Mapping canopy gaps in an indigenous subtropical coastal forest using high resolution worldview-2 data

Oupa Malahlela [a, b] *, Moses Azong Cho [a, b] and Onisimo Mutanga [b]

[a] *Earth Observation Research Group, Natural Resources and Environment, Council for Scientific and Industrial Research, Pretoria 0001, South Africa*

[b] *Geography Department, University of KwaZulu-Natal, Pietermaritzburg Campus, Scottsville 3209, South Africa*

*Corresponding author. Current address: P O Box 395, Pretoria, 0001, South Africa. Tel: +27 128412233. Email address: OMalahlela@csir.co.za
M. A. Cho. Current address: P O Box 395, Pretoria, 0001, South Africa. Tel: +27 128413669. Email address: MCho@csir.co.za
O. Mutanga. Current address: P. Bag X01, Scottsville 3209, Pietermaritzburg, South Africa. Email address: MutangaO@ukzn.ac.za

Abstract

Invasive species usually colonize canopy gaps in tropical and sub-tropical forests, which results in loss of native species. Therefore, an understanding of the location and distribution of canopy gaps will assist in predicting the occurrence of invasive species in such canopy gaps. We tested the utility of WorldView-2 with eight (8) spectral bands at 2 m spatial resolution to delineate forest canopy gaps in a subtropical Dukuduku coastal forest in South Africa. We compared the four (4) conventional visible-near infrared bands with the eight (8) band WorldView-2 image. The 8-band WorldView-2 image yielded higher overall accuracy of 86.90% (kappa = 0.82) than the resampled conventional 4 band image which yielded an overall accuracy of 74.64% (kappa = 0.63) in pixel-based classification. We further compared the vegetation indices which were derived from four conventional bands with those derived from WorldView-2 bands. The Enhanced Vegetation Index (EVI) yielded the highest overall accuracy in the category of conventional indices (85.59% at kappa = 0.79), while the modified Plant Senescence Reflectance Index (mPSRI) involving the red-edge band showed the highest overall accuracy (93.69%) in the category of indices derived from an eight band WorldView-2 imagery in object-based classification. Overall, the study shows that the unique high resolution WorldView-2 data can improve the delineation of canopy gaps as compared to the conventional multispectral bands.

Keywords: Enhanced vegetation index, invasive species, modified plant senescence reflectance index,

1. Introduction

Globally, subtropical coastal forest constitutes one of the smallest vegetation biomes. In South Africa, subtropical forest patches are reported to be disproportionally rich in biodiversity when compared to other dominant biomes such as the savannahs (Geldenhuys, 1989). Over the years, the forests have been fragmented into patches of various sizes as a result of anthropogenic activities such as subsistence farming, commercial agriculture and human settlement expansions (Fourcade, 1889; Geldenhuys, 1989). Some of the patches are intensively managed e.g. the Dukuduku coastal forest (Van Gils *et al*., 2006) to avert further degradation and loss of biodiversity. However, the sustainability of indigenous biodiversity in the remnant patches is threatened by the presence of alien invasive species (Van Gils *et al*., 2006; Moore, 2004). These species take advantage of canopy gaps that occur within the patches for their establishment and proliferation. Canopy gaps can be formed from the fall of dead trees (Kupfer and Runkle, 1996; Brokaw and Grear, 1991), selective timber harvesting (Suarez *et al*., 1997), or tree fall from disturbance events such as strong winds (Brokaw, 1982; Whitmore, 1989). Naturally, in a stable subtropical forest, canopy gaps are closed up by the regeneration of indigenous species through a process of succession (Orians, 1982). However, in South Africa, forest canopy gaps in coastal forests may be invaded by light-loving alien invasive species such as *Chromolaena odorata* (Weiss and Noble, 1984; Goodall and Erasmus, 1996). This species has an allelopathic influence to the seedlings of the indigenous species, and therefore hindering their recruitment by shading the seedlings and altering the chemical composition of the soil beneath it (Codilla and Metillo, 2011). Therefore, knowledge of the location and distribution of canopy gaps is crucial for controlling the proliferation of invasive species in the remaining subtropical forest patches in South Africa.

Conventionally, the delineation of forest canopy gaps is done through field surveys (Brokaw and Grear, 1991). However, field based methods are limited to areas that are easily accessible, and are seldom used for larger and wider areas (Runkle, 1982). On the other hand, remote sensing has been recommended as a more cost effective and less laborious alternative to field-based methods for broad areas (Woodcock *et al*., 2001). The detection of forest canopy gaps has been studied previously using Landsat imagery with a spatial resolution of 30 meters (Negrón-Juárez, 2011; Asner *et al*., 2004). The limitation of this type of coarse resolution sensor lies in its inability to map canopy gaps that are less than 30 m in size (Clark *et al*., 2004). The use of high spatial resolution multispectral sensors such as *Système Pour l'Observation de la Terre* (SPOT) (10 m spatial resolution) and IKONOS (4m spatial resolution) has mitigated the spatial resolution problem by improving accuracy for characterizing vegetation (Clark *et al*., 2004). However, several studies (Knipling, 1970; Mutanga and Skidmore 2004; Chen *et al*. 2009; Cho *et al*., 2009) have documented that these multispectral sensors suffer from the saturation of the visible-near infrared signal in dense vegetation. This problem could potentially make it difficult to discriminate between tree canopies and vegetated gaps in closed canopy coastal forest (Figure 1) (Weiss and Baret, 1999).

Recently, hyperspectral data and Light detection and ranging (LiDAR) have been utilized to mitigate the saturation problem common in conventional multispectral sensors (Treitz *et al*., 2003; Ustin and Trabuco, 2000). The Normalized Difference Vegetation Index (NDVI) derived from broad-band imagery has been shown to saturate in in high density vegetation e.g. leaf area index > 3, a typical habitat condition of vegetated tropical forest gaps. The red edge bands (700 – 740 nm) present in hyperspectral sensors have been shown to solve the saturation problem, e.g. the red edge NDVI has been used to improve the estimation of vegetation biomass at high canopy density (Mutanga and Skidmore, 2004; Smith *et al*., 2004; Sellers, 1985; Cho and Skidmore 2009). This is due to the fact increasingvegetation density causes saturation in the red (680 nm) absorption trough but a broadening of the absorption trough, causing shifts in the red edge slope towards longer wavelengths (Dawson and Curran, 1998).

There have been a number of successful studies that have incorporated LiDAR technology for mapping forest gaps in tropical forests. For example, Kellner *et al*. (2009) used LiDAR technology to study the structural characteristics of canopy gaps in tropical rainforest gaps. On the other hand, Vepakomma *et al*. (2008) have delineated forest gaps in the boreal forest using LiDAR technology, resulting in mapping accuracy of 96%. More recently, Asner *et al.* (2013) highlighted the significance of LiDAR technology in studying the distribution of canopy gaps in the Amazon area. However, both the hyperspectral and LiDAR technologies have not been fully explored due to the high cost associated with acquiring these data (Asner *et al*., 2004). Moreover, the question of whether all hundreds of contiguous spectral bands of hyperspectral data are needed to delineate vegetated forest canopy gaps arises, given the reported redundancy in hyperspectral data (Cho *et al*., 2012; Mutanga *et al*., 2012).

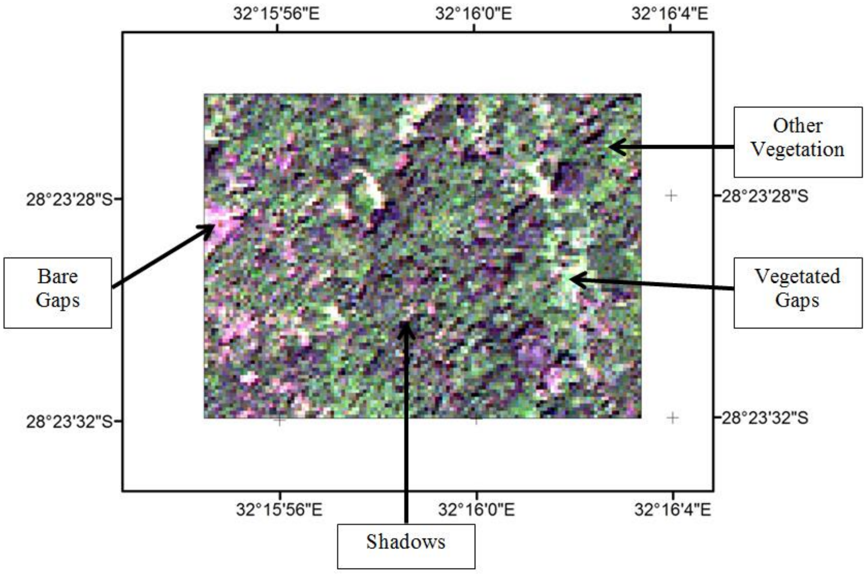

Figure 1. A subset of WorldView-2 image showing forest canopy gaps (vegetated and non-vegetated) surrounded by tree canopies and shadow gaps.

