Int J Appl Earth Obs Geoinformation 81 (2019) 27–36

Contents lists available at ScienceDirect

Int J Appl Earth Obs Geoinformation

journal homepage: www.elsevier.com/locate/jag

Optimal dates for assessing long-term changes in tree-cover in the semi-arid biomes of South Africa using MODIS NDVI time series (2001–2018)

Moses Azong Cho[a,b,*], Abel Ramoelo[c,d]

[a] Natural Resources and Environment Unit, The Council for Scientific and Industrial Research CSIR, PO Box 395, Pretoria, South Africa
[b] Department of Plant and Soil Science, University of Pretoria, South Africa
[c] University of Limpopo, South Africa
[d] Scientific Services, Conservation Science Department, South African National Parks, P.O. Box 647, Pretoria, South Africa

ARTICLE INFO

Keywords:
Tree-cover change
MODIS
NDVI time series

ABSTRACT

The varying proportions of tree and herbaceous cover in the grassland and savanna biomes of Southern Africa determine their capacity to provide ecosystem services. The asynchronous phenologies e.g. annual NDVI profiles of grasses and trees in these semi-arid landscapes provide an opportunity to estimate percentage tree-cover by determining the period of maximum contrast between grasses and trees. First, a 16-day NDVI time series was generated from MODIS NDVI data, i.e. MOD13A2 16-day NDVI composite data. Secondly, percentage tree-cover data for 100 sample polygons (4 × 4) pixels for areas that have not undergone change in tree cover between 2001 and 2018 were derived using high resolution Google Earth imagery. Next, a time series consisting of the coefficients of determination (R^2) for the NDVI/tree-cover linear regression were computed for the 100 polygons. Lastly, a threshold $R^2 > 0.5$ was used to determine the optimal period of the year for mapping tree-cover. It emerged that the narrow period from Julian day 161–177 (June 10–26) was the most consistent period with $R^2 > 0.5$ in the region. 18 tree-cover maps (2001–2018) were generated using linear regression model coefficients derived from Julian day 161 for each year. Kendall correlation coefficient (tau) was used to determine areas of significant ($p < 0.05$ and $p < 0.01$) increasing or decreasing trend in tree-cover. Areas (polygons) that showed increasing tree-cover appeared to be more widespread in the trend map as compared to areas of decreasing tree-cover. An accuracy assessment of the map of increasing tree-cover was conducted using Google Earth high resolution images. Out of 330 and 200 mapped polygons verified using $p < 0.05$ and 0.01 thresholds, respectively, 180 (54% accuracy) and 132 (65% accuracy) showed evidence of tree recruitment. Farm abandonment appeared to have been the most important factor contributing to increasing tree-cover in the region.

1. Introduction

Grassland and savanna biomes in South Africa cover about 60% of the land Wessels et al., 2011. The varying proportions of tree and herbaceous cover in these biomes determine their ability to provide ecosystem services including grazing and browsing resources, food, medicine and habitats to important wild fauna (Ramoelo et al., 2012; Scholes and Archer, 1997). The variability of grass and trees-cover is controlled by a number of environmental and anthropogenic factors. For example, increasing levels of anthropogenic CO_2 are expected to favour tree growth (Kgope et al., 2010; Buitenwerf et al., 2012). The consequence of this could be the expansion of woodlands into grass dominated systems, a phenomenon known as bush or woody encroachment. Human activities including deforestation, overgrazing, fencing-off of land for wildlife conservation, clearing of land for agriculture or farm abandonment are equally changing the dynamics of trees and grasses or forest cover in the region (Mograbi et al., 2017). Skowno et al. (2017) have reported a net increase in the extent of woodlands in South Africa between 1990 and 2013 at the expense of grasslands, most of this occurring in the savanna biome. The increasing tree cover in the South African savanna threatens the grazing potential of this system by suppressing the productivity of herbaceous plant species and reducing land accessibility by livestock and wildlife (Buitenwerf et al., 2012). Lesoli et al. (2013) have argued that performance measures, monitoring and adaptive management are necessary to control bush encroachment.

Remote sensing is now a popular means of establishing long-term spatio-temporal patterns of forest cover at the regional to continental

* Corresponding author at: Natural Resources and Environment Unit, The Council for Scientific and Industrial Research CSIR, PO Box 395, Pretoria, South Africa.
E-mail address: mcho@csir.co.za (M.A. Cho).

https://doi.org/10.1016/j.jag.2019.05.014
Received 12 February 2019; Received in revised form 13 May 2019; Accepted 13 May 2019
Available online 20 May 2019

scale (Defries et al., 2000; Shimada et al., 2014). However, there is no remote sensing system for periodic monitoring of bush encroachment in South Africa and in Southern Africa as a whole. To this point, the study by Skowno et al. (2017) is the most comprehensive mapping of bush encroachment in South Africa. The study was however limited in temporal resolution as it made use of land cover maps for only two years in a 23 years period; land-cover maps of 1990 and 2013. Moreover, the land-cover maps were obtained from another source; Geo Terra Image Pty Ltd. Skowno et al. (2017) considered woodlands as areas having a tree-cover percentage of greater than 20% because of poor classification of sparse woodlands in the land-cover maps. It should be noted the above threshold of woodlands is not consistent with the definition of forest by the Food and Agricultural Organisation's FAO, which defines forest as an area of tree-cover greater than 10% and of more than 0.5 ha.

Several other studies have reported poor accuracies of woodland classification in Southern Africa (Sedano et al., 2005; Gong et al., 2013; Fritz et al., 2010) in existing global tree-cover maps (Hansen et al., 2000; Loveland et al., 2000; Friedl et al., 2002; Tateishi et al., 2011). For example, Sedano et al. (2005) established errors of greater than 50% in the Miombo woodlands in Mozambique for the Moderate Resolution Imaging spectrometer MODIS - MOD12Q1 land- cover product and concluded that global land-cover products lack the detail needed for resource management at the national and provincial level. In the first ever 30 m resolution global land-cover map produced from Landsat TM data by Gong et al. (2013), the lowest classification accuracy for woodlands was observed in the Southern half of Africa. Sexton et al. (2013) observed a systematic bias in tree-cover prediction, characterised by over prediction in sparsely cover areas and under-prediction in dense tree-cover area after validation of the Vegetation Continuous Fields Tree-Cover data in Southern Africa. In part, the low classification accuracy of open woodlands in Southern African in the global forest maps could be attributed to classification errors resulting from high within-class variability (Landgrebe, 1997; Cho et al., 2010) because of the wide definition of forest (10–100% tree cover) according to the FAO's definition.. The problem of high within-forest class variation can be mitigated by initially considering several classes of woodlands from sparse to dense which could be later merged into one class as in Skowno et al. (2017) and as recommended by Cho et al. (2010). High classification errors for the open woodland class could also occur if the training data used for the classification is biased towards dense forest areas. Another drawback of determining forest or woody-cover change through the land-cover classification approach is that it is generally time consuming and therefore not practical for assessing contiguous trends e.g. yearly forest-cover change.

An alternative method to the 'land-cover classification' approach for assessing forest-cover change is the continuous mapping approach. Continuous mapping of percentage tree-cover usually involves the use of parametric or non-parametric regression to relate tree-cover percentage with predictor variables derived from remote sensing data (Kobayashi et al., 2016). The tree-cover map can then be categorised into the desired number of forest classes. Kobayashi et al. (2016) used a regression tree algorithm with MODIS data to produce a global tree-cover percentage map and reported improved accuracies when compared to two existing global tree-cover datasets. The most commonly used remote sensing predictor is the Normalised Difference Vegetation Index (NDVI) (Scanlon et al., 2002; Yang et al., 2012). There are however a number of challenges involved with the use of NDVI for tree-cover assessment; (i) NDVI varies across the growing season from the start to the end of the season, rendering the use of images for different periods impractical in tree-cover assessment; (ii) cloud cover presents a huge challenge, (iii) grass and soil background reflectance affects the reflected signal from trees and (iv) NDVI saturates at high canopy cover. The most viable alternative to improving tree cover assessment at the broad landscape is the use of Radio Detector and Ranging (RADAR) sensors as they make use of microwaves which penetrate clouds (Naidoo et al., 2016) and are sensitive to the structure of the vegetation. However, the lack of long-term time series of freely available RADAR data has limited their application in establishing contiguous long-term trends in tree-cover. Light Detection and Ranging LiDAR data provides the best opportunity for accurate mapping of tree-cover (Naidoo et al., 2016). However, the paucity of spaceborne LiDAR data has stymied its application for regional forest or tree-cover assessment.

