Exploring the utility of the additional WorldView-2 bands and support vector machines in mapping land use/land cover in a fragmented ecosystem, South Africa

Galal Omer[1,2], Onisimo Mutanga[1], Elfatih M. Abdel-Rahman[1,2], Elhadi Adam[3]

[1] School of Agricultural, Earth and Environmental Sciences, Pietermaritzburg Campus, University of KwaZulu-Natal, Scottsville P/Bag X01, Pietermaritzburg 3209, South Africa.
[2] University of Khartoum, Khartoum North 13314, Sudan; Email: gekhaled189@gmail.com.
[3] Schools of Geography, Archaeology and Environmental Studies, University of the Witwatersrand – Johannesburg, Wits 2050, Johannesburg, South Africa.

Abstract

Land use/land cover (LULC) classification is a key research field in environmental applications of remote sensing on the earth's surface. The advent of new high resolution multispectral sensors with unique bands has provided an opportunity to map the spatial distribution of detailed LULC classes over a large fragmented area. The objectives of the present study were: (1) to map LULC classes using multispectral WorldView-2 (WV-2) data and SVM in a fragmented ecosystem; and (2) to compare the accuracy of three WV-2 spectral data sets in distinguishing amongst various LULC classes in a fragmented ecosystem. WV-2 image was spectrally resized to its four standard bands (SB: blue, green, red and near infrared-1) and four strategically located bands (AB: coastal blue, yellow, red edge and near infrared-2). WV-2 image (8bands: 8B) together with SB and AB subsets were used to classify LULC using support vector machines. Overall classification accuracies of 78.0% (total disagreement = 22.0%) for 8B, 51.0% (total disagreement = 49.0%) for SB, and 64.0% (total disagreement = 36.0%) for AB were achieved. There were significant differences between the performance of all WV-2 subset pair comparisons (8B versus SB, 8B versus AB and SB versus AB) as demonstrated by the results of McNemar's test (Z score ≥1.96). This study concludes that WV-2 multispectral data and the SVM classifier have the potential to map LULC classes in a fragmented ecosystem. The study also offers relatively accurate information that is important for the indigenous forest managers in KwaZulu-Natal, South Africa for making informed decisions regarding conservation and management of LULC patterns.

Keywords: land use/cover classification, fragmented ecosystem, WorldView-2, support vector machines

1. Introduction

Land use/land cover (LULC) is a fundamental variable that influences and links with many parts of human and physical systems and is a vital data component for many aspects of environmental change (Foody, 2002, Otukei and Blaschke, 2010). The changes in LULC have significant effects on basic ecosystem processes including biogeochemical cycling and land degradation (Penner, 1994, Foley *et al.*, 2005, Otukei and Blaschke, 2010). Similarly, the LULC classification plays an important role in environmental monitoring, and management (Otukei and Blaschke, 2010). Despite this important role, LULC mapping is still facing a complex challenge in nexus to ambiguous classes used (Cingolani *et al.*, 2004, Otukei and Blaschke, 2010). Additionally, fragmented ecosystems in many parts of Africa are characterized by the removal and clearing of the forest for pasture, agriculture and settlements leading to vegetation species loss (van Wyk *et al.*, 1996, Cho *et al.*, 2013). In most cases, indigenous forests are fragmented into patches of various sizes and shapes surrounded by a matrix of different LULC classes (Benitez-Malvido, 1998, Cho *et al.*, 2013). In this context, information relating to the dynamics, distribution and productivity of LULC is not only beneficial to the source of economic security (Eldeen, 2005, van Wyk, 2008), but is also needed for fragmented ecosystems inventory, management and monitoring (Cingolani *et al.*, 2004, Pignatti *et al.*, 2009, Cho *et al.*, 2013). In order to meet the management and monitoring requirements of fragmented ecosystems, more specific information from LULC surveys and inventories is needed. However, it is quite difficult and challenging to produce LULC maps using traditional field survey approaches.

