Principal Component Analysis and Change Detection
Southern Belize : 2000 - 2011

Abstract

Principal Components Analysis (PCA) can be used by image analysts as a data reduction technique whereby the information content from a number of bands is compressed into a few principal components. In other words, PCA can be used to reduce the dimensionality of the data while minimizing loss of information. In addition, PCA images placed in a two- or three-band color-composite may be more easily interpreted than conventional color-infrared composites.

Methods

A composite was created using bands 3,4, and 5 from both Landsat 5 images resulting in a 6 band multi-date composite (Table 1). Bands 3,4, and 5 were chosen for their differences in spectral qualities; Band 3 is the visible Red channel, band 4 is the Near Infrared channel and band 5 is the Mid Infrared hannel. Using these 3 spectral channels reduced data redundancy that most often occurs between similar spectral channels of an image, also reducing data volume and processing time. Other researchers have demonstrated the benefits to reducing the spectral redundancy of an image in this way (Hayes and Sader, 2001).

Composite Bands	**Date - Band**	**Spectral**
1	2001 - 3	Visible
2	2001 - 4	NIR
3	2001 - 5	SWIR
4	2011 - 3	Visible
5	2011 - 4	NIR
6	2011 - 5	SWIR

Table 1: Multi-date composite

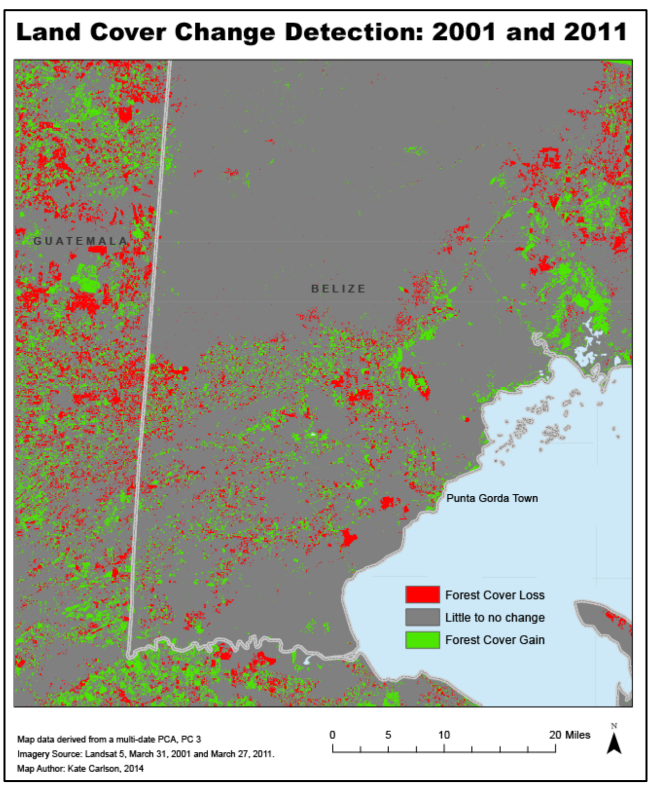
During the first pass of the PCA algorithm, a covariance matrix of the input bands is computed. This covariance matrix is then used during the second pass of the algorithm to compute the principal components or eigenvectors (Jensen, 2007). The multi-date composite was then transformed into a principal component dataset consisting of 6 principal component images (or bands), eigenvalues describing the percentage of spectral variance each principal component band contains, and a table of eigenvector factor scores that indicate how much each band contributed to each of the 6 principal components. Factor loadings were computed for each principal component (Table 2).

Band	PC 1	PC 2	PC 3	PC 4	PC 5	PC 6
2001(B3) – 1	0.467876274	-0.677417698	0.348274946	0.012662069	0.386377859	0.226877352
2001(B4) – 2	0.953184563	0.263930678	-0.035239921	-0.143102583	0.007491865	0.001826225
2001(B5) – 3	0.9329735	-0.28173076	0.222596018	-0.0189497	-0.014421652	-0.008523714
2011(B3) – 4	0.571428476	-0.692837595	-0.227640441	-0.009176281	0.344144256	-0.151969336
2011(B4) – 5	0.949681659	0.294042441	0.016024569	0.106347658	0.008720082	-0.001200262
2011(B5) – 6	0.9338168	-0.321107983	-0.155540557	0.019765844	-0.015606977	0.007017686

Table 2: Factor Loadings

Analysis

The factor loadings are evaluated to determine which of the 6 principal components demonstrates land cover changes between 2001 and 2011. The algebraic signs are used as indicators of change by evaluating the factor load value of a band for 2001 and the associated band for 2011 (example, 2001(B3) and 2011(B3)). The greater the numeric difference between the 2 dates will be an indication of which principal component that has the highest ability to threshold changes in the land cover; forest clearing, no-change, and regrowth (Hayes and Sader, 2001). Principal Component 3 has been chosen to most effectively show these types of land cover changes (Figure 2c). After investigating the pixel values, threshold indicators were chosen that best represent areas that lost forest cover in black and areas that gained forest cover shown in white.



Map Figure 1

Sources:

Jensen, John (2007). *Remote Sensing of the Environment: An Earth Resource Perspective*. Upper Saddle River, NJ: Pearson/Prentice Hall.

Hayes, Daniel and Sader, Steven (September 2001). Comparison of Change Detection Techniques for Monitoring Tropical Forest Clearing and Vegetation Regrowth in a Time Series. *Photogrammetry Engineering & Remote Sensing, vol. 67, No. p, pp. 1067-1075*.