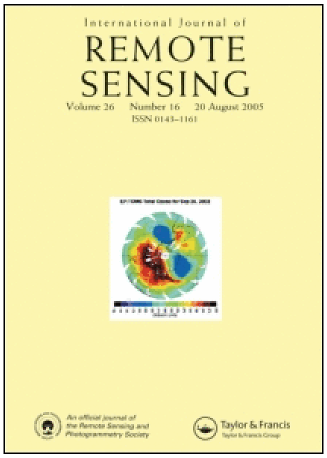
Principal component analysis of multitemporal image pairs

Article *in* International Journal of Remote Sensing · May 1985

DOI: 10.1080/01431168508948491



International Journal of Remote Sensing
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713722504

Principal components analysis of multitemporal image pairs

S. E. Ingebritsen [a]; R. J. P. Lyon [a]
[a] Remote Sensing Laboratory, Department of Applied Earth Sciences, Stanford University, Stanford, California, U.S.A

Online Publication Date: 01 May 1985
To cite this Article: Ingebritsen, S. E. and Lyon, R. J. P. (1985) 'Principal components analysis of multitemporal image pairs', International Journal of Remote Sensing, 6:5, 687 - 696
To link to this article: DOI: 10.1080/01431168508948491
URL: http://dx.doi.org/10.1080/01431168508948491

INT. J. REMOTE SENSING, 1985, VOL. 6, NO. 5, 687–696

Principal components analysis of multitemporal image pairs

S. E. INGEBRITSEN and R. J. P. LYON

Remote Sensing Laboratory, Department of Applied Earth Sciences, Stanford University, Stanford, California, 94305, U.S.A.

(Received 2 January 1984; in final form 26 March 1984)

Abstract. Principal components analysis of eight-channel data sets consisting of multitemporal LANDSAT MSS image pairs often generates higher-order principal components that are related to changes in 'brightness' and 'greenness'. This is the expected result of such analysis in a wide variety of biological and geological environments where the original imagery is intrinsically two dimensional; the two dimensions are 'brightness' and 'greenness'; and the change in land cover between images exceeds some threshold value.

1. Introduction

Principal components analysis (PCA) is a statistical technique that transforms a multivariate data set consisting of intercorrelated variables into a data set consisting of variables that are uncorrelated linear combinations of the original variables. The transformed variables are referred to as principal components (PCs). PCs are chosen in such a way that the first PC expresses the maximum possible proportion of the variance in the original data set; subsequent PCs account for successively smaller proportions of the remaining variance.

In remote sensing, PCA is often used as a data-compression technique. Since colour display facilities are defined by three channels, it is often difficult to appreciate visually all of the information available in multispectral data. PCA may allow the available information to be displayed in two or three dimensions.

Another application of PCA in remote sensing is to create combinations of n spectral bands that generate physically significant indices in n space.

The applicability of PCA to yet another remote-sensing problem, that of detecting and monitoring temporal change, has been demonstrated by Lodwick (1979, 1981) and Byrne *et al.* (1980). These studies involved geographically registered multitemporal LANDSAT multispectral scanner (MSS) imagery. Lodwick (1979) found that seasonal changes could be measured using the first two PCs, by differencing the PC scores between successive images or by linear regression across a number of images. Byrne *et al.* (1980) superimposed two LANDSAT images of the same area and treated them as a single eight-channel data set. PCA of this augmented data set generated higher-order PCs that contained useful information about temporal change.

This paper documents the type of behaviour observed by Byrne *et al.* (1980) in two biologically and geologically disparate environments, and explains this behaviour in terms of the nature of LANDSAT MSS imagery and the PCA process.

2. Study areas

The work described in this paper was performed during a U.S. Department of the Interior-sponsored† study of the use of satellite imagery to detect and monitor environmental and land-use change. The study areas included the area in and around Midnite Mine, a large open-pit uranium mine located approximately 50 km northwest of Spokane, Washington, and a wetland area in the Carson Desert, Nevada.

Midnite Mine is in the Okanogan Highlands, an area of mature topography and moderate relief. The Okanogan Highlands are comprised of a wide variety of geological materials, but within the Midnite Mine study area there are three major rock units: a quartz monzonite, a phyllite and a calc-silicate unit. The vegetation in this study area is predominantly an association of Western Yellow Pine (*Pinus ponderosa*), Bitterbrush (*Porchia tridentata*) and annual grasses.

