L-band Synthetic Aperture Radar imagery performs better than optical datasets at retrieving woody fractional cover in deciduous, dry savannahs

Laven Naidoo[a,b], Renaud Mathieu[a,b], Russell Main[a,b], Konrad Wessels[b,c], Gregory P. Asner[d]

[a]Ecosystem Earth Observation, Natural Resources and the Environment, CSIR, Pretoria, South Africa

[b]Department of Geography, Geoinformatics and Meteorology, University of Pretoria, Pretoria, South Africa

[c]Remote Sensing Unit, Meraka Institute, CSIR, Pretoria, South Africa

[d]Department of Global Ecology, Carnegie Institution for Science, Stanford, CA, USA

Address all correspondence to: Laven Naidoo, Ecosystem Earth Observation, Natural Resources and the Environment, CSIR, Pretoria, South Africa; E-mail: LNaidoo@csir.co.za; Tel: (+27)12 841 2233

Abstract. Woody canopy cover (CC) is the simplest two dimensional metric for assessing the presence of the woody component in savannahs, but detailed validated maps are not currently available in southern African savannahs. A number of international EO programs (including in savannah landscapes) advocate and use optical LandSAT imagery for regional to country-wide mapping of woody canopy cover. However, previous research has shown that L-band Synthetic Aperture Radar (SAR) provides good performance at retrieving woody canopy cover in southern African savannahs. This study's objective was to evaluate, compare and use in combination L-band ALOS PALSAR and LandSAT-5 TM, in a Random Forest environment, to assess the benefits of using LandSAT compared to ALOS PALSAR. Additional objectives saw the testing of LandSAT-

5 image seasonality, spectral vegetation indices and image textures for improved CC modelling. Results showed that LandSAT-5 imagery acquired in the summer and autumn seasons yielded the highest single season modelling accuracies (R^2 between 0.47 and 0.65), depending on the year but the combination of multi-seasonal images yielded higher accuracies (R^2 between 0.57 and 0.72). The derivation of spectral vegetation indices and image textures and their combinations with optical reflectance bands provided minimal improvement with no optical-only result exceeding the winter SAR L-band backscatter alone results (R^2 of ~0.8). The integration of seasonally appropriate LandSAT-5 image reflectance and L-band HH and HV backscatter data does provide a significant improvement for CC modelling at the higher end of the model performance (R^2 between 0.83 and 0.88), but we conclude that L-band only based CC modelling be recommended for South African regions.

1   Introduction

Savannahs are mixed tree-grass communities which cover half of the African continent and one fifth of the global land surface[1]. The woody component has considerable impact on natural and anthropogenic processes, for instance on the fire regime, biomass production, nutrient cycling, soil erosion and the water cycle[2] while providing numerous ecosystem resources, such as fuelwood, building material and non-timber products[3]. At regional scale, the quantification of carbon captured in woody plants plays an important role in understanding the global carbon

cycle and fluxes between sinks and sources[4]. Monitoring regional woody resources is essential

to its sustainable management, which are threatened by deforestation or degradation (e.g.

fuelwood extraction), but also by bush encroachment [5,6].

The woody component can be represented by a variety of structural parameters such as

height, fractional cover, above ground biomass, basal area or canopy volume. Woody canopy

cover is the simplest two dimensional metric for quantifying the woody component. It can be

defined as the percent area vertically projected on a horizontal plane by woody plant

canopies[7], and is often referred to as fractional canopy cover (CC hereafter). In heterogeneous

Southern African savannahs, CC varies considerably across structural classes, e.g. from tall

closed forests to short closed, bush encroached shrubs to sparsely distributed tall trees with a

short shrub understory[8]. In South Africa and southern Africa there is no locally calibrated and

validated detailed (<1ha) maps of CC, despite it being recognised as an essential biodiversity

variable[9].

Remote sensing is the most appropriate tool for assessing woody structure across large

areas. This is due to its ability to sense the high spatio-temporal variability, species diversity

and phenological status, over large geographical scales – a defining but challenging set of

characteristics typical of African Savannahs[10]. Synthetic Aperture Radar (SAR) sensors are

particularly suited and extensively used for woody structural mapping, because of their

capacity to capture within-canopy properties[11,12-13], and lack of sensitivity to cloudy or hazy

conditions.   A combination of polarisation (e.g. Horizontal-Vertical or HV polarisation),

scattering interactions (e.g. volumetric scattering), wavelength and dielectric constant of the

vegetation permits such an interaction by SAR backscatter. In terms of wavelength, the longer

L-band (15 to 30cm) has been proven to be most effective both in forests and savannahs[14,15,16],

because L-band microwaves can penetrate deeper into the vegetation and interact with the main branches and trunks[16]. X-band was considered the least effective for predicting CC in savannahs[15] while C-band did yield positive results[15,17] though not as effective as L-band data.

Although not known to be adept in sensing three dimensional vegetation structure (e.g. aboveground biomass), multi-spectral optical sensors are well suited for mapping two dimensional structure such as canopy cover at various spatial scales in tropical forests, savannahs, and shrublands/grasslands[19,21,22]. While the SAR technology senses the canopy geometry, the mapping of canopy cover with optical sensors relies mostly on the contrast between the "greenness" of tree canopies and the background with the absorption peaks of the red and reflectance peaks of the green channels playing a big role in detecting this chlorophyll related "greenness". Thus, identifying the time period during the annual vegetation cycle at which a maximum contrast is achieved between green tree canopy and dry grass is important[23]. In addition, textural image products, which provide information regarding the local variance, can be used as a measure of the canopy roughness, gaps, and associated shadow. Optical sensor technologies - mostly LandSAT-based - have been adopted into successful national programmes for monitoring woody canopy cover changes. These include the Australian Statewide Landcover and Trees Study (SLATS)[21], the Australian National Carbon Accounting System – Land Cover Change Program (NCAS-LCCP)[20], the LandSAT Ecosystem Disturbance Adaptive Processing System (LEDAPS) for monitoring North American forest disturbance[25], and the Amazon Deforestation Monitoring Project (PRODES)[18]. Such monitoring programme is not yet in place for the savannahs of Southern Africa. Techniques, generally ranging from parametric methods (e.g. multi-linear regression[17,21]) to the more specialized such as non-parametric machine learning algorithms[15,21,47] (e.g. decision trees) and linear and non-

linear spectral unmixing[19,49], of coarser optical pixels, have been implemented for extracting fractional canopy cover at the regional scale.

Given the sensitivity of optical sensors to photosynthetically active vegetation and the sensitivity of SAR backscatter to vegetation structure, their possible integration may yield improved woody structure estimates via the provision of complementary information which neither sensor type could provide alone. The integration of SAR and optical technologies has been successfully applied in previous studies[26,27,47] in dense forested environments, savannahs and plantations. None of these studies has taken into account the effects of phenology on optical imagery, especially in savannah environments with complex tree and grass phenological seasonal changes. With this in mind and to support the development of a national system for monitoring CC in South African savannahs, the objective of this study was to evaluate how well CC could be retrieved, at the 30m spatial resolution, using multi-temporal datasets of SAR (L-band ALOS PALSAR) and optical (LandSAT-5 TM) sensor data, both independently and in combination. There were two main sets of research questions. The first set of questions focused on how the accuracy of CC predictions compared when using LandSAT versus L-band dual-polarised SAR input data, whether the integration of additional optical predictor features (e.g. textures and vegetation indices) or the integration of optical LandSAT and L-band SAR data produced any noticeable improvements in CC modelled predictions. The second research question sought to ascertain the season(s) in which LandSAT-5 data predicted CC with the highest accuracies. This question is related to the fact that savannah vegetation undergoes distinct seasonal phenological changes during which the green fractional cover of grasses and woody plants varies considerably[10]. We hypothesized that the season when trees are completely covered in green foliage, while grasses are dry (Autumn/Spring), should be the best

period to retrieve CC. The identification of phenologically optimised optical imagery may improve CC estimation, when integrated with SAR data. CC derived from very high resolution airborne LiDAR data were used as training and validation of the models.

