Detection of Breakpoints in Global NDVI time series

Rogier de Jong [a,b,*], Jan Verbesselt [a], Michael E. Schaepman [c], Sytze de Bruin [a]

[a] Laboratory of Geo-Information Science and Remote Sensing, Wageningen University, The Netherlands
[b] ISRIC – World Soil Information, Wageningen, The Netherlands
[c] Remote Sensing Laboratories, University of Zurich, Switzerland

Abstract – Continuous global time series of vegetation indices, which are available since early 1980s, are of great value to detect changes in vegetation status at large spatial scales. Most change detection methods, however, assume a fixed change trajectory – defined by the start and end of the time series – and a linear or monotonic trend. Here, we apply a change detection method which detects abrupt changes within the time series. This Breaks For Additive Season and Trend (BFAST) approach showed that large parts of the world are subjected to trend changes. The timing of the breakpoints could in some cases be related to satellite changes, but also to large-scale natural influences like the Mt. Pinatubo eruption. Shifts from greening to browning (or vice versa) occurred in 15% of the global land surface, which demonstrates the importance of accounting for trend breaks when analyzing long-term NDVI time series.

Keywords: GIMMS NDVI, time series analysis, trend breaks, BFAST

1. INTRODUCTION

Land management and policy making requires information on vegetation status on ever larger temporal and spatial scales. Time series of Normalized Difference Vegetation Index (NDVI) imagery provide repeatable measurements at scales at which climate- and human-induced changes take place and go back until the early 1980s. Detecting changes within the time series is an important step to be taken before attributing drivers of change and assess the impact on environmental services provided by the land.

Many regional and few global studies assessed vegetation changes using NDVI time series. The direction and rate of change – together referred to as *trend* – are usually determined by the slope of a linear regression model. The trajectory of change is predetermined by the time-span of the dataset and trends are assumed linear or monotonic. In many areas in the world, however, trends are known or expected not to behave according to these assumptions, but show in fact changes (e.g. Slayback *et al.*, 2003). If this happens, methods which are not designed to account for this are likely to underestimate the actual rate of change, especially if the direction (greening vs browing) alters.

Changes within NDVI time series can be divided into three major classes (Verbesselt *et al.*, 2010a): seasonal changes and gradual or abrupt trend changes. Here, we apply a change detection method which is able to detect abrupt changes and we use the terms trend breaks or breakpoints as synonyms. A global NDVI dataset (1981–2006) is used to investigate in which regions trend breaks occurred and especially where these involved shifts from greening to browning (or vice versa). Outliers due to cloudiness were removed from the NDVI data and subsequently change detection was applied. Trend slopes between consecutive breakpoints were tested for significant difference from zero. We conclude that 43% of the earth surface has been subjected to NDVI trend changes and that 15% of the surface experienced sign changes of the slope, i.e. changes from greening to browning or vice versa. As such, we demonstrate the importance of accounting for trend breaks.

2. DATA AND METHODS

2.1. NDVI data

In an effort to monitor fluctuations in vegetation and understand interactions with the environment, the National Oceanic and Atmospheric Administration (NOAA) has been collecting images of vegetation condition using Advanced Very High Resolution Radiometer (AVHRR) sensors. The non-linear combination of red and near infrared radiance $(NIR - RED)/(NIR + RED)$, known as NDVI, exhibits a strong relationship between spectral radiance and green biomass and is commonly used for vegetation assessments. Several AVHRR NDVI products are available, of which we use the Global Inventory for Mapping and Modeling Studies (GIMMS). The data spans from 1981 through 2006, which is the longest run available. It has a temporal resolution of two weeks (24 scenes/year) and a spatial resolution of ~8 kilometers. Errors in NDVI introduced from orbital drift were largely (~90%) eliminated and it has been corrected for atmospheric effects (Tucker *et al.*, 2005). The transitions between platforms my cause abrupt changes in the data (de Beurs and Henebry, 2005), but these are expected not to affect trend slopes (i.e. gradual changes) in vegetation index (Kaufmann *et al.*, 2000). A maximum value compositing (MVC) technique was used to minimize cloud contamination during GIMMS processing and we applied an harmonic analysis algorithm (Roerink *et al.*, 2000) to remove remaining noise in areas with persistent cloud cover. In this way, the original data is maintained – only outliers are replaced – and the risk of detecting trend changes caused by persistent cloud cover is minimized.