The Carson Desert study area lies within the largest intermontane basin in Nevada. It includes Stillwater and Fallon National Wildlife Refuges, which incorporate much of the wetland formed by the terminal sink of the Carson river. The study area is mantled by unconsolidated sediments of Lake Lahontan age and younger. The vegetative cover, which is relatively sparse, includes Greasewood (*Sarcobatus* sp.), Salt Cedar (*Tamarix* sp.) and marsh grasses.

3. Intrinsic dimensionality and *n*-space indices

The term intrinsic dimensionality is often used to refer to the smallest number of dimensions which could be used to accurately represent a data set. The intrinsic dimensionality of LANDSAT MSS data is approximately two (Swain and Davis 1978).

Kauth and Thomas (1976) showed that a pair of indices in four space—'brightness' and 'greenness'—contained almost all of the variance in LANDSAT agricultural scenes. Similar indices in four space are generated by PCA of LANDSAT images. The 'brightness' generated by PCA of LANDSAT imagery is not strictly analogous to the brightness index of Kauth and Thomas, since their index is based on soil data only, while the 'brightness' component generated by PCA of a heterogeneous area is a more generalized measurement. However, since the mean vector of soil reflectance lies near the diagonal in four space along which the normalized reflectance of all LANDSAT bands is equivalent, the effect is often similar.

The concept of brightness and greenness indices is most readily represented in two dimensions. Richardson and Wiegand (1977) showed that a plot of LANDSAT MSS band 5 (visible red) and 7 (near-infrared (IR)) data for soils would fall on a straight line (see figure 1). As the amount of vegetation present increases, red reflectance decreases and the near-IR reflectance increases. The perpendicular distance of a point away from the soil line is a measure of the amount of green vegetation present, and is referred to as the perpendicular vegetation index (PVI).

The concept of a soil 'line' is as fundamental in four dimensions as in two, but more difficult to represent. Four-dimensional greenness, like the PVI, is defined by the orthogonal distance from the soil 'line'. The four-dimensional brightness index establishes the soil data space, and the greenness index emphasizes departures from it. Due to the characteristics of typical soil and vegetation spectra, these departures

† Funding originally came from the Conservation Division of the U.S. Geological Survey. During the contract period the Conservation Division briefly became an independent agency within the Department of the Interior, the Minerals Management Service, before being absorbed by the Bureau of Land Management.

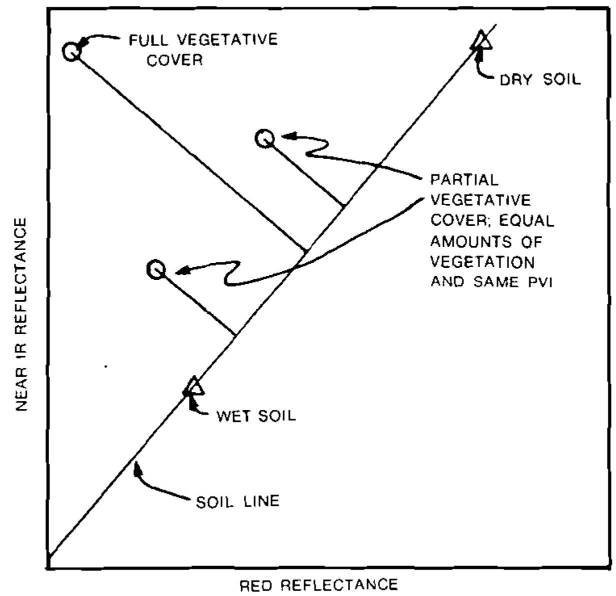


Figure 1. A soil line and vegetation points in two-dimensional space (after Jackson 1983).

tend to be negative in the visible bands and positive in the near-IR (see figure 2). The factor loadings for eigenvectors representing brightness are uniformly high and positive, while those related to greenness are negative in visible wavelengths and positive in the near-IR.

PCA was applied to several LANDSAT images of the Midnite Mine and Carson Desert study areas. In every case, the first PC was related to brightness, the second PC greenness. The third and fourth PCs showed mostly noise.