2   Study Area

The region under study includes the southern portion of the Greater Kruger National Park Region, South Africa [23° 39'S to 25° 19'S, 30° 57'E to 32° 11'E]. This region consists of communal rangelands (Bushbuckridge Municipality District), private game reserves (Sabi Sands) and national or provincial parks (southern Kruger National Park, Andover) (Figure 1). The
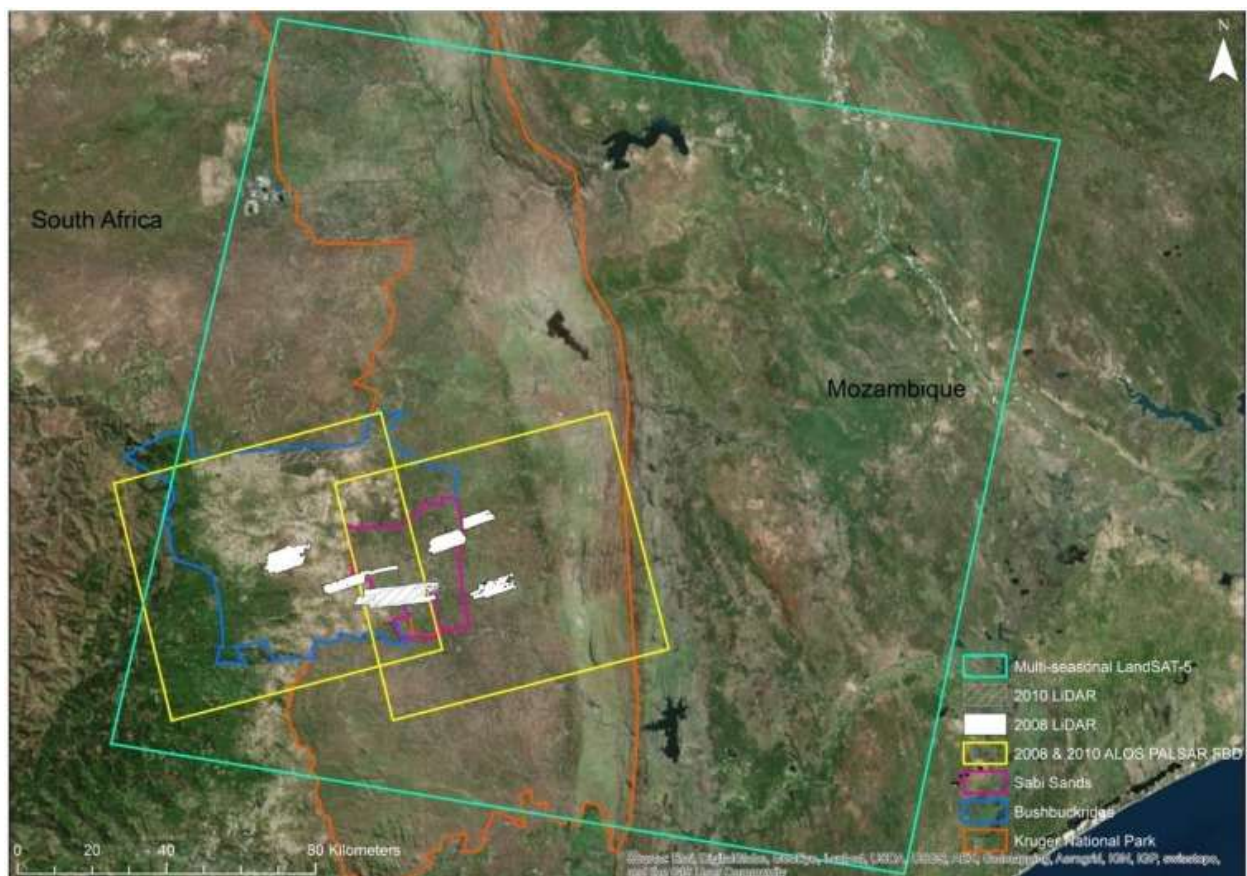
Fig. 1. The Southern Kruger National Park study area and coverage of remote sensing datasets.

region covers an extensive range of geologies (e.g. granite, basalt, gabbro), vegetation types (Clay Thorn Bushveld, Mixed Bushveld, Sweet Lowveld Bushveld and Open Grassland[29]), rainfall (mean annual precipitation of 1200mm in the west to 550mm in the east), management regimes (communal and protected) and disturbance regimes (fire, elephant damage, grazing and browsing patterns of herbivores, fuelwood harvesting). The general topography is flat to gently undulating with an average elevation of 400m[42]. The area undergoes short, dry winters and possesses a wet season (convective rainfall) with an annual rainfall between 235mm to 1000mm. Temperatures are mild (average of 22°C) with no occurrence of frost.

3   Materials and Methodology

3.1 Remote Sensing Data

HH/HV ALOS PALSAR L-band intensity scenes and multi-seasonal LandSAT-5 (bands 1-5, 7) scenes were collected. The L-band imagery (2 images for each year) was acquired in winter (25th of August and 23rd September (very early spring while landscape is dry and leaf-off) 2008; 14th and 31st August 2010) when the environment was dry and the trees devoid of leaves[17] (table 1). This will reduce possible backscatter error associated with soil moisture. LandSAT-5 scenes were inventoried (U.S Geology Survey LandSAT Earth Explorer portal) between 2007 and 2011 (to match the LiDAR dataset available in 2008 and 2010, with an acceptable difference of plus and minus one year) and acquired in various seasons to assess the potential effects of differential phenology between trees and grasses, specifically, spring (September-November), summer (December-March), autumn (April-May) and winter (June-August) (table 1). In summer, both tree leaves and grasses are green while in autumn, grasses are dry with trees remaining green but beginning to lose leaves. In winter, most trees have lost leaves and

grasses are dry while in spring, grasses are fairly dry while the trees first undergo a green flush of leaves[10]. Only LandSAT-5 imagery with an overall scene cloud cover of ≤6% was considered. Several years were considered to assess the possible model inconsistencies which may results from high inter-annual rainfall variability, and associated variability of greenness and phenology. Extensive airborne 2008 and 2010 LiDAR dataset (total coverage of c.a. 35000 ha and 10000 ha respectively) were acquired for this study (table 1; figure 1) by the Carnegie Airborne Observatory (CAO) Alpha sensor[28] during April-May of 2008 and 2010.

Table 1: LandSAT-5, ALOS PALSAR and LiDAR data inventory

Sensor	scene ID	Season	Date of Acquisition
LandSAT-5 TM	LT51680772007047JSA00	Summer	16/02/2007
LandSAT-5 TM	LT51680772007143JSA00	Autumn	23/05/2007
LandSAT-5 TM	LT51680772007175JSA00	Winter	24/06/2007
LandSAT-5 TM	LT51680772007223JSA00	Winter	11/08/2007
LandSAT-5 TM	LT51680772008034JSA01	Summer	03/02/2008
LandSAT-5 TM	LT51680772008098JSA01	Autumn	07/04/2008
LandSAT-5 TM	LT51680772008242JSA00	Winter	29/08/2008
LandSAT-5 TM	LT51680772008274JSA02	Spring	30/09/2008
LandSAT-5 TM	LT51680772009084JSA00	Summer	25/03/2009
LandSAT-5 TM	LT51680772009132JSA00	Autumn	12/05/2009
LandSAT-5 TM	LT51680772010023JSA00	Summer	23/01/2010
LandSAT-5 TM	LT51680772010119JSA00	Autumn	29/04/2010
ALOS PALSAR	ALPSRP137816680	Winter	25/08/2008
ALOS PALSAR	ALPSRP142046680	Spring	23/09/2008
ALOS PALSAR	ALPSRP242696680	Winter	14/08/2010
ALOS PALSAR	ALPSRP245176680	Winter	31/08/2010
CAO LiDAR	CAO 2008	Autumn	April-May 2008
CAO LiDAR	CAO 2010	Autumn	April-May 2010

3.2 LiDAR Data Processing

Digital Elevation Models (DEM) and top-of-canopy surface models (CSM) were created by processing the raw LiDAR point clouds using REALM (Optech Inc., Vaughn, Canada) and TerraScan/TerraMatch (Terrasolid Ltd., Jyvaskyla, Finland) software. Canopy height models

(CHM, pixel size of 1.12m) were computed by subtracting the DEM from the CSM. The LiDAR fractional woody cover metric were created by first applying a height threshold of 0.5m to generate a binary product of "no woody canopy" versus "woody canopy", which was then aggregated at 25m pixel size (CC in percent). The height LiDAR data were validated against field height measurements of approximately 800 trees (R^2 = 0.93, standard error 0.73m)[31], as well as the CC LiDAR data against field CC measurements for 37 sites (R^2=0.79; Root Mean Square Error=12.4%)[15].

