance spectra using principal components analysis", Proc. Lunar Planet. Sci. Conf. 15, *J. Geophys Res.*, 80, 797-804.

Willden, R., 1964, "Geology and mineral deposits of Humbolt County, Nevada", *Bull. Nevada Bureau Mines and Geol.*, 59, 154pp.