Traditional approaches are complex and they require intensive fieldwork which is a costly and time-consuming process that often lacks the necessary geometric accuracy, particularly in highly heterogeneous and fragmented ecosystems. In this regard, remote sensing is a particularly useful tool, as it has successfully been used for tree species and land cover classification (Clark *et al.*, 2005, Larsen, 2007). Moreover, multispectral sensors such as Landsat and SPOT cover large areas of the earth's surface at repeated time intervals, making remote sensing a perfect alternative to traditional approaches for LULC. Recently, the developments of high spatial resolution multispectral sensors such as IKONOS have brought unique opportunities for classifying and monitoring LULC (Pu and Landry, 2012). Multispectral data either have high spatial resolution but offer only a few bands like blue, green, red, and near infrared (NIR), or they offer relatively more bands but with lower spatial resolution. The low spatial resolution multispectral sensors might not accurately map LULC classes in a heterogeneous and fragmented

ecosystem (Foody, 2002, Cho *et al.*, 2012). Multiple objects within a pixel in such a case can lead to spectral confusion and poor distinction amongst discrete and continuous cover types, resulting in ambiguous LULC classes (Cingolani *et al.*, 2004). On the other hand, multispectral data of very fine spatial resolution could not capture the intra-class variability accurately when course LULC classes are mapped. As long as the spectral resolution is concerned, the advent of hyperspectral data can overcome the limitations of multispectral data by providing spectral data of many and contiguous wavebands (Vaiphasa *et al.*, 2007) for more accurate and reliable LULC maps (Pal, 2006, Petropoulos *et al.*, 2012). However, the use of hyperspectral data has its own limitations in terms of cost, availability, processing, and high dimensionality (Vaiphasa *et al.*, 2007, Dalponte *et al.*, 2009).

Recently, high spatial resolution multispectral sensors such as RapidEye, WorldView-2 (WV-2) and Sentinel-2 were designed with a relatively fewer strategically located bands (additional bands) to overcome the limitations of their spectral capabilities over other high spatial resolution multispectral sensors of conventional bands (standard bands) such as QuickBird. The potential of WV-2 data, for instance, has been demonstrated in a number of diverse studies that include among others, predicting and mapping forest structural parameters (Ozdemir and Karnieli, 2011), urban land cover mapping (Zhou *et al.*, 2012), and discriminating commercial forest species (Peerbhay *et al.*, 2014). These studies discussed the utility of the eight available spectral bands of WV-2 imagery on mapping and predicting a feature of interest and concluded that WV-2 data have considerably improved the classification and prediction accuracies compared to conventional sensors. For instance, Pu and Landry (2012) explored the use of WV-2 and IKONOS data sets on mapping tree species and found that the bands available in WV-2 significantly increased the classification accuracy compared to the bands available in IKONOS sensor. These studies have limitations related to paucity of knowledge on the performance of WV-2 spectral subsets in vegetation and LULC types with advanced classification methods.

Various classification methods have been implemented in order to map vegetation species and LULC classes using WV-2 data. These methods include discriminant analysis (Pu and Landry, 2012), decision trees (Heumann, 2011), maximum likelihood and minimum distance to the mean classifiers (Cho *et al.*, 2011, McCarthy and Halls, 2014). All these classifiers have used supervised classification methods with conventional multispectral data. Amongst these classifiers, the maximum likelihood and minimum distance to the mean have been the most widely used classifiers (Kavzoglu and Mather, 2003, Otukei and Blaschke, 2010). The two classifiers have the ability to generate acceptable accuracy, simplicity and availability in most image processing packages (Zhang *et al.*, 2007, Cho *et al.*, 2011). However, all these classifiers

have their own limitations that are related particularly to distributional assumptions and to mapping areas with limited training samples (Kavzoglu and Mather, 2003, Cho *et al.*, 2012).

To tackle these problems, powerful classification methods are essentially used for mapping LULC (Lu and Weng, 2007). These classification methods include, for example, neural networks, random forest, support vector machines (SVM) and decision trees (Civco, 1993, Heumann, 2011, Pal, 2003, Pal, 2006). Amongst these methods, attention has been paid to the use of SVM due to its superior image handling capability (Vapnik, 1998). Numerous studies have used SVM classifier with multispectral imagery for LULC mapping (Kavzoglu and Colkesen, 2009, Otukei and Blaschke, 2010, Petropoulos *et al.*, 2012, Adam *et al.*, 2014).These researchers have found relatively better or similar performances obtained by this classifier as compared to other classifiers when multispectral and hyperspectral data were used. The exploration of the utility of WV-2 strategically located additional bands for improving the accuracy of LULC maps in a fragmented ecosystem is needed. To the best of our knowledge, there is a lack of information on how SVM classifier can perform on delineating LULC class in a fragmented ecosystem using different WV-2 spectral subsets. Therefore, the objectives of the present study were; (1) to map different LULC classes using WV-2 data and SVM classifier in a fragmented ecosystem; and (2) to compare the accuracy of three WV-2 spectral subsets in distinguishing amongst various LULC classes in a very fragmented ecosystem.

