20th International Congress on Modelling and Simulation, Adelaide, Australia, 1–6 December 2013
www.mssanz.org.au/modsim2013

Sensitivity of the BFAST algorithm to MODIS satellite and vegetation index

L.M. Watts [a] **and S.W. Laffan** [a]

[a] *School of Biological, Earth and Environmental Sciences, University of New South Wales, Sydney, NSW, 2052, Australia*
Email: laura.watts@unsw.edu.au

Abstract: The Breaks for Additive Seasonal and Trend (BFAST) algorithm combines the additive decomposition of time series with abrupt change detection. It potentially allows for the effective use of the temporal detail available in satellite image time series for examining vegetation response patterns across regional extents, accounting for variation at the seasonal scale while detecting changes in the long term trends. While BFAST has been validated using NDVI time series (Verbesselt *et al.* 2010a), its sensitivity to different parameters and input data has not yet been assessed. Understanding the effects of the data source and type and the variation of the algorithm's user-defined parameters on the timing and number of abrupt changes it detects will allow it to be used more effectively. This study aims to assess the effects of the satellite used for data collection, the vegetation index (EVI or NDVI) and varying the value of a BFAST argument called the h parameter, which controls the potential number of trend breaks detected, on abrupt change detection for a study area in the Paroo region of north-western New South Wales, Australia.

Moderate Resolution Imaging Spectroradiometer (MODIS) EVI and NDVI time series were decomposed for 165 sample points chosen to include a range of land cover types in the study region. The effect of the MODIS satellite (Aqua or Terra) used to collect data was assessed by comparing the number of breaks detected and their timing between time series derived from each satellite. The effect of changing the h parameter was assessed by comparing the similarity in the length of time periods with a significant trend component slope between different values of h (1/3, 1/5 and 1/7).

The timing of detected breaks was affected by the satellite used to collect data, despite the visual similarity of Aqua and Terra time series. Greater certainty in the timing of breaks was achieved when using smaller values for the h parameter. Of the three factors tested, the vegetation index had the greatest impact on the timing and number of breaks in the long term trend detected by BFAST. The effect of the vegetation index was dependent upon the h parameter used, and the effects of both the h parameter and the satellite varied between EVI and NDVI. These results suggest a moderate sensitivity of the BFAST algorithm to all three of these factors, and also an interaction between them. This should be taken into consideration when using BFAST for long term vegetation change detection.

Keywords: *MODIS, EVI, NDVI, time series, change detection, vegetation dynamics*

1. INTRODUCTION

Understanding long term trends in vegetation response to available moisture, including abrupt changes caused by disturbance events, is important in environmental fields such as climate change modelling, carbon accounting and natural resource management. Remotely sensed imagery provides an effective means of detecting change across regional extents. Time series analysis of vegetation indices is commonly used to model changes in vegetation response to changing climatic conditions and disturbances (de Jong *et al.* 2011; Weiss *et al.* 2004).

Few change detection methods are available which exploit the full temporal detail available in image time series derived from modern satellites (Verbesselt *et al.* 2010b). Most available methods either model the entire time series to simplify the identification of phenological indicators (Bradley *et al.* 2007; Galford *et al.* 2008; Jönsson and Eklundh 2002), or extract these key dates and other statistics from a decomposed time series (Hill and Donald 2003; Reed *et al.* 1994).

In response to the limitations of available time series analysis methods, the Breaks for Additive Seasonal and Trend (BFAST) algorithm has been developed to identify long term trends and abrupt changes (breaks) in time series while explicitly accounting for seasonal variation (Verbesselt *et al.* 2010a). It allows for the detection of trend changes which may otherwise be masked by seasonal variability. BFAST can also be used to detect shifts in land surface phenology while accounting for the underlying trend (Verbesselt *et al.* 2010b).

BFAST combines change detection with the additive decomposition of the signal into trend, seasonal and remainder components (Figure 1). The algorithm iteratively fits piecewise linear trend and seasonal models to a time series. The model is of the general form: $Y_t = T_t + S_t + e_t$ where Y_t is the observed data at time t, T_t is the trend component, S_t is the seasonal component and e_t is the remainder, or residual component.

