

Estudio multitemporal de imágenes usando la transformación de componentes principales

Siljeström P.; Moreno, A.; García L.V. & Clemente, L.

Instituto de Recursos Naturales y Agrobiología de Sevilla

RESUMEN

La transformación de componentes principales se utiliza para realzar zonas que han sufrido cambios normalmente asociados a variaciones en la cobertura vegetal, que, según numerosos autores, se hacen particularmente visibles en los componentes de orden elevado. Dichos componentes son posteriormente clasificados por técnicas no supervisadas en grandes clases dinámicas, comparándose los resultados con otras clasificaciones realizadas a partir de cada una de las imágenes estudiadas, como se ha venido haciendo clásicamente.

ABSTRACT

The principal components transformation is used to highlight areas that show a change in time, normally associated with fire damage and vegetation regrowth. These variations are enhanced in the high order components, following previous studies. The components are classified by unsupervised techniques to yield thematic maps on which dynamic cover type classes are recorded. This maps are compared with others based on the individual images, normally used in classical methods.

Introducción

La transformación de componentes principales (o de Karhunen-Loeve) es una técnica de reducción de información redundante. Dicha técnica permite una descripción de datos multidimensionales de tal forma que los ejes de las variables estén decorrelacionados, conteniendo el primer eje o componente la máxima varianza, que va disminuyendo en los componentes sucesivos (Jenson & Waltz, 1979; González & Wintz, 1987).

La decorrelación producida en este proceso ha demostrado tener un gran interés en el realce de pequeñas zonas localizadas que han sufrido un cambio temporal (Byrne & Crapper, 1979; Byrne et al, 1980). En efecto, en regiones que han permanecido relativamente constantes, la correlación entre imágenes de diferentes fechas es elevada. Por el contrario, una correlación baja o nula entre distintas fechas en un mismo área indica que ha habido un cambio. Teniendo en cuenta que la mayor proporción de varianza en una secuencia multitemporal está asociada con tipos de cober-

turas correlacionadas o constantes, regiones que han sufrido cambios se verán realzadas en los componentes de orden elevado, generados a partir de una transformación de componentes principales realizada sobre datos multitemporales y multispectrales (Richards, 1984).

Las técnicas clásicas utilizadas para la identificación de cambios en un área (clasificación y operaciones aritméticas) suponen la uniformidad multitemporal entre las bandas, sus condiciones atmosféricas, y de radiación o incluso entre los procedimientos de calibración de los datos del satélite (Byrne et al, 1980). A través del análisis de componentes principales, estas variaciones, incluyendo las áreas que no han cambiado, se recogen en el primer y segundo componente y son más sencillas de eliminar desenmascarando las zonas que sí han variado.

Área de estudio

El área de estudio se sitúa dentro del Parque Nacional de Doñana, localizado en la margen derecha de la desembocadura del Guadalquivir (SW España). A partir de la creación del Parque en 1969 se tienen datos de numerosos fuegos, en su mayoría provocados.

En este trabajo se estudia el fuego que asoló 300 Ha de matorral en 1981 (Encinillas Altas), a través de una imagen TM tomada en junio de 1984 (Fig. 1, A). En dicha imagen, que corresponde a una subescena de la banda 5, la zona quemada aparece en tonos gris claroblanco.

La segunda imagen estudiada, fue tomada en febrero de 1986 y recoge un área incendiada en 1985 de unas 900 Ha (Domínguez, 1993) en el mismo ecosistema, y limitando con el fuego anterior (Fig. 1, B). En esta imagen aparece la zona quemada en tonos claros, al igual que el caso anterior, mientras que la cicatriz de 1981 es prácticamente irreconocible gracias al crecimiento de la vegetación.