The development of high spatial and spectral resolution multispectral sensors such as WorldView-2 (Ozdemir and Karnieli, 2011) and RapidEye (Ramoelo *et al*., 2012) has opened new opportunities for diverse vegetation characterization in terms of its biophysical and biochemical properties. The red edge band present in these sensors (Figure 2) has been successfully used to estimate leaf nitrogen (Ramoelo *et al*., 2012; Cho *et al.*, 2013) and biomass in a dense canopy environment (Mutanga *et al*., 2012). WorldView-2 sensor consists of spectral bands that are strategically designed to maximize the sensitivity of signal to plant characteristics such as biomass, health and productivity. The question that arises is whether the presence of red edge band in WorldView-2 can provide improved discrimination of forest canopy gaps in a closed canopy coastal forest when compared to the conventional red, green, blue and NIR bands present in conventional sensors such as SPOT, IKONOS or Landsat?

In addition, canopy gap delineation requires classification techniques that will aid in separating canopy gaps from the rest of the forest canopies. A number of studies have delineated forest gaps using conventional pixel-based classifiers such as maximum-likelihood and spectral angle mapper classifiers (Fox *et al*., 2000; Negrón-Juárez *et al*., 2011; Betts *et al*., 2005). Pixel-based classification methods have their own short-comings when applied to closed canopy forests. These techniques only consider spectral information and not the geometry and size of individual canopy gaps (Mallinis *et al*., 2008; Kim *et al*., 2009). The question arises whether object-based image analysis can assist in minimizing the short-comings of pixel-based classifications in delineating forest gaps, and thereby improve classification. An object-based classifier has some advantages in that it requires less computational space compared to methods such as Artificial Neural Network and Random Forest (Blaschke and Strobl, 2001).

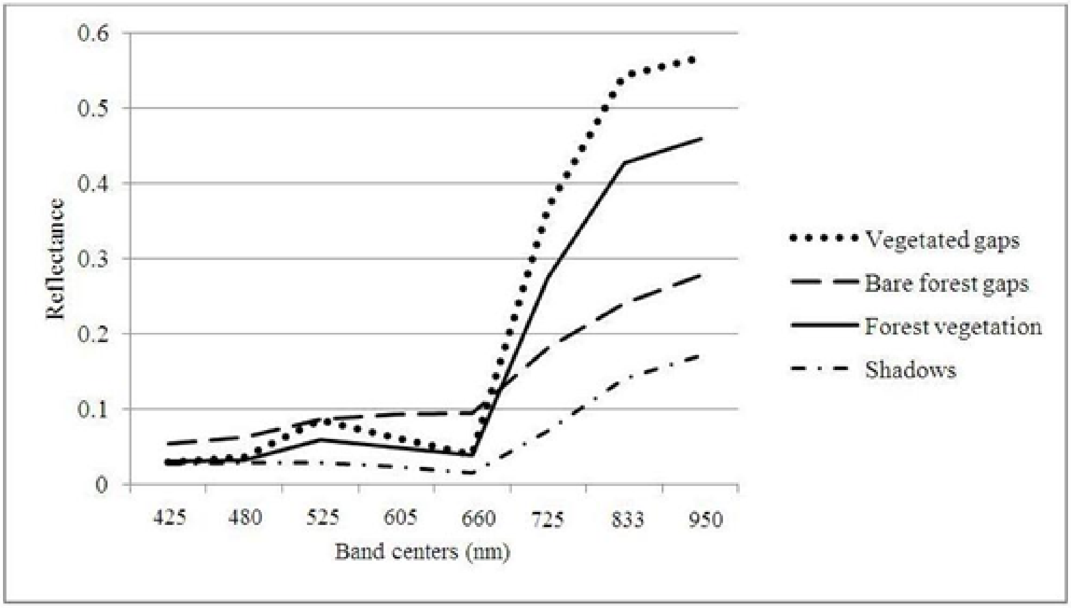

Figure 2. Spectral profile of five averaged spectra of four classes from WorldView-2 imagery of the Dukuduku forest. This figure shows that four forest classes are distinguishable from the red edge to NIR regions, by observing their spectral profiles.

The aims of this study were to:

a) assess the suitability of WorldView-2 multispectral bands in delineating forest gaps, when compared to conventional visible-near infrared bands common in conventional sensors.

b) determine the best vegetation indices for delineation of forest gaps in closed canopy coastal forest using object-based classification

c) investigate the performance of an object-based classification technique over pixel-based classification methods for delineating forest canopy gaps.

2. Methods

2.1 *Study area*

The study was undertaken in Dukuduku indigenous coastal forest located near St. Lucia, north eastern part of KwaZulu-Natal, South Africa. It is located within the Mtubatuba Local Municipality between the geographical coordinates of 28°38’33”S and 32°31’67” E, about 226 km north of Durban (Figure 3). The forest covers a land area of about 3 200 hectares.

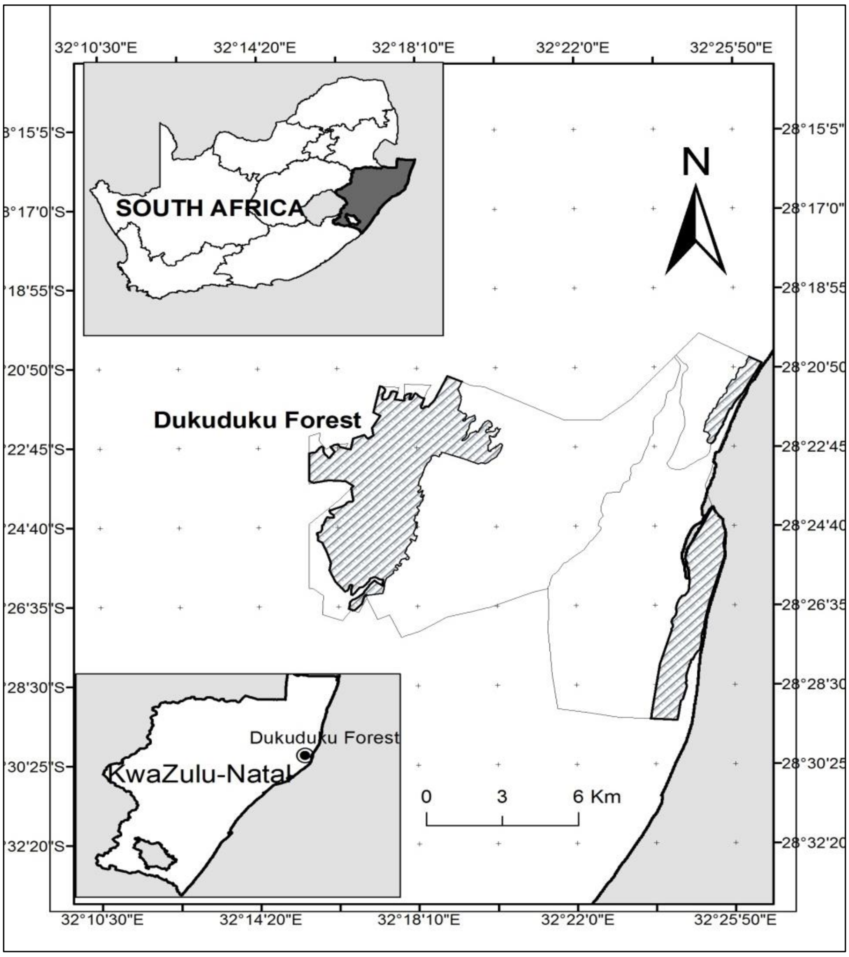

Figure 3. Location of the study area adjacent to the Indian Ocean, north-eastern sea-shore of KwaZulu-Natal, South Africa.

The area was chosen for the study because it is the largest remaining patch of indigenous forest on the north-eastern coastal shoreline of KwaZulu-Natal. Alien species invasion especially by *Chromolaena odorata* poses a major threat to the indigenous forest in this area.

This forest is surrounded by sugar plantation farms, *Eucalyptus* plantations and villages that practise subsistence farming.