Gong et al. (2013) suggested that improvement of land cover assessment from frequently available image would depend on the selection of data from suitable seasons and atmospheric condition, and on the development of new algorithms. In a study of land surface phenology (LSP) in the grassland and savanna biomes of South Africa, Cho et al. (2017) established differing patterns of LSP for grass and tree dominated areas. For example, tree-dominated areas showed a longer growing season and largely determine the period corresponding to the end of the growing season. The asynchronous phenologies of grasses and trees in the semi-arid landscapes of Southern Africa provide an opportunity to determine the optimal date or period for tree-cover assessment. Gaughan et al. (2013) recognised the importance of phenology in mapping tree-cover in Southern African savanna and derived a methodology that made use of NDVI green-up phenological metrics to estimate tree-cover. Cho et al. (2017) however found out that across a much broader savanna landscape, the end of growing season metrics showed a higher relationship with tree-cover than green-up metrics. Furthermore, Gaughan et al. (2013) argued that the tree-grass contrasts in Southern African savannas are highest in the dry season July-October when trees retain green biomass but grasses have senesced. The above assertion may not be true across the semi-arid savannas because many tree species in these systems are deciduous and are without leaves during the months of July-October. The current study is premised on the assumption that the optimal period for tree-cover estimation in the semi-arid savannas in Southern Africa could occur after the grasses have senesced and before the onset of leaf shedding by the trees. The exact period has not been determined. Therefore, the aim of this study was to determine the optimal period of the year for estimating tree-cover from MODIS NDVI time series and hence to establish a rapid procedure for assessing long-term changes in tree-cover in the grassland and savanna biomes of South Africa.

2. Material and methods

2.1. Study area

The study area consists of MODIS tile h20v11, clipped to South African boundaries (Fig.1). The region falls between latitude 22° - 29 °S and longitude 22° - 32 °E. Two main biomes dominate this MODIS tile, namely the savanna and grassland (Mucina, et al. 2006). The MOD13A2 16-day composite NDVI at 250 m resolution data from 1 Jan 2001 to 10 June 2018 were downloaded from the USGS explorer website.

2.2. Tree-cover percentage training data

The method to estimate tree-cover percentage was described in Cho et al. (2017). 150 random polygons 4 by 4 pixels or 1000 m by 1000 m were created on the MODIS image scene, converted to keyhole markup language (kml) file format and overlaid on Google Earth high resolution image in the Google Earth software (Google Inc). Only polygons that have not undergone changes in tree-cover from 2001 to recent times (generally 2016–2018) were retained for the study. This was achieved by examining historical images of each polygon. Plantation forest, crop fields and settlements were not included. Furthermore, only images acquired during the summer months (November to April) were used because tree canopy cover in the savanna is known to vary with season. The first 100 polygons that were assessed for tree-cover were used in the study. Three-band (red-green-blue) raster image scenes JPEG

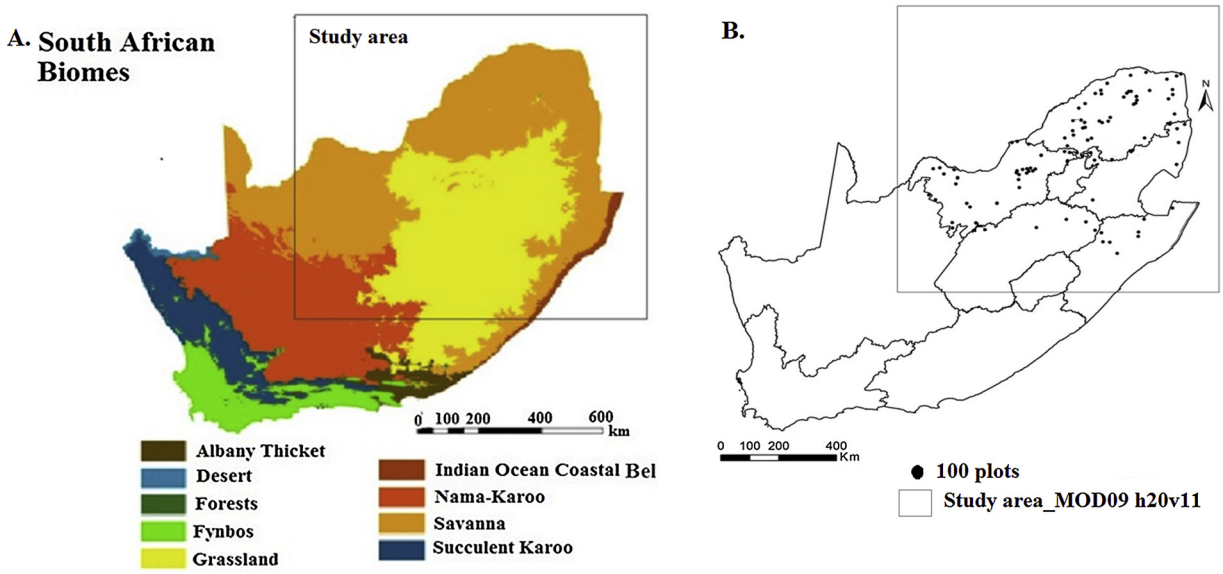

Fig. 1. Biomes of South Africa A Mucina and Rutherford (2006) and 100 random plots selected to analyse the relationship between NDVI and percentage tree cover B.