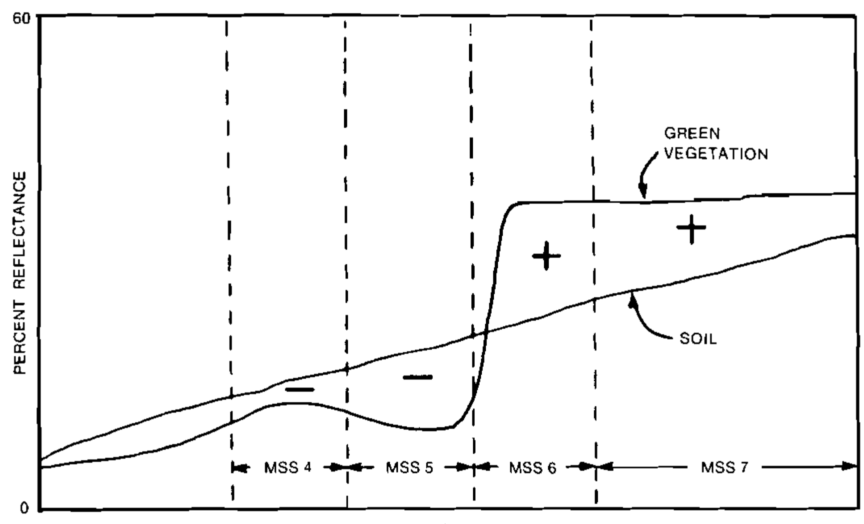


Figure 2. Typical reflectance spectra for soil and green vegetation. The brightness index establishes the soil data space and the greenness index emphasizes departures from it (+ and −). MSS 4, 5, 6 and 7 are the LANDSAT MSS bandpasses.

4. Change detection with multitemporal LANDSAT imagery

PCA can be used to detect and identify temporal change when registered LANDSAT MSS images are merged and treated as a single data set. In many cases, PCA of such a merged data set will generate higher-order PCs representing *changes* in brightness and greenness, which we refer to as the Δbrightness and Δgreenness PCs. This is the expected result if

(1) the two original images both have an intrinsic dimensionality of two,
(2) the two dimensions are brightness and greenness and
(3) the change in land cover and/or the condition of the vegetation between the two dates exceeds some threshold value.

The first two criteria are met in a wide variety of biological and geological environments. A threshold level of change between dates may be related to natural seasonal variations in the density and vigour of vegetation, changes in soil moisture, change related to human activity, or a combination of these factors. If all three criteria are met there will be four meaningful PCs in the transformed data set. These first four PCs will represent 'stable' brightness, 'stable' greenness, Δbrightness and Δgreenness. The Δbrightness PC is roughly analogous to an albedo-difference image, and Δgreenness is related to changes in the vegetative cover. The 'stable' brightness and greenness components represent an averaging of the results from the two dates.

The explanation of this behaviour is straightforward. If the variance–covariance or correlation matrices from both of the *original* LANDSAT images have an intrinsic dimensionality of two, the augmented matrix formed by merging the two images must have an intrinsic dimensionality of four or less. If the two dimensions of each of the original images are brightness and greenness, there will be PCs in the augmented eight-channel data set representing 'stable' brightness and greenness. If there is no change between the two dates the augmented matrix, like the original matrices, will be intrinsically two-dimensional, and 'stable' brightness and greenness will be the only meaningful PCs. Given a finite amount of change, differences in brightness and greenness between dates introduce additional sources of variance in the augmented data set. As a result, the augmented matrix often has an intrinsic dimensionality of four, with the additional dimensions being Δbrightness and Δgreenness. In the multitemporal data set, Δbrightness can be conceptualized as being orthogonal to 'stable' brightness within a brightness plane. Similarly, Δgreenness may be conceptualized as lying in a greenness plane orthogonal to 'stable' greenness.

There is no generally accepted way to diagnose the intrinsic dimensionality of variance–covariance or correlation matrices, though the condition number λ_1/λ_N, where λ_1 and λ_N are the eigenvalues associated with the first and last PCs, has been used to test for singularity.