The intercept and slope of the trend component model are used to derive the magnitude and direction of breaks. One of the parameters for BFAST is the 'h' parameter. It determines the minimal segment size between potential breaks in the trend component and is equal to the number of observations in a segment of this size divided by the total length of the time series.

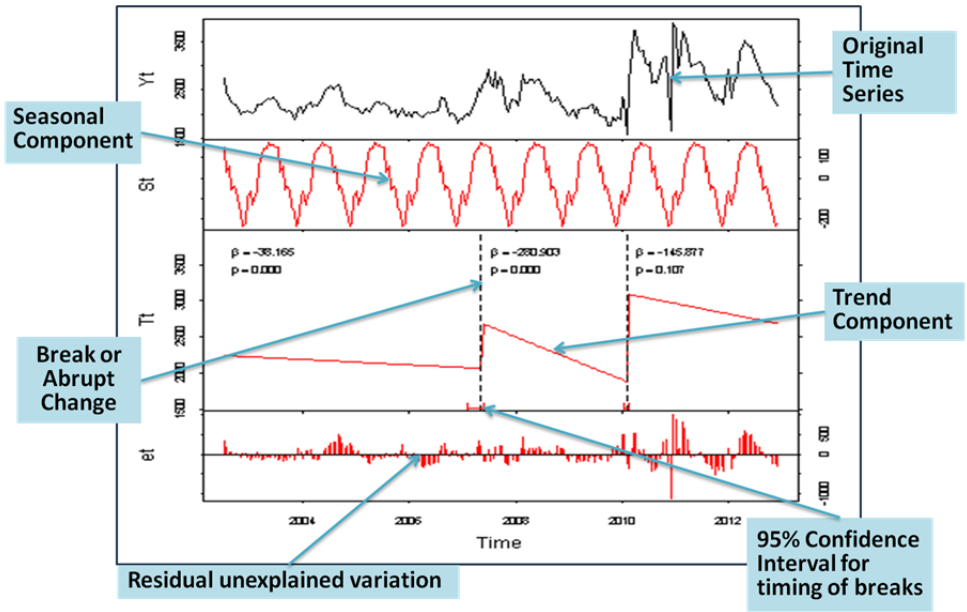

Figure 1. The components of the graphical output from time series decomposition using the BFAST algorithm. The slope (β) and P values shown relate to the linear model fit to the deseasonalised time series ($Yt - St$) (not shown), which is used to derive the trend component.

Vegetation Indices such as the Enhanced Vegetation Index (EVI) and the Normalised Difference Vegetation Index (NDVI) are used as indicators of vegetation vigour. NDVI saturates in areas with high green biomass and is susceptible to interference from soil background reflectance (Carlson and Ripley 1997; Huete 1988). EVI was developed to improve sensitivity in high biomass regions and includes a background reflectance correction factor (Huete *et al*. 2002).

Vegetation Index products derived from data collected by the Moderate Resolution Imaging Spectroradiometer (MODIS) sensor are available at high temporal resolutions appropriate for regional time series analysis and are increasingly used in vegetation change studies (Jacquin *et al.* 2010; Sakamoto *et al.* 2005). The MODIS instrument is on board two of NASA's satellites, Aqua and Terra. The Terra satellite crosses the equator moving from north to south at 10.30 am, whereas the Aqua satellite crossing is from south to north at 1:30pm. Data collection by Terra began in 2000, two years earlier than Aqua. Previous research has found evidence that the blue band of the Terra MODIS sensor is affected by degradation (Wang *et al.* 2012). This may impact on change detection, particularly using EVI, as its calculation includes blue band reflectance. The elimination of this effect is a possible advantage of using Aqua data over Terra.