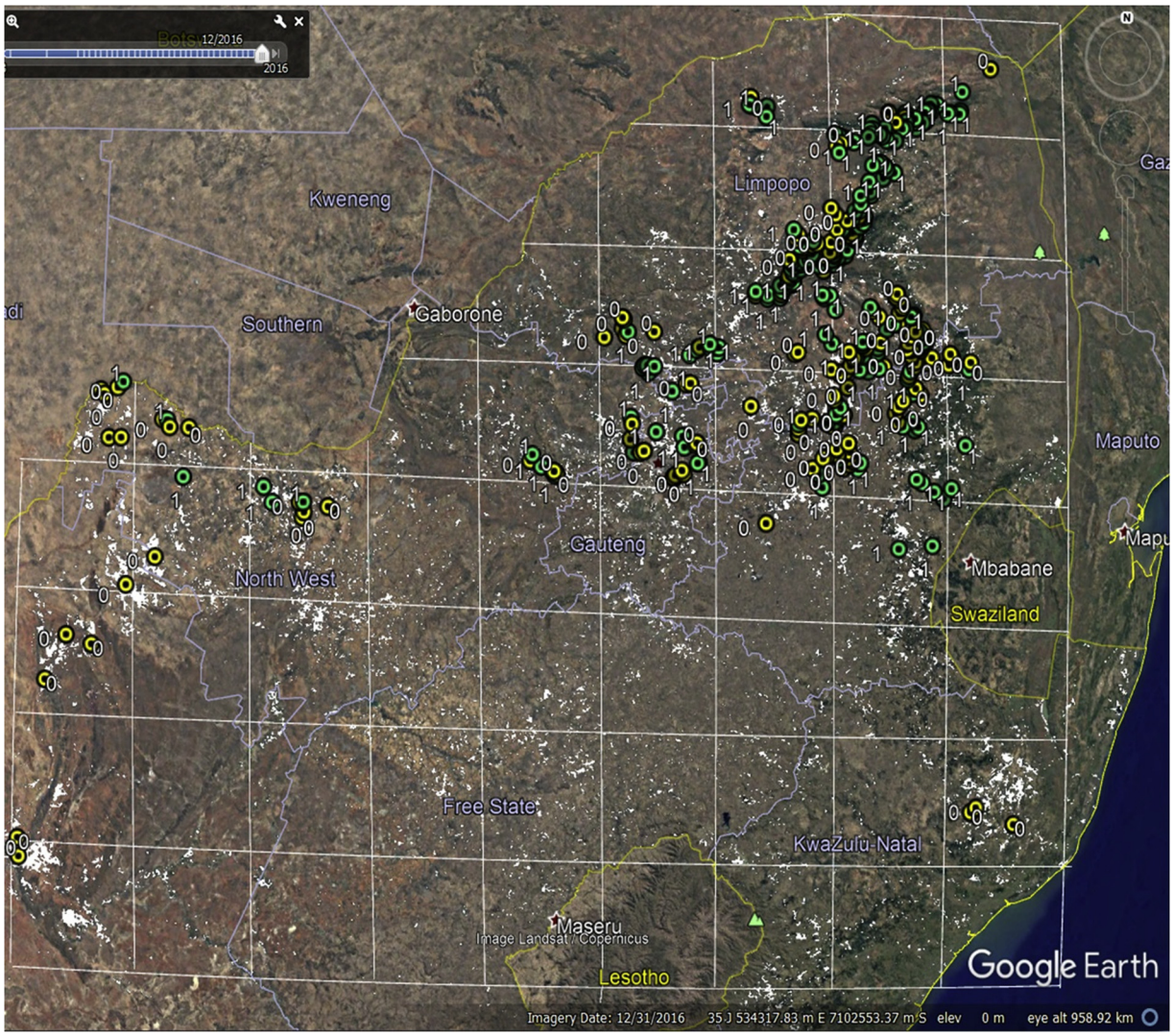

Fig. 2. Validation of increasing tree cover map using Google Earth imagery. 1green = evidence of increasing tree-cover and 0 yellow = no evidence of change in tree cover. White polygons denote areas with significant increasing trend in tree-cover ($p < 0.05$) increasing trend in tree-cover (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

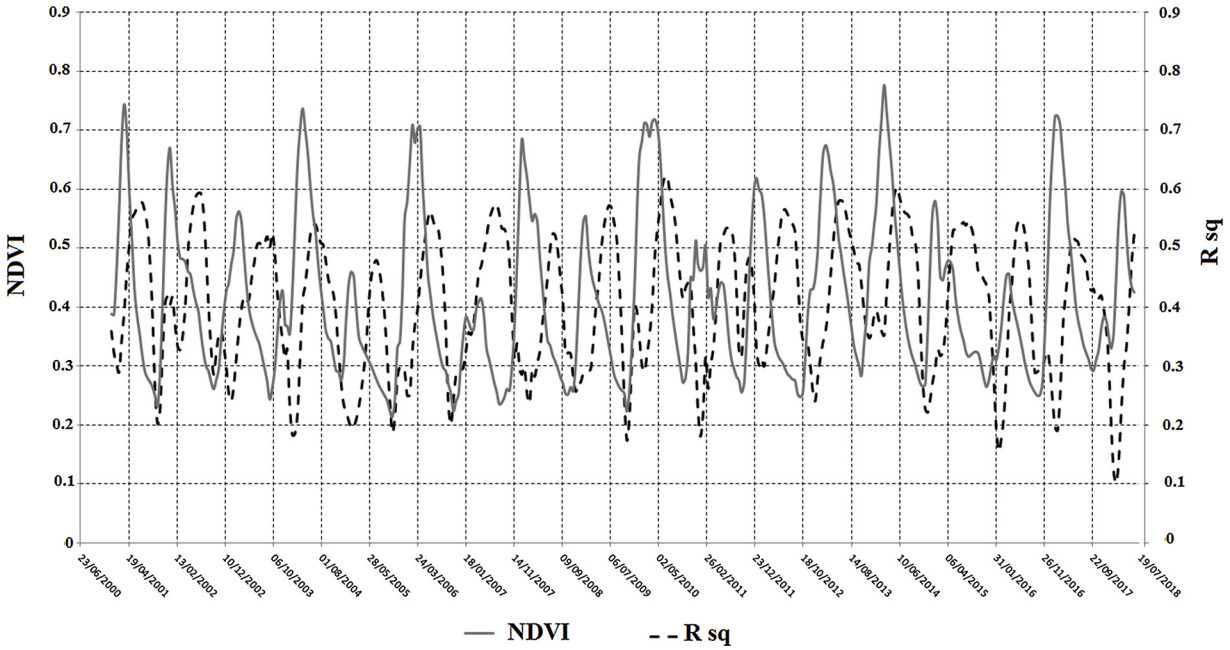

Fig. 3. Coefficient of determination R-sq between NDVI and tree cover overlaid on 16-day composite NDVI time series.

Table 1
Bootstrapped 30 bootstrap samples linear regression model coefficients and accuracies for the regression between NDVI and tree cover percentage for 100 plots.

Year	linear model: $y = a + bx$ coefficients		Calibration model standard error		Calibration model R^2		Model validation 30 bootstrap samples	
	a	b	Mean SE	± 95% CI	Mean R^2	± 95% CI	Mean SE	± 95% CI
2001	−39.7583	0.0176	16.44	0.15	0.58	0.01	17.23	0.44
2002	−34.3645	0.0192	16.43	0.21	0.58	0.01	16.08	0.63
2003	−30.1024	0.0179	18.09	0.24	0.49	0.01	18.55	0.65
2004	−28.6479	0.0148	17.20	0.20	0.53	0.01	17.52	0.54
2005	−30.2968	0.0177	18.60	0.27	0.45	0.02	19.07	0.76
2006	−37.0047	0.0173	16.89	0.21	0.55	0.01	16.62	0.61
2007	−31.2410	0.0183	16.85	0.14	0.55	0.01	17.51	0.40
2008	−29.9520	0.0166	18.80	0.20	0.45	0.01	18.90	0.57
2009	−36.5464	0.0177	16.63	0.15	0.56	0.01	17.42	0.45
2010	−42.2027	0.0172	15.70	0.20	0.61	0.01	15.51	0.57
2011	−35.4115	0.0170	17.28	0.25	0.51	0.01	18.68	0.66
2012	−39.0064	0.0209	16.92	0.17	0.55	0.01	16.82	0.49
2013	−29.2562	0.0159	16.33	0.16	0.58	0.01	16.66	0.44
2014	−37.3127	0.0178	16.39	0.21	0.56	0.01	16.28	0.60
2015	−33.2049	0.0182	17.17	0.20	0.54	0.01	17.53	0.61
2016	−35.7633	0.0174	17.25	0.25	0.53	0.01	17.35	0.71
2017	−31.7576	0.0158	17.54	0.25	0.50	0.01	18.05	0.71
2018	−31.3059	0.0144	17.08	0.25	0.53	0.02	18.44	0.72
Average	−34.0631	0.0173	17.09	0.21	0.54	0.012	17.45	0.59
± 95% CI	1.8551	0.0007	0.37	0.02	0.02	0.00	0.47	0.05

format of the selected polygons were saved in Google Earth and later imported into ENVI software (ITT Visual information solutions), where they were classified into two classes (tree and non-tree) using minimum distance supervised classification. The tree-cover percentage per polygon was computed as a ratio of the area of the tree-class to the total area of the polygon.