Heuristically, we would expect the value of

$$\mathrm{NSR} = 100\left(\frac{\sum_{j=P+1}^{N} \lambda_j \Big/ (N-P)}{\sum_{j=1}^{P} \lambda_j \Big/ P}\right)^{1/2}$$

where N is the dimension of the matrix (full rank), P is the ith component and λ_j is the eigenvalue associated with the jth component, to reach a minimum where i is equal to the intrinsic dimensionality of the matrix (see, for example, Jupp 1978). This noise to signal (NSR) ratio will tend to identify the point at which the variance accounted for by each succeeding PC reaches a low, nearly constant noise level—i.e., the point at which the variance in the last $(N-P)$ dimensions is approximately spherical. Note that the NSR is undefined at $i=N$, so this diagnostic is not applicable if the intrinsic dimensionality is N or $(N-1)$.

Table 1 shows the results of applying the NSR diagnostic to single-date and merged (eight-channel) images of the Midnite Mine and Carson Desert areas. With one exception, the matrices from the single-date images are two-dimensional, and in every case the first two PCs are identifiable as brightness and greenness, on the basis of the factor loadings (table 2) and inspection of the corresponding PC score imagery. With one exception, the augmented matrices from the merged images are four dimensional, and in every case the first four PCs are related to 'stable' brightness and greenness, Δbrightness and Δgreenness. The 'stable' brightness PCs

Table 1. The NSR test applied to four-dimensional correlation matrices derived from single-date images, and to eight-dimensional matrices derived from multitemporal image pairs.

PC	λ	Percent of variance	Physical significance	NSR
		Midnite Mine study area		
		8 August 1973		
1	3·282	82·0	Brightness	27·01
2	0·603	15·0	Greenness	17.27, 'rank' 2
3	0·075	1·9		17·54
4	0·041	1·0		—
		22 July 1978		
1	2·955	73·9	Brightness	34·33
2	0·953	23·8	Greenness	15·33, 'rank' 2
3	0·056	1·4		16·29
4	0·036	0·9		—
		29 July 1980		
1	3·182	79·6	Brightness	29·47
2	0·739	18·5	Greenness	14·20, 'rank' 2
3	0·049	1·2		15·21
4	0·031	0·8		—
		8 August 1973 and 29 July 1980		
1	5·644	70·5	Brightness	24·42
2	1·092	27·3	Greenness	25·01
3	0·809	10·1	Δbrightness	19·03
4	0·263	3·3	Δgreenness	15·69, 'rank' 4
5	0·075	0·9		15·75
6	0·047	0·6		16·30
7	0·041	0·5		—
8	0·029	0·4		—
		22 July 1978 and 29 July 1980		
1	5·628	70·4	Brightness	24·54
2	1·469	18·4	Greenness	20·60
3	0·538	6·7	Δbrightness	16·96
4	0·196	2·4	Δgreenness	14·73, 'rank' 4
5	0·057	0·7		15·45
6	0·049	0·6		—
7	0·036	0·4		—
8	0·029	0·4		—

Table 1 (*continued*).

PC	λ	Percent of variance	Physical significance	NSR
		Carson Desert wetland area 26 May 1973		
1	3·710	92·8	Brightness	16·13
2	0·268	6·7	Greenness	7·33, 'rank' 2
3	0·012	0·3		8·22
4	0·009	0·2		—
		6 August 1973		
1	3·623	90·6	Brightness	18·63
2	0·346	8·7	Greenness	11·85
3	0·018	0·4		10·08
4	0·014	0·3		—, 'rank' 4 (?)
		26 May 1973 and 6 August 1973		
1	6·685	83·6	Brightness	16·76
2	0·650	8·1	Δbrightness	17·38
3	0·529	6·6	Greenness	10·17
4	0·085	1·1	Δgreenness	8·01, 'rank' 4
5	0·019	0·2		8·23
6	0·014	0·2		8·38
7	0·010	0·1		—
8	0·009	0·1		—
		Midnite Mine study area 29 July 1980 and Carson Desert wetland area 6 August 1973 (representing an area in which land cover is totally changed between dates)		
1	4·252	53·2	Brightness	35·48
2	2·556	31·9	Δbrightness	24·16
3	0·746	9·3	Δgreenness	18·82
4	0·337	4·2	Greenness	11·77
5	0·048	0·6		11·35
6	0·031	0·4		10·78, 'rank' 6
7	0·018	0·2		10·84
8	0·013	0·2		—

are characterized by uniformly positive factor loadings, and the 'stable' greenness PCs by negative loading for the visible bands and positive loading on the near-IR bands, particularly band 7. The Δbrightness PCs are characterized by positive loading on the data from one date and negative loading on data from the other date, and the Δgreenness PCs are characterized by a greenness-type loading pattern on the data from one date and an 'inverted greenness' loading pattern on data from the other date (see table 2).