BFAST has considerable potential for detecting abrupt changes in long term greening and browning trends across regional extents. However, it is first necessary to consider the potential effects of the source and type of time series data and the number of breaks which can be detected, which is controlled by the h parameter. The aim of this study is to assess the effects of the satellite used for data collection, the vegetation index (EVI or NDVI) and varying the value of the h parameter on the detection of abrupt changes by the BFAST algorithm for a study area in the Paroo region of north-western New South Wales.

2. THE PAROO REGION

The Paroo region is in far north-western New South Wales and south-western Queensland, Australia (Figure 2). With an annual average rainfall of less than 300 mm (Australian Bureau of Meteorology 2013) this region lies within a semi-arid to arid zone. The Paroo catchment is approximately 73,600 km^2 in area (Kingsford and Porter 1999). The Paroo River is ephemeral, with a highly variable flow regime (Young and Kingsford 2006). It is the last remaining unregulated river in the Murray-Darling Basin, and its naturally variable flows drive the region's ecology.

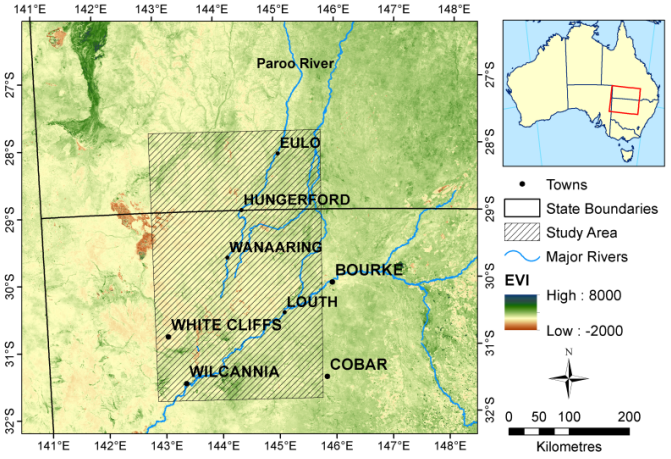

Figure 2. The Paroo Region of north-western New South Wales. The satellite image shows EVI data for 09 May 2010.

Vegetation cover in the study region is sparse compared with humid regions. Dominant groundcover species include canegrass, sedges and lignum. The impact of human activities is minimal compared with other catchments in NSW due to low stocking rates and limited water resource development (Brock *et al.* 2006). The Paroo region was chosen for this study because of its low local relief and relatively simple vegetation communities, which would likely simplify the identification and comparison of abrupt changes.

3. METHODS

3.1. Data Preparation and Sampling

The datasets used in this study were MODIS 16-day vegetation index products from collection 5, downloaded from the MODIS website (http://modis.gsfc.nasa.gov). For the Terra product (MOD13Q1) images covered the period from 18-Feb-2000 to 18-Dec-2012. The Aqua product (MYD13Q1) covered the period 4-Jul-2002 to 26-Dec-2012. The MODIS vegetation index values range from -2000 to 10,000, as they are stored as integers to reduce file sizes. The images were mosaicked, and the study region extracted for each time slice. The vegetation index value for each time slice at a set of sample locations was extracted and stored in table format.

Two sample datasets were created, both chosen to sample different land cover types. The first comprised 30 locations, identified using satellite imagery available through Google Earth, referred to hereafter as the 30 test samples. The second consisted of 135 sample points including sites visited during reconnaissance fieldwork, and groundcover observation sites from the NSW Office of Environment and Heritage Remote Sensing Section. These points are hereafter referred to as the field samples.

3.2. Satellite Comparison

Values from vegetation index products derived from both the Terra and Aqua satellites were extracted based on the 30 test sample points. Only values from the period 2002-2012 were sampled from the Terra images to match the length of the Aqua time series. Although the products from both satellites have a 16-day resolution, there was an 8-day offset in the observation dates between the two satellites, due to a difference in the timing of compositing periods. For each sample location, the two time series were correlated with one another and the timing of breaks detected by the BFAST algorithm was compared. All comparisons were performed at the pixel level.

A Jaccard similarity coefficient (Equation 1) was used to compare the timing of detected breaks.

$$J(A,B) = \frac{|A \cap B|}{|A \cup B|} \quad (1)$$

where A and B are the sets of breaks for the samples being compared. An index value of one means the two samples are identical, while a value of zero means there is no overlap.