2.3. NDVI vs tree-cover relationship and determining the best dates for tree-cover assessment

The MOD13A2 NDVI time series data was smoothened using a Savitzky-Golay (Savitzky and Golay, 1964) smoothing filter (using a 3 window filter and a 5th order polynomial function). Subsequently, a time series of the coefficients of determination R^2 between NDVI and tree-cover percentage was established for each 16-day composite MODIS NDVI data. To determine the best period for tree-cover assessment, we first ranked the R^2 from the highest to the lowest value. Secondly, only dates corresponding to $R^2 > 0.5$ were considered for further analysis. The average, Median and mode dates with $R^2 > 0.5$ were then determined. The statistical analysis revealed Julian days 161 June 10 and 177 June 26 as the most frequent (the mode) date and median dates with high R^2 between the NDVI and tree-cover. A tree-cover prediction model for each year was established using a linear regression analysis between NDVI of Julian day 161 and tree-cover percentage. Bootstrapped regression analyses with 30 replicates were conducted to determine the linear regression model coefficients, calibration model R^2, and standard errors (SE) of calibration and prediction. That is, the bootstrapping process involved random sampling with replacement of 2/3 and 1/3 of the data for the calibration and validation models, respectively.

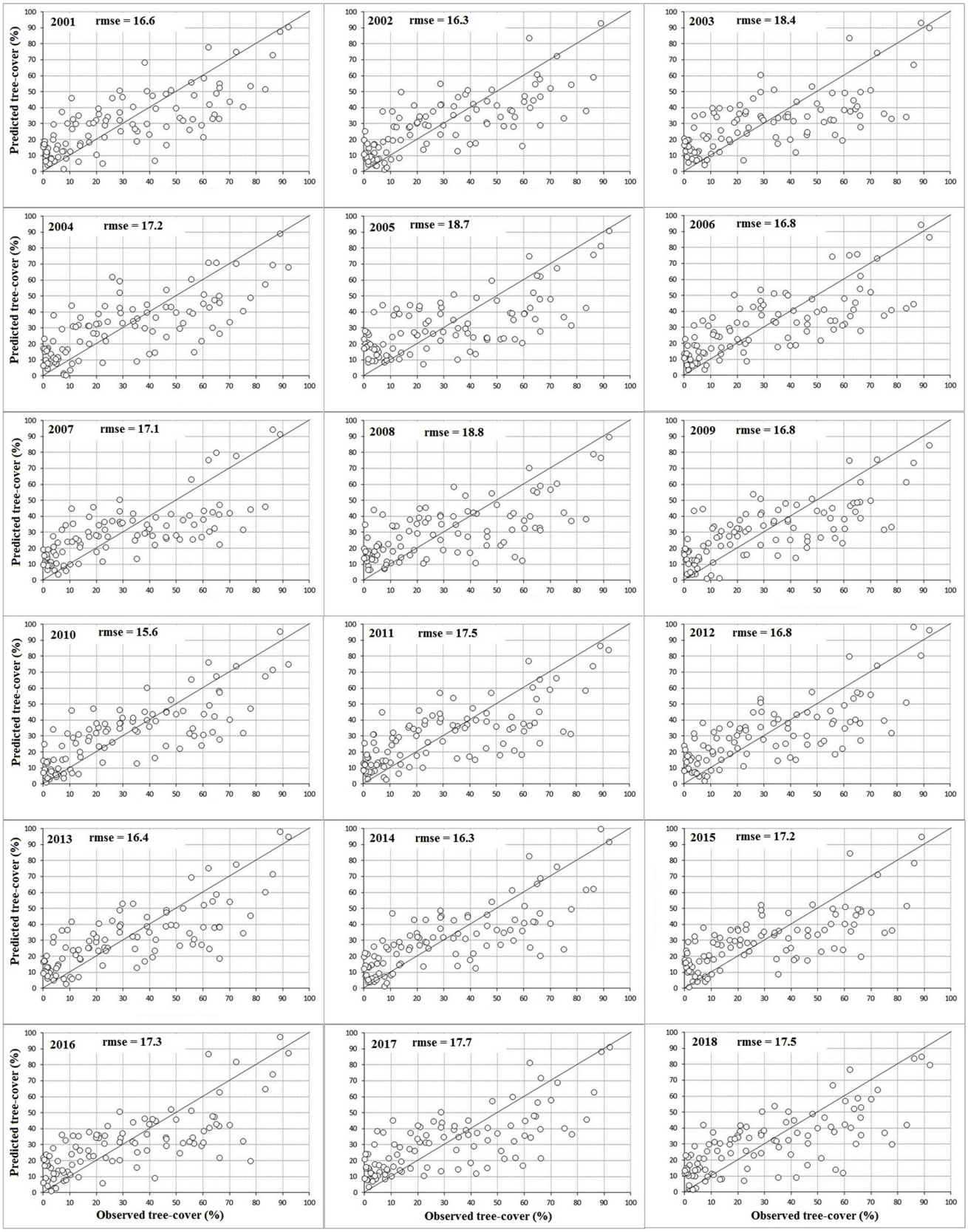

Fig. 4. Scatter plot between observed and predicted tree-cover for the various years.2001–2018.

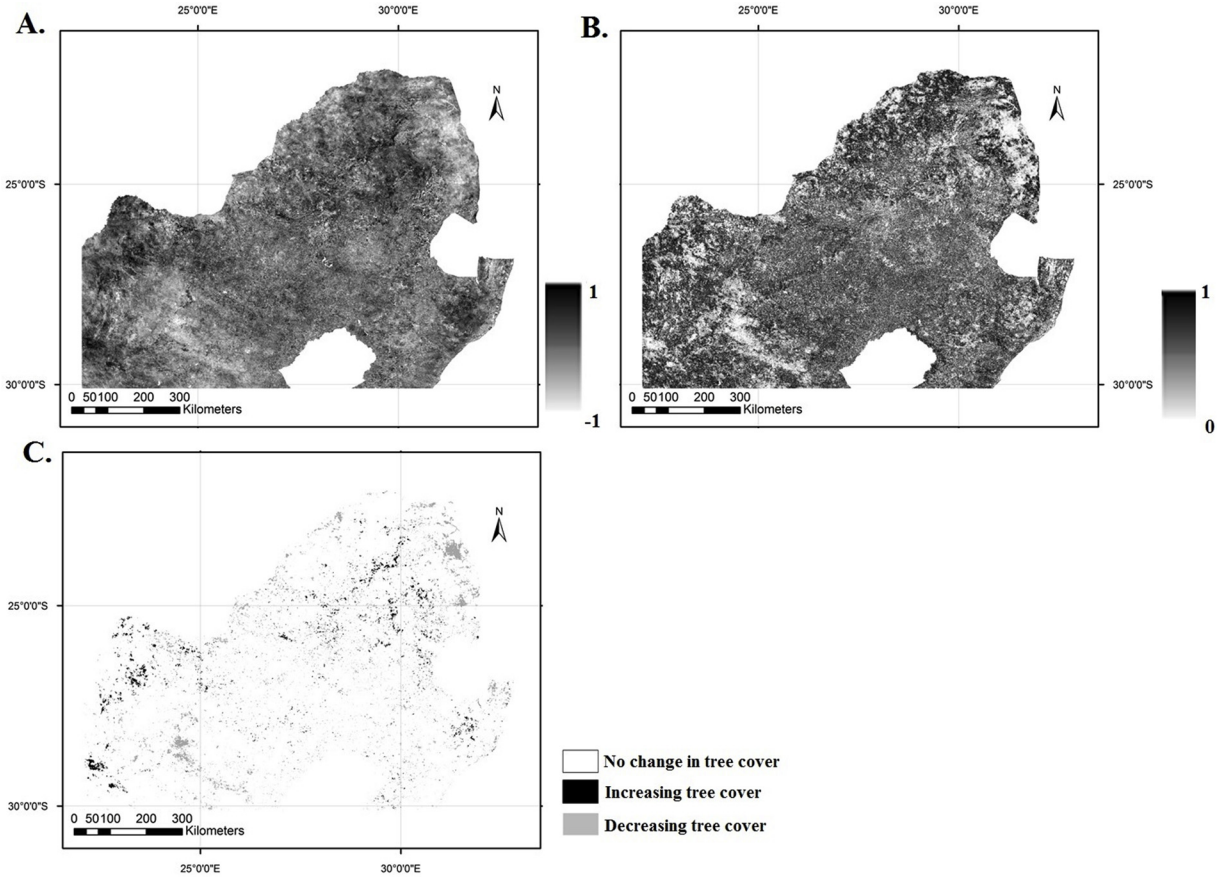

Fig. 5. Trend analysis of tree-cover change 2001–2018 using Kendall correlation analysis. A = Kendall Tau coefficient, B = *p*-values and C = threshold of $p < 0.05$ highlighting the areas of significant increasing or decreasing tree-cover percentage.