The order in which the stable greenness, Δbrightness, and Δgreenness PCs appear depends upon the amount and nature of change between dates. Stable brightness is the first PC in each of these cases.

Figures 3 and 4 are Δbrightness and Δgreenness PC images of the Midnite Mine study area generated from LANDSAT images acquired on 8 August 1973 and 29 July 1980. The imagery has been geographically registered and resampled using bicubic interpolation. The resampled pixel size is approximately 17 m × 21 m. The boundaries sketched on the imagery represent the limits of the mined area in 1973 (inner line) and 1980 (outer line). In the Δbrightness image (figure 3), the area

Table 2. Factor loading patterns.

LANDSAT band	Principal component 1	2	3	4	5	6	7	8
	Midnite Mine study area 8 August 1973							
4	0·92	−0·36	−0·05	−0·14				
5	0·92	−0·36	−0·04	0·14				
6	0·95	0·22	0·22	−0·01				
7	0·82	0·55	−0·15	0·00				
	22 July 1978							
4	0·89	−0·43	−0·05	−0·13				
5	0·88	−0·46	−0·04	0·13				
6	0·93	0·32	0·18	0·00				
7	0·72	0·68	−0·14	0·01				
	29 July 1980							
4	0·91	−0·39	−0·05	−0·12				
5	0·91	−0·39	−0·04	0·13				
6	0·96	0·24	0·18	0·00				
7	0·78	0·62	−0·12	0·00				
	8 August 1973 and 29 July 1980							
4	0·87	−0·22	−0·39	0·17	−0·05	0·03	−0·13	0·02
5	0·87	−0·22	−0·38	0·18	−0·04	−0·01	0·14	−0·03
6	0·89	0·26	−0·27	−0·14	0·22	−0·01	−0·01	0·01
7	0·74	0·54	−0·17	−0·27	−0·15	0·00	0·01	0·00
4	0·84	−0·44	0·26	−0·16	0·00	−0·05	−0·03	−0·11
5	0·84	−0·45	0·25	−0·12	−0·02	−0·03	0·03	0·12
6	0·89	0·13	0·40	0·08	0·02	0·17	0·02	−0·01
7	0·75	0·48	0·35	0·25	0·00	−0·12	−0·01	0·01
	22 July 1978 and 29 July 1980							
4	0·89	−0·34	−0·22	0·18	−0·04	0·04	−0·12	0·04
5	0·88	−0·37	−0·21	0·18	−0·04	−0·01	0·12	−0·05
6	9·86	0·37	−0·31	−0·05	0·17	−0·06	0·00	0·01
7	0·63	0·70	−0·26	−0·15	−0·14	0·04	0·01	0·00
4	0·88	−0·38	0·14	−0·21	−0·01	−0·05	−0·05	−0·11
5	0·89	−0·39	0·14	−0·18	−0·02	−0·03	0·04	0·12
6	0·91	0·18	0·32	0·02	0·06	0·16	0·02	−0·01
7	0·74	0·52	0·38	0·19	−0·04	−0·11	−0·02	0·01
	Carson Desert wetland area 26 May 1973							
4	0·95	−0·30	−0·07	0·01				
5	0·98	−0·18	0·07	−0·05				
6	0·99	0·14	0·03	0·07				
7	0·93	0·35	−0·04	−0·04				
	6 August 1973							
4	0·94	−0·32	−0·04	−0·07				
5	0·97	−0·22	0·00	0·09				
6	0·99	0·13	0·11	−0·03				
7	0·90	0·43	−0·07	0·00				

Table 2 (*continued*).