The timing of a pair of breaks was considered to be the same if their 95% confidence intervals overlapped. The intersection for the samples being compared was calculated as the number of breaks with the same timing, and the union as the sum of all unique break dates across the samples. A Jaccard Index was calculated for each sample, and across all samples. The proportion of samples with a value of one was calculated, as was the agreement between samples based on the number of breaks detected regardless of their timing.

3.3. Assessing the Effects of the h Parameter and Vegetation Index

Terra satellite data were used to assess the effects of the h parameter and the vegetation index used, as Terra was able to provide the longer time series. The h parameter is expected to influence the number of breaks detected. Each test and field sample was decomposed using h parameter values of 1/3, 1/5 and 1/7. Using h = 1/3, the minimum length of trend components between detected breaks is four years.

A variation on the Jaccard similarity approach was used for comparing the h parameters, because agreement between samples is likely to be a function of the number of breaks detected alone. Using the P values from the linear modelling of the deseasonalised data, periods in the time series with significant trend component slopes were extracted for each sample. The intersection was defined as the sum of these periods that were common to all h parameters. The union was calculated as the sum of significant periods observed using any one of the h parameters tested (Figure 3). Significance levels of both 0.05 and 0.01 were used. A pairwise comparison between the three h parameters was also performed.

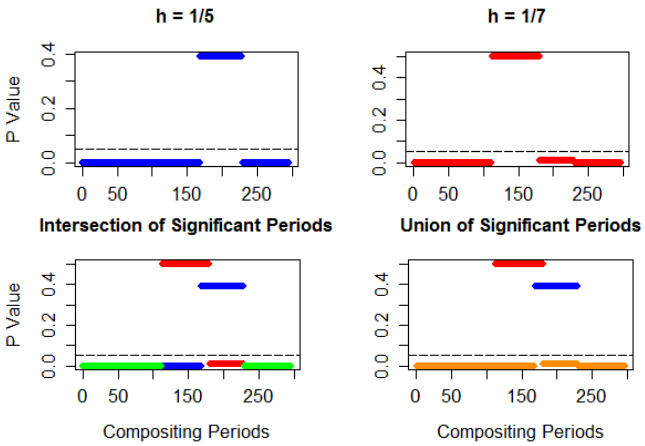

Figure 3. The intersection (green lines) and union (orange lines) of significant trend periods for Sample 1 of the 30 test sample points, using h = 1/5 and h = 1/7. The dashed line shows P = 0.05. The study period consists of 296 MODIS compositing periods.

Both EVI and NDVI time series were decomposed using BFAST to assess the satellite and h parameter effects described above. The effect of the vegetation index on the timing of detected breaks was also assessed directly by comparing break timing between decomposed EVI and NDVI time series for the test and field samples using the Jaccard Index approach.

4. RESULTS

4.1. The Effect of the Satellite

The time series derived from the Terra and Aqua satellites were well correlated with each other. The average Pearson's r value was higher using NDVI (0.90) than EVI (0.84). The two time series were also visually similar for most sample points (Figure 4). Measures of similarity in the timing of breaks showed greater than 70% agreement between the two satellites for NDVI. Jaccard Indices showed a 10% higher agreement for NDVI than for EVI, and a higher percentage of samples with identical break timing (Table 1).

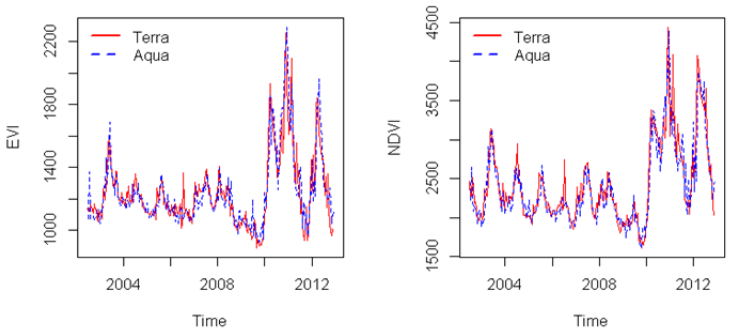

Figure 4. EVI and NDVI time series derived from both satellites for Sample 6 of the 30 test samples.

