Dr. Eugene L. Maxwell
Colorado State University
Fort Collins, CO 80523

Multivariate System Analysis of Multispectral Imagery

Multivariate analysis methods were used to evaluate variables, preprocessing results, and classification accuracy.

Introduction

This research project investigated the feasibility of providing useful information for rangeland managers from remote sensing imagery, in particular, imagery obtained by the Earth Resources Technology Satellite (ERTS). A systems analysis was used to gain a greater understanding of the parameters governing the relationships between actual ground conditions and the imagery obtained by the satellite. This led to data preprocessing, including ratioing of spectral bands, filtering, and transformation of variables. This in turn improved the interpretation of the data in terms of range conditions.

Abstract: *Remote sensing, particularly from a satellite, is potentially an effective and economic means to gather information about natural resources. Routine applications, however, have been fraught with many difficulties and disappointing results. A system analysis of remote sensing as a source of information has revealed the sources of many problems. More importantly, these analyses have shown methods for data preprocessing which should greatly improve results. The application of ERTS data to rangeland management problems was the ultimate goal of this research.*

Several noise sources (causing random fluctuations of radiance values) were identified as significantly degrading the quality of ERTS data. These included source (scene) noise, atmospheric propagation changes, radiometric errors, electronic system noise, data processing noise, and data analysis errors. Systems analysis suggested several data preprocessing methods to improve data quality. Preprocessing of data included (1) cleaning of data, (2) ratioing of variables, (3) transformation of variables, and (4) filtering.

Multivariate analysis methods were used to evaluate variables, preprocessing results, and classification accuracy. The most significant variables for rangeland analysis were the ratio of ERTS band 7 to band 5 and band 5. Cleaning of training data greatly reduced classification errors by more accurately determining class signatures. Filtering reduced classification errors for vegetation types from 18.8 per cent to 3.6 per cent. The two-dimensional moving average filter was particularly effective in reducing atmospheric fluctuations and noise introduced by the satellite sensor system. Canonical transformation of variables eliminated correlation between variables, concentrated between groups variance in the first canonical variates, but did not significantly improve classification results.

To support the systems analysis, a field measurement program was undertaken to

Photogrammetric Engineering and Remote Sensing,
Vol. 42, No. 9, September 1976, pp. 1173-1186.

provide basic information on the ground and atmospheric conditions at the time of each ERTS overflight. Color photographs of surface and atmospheric conditions, and quantitative data for vegetation cover and biomass, were also obtained.

A companion paper (Maxwell, 1976) reports on the management application aspects of the project. This paper concentrates on the multivariate system analyses and the implications of the results therefrom.

SYSTEM ANALYSIS DEFINED

A general definition of systems analysis is provided by Quade (1968). "A systematic approach to helping a decision maker choose a course of action by investigating his full problem, searching out objectives and alternatives, and comparing them in the light of their consequences, using an appropriate framework—insofar as possible analytic—to bring expert judgment and intuition to bear on the problem." This definition of systems analysis does not, of course, define what we mean by the word "system." For our purposes, a system can be defined as an arrangement of components and parameters which are interrelated, and related to outside components and parameters, such as to form a complete functioning entity. The development of a *model* of the system and the analysis of that model will usually provide greater insight into the functioning of the system and its various components.

NEED FOR SYSTEMS ANALYSIS

If the data obtained from remote sensing systems provided an accurate, undistorted measure of the reflectance of materials and scenes, there would be little need for systems analyses. In fact, remote sensing data is a measure of the radiance from a scene which has been modified by atmospheric transmission, noise from a variety of sources, and measurement error. Furthermore, when an attempt is made to identify the radiance from a specific scene type (e.g., wheat) it is found to be a function of type variations, spectral variations of incident energy, solar zenith angle, soil type and conditions, slope and aspect, look angle, and other factors. This data variability has produced disappointing results from many attempts to use remote sensing data and is a major cause of signature extension problems.

The application of systems analysis techniques can provide greater insight into the sources of data variability. More importantly, systems analyses will lead to improved data processing, reduction of noise, and more accurate interpretation of data in terms of scene characteristics. Ultimately, a better understanding of remote sensing systems should result in the design of better hardware and software.

EXPERIMENTAL PROGRAM

This section presents a brief description of the field site, the field measurements, and the data reduction and refining methods. This will provide an overview for reference when considering the more detailed information to follow.

THE FIELD SITE

The test fields used on this effort are located within the Pawnee National Grasslands in northeastern Colorado. This is the location of the International Biological Program (IBP) Grassland Biome Site. The vegetation at this site is typical of the shortgrass prairie of the Great Plains. Blue grama *(Bouteloua gracilis)* is the dominant species of grass. Six areas were closely monitored to obtain ground verification data for canopy cover, biomass, and phenology. These areas include a heavily grazed blue grama field (HBOGR), a lightly grazed blue grama field (LBOGR), a pitted blue grama field (PITTED), a western wheatgrass swale (ASWALE), a crested wheatgrass field (CRESTD), and a fourwing saltbush field (FRWING). In addition to these closely monitored fields, a sandy arroyo (SAND) was included because it contained virtually no vegetation. Wheatfields (WHEAT) and fallow ground (GROUND) lying between the fields were also included since they are commonly found near natural grassland regions.

FIELD MEASUREMENTS

A systematic sampling of the six grassland test sites was undertaken for the July 10, July 28, and August 15, 1973 ERTS overpass dates. Circular quadrats, 1000 square centimeters in area, were used in a double sampling procedure to estimate canopy cover and green standing crop biomass for each species present. The double sampling procedure involved ocular estimates of canopy cover and green biomass within each quadrat, and clipping and weighing of all vegetation in every fifth quadrat. A regression analysis was then used to correct the ocular estimates of biomass. Twenty quadrats were sampled in each of three stands for each of the six grassland test fields.

SELECTION OF ERTS TRAINING DATA

The computer classification of an ERTS scene requires that the signature (mean vector and covariance matrix for the maximum likelihood method) for each of the classes used must be available for programming (training) the computer to recognize the classes. Supervised classification procedures, such as those used here, require the selection from image data of a representative sample of each class population (training data) to be used for calculation of mean vectors and covariance matrices. For range management applications it was desired to classify areas in terms of both vegetation type (ecosystems) and quantity of standing crop (biomass).