2.2 *Image acquisition and pre-processing*

The WorldView-2 image with 8 multispectral bands (Table 1) acquired on the 1 December 2010 was used for the study. WorldView-2 has four new additional bands that are not present in well-known sensors such as SPOT, IKONOS or Landsat. This sensor has spectral bands that are strategically located to aid in vegetation analysis. The spectral bands have wavelength ranges of 400 – 450 nm (absorbed by chlorophyll), 450 – 510 nm (absorbed by chlorophyll), 510 – 580 nm (sensitive to plant health such as greenness), 585 – 625 nm (absorbed by carotenoids – detects ‘yellowness’ of vegetation), 630 – 690 nm (absorbed by chlorophyll), 705 – 745 nm (sensitive to subtle variations vegetation greenness), 770 – 895 nm (sensitive to leaf mass and moisture content), 860 – 1040 nm (sensitive to leaf mass and moisture content) (Ustin *et al*., 2009). The WorldView-2 (WV-2) imagery was geometrically corrected by the supplier (Updike and Comp, 2010). To assess the accuracy of geometric correction, coordinates of some iconic points on the ground, including road junctions and isolated tree canopies were extracted from the image and loaded onto handheld GPS, with maximum spatial accuracy of 4m.

Table 1. WorldView-2 bands and their respective band centers in nanometers, compared to bands present/absent in SPOT and Landsat.

Band Name	Wavelength (μm)	Band Centers (nm)	Present in	
			SPOT	*Landsat*
Coastal Blue	0.40 – 0.45	425	NO	NO
Blue	0.45 – 0.51	480	NO	YES
Green	0.51 – 0.58	545	YES	YES
Yellow	0.59 – 0.63	605	NO	NO
Red	0.63 – 0.69	660	YES	YES
Red-Edge	0.70 – 0.75	725	NO	NO
Near Infrared-1	0.77 – 0.89	833	YES	YES
Near Infrared-2	0.86 – 1.04	950	NO	NO

Atmospheric correction was carried out using ATCOR 2/3 version module (developed and distributed by ReSe Applications). This software was used for its availability and for being one of the widely used image atmospheric correction algorithms (San and Suzen, 2010). An ATCOR2/3 is based on a MODTRAN 5 code (Berk *et al.*, 1998). A MODTRAN 5 code (MODerate resolution atmospheric TRANsmission) is an algorithm designed to model atmospheric propagation of electromagnetic radiation by calculating the databases of atmospheric look-up tables for the spectral regions of between 0.2 and 100 μm (Berk *et al*., 1998) . The atmospheric conditions specified in ATCOR software for this image processing was the ‘tropical rural’ conditions. Figure 4 shows the stages through image analysis and eventual gap delineation using WorldView-2 imagery of the Dukuduku forest.

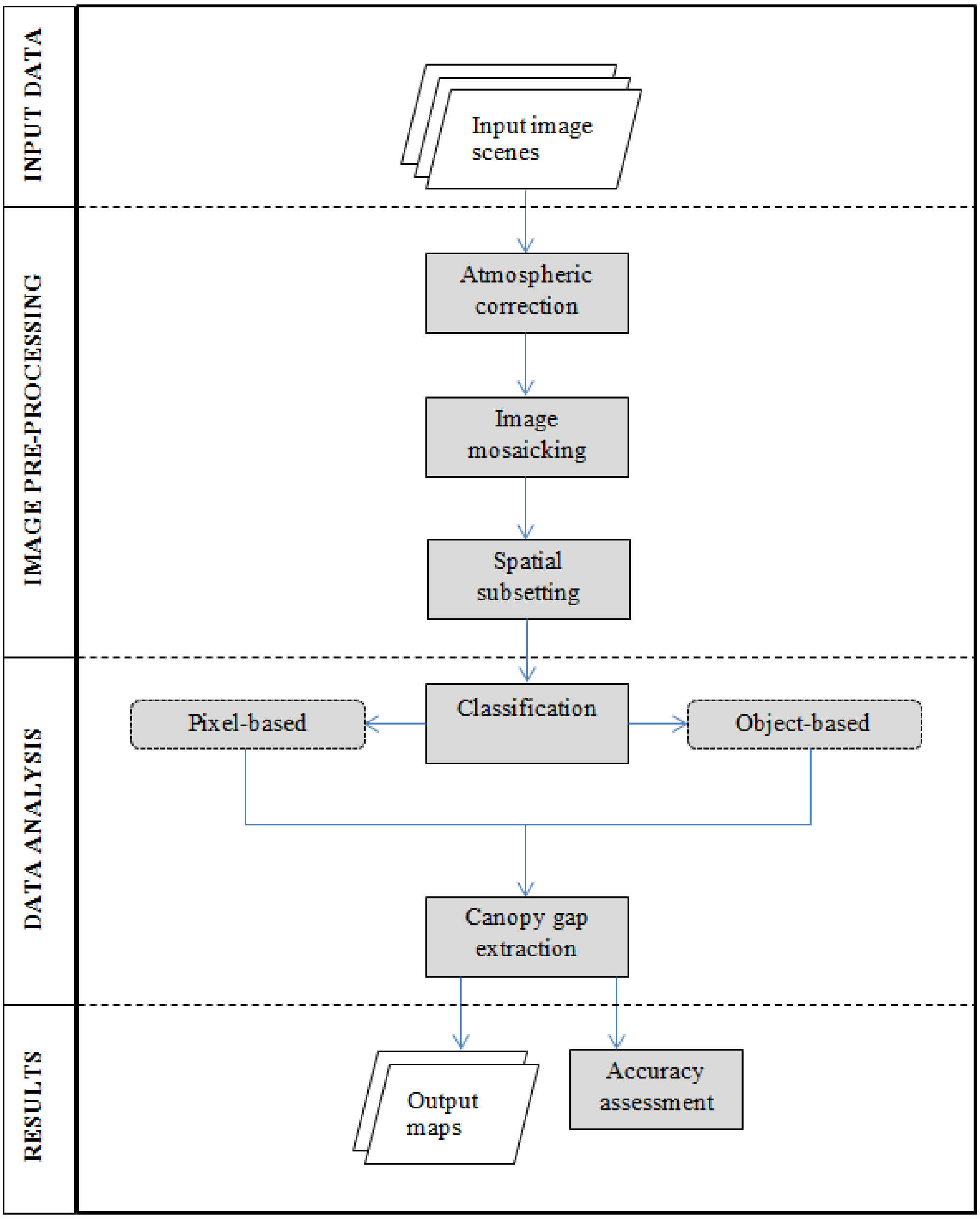

Figure 4. A flow diagram showing the processing scheme for delineating forest gaps adopted in the present study.

2.3 *Field data collection*

Two field data collection trips were undertaken in order to record data on forest gaps, invasive plant species, and the surrounding indigenous vegetation. The first field data collection trip was undertaken between 24 July 2011 and 4 August 2011, while the second one was from the 21 October to 2 November 2011. These dates were primarily dictated by the condition of the atmosphere to avoid rainy weather at the coast and the logistical constraints, and not necessarily the phenological characteristics of the forest. Thirteen line transects were randomly pre-selected across the entire forest, which represented the general trend of vegetation characteristics and to maximize area that is covered within the forest boundaries (Eberhardt, 1986; Battles *et al*., 1996). Each transect was had a minimum length of 1 km, and the canopy gaps were recorded within 10 m width of the transects. Forest canopy gaps were identified along these transects (Figure 5), which included vegetated forest gaps (with/without invasive plant species), bare forest gaps, and the individual tree crowns of forest vegetation. Transects were numbered accordingly from 1 to 13, with transect 12 sampled for forest gaps along the edges. Sampling was done starting from deep inside the forest towards forest edges, according to individual transect orientation. We have adopted line transects method for its simplicity, its popularity for ecological modelling (Forbes and Gross, 1921; Buckland *et al*. 2001), its efficiency and it is relatively inexpensive for many biological populations (Anderson *et al*. 1979).