Metodología: registrado de imágenes

La investigación se llevó a cabo con un sistema de análisis de imagen Intergraph y un software denominado ISI-2. Este paquete de programas permite efectuar un registro «imagen a imagen» de forma interactiva, seleccionando puntos de control en las dos escenas multitemporales completas (de unos 870 km² cada una). Gracias a la gran variedad que presentan los ecosistemas del Parque Nacional de Doñana (sistemas de dunas móviles, lagunas, marismas, etc), se seleccionaron 48 puntos de control. Se aplicó a ambas imágenes un modelo de correlación afín o de primer orden con 90 grados de libertad, resultando un error standard de 0.0505. Posteriormente se usó como método de remuestreo el modelo de

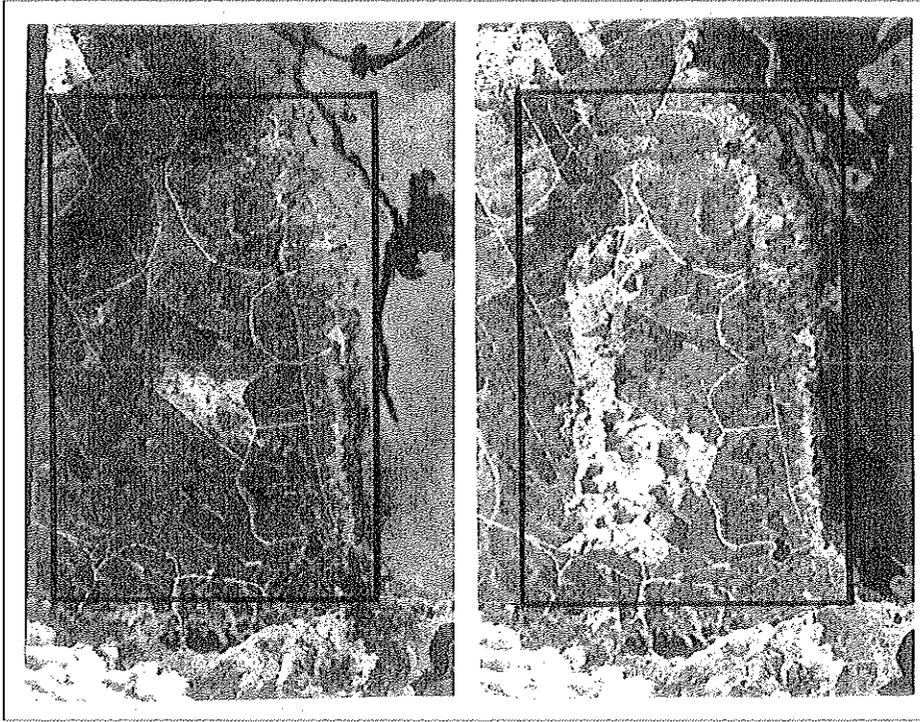


FIGURA 1

Reconocimiento del área de estudio en la banda 5 TM tomada en julio de 1984 (A) y en febrero de 1986 (B)

«nearest neighbour», obteniendo como resultado final un error standard en la predicción de menos de 1/4 de pixel tanto en filas como en columnas.

A partir de la imagen registrada se recortó una subescena de 320 x 200 pixels (unos 58 km²) en cada una de las 8 bandas escogidas. Dichas bandas, correspondían a la 3, 4, 5 y 7 de imágenes TM tomadas en junio de 1984 y febrero de 1986 (Fig. 1), respectivamente. Sobre ellas se aplicó el análisis de CP y, a partir de los componentes más útiles, se realizó un mapa temático de clases dinámicas (que presenten algún tipo de variación), comparándolo con otras clasificaciones realizadas sobre cada una de las imágenes.

Análisis de componentes principales. Resultados

Como primer paso, se unieron las subescenas registradas para formar un conjunto de datos multitemporales de 8 dimensiones, a partir del cual se generaron los componentes principales.

Las imágenes correspondientes a los componentes PC1, PC2, PC4 y PC6 se muestran en la Fig. 2 (A, B, C y D, respectivamente). El resto de los componentes no presentan rasgos significativos para este estudio, por lo que no se consideran. La correspondencia de los distintos tonos de gris en las imágenes de los CP puede apreciarse fácilmente estudiando los valores de las cargas factoriales de los autovectores. Cada componente es una combinación lineal de las 8 bandas originales, donde los valores de las cargas factoriales de los autovectores muestran la correlación entre dichos componentes y las distintas bandas (González & Wintz, 1987). Estos autovectores, con sus autovalores asociados (que representan la varianza de cada componente), se recogen en la Tabla 1.