Table 1. Similarity metrics for the timing of breaks detected in time series from MODIS Terra compared with MODIS Aqua data using both EVI and NDVI (n = 30, h = 1/5).

	Jaccard Index	Samples with identical breaks (%)	Samples with equal no. of breaks (%)
EVI	0.67	50	70
NDVI	0.77	70	77

4.2. The Effects of the h parameter and Vegetation Index

As expected, decreasing the h parameter value allowed BFAST to detect a greater number of breaks (Figure 5). The absolute value of the largest step change in the trend component generally increased as the h parameter decreased (Figure 6). This would be expected, as a decrease in h parameter corresponds with a shorter run for trend components, causing an increased slope. The larger the gradient of the trend components, the larger the step changes between components are likely to be.

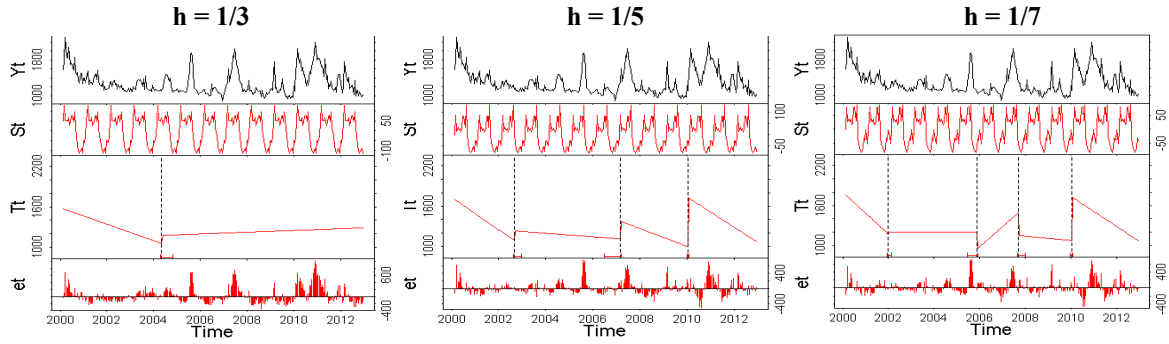

Figure 5: Decreasing the h parameter allowed more breaks to be detected in the same time series.

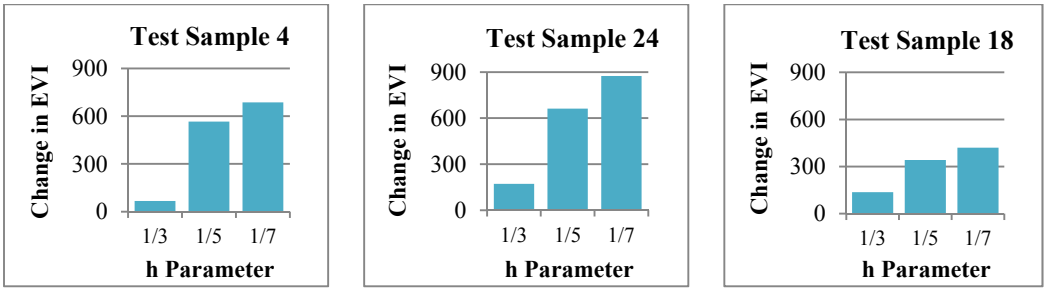

Figure 6: An inverse relationship was commonly found between the h parameter and the magnitude of the largest step change (break) in the trend component.

For the 165 sample points tested, agreement based on Jaccard Indices was higher for the pairwise comparison of the h parameters than across all three values. The h parameters 1/5 and 1/7 were the most similar based on common significant trend periods for both vegetation indices and significance (α) levels tested (Table 2). Similarity values were generally between 3% and 9% higher for NDVI than for EVI. This trend was reversed for the comparison between 1/5 and 1/7, however, with EVI showing slightly higher agreement values.