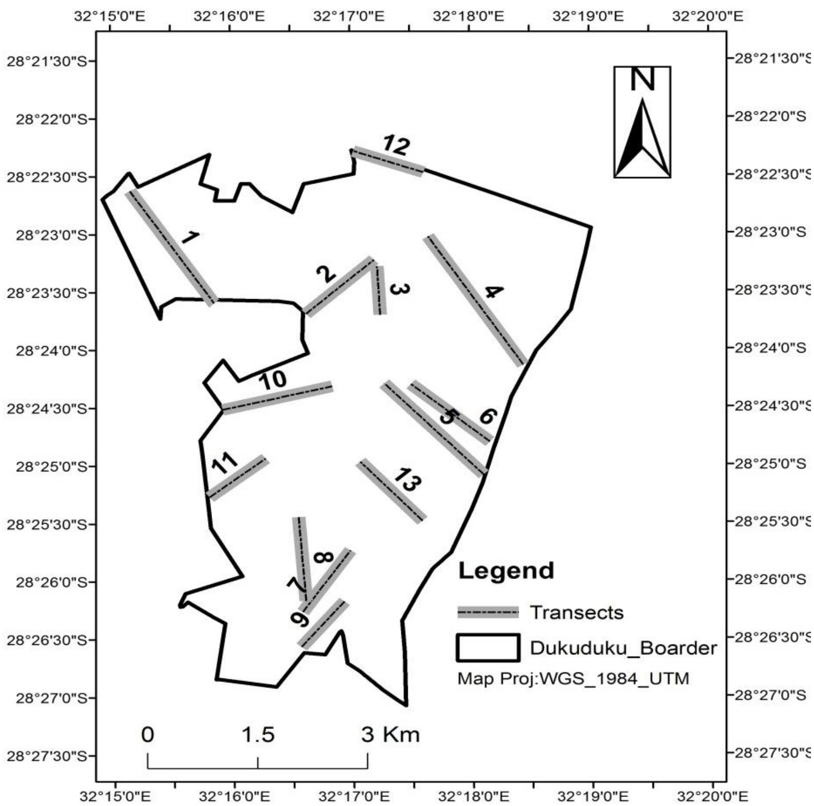

Figure 5. Distribution of line transects followed during field sampling

A standard Global Positioning System (GPS) named Garmin Vista eTRax™ was used to record the location of each forest gap and the surrounding vegetation type. A total of 276 samples or cases were collected. The data were randomly split into 60% calibration (n = 165) and 40% (n = 111) for validation. We have decided on this split in order to maximize the model training performance and to avoid model over-fitting when validation dataset is too large. The calibration dataset was used for training the classifiers. The validation dataset was used to assess the accuracy of the classification techniques.

2.4 *Image processing*

In image processing two classification approaches on WorldView-2 image were tested for this study, namely pixel-based classification and object-based image analysis. For pixel-based image analysis, classification methods such as maximum likelihood (MLC), support vector machines (SVM) and Random Forests (RF) were explored to determine the best commonly used pixel-based classifier for forest gap delineation. In object-based analysis, a multi-resolution segmentation algorithm was explored for creating image objects at different scale parameters (10, 25, and 35), shape and compactness. The shape range between 0.0- 0.9 was tested, while compactness factor was tested in the same manner as shape in eCognition software.The higher the value assigned to shape factor (value >0) the more the shape of an object in an image is considered. In addition, the threshold values from computed vegetation indices were used to discriminate forest vegetation into separate classes.

2.4.1 *Pixel-based classification*

Three pixel-based classification methods were tested for the study, namely, maximum likelihood (MLC), support vector machines (SVM) and Random Forest (RF) classifiers. MLC classifier is based on a normalized (Gaussian) estimate of the probability density function of each class. It is known to be the most powerful classification method when accurate training data is provided and one of the most widely used algorithms (Zhou and Robson, 2001; Lillesand *et al*, 2004). SVM are algorithms based on the statistical learning theory and have the aim of determining the location of decision boundaries that provide the optimal separation of classes (Vapnik, 1995). On the other hand, RF algorithm is increasingly being used today due to its high prediction accuracy and information on variable importance for classification (Touw *et al*, 2012). RF are non-parametric, relatively robust to outliers and noise algorithms that train an ensemble of individual decision trees based on samples, their class designation, and variables for classification and regression (Touw *et al*., 2012).

MLC and SVM classifiers were performed in ENVI 4.8 software with IDL (Exelis Visual Information Solutions, Boulder, Colorado). RF classification was done in EnMap, through the IDL module. For the RF, the default settings were accepted for classification, with 100 trees chosen in EnMap for classification in RF.

2.4.2 *Object-based classification*

In object-based image analysis (OBIA) the process of classification began with image segmentation where similar pixels are merged together using homogeneity criteria such as spectral similarity, weight or compactness (Baatz *et al*., 2004). The performance of OBIA relies on the quality of individual image segments and the accuracy of the segmentation process, and this depends on segmentation parameters such as scale, shape and compactness (Blaschke and Strobl, 2001). A multi-resolution segmentation algorithm was used to begin the segmentation process, where three scale parameters (10, 25, and 35) were tested. The scale parameter of 10, a compactness factor of 0.1 and a shape factor of 0.7 were selected because the objects created were not very different in size and shape from what was observed in the field, and the accuracy of class discrimination decreased with increasing scale parameter and compactness (Figure 6). A multi-resolution algorithm is embedded in eCognition software, which is an object-based processing program made available in 2000 from Definiens Imaging GmbH (Blaschke and Strobl, 2001). Vegetation indices were calculated in eCognition and were used to discriminate four (4) forest classes, using individual index's decision tree, where thresholds were defined at each level to allocate segmented objects to a particular class. These classes were decided by considering the fact that the indigenous forest is a protected area, and that some other possible classes were not of interest.

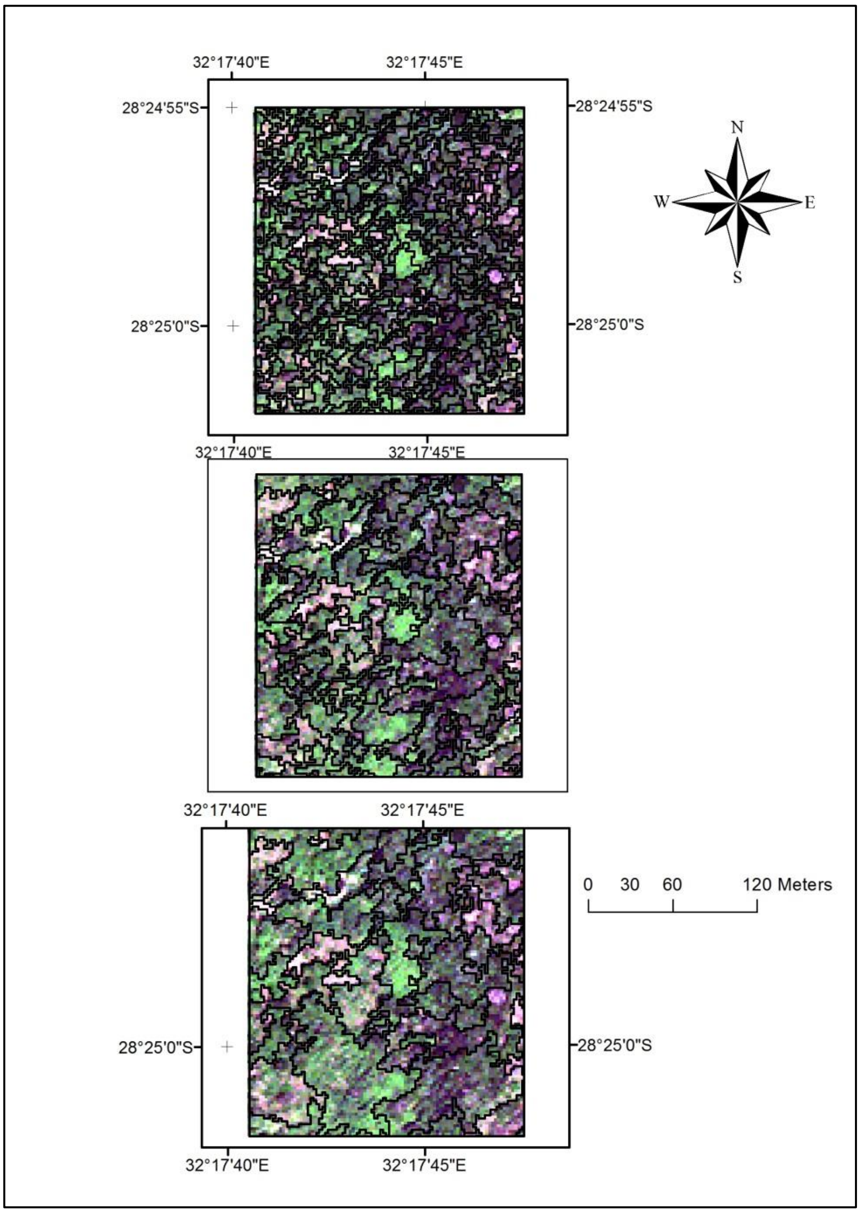

Figure 6: Segmentation of WorldView-2 image of Dukuduku forest at different scale parameters (A = 10, B = 25 and C = 35).

