Supplementary Material

Introduction

In this document we explain in detail the PC axis selection. In addition, the seasonality of a PC component is quantified to further support our method.

PC axis selection

To automatically monitor forest change, a PC axis is automatically selected per pixel. We sampled 100 locations where the forest remain steady from each study site to find a criterion to automatically select the PC that is sensitive to deforestation. Firstly, the PCA is applied to the combined time series of locations where the validation data indicates no deforestation (100 points for the Bolivian site and 100 points for the Brazilian site). The data is organised as $X = \left[X_{ij}\right]_{(N\times 6)}$, which contains grayscale values at times $i = 1, ..., N$ for spectral bands $j = 1, ..., 6$, now with N the sum of the lengths of all time series ($100 \times 444 = 44400$ for the Bolivian site and $100 \times 225 = 25500$ for the Brazilian site).

The PC loadings of the first 4 PCs are shown in Fig. 1. For the Bolivian site, the loadings of the validation time series that contain no deforestation events are shown in Fig. 1a. The loadings of the first PC (PC1) capture the positive correlations between bands 1-3, 5 and 7. PC1 may be interpreted as the general brightness. The loading of band 4 is low and has an opposite sign to the other bands, indicating high correlation between bands 1-3, 5 and 7, and low or negative correlation between band 4 and each of the other bands. This corresponds to the spectral reflectance of dry and green vegetation to different wavelengths

(a) Bolivian site (b) Brazilian site

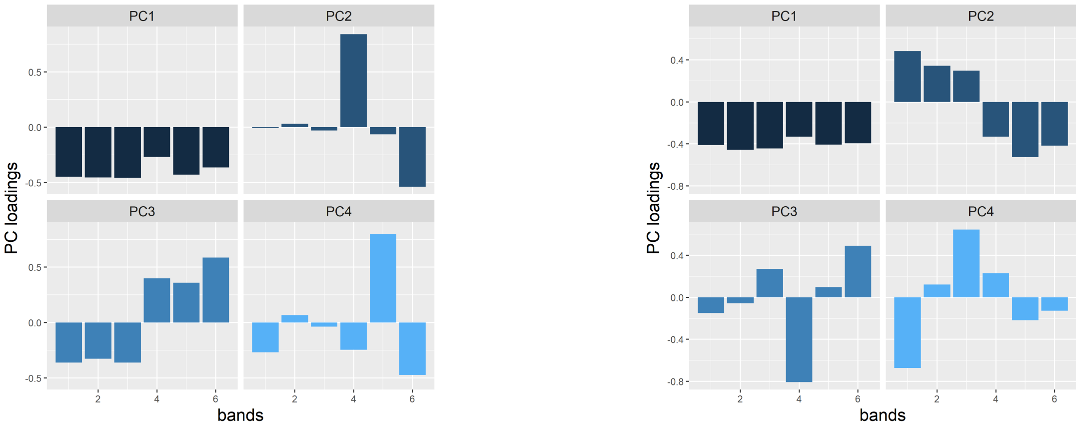

Figure (1) The PC loadings for all the validation time series containing no deforestation events at each site.

in (Fig. 2, Clark et al., 2003), Clark et al. (2003) found that the dry vegetation reflectance is higher than green vegetation reflectance to the visible and short-wave infrared (SWIR) light, but lower for the NIR light (0.76 - 0.9 μm).

The second PC (PC2) shows that the loadings of bands 1, 3 and 5 are close to zero. The highest loading is that for band 4, followed by band 2 and band 7 which has a different sign. Fig. 2 shows that from dry to green vegetation, the reflectance of the NIR light rise drastically, while dropped the most to the SWIR light of band 7 (2.08 - 2.35 μm). Thus we interpret PC2 as the vegetation greenness fluctuation, which often contains seasonality that is affected by climate and can hinder the identification of deforestation.

The third PC (PC3) indicates a contrast between visible and IR (near infrared and shortwave infrared) bands. As the IR bands are sensitive to water, PC3 relates to the wetness. The seasonality effects explained by PC2 and the general brightness explained by PC1 are exempt from PC3. We thus hypothesise the PC3 to be suitable for deforestation monitoring.

The PC loadings of the time series that consist no deforestation events at the

Brazilian site (Fig. 1b) support the finding above. The PC1 captures all the positive correlations between bands, with band 4 also highly correlated to other bands. As the forest of this site experiences weak seasonality, the interpretation regarding a lack of correlation between band 4 and other bands in PC1 of the Bolivian site is justified. The loadings indicate a contrast between the IR bands and the visible bands is strongest in PC2. The high PC loading of band 4 is presented in PC3.