A trend of decreasing mean confidence interval width with decreasing h parameter can also be noted (Table 3). The timing of the earliest break detected was used for this comparison because at least one break was detected for most samples across all h parameters. Confidence intervals were narrower on average for NDVI than for EVI, but the difference between the two indices decreased as the h parameter decreased.

Agreement between the timing and number of breaks detected when comparing EVI with NDVI was also influenced by the h parameter (Table 4). Similarity in break timing between the two indices was highest using the intermediate h value of 1/5.

Table 2. Agreement between sample time series decomposed using different h parameters, based on overlapping significant trend periods (n = 165).

	α = 0.01		α = 0.05	
h Parameter	EVI	NDVI	EVI	NDVI
all	0.43	0.48	0.49	0.55
1/3 and 1/5	0.55	0.57	0.59	0.64
1/3 and 1/7	0.54	0.62	0.59	0.68
1/5 and 1/7	0.69	0.68	0.76	0.74

Table 3: Mean 95% confidence interval width (days) for the first break detected using each h parameter (n = 165)

h Parameter	EVI	NDVI
1/3	209.28	160.79
1/5	173.17	123.29
1/7	119.27	80.78

Table 4: Similarity metrics for the timing of breaks in the comparison between EVI and NDVI time series for each h parameter (n = 165).

h Parameter	Jaccard Index	Samples with identical breaks (%)	Samples with equal no. of breaks (%)
1/3	0.55	57.58	69.09
1/5	0.75	61.82	74.55
1/7	0.65	46.06	67.88

5. DISCUSSION AND CONCLUSIONS

The results of the satellite comparison suggest that, especially for EVI, the satellite used could have an impact on the timing of breaks detected by BFAST and to a lesser extent, the number of breaks. Despite the visual similarity between the two time series, there was a 23% difference in break timing between the two satellites using NDVI, which increases to 33% using EVI (see Table 1). For EVI, only half the samples tested showed complete agreement in terms of break timing.

Varying the h parameter changed the number of breaks detected as would be expected, however, increasing the number of breaks allowed did not cause the maximum possible number to be detected for all samples. There is some evidence to suggest an advantage of using h values of 1/5 or smaller. The narrower confidence intervals for breaks detected using lower values of h indicate breaks are detected with greater certainty. There is also a greater difference in the timing of significant trend periods between time series decomposed using h = 1/3 and both 1/5 and 1/7 than between 1/5 and 1/7 (Table 2). This suggests the smaller values of h are more likely to detect similar trends in a time series, and may indicate that 1/3 (corresponding to trend components of four years in length for this dataset) is too large a value to allow BFAST to detect all the variability present in vegetation index time series for the Paroo region. It is likely that the optimal h parameter will vary between datasets and locations.

The effects of both the satellite and the h parameter on the timing of breaks detected by BFAST were greater for EVI than for NDVI. The average confidence interval width for the first break was also smaller for NDVI, suggesting greater certainty in break timing using this index. Both the similarity in agreement metrics between h = 1/5 and 1/7 (Table 2) and the high agreement between EVI and NDVI observed using h = 1/5 (Table 4) suggest intermediate to small values of h can reduce the effect of the vegetation index on the timing of breaks detected by BFAST. The low agreement metrics shown in Table 4 also suggest a strong vegetation index effect. These differences between the two indices may be related to the inclusion of the blue band in the calculation of EVI. As the band covering the shortest wavelengths, more noise may be present than in the red and infrared bands. Apart from the narrower confidence intervals for NDVI, these results do not necessarily suggest an advantage of using one index over the other, but simply that break detection by BFAST is somewhat dependent on the index used.

The timing and number of breaks detected by BFAST are sensitive to the satellite used for data collection, the value of the h parameter, and, in particular, the vegetation index used. Further research is needed to better understand the effects of the h parameter on the detection of breaks, and whether the optimal parameter for use varies spatially or temporally.