We tested established vegetation indices that were derived from spectral bands present in conventional satellites such as Landsat and SPOT, and those that can be derived from WorldView-2 new bands for delineation of forest gaps. Vegetation indices were selected from those that were sensitive to broadband greenness, narrowband greenness and plant senescence (Asner *et al*., 2002). Table 2 shows vegetation indices that are derived from conventional Red, Green, Blue and Near-Infrared bands, and also the vegetation indices derived from new WV-2 bands.

Tables 2. Vegetation indices explored for delineating forest canopy gaps, derived from conventional sensors and WorldView-2 sensor.

(a)

	Index	Equation	Reference
Traditional Sensor (Landsat/SPOT)	Normalized Difference Vegetation Index ($NDVI_{660}$)	$\mathrm{NDVI}_{660} = \frac{\rho_{NIR} - \rho_{Red}}{\rho_{NIR} + \rho_{Red}}$	Jackson et al. (1983)
	Enhanced Vegetation Index (EVI)	$EVI = 2.5\left(\frac{\rho_{NIR} - \rho_{Red}}{\rho_{NIR} + 6\rho_{Red} - 7.5\rho_{Blue} + 1}\right)$	Huete et al.(1997)
	Atmospherically Resistant Vegetation Index(ARVI)	$ARVI = \frac{\rho_{NIR} - (2\rho_{Red} - \rho_{Blue})}{\rho_{NIR} + (2\rho_{Red} - \rho_{Blue})}$	Kaufman and Tanre (1992)
	Green Normalized Difference Vegetation Index ($NDVI_{545}$)	$NDVI_{545} = \frac{\rho_{833} - \rho_{545}}{\rho_{833} + \rho_{545}}$	Gitelson and Merzlyak (1994)

(b)

	Index	Equation	Reference
WorldView-2 Sensor	Modified plant Senescence Reflectance Index (mPSRI)	$\mathrm{mPSRI} = \frac{\rho_{\mathrm{Red}} - \rho_{\mathrm{Blue}}}{\rho_{\mathrm{Red\ Edge}}}$	After Merzlyak et al. (1999)
	Normalized Pigment Chlorophyll Index (NPCI)	$\mathrm{PCI} = \frac{\rho_{660} - \rho_{425}}{\rho_{660} + \rho_{425}}$	Peñuelas et al. (1995)
	Red Edge Normalized Difference Vegetation Index ($NDVI_{725}$)	$\mathrm{NDVI}_{725} = \frac{\rho_{835} - \rho_{725}}{\rho_{835} + \rho_{725}}$	Gitelson and Merzlyak (1994)
	Yellowness Index (YI)	$\mathrm{YI} = \frac{\rho(\lambda_{-1}) - 2\rho(\lambda_0) + \rho(\lambda_{+1})}{(\lambda_{660} - \lambda_{605})^2}$	Adams et al.(1999)
	Near Infrared Normalized Vegetation Index ($NDVI_{NIR}$)	$NDVI_{NIR} = \frac{\rho_{NIR1} - \rho_{NIR2}}{\rho_{NIR1} + \rho_{NIR2}}$	Tested for the study
	Yellow Normalized Difference Vegetation Index ($NDVI_{605}$)	$NDVI_{605} = \frac{\rho_{NIR} - \rho_{605}}{\rho_{NIR} + \rho_{605}}$	Tested for the study

2.4.3 *Accuracy Assessment*

Classification accuracies were used to assess the reliability of the results; namely, producer, user and overall accuracies. Producer accuracy is derived from calculating the total number of correctly classified cases in one class divided by the total number of cases of that class as indicated by reference data (Congalton, 1991). The user accuracy is derived from calculating the total number of correctly classified cases of one category divided by the total number of cases classified in that category (Story and Congalton, 1986). Finally, overall accuracy is computed by dividing the number of correctly classified cases by the total number of cases in the error matrix. Error matrix tables were computed as part of the accuracy assessment procedure. The procedure of computing error matrices is a very effective way to present accuracy in that accuracies are described along with both errors of inclusion (commission errors) and errors of exclusion (omission errors) present in the classification (Congalton, 1991).

2.4.4 *Comparing Classifier Performance*

In order to compare the performance among the classifiers, a McNemar's test was applied on the results of each classifier against another. The McNemar's test was used to test for the performance of the classifiers since the same samples were used for classification tests, and were therefore not independent as would be required for the Kappa difference test (Foody, 2004). The McNemar's test is preferable since it is a parametric test and very simple to understand. The test is based on a chi-square (χ^2) statistic, computed from two error matrices as follows:

$$\chi^2 = \frac{(f_{12} - f_{21})^2}{f_{12} + f_{21}} \qquad \text{Eq: 1}$$

where f_{12} denotes the number of cases that are wrongly classified by classifier 1 but correctly classified by classifier 2, and f_{21} denotes number of cases that are correctly classified by classifier 1 but wrongly classified by classifier 2. Additional f_{11} and f_{22} were included to indicate the number of cases wrongly classified by both classifiers, and the number of cases correctly classified by both classifiers, respectively.

3. Results

3.1 *Pixel-based classification*

MLC classifier showed the highest overall classification accuracies (86.90%) when compared to SVM (80.18 %) and RF (84.68%) classifiers. MLC also showed the highest average producer and user accuracies for all four classes (86.33% and 87.54%, respectively) when compared to SVM (80.28% and 82.26%, respectively) and slightly higher than RF (85.68% and 80.73%, respectively) (Table 3). The MLC also showed higher overall (86.90%) for the 8-band WorldView-2 imagery when compared to the spectrally resampled 4 band image similar to SPOT, IKONOS and Landsat (74.64%) as seen in table 3. The classification results of MLC (8 bands) show that the highest confusion (4.35%) was found between vegetated gaps and other forest vegetation class (Table 6), while marginal error was found between all other classes (<2.17%). The lowest classification accuracies were obtained using MLC applied on 4 spectral bands resampled from WorldView-2 image.

Table 3. Classification accuracies of three (3) pixel-based methods applied to WorldView-2 image (with 8 bands), and compared to spectrally resampled 4 band image e.g. Landsat.

Classifier	Kappa Coefficient	Mean Producer Accuracy (%)	Mean User Accuracy (%)	Overall Accuracy (%)
Maximum Likelihood (8 bands)	0.82	86.33	87.54	86.90
Maximum Likelihood (4 bands)	0.63	70.53	70.83	74.64
Random Forests	0.78	85.68	80.73	84.68
Support Vector Machines	0.72	80.28	82.26	80.18

3.2 *Object-based classification*

The results of vegetation indices were divided into two groups: (i) vegetation indices that can be derived from conventional sensors (typical Red, Green, Blue and Near-Infrared bands), and (ii) vegetation indices that can be derived from the 8-band WorldView-2 imagery. Amongst the vegetation indices derived from conventional sensors, the enhanced vegetation index (EVI) computed from conventional R.G.B and NIR bands yielded the highest average producer (85.07%), average user (79.73%) and overall classification accuracies (85.59%) on all forest classes. On the contrary, the atmospherically resistant vegetation index (ARVI) yielded the lowest average user and overall classification accuracies (60.42% and 67.57%, respectively).

Amongst the indices derived from WorldView-2 bands, the modified plant senescence reflectance index (mPSRI) showed the highest average producer (92.10%) and average user accuracy (93.50%) for all classes and overall classification accuracy (93.69%). The difference between high performing vegetation index from conventional bands (overall classification accuracy of 85.59 %) and that from WorldView-2 bands (overall classification accuracy of 93.69%) is 8.1%. Table 4 shows the comparison between object-based classifiers and pixel-based methods.