The selection of training data involved first the location of each of the test sites on a computer generated graymap of the Pawnee portion of an ERTS image. From two to five training fields (subsets) were selected from each test field (class) for each of the three image dates. The number of pixels within the training fields for each class and each date varied from 30 to 100.

The formation of training data for biomass classes was based on the field measurement of biomass at the time of each ERTS overpass. Actually, the training data previously selected for recognition of vegetation types were regrouped into biomass classes as shown in Table 1. Notice that data from different vegetation types and different dates have been combined to form these biomass classes.

System Descriptions

Three system descriptions of components and parameters are given to provide the relationship between grassland conditions and a multispectral image. The first of these descriptions is an attempt at a more or less realistic, physical description of the system. The second makes use of parameters and components more readily observable by the system analyst. The third system, which was actually analyzed, is greatly simplified and includes only the components and parameters generally available to most researchers.

PHYSICAL SYSTEM DESCRIPTION

A diagram of the physical system, which is composed of several subsystems, is shown in Figure 1. This diagram shows the incoming radiation penetrating the atmosphere, which includes occasional clouds and haze. The clouds and haze cause spectral variations in the amount of energy reaching the earth's surface. The irradiance of the plants and soil also is a function of the angle of incidence of the incoming radiation, which is affected by topography as illustrated, and by time of day and season which has not been illustrated. Each of these factors will affect the reflected energy, as will the different soil types and plant species.

The reflected radiation must again penetrate the atmosphere, including various amounts of cloud and haze. Once the electromagnetic energy reaches the satellite, an optical system establishes the spectral and spatial resolution. As indicated in the exploded view of the ERTS system, the multispectral scanner (MSS) divides the electromagnetic spectrum into four bands. Each of these bands of energy are detected by one of six sensors. On ERTS, six lines are scanned simultaneously, requiring six sensors per band. The different sensitivity or calibration of these sensors produces radiometric errors which often results in a horizontal striation on the images. This has the effect of adding noise to the reflected energy, if we define noise as any apparent variation in reflectance which is produced by something other than a change in the surface features of interest.

TABLE 1. FORMATION OF BIOMASS CLASSES. (WET GREEN BIOMASS).

Biomass Class	lbs/acre	Veg. Classes Included	Image Dates Included (1973)
ABLE	0– 100	SAND, GROUND	7/10, 7/28, 8/15
BAKER	100– 500	HBOGR, LBOGR PITTED, ASWALE	7/10
CHARLS	500–1000	HBOGR, PITTED	7/28, 8/15
DOG	1000–1500	LBOGR, PITTED HBOGR	7/28, 8/15
EASY	1500–2000	LBOGR, ASWALE	7/28, 8/15
FOX	2000–3000	FRWING	7/10, 7/28
GEORGE	>3000	FRWING	8/15

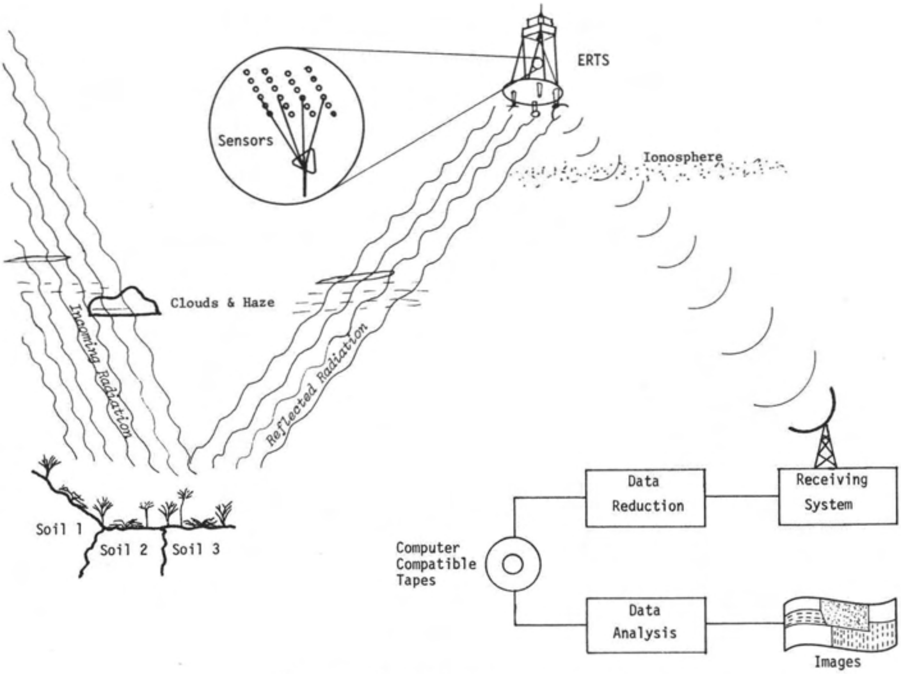


FIG. 1. Physical system diagram.

The energy which is detected and processed within the satellite is next transmitted to earth for detection and processing by the receiving system and data reduction system. Propagation of the microwave energy from the satellite to the earth will be affected somewhat by the ionosphere, but the amount of noise added will probably be slight. There is also the opportunity for addition of system noise from the receiving and data reduction equipments.

It should be noted that noise is added in the data transmission system when the continuous range of reflected energies detected by the satellite is reduced to a seven bit digital signal for MSS bands 4, 5, and 6 and a six bit digital signal for MSS band 7. Reduction of a signal to these limited dynamic ranges results in rounding errors, which are a form of noise.

Finally, computer compatible tapes (CCT's) are produced, which are then analyzed and interpreted to form classification maps or images. This data analysis, including preprocessing of the data, can be used to reduce some of the noise which has been added to the original signals as noted above, but it also is likely to add noise of its own, particularly in the form of interpretation errors.

OBSERVABLE SYSTEM DESCRIPTION

All of the components and parameters illustrated in Figure 2 are not usually observed when using data from the ERTS system. With the use of additional sensors on the ground and in the satellite system, however, they could be observed. This would significantly improve the quality and quantity of information obtained from the ERTS system. It is worthwhile, therefore, to discuss this "potentially" observable system, prior to considering the actual system which we have at our disposal.