LANDSAT band	Principal component 1	2	3	4	5	6	7	8
			26 May 1973 and 6 August 1973					
4	0·90	−0·32	−0·25	−0·13	0·03	0·03	−0·05	0·00
5	0·94	−0·29	−0·14	−0·08	−0·03	−0·04	0·06	−0·03
6	0·95	−0·26	0·15	0·06	−0·02	−0·01	0·00	0·08
7	0·89	−0·25	0·34	0·15	0·02	0·02	−0·01	−0·05
4	0·90	0·25	−0·33	0·09	0·05	0·05	0·04	0·01
5	0·93	0·26	−0·22	0·09	−0·01	−0·07	−0·04	−0·01
6	0·94	0·32	0·09	−0·05	−0·10	0·05	−0·01	−0·01
7	0·86	0·31	0·38	−0·13	0·06	−0·02	0·01	0·01
	Midnite Mine study area 29 July 1980 and Carson Desert wetland area 6 August 1973 (representing an area in which land cover is totally changed between dates)							
4	0·64	0·65	−0·39	−0·05	−0·05	−0·12	0·00	0·00
5	0·64	0·66	−0·32	−0·03	−0·03	−0·13	−0·01	0·00
6	0·68	0·67	0·25	−0·01	0·17	−0·01	0·00	0·00
7	0·55	0·55	0·62	0·06	−0·12	0·00	0·00	0·00
4	0·80	−0·51	0·06	−0·31	0·00	0·00	−0·04	−0·07
5	0·82	−0·52	0·03	−0·21	−0·01	0·01	−0·01	0·09
6	0·86	−0·49	−0·03	0·13	0·00	0·00	0·11	−0·02
7	0·80	−0·43	−0·09	0·42	0·01	−0·01	−0·07	0·00

affected by mining between 1973 and 1980 appears anomalously bright, because the reflectivity of the exposed bedrock is higher than that of the original combination of vegetation and soil. In the Δgreenness image areas where greenness decreased between 1973 and 1980 are dark, while those where the greenness increased are bright. The areas mined between 1973 and 1980 are dark, due to the removal of vegetative cover. The area that had already been mined in 1973 is a medium tone because little or no green vegetation was present on either date. Several of the drainages in the study area have a lower greenness index in 1980 than in 1973; although there is a water pollution problem associated with Midnite Mine, this change is probably not related to mining activities because it is not confined to the area below the mine.

The pattern of temporal change defined by the Δbrightness component image shown in figure 3 is similar to the pattern defined by a difference image from the same dates. The Δgreenness component image in figure 4 is somewhat similar to an near-IR/red ratio difference image, although the ratio difference image is sensitive to changes in slope and slope aspect that do not affect the Δgreenness component image.

5. Conclusion

PCA of merged multitemporal LANDSAT MSS image pairs commonly provides information about the spatial distribution, magnitude and nature of temporal change by generating PCs that represent changes in brightness and greenness. A Δbrightness component image may be regarded as analogous to an albedo-difference image, and a Δgreenness component image provides information about change in the vegetative cover.

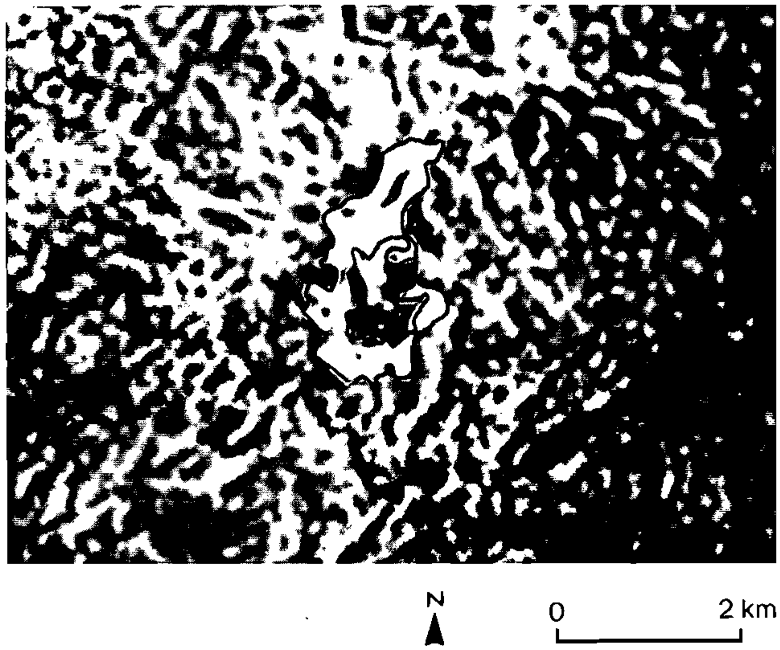


Figure 3. Δbrightness component image of the Midnite Mine study area, 8 August 1973 to 29 July 1980.