2. Methodology

2.1 Study area

The study area is located in Dukuduku forest, in the northern part of KwaZulu-Natal at the entrance to the iSimangaliso Wetland Park, South Africa (28º25'S, 32º17'E) (Figure 1). The area covers approximately 19887 hectares. The subtropical climate dominating the study area has warm moist summers and mild dry winters. The mean daily maximum temperatures are 26°C in January and 21°C in July, while mean daily minimum temperatures are 19°C in January and 9°C in July (Von Maltitz *et al*., 2003). The rainy season falls between November and March with a mean annual rainfall of 1250 mm (Von Maltitz *et al*., 2003). The Dukuduku forest is undergoing fragmentation as a result of indigenous forest clearance for agriculture, plantation forestry, and settlements (Cho *et al.*, 2012, Cho *et al.*, 2013). Some other anthropogenic activities in the area have also led to a fragmented ecosystem. The area is covered by various natural indigenous vegetation species with different age groups and other forms of LULC classes. These include sugarcane farms and commercial plantation forests occurring on the artificially drained

floodplain to the south of the area and grassland to the north (van Wyk *et al.*, 1996, Cho *et al.*, 2012). Some of the natural vegetation has been removed and cleared for other land uses (van Wyk *et al.*, 1996, Cho *et al.*, 2012, Cho *et al.*, 2013). Therefore, the study area comprises a wide range of land use and fragmented forest classes and thus is appropriate to meet the objectives of this study.

Figure 1. Location of study area in KwaZulu-Natal Province (KZN) of South Africa

2.2 Image acquisition and pre-processing

A cloud-free WV-2 multispectral image covering the study area was acquired on the 1st December 2013. WV-2 image consists of eight multispectral bands in the 400-1040 nm spectral range with a spatial resolution of 2 m and swath width of 16.4 km at nadir. The spectral bands of WV-2 are coastal blue (400–450 nm), blue (450–510 nm), green (510–580 nm), yellow (585–625 nm), red (630–690 nm), red edge (705–745 nm), NIR-1 (770–895 nm), and NIR-2 (860–1040 nm). The image was atmospherically corrected and transformed to canopy reflectance using the Quick Atmospheric Correction (QUAC) extension in Environment for Visualizing Images (ENVI 4.7) software (ENVI, 2006). The image was then referenced to the Universal Transverse Mercator (UTM zone 36 South) projection using WGS-84 Geodetic datum. The acquired image was geometrically corrected by DigitalGlobe. After the geometric and atmospheric correction, the WV-2 image was spectrally resized to four standard bands (blue, green, red and NIR 1: SB) and four strategically located new bands (coastal blue, yellow, red edge and NIR 2: AB). These subsets together with all eight bands (8B) of WV-2 were compared for mapping LULC classes using SVM supervised classifier.

2.3 Field data collection

The field campaign was carried out on 7th December 2013, within a week of the WV-2 imagery acquisition. This was done in order to collect ground reference data of eight LULC classes, namely: dune forest (DF), indigenous forest (IF), fragmented forest (FF), *Eucalyptus spp* (EP), *Pinus spp* (PN), mature sugarcane (MS), young sugarcane (YS), and grassland (GL) using a handheld Leica GS20 GPS (geographical positioning system) with sub-meter accuracy. During the field visit, a total of 75 sample data points were collected for each class. The ground reference data were collected using random sampling protocol to adequately sample LULC classes based on their representative sizes within the study area. The reference data were then divided randomly into training (70%) and test (30%) dataset using Hawth's Analysis tool in ArcGIS 9.3. The SVM classifier was trained on 70% of a randomly selected holdout sample and final accuracy was assessed using the remaining 30% samples.