En la interpretación de estos valores hay que tener en cuenta que el signo negativo no tiene mayor importancia que la de deberse a una inversión en la polarización introducida al escoger la opción automática en el software utilizado.

El primer componente presenta valores proporcionales al brillo total de la imagen (albedo), mientras que los siguientes comienzan a realzar los cambios. Realmente es el componente de 6º orden (Fig. 2, D) el que presenta la relación más fuerte con los rasgos de los incendios de que trata el presente trabajo. Pixels con cobertura vegetal constante (vegetación intacta en 1981 y 1986, por ej.) aparecen en un tono gris medio, mientras que aquellos que han variado (ya sea de vegetados a quemados o viceversa) presentan tonos tendentes al blanco o al negro. En efecto, el incendio de 1981 aparece en tonos muy claros, casi blanco, mientras que el ocurrido en 1986 presenta tonos muy oscuros. Por otro lado, la zona de contacto entre el matorral y la marisma (Vera) (Siljeström y Clemente, 1987) se caracteriza por la frecuencia de pastizales densos en la época húmeda (imagen de febrero 86) y por su aridez en la época seca (imagen de junio 84). Este cambio estacional de la vegetación en la vera puede asemejarse al de zonas quemadas (sin vegetación y posteriormente regeneradas, por lo que no es de extrañar que al realizar un análisis de PC se confunda con ellas.

El primer componente principal (Fig. 2, A), que recoge la mayor proporción de la varianza (59.93%) realza el fuego ocurrido en 1985 (por presentar una mayor superficie), apareciendo en tonos muy oscuros, mientras que el incendio ocurrido en 1981 es prácticamente indiferenciable.

El segundo componente (Fig. 2, B), por otro lado, realza el fuego de 1981 ya que éste significa el máximo de varianza (30.12%) en lo que queda de la imagen, desapareciendo la cicatriz del fuego de 1985.

El cuarto componente (Fig. 2, C) muestra el incendio de 1985 en tono gris medio, ligeramente más claro que el fuego de 1981, confundiendo ambos. Sin embar-