2.4. Trend analysis of tree-cover percentage 2001–2018

The non-parametric Kendall rank statistical test was used to assess the significance of any increasing or decreasing trend in tree cover-over percentage in the 18 year period for each pixel. The Kendall rank correlation coefficient or tau ranges from -1 and 1, with values greater and less than zero indicating increasing and decreasing trend, respectively. An IDL Interactive Data Language script was written to compute the tau and associated *p*-values for each pixel in the stack of 18 tree-cover maps. In order to delineate pixels that showed increasing or decreasing tree-cover, we explored two scenarios of *p*-values i.e. significance levels; '$p < 0.05$' and '$p < 0.01$'. The delineated pixels generally occurred in clusters. They were then vectorised as polygons in ENVI, saved as shapefiles and later converted into the kml file format.

2.5. Accuracy assessment of tree-cover trends

The accuracy of predicted trends in tree-cover was assessed using Google Earth high resolution images. A significant trend in increasing tree-cover may result from two kinds of processes, namely tree recruitment and expanding canopies of existing trees (i.e. increasing leaf area index). While a significant trend of decreasing tree-cover could occur from gradual thinning of the woodland from activities such as wood harvesting, tree felling by elephant or bush fires. We only thoroughly assessed the accuracy of polygons mapped as showing significant increasing trend in tree-cover because this was a more widespread process as compared to the decreasing tree-cover trend. A 100 km resolution grid was manually created on the Google Earth scene of the study area to guide the accuracy assessment process (Fig. 2). Carefully, moving from one grid cell to the other, we looked for evidence of increasing tree-cover amongst the polygons using the historical sliding bar of Google Earth. The accuracy assessment was limited to the identification of visible signs of tree recruitment in the mapped polygons. It was difficult to assess increasing tree-cover which may result from increasing tree LAI or densification of the tree canopy. Furthermore, we avoided very small polygons and where high resolution historical images were not available. The numbers one (1) and Zero (0) were respectively assigned to polygons that showed evidence of tree recruitment and no evidence of tree recruitment. The prediction accuracy was assessed as a ratio of the total number of ones to the total number of polygons assessed.

3. Results

3.1. Determining the optimal period of the year for tree-cover assessment

The NDVI time series and the resulting of NDVI/tree-cover regression R^2s showed asynchronous profiles or patterns (Fig. 3). The two profiles peak at different times of the year. The NDVI profile generally lags behind the R^2 profile by about three months. For example, the NDVI growing season curve (July to June) generally peaks in the month of February, while the R^2 curve peaks in early June. Statistical analysis of dates with the $R^2 > 0.5$ for each year revealed Julian day 161 or June 10 as the most frequent period i.e. the mode date with $R^2 > 0.5$ between NDVI and tree-cover percentage. Julian day 177 or June 26 was identified as the median date. We therefore concluded that the narrow period window from June 10–26 of only 16 days is the optimal period for assessing tree-cover percentage in the semi-arid grassland and savanna biomes of South Africa.

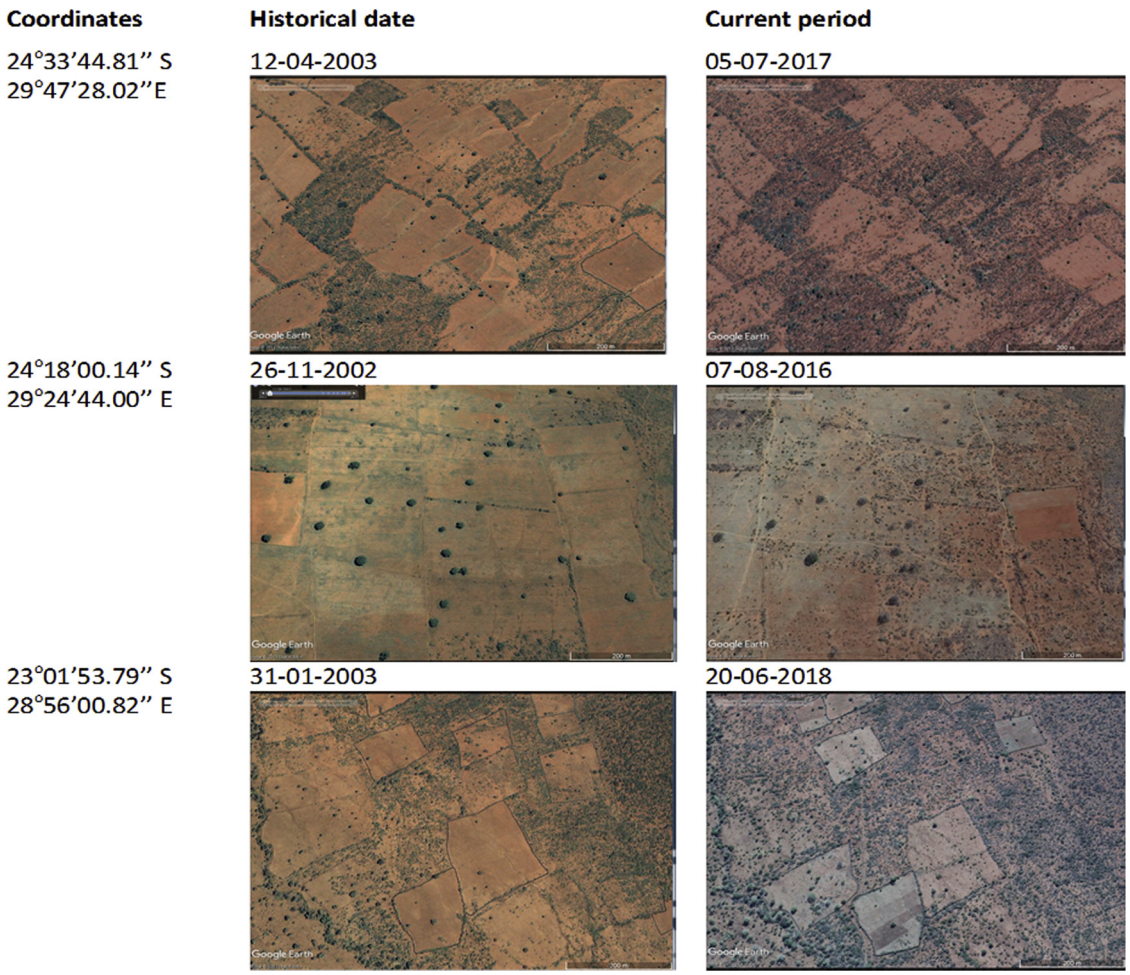

Fig. 6. Google Earth image scene showing increased in tree-cover in agricultural landscapes.

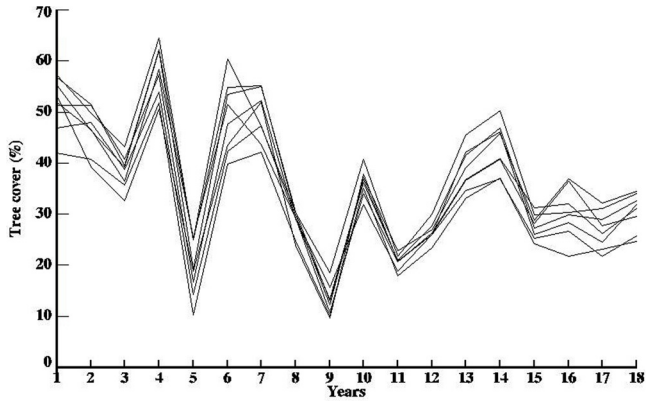

Fig. 7. Tree-cover profiles of an area in the Kruger National Park that showed significant decreasing trend in tree-cover between 2001 and 2018. The area was burnt in the winter of 2008.