In a next step, we masked out the non-vegetated areas (yearly mean NDVI < 0.10), non-terrestrial pixels and pixels corresponding to the International Geosphere-Biosphere Programme (IGBP) land cover classes wetlands, urban and ice (Loveland *et al.*, 2000). Tropical evergreen forest was also excluded, as the use of NDVI in these regions is disputed because of saturation of the signal and corresponding low signal to noise ratios. For practical reasons (i.e. computation time), the dataset needed to be further reduced to a size that allowed processing on 24 nodes of a high performance cluster. Accordingly, a 3x3 kernel was applied to process the center pixel out of 9. We preferred this approach over resampling to a coarser resolution because the original GIMMS data – which we believe is the best possible representation of the original 1 km^2 data – is preserved. The resulting dataset consisted of 229 388 pixels.

2.2. Detecting breakpoints

NDVI time series typically contain a strong seasonal component linked with the growing seasons of vegetation being monitored. Most of the existing change detection techniques are unable to account for seasonal variation and analyze time series by aggregating data by year or season or compare specific periods (e.g. summer) between years (Coppin *et al.*, 2004). We investigate the application of a generic change detection method that accounts for seasonality and enables the detection of trend change within the time series (Verbesselt *et al.*, 2010a,b). The key concepts of this method for detecting Breaks For Additive Seasonal and Trend (BFAST) is explained below, but for technical elaboration the reader is referred to above-mentioned publications.

BFAST is an iterative algorithm that combines the decomposition of time series into seasonal, trend, and remainder component with methods for detecting changes. An additive decomposition model is used to iteratively fit a piecewise linear trend and a seasonal model (Haywood and Randall, 2008). The general model is of the form:

$$Y_t = T_t + S_t + e_t \; : \; t \in T \qquad (1)$$

where, at time t (in the time series T), Y_t is the observed NDVI value, T_t is the trend component, S_t the seasonal component en e_t the remainder component which contains the variation beyond what is explained by T_t and S_t.

Each time before fitting the T_t and S_t components, it is tested whether abrupt changes are occurring. The ordinary least squares residuals-based moving sum (MOSUM) test, is selected for this purpose (Zeileis and Kleiber, 2005). If the test indicates significant change (α = 0.05), the breakpoints are estimated using the method of Bai and Perron (2003). This method minimizes the Bayesian Information Criterion (BIC) to determine the optimal number of breaks and iteratively minimizes the residual sum of squares to estimate the optimal break positions. The motivation for this procedure is given by Verbesselt *et al.* (2010a), who also provide a validation for the algorithm using both simulated time series and MODIS NDVI data for Australian pine plantations. For illustration, Figure 1 shows the decomposition and breakpoints for a simulated NDVI time series of 10 years. The seasonal component is simulated by a sine function (amplitude 0.25 and period 1 year), the trend component consists of three segments (with slopes β_1= 0.005; β_2 = -0.025 and β_3 = 0.075) and the remainder is normally distributed with a standard deviation of 0.05.

The only parameterization required for the BFAST procedure is the maximum number of breakpoints m_{max} and the minimal time between breakpoints h. Preventing the algorithm from identifying practically insignificant changes, we set h to 3 years and m_{max} to 3 breaks. All analyses were performed using R statistical software.

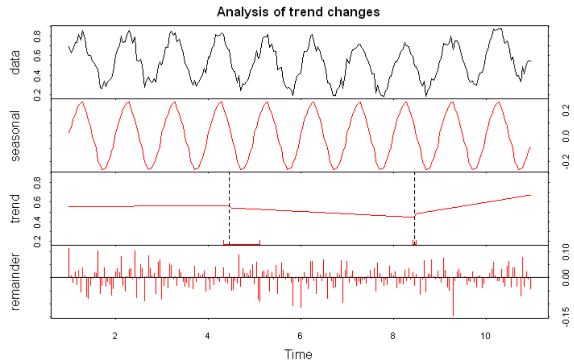

Figure 1. Example of trend break analysis for a simulated NDVI time series of 10 years. The seasonal and remainder components vary around zero and, as a result, the trend component shows the trend in NDVI. The trend component shows a greening and a browning segment. The slope of the remaining, first, segment is not significant (α = 0.10).