Table 4. Comparison of the best classifiers in per-pixel based group and object-based group, comprising of conventional bands and new WorldView-2 bands respectively. PA = producer accuracy, UA = user accuracy

Class name	Pixel-based classification MLC (4 bands)		Pixel-based classification MLC (8 bands)		Object-based classification (EVI)		Object-based classification (mPSRI)	
	PA (%)	UA (%)	PA (%)	UA (%)	PA (%)	UA (%)	PA (%)	UA (%)
Bare gaps	83.5	73.6	95.4	97.4	79.4	100	96.2	96.3
Vegetated gaps	73.2	61.3	84.1	90.4	78.5	57.9	84.2	84.2
Shadow gaps	67.1	62.6	84.6	81.2	92.3	63.2	90.4	100
Others	71.2	85.6	81.0	81.0	90.0	97.8	97.7	93.5
Overall accuracy (%)	74.64		86.90		85.59		93.69	

Table 5. Comparison of the classifier performance for both pixel-based and object-based classifiers using McNemar's test

Models Compared	f_{11}	f_{12}	f_{21}	f_{22}	Total	chi-sq. ($\chi 2$)	p value	df
mPSRI vs. MLC (8 bands)	6	0	9	96	111	9.00	<0.05	1
mPSRI vs. MLC (4 bands)	5	1	28	77	111	25.00	<0.05	1
mPSRI vs. EVI	5	1	16	89	111	13.23	<0.05	1
mPSRI vs. RF	4	1	13	93	111	10.29	<0.05	1
mPSRI vs. SVM	4	0	19	88	111	19.00	<0.05	1
MLC (8 bands) vs. MLC (4 bands)	14	1	19	77	111	16.20	<0.05	1
MLC (8 bands) EVI	10	5	10	86	111	1.66*	<0.05	1
MLC(8 bands) vs. RF	8	7	9	87	111	0.25*	<0.05	1
MLC(8 bands) vs. SVM	13	2	10	86	111	5.33	<0.05	1
EVI vs.RF	7	14	10	80	111	0.67*	<0.05	1
EVI vs. SVM	11	6	15	79	111	3.85	<0.05	1
EVI vs. MLC(4 bands)	13	8	20	70	111	5.14	<0.05	1
RF vs SVM	6	11	15	79	111	0.62*	<0.05	1
RF vs MLC(4 bands)	12	4	23	72	111	13.37	<0.05	1

Figure 7 shows the results of the delineated gaps and the confusion matrix is shown by Table 6. The mPSRI showed average producer accuracy of 90.25 % and average user accuracy of 90.26 % for forest gaps.

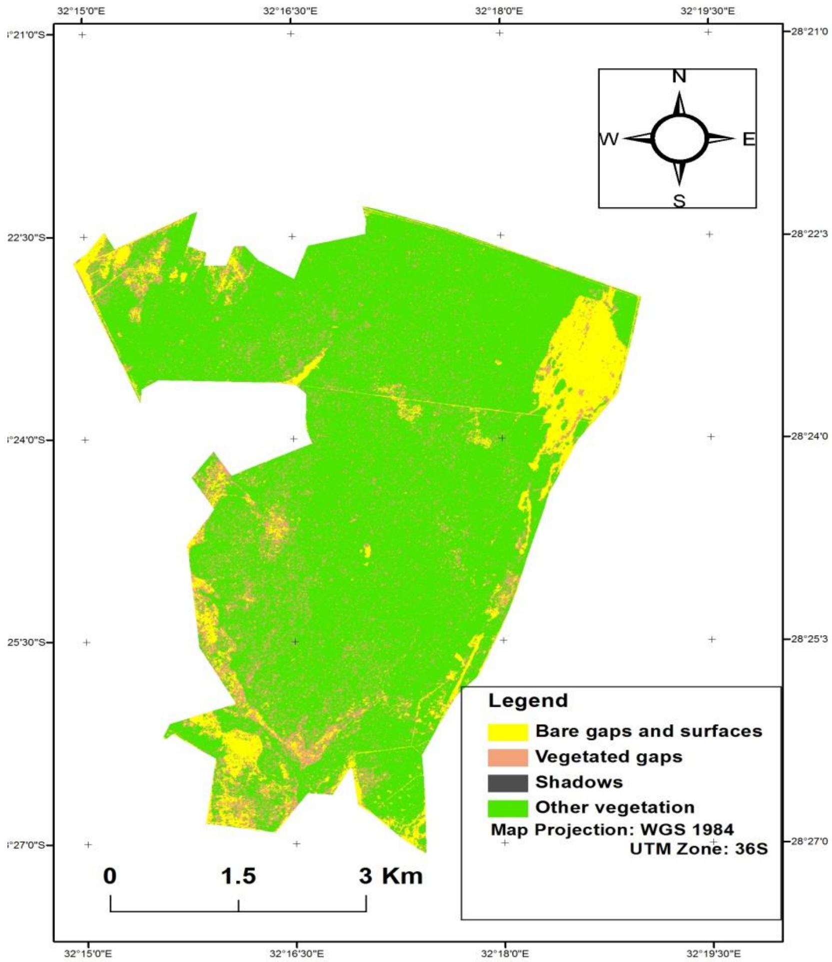

Figure 7: Delineated forest gaps resulting from mPSRI in object-based classification

Table 5 shows McNemar's test results with the number of cases correctly and incorrectly classified by pixel based methods (8 band WorldView-2 image and conventional 4 band image), and object-based methods (Enhanced Vegetation Index and modified Plant Senescence Reflectance Index). The results showed that there was a statistical difference between pixel-based and object-based classification techniques. Most of the comparisons showed statistical difference amongst each other at $\rho < 0.05$ and at 1 degree of freedom. There is no significant difference between MLC (8 bands) vs EVI, MLC (8 bands) vs RF, EVI vs RF, and RF vs SVM. Table 7 shows the results of three (3) best vegetation indices derived from (i) conventional R.G.B and NIR bands common in Landsat, and (ii) WorldView-2 sensor. This table indicates that higher overall accuracies were obtained from mPSRI (93.69%) derived from new WorldView-2 bands than the conventional EVI (85.59%). The EVI yielded the highest producer accuracy (85.05%) in the list of conventional vegetation indices, while the mPSRI yielded both the highest producer accuracy (92.18%) and user accuracy (93.50%) in the list of new WorldView-2 indices.

Table 6: Confusion matrix resulting from mPSR Index in object-based image analysis

Classified Image	Class	Bare gaps	Vegetated gaps	Shadows	Others	Total	*User Accuracy*
		Reference Image					
	Bare gaps	26	1	0	0	27	96.30
	Vegetated gaps	1	16	1	1	19	84.21
	Shadows	0	0	19	0	19	100.00
	Others	0	2	1	43	46	93.48
	Total	27	19	21	44	111	
	Producer Accuracy%	96.29	84.21	90.48	97.73		
	Kappa Index					0.91	
	Overall Accuracy					93.69	

Table 7: Comparison of classification results obtained from vegetation indices derived from conventional sensor and those from WorldView-2 imagery. (PA= mean producer accuracy, UA= mean user accuracy, OA= overall accuracy)

	Index	PA (%)	UA (%)	OA (%)
Conventional R.G.B, NIR Indices	EVI	85.05	79.73	85.59
	$NDVI_{545}$	83.50	80.70	84.69
	$NDVI_{660}$	75.13	73.85	82.93
New WorldView-2 Indices	$NDVI_{725}$	75.42	80.79	82.90
	mPSRI	92.18	93.50	93.69
	NPCI	73.15	74.35	78.37

4. Discussion and Conclusions

Canopy gaps form an important part of forests and have been mapped using different methods (Runkle, 1982: Emborg, 1998; Vepakomma *et al*., 2008) but rarely in subtropical forests (Brokaw, 1985). Most of the successful studies focused on delineating forest gaps from combined optical and hyperspectral data (Hodgson and Bresnahan, 2004). To the best of our knowledge, high resolution multispectral data alone has not been used for delineating forest canopy gaps in closed canopy environment. Findings from this study highlight the possibility of using high spatial resolution WorldView-2 imagery for delineating forest canopy gaps in the subtropical forest environment. The suitability of 8 band WorldView-2 imagery to delineate canopy gaps was assessed and compared with resampled conventional 4 bands (visible-near Infrared) common in Landsat imagery. Higher classification accuracies were achieved from an 8-

band WorldView-2 image when compared to the conventional 4 band imagery (red, green, blue and near infrared bands) similar to those found in SPOT, IKONOS and Landsat. In addition, the best three highest performing indices derived from WorldView-2 imagery (NPCI, mPSRI and $NDVI_{725}$) yielded the highest average user accuracy (82.83%) for all forest classes (Table 7) compared to the three best performing indices derived from conventional sensors (EVI, $NDVI_{545}$ and $NDVI_{660}$) (78.07%) in object-based image analysis. These findings therefore support the assertion that the utility of WorldView-2 sensor provides improved estimations of vegetation biophysical characteristics in subtropical environments such as biomass and tree species discrimination (Mutanga *et al*., 2012; Cho *et al*., 2012). The improved results utilizing the red edge band (centered at 725 nm) in WorldView-2 sensor might be attributed to the fact that the reflectance in this shows less saturation in dense vegetation when compared to the red band (660-680 nm) reflectance that is common in conventional sensors such as Landsat and IKONOS (Mutanga and Skidmore, 2004).