The incoming radiation is now represented as a radiation source and propagation effects are shown as parameters controlling the flow of radiation to the scene. As mentioned previously, propagation effects are spectrally dependent; therefore, the radiation energy is split into four paths. This propagation effect could be observed by a spectral radiometer measuring the incoming radiation in the same spectral bands as those received by the ERTS system. This has been done occasionally on special research projects.

The noise produced by surface variations such as topography, soil types, and vegetation changes is now shown as a source noise added to the reflected multispectral signal.

The extent to which we can observe and identify these source noises, which may also be spectrally dependent, will be discussed a little later. It may at times be difficult to decide whether such source variations are noise or desired signal.

Slater (1974) and Duggin (1974) have also investigated the limitations imposed on remote sensing system capabilities by source noise and atmospheric noise (visible propagation variations). Short term atmospheric variations produce a form of atmospheric noise in the frequency range from 0.1 Hz to 1 kHz. Slater (1974) has summarized the work of several others who show fluctuations in received radiant energy from 0 to 10 percent of the mean level, i.e., the coefficient of variation is 0 to 0.1. This higher frequency atmospheric noise is conjectured to be due to fast moving subvisual to barely visual high cirrus clouds, or to streams of particulate material in the atmosphere. Whatever the cause, the effects should be similar in magnitude and exactly in phase in all spectral bands, thereby affording the opportunity to reduce this noise by ratioing two of the spectral bands. Duggin's studies of variations in surface reflections (source noise) show coefficients of variation ranging from 0.05 to 0.11.

We have indicated the spectral correlation between the bands by the ellipse enclosing the four channels of information or signal flow. This correlation can be easily observed by computing the correlation between the four channels of data on the computer compatible tapes.

The effects of clouds, haze, and other atmospheric variations on the propagation of the electromagnetic energy from the ground to the satellite could probably be observed by monitoring the diffuse sky radiation coming from the direction of the satellite.

The intensity of the radiation incident on the satellite sensors may be expressed as

$$N_{\lambda_i C_i} = \frac{1}{\pi}\left[\rho\ (\lambda, C, \theta, \phi)\ H\ (\lambda, \tau, \theta)\ \tau\ (\lambda, \phi)\right] + N_a\ (\lambda, \phi) + N_n\ (\lambda, C, \tau, L) \qquad (1)$$

where

$N_{\lambda_i C_i}$ is the radiation intensity at wavelength λ_i for class C_i;

$\rho(\lambda, C, \theta, \phi)$ is the reflectivity of the scene which is a function of wavelength, class, solar zenith angle θ, and look angle ϕ;

$H(\lambda, \tau, \theta)$ is the scene irradiance which is a function of λ, θ, and the atmospheric transmittance τ;

$\tau(\lambda, \phi)$ is the atmospheric transmittance from the scene to the satellite;

$N_a(\lambda, \phi)$ is radiation from the atmosphere; and

$N_n(\lambda, C, \tau, L)$is source noise which is a function of scene location L as well as λ, C, and τ.

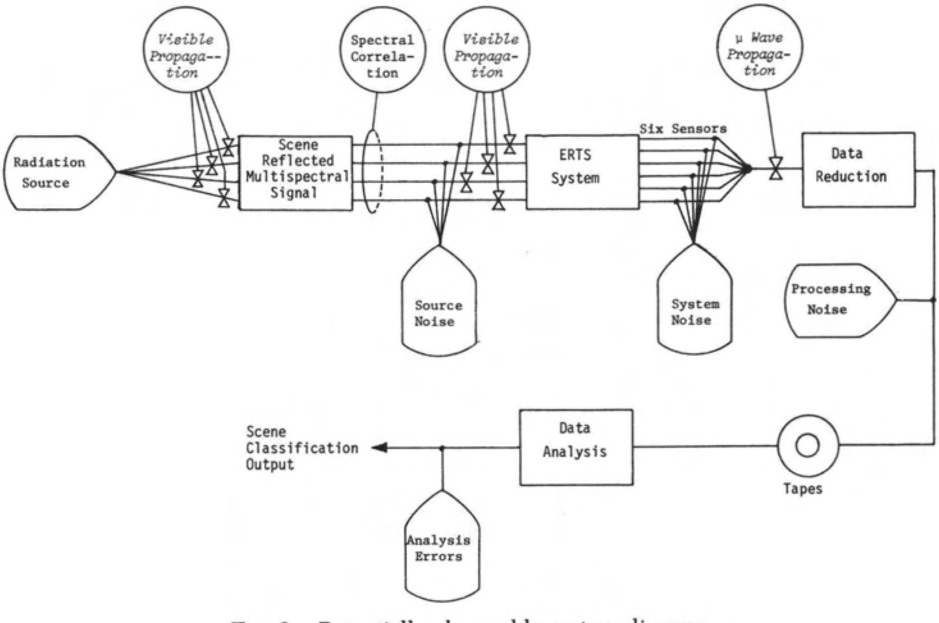


FIG. 2. Potentially observable system diagram.

The complexity of the situation is further increased by the satellite system.

In addition to modification by the electronic noise of the system, the signal from the sensors is a function of the specific sensor in use. These are called radiometric effects. Within the ERTS system, any noise added by that system is observable only when it is correlated with the selection of the six sensors.

We see, therefore, that we have source, propagation, system, and processing noises added to the multispectral signal. Radiometric system noise is easiest to separate since that noise can be identified with the six sensors. Source noise and processing noise may not be observable separately without the preparation of extremely detailed information on the actual scene variations. In the absence of that sort of data they will be treated as a combined noise which may be observed by noting the effects of various preprocessing schemes such as filtering.

Analysis errors, which may be the result of noise and propagation effects as well as actual malfunctions of the data analysis system, can be observed only when sufficient ground truth is available to accurately identify correct versus incorrect classifications.

In the final analysis, our measure of the effectiveness of any remote sensing system will be based on the quantity and usefulness of the information provided. The quantity of information transmitted by a communications system is given by

$$C = W \log (1 + S/N) \qquad (2)$$

where

C is the system channel capacity in bits/sec.,
W is channel bandwidth in Hertz, and
S/N is the signal-to-noise ratio.