3.3 SAR Data Processing

The level 1.1 PALSAR intensity datasets were processed in GAMMA[TM] SAR for multi-looking, radiometric calibration (from raw digital numbers to sigma nought backscatter), geocoding and topographic normalization. Multi-looking factors of 2 and 8 was applied to the range and azimuth, respectively, to best remove unwanted speckle. This was sufficient to have the majority of the speckle removed, while preserving image detail. A 20m DEM derived from 1:50 000 South African topographic maps was used for the geocoding and topographic normalization. As a final step the imagery was resampled, via bicubic-log spline interpolation function, by using a DEM oversampling factor of 1.6, to achieve a fixed spatial resolution of 12.5m to create images with a finer spatial detail.

3.4 LandSAT-5 Optical Data Processing and Derived Products

The LandSAT imagery underwent atmospheric correction with the use of ATCOR 2 which converted the raw digital number data to top of canopy (TOC, assuming flat terrain) reflectance using a Modtran-5 radiative transfer code. The default post May 2003 calibration file was used. Dry rural, fall (spring) rural, mid-latitude summer and winter rural atmospheric models were

utilised with the visibility distance set between 9.0km and 59km depending on the season and year (historical Skukuza visibility data obtained from http://weatherspark.com).

The TOC reflectance data were used as the main optical input variables to be tested. Additional vegetation indices and image textures were derived from the best performing LandSAT seasonal image for further analyses. This included a number of grey-level co-occurrence textural indicators (GLCM)[39], and spectral vegetation indices which are known to be sensitive to vegetation structure (table 2)[32-38] and, for some of these, to be less sensitive to environmental factors, such as soil background or atmospheric conditions, or saturation effects. GLCM texture parameters, such as variance and entropy, were selected as they were reported to be strongly correlated with vegetation structure[24,40] and in some case even better correlated than spectral indices[41]. Preliminary results (not shown) showed that variance,

Table 2: Reflectance, indices and textural optical products derived from LandSAT-5 data

Type	Product	Formulae or desciption if not applicable	Reference
Reflectance	Raw TOC reflectance	Band 1 (450-520nm) - Blue Band 2 (520-600nm) - Green Band 3 (630-690nm) - Red Band 4 (760-900nm) - NIR Band 5 (1550-1750nm) - MIR Band 7 (2080-2350nm) - MIR	
Vegetation Index	Enhanced Vegetation Index (EVI)	$2.5 \times \dfrac{(NIR - Red)}{(NIR + (6 \, X \, Red) - (7.5 \, X \, Blue) + 1)}$	Huete et al. (1997)
Vegetation Index	Modified Simple Ratio (MSR)	$\dfrac{\left(\frac{NIR}{Red}\right) - 1}{\sqrt{\frac{NIR}{Red}} + 1}$	Sims & Gamon (2002)
Vegetation Index	Non-linear Vegetation Index (NLI)	$\dfrac{NIR^2 - Red}{NIR^2 + Red}$	Goel & Qin (1994)
Vegetation Index	Soil-Adjusted Vegetation Index (SAVI)	$\dfrac{NIR - Red}{NIR + Red + 0.5} \times (1 + 0.5)$	Huete (1988)
Vegetation Index	Simple Ration (SR)	$\dfrac{NIR}{Red}$	Jordan (1969)
Vegetation Index	Normalised Difference Vegetation Index (NDVI)	$\dfrac{NIR - Red}{NIR + Red}$	Rouse et al. (1974)
Vegetation Index	Moisture Vegetation Index (MVI band 7)	$\dfrac{NIR - MIR7}{NIR + MIR7}$	Sousa & Ponzoni (1998)
GLCM Textures	Variance, Entropy, Dissimilarity & Contrast (3 X 3 window)	Applied to bands 1-7	Haralick et al. (1973)

TOC= Top of Canopy; NIR = Near Infrared; MIR = Middle Infrared

entropy, dissimilarity and contrast textures, derived from the bands 1 to 5 and 7, were more correlated with CC.

3.5 Modelling Algorithms, Modelling Scenarios, Model Validation and CC Mapping

A fixed grid of 105m X 105m cells (best trade-off between mapping details and model performance[17,42]), with a 50m spacing to avoid spatial autocorrelation of CC[31], was used to extract mean SAR, optical and LiDAR CC values. Cells occupying water bodies, roads, rivers, informal settlements and clouds (in the LandSAT imagery) were excluded from the analysis. A random forest machine learning algorithm[30] was applied in the R rattle software with 35% of the data being used for model training and 65% for model validation. Due to its use of multiple decision trees, bagging and internal cross-validation mechanisms, RF is considered as a major improvement over other traditional decision tree types. The algorithm is easy to implement and is robust as it only requires two main user-defined inputs (number of trees built in the 'forest' and the number of possible splitting variables for each node[43]).

Before its final implementation, efforts were made to test the RF generalisation by introducing an additional independent test dataset for model tuning before validation. During the tuning phase, the total number of trees ('ntree') in the forest and the RF tree complexity were varied to test their influence on accuracy whilst trying to limit the RF complexity. RF tree complexity included the minimum number of terminal nodes ('nodesize') and the maximum number of terminal nodes that the trees can have in the forest ('maxnodes')[30]. After repeating the process three times, results showed that an 'unpruned' (i.e. no limitation on a tree's depth and number of terminal nodes) tree architecture with 200 trees within the forest, yielded optimum results. Hence, the RF models were created with 'ntrees' = 200 and 'mtry' = square

root of # model predictors (a rule of thumb for 'mtry'[44]) with the trees being allowed to grow unpruned.

For the modelling process, several scenarios were assessed. The LandSAT images were first tested individually in order to ascertain the best season for predicting CC. Multi-seasonal LandSAT scenarios were also tested by combining all available seasonal images for each year. Seven additional scenarios using combined reflectance, texture and vegetation indices were assessed, but this was only performed for the best performing optical scenario mentioned above. Due to the large number of vegetation indices and textures used in this study, which may display high degrees of co-linearity, a RF variable importance measure called the permutation accuracy or %IncMSE (percentage increase in mean squared error) was considered to select only the top three indices and texture variables for inclusion in the modelling scenarios. %IncMSE records the percentage increase in the mean squared errors in the model when a particular variable is assigned random values while the remaining variables are left unchanged[44]. The higher the resultant error, the more important that particular variable is to the model. 2008 and 2010 L-band SAR dataset-only scenarios served as the scenario of comparison for the optical-only tests.

The SAR datasets were then integrated with the five best performing seasonal LandSAT-5 images and the combined multi-seasonal LandSAT-5 datasets for each year to quantify the improvement gained from combining SAR and optical data. The scenarios were assessed and compared using the coefficient of determination (R^2), Root Mean Square Error (RMSE) and Standard Error of Prediction (SEP). SEP refers to the standard deviation of the prediction errors and is a measure of the unexplained variation of a model. XY scatterplots derived from observed CC (LiDAR-derived) versus predicted CC (RF model-derived) were included to highlight

the main accuracy differences between, optical-only, SAR-only and integrated SAR and optical scenarios. Finally, for the expecting discussion on vegetation phenology, monthly EVI values was derived from 8-day composite 500m MODIS MCD43 BRDF-corrected surface reflectance data[46], between 2005 to end of 2012, which was selected over tree dominated and grass dominated environments (with the help from LiDAR). A multi-temporal EVI difference product between these grass and tree dominated environments was also derived for discussion.