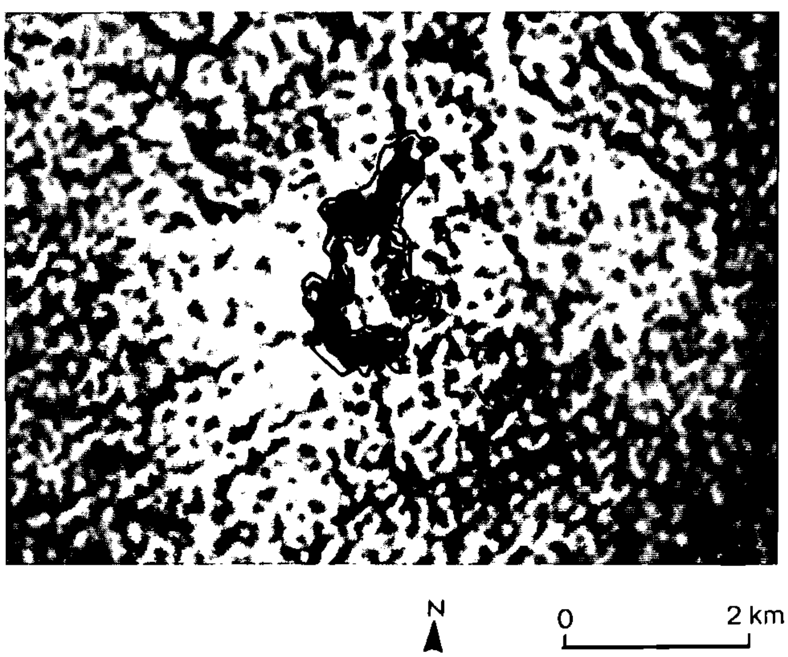


Figure 4. Δgreenness component image of the Midnite Mine study area, 8 August 1973 to 29 July 1980.

With more spectrally complicated data than that provided by the LANDSAT MSS, it might be more difficult to generate separate stable and change components using simple PCA in this manner. However, data from more sophisticated sensors could be simplified (e.g., in the case of Thematic-Mapper-type data by using bands 2, 3 and 4 only) to obtain a similar result.

Work in progress suggests that PCA of LANDSAT time-difference data often produces PCs that are highly correlated with the Δbrightness and Δgreenness PCs produced from the equivalent image pair.

Acknowledgments

The authors benefited greatly from discussions with David Jupp (CSIRO, Australia) on this topic, and would also like to thank Chris Elvidge (Department of Applied Earth Sciences, Stanford University) and Paul Switzer (Department of Statistics, Stanford University) for their critical reviews of the original manuscript.

References

BYRNE, G. F., CRAPPER, P. F., and MAYO, K. K., 1980, Monitoring land-cover change by principal components analysis of multitemporal Landsat data. *Remote Sensing Environ.*, **10**, 175.

JACKSON, R. D., 1983, Spectral indices in *n*-space. *Remote Sensing Environ.*, **13**, 409.

JUPP, D. L. B., 1978, Estimation of the magnetotelluric impedance functions. *Phys. Earth Planet. Interiors*, **17**, 75.

KAUTH, R. J., and THOMAS, G. S., 1976, The tasselled cap—a graphic description of the spectral-temporal development of agricultural crops as seen by Landsat. *Proceedings of the Symposium on Machine Processing of Remotely Sensed Data*, Purdue University, pp. 45–51.

LODWICK, G. D., 1979, Measuring ecological changes in multitemporal Landsat data using principal components. *Proceedings of the International Symposium on Remote Sensing of the Environment*, Ann Arbor, Michigan, pp. 1131–1141.

LODWICK, G. D., 1981, A computer system for monitoring environmental change in multitemporal Landsat data. *Can. J. remote Sensing*, **7**, 24.

RICHARDSON, A. J., and WIEGAND, C. L., 1977, Distinguishing vegetation from soil background information. *Photogramm. Engng remote Sensing*, **43**, 1541.

SWAIN, P. H., and DAVIS, S. M., 1978, *Remote Sensing: The Quantitative Approach* (New York: McGraw-Hill).