Similarly, I suggest that the information capacity from an image might be expressed as

$$IC = Q \log (1 + S/N) \qquad (3)$$

where

IC is system image capacity in bits/pixel and
Q is a system quality factor.

The system quality factor Q will be related to such parameters as dynamic range, number of spectral bands, total bandwidth covered by all bands, and the correlation between bands. These parameters are included or implied in Figure 2.

This description of the potentially observable system will be particularly useful when considering data preprocessing methods. The processing of data, however, must also be related to the analyzable system shown in Figure 3.

ANALYZABLE SYSTEM DESCRIPTION

The analyzable system diagram shows four channels of signal, plus noise, entering a signal preprocessing system. For the research discussed here, two additional channels of information are obtained by ratioing two pairs of the original channels. In other words, the six channels of information coming from the signal preprocessing are not related in any way to the six sensors used in the ERTS system. Finally, we have a system for performing signal processing which produces a classification output. By noting the effects of signal preprocessing and by using multivariate system analyses, not only can we analyze the system in Figure 3, but we can learn a great deal about the system in Figure 2. This use of multivariate system analysis methods, to learn more about the

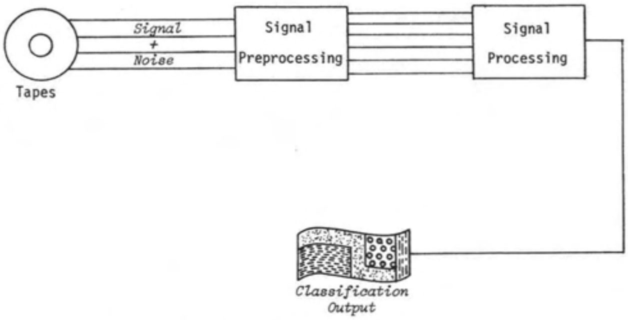


FIG. 3. Analyzable system diagram.

observable and physical systems and in the process to improve our classification results, is what this research and this paper are all about. Specifically, these systems descriptions were used to develop data preprocessing and data processing methods.

Preprocessing Methods

The potentially observable system shown in Figure 2 suggests the use of several preprocessing methods to reduce the effects of spectral correlation, atmospheric propagation changes, and the various noise sources. Methods considered include cleaning of training data; and ratioing, filtering, and transformation of all data.

CLEANING TRAINING DATA

Usually training data are obtained by simply selecting one or more rectangular training fields within a larger region, previously identified as a given class. This method ignores the possibility that some of the pixels within these rectangles may not be of that class or may be excessively noisy. The system may work reasonably well for homogenous classes, such as agricultural crops, but it breaks down completely for natural vegetation classes which are inherently heterogeneous.

It could be argued that this source noise (class heterogeneity) is really not noise but is part of the class signature. I agree when considering the tendency of specific species to occur in stands within an ecosystem, but I disagree when considering rocky outcrops in a grassy field or a clump of aspen in a spruce-fir zone. Obviously, cleaning of training data, removal of anomalous data, must be done with care and considerable insight. But it must be done.

The argument for removing data abnormally affected by the atmosphere, radiometric errors, and other system noise is more straightforward. It is obvious that if enough noise is added to any signal, all characteristics of the signal will eventually disappear.

Signature extension problems are certainly in part caused by noise of the types discussed here. I submit that proper cleaning of training data and compensation for satellite system changes and topography effects could greatly reduce signature variations.

The cleaning of training data was accomplished iteratively as follows. The mean vector and covariance matrix (signature) was computed for each class based on the original data from rectangular fields. Then the *posteriori* probabilities of each pixel belonging to each class were computed. Pixels were removed if they had a low probability of belonging to their original class and/or a high probability of belonging to another class. New signatures were then computed and more pixels removed, etc., etc. Usually 3 or 4 iterations were enough to provide adequate cleaning, which was indicated by high *posteriori* probabilities for the remaining pixels, and no pixels which met the criteria (probability thresholds) for removal.

RATIOING

Chlorophyll absorbs electromagnetic energy most efficiently at wavelengths of 0.4 to 0.5 micrometers and 0.65 to 0.69 micrometers. Furthermore, green vegetation has a very high reflectance coefficient in the near infrared from 0.75 to 1.2 micrometers. These unique absorption and reflectance characteristics for green vegetation have been investigated by Miller and Pearson (1971) and Tucker (1973). They have shown that a ratio of the near infrared and chlorophyll absorption bands is well correlated with the amount of green biomass within the scene. Since ERTS band 5 (0.6 to 0.7 micrometers) contains the region of strongest chlorophyll absorption and ERTS band 7 (0.8 to 1.1 micrometers) is a spectral band characterized by strong vegetation reflectance, it was considered that this ratio, 7/5, might be an important variable for biomass classification. Also, since ERTS band 4 (0.5 to 0.6 micrometers) does not contain either of the primary chlorophyll absorption bands, the ratio 5/4 might also be an important variable.

From a systems viewpoint (see Figure 2) the main advantage which ratios should give for vegetation classification is an improved signal to noise ratio. We note that an increase in green vegetation will result in an increase in the ratio, 7/5, which will be greater than the change in either of the original variables by themselves. On the other hand, this ratio should effectively reduce signal fluctuations caused by the effects of source noise, changes in atmospheric conditions from one image date to the next, and rapid atmospheric propagation variations along the scanline. For instance, changes in the intensity of the received signal due to propagation effects (clouds, haze, etc.) should be similar in each of the bands. The changes will not be identical in magnitude, since propagation effects are spectrum dependent, but they should be in-phase. Thus, a decrease in one band at a given point in time and space will be associated with a decrease in the other band and ratioing will result in some reduction in noise.

Also, changing soil and changing amounts of dead vegetation will, in general, cause a similar spectral reflectance change for all ERTS bands. Again, the changes will not be identical for all spectral bands, but ratioing will cause some reduction in source noise. In effect, a ratio of Equation 1 for two wavelengths should result in a cancellation of in-phase noise components of each of the terms.