4   Results

4.1 Individual and multi-seasonal LandSAT-5 reflectance compared to SAR

When examining the individual seasonal LandSAT-5 reflectance accuracies (table 3), the season

Table 3: Individual seasonal, multi-seasonal LandSAT-5 and individual SAR RF modelled CC validation results

Dataset	Acquisition Date	Season of Imagery	R²	RMSE (%)	SEP (%)	Total No. Obs*
Individual LandSAT-5 TM	16/02/2007[1]	Summer	0.47	12.64	52.02	8804
	23/05/2007[1]	Autumn	0.34	13.96	58.46	8804
	24/06/2007[1]	Winter	0.32	14.25	58.76	8804
	11/08/2007[1]	Winter	0.32	14.10	58.69	8733
	03/02/2008[1]	Summer	0.53	11.84	49.24	8804
	07/04/2008[1]	Autumn	0.46	12.89	52.64	8010
	29/08/2008[1]	Winter	0.37	13.60	56.73	8804
	30/09/2008[1]	Spring	0.40	13.19	53.2	8339
	25/03/2009[1]	Summer	0.44	12.76	52.86	8804
	12/05/2009[1]	Autumn	0.50	12.04	49.6	8697
	23/01/2010[2]	Summer	0.64	14.77	46	2098
	29/04/2010[2]	**Autumn**	**0.65**	**13.55**	**44.43**	**3201**
Multi-seasonal LandSAT-5 TM	2007[1]	All available images	0.58	11.27	47.23	8733
	2008[1]	All available images	0.64	10.53	43.31	8010
	2009[1]	All available images	0.57	11.36	46.92	8697
	2010[2]	**All available images**	**0.72**	**12.84**	**39.75**	**2098**
SAR	**25/08/2008[1]**	**Winter**	**0.80**	**7.88**	**32.08**	**8804**
	14/08/2010[2]	**Winter**	**0.81**	**10.17**	**33.16**	**3201**

* Variable depending on LiDAR coverage per year and cloud cover; [1] 2008 LiDAR dataset for reference dataset; [2] 2010 LiDAR dataset for reference dataset

which yielded the highest model accuracies varied between years; summer was best in 2007 and 2008, and autumn in 2009 and 2010. Amongst all the individual datasets, the April 2010 LandSAT-5 reflectance (autumn) dataset yielded the highest model accuracies. The winter datasets (2007, 2008) yielded the poorest modelled CC results. Overall the performance of single LandSAT datasets was poor with a SEP varying between 44 and 58%. Combining all the multi-seasonal images for each year improved the accuracies by an RMSE of ~1-2% and SEP of ~4-6% compared to the best individual seasonal image for that year. However, both individual seasonal and combined multi-seasonal image yielded significantly lower accuracies than those of the individual SAR images. For instance, the 2008 and 2009 SAR models had a SEP of 15 and 10% lower, compared to the best LandSAT season of that specific year. Moreover, both SAR models produced consistent results, with a similar R^2 and SEP.

Table 4: Reflectance, indices and textural LandSAT-5 (Autumn 2010 image) product RF modelled CC validation results

2010 Optical Product(s)	R^2	RMSE (%)	SEP (%)	Total No. Obs
Reflectance only	0.65	13.55	44.43	3201
Textures only*	0.03	23.66	77.96	3201
Indices only*	0.45	17.22	57.16	3201
Reflectance +Textures*	0.67	13.30	43.74	3201
Reflectance + Indices*	0.66	13.52	44.93	3201
Indices* + Textures*	0.47	17.06	55.87	3201
Reflectance + Textures* + Indices*	0.68	12.98	43.53	3201
2010 SAR only[1]	**0.81**	**10.17**	**33.16**	**3201**

* Top 3 vegetation indices/texture metrics used based on %IncMSE

4.2 Optical reflectance, textures and indices compared to and integrated with SAR data results

The top three texture metrics and spectral vegetation indices were added as additional features to the best performing LandSAT-5 reflectance dataset (April 2010) (table 4). The optical reflectance-only scenario yielded the best results, followed by the derived vegetation

indices, and the textures-only produced by far the poorest results. However, the combination of reflectance and textures yielded better results (albeit small) than the reflectance and indices combination which suggested that image textures do provide more additional information in comparison to the indices. Combining all three datasets (reflectance, textures and indices) provided the highest overall accuracy, however improvement was marginal compared to the optical reflectance-only scenario. These results were consistent for other years (2007, 2008 and 2009, not shown). Combining the best seasonal LandSAT-5 reflectance dataset per year (2008 and 2010) with SAR data brought about modest, but significant improvements (improved SEP of ~4-5%) in comparison to SAR-only scenarios (table 5). The addition of three most important vegetation indices and textures to the previously mentioned scenario did not improve the results. Also, the difference in accuracy between the best seasonal reflectance and combined multi-seasonal images, integrated with SAR datasets, were minimal (improved SEP of

Table 5: Integrated SAR and best performing LandSAT-5 reflectance RF modelled CC validation results
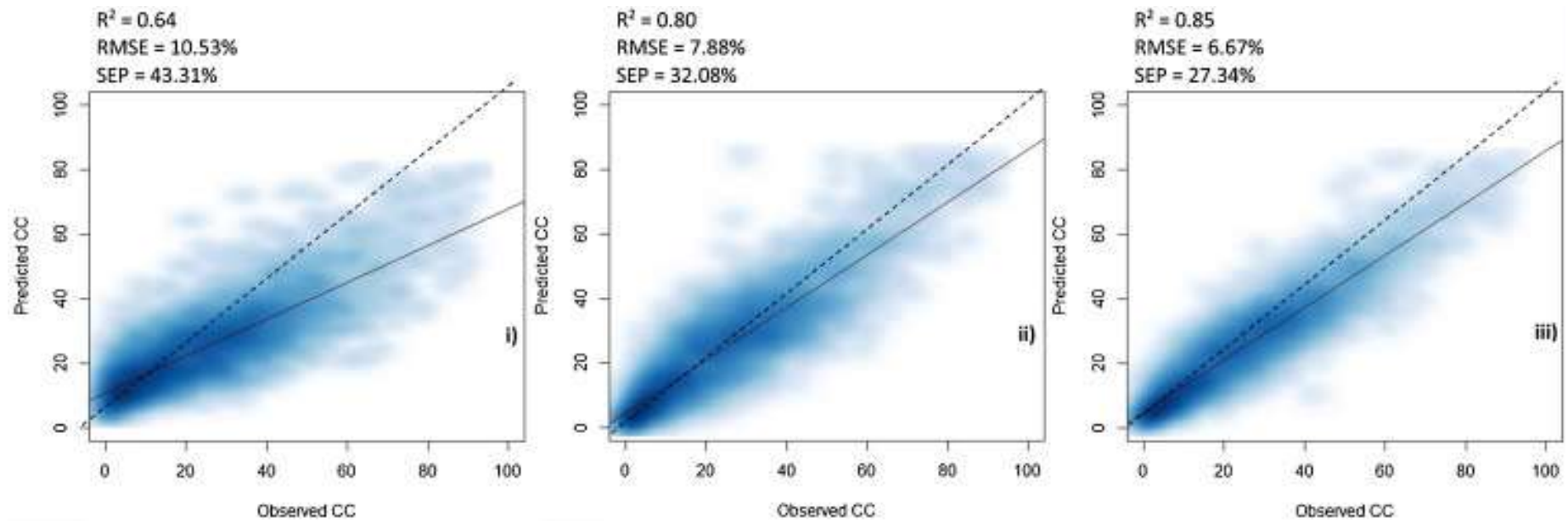
Dataset	Acquisition Year	Season of Imagery	R^2	RMSE (%)	SEP (%)	Total No. Obs
SAR + Best LandSAT-5 TM	2007[1]	SAR + Summer	0.84	6.89	28.73	8733
	2008[1]	SAR + Summer	0.85	6.84	28.24	8010
	2009[1]	SAR + Autumn	0.83	7.09	29.82	8697
	2010[2]	**SAR + Autumn**	**0.88**	**8.51**	**26.15**	**3201**
SAR + Optical * Products	**2010[2]**	**SAR + Autumn**	**0.88**	**8.15**	**26.90**	**3201**
SAR + Multi-seasonal LandSAT-5 TM	2007[1]	All available images	0.85	6.75	28.37	8733
	2008[1]	All available images	0.85	6.67	27.34	8010
	2009[1]	All available images	0.84	6.91	28.79	8697
	2010[2]	**All available images**	**0.89**	**8.32**	**25.64**	**2098**
SAR	**2008[1]**	**Winter**	**0.80**	**7.88**	**32.08**	**8804**
	2010[2]	**Winter**	**0.81**	**10.17**	**33.16**	**3201**