REFERENCES

Australian Bureau of Meteorology (2013) Commonwealth of Australia, Canberra, accessed on 24 July 2012, < http://www.bom.gov.au>.

Bradley, B.A., Jacob, R.W., Hermance, J.F., & Mustard, J.F. (2007). A curve fitting procedure to derive inter-annual phenologies from time series of noisy satellite NDVI data. *Remote Sensing of Environment,* 106(2), 137-145.

Brock, M.A., Capon, S.J., & Porter, J.L. (2006). Disturbance of plant communities dependent on desert rivers. In R.T. Kingsford (Ed.), *Ecology of Desert Rivers*, pp. 100-132. Cambridge University Press, Cambridge, UK.

Carlson, T.N., & Ripley, D.A. (1997). On the relation between NDVI, fractional vegetation cover, and leaf area index. *Remote Sensing of Environment,* 62(3), 241-252.

de Jong, R., de Bruin, S., de Wit, A., Schaepman, M.E., & Dent, D.L. (2011). Analysis of monotonic greening and browning trends from global NDVI time-series. *Remote Sensing of Environment,* 115(2), 692-702.

Galford, G.L., Mustard, J.F., Melillo, J., Gendrin, A., Cerri, C.C., & Cerri, C.E.P. (2008). Wavelet analysis of MODIS time series to detect expansion and intensification of row-crop agriculture in Brazil. *Remote Sensing of Environment,* 112(2), 576-587.

Hill, M.J., & Donald, G.E. (2003). Estimating spatio-temporal patterns of agricultural productivity in fragmented landscapes using AVHRR NDVI time series. *Remote Sensing of Environment*, 84(3), 367-384.

Huete, A., Didan, K., Miura, T., Rodriguez, E.P., Gao, X., & Ferreira, L.G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment,* 83(1-2), 195-213.

Huete, A.R. (1988). A soil-adjusted vegetation index (SAVI). *Remote Sensing of Environment,* 25(3), 295-309.

Jacquin, A., Sheeren, D., & Lacombe, J.-P. (2010). Vegetation cover degradation assessment in Madagascar savanna based on trend analysis of MODIS NDVI time series. *International Journal of Applied Earth Observation and Geoinformation,* 12, Supplement 1, S3-S10.

Jönsson, P., & Eklundh, L. (2002). Seasonality extraction by function fitting to time-series of satellite sensor data. *Geoscience and Remote Sensing, IEEE Transactions on,* 40(8), 1824-1832.

Kingsford, R.T., & Porter, J. (1999). Wetlands and Waterbirds of the Paroo and Warrego Rivers. In R.T. Kingsford (Ed.), *A Free-flowing River: The Ecology of the Paroo River*, pp. 23-50. National Parks and Wildlife Service, Sydney.

Reed, B.C., Brown, J.F., VanderZee, D., Loveland, T.R., Merchant, J.W., & Ohlen, D.O. (1994). Measuring phenological variability from satellite imagery. *Journal of Vegetation Science,* 5(5), 703-714.

Sakamoto, T., Yokozawa, M., Toritani, H., Shibayama, M., Ishitsuka, N., & Ohno, H. (2005). A crop phenology detection method using time-series MODIS data. *Remote Sensing of Environment,* 96(3-4), 366-374.

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

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

Wang, D., Morton, D., Masek, J., Wu, A., Nagol, J., Xiong, X., Levy, R., Vermote, E., & Wolfe, R. (2012). Impact of sensor degradation on the MODIS NDVI time series. *Remote Sensing of Environment,* 119, 55-61.

Weiss, J.L., Gutzler, D.S., Coonrod, J.E.A., & Dahm, C.N. (2004). Long-term vegetation monitoring with NDVI in a diverse semi-arid setting, central New Mexico, USA. *Journal of Arid Environments,* 58(2), 249-272.

Young, W.J., & Kingsford, R.T. (2006). Flow variability in large unregulated dryland rivers In R.T. Kingsford (Ed.), *Ecology of Desert Rivers*, pp. 11-46. Cambridge University Press, Cambridge, UK.