From these considerations, we must conclude that ratioing of variables will result in the cancellation of in-phase fluctuations of the original variables. That this has occurred for the ratio of channel 7 to channel 5 is obvious from Table 2. The sand, wheat, and ground classes were chosen for this example because of their homogeneity.

FILTERING

Virtually all the sources of noise and propagation fluctuations represented on Figure 2 will be reduced by filtering the data. The only exception to this would be atmospheric propagation changes from one date to the next.

A simple digital filter was employed for this effort. It is the moving average (MA) filter which is a discrete form of convolution integral. The equation for a two-dimensional version of the MA filter is

$$V'(x_l, y_k) = \sum_{i=n}^{+n} \sum_{j=n}^{+n} w_{ij}\, V(x_{l-i}, y_{k-j}) \Big/ \sum w_{ij} \tag{4}$$

where

- V is the reflectance value at position x, y;
- V' is the filtered reflectance value at position x, y;
- w_{ij} are weighting coefficients; and
- n is an integer (1, 2, 3, . . .) which determines the area over which data are averaged.

This is a type of filter where the weighting coefficients are impulse response coefficients. Note also that this filter is closely related to an analog low pass filter with a bandwidth determined by n. Figure 4 illustrates this filtering method by representing the selection of data to be averaged with an input mask of coefficients to be moved over the entire array of data.

A portion of the Pawnee Test Site is shown on Figure 5, which is a microfilm graymap for ERTS band 7, July 28, 1973. This part of the July 28 image was chosen for a filtering experiment because it exhibits several of the noise sources given on the observable system diagram (Figure 2). For instance, we can clearly see the shadow of a jet contrail and the effect of thin cirrus clouds shading the region to the north and west of the contrail shadow. One can say, therefore, that the upper left-hand portion of Figure 5 has been affected by visible propagation flucutations, both for the incoming direct radiation from the sun and the reflected outgoing radiation to the satellite. It appears that the lower right-hand portion has been affected mostly by the reflected radiation penetrating through the cirrus clouds. One can expect, therefore, that all the data in Figure 5 have been extensively affected by atmospheric noise fluctuations of the type discussed by Slater (1974) and Duggin (1974).

The striations produced by the radiometric system noise is very evident when one looks across the figure at a shallow angle. Note also that the severe atmospheric noise in the upper left-hand portion overrides and

TABLE 2. COEFFICIENTS OF VARIATION FOR ERTS BANDS 5 AND 7 AND THE RATIO 7/5. (JULY 28, 1973 DATA).

Variable	Sand	Wheat	Ground
5	0.070	0.040	0.048
7	0.083	0.043	0.054
7/5	0.032	0.032	0.031

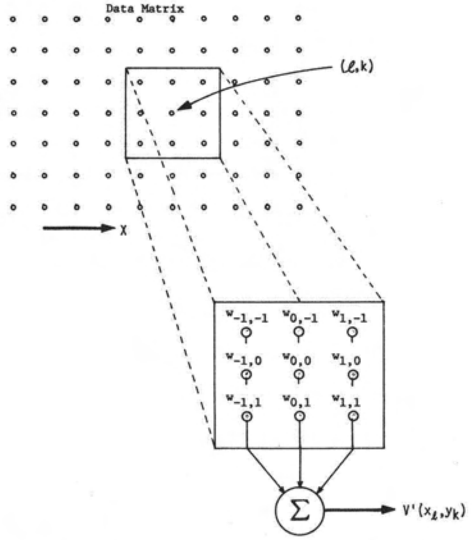
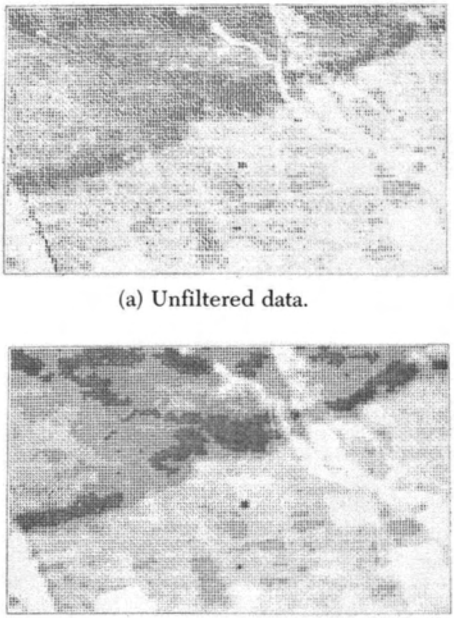


FIG. 4. An illustration of a 3 × 3, two-dimensional moving average filter mask (see Equation 4).



(a) Unfiltered data.

(b) Filtered data.

FIG. 5. Computer generated graymaps of the central portion of the Pawnee Test Site, July 28, 1973, ERTS Band 7.

reduces the visibility of the horizontal striations.

The data in Figure 5 were filtered according to Equation 4 using $n = 1$; $w_{0,0} = 1.0$; $w_{-1,0} = w_{1,0} = 0.6$; $w_{0,-1} = w_{0,1} = 0.42$; and $w_{-1,-1} = w_{1,-1} = w_{-1,1} = w_{1,1} = 0.34$. The relative value of these weighting factors (see Figure 4) is based on the 56-meter spacing between pixels along scanlines, which results in some overlap, and the 79-meter spacing between scanlines.

The effect of the filtering is clearly seen in Figure 5. The boundaries of fields are more clearly defined and the high frequency noise which resulted in the splotchy or textured appearance has been almost entirely removed. The striations are still visible, but their intensity has been reduced.

The only negative aspect of this filtering can be noted by comparing the widths of the sandy arroyo and U. S. highway 85, filtered and unfiltered. U. S. 85 is the dark line going across the lower left-hand corner of the figure. The filtering has tended to increase the width of these very narrow, high contrast features. This undesirable broadening of narrow features could be eliminated by using a band pass filter, which could be accomplished with a Fourier transform technique. The results of the filtering on the classification of vegetation will be discussed later.

LINEAR TRANSFORMATION OF VARIATES

Transformation of variates may be undertaken for many reasons; including analysis of cause and effect relationships; reduction of the number of variables; obtaining new orthogonal, independent variables; and maximizing and separation of groups (classes) along new coordinate axes. The purposes of transformations for this research were two-fold: (1) to maximize the separation of classes along new axes and (2) to satisfy the assumption of independent variables associated with most multivariate analyses.