[1]Utilized 2008 LiDAR as reference dataset and 2008 SAR as input variables; [2]Utilized 2010 LiDAR as reference dataset and 2010 SAR as input variables; * Optical Products refers to the Reflectance + Textures + Indices scenario in Table 4

Fig. 2. Regional scale CC map using the best performing RF integrated L-band and single date LandSAT-5 band reflectance model (2010 L-band & 2010 Autumn LT5 image; coverage excludes extensive cloud cover).

Fig. 3. Predicted CC (RF model derived) versus observed CC (LiDAR-derived) scatterplots for: i) 2008 Multi-seasonal LandSAT-5 Reflectance-only, ii) 2008 SAR-only and iii) integrated 2008 Multi-seasonal LandSAT-5 Reflectance and SAR modelled validation results.

0.5-1%). The year 2010 obtained the highest accuracies (R^2=0.89; RMSE=8.32%; SEP=25.64% for the integrated SAR and multi-seasonal dataset). The best trade-off between accuracy and complexity were given by the 2010 integrated SAR and autumn season reflectance model (R^2=0.88; RMSE=8.51%; SEP=26.15%), as it used a single SAR and single LandSAT-5 image. This model was used to create the regional CC map (figure 2).

The observed CC versus predicted CC XY scatterplots (figures 3i-iii) supported the main findings from LandSAT-5 reflectance-only, SAR-only and integrated SAR backscatter and LandSAT-5 reflectance analyses. The 2008 multi-seasonal LandSAT-5 reflectance only scatterplot (figure 3i) illustrated noticeable overestimation below 25% observed CC mark with major underestimation beyond this point, according to the 1:1 line. In comparison, the 2008 SAR-only scatterplot (figure 3ii) illustrated drastic improvements in reducing the severity of CC overestimation and underestimation. The integration of the SAR and multi-seasonal reflectance scatterplot (figure 3iii) however, yielded a similar trend to the SAR-only scatterplot with a slightly tighter clustering of points around the 1:1 line.

5 Discussion

A previous study demonstrated that L-band ALOS PALSAR data has the potential for accurate tree cover mapping in South African savannahs[15]. Since LandSAT data are freely available and routinely used for regional forest monitoring[20,21,25], this study sought to compare and integrate optical imagery (LandSAT-5) and SAR data (ALOS PALSAR L-band) across various seasons in order to determine if it improves the accuracy of woody cover mapping.

We hypothesized that the season when trees are covered in green foliage, while grasses are dry, should be the best period to retrieve CC with LandSAT data, since this is when there would
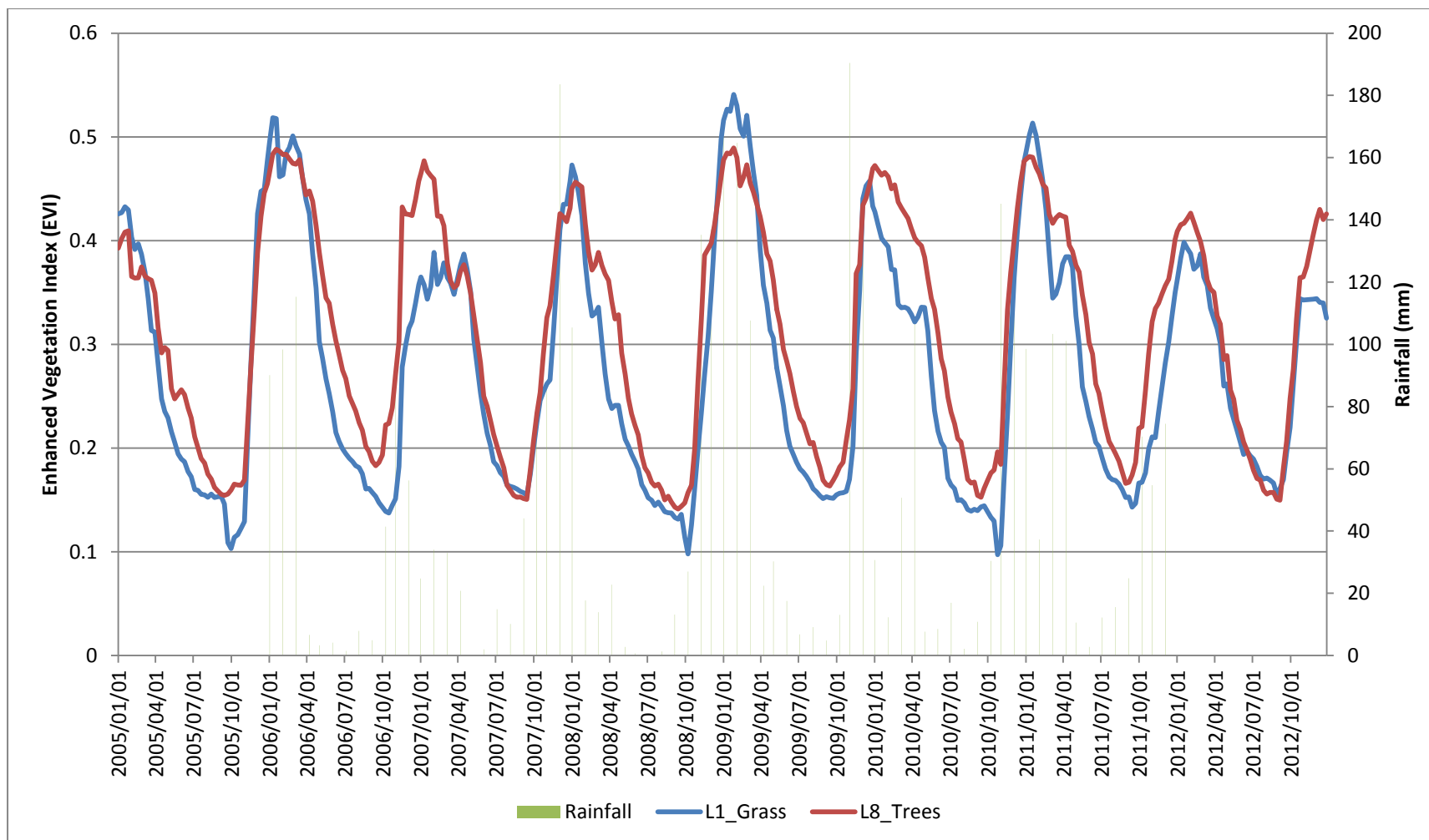
Figure 4: Temporal fluctuations of mean EVI values (extracted from MODIS data) over a predominant grassland site (L1) and a predominant woodland site (L8) between the beginning of 2005 and end of 2012. Rainfall measurements between beginning of 2006 and end of 2011 have also been included. The monthly average rainfall data was extracted from Graskop, Skukuza and Phalaborwa weather stations, in Mpumalanga, and was provided by the South African Weather Services