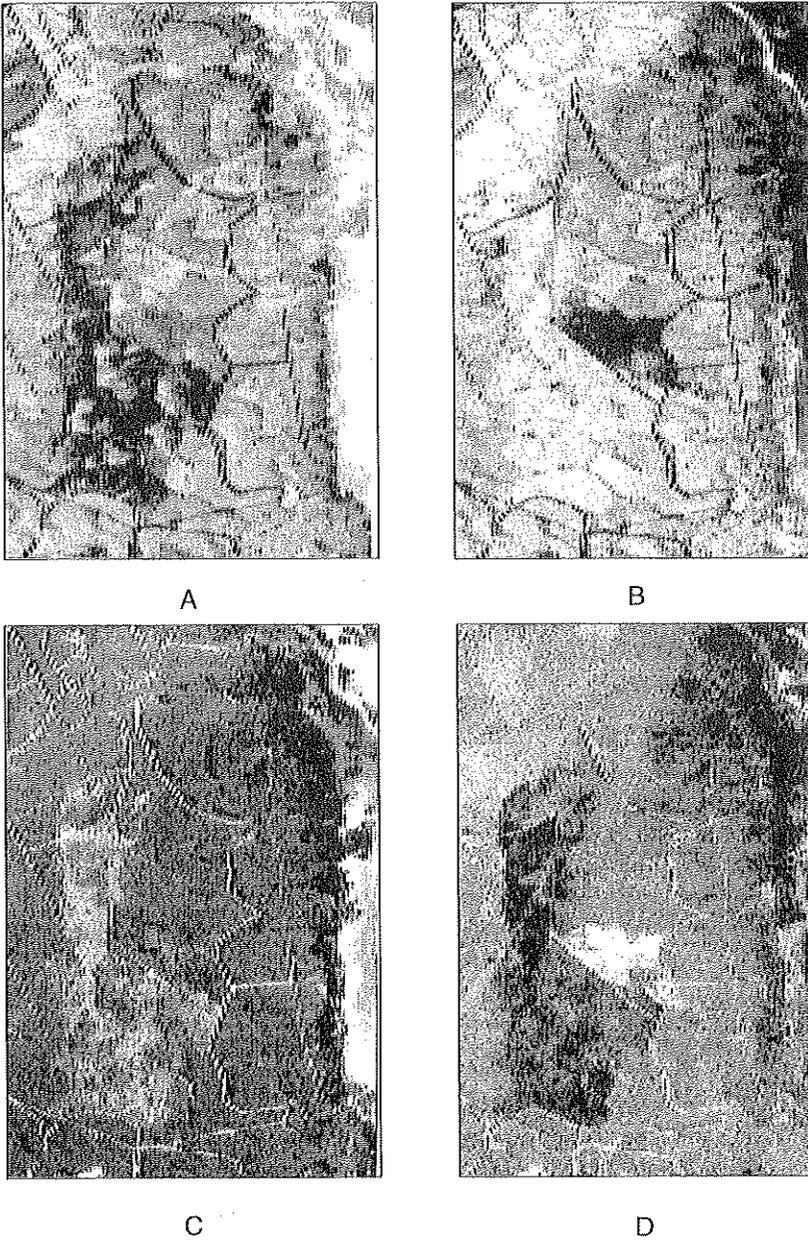


FIGURA 2
Componentes principales más significativos: PC1 (A), PC2 (B), PC4 (C) y PC6 (D)

go, el matorral circundante aparece en tonos más oscuros, delimitando razonablemente bien el total del área quemada.

Se hizo una composición de falso color asignándole el rojo al PC2, verde al PC5 y azul al PC6, y estirando los respectivos histogramas convenientemente (Fig. 3, A). La imagen resultante recoge claramente los diferentes tipos de cubiertas vegetales atendiendo a su variación a lo largo del tiempo. En efecto, el área con vegetación intacta en 1981 e incendiada en 1985 de color amarillo intenso, y la zona quemada en 1981 y posteriormente regenerada de color azul fuerte. El matorral circundante presenta distintos tonos de rojorosa que se saturan hacia la Vera donde adopta un color rojo intenso. Finalmente, la marisma aparece muy bien diferenciada, en colores que varían desde el verde al turquesa.

Medias y desviaciones típicas en cada banda utilizada								
Bandas	1984				1986			
	3	4	5	7	3	4	5	7
x	24.44	32.95	50.76	25.17	21.75	30.39	45.14	21.35
σ	5.13	6.69	12.68	8.03	6.20	7.44	18.17	11.01
Autovalores								
PC	1	2	3	4	5	6	7	8
Autoval	504.45	253.54	40.79	18.55	11.17	7.39	3.41	2.39
σ^2 (%)	59.93	30.12	4.84	2.20	1.32	0.88	0.41	0.28
Cargas factoriales de los autovectores								
	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
TM3-84	-0.15	-0.82	-0.09	0.39	0.18	0.28	0.10	0.15
TM4-84	-0.03	-0.75	-0.56	-0.02	0.32	-0.09	-0.03	-0.06
TMS-84	-0.23	-0.96	0.09	-0.09	-0.05	-0.08	0.01	0.04
TM7-84	-0.28	-0.92	0.13	0.05	-0.07	0.18	0.07	-0.11
TM3-86	-0.82	-0.13	-0.15	0.45	-0.18	-0.16	0.12	-0.07
TM4-86	-0.72	0.11	-0.62	-0.12	-0.23	0.09	-0.02	0.04
TMS-86	-0.98	0.13	0.05	-0.07	0.04	0.02	0.03	-0.01
TM7-86	-0.97	0.12	0.06	0.14	0.02	-0.05	-0.12	0.04

TABLA I
Resultado de Análisis de Componentes Principales

Clasificación de imágenes

Para comparar la utilidad de los mapas temáticos generados a partir de los componentes principales, se realizaron, clasificaciones no supervisadas sobre las cuatro bandas escogidas dentro de cada imagen. En cada clasificación se intentaron separar los principales tipos de coberturas, tanto si habían variado como si no. Así, conflictos entre clases que pudieran ser solventados por los componentes principales (que recogen únicamente clases que han variado) se harían evidentes. Es más, al ser la comparación entre clasificaciones individuales un método clásico y muy extendido en la evaluación de cambios, puede valorarse la aportación que ofrece la transformación de CP en la diferenciación de clases para este tipo de estudios.