Both Principal Components and Canonical transformations result in independent, orthogonal variables, but only the Canonical transformation is designed to maximize the separation between classes. Matrix notation will be used to simplify the discussion.

The Canonical transformation operates on data which have been subdivided into groups representing, supposedly, different classes or populations. It assumes each group is from a Normal population and that the covariance matrices of all groups are equal. This "within groups" covariance matrix is designated Γ and the "between groups" covariance matrix, which accounts for the variance between groups, is designated Ξ.

Now if the Canonical transformation is defined by

$$\boldsymbol{Y} = \boldsymbol{C}(X - \mu) \tag{5}$$

where

$\boldsymbol{X}$ is the matrix of n dimensional data samples,
μ is the mean vector for the data, and
$\boldsymbol{C}$ is the transformation matrix,

then maximization of group separation along canonical axes requires the maximization of

$$\boldsymbol{C}\,\Xi\,\boldsymbol{C}^{\mathrm{T}} \tag{6}$$

subject to the restriction

$$\boldsymbol{C}\,\Gamma\,\boldsymbol{C}^{\mathrm{T}} = \boldsymbol{I} \tag{7}$$

which insures independent y variates, each having unit variance. These dual requirements result in a characteristic equation of the form

$$\left|\Gamma^{-1/2}\,\Xi\,\Gamma^{-1/2} - \lambda\,I\right| = 0 \tag{8}$$

Solving Equation 8 yields eigenvalues, λ_i, which may be used to solve Equation 7 for the eigenvectors which form the matrix $\boldsymbol{C}$. This matrix used in Equation 5 accomplishes a Canonical transformation. For more detailed information the reader is referred to Seal (1964).

Processing Results

The data processing methods could be placed under the general category of pattern recognition, using supervised classification. The few examples presented in this paper have been carefully selected to illustrate the effects of data preprocessing, and the overall value of multivariate systems analysis.

System Description and Analysis Methods

The stochastic system analyzed is shown diagramatically in Figure 3. Computer compatible tapes contained four spectral bands of reflectance signals plus noise. Signal preprocessing included cleaning of training data, ratioing of channels, filtering, and Canonical transformation of variables. Signal processing was accomplished with Colorado State University's pattern recognition program, RECOG. RECOG (similar to Purdue's LARSYS program, as employed here), assumes the data are normally distributed, and employs a Bayesian decision rule of the form (see Duda and Hart, 1973 for a good discussion of Bayes' rule)

$$p(C_i|x) = \frac{p(x|C_i)\, p(C_i)}{\sum\limits_i p(x|C_i)\, p(C_i)} \qquad (9)$$

where

C_i is class i and

x is a data sample.

The $p(x|C_i)$ is determined from estimates of the mean, μ_i, and the covariance matrix, Σ_i, for each class; these estimates obtained from training data. The class probabilities, $p(C_i)$, were assumed equal.

Results with Limited Preprocessing

These results do include the effect of preprocessing training data, through the cleaning process previously described, but this does not apply to the bulk of the data to be classified. Also, these results will include the use of ratioed variables, 7/5 and 5/4. The results of filtering and data transformation, however, will be reserved for the following sections.

The final classification results for the July 28 vegetation type training fields are shown

Table 3. Classification Results for Vegetation Type Training Data, Before and After Data Cleaning. (July 28, 1973)

	Before Cleaning Classes									Per Cent
Classes	1	2	3	4	5	6	7	8	9	Correct
1) HBOGR	48	27	20	0	1	3	0	0	0	48
2) LBOGR	25	68	2	0	4	0	0	0	0	68
3) PITTED	8	0	38	0	0	2	0	2	4	72
4) FRWING	3	6	0	62	23	0	0	0	0	66
5) ASWALE	1	5	0	5	37	0	0	0	0	77
6) CRESTD	3	0	4	0	0	21	5	0	3	64
7) SAND	0	0	5	0	0	2	18	1	1	67
8) WHEAT	0	0	7	0	0	1	0	47	13	69
9) GROUND	0	0	2	0	0	4	2	8	50	76
	After Cleaning Classes									Per Cent
Classes	1	2	3	4	5	6	7	8	9	Correct
1) HBOGR	49	9	1	0	0	0	0	0	0	83
2) LBOGR	5	70	0	0	0	0	0	0	0	93
3) PITTED	1	0	34	0	0	1	0	0	0	94
4) FRWING	0	0	0	42	8	0	0	0	0	84
5) ASWALE	0	0	0	1	39	0	0	0	0	98
6) CRESTD	0	0	0	0	0	25	0	0	0	100
7) SAND	0	0	0	0	0	0	15	0	0	100
8) WHEAT	0	0	0	0	0	0	0	44	0	100
9) GROUND	0	0	0	0	0	0	0	0	51	100

Note: Summing along lines gives the total number of original cases (training data) in each class. Summing columns gives the total number of cases assigned to each class by the decision algorithm.

TABLE 4. RELATIVE VALUE OF VARIABLES FOR VEGETATION TYPE CLASSIFICATIONS—JULY 28, 1973. (FROM THE STEPWISE DISCRIMINANT ANALYSIS).

Step Number	Variable Entered	Initial F Value	F Value To Enter
1	5	1275	1275
2	7/5	1111	261
3	7	200	102
4	4	582	53
5	5/4	435	34
6	6	364	5

TABLE 5. WITHIN GROUPS (POOLED) CORRELATION MATRIX FOR VEGETATION TYPE TRAINING DATA—JULY 28, 1973.

	Variables					
Variables	4	5	6	7	7/5	5/4
4	1.00					
5	0.68	1.00				
6	0.64	0.62	1.00			
7	0.60	0.64	0.75	1.00		
7/5	−0.06	−0.37	0.14	0.40	1.00	
5/4	−0.30	0.48	0.07	0.14	−0.42	1.00

on Table 3 before and after cleaning. The importance of each of the six variables is indicated in Table 4. The initial F values are the best indicators of the relative amount of between-group variance which is accounted for by each variable. The reduction in F values after each variable is entered is due to the correlation between variables. The pooled correlation matrix for these data is given in Table 5. The relatively low correlation between the ratio variables and the variables from which they were formed is indicative of high noise levels, as anticipated.