be the highest spectral contrast between the two life-forms[10,23]. RF modelling results indicated that summer and autumn seasons yielded the highest accuracies, for particular years, with the winter imagery consistently yielding the poorest results. The fluctuations in tree and grass greenness by way of MODIS EVI time series[45] were explored to interpret these seasonal results (figure 4). Monthly EVI values (aggregated from 8 day image composites) were extracted from 500m MODIS MCD43 BRDF-corrected surface reflectance data[46] for grass dominated gabbro and tree dominated granite landscapes, for the years 2005 to 2013. Generally, EVI values follow a distinct rainfall-driven cyclical but variable pattern for trees and grasses, with EVI values peaking during summer (January-February) but falling noticeably to the lowest point in each year during the late winter and early spring (July-September). As evidenced in figure 4, trees green up earlier than the grasses which only start greening up after the first rains and senesce more rapidly than trees[10]. Hence, trees have a longer green period. In addition, savannah trees have a less variable inter-annual phenological cycle since trees use long-term, accumulated water reserves, and are constrained by its root system distribution, in contrast to grasses that rely on variable short term resources such as summer rainfall. Rainfall fluctuations induce more variable grass phenology which in turn complicates the separation between tree and grass. For instance, in 2009 (wet year) the grass dominated landscape reached a higher maximum EVI value compared to the tree dominated landscape, while the opposite was observed during a typically dry year in 2007.

We plotted the grass EVI minus tree EVI graph difference through time to ascertain the periods when the difference between tree and grass greenness were the greatest (Figure 5). These were the most pronounced during brief moments in late spring or during some autumn periods, or in some cases brief summer periods in a dry year (as the case in 2007 and 2010),
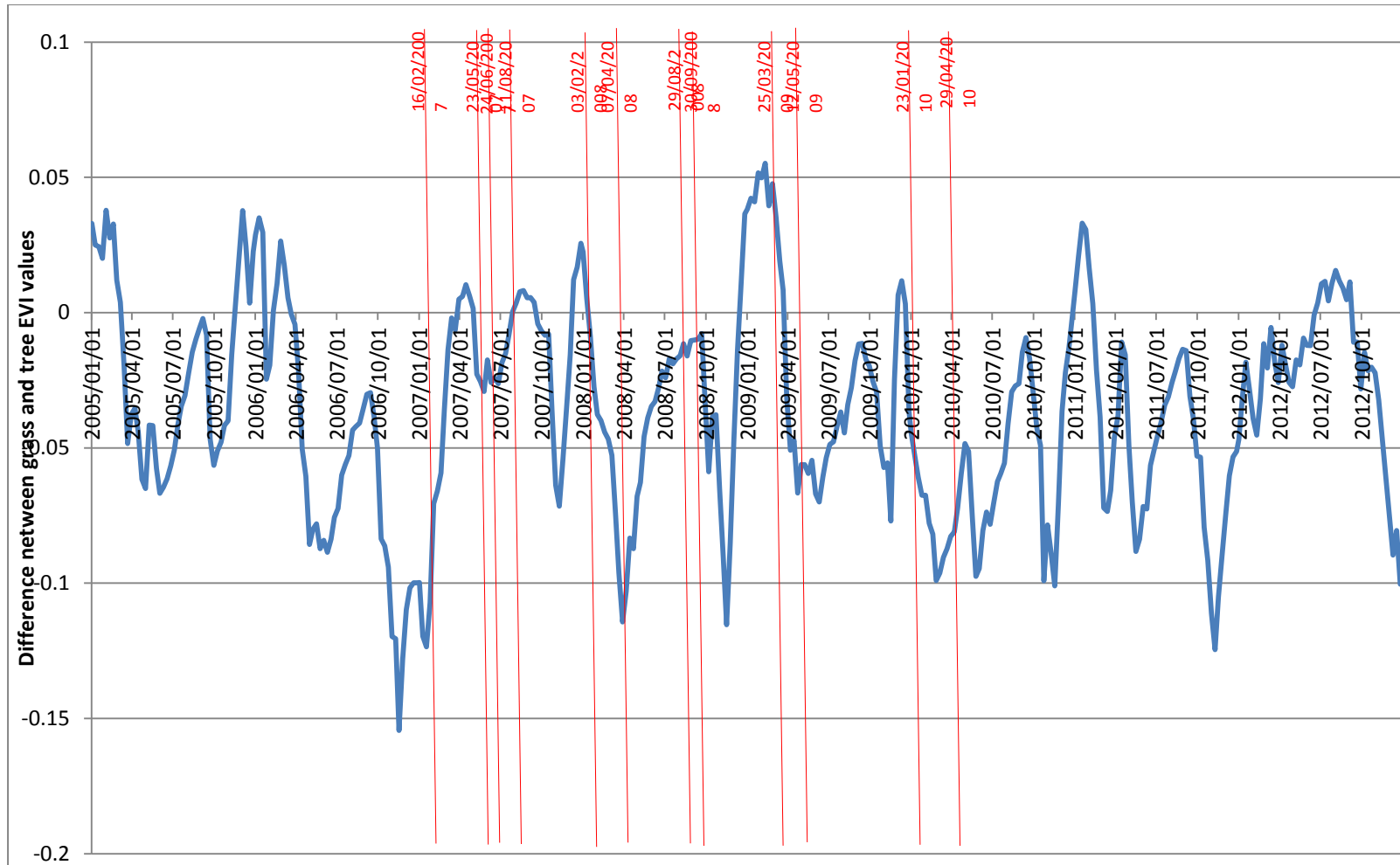
Figure 5: Temporal differences of mean grass and tree EVI values (extracted from MODIS data) over a predominant grassland site and a predominant woodland site between the beginning of 2005 and end of 2012. Red lines indicate the multi-seasonal LandSAT-5 image acquisition dates

while peak differences vary greatly between years and even in some cases were small (e.g. years 2009 and 2010). The best performance obtained from the summer Landsat images in 2007 were most likely caused by dry conditions which resulted in larger differences in the spectral characteristics of grasses and trees. This was not the case when conditions were wetter, with greener grasses (e.g. in 2009). The above patterns were not observed in every year but there was a significant trend between the corresponding tree versus grass difference in EVI and modelled SEP values of the seasonal LandSAT-5 images (R^2 = 0.37; p < 0.05). The poor results obtained with the only spring image available (year 2008) appeared to be linked to the image timing which was acquired too early during the spring season while trees had not started to flush leaves (Figure 4-5). In winter, since most of the trees are deciduous and shed their leaves when grasses are dry, the EVI contrast is consistently the smallest and expectedly produced the poorest results. This contrasts with the Australian example where the dominance of evergreen tree canopies with dry grass, during the prolonged winter periods, supported the successful use of Landsat for mapping tree cover in the SLATS and NCAS-LCCP programmes[20,21]. The brief transitional periods experienced in our South African landscapes during which the contrast between green trees (high tree EVI) and dry grass (low grass EVI) is high are difficult to target, as none of the historic LandSAT image acquisition dates actually fell within the period of largest EVI difference (Figure 5). In addition, the unavoidable presence of clouds, which at times occur irrespective of season, confounds matters further.

Attempts made to improve on the single date LandSAT-5 modelling results by using multi-seasonal approaches yielded slightly better SEP (improvement between 1 and 5%). The combination of images acquired at different seasons provided complimentary spectral information that is not present in a single season (e.g. the dry grass signal not present in

summer seasons). In addition, the combination of image textures with spectral reflectance contributed more towards improving modelling accuracies (improved SEP of 1%) than the incorporation of vegetation indices, which may have contained more redundant spectral information. Image textures are sensitive to local brightness variations arising from tree canopy structural properties, for instance shadow, which are linked to the woody canopy cover[40]. However, improvement from textural metrics remained marginal compared to the 10-13% classification accuracy gain obtained by using texture with LandSAT images for mapping western African forests reported in [47].