2.3 Analysis of trends and breakpoints

For each segment i (i.e. period between consecutive breakpoints) the slope β_i obtained from the linear model is assessed for significance using an analysis of variance (ANOVA) test under the hypothesis H_0: $\beta_i = 0$ and H_A: $\beta_i \neq 0$. Only significant slopes (α = 0.10) were adopted as indicators for greening ($\beta_i > 0$) or browning ($\beta_i < 0$).

It has been suggested that platform changes cause discontinuities in the data. These might cause trend changes in the NDVI time series. For this reason, the positions in time of the breakpoints were compared to the dates of instrument change within the GIMMS dataset by evaluating the frequency around these dates against the expected frequency given a random distribution over time.

3. RESULTS

Following the described approach, large parts of the Earth's surface were detected to have experienced NDVI trend changes during the 1981 – 2006 period. Out of the processed pixels, 40.9% show zero, 21.3% one, 17.7% two and 20.1% three breakpoints. Most breakpoints were detected in Australia, Argentina, large parts of Europe, the northern parts of Mexico, the Sahel and Africa's south-west. According to the IGBP classification, the majority of all regions with multiple breakpoints is shrub land (followed by grassland) and many of these regions are semi-arid.

Most breakpoints separate segments with a significant slope from segments with an insignificant slope. However, for ~15% of the total area the slopes show both a period of significant greening and a period of significant browning. Figure 2 shows a map of the occurrence of these mixed trends, irrespective of the change trajectories or trend slopes involved, and Table 1 summarizes all detected trends. Large areas which since 1981 have only shown greening are found on the Northern Hemisphere, conspicuously (north-)eastern Europe and parts of the Siberian, Alaskan and Canadian tundra. In the same way, browning trends are found in northern Argentina, parts of sub-equatorial Africa, China and the northern boreal forests. It appeared that almost a third of the global land area experienced significant trends for the entire time-span of the AVHRR record (26 years), while this increases to 57% if trend breaks are taken into account. In this respect, greening has a larger share than browning: it represents a larger area and 80% of the greening slopes are significant for the entire period of 26 years (versus 60% of the browning slopes).

The application of BFAST provided estimates of the time of abrupt changes. Figure 3 shows the frequency distribution of these breaks in bins of six months. Also, the expected number of breaks given a random distribution over time is shown (dashed line). It appears that several sensor changes coincide with periods in which many trend breaks were detected. Two peaks reach almost 170% of the expected number and have a higher frequency than any other period of six months between 1981 and 2006. These dates correspond to the transition dates between NOAA platforms 9, 11 and 14.

4. DISCUSSION

The presented approach is capable of separating between periods of browning and periods of greening, which appear to have occurred in several regions, including many semi-arid ecosystems. An example is the Sahel where positive slopes are found in the first part of the time series and negative slopes for the better part of the 90s and later. For this and many other regions in the world ample studies have related the NDVI trends to possible drivers. It is beyond the scope of this paper to review the findings.

Table 1. Results from trend change analysis expressed in area (million km^2 and percentage of total). For greening and browning trends a subdivision is made into area where a significant trend ($\alpha = 0.10$) was found and area where the trend was found for the entire time-span of 26 years.

Description	**Area**	***Pct***
Masked	40.49	*27.2%*
No significant trend	23.62	*15.9%*
Only greening	40.15	*27.0%*
of which 26 yrs	31.50	*21.2%*
Only browning	23.03	*15.5%*
of which 26 yrs	14.83	*10.0%*
Both greening and browning	21.55	*14.5%*
Total	148.84	

Atmospheric correction of the GIMMS dataset accounted for aerosols injected into atmosphere by volcano eruptions, like the Mount Pinatubo eruption in June 1991 (Slayback *et al.*, 2003). However, discontinuities in the NDVI time series might result from actual vegetation response to a period of global cooling. An overrepresentation of breaks (140%) was found after the mentioned eruption (Figure 3). The spatial distribution of breakpoints does not show a higher concentration at closer distances to the volcano, but cooling effects attributable to the eruption have been reported around the world (Soden *et al.*, 2002).

The results from the trend break analysis showed that almost 15% of the global area has to do with trend shifts from greening to browning or vice versa. In these regions, common change detection methods are likely to average out this mixed trend effect, resulting in underestimation of the trend significance. These areas might be labeled stable while changes have occurred in reality. The longer a time series is, the more likely it is that this effect conceals actual trends.