La clasificación, no supervisada, se basó en el clustering por el método isodata implementado en el sistema Intergraph, utilizándose el clasificador de máxima verosimilitud por ofrecer los mejores resultados. En una primera aproximación para la escena de junio 1984 se generaron 12 clases que fueron agrupadas en 4: zona quemada (rojo), vegetación intacta (verde), Vera (borde de marisma) (amarillo) y marisma/otras (negro).

En la clasificación de la imagen de junio de 1984 (Fig. 3, B) puede observarse una cierta confusión entre la Vera (borde de marisma constituida por pastizales resecos y sin vegetación) y la zona quemada (de características físicas y espectrales similares), aunque esta pueda delimitarse razonablemente bien.

La Fig. 3, C muestra el resultado de la clasificación no supervisada aplicada sobre las cuatro bandas de la imagen de febrero-1986. Igual que en el caso anterior, las clases espectrales resultantes se redujeron a las cuatro anteriormente mencionadas. En este caso, la confusión se produce entre las mismas clases que en el anterior (zona quemada/borde de marisma) aunque en mucha menor proporción ya que, al corresponder la imagen a febrero (época de encharcamiento de la marisma), el borde se inunda y sus características, tanto morfológicas como espectrales, varían radicalmente.

Por otro lado, se aplicó el mismo tipo de clasificación a las imágenes de los cuatro componentes estudiados (PC1, PC2, PC4 y PC6)(Fig. 3,D). Estos cuatro componentes engloban un alto porcentaje de la varianza de la escena (93.13%), a lo que se añade lo bien delimitados que aparecen los incendios de 1985 y 1981 en los PC1 y PC2, respectivamente. Aun así aparece una cierta confusión entre los pixels correspondientes al incendio de 1985 con zonas de vegetación intacta, así como entre los primeros con los correspondientes al borde de marisma (Vera).

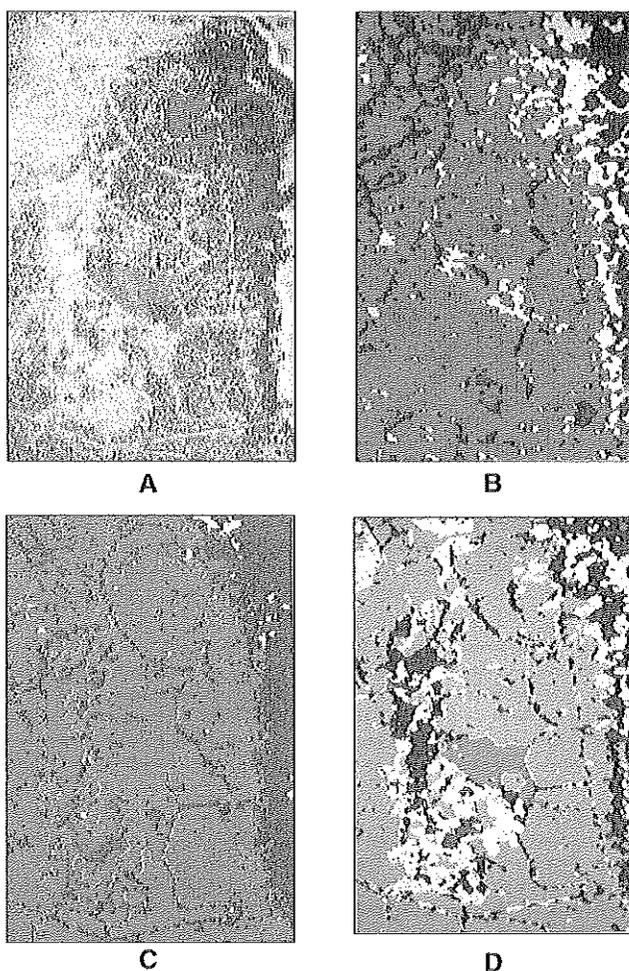


FIGURA 3

Composición RGB realizada con PC2, PCS y PC6, respectivamente (A) y clasificaciones no supervisadas de la imagen de junio de 1984 (B), de febrero de 1986 (C) y de los PC1, PC2, PC4 Y PC6 (D). (A. 99)

Conclusiones

En primer lugar se presenta en la Tabla 2 las prioridades estimadas a la hora de usar las diferentes técnicas analizadas para la caracterización de la zona de estudio.