None of the best performing LandSAT-5 only models had accuracies that measured up to those obtained with a single winter L-band SAR image (R^2 = 0.72 versus R^2 = 0.81). Again, these results are in contrast with[47] where one single optical LandSAT image mapped with 73% accuracy contrasted forest types, such as evergreen, semi-deciduous, secondary forest, savannahs, while ALOS PALSAR-based classification was significantly less accurate with 49%. The highest modelled CC accuracies achieved by LandSAT derived optical products (R^2=0.72; RMSE=~12%) in our study was significantly lower than the CC accuracies achieved in the Australian Statewide Landcover and Trees Study (SLATS) Program[21] and the Australian National Carbon Accounting System – Land Cover Change Program (NCAS-LCCP)[20] which mapped savannah and forested landscapes with R^2>0.79 and RMSE<10%. The comparatively limited accuracies achieved by LandSAT-5 in this work, for a representative, but relatively small area of savannahs in South Africa, indicated that the implementation of a CC monitoring system based solely on LandSAT-derived data would not be adequate in South African Savannahs. L-band SAR data prove to be a much more effective alternative for reliable and consistent CC mapping and monitoring in this open forest environment.

The integration of the SAR dataset with the best single season and multi-seasonal LandSAT-5 reflectance yielded models with the highest accuracies, which corroborates findings in previous studies[26,27,47]. For instance, the SAR-only and the multi-seasonal LandSAT only models explained 81% and 72% of the CC variance, respectively, while a model combining the two explained 89% of the variance. The significant increase in accuracy (7.5% improvement of SEP) at the high end of the model performance demonstrated that optical reflectance data provided additional information which is complementary to that captured by the SAR backscatter. Significant complementarity between SAR and LandSAT data was demonstrated by [48] with the combined datasets yielding highly accurate results within the Australian NCAS-LCCP (global classification accuracy of 90%).

6   Concluding remarks

This study provides important insights for monitoring woody cover in South African savannahs (35% of land surface, and 95% of forested landscapes). Due to the narrow temporal 'window' during which trees and grass may differ sufficiently in phenological greenness, CC mapping and monitoring in savannahs based solely on Landsat data is not recommended. Extensive cloud cover during the summer or even autumn seasons would further compound this problem. CC mapping should in priority rely on L-band SAR datasets, which need to be prioritised for future acquisition in the region. The recent launch of the ALOS PALSAR-2 sensor, or the future launch of SAOCOM or NISAR, will ensure long-term provision of L-band SAR data. There was significant, yet modest, improvement (R^2 of ~0.08, ~1.9% of RMSE and ~7.5% of SEP) in accuracy when optical reflectance bands were combined with the L-band backscatter. When available the SAR datasets could be complemented with LandSAT mosaics acquired in

summer/autumn, but multi-temporal LandSAT imagery is unlikely to provide improved performance, similarly to the use of vegetation indices and texture metrics. The authors recommend that further testing of the performance of LandSAT imagery, alone and in combination with SAR data, be conducted in other southern African vegetation types where tree canopies are evergreen, such as in commercial plantations, indigenous forests and thickets. It is also recommended that a system based on L-band SAR datasets, with supporting airborne LiDAR data for model calibration and validation, should be applied to other bioregions (e.g. afromontane and coastal indigenous forests) before a national CC monitoring programme can be established in the future.

7   Acknowledgements

References

1. Scholes, R.J., B.H. Walker. *An African Savanna: Synthesis of the Nylsvley Study*. Cambridge: Cambridge University Press (1993)

2. Sankaran, M., J. Ratnam, N. Hanan. "Woody cover in African savannas: the role of resources, fire and herbivory". *Global Ecology and Biogeography, 17*, 236-245 (2008)

3. Shackleton, C.M., S.E. Shackleton, E. Buiten, N. Bird. "The importance of dry woodlands and forests in rural livelihoods and poverty alleviation in South Africa". *Forest Policy and Economics*, **9**, 558-577 (2007)

4. Valentini, R., A. Arneth, A. Bombelli, S. Castaldi, R. Cazzolla Gatti, F. Chevallier, P. Ciais, E. Grieco, J. Hartmann et al. "A full greenhouse gases budget of Africa: synthesis, uncertainties, and vulnerabilities". *Biogeosciences*, **11**, 381-407 (2014)

5. Mograbi, P.J., B.F.N. Erasmus, E.T.F Witkowski, G.P. Asner, K.J. Wessels, R. Mathieu, D.E. Knapp, R.E. Martin, R. Main. "Biomass increases go under cover: Woody vegetation dynamics in South African rangelands". *PLoS ONE*, **10** (5), 1-21 2015

6. Shackleton, C.M., N.J. Griffin, D.I. Banks, J.M. Mavrandonis, S.E. Shackleton. "Community structure and species composition along a disturbance gradient in a communally managed South African savanna". *Vegetatio*, **115** (2), 157-167 (1994)

7. Jennings, S.B., N.D. Brown, D. Sheil. "Assessing forest canopies and understorey illumination: canopy closure, canopy cover and other measures". *Forestry, 72*, 59-73 (1999)

8. Edwards, D. "A broad-scale structural classification of vegetation for practical purposes". *Bothalia*, **14** (3&4), 705-712 (1983)

9. Pereira, H.M., S. Ferrier, M. Walters, G.N. Geller, R.H.G. Jongman, R.J. Scholes, M.W. Bruford, N. Brummitt, S.H.M. Butchart, A.C. Cardoso et al. "Essential Biodiversity Variables". *Science*, **339**, 277-278 (2013)

10. Archibald, S., R.J. Scholes. "Leaf green-up in a semi-arid African savannah – Separating tree and grass responses to environmental cues". *Journal of Vegetation Science*, **18**, 583-594 (2007)

11. Le Toan, T., S. Quegan, M.W.J. Davidson, H. Balzter, P. Paillou, K. Papathanassiou, S. Plummer, F. Rocca, S. Saatchi, H. Shugart, L. Ulander. "The BIOMASS mission: Mapping global forest biomass to better understand the terrestrial carbon cycle". *Remote Sensing of Environment*, **115**, 2850-2860 (2011)

12. Sun, G., K. Jon Ranson, Z. Guo, Z. Zhang, P. Montesano, D. Kimes. "Forest biomass mapping from LiDAR and Radar synergies". *Remote Sensing of Environment*, **115**, 2906-2916 (2011)

13. Santoro, M., A. Shvidenko, I. McCallum, J. Askne, C. Schmullius. "Properties of ERS-1/2 coherence in the Siberian boreal forest and implications for stem volume retrieval". *Remote Sensing of Environment*, **106**, 154-172 (2007)

14. Carreiras, J.M.B., J.B. Melo, M.J. Vasconcelos. "Estimating the above-ground biomass in Miombo savannah woodlands (Mozambique, East Africa) using L-band synthetic aperture radar data". *Remote Sensing Open Access*, **5**, 1524-1548 (2013)

15. Naidoo, L., R. Mathieu, R. Main, W. Kleynhans, K. Wessels, G. Asner, B. Leblon. "Savannah woody structure modelling and mapping using multi-frequency (X-, C- and L-band) Synthetic Aperture Radar data". *ISPRS Journal of Photogrammetry and Remote Sensing*, **105**, 234-250 (2015)

16. Mitchard, E.T.A., S.S. Saatchi, I.H. Woodhouse, G. Nangendo, N.S. Ribeiro, M. Williams, C.M. Ryan, S.L. Lewis, T.R. Feldpausch, P. Meir. "Using satellite radar backscatter to predict above-ground woody biomass: A consistent relationship across four different African landscapes". *Geophysical Research Letters*, **36**, 1-6 (2009)

17. Mathieu, R., L. Naidoo, M.A. Cho, B. Leblon, R. Main, K. Wessels, G.P. Asner, J. Buckley, J. Van Aardt, B.F.N. Erasmus, I.P.J. Smit. "Toward structural assessment of semi-arid African savannahs and woodlands: the potential of multitemporal polarimetric RADARSAT-2 fine beam images". *Remote Sensing of Environment*, **138**, 215-231 (2013)

18. Hansen, M.C., T.R. Loveland. "A review of large area monitoring of land cover change using Landsat data". *Remote Sensing of Environment*, **122**, 66-74 (2012)