The maximum likelihood classifier applied on a spectrally resampled 4-band imagery, common in SPOT, IKONOS and Landsat, showed a drop in average user and overall classification accuracies (from 86.33% to 70.83%, and from 86.90% to 74.64% respectively), which is an indication that new WorldView-2 bands provide spectral enhancements to the common RGB and NIR bands (Mutanga *et al*., 2012; Cho *et al*., 2012; Ozdemir and Karnieli, 2011). The pixel-based overall classification accuracy resulting from WorldView-2 bands is 15.50% higher than the overall classification accuracy derived from conventional RGB and NIR bands. This evidence also highlights the spectral saturation that poses a major challenge when using conventional bands (Cho *et al*., 2008) and shows that the presence of new WorldView-2 bands can minimize this problem. This also signifies the spectral enhancement provided by additional bands such as coastal, yellow and NIR-2. On the other hand, there was no statistical difference between the observed results from the classification by RF (commonly used algorithm) and MLC, due to the conditions under which MLC performs. The MLC requires that the cells in each class in the multidimensional space be normally distributed so as to allow the decision based on the Bayesian theorem for classification (Jeon and Landgrebe, 1999).

The mPSRI which is derived from new bands of WorldView-2 yielded the highest overall classification accuracy than all the selected indices. The mPSRI index is derived from a plant senescence reflectance index (PSRI) proposed by Merzlyak *et al*. (1999), where the red edge band was used instead of a red band. Although this index was initially proposed for estimating stages of leaf senescence and fruit ripening, our study indicated that it can also be used to delineate forest gaps in closed canopy forest. Additionally, we have observed the increased average user and overall classification accuracies of red edge NDVI over the conventional NDVI. The saturation problem that is prevalent in conventional sensors is minimized when the red edge band is used in vegetation indices such as NDVI, and this characteristic was crucial for our study since the confusion between vegetated gaps and tree crown was minimized (Mutanga and Skidmore, 2004). This confirms our hypothesis that separability of forest gaps from forest tree crowns can be increased by using indices that are derived from WorldView-2 than those derived from conventional sensors.

Although the results obtained from high resolution WorldView-2 data offer a promising hope to the delineation of forest canopy gaps in tropical indigenous forest, they cannot, however be comparable to those obtained from LiDAR technology. This is primarily so due to the fact that

LiDAR data also addresses the tree height characteristic in which optical sensors fail to capture (Dubuyah and Drake, 2000; Harding *et al*., 2001; Nelson *et al*., 1997). However, although LiDAR technology provides very accurate measurements the application of this technology is limited by its high data acquisition costs and high data dimensionality (Mutanga *et al*., 2012).

Based on the results, we conclude that the use of 8-band WorldView-2 imagery increases classification accuracies (average producer, user and overall) for delineating forest canopy gaps when compared to the conventional VNIR bands present in SPOT, IKONOS and Landsat. We also conclude that vegetation indices derived from new WorldView-2 red-edge band (NDVI725 and mPSRI) yielded higher average user accuracy than those that are derived from conventional sensors, highlighting the significance of new WorldView-2 bands.

Acknowledgments

We wish to thank the Council for Scientific and Industrial Research's Natural Resources and Environment (CSIR-NRE) unit, of South Africa, for its financial assistance through its research grants. Many thanks also to the Department of Agriculture, Forestry and Fisheries for the support on the study area. Authors wish to thank Sibuyiselo Gumede of Khula village who assisted in field data collection. To Tino, a dearest son.

References

Adams, M. L., Philpot, W. D. and Norvell, W. A., 1999, Yellowness index: An application of spectral second derivatives to estimate chlorosis of leaves in stressed vegetation. *International Journal of Remote Sensing 20***,** pp. 3663 – 3675.

Anderson, D.R, Laake, J.L., Crain, B.R. and Buraham, K.V., 1979, Guidelines for line transect sampling of biological populations. *Journal of Wildlife Management* 43, pp. 70 –78.

Asner, G. P., Keller, M, Pereira, R., Zweede, J. C. and Silva, J. N. M., 2004, Canopy damage and recovery after selective logging in Amazonia: field and satellite studies *Ecological Applications* 14, Suppl. 280 – 98.

Asner, G.P., Kellner, J.R., Kennedy-Bowdoin, T.Y., Knapp, D.E., Anderson, C., and Martin, R.E., 2013, Forest Canopy Gap Distribution in the Southern Peruvian Amazon. PLoS ONE 8, e60875.

Battles, J.J., Dushoff, J.G. and Fahey, T.J., 1996, Line intersects sampling of forest canopy gaps. *Journal of Forest Science* 42, pp. 131 – 138.

Betts, H.D., Brown, L.J. and Stewart, G.H., 2005, Forest canopy gap detection and characterization by the use of high-resolution Digital Elevation Models. *The New Zealand Journal of Ecology* 29, pp. 95 – 103.

Blaschke, T. and Strobl, J., 2001, Whats wrong with pixels? Some recent developments interfacing remote sensing and GIS. *GeoBIT/GIS* 14, pp. 12 – 17.

Brokaw, N. and Grear, J., 1991, Forest structure before and after hurricane Hugo at three elevations in Liquillo Mountains, Puerto Rico. *Biotropica* 23, pp. 386 – 392 .

Brokaw, N.V.L., 1982, The definition of tree fall gap and its effect on measure of forest dynamics. *Biotropica* 11, pp. 158 – 160 .

Buckland, S. T., Anderson, D. R., Burnham, K. P., Laake, J. L., Borchers, D. L., and Thomas, L., 2001, Introduction to Distance Sampling, Oxford: Oxford University Press.

Chen, Q., 2009. Improvement of the Edge-based Morphological (EM) method for lidar data filtering. *International Journal of Remote Sensing* 30, pp. 1069 – 1074.

Cho, M.A., Mathieu, R., Wessels, K., Naidoo, L., Asner, G.P., Van Aardt, J., Ramoelo, A., Debba, P., Main, R., Smit, I.P., and Erasmus, B., 2012, Mapping tree species composition in South African savannas using an integrated airborne spectral and LiDAR system. *Remote Sensing of Environment* 125, pp. 514 – 526.

Cho, M.A., Skidmore, A.K. and Atzberger, C., 2008, Towards red-edge positions less sensitive to canopy biophysical parameters for leaf chlorophyll estimation using properties optique spectrales des feuilles (PROSPECT) and scattering by arbitrarily inclined leaves (SAILH) simulated data. *International Journal of Remote Sensing* 29, pp. 2241 – 2255.

Clark, M. L., Clark, D.B. and Roberts, D.A., 2004, Small-footprint lidar estimation of sub-canopy elevation and tree height in a tropical rain forest landscape. *Remote Sensing of Environment* 91, pp. 68 – 89.

Codilla, L.T., and Metillo, E.B., 2011, Distribution and Abundance of the Invasive Plant Species *Chromolaena odorata* L. in the Zamboanga Peninsula, Philippines. *International Journal of Environmental Science and Development* 2, pp. 406 – 410.

Dawson, T.P., and Curran, P.J., 1998, A new technique for interpolating the reflectance red edge position. *International Journal of Remote Sensing* 19, pp. 2133 – 2139.

Dubayah, R.O., and Drake, J.B., 2000, Lidar remote sensing for forestry. *Journal of Forestry* 98, pp. 44 – 46

Eberhardt, L.L., 1986, A preliminary appraisal of line transects. *The journal of Wildlife Management* 32, pp. 82 – 88.

Emborg, J., 1998, Understorey light conditions and regeneration with respect to the structural dynamics of a near-natural temperate deciduous forest in Denmark. *Forest Ecology and Management* 106, pp. 83 – 95.

Exelis Visual Information Solutions, Boulder, Colorado. *Environment for Visualizing Images* (Used in 2012)

Forbes, S.A. and Gross, A.O., 1921, The orchard birds of an Illinois summer. III. *Nature, History and Survey Bulletin* 14, pp. 1 – 8.

Fourcade, H.G., 1889, Report on the Natal forests, 197 pp. W. Watson, Pietermaritzburg.

Fox, T.J., Knutson, M.G. and Hines, R.K., 2000, Mapping forest canopy gaps using air-photo interpretation and ground surveys. *Wildlife Society Bulletin* 28, pp. 882 – 889.