2.4 Statistical analysis

The effectiveness of SVM classifier to map LULC classes was investigated in this study. The classifier was trained on 70% (n = 53) of a randomly selected holdout sample and final accuracy assessments were evaluated using the remaining 30% (n = 22) of the dataset. When the

training positions and classes were allocated, classification signatures were created for the eight LULC classes in the study area. After assessing and adjusting the signatures, SVM supervised classification method was then employed to classify the WV-2 image. SVM parameters were optimized and then input into the ENVI software to map the classes on WV-2 image. The e1071 library version 2.15.2 in R statistical packages (R Development Core, 2012) was employed for SVM parameters optimization.

2.4.1 Support vector machines (SVM) classifier

SVM was originally introduced by Cortes and Vapnik (1995) as a binary classifier. However, real remote sensing problems usually include identification of multiple classes. Amendments are made to the simple SVM binary classifier to run as a multi-class classifier (Burges, 1998). SVM is a distribution-free algorithm that requires few training data points without encounter any over-fitting problems due to the use of kernel functions (Cortes and Vapnik, 1995). In the present study, WV-2 subsets (8B, SB and AB) were used for defining the feature space of SVM. A radial basis function (RBF) was utilized to apply the kernel method of SVM. Interested readers are referred to among others, Cortes and Vapnik, (1995), Mountrakis *et al.* (2011) and Abdel-Rahman *et al*. (2014) for detailed information about the concept, theoratical and mathematical formulation of SVM algorithm.

2.4.2 Accuracy assessment

A confusion matrix was constructed to compare the true class with the class assigned by SVM and to calculate the overall accuracy (OA), producer's accuracy (PA) and user's accuracy (UA). OA has the advantage of being directly interpretable as the percentage (%) between the number of correctly classified samples and the number of test samples. PA refers to the probability of a certain class being correctly classified, while UA represents the likelihood that a sample belongs to specific class and the classifier accurately assigns it such a class. In addition, two parameters were calculated from the cross-tabulation matrix to evaluate the reliability of SVM classifier. These include quantity disagreement (QD) and allocation disagreement (AD) which were developed by Pontius and Millones (2011). The quantity disagreement is the amount of the contrast between the number of test data and predicted data, while the allocation disagreement describes the number of expected classes that have less than optimal spatial location in comparison to the test data. Depending on the accuracy metrics achieved for each WV-2 data sets in each accuracy assessment method, a statistical analysis can be performed to test if there was any significant difference between the classification results of three WV-2

spectral subsets. Hence, McNemar's test was performed to test whether there were any significant differences amongst the confusion matrices of the three WV-2 spectral dataset. McNemar's test is a nonparametric test based upon standardized normal test statistic calculated from error matrices of SVM classifier given as follows (Foody, 2004, Leeuw *et al.*, 2006):

$$Z = \frac{f12 - f21}{\sqrt{f12 + f21}}$$

[1]

where $f12$ denotes the number of samples that are misclassified on the first confusion matrix but correctly classified on the second confusion matrix. $f21$ denotes the number of samples that are misclassified on the second confusion matrix but correctly classified on the first confusion matrix. A difference in accuracy between the confusion matrices of different WV-2 spectral subsets is statistically significant ($p \leq 0.05$) if Z value is more than 1.96 (Foody, 2004, Leeuw *et al.*, 2006).

3. Results

3.1 Optimization of support vector machines

The results of grid search and 10-fold cross validation method indicated optimal values of λ and *C* for SVM, respectively, of 0.1 and 10 for 8B, 0.1 and 1000 for SB, and 0.1 and 100 for AB. When these optimal values were input into SVM classifier, minimum classification error of 32.0%, 35.3%, and 36.3% for 8B, SB and AB, respectively were obtained. The optimal values of λ and *C* were input into SVM classification algorithm to map different LULC classes in the study area using the three WV-2 spectral data sets.