19. Foody, G.M., R.M. Lucas, P.J. Curran, M. Honzak. "Mapping tropical forest fractional cover from coarse spatial resolution remote sensing imagery". *Plant Ecology*, **131**, 143-154 (1997)

20. Lehmann, E.A., J.F. Wallace, P.A. Caccetta, S.L. Furby, K. Zdunic. "Forest cover trends from time series Landsat data for the Australian continent". *International Journal of Applied Earth Observation and Geoinformation*, **21**, 453-462 (2013)

21. Armston, J.D., R.J. Denham, T.J. Danaher, P.F. Scarth, T.N. Moffiet. "Prediction and validation of foliage projective cover from Landsat-5 TM and Landsat-7 ETM+ imagery". *Journal of Applied Remote Sensing*, **3**, 1-28 (2009)

22. Boggs, G.S. "Assessment of SPOT 5 and QuickBird remotely sensed imagery for mapping tree cover in savannas". *International Journal of Applied Earth Observation and Geoinformation*, **12**(4), 217-224 (2010)

23. Zeidler, J., M. Wegmann, S. Dech. "Spatio-temporal robustness of fractional cover upscaling: a case study in semi-arid Savannah's of Namibia and Western Zambia". *Proc. SPIE 8538, Earth*

Resources and Environmental Remote Sensing/GIS Applications III, 85380S (October 25, 2012); 1-10 (2012) [doi: 10.1117/12.970623]

24. Nichol, J.E., L.R. Sarker. "Improved biomass estimation using the texture parameters of two high-resolution optical sensors". *IEEE Transactions on Geoscience and Remote Sensing*, **49** (3), 930-948 (2011)

25. Ju, J., D.P. Roy, E. Vermote, J. Masek, V. Kovalskyy. "Continental-scale validation of MODIS-based and LEDAPS Landsat ETM+ atmospheric correction methods". *Remote Sensing of Environment*, **122**, 175-184 (2012)

26. Moghaddam, M., J.L. Dungan, S. Acker. "Forest variable estimation from fusion of SAR and multispectral optical data". *IEEE Transactions on Geoscience and Remote Sensing*, **40** (10), 2176-2187 (2002)

27. Lucas, R.M., N. Cronin, M. Moghaddam, A. Lee, J. Armston, P. Bunting, C. Witte. "Integration of radar and Landsat-derived foliage projected cover for woody regrowth mapping", Queensland, Australia. *Remote Sensing of Environment*, **100** (3), 388-406 (2006)

28. Asner, G.P., D.E. Knapp, T. Kennedy-Bowdoin, M.O. Jones, R.E. Martin, J. Boardman, C.B. Field. "Carnegie Airborne Observatory: in-flight fusion of hyperspectral imaging and waveform LiDAR for 3D studies of ecosystems". *Journal of Applied Remote Sensing*, **1**, 1-27 (2007)

29. Mucina, L., M.C. Rutherford (Eds.). *The vegetation of South Africa, Lesotho and Swaziland*. South African National Biodiversity Institute, Pretoria (2006)

30. Breiman, L. "Random forests". *Machine Learning*, **45** (1), 5-32 (2001)

31. Wessels, K.J., R. Mathieu, B.F.N. Erasmus, G.P. Asner, I.P.J. Smit, J.A.N. van Aardt et al. "Impact of communal land use and conservation on woody vegetation structure in the Lowveld savannas of South Africa". *Forest Ecology and Management*, **261**, 19-29 (2011)

32. Huete, A.R., H.Q. Liu, K. Batchily, W. van Leeuwen. "A comparison of vegetation indices global set of TM images for EOS-MODIS". *Remote Sensing of Environment*, **59** (3), 440-451 (1997)

33. Sims, D.A., J.A. Gamon. "Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages". *Remote Sensing of Environment*, **81** (2-3), 337-354 (2002)

34. Goel, N.S., W. Qin. "Influence of canopy architecture on various vegetation indices and LAI and FPAR: a computer simulation". *Remote Sensing Reviews*, **10**, 309-347 (1994)

35. Huete, A.R., R.D. Jackson. "Soil and atmosphere influences on the spectra of partial canopies". *Remote Sensing of Environment*, **25** (1), 89-105 (1988)

36. Jordan, C.F. "Derivation of leaf area index from quality of light on the forest floor". *Ecology*, **50** (4), 663-666 (1969)

37. Rouse, J.W., R.H. Haas, J.A. Schell, D.W. Deering. "Monitoring vegetation systems in the Great Plains with ERTS". In: S.C. Fraden, E.P. Marcanti, M.A. Becker (eds.), *Third ERTS-1 Symposium*, 10-14 Dec. 1973, NASA SP-351, Washington D.C. NASA, 309-317 (1974)

38. Sousa, C.L., F.J. Ponzoni. "Avaliacao de indices de vegetacao e de bandas TM/Landsat para estimative de volume de madeira em floresta implantada de Pinus spp". *IX Simposio Brasileiro de Sensoriamento Remoto*, p.11 (1998)

39. Haralick, R.M., K. Shanmugam, I. Dinstein. "Textural Features for Image Classification", *IEEE Transactions on Systems, Man and Cybernetics*, **3** (6), 610-621 (1973)

40. Asner, G.P., M. Keller, R. Pereira, J.C. Zweede. "Remote sensing of selective logging in Amazonia assessing limitations based on detailed field observations, Landsat ETM+, and textural analysis". *Remote Sensing of Environment*, **80**, 483-496 (2002)

41. Lu, D. "Aboveground biomass estimation using Landsat TM data in the Brazilian Amazon". *International Journal of Remote Sensing*, **26** (12), 2509-2525 (2005)

42. Urbazaev, M., C. Thiel, R. Mathieu, L. Naidoo, S.R. Levick, I.P.J. Smit, G.P. Asner, C. Schmullius "Assessment of the mapping of fractional woody cover in southern African savannas using multi-temporal and polarimetric ALOS PALSAR L-band images". *Remote Sensing of Environment*, **166**, 138-153 (2015)

43. Prasad, A.M., L.R. Iverson, A. Liaw. "Newer classification and regression tree techniques: bagging and random forests for ecological prediction". *Ecosystems*, **9** (2), 181-199 (2006)

44. Liaw, A., M. Wiener. "Classification and regression by randomForest". *R News*, **2** (3), 18-22 (2002)

45. Jin, C., X. Xiao, L. Merbold, A. Arneth, E. Veenendaal, W.L. Kutsch. "Phenology and gross primary production of two dominant savanna woodland ecosystems in Southern Africa". *Remote Sensing of Environment*, **135**, 189-201 (2013)

46. Schaaf, C.B., F. Gao, A.H. Strahler, W. Lucht, X. Li, T. Tsang, N.C. Strugnell, X. Zhang, Y. Jin, J.P. Muller, P. Lewis et al. "First operational BRDF, albedo nadir reflectance products from MODIS". *Remote Sensing of Environment*, **83**, 135-148 (2002)

47. Laurin, G.V., V. Liesenberg, Q. Chen, L. Guerriero, F. Del Frate, A. Bartolini, D. Coomes, B. Wilebore, J. Lindsell, R. Valentini. "Optical and SAR sensor synergies for forest and land cover mapping in a tropical site in West Africa". *International Journal of Applied Earth Observation and Geoinformation*, **21**, 7-16 (2013)

48. Lehmann, E.A., P. Caccetta, K. Lowell, A. Mitchell, Z. Zhou, A. Held, T. Milne, I. Tapley. "SAR and optical remote sensing: Assessment of complementarity and interoperability in the context of a

large-scale operational forest monitoring system". *Remote Sensing of Environment*, **156**, 335-348 (2015)

49. Chen, X., L. Vierling, E. Rowell, T. DeFelice. "Using lidar and effective LAI data to evaluate IKONOS and Landsat 7 ETM+ vegetation cover estimates in a ponderosa pine forest. *Remote Sensing of Environment*, **91**, 14-26 (2004)