Se concluye, que el análisis de componentes principales muestra de forma global las diferencias entre ambas subescenas, siempre y cuando esté bien hecho.

Esto significa, que para que esta transformación tenga validez, la subescena elegida debe contener una zona que haya cambiado dentro de un área razonablemente grande que no lo haya hecho. Esto es importante para asegurarse que la inclinación del primer eje (y por tanto de los siguientes) quede determinada por la varianza asociada a coberturas vegetales. El análisis de componentes principales demuestra, por tanto, una gran utilidad en situaciones donde la varianza atribuible a coberturas estáticas domina a aquella atribuible a coberturas dinámicas, como ocurre en este caso de áreas incendiadas. La realización de mapas temáticos sobre las imágenes individuales, así como las procedentes de los componentes principales muestran una vez más este potencial, ya que dicha transformación «resume» y resalta los principales cambios.

Por tanto esta técnica puede considerarse complementaria a la metodología clásica utilizada, desde el momento en que permite eliminar la varianza proveniente tanto de las condiciones externas a la subescena (atmosféricas, de calibración del sensor, etc.), como de ciertas características internas (cobertura vegetal que no ha variado en el tiempo).

Delimitación del incendio de 1985:

PC 1 > PC6 > PC4 >> PC2.

Delimitación del incendio de 1981:

PC2 > PC6 >> PC 1 = PC4.

Regeneración de la vegetación quemada en 1981:

PC6 > PC2 > PC1 > PC4.

Distinción de las unidades morfoedáficas:

PC4 > PC1 > PC6 >> PC2.

Distinción de ambos fuegos y su evolución con el tiempo:

PC6 >>> PC 1 > PC4 > PC2.

TABLA 2

Prioridades en la bondad de resultados obtenidos COII diferentes metodologías

Bibliografía

Byrne, G.R. and Crapper, P.F.: 1979. An example of detection of changes between successive Landsat images by numerical methods in an urban area. *Proceedings of the First Australasian Conference on Remote Sensing (Landsat '79)*, Sydney.

Byrne, G.R.; Crapper, P.F. and Mayo, K.K.: 1980. Monitoring landcover changes by principal component analysis of multitemporal Landsat data. *Remote Sensing of Environment*, 10: 175-184.

Domínguez, L.: 1993. ICONA. Comunicación personal.

- González, R.C. and Wintz, P.:** 1987. *Digital Image Processing*. Addison Wesley, Reading, WA.
- Jenson, S.K. and Waltz, F.A.:** 1979. Principal Component Analysis and Canonical Analysis in Remote Sensing. *Proc. Am. Soc. Photogr., 45th Ann. Meetillg* :337-348.
- Richards, J.A.:** 1984. Thematic Mapping from Multitemporal Image Data Using the Principal Components Transformation. *Remote Sensing of Environment*, 16: 35-46.
- Siljeström, P. and Clemente, L.:** 1987. Evolución edáfica en la vera arcillosa del Parque Nacional de Doñana. *Anales de Edafología y Agrobiología*, vol.XLVI. 9/10: 1089-1101.
- Siljeström, P.; García, L.V. and Clemente, L.:** 1993a. Identificación de patrones geomorfológicos a través de la interpretación de las razones entre bandas TM. *El Cuaternario de España y Portugal* edit. ITGE, vol. 1:235-241.
- Siljeström, P.; Moreno, A.; García, L.V. and Clemente, L.:** 1993b. Diferenciación de unidades edáficas en base a imágenes TM. *Revista de Teledetección*. En prensa nº 7.