Geldenhuys, C. J., 1989, Biogeography of the mixed evergreen forests of southern Africa. Ecosystems Programmes Occasional Report no. 45. FRD, Pretoria.

Gitelson, A.A. and Merzlyak, M.N., 1994, Spectral Reflectance Changes Associated with Autumn Senescence of Aesculus Hippocastanum L. and Acer Platanoides L. Leaves. Spectral Features and Relation to Chlorophyll Estimation. *Journal of Plant Physiology* 143, pp. 286 – 292.

Goodall, J.M. and Erasmus, D.J., 1996, Review of the status and integrated control of the invasive weed, *Chromolaena odorata,* in South Africa. *Agriculture, Ecosystem and Environment* 56, pp. 151 – 164.

Harding, D.J.,. Lefsky, M.A., Parker, G.G., and Blair, J.B., 2001, Laser altimeter canopy height profiles: Methods and validation for closed-canopy, broadleaf forests. *Remote Sensing of Environment* 76, pp. 283 – 297

Hodgson, M. E. and Bresnahan, P., 2004, Accuracy of airborne Lidar-derived elevation: empirical assessment and error budget. *Photogrammetric Engineering and Remote Sensing* 70, pp. 331 – 339.

Huete, A.R., H. Liu, Q. Batchily, K. and Van Leeuwen, W.J., 1997, A Comparison of Vegetation Indices over a Global Set of TM Images for EOS-MODIS. *Remote Sensing of Environment* 59, pp. 440 – 451.

Jackson, R. D., Slater, P. N. and Pinter, P. J., 1983, Discrimination of growth and water stress in wheat by various vegetation indices through clear and turbid atmospheres. *Remote Sensing of Environment* 13, pp. 187 – 208.

Jeon, B. and Landgrebe, D.A., 1999, Partially supervised classificatoin using weighted unsupervised clustering, *IEEE Transsctions in Geoscience and Remote Sensing* 37, pp. 1073 – 1079.

Kaufman, Y. J. and Tanre, D., 1992, Atmospherically resistant vegetation index (ARVI) for EOS-MODIS. *IEEE Transactions on Geoscience and Remote Sensing* 30**,** pp. 261 – 270.

Kellner J.R, Clark, D.B, and Hubbell, S.P., 2009, Pervasive canopy dynamics produce short-term stability in a tropical rain forest landscape. *Ecology Letters* 12, pp. 155 – 164

Kim, S., Schreuder, G.F., Briggs, D., and Hinckley, T., 2009, Individual tree species identification using LiDAR-derived crown structures and intensity data. Doctoral thesis. University of Washington, College of Forest Resources, pp. 1 – 137.

Knipling, E.B., 1970, Physical and physiological basis for the reflectance of visible and near-infrared radiation from vegetation. *Remote Sensing of Environment* 1, pp. 155 – 159..

Kupfer, J.A. and Runkle, J.R., 1996, Early gap successional pathways in a *Fagus-Acer* forest preserve: patterns and determinants. *Journal of Vegetation Science* 7, pp.247 – 256.

Mallinis, G., Koutsias, N., Tsakiri-Strati, M. and Karteris, M., 2008. Object-based classification using Quickbird imagery for delineating forest vegetation polygons in a Mediterranean test site. *ISPRS Journal of Photogrammetry & Remote Sensing* 63, pp. 237 – 250.

Merzlyak, J.R., Gitelson, A.A, Chivkunova, O.B. and Rakitin, V.Y., 1999, Non-destructive Optical Detection of Pigment Changes During Leaf Senescence and Fruit Ripening. *Physiologia Plantarum* 106, pp. 135 – 141.

Moore, A.B., 2004, Alien Invasive species: Impacts on Forests and Forestry. – Forest Health and Biosecurity Working Papers FAO, pp. 66. Rome, Italy.

Mutanga, O. and Skidmore, A.K., 2004, Narrow band vegetation indices solve the saturation problem in biomass estimation. *International Journal of Remote Sensing* 25, pp. 3999 – 4014.

Mutanga, O., Adam, E. and Cho, M.A., 2012, High density biomass estimation for wetland vegetation using WorldView-2 imagery and random forest regression algorithm. *International Journal of Applied Earth Observation and Geoinformation* 18, pp. 399 – 406.

Negrón-Juárez, R., I., Chambers, J.Q. and Marra, D.M., 2011, Detection of subpixel treefall gaps with Landsat imagery in Central Amazon. *Remote Sensing of Environment* 115, pp. 3322 – 3328.

Nelson, R., Oderwald, R., and Gregoire, T.G., 1997, Separating the ground and airborne laser sampling phases to estimate tropical forest basal area, volume, and biomass. *Remote Sensing of Environment* 60, pp. 311 – 326.

Orians, G. H., 1982, The influence of tree-falls in tropical forests in tree species richness. *Tropical Ecology* 23, pp. 255–279.

Ozdemir, I. and Karnieli, A., 2011, Predicting forest structural parameters using the image texture derived from WorldView-2 multispectral imagery in a dryland forest, Israel. *International Journal of Applied Earth Observation and Geoinformation* 13, pp. 701–710.

Peñuelas, J., Baret, F. and Filella, I., 1995, Semi-empirical indices to assess carotenoids/chlorophyll a ratio from leaf spectral reflectance. *Photosynthetica* 31**,** pp. 221 - 230

Perumal, K. and Bhaskaran, R., 2010, Supervised classification performance of multispectral images. *Journal of Computing* 2. ISSN 2151-9617.

Runkle, J.R., 1982, Patterns of disturbance in some old-growth mesic forests of Eastern North-America. *Ecology* 63, pp. 1533–1546.

San, B.T., and Suzen, M.L., 2010, Evaluation of different atmospheric correction algorithms for EO-1 Hyperion imagery. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science* 38, Kyoto Japan.Sellers, P.J., 1985, Canopy reflectance, photosynthesis and transpiration. *International Journal of Remote Sensing* 6, pp. 1335–1372.

Smith, K.L., Steven, M.D. and Colls, J.J., 2004, Use of hyperspectral derivative ratios in the red edge region to identify plant stress responses to gas leaks. *Remote Sensing of Environment* 92, pp. 207 – 217.

Suarez, A.V., Pfennig, K.S. and Robinson, S.K., 1997, Nesting success of a disturbance-dependent sonbird on different kinds of edges. *Conservation Biology* 11, pp. 928 – 935.

Touw, W.G., Bayjanov, J.R., Overmars, L., Backus, L., Boekhorst, J., Wels, M., and Van Hijum, S.A.F.T., 2012, Data mining in the life sciences with random Forest: a walk in the park or lost in the jungle? *Briefing in Bioinformatics* pp. 1 – 12.

Treitz, P., Lim, K., Wulder, M., Stonge, B. and Flood, M., 2003, LiDAR remote sensing of forest structure. *Progress in Physical Geography* 27, pp. 88 – 106..

Ustin, S.L. and Trabucco, A., 2000, Using Hyperspectral Data to Assess Forest Structure. *Journal of Forestry* 98, pp. 47 – 49.

Van Gils, H., Mwanangi, M. and Rugege, D., 2006, Invasion of an alien shrub across four land management regimes, west of St Lucia, South Africa. *South African Journal of Science* 102, February.

Vapnik, V.N., 1995, The Nature of Statistical Learning Theory (New York: Springer-Verlag)

Vepakomma, U., Onge, B.S. and Kneeshaw, D., 2008, Spatially explicit characterization of boreal forest gap dynamics using multi-temporal lidar data. *Remote Sensing of Environment* 112, pp. 2326–2340.

Weiss, P.W. and Noble, I.R., 1984, Status of coastal dune communities invaded by *Chrysanthemoides monilifera*. *Australian Journal of Ecology* 9, pp. 93–98.

Whitmore, T.C., 1989, Canopy gaps and the two major groups of forest trees. Special feature – Treefall gaps and forest dynamics. *Ecology* 70, pp. 536–538.

Woodcock, C. E., Macomber, S. A., Pax-Lenney, M. and Cohen, W. C., 2001, Monitoring large areas for forest change using Landsat: Generalization across space, time and Landsat sensors. *Remote Sensing of Environment* 78, pp. 194−203.

Zhou, Q. and Robson, M., 2001, Automated rangeland vegetation cover and density estimation using ground digital images and a spectral-contextual classifier. *International Journal of Remote Sensing* 22 pp.3457 – 3470.