3.2 Accuracy assessment

Figure 2 shows LULC maps obtained using SVM classifier. The main visual difference between the maps is that a relatively homogeneous map was produced when the 8B was used as compared with other spectral subsets. The maps also show that the Dukuduku indigenous forest was mainly surrounded by grassland and commercial forest plantation, while the grassland on the north eastern part of the study area was fragmented. Most of sugarcane farms in the study area were at a mature growth stage.
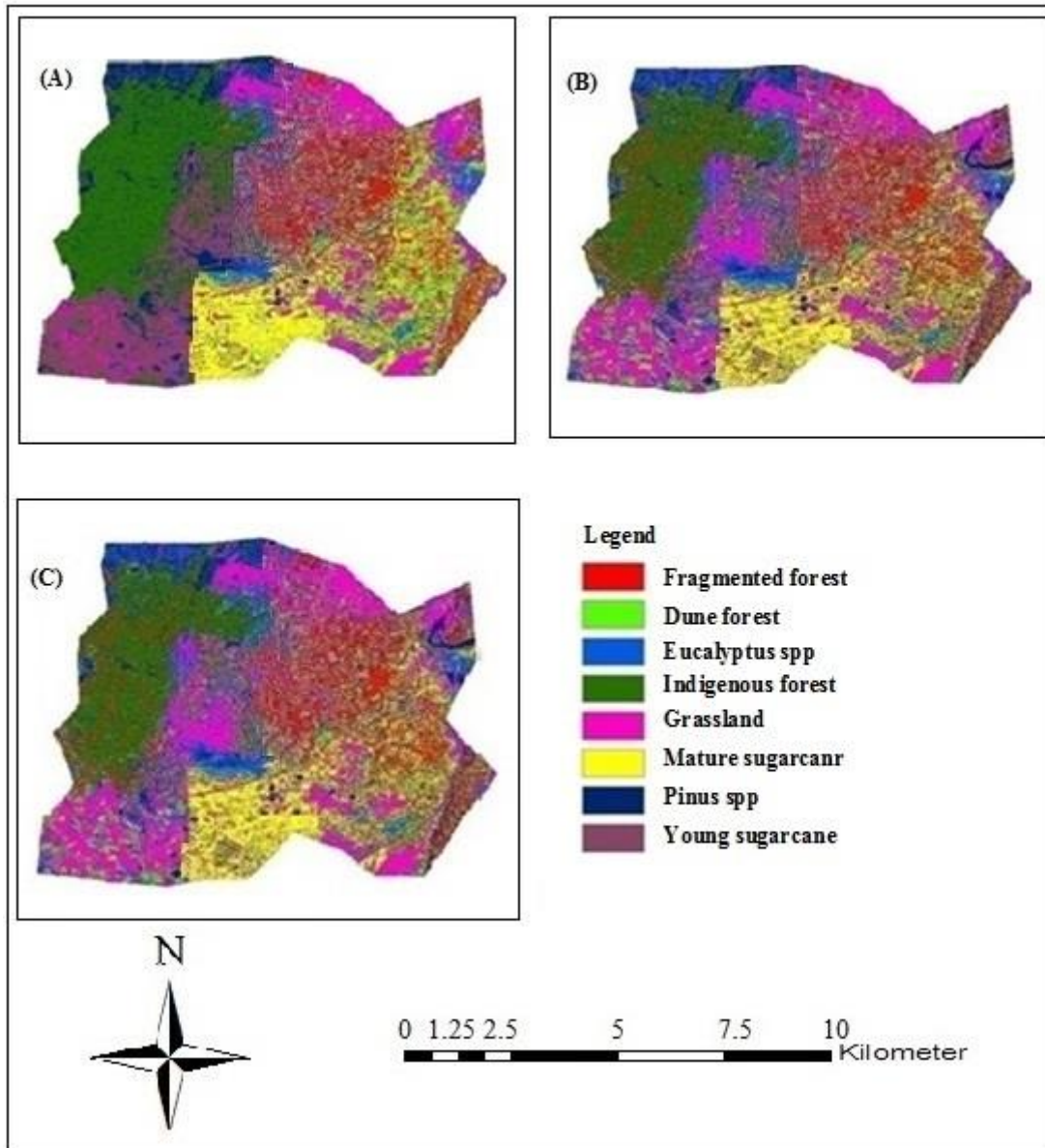
Figure 2. Land use/cover classification maps obtained using support vector machines classifier: (a) all eight WorldView-2 bands, (b) four standard WorldView-2 bands and (C) four additional WorldView-2 bands

Furthermore, the overall accuracy assessment for mapping LULC class was 78.0% (total disagreement = 22.0%), 51.0% (total disagreement = 49.0%) and 64.0% (total disagreement = 36.0%) using 8B, SB and AB respectively (Tables 1, 2 and 3). SVM classifier obtained quantity disagreement values of 5.0%, 13.0% and 14% for 8B, SB and AB respectively (Tables 1, 2 and 3). The tables also show relatively high allocation disagreement values of 17.0%, 36.0% and 22.0%.