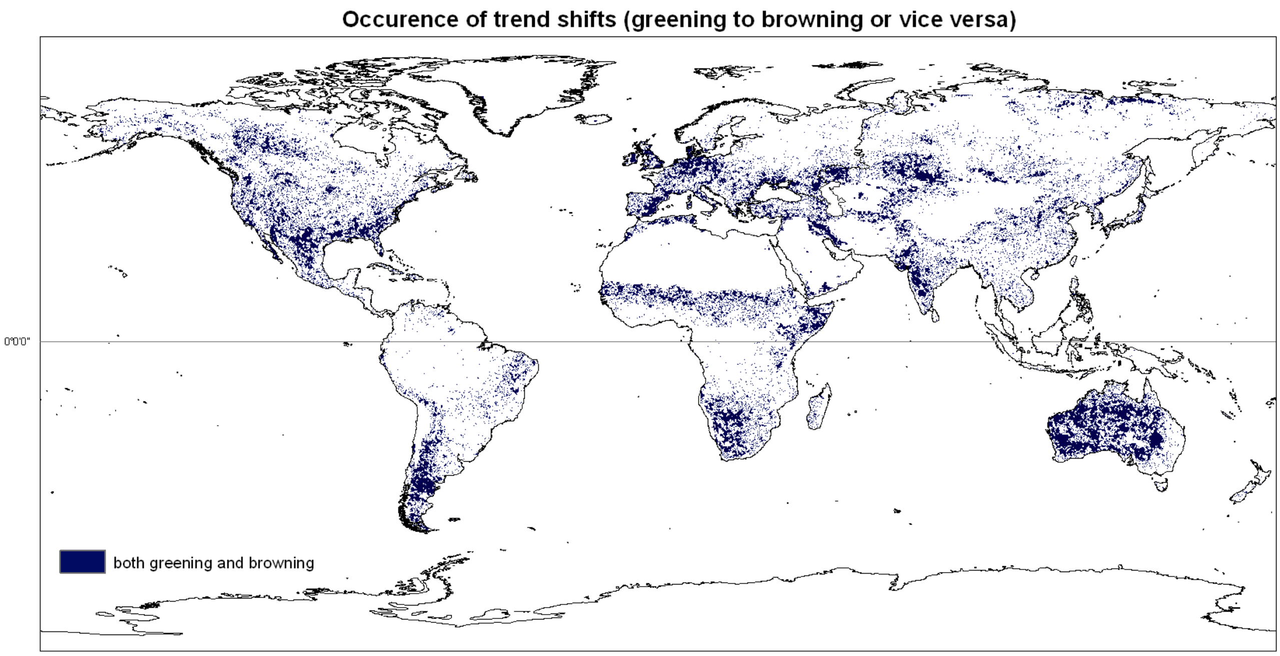

Figure 2. Spatial distribution of detected trend shifts (greening to browning or vice versa).

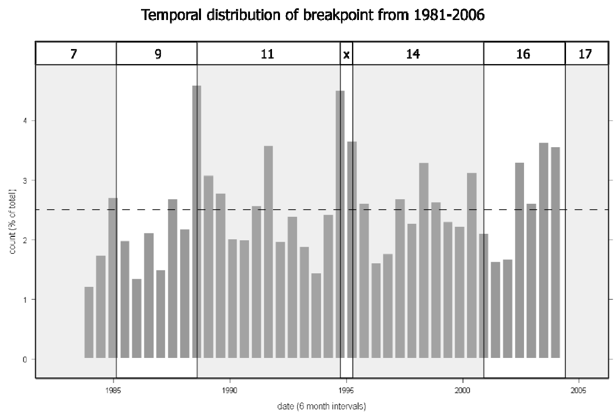

Figure 3 Relative frequency (y-axis) of detected breakpoints. Dates are aggregated to bins of 6-month between January 1984 and December 2003 (x-axis). The dashed line represents the expected frequency if breakpoints would be randomly distributed over time. The upper bar shows the timeline of the different NOAA platforms used to acquire the data, in which x denotes the NOAA-9 descending node period, which was used due to malfunction of NOAA-11 and failure of NOAA-13 to achieve orbit (Tucker *et al.*, 2005).

Platform changes are known to cause discontinuities within the GIMMS data (Cracknell, 1997), but these were found not to affect NDVI trends in most biomes. The likeliness of contamination because of these transitions is highest in biomes with a sparse vegetation cover and relatively light-colored soils (Kaufmann *et al.*, 2000). This is consistent with the high representation of detected trends in shrub land and grassland biomes around some of the transition dates (Figure 3). It is therefore conceivable that the timing of the breaks is influenced by platform changes, although this would not necessarily affect the detected slopes. These might be affected by gradual changes in solar zenith angle due to orbital drift of the platforms, but the GIMMS data has been thoroughly corrected for this effect (Tucker *et al.*, 2005).