As noted previously, training data for biomass classes were formed from a mixture of vegetation types and data from all three image dates. No further refining of this data was possible, because there was a wide range of biomass values included in each class, which could be expected to produce classification errors. Furthermore, the field sampling program was not adequate to accurately monitor biomass variations within the training fields. Hence, the relatively poor training data classification results given in Table 6 were expected. The close grouping of errors about the diagonal is indicative of noisy data, in this case, mostly source noise as previously defined. The relative value of the variables for biomass classification is indicated by the F Values given in Table 7. The importance of the ratio 7/5 is apparent and consistent with the discussion of ratioing.

A very useful multivariate procedure provides a plot of the data against the first two canonical variates. Such a plot for the biomass training data is given in Figure 6. Even though the processing at this point was not using transformed variates, a canonical plot is useful to show group separations from the best two-dimensional view. It is interesting to note that the biomass classes plotted against the first two canonical variates follows an orderly pattern when all six variables are used. With only four variables (no ratios) the data are more scattered and less orderly. The reduction in scatter with ratios is undoubtedly the result of a reduction of source and atmospheric (propagation) noise as hypothesized. The more orderly behavior probably results from a reduction in the effect of mean reflectance changes from one date to the next and the greater sensitivity of

TABLE 6. CLASSIFICATION RESULTS FOR BIOMASS TRAINING DATA. (SEE TABLE 9 FOR THE BIOMASS CODE)

	Classes							Per Cent
Classes	1	2	3	4	5	6	7	Correct
1) ABLE	164	15	8	1	0	0	0	87
2) BAKER	25	172	41	2	0	0	0	72
3) CHARLS	0	12	61	25	0	0	0	62
4) DOG	0	11	41	110	22	0	0	60
5) EASY	0	0	1	15	145	0	5	87
6) FOX	0	0	0	0	19	86	2	80
7) GEORGE	0	0	0	0	0	10	39	80

Note: Summing along lines gives the total number of original cases (training data) in each class. Summing columns gives the total number of cases assigned to each class by the decision algorithm.

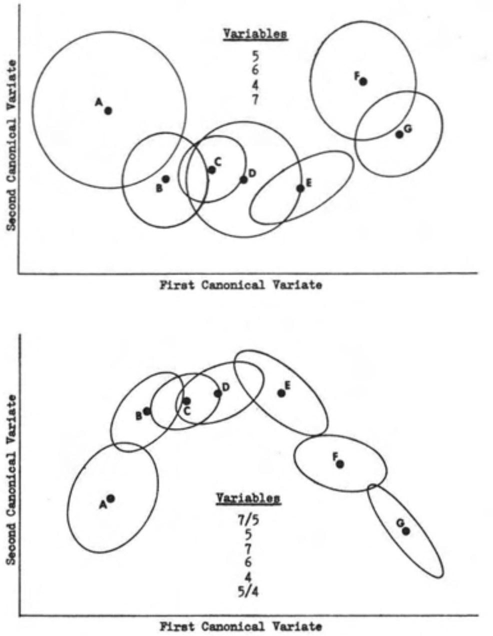


FIG. 6. Plots of the first two canonical variates for biomass training classes. Ellipses represent approximately 1σ boundaries.
Variables, ERTS bands, are listed in the order of decreasing F values.

variable 7/5 to biomass change. The use of the ratio 7/5 is strongly supported by these results.

RESULTS FOR CANONICALLY TRANSFORMED DATA

The training data for biomass classes were subjected to a canonical transformation. Following the transformation, the covariance and correlation matrices were reduced to exact identity matrices, showing the complete independence of the canonical variates. The F values for the new canonical variates are given in Table 8. Note that initial values and values to enter are now equal since there is no correlation between these variables. Comparing Tables 7 and 8 we note that canonical variate 1 accounts for more of the data dispersion than did variable 7/5. The effect of this high F value for the first canonical variate can best be seen by comparing classification results for variable 7/5 and the first canonical variable (used by themselves) in Table 9. A significant improvement in classification results is noted.

The final classification results using all of the transformed variates showed only a slight improvement in classification accuracy. This can probably be attributed to both the high F values of the original variables and the use of all variables for classification. In other words, the original variables are quite capable of separating the classes.

Furthermore, regardless of the correlation between them, when all six variables are used in a Bayesian classification system, virtually all of the information is extracted. Canonical transformation does not increase the separation between classes, it simply maximizes the between-groups variance for the first variates. This concentrates class in-

TABLE 7. RELATIVE VALUE OF VARIABLES FOR BIOMASS CLASSIFICATION.

Step Number	Variable Entered	Initial F Value	F Value To Enter
1	7/5	1080	1080
2	5	771	342
3	6	325	139
4	7	233	56
5	4	343	31
6	5/4	573	53

TABLE 8. RELATIVE VALUE OF CANONICAL VARIABLES FOR BIOMASS CLASSIFICATION.

Step Number	Canonical Variable Entered	Initial F Value	F Value To Enter
1	1	1540	1540
2	2	439	439
3	3	183	183
4	4	56	56
5	5	5	5
6	6	0	not entered

TABLE 9. A COMPARISON OF CLASSIFICATION ACCURACY USING ONLY THE ORIGINAL RATIO VARIABLE 7/5 AND THE FIRST CANONICAL VARIABLE.
(SEE TABLE 1 FOR THE BIOMASS CODE)

Biomass Class	Per Cent Errors With Variable 7/5	Per Cent Errors With 1st Canonical Variable
ABLE	30	25
BAKER	44	33
CHARLS	30	33
DOG	52	45
EASY	42	30
FOX	33	28
GEORGE	16	22
Average Error	35.3 Per Cent	30.9 Per Cent

TABLE 10. A COMPARISON OF CLASSIFICATION RESULTS FOR FILTERED AND UNFILTERED TRAINING DATA FOR JULY 28, 1973. ALL SIX VARIABLES WERE USED.