Based on these interpretation of the loadings, we select the PC axis that maximise contrast between the visible and IR bands, i.e., PC3 of Fig. 1a and PC2 of Fig. 1b to monitor deforestation. The PC axis selection criterion (eq.2 of section 3.2 of the manuscript) is established based on this character.

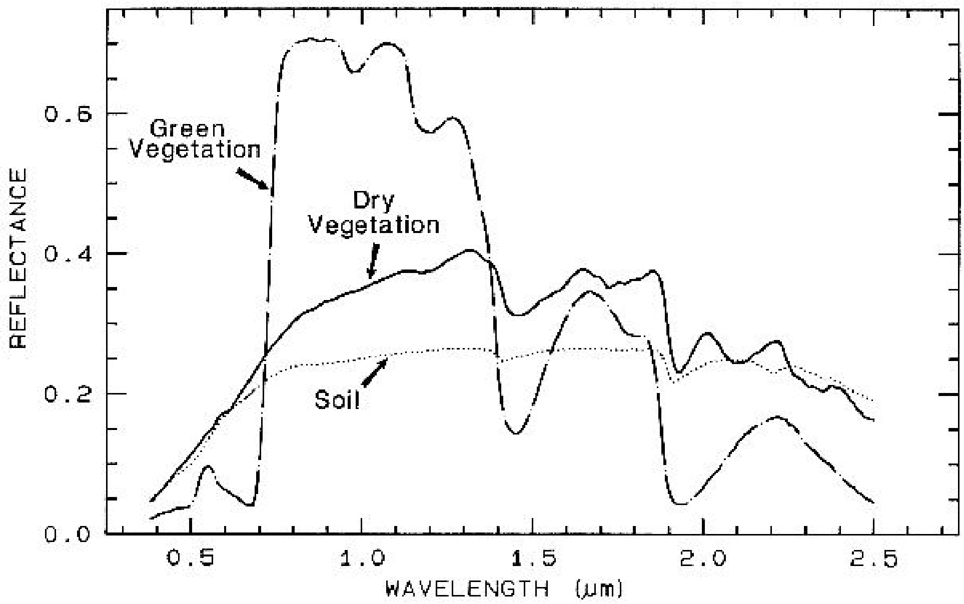

Figure (2) The spectral reflectance of healthy vegetation, brown vegetation, and soil, source: Clark et al. (2003)

Seasonality analysis

This study hypothesises that the seasonality is contained in some PC components and is filtered out from SRI. To evaluate to which extent a PC captures a seasonal signal and separates it from the deforestation signal, the amount of seasonality in the PC that indicates vegetation fluctuation (see "PC axis selection", called PC-greenness) was quantified. We first detected annual seasonality of a time series using a first-order harmonic model. The coefficient of determination (R^2) was used to assess the variance explained by the harmonic model. Then, we attempted to discover periodicity of a time series by calculating the smoothed periodogram, which gives the spectral density at different frequencies (Chatfield, 2016), based on AR (autoregressive) models fitted to each time series. After omitting the missing data, time series were aggregated to monthly values for this.

We used the validation points from the Bolivian site that contain no deforestation to inspect the seasonality in PC-greenness. PCA was applied to each pixel. When fitting a first order harmonic model to the PC-greenness, the median of the R^2 over all the time series is about 0.5, meaning that in half of the cases at least 50% of the variation of the PC-greenness can be explained by annual seasonality. To avoid missing data in the periodogram analysis, we aggregated the time series from 2003 - 2014 monthly, as the data of this period is the most dense. The missing values of the monthly time series were linearly interpolated. For 90% of the time series, the highest spectral density are at an annual frequency. The PC greeness and all the periodograms can be reproduced using scripts available at https://github.com/mengluchu/multibandsBFAST/seasonality.

References

Chatfield, C., 2016. The analysis of time series: An introduction. CRC press.

Clark, R. N., Swayze, G. A., Livo, K. E., Kokaly, R. F., Sutley, S. J., Dalton, J. B., McDougal, R. R., Gent, C. A., 2003. Imaging spectroscopy: Earth and planetary remote sensing with the USGS tetracorder and expert systems. Journal of Geophysical Research: Planets 108 (E12).