Table 1. Classification confusion matrix of support vector machines (SVM) classifier using WorldView-2 8B for the 30% test data sets. The confusion matrix includes overall accuracy (OA), quantity disagreements (QD) and allocation disagreements (AD)

Class	FF	EP	IF	GL	MS	DF	PN	YS	Total
FF	15	0	0	3	1	5	0	0	24
EP	1	20	5	0	0	1	0	0	27
IF	0	1	13	2	1	0	1	1	19
GL	1	0	2	17	1	0	1	0	22
MS	3	0	1	0	16	0	0	0	20
DF	2	1	0	0	1	16	0	0	20
PN	0	0	1	0	2	0	20	1	24
YS	0	0	0	0	0	0	0	20	20
Total	22	22	22	22	22	22	22	22	176
OA (%)	78.0								
QD (%)	5.0								
AD (%)	17.0								

FF = Fragmented forest, EP = *Eucalyptus spp*, IF = Indigenous forest, GL= Grassland, MS = Mature sugarcane, DF = Dune forest, PN=*Pinus spp*, YS = Young sugarcane.

Table 2. Classification confusion matrix of support vector machines (SVM) classifier using WorldView-2 SB for the 30% test data sets. The confusion matrix includes overall accuracy (OA), quantity disagreements (QD) and allocation disagreements (AD)

Class	FF	EP	IF	GL	MS	DF	PN	YS	Total
FF	13	0	3	5	4	4	0	0	29
EP	0	13	3	0	0	4	0	0	20
IF	0	4	12	6	0	1	1	0	24
GL	0	0	0	2	0	0	0	0	2
MS	1	0	3	5	8	4	3	2	26
DF	6	4	1	1	0	9	0	2	23
PN	0	1	0	0	6	0	17	2	26
YS	2	0	0	3	4	0	1	16	26
Total	22	22	22	22	22	22	22	22	176
OA (%)	51.0								
QD (%)	13.0								
AD (%)	36.0								

FF = Fragmented forest, EP = *Eucalyptus spp*, IF = Indigenous forest, GL= Grassland, MS = Mature sugarcane, DF = Dune forest, PN=*Pinus spp*, YS = Young sugarcane.

Table 3. Classification confusion matrix of support vector machines (SVM) classifier using WorldView-2 AB for the 30% test data sets. The confusion matrix includes overall accuracy (OA), quantity disagreements (QD) and allocation disagreements (AD)

Class	FF	EP	IF	GL	MS	DF	PN	YS	Total
FF	19	0	2	4	4	0	0	0	29
EP	0	20	2	0	0	1	0	0	23
IF	0	0	11	3	0	3	0	0	17
GL	0	0	1	8	1	0	0	0	10
MS	1	1	0	3	9	1	2	2	19
DF	0	1	1	0	0	13	0	3	18
PN	0	0	0	2	8	1	19	2	32
YS	2	0	5	2	0	3	1	15	28
Total	22	22	22	22	22	22	22	22	176
OA (%)	64.0								
QD (%)	14.0								
AD (%)	22.0								

FF = Fragmented forest, EP = *Eucalyptus spp*, IF = Indigenous forest, GL= Grassland, MS = Mature sugarcane, DF = Dune forest, PN=*Pinus spp*, YS = Young sugarcane.

Generally, all LULC classes achieved over 70% producer's and user's accuracies, with exception of FF (PA = 63% and UA = 68%) and IF (PA of 68% and UA of 59%) when 8B subset was used (Figure 3). Our results indicate that the values of the UA were less than 50% for dune forest, grassland, and mature sugarcane classes on SB subset (Figure 4) and for grassland and mature sugarcane on AB subset (Figure 5).

According to McNemar's test, there was significant difference (Z ≥ 1.96) at 95% confidence level amongst the confusion matrices of SVM classifier using WV-2 8B, SB and AB subsets (Table 4). Table 5 shows areas under each LULC class obtained from WV-2 8B, SB and AB subsets using SVM classification algorithm. The incomparable areas obtained by WV-2 subsets also confirm the dissimilar performance of SVM classification algorithm. The study area is dominated by indigenous forest, commercial plantation and grassland.