5. CONCLUSIONS

This study illustrated the capability of the BFAST approach to detect changes in long-term NDVI time series and revealed both a temporal and a spatial pattern in breakpoints. These breaks mainly separated insignificant from significant trends, but in almost 15% of the global area the breaks involved a shift form greening to browning or vice versa. In these cases it is specifically important to account for trend breaks in NDVI trend analysis, as this mixed trend effect might render linear trend analysis unsuitable.

The timing of the breaks needs careful interpretation as platform changes might cause an overrepresentation of breakpoints around the associated dates. The intermediary trend slopes are not likely to be affected by these platform changes and analysis of slopes which significantly differ from zero provides insight in which periods contributed most to greening and/or browning in the target area.

ACKNOWLEDGEMENTS

We are grateful to the GIMMS group for providing their data and to the ESG-HPC people of Wageningen University for facilitating model calculations. In special we would like to thank Wietse Franssen for his assistance.

REFERENCES

J. Bai and P. Perron, "Computation and analysis of multiple structural change models," Journal of Applied Econometrics, vol 18(1), p.p. 1-22, 2003.

K.M. de Beurs and G.M. Henebry, "A statistical framework for the analysis of long image time series," International Journal of Remote Sensing, vol 26(8), p.p. 1551-1573, 2005.

P. Coppin, I. Jonckheere, K. Nackaerts, B. Muys and E. Lambin, "Digital change detection methods in ecosystem monitoring: a review," International Journal of Remote Sensing, vol 25(9), p.p. 1565-1596, 2004.

A.P. Cracknell (1997). The advanced very high resolution radiometer (AVHRR). London, Taylor & Francis.

J. Haywood and J. Randall (2008). Trending seasonal data with multiple structural breaks. NZ visitor arrivals and the minimal effects of 9/11. Research report 08/10. Victoria, New Zealand, University of Wellington. vol p.p. 26.

R.K. Kaufmann, L. Zhou, Y. Knyazikhin, V. Shabanov, R.B. Myneni and C.J. Tucker, "Effect of orbital drift and sensor changes on the time series of AVHRR vegetation index data," IEEE Transactions on Geoscience and Remote Sensing, vol 38(6), p.p. 2584-2597, 2000.

T.R. Loveland, B.C. Reed, J.F. Brown, D.O. Ohlen, Z. Zhu, L. Yang and J.W. Merchant, "Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data," International Journal of Remote Sensing, vol 21, p.p. 1303-1330, 2000.

G.J. Roerink, M. Menenti and W. Verhoef, "Reconstructing cloudfree NDVI composites using Fourier analysis of time series," International Journal of Remote Sensing, vol 21(9), p.p. 1911 - 1917, 2000.

D.A. Slayback, J.E. Pinzon, S.O. Los and C.J. Tucker, "Northern hemisphere photosynthetic trends 1982-99," Global Change Biology, vol 9(1), p.p. 1-15, 2003.

B.J. Soden, R.T. Wetherald, G.L. Stenchikov and A. Robock, "Global cooling after the eruption of Mount Pinatubo: A test of climate feedback by water vapor," Science, vol 296(5568), p.p. 727-730, 2002.

C. Tucker, J. Pinzon, M. Brown, D. Slayback, E. Pak, R. Mahoney, E. Vermote and N. El Saleous, "An extended AVHRR 8km NDVI dataset compatible with MODIS and SPOT vegetation NDVI data," International Journal of Remote Sensing, vol 26(20), p.p. 4485-4498, 2005.

J. Verbesselt, R. Hyndman, G. Newnham and D. Culvenor, "Detecting trend and seasonal changes in satellite image time series," Remote Sensing of Environment, vol 114(1), p.p. 106-115, 2010a.

J. Verbesselt, R. Hyndman, A. Zeileis and D. Culvenor, "Phenological change detection while accounting for abrupt and gradual trends in satellite image time series," Remote Sensing of Environment, vol 114(12), p.p. 2970-2980, 2010b.

A. Zeileis and C. Kleiber, "Validating multiple structural change models-a case study," Journal of Applied Econometrics, vol 20(5), p.p. 685-690, 2005.