Class	Per Cent Errors Unfiltered	Per Cent Errors Filtered
HBOGR	13	3
LBOGR	30	13
PITTED	11	0
FRWING	26	4
ASWALE	13	4
SAND	37	0
GROUND	6	0
CONTRL	14	5
Average Error	18.8 Per Cent	3.6 Per Cent

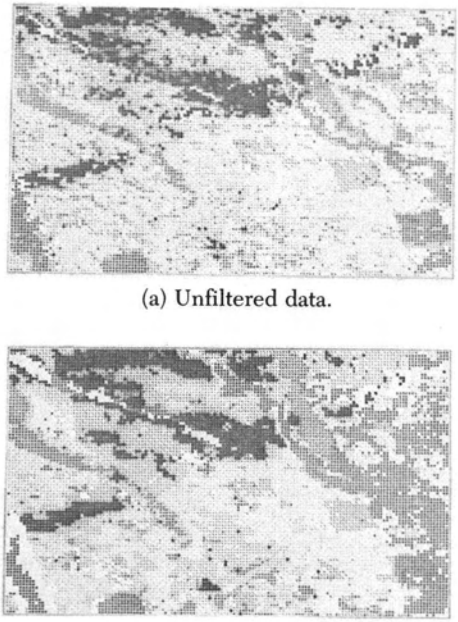
TABLE 11. A COMPARISON OF INITIAL *F* VALUES FOR THE FILTERED AND UNFILTERED TRAINING DATA FROM THE REGION SHOWN ON FIGURES.

Variable	Initial *F* Values Unfiltered Data	Filtered Data
4	329	1040
5	470	1411
6	406	1088
7	371	704
7/5	273	649
5/4	301	1014

formation in the first variates, which would be more important for a deterministic (regression) model than for the multivariate stochastic model employed here.

RESULTS FOR FILTERED DATA

The smoothing effect of filtering was evident on Figure 5. Now the effect of filtering on classification results will be considered.


(a) Unfiltered data.

(b) Filtered data.

FIG. 7. Classification maps for the area shown on Figure 5. July 28, 1973 data. Classes from black to white are: CONTRL, GROUND, SAND, ASWALE, FRWING, PITTED, LBOGR, HBOGR, and not classified.

Most of the vegetation types previously noted were found within the small portion of the Pawnee Test Site used for the filtering experiment. To these classes was added CONTRL for the contrail shadow. Filtered and unfiltered training data were selected for all of the classes. Neither of these sets of training data were cleaned by removal of anomalous samples, however, except when it was known the anomaly resulted from errors in determining training field boundaries. Noisy samples were not removed.

The classification results for these filtered and unfiltered training data are given in Table 10. The unfiltered data have an average of 18.8 per cent errors while the filtered data average 3.6 per cent errors. This is a factor of 5 improvement. The pixels used for these results were, of course, identical in every other respect.

The effect of filtering is further revealed by a comparison of the initial *F* values given in Table 11. This three-to-one increase in *F* values is a measure of the significance of the variables for separating the classes. The increase is due to a reduction in the within-groups covariance values (a reduction in noise) by a factor of three.

The training data (filtered and unfiltered) were used to calculate mean vectors and covariance matrices for Bayesian classification processing. The results are given on Figure 7. In general, it is possible to identify large contiguous areas of a given class over more of the region for the filtered data. The noisiness evident for the unfiltered data has been greatly reduced. Also, the undesirable broadening of narrow features is apparent. This could be eliminated or reduced by use of more sophisticated filtering methods. The horizontal striations produced by radiometric sensor "noise" are virtually eliminated from the filtered results. All in all, the filtering seems to have been very effective and desirable.

Conclusions

The development of systems models suggested several possible sources of noise. Furthermore, these models and the system characteristics noted therein identified several possible methods for preprocessing ERTS data.

The preprocessing methods, including ratioing of variables, cleaning of training data, linear transformation of variables, and spatial filtering, were effective in improving classification results. Ratioing of variables effectively reduced random fluctuations of reflectance values caused by source variations and changing atmospheric conditions. Ratioing was particularly useful for biomass classification since the ratio of ERTS band 7 to band 5 enhanced the effect of biomass changes.

A linear canonical transformation of variables provided only a small reduction in classification errors when all six variables were used. When only one or two variables are used for classification, however, significant improvements are noted for canonical variables because of the concentrations of between-groups variance in the first variates.

A two-dimensional spatial filter, moving average type, significantly reduced source, system, and atmospheric noise, resulting in a five-to-one reduction in classification errors. More sophisticated filtering methods will undoubtedly yield even greater improvement.

Obviously, the use of systems analysis to gain greater insight into the operation and application of remote sensing systems could yield many benefits including

(1) The design of better systems,
(2) The design of experiments to yield more information,
(3) Improved interpretation of remote sensing data, and
(4) More effective application of remote sensing systems to environmental monitoring and resource management.

References

Duda, Richard D. and Peter E. Hart, 1973. *Pattern Classification and Scene Analysis,* John Wiley and Sons, New York.

Duggin, M. J., 1974. On the natural limitations of target differentiation by means of spectral discrimination techniques, 9th Symp. on Rem. Sens. of Envir., Ann Arbor, Mich., April, 1974.

Maxwell, E. L., 1976. A remote rangeland analysis system, *Jour. of Range Management,* 29(1), Jan. 1976, pp. 66-73.

Miller, L. D. and R. L. Pearson. 1971. Aerial mapping program of the IBP Grassland Biome: Remote sensing of the productivity of shortgrass prairie as input into biosystem models. *Proceedings* of the 7th Int'l. Sym. on Rem. Sens. Vol. I, Univ. of Mich. pp. 165-207.

Quade, E. S. and W. I. Boucher, ed., 1968. *Systems Analysis and Policy Planning, Applications in Defense,* Amer. Elsevier Pub. Co., New York, 1965 (pp. 15-17).

Seal, H. L., 1964. *Multivariate Statistical Analysis for Biologists,* Methuen and Co. Ltd., London.

Slater, P. N. 1974. Specifications for photographic and electro-optical remote sensing systems, 18th annual technical meeting of the Soc. of Photo-optical Instr. Engrs., San Diego, Calif. Aug. 1974.

Tucker, C. J., 1973. *The remote estimation of a grassland canopy.* M. S. Thesis, Colorado State University, Fort Collins, Colo.