3.2. Mapping tree-cover percentage for each year

Before we proceeded to map tree-cover for each year, we first explored whether a multivariate regression approach such as Random Forest (RF) involving all seven bands of MODIS visible to SWIR bands for Julian day 161 would improve on the NDVI/tree-cover linear regression results. The MOD15A images (500 m resolution) were used in this case. The accuracies of the RF regression models were rather lower than those of the linear regression involving NDVI. For example, the 2010 data yielded a calibration model standard error SE of 17.96 ± 0.92 (± 95% CI) as compared to SE = 17.02 ± 0.37 (± 95% CI) for the NDVI/tree-cover linear regression model, and a validation SE = 20.86 ± 1.25 (± 95% CI) compared to SE = 17.45 ± 0.59 (± 95% CI) for the NDVI model. The lower accuracy of the MODIS 500 m image might be due to the coarser resolution of the image. Only the full results of the NDVI/tree-cover linear regression models are represented in Table 1. We therefore used the bootstrapped linear regression model coefficients involving NDVI shown in Table 1 to produce tree-cover maps for all 18 years. The mean R^2 for the calibration models was 0.54 ± 0.02. As could be observed from the 18 scatter plots between the observed and predicted tree-cover data Fig. 4, there were systematic biases in the estimation of tree-cover, whereby low values were predicted high and high values predicted low. In general, tree-cover values below 10% were generally predicted with a positive bias by about 10 order of magnitude i.e. SE of +10%. This would imply that predicted tree-cover values of below 20% would generally be considered as areas of sparse woody cover i.e. below 10% of actual tree-cover or non-forest according to the FAO definition of forest.

3.3. Assessing tree-cover change 2001 – 2018

Using a '$p < 0.05$' significance level for the Kendall correlation (Fig. 5), we found evidence of tree recruitment in 180 polygons out of a total of 330 polygons assessed, corresponding to an accuracy of 54%. While for the 'p < 0.01' scenario, we found evidence of tree recruitment in 132 out of 200 polygons assessed accuracy = 65%. This

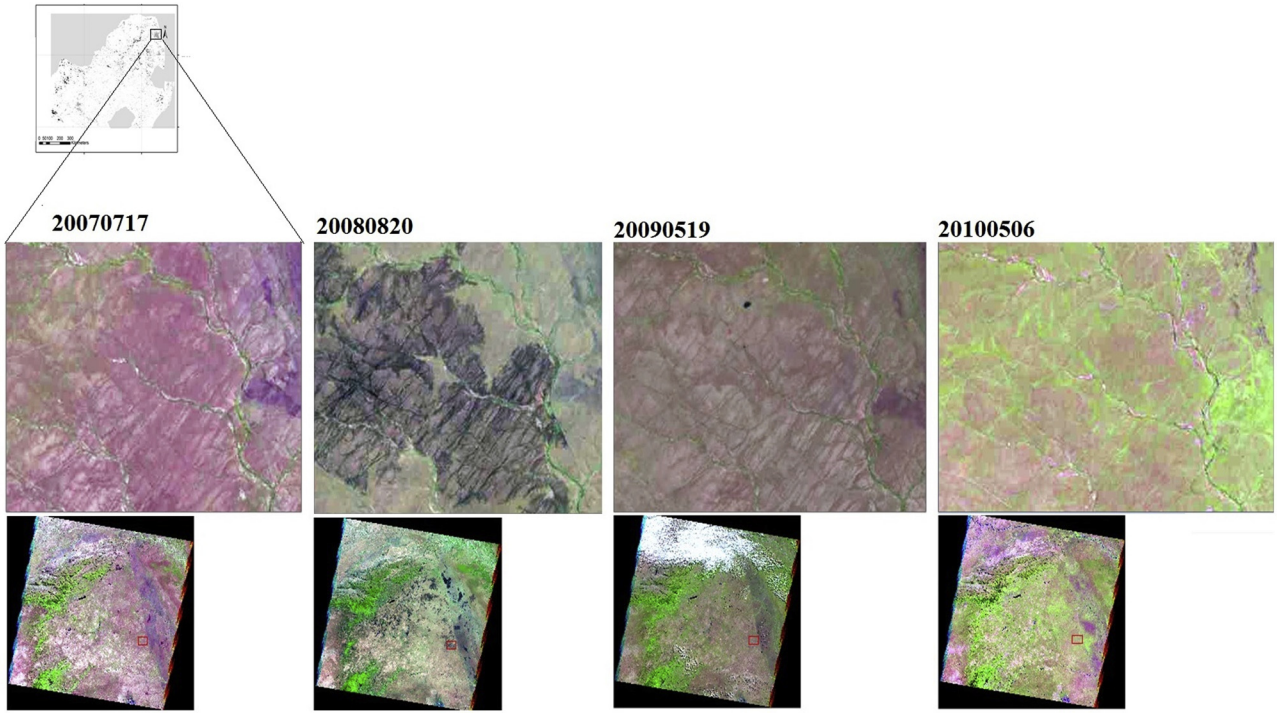

Fig. 8. Landsat TM5images 2007, 2008, 2009 and 2010 of an area in the Kruger National Park that showed significant decreasing trend in tree-cover between 2001 and 2018. A fire scar can be observed in 2008.

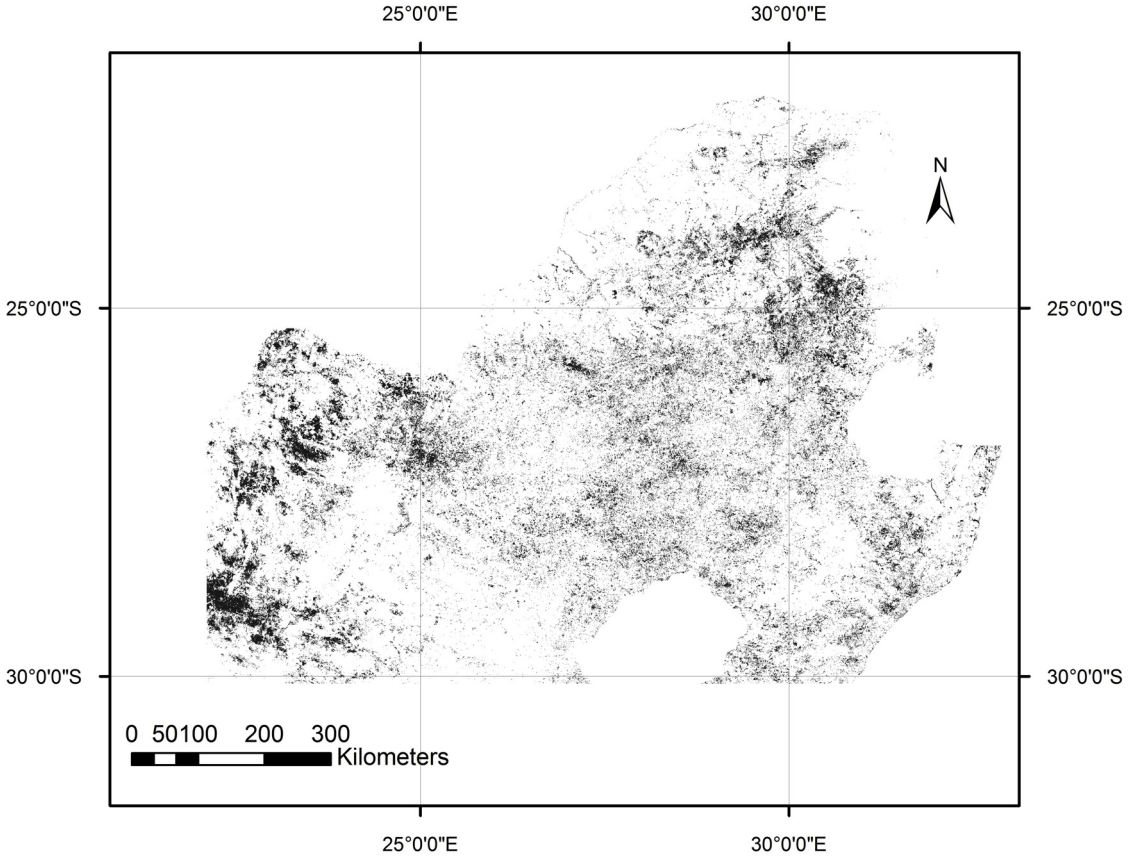

Fig. 9. Areas of significant $p < 0.05$ increasing or decreasing trend in NDVI derived from Kendall correlation analysis.

improved accuracy using the '$p < 0.01$' threshold was not achieved without some consequences, as 48 polygons that showed evidence of tree recruitment under the '$p < 0.05$' scenario were omitted. Most of the tree recruitment is occurring in the province of Limpopo and could be largely attributed to farm abandonment as observed from the Google Earth high resolution scenes e.g. Fig. 6

Significant decreasing tree-cover trends were observed in some large polygons in the Kruger National Park. The time series of the predicted tree-cover for this area Fig. 7 showed some strong irregular patterns though with an overall significant negative trend in tree cover. We further investigated the area using Landsat 5 T M images and observed that the sharp decrease in tree-cover observed in 2009 for example might have been caused by bush fire in the preceding year 2008 as a burnt scar can be observed in the 2008 image Fig. 8. A similar phenomenon was observed between 2004 and 2005.

It would be noted that it was worthwhile to convert the NDVI maps of Julian day 161 to tree-cover maps in order to assess increasing or decreasing trends in tree-cover for the 18 years period because similar analysis with NDVI yielded poor results. The NDVI trend analysis tends to show general trends in vegetation cover as significant trends could be observed in the grassland biome that are not related to increasing or decreasing tree-cover Fig. 9.

4. Discussion

Understanding the spatio-temporal dynamics of tree-cover in the savannas of Africa is critical for managing these ecosystems to sustain the provision of several key ecosystem services. The semi-arid savannas in Southern Africa are highly heterogeneous in space and time and challenging to map (Gaughan et al., 2013). In this study, we have explored the seasonal variability of vegetation green density using NDVI time series data with the aim of identifying the optimal period of the year for tree-cover estimation in the semi-arid savannas of Southern Africa. The assumption of the existence of an optimal period for tree cover assessment was rooted in the fact that several studies have shown that trees and grasses in Southern African savannas show asynchronous phenological profiles from the start to the end of the growing season (Archibald and Scholes, 2007; Chidumayo, 2001; Cho et al., 2017). For example, Cho et al. (2017) showed that the day corresponding to the end of the growing season in these systems explained 40% of the variance in tree-cover. It was therefore logical to assume that a period exists when the green contrast between the grasses and trees is highest and that such a period would be optimal for tree-cover assessment using NDVI, which is a measure of green density.

An optimal period exists for estimating tree-cover percentage in the semi-arid grassland and savanna biomes of South Africa from NDVI and the period is actually a narrow window from June 10–26 of just 16 days. This period is a month later than the period identified as the end of the growing season (May 9) for tree dominated areas in the same study area (Cho et al., 2017). June marks the beginning of leaf shedding by most deciduous and semi-deciduous trees and by July most deciduous trees are without leaves (Janecke and Smit, 2011). The proportion of deciduous trees in a pixel might have an influence on the model accuracy. Studies of tree-cover assessment in South Africa have hardly taken note of the narrow window of dates within which this important vegetation parameter can be accurately assessed using commonly available multispectral images. The good news is that the rarity of cloud cover in the month of June in the summer rainfall biomes of Southern African makes it largely possible to acquire cloud-free images during this narrow window of opportunity to assess tree-cover. However, irrigated fields and wetland areas could be wrongly predicted as having high tree-cover in June.

Using an optimal period to estimate tree-cover from NDVI did not mitigate some of the inherent shortcoming of the index in the Southern African savannas generally characterised by low tree cover. Only about have of the variance in tree-cover was explained by NDVI for all the years mean $R^2 = 0.54$. The linear regression models still performed poorly in estimating low tree-cover and saturated at high canopy cover. Even the use of a non-parametric multi-band regression algorithm, Random Forest in our case did not improve on the NDVI results. Naidoo et al. (2016) observed similar accuracies for estimating tree-cover in autumn in the South African Savanna using Landsat TM imagery e.g. R^2s of 0.34, 0.46, 0.50 and 0.65 for 23/05–2007, 07/04/2008, 12/05/2009 and 29/04/2009, respectively. Yang et al. (2012) obtained an $R^2 = 0.6$ for estimating tree cover in the semi-arid woodlands of Navada, USA using Landsat TM data. The availability of Sentinel-2, a new high resolution and freely available multispectral sensor with bands in the visible to short-wave infrared and more importantly in red-edge region offer promise for resolving the saturation problem of NDVI (Mutanga et al., 2004; Cho et al., 2009) and should be investigated. Furthermore, the availability of Sentinel-1 data will address the lack of time series of freely available RADAR data.

Google Earth high resolution images were very instrumental in this study, both for obtaining training data and in the validation of tree recruitment. There is however no complete wall-to-wall coverage of high resolution images for each year, which hindered the accuracy assessment of the tree-cover change maps. Appropriate high resolution images for every year could also be instrumental in assessing increase in tree-cover resulting from expansion of canopy cover or decreasing tree cover resulting from gradual thinning of the forest. We observed that tree recruitment is common in the province of Limpopo because of farm abandonment. Blair et al. (2018) have reported widespread cropland abandonment in the region between 1950 and 2010. They reported that farmers in the region attributed farm abandonment to rainfall variability and droughts. Increasing bush density in the savanna reduces land accessibility by wildlife and livestock and might also negatively affect the rangeland carrying capacity because of increased competition between trees and herbaceous vegetation for light, water and nutrients. On the other hand, increasing tree cover could increase the landscape's resilience to the impacts of droughts.

5. Conclusions

The most important conclusions from the study are:

- The narrow window between June 10 and 26 is the optimal period for assessing tree-cover percentage in the semi-arid grassland and savanna biomes of South Africa from optical remote sensing data.
- Farm abandonment appears to have been the most important factor contributing to increasing tree-cover change in the savanna landscape in South Africa.

This study shows that timing of image acquisition in the semi-arid biome of South Africa is critical for accurate assessment of tree-cover as only a narrow window of opportunity exist to achieve accurate assessment. Furthermore, the method proposed in this study can be appropriate for long-term assessment of tree-cover trends as opposed to short-term assessment given the modest accuracy recorded with the use of NDVI. The study needs to be extended to cover all of the Southern Africa in order to ascertain the optimal dates for tree-cover assessment in entire semi-arid biomes of Southern Africa.

Acknowledgments

Funding for this study was provided by the Council for Scientific and Industrial Research CSIR Parliamentary Grant and ECOPOTENTIAL project which received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No. 641762.

References

Archibald, S., Scholes, R.J., 2007. Leaf green-up in a semi-arid African savanna -separating tree and grass responses to environmental cues. J. Veg. Sci. 184, 583–594.

Blair, D., Shackleton, C., Mograbi, P., 2018. Cropland abandonment in south african smallholder communal lands: land cover change 1950–2010 and farmer perceptions of contributing factors. Land 74, 121.

Buitenwerf, R., Bond, W.J., Stevens, N., Trollope, W.S.W., 2012. Increased tree densities in South African savannas: &50 years of data suggests CO2 as a driver. Glob. Chang. Biol. 182, 675–684.

Chidumayo, E.N., 2001. Climate and phenology of savanna vegetation in Southern Africa. J. Veg. Sci. 123, 347–354.

Cho, M.A., Skidmore, A.K., Sobhan, I., 2009. Mapping beech *Fagus sylvatica* L. Forest structure with airborne hyperspectral imagery. Int. J. Appl. Earth Obs. Geoinf. 11, 201–211.

Cho, M.A., Debba, P., Mathieu, R., Naidoo, L., Van Aardt, J., Asner, G.P., 2010. Improving discrimination of savanna tree species through a multiple-endmember spectral angle mapper approach: canopy-level analysis. IEEE Trans. Geosci. Remote. Sens. 482, 4133–4142.

Cho, M.A., Ramoelo, A., Dziba, L., 2017. Response of land surface phenology to variation in tree cover during green-up and senescence periods in the Semi-Arid Savanna of Southern Africa. Remote Sens. 9, 689.

Defries, R.S., Hansen, M.C., Townshend, J.R.G., Janetos, A.C., Loveland, T.R., 2000. A new global 1-km dataset of percentage tree cover derived from remote sensing. Glob. Chang. Biol. 62, 247–254.

Friedl, M.A., McIver, D.K., Hodges, J.C.F., Zhang, X.Y., Muchoney, D., Strahler, A.H., Woodcock, C.E., Gopal, S., Schneider, A., Cooper, A., Baccini, A., Gao, F., Schaaf, C., 2002. Global land cover mapping from MODIS: algorithms and early results. Remote Sens. Environ. 831, 287–302.

Fritz, S., See, L., Rembold, F., 2010. Comparison of global and regional land cover maps with statistical information for the agricultural domain in Africa AU - Fritz, Steffen. Int. J. Remote Sensing 319, 2237–2256.

Gaughan, A.E., Holdo, R.M., Anderson, T.M., 2013. Using short-term MODIS time-series to quantify tree cover in a highly heterogeneous African savanna. Int. J. Remote Sens. 3419, 6865–6882.

Gong, P., Wang, J., Yu, L., Zhao, Y., Zhao, Y., Liang, L., Niu, Z., Huang, X., Fu, H., Liu, S., Li, C., Li, X., Fu, W., Liu, C., Xu, Y., Wang, X., Cheng, Q., Hu, L., Yao, W., Zhang, H., Zhu, P., Zhao, Z., Zhang, H., Zheng, Y., Ji, L., Zhang, Y., Chen, H., Yan, A., Guo, J., Yu, L., Wang, L., Liu, X., Shi, T., Zhu, M., Chen, Y., Yang, G., Tang, P., Xu, B., Giri, C., Clinton, N., Zhu, Z., Chen, J., Chen, J., 2013. Finer resolution observation and monitoring of global land cover: first mapping results with Landsat TM and ETM+ data. Int. J. Remote Sens. 347, 2607–2654.

Hansen, M.C., Defries, R.S., Townshend, J.R.G., Sohlberg, R., 2000. Global land cover classification at 1 km spatial resolution using a classification tree approach AU - Hansen, M. C. Int. J. Remote Sens. 216-7, 1331–1364.

Janecke, B.B., Smit, G.N., 2011. Phenology of woody plants in riverine thicket and its impact on browse availability to game species. Afr. J. Range Forage Sci. 283, 139–148.

Kgope, B.S., Bond, W.J., Midgley, G.F., 2010. Growth responses of African savanna trees implicate atmospheric [CO2] as a driver of past and current changes in savanna tree cover. Austral Ecol. 354, 451–463.

Kobayashi, T., Tsend-Ayush, J., Tateishi, R., 2016. A new global tree-cover percentage map using MODIS data. Int. J. Remote Sens. 374, 969–992.

Landgrebe, D., 1997. On Information Extraction Principles for Hyperspectral Data: a White Paper. School of Electrical and Computer Engineering, Purdue University, West Lafayette, pp. 1–25.

Lesoli, M.S., Gxasheka, M., Solomon, T.B., Moyo, B., 2013. Integrated Plant Invasion and Bush Encroachment Management on Southern African Rangelands. INTECH.

Loveland, T.R., Reed, B.C., Brown, J.F., Ohlen, D.O., Zhu, Z., Yang, L., Merchant, J.W., 2000. Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. Int. J. Remote Sens. 216-7, 1303–1330.

Mograbi, P.J., Asner, G.P., Witkowski, E.T.F., Erasmus, B.F.N., Wessels, K.J., Mathieu, R., Vaughn, R., 2017. Humans and elephants as treefall drivers in African savannas. Ecography 4011, 1274–1284.

Mucina, L., Rutherford, M.C., 2006. The Vegetation of South Africa, Lesotho and Swaziland. Pretoria: Strelitzia, 19, South African Biodiversity Institute.

Mutanga, O., Skidmore, A.K., Prins, H.H.T., 2004. Predicting in situ pasture quality in the Kruger National Park, South Africa, using continuum-removed absorption features. Remote Sens. Environ. 893, 393–408.

Naidoo, L., Mathieu, R., Main, R., Wessels, K., Asner, G.P., 2016. L-band Synthetic Aperture Radar imagery performs better than optical datasets at retrieving woody fractional cover in deciduous, dry savannahs. Int. J. Appl. Earth Obs. Geoinf. 52, 54–64.

Ramoelo, A., Skidmore, A.K., Cho, M.A., Schlerf, M., Mathieu, R., Heitkönig, I.M.A., 2012. Regional estimation of savanna grass nitrogen using the red-edge band of the spaceborne RapidEye sensor. Int. J. Appl. Earth Obs. Geoinf. 190, 151–162.

Savitzky, A., Golay, M.J.E., 1964. Smoothing and differentiation of data by simplified least squares procedures. Anal. Chem. 368, 1627–1639.

Scanlon, T.M., Albertson, J.D., Caylor, K.K., Williams, C.A., 2002. Determining land surface fractional cover from NDVI and rainfall time series for a savanna ecosystem. Remote Sens. Environ. 822, 376–388.

Scholes, R.J., Archer, S.R., 1997. Tree-grass interactions in savannas. Annu. Rev. Ecol. Syst. 281, 517–544.

Sedano, F., Gong, P., Ferrão, M., 2005. Land cover assessment with MODIS imagery in southern African Miombo ecosystems. Remote Sens. Environ. 984, 429–441.

Sexton, J.O., Song, X.-P., Feng, M., Noojipady, P., Anand, A., Huang, C., Kim, D.-H., Collins, K.M., Channan, S., DiMiceli, C., Townshend, J.R., 2013. Global, 30-m resolution continuous fields of tree cover: Landsat-based rescaling of MODIS vegetation continuous fields with lidar-based estimates of error AU - Sexton, Joseph O. Int. J. Digit. Earth 65, 427–448.

Shimada, M., Itoh, T., Motooka, T., Watanabe, M., Shiraishi, T., Thapa, R., Lucas, R., 2014. New global forest/non-forest maps from ALOS PALSAR data 2007–2010. Remote Sens. Environ. 155, 13–31.

Skowno, A.L., Thompson, M.W., Hiestermann, J.B.R., West, A.G., Bond, W.J., 2017. Woodland expansion in South African grassy biomes based on satellite observations 1990–2013: general patterns and potential drivers. Glob. Chang. Biol. 236, 2358–2369.

Tateishi, R., Uriyangqai, B., Al-Bilbisi, H., Ghar, M.A., Tsend-Ayush, J., Kobayashi, T., Kasimu, A., Hoan, N.T., Shalaby, A., Alsaaideh, B., Enkhzaya, T., Gegentana Sato, H.P., 2011. Production of global land cover data – GLCNMO AU - Tateishi, Ryutaro. Int. J. Digital Earth 41, 22–49.

Wessels, K.J., Mathieu, R., Erasmus, B.F.N., Asner, G.P., Smit, I.P.J., Van Aardt, J.A.N., Main, R., Fisher, J., Marais, W., Kennedy-Bowdoin, T., Knapp, D.E., Emerson, R., Jacobson, J., 2011. Impact of communal land use and conservation on woody vegetation structure in the Lowveld savannas of South Africa. For. Ecol. Manage. 2611, 19–29.

Yang, J., Weisberg, P.J., Bristow, N.A., 2012. Landsat remote sensing approaches for monitoring long-term tree cover dynamics in semi-arid woodlands: comparison of vegetation indices and spectral mixture analysis. Remote Sens. Environ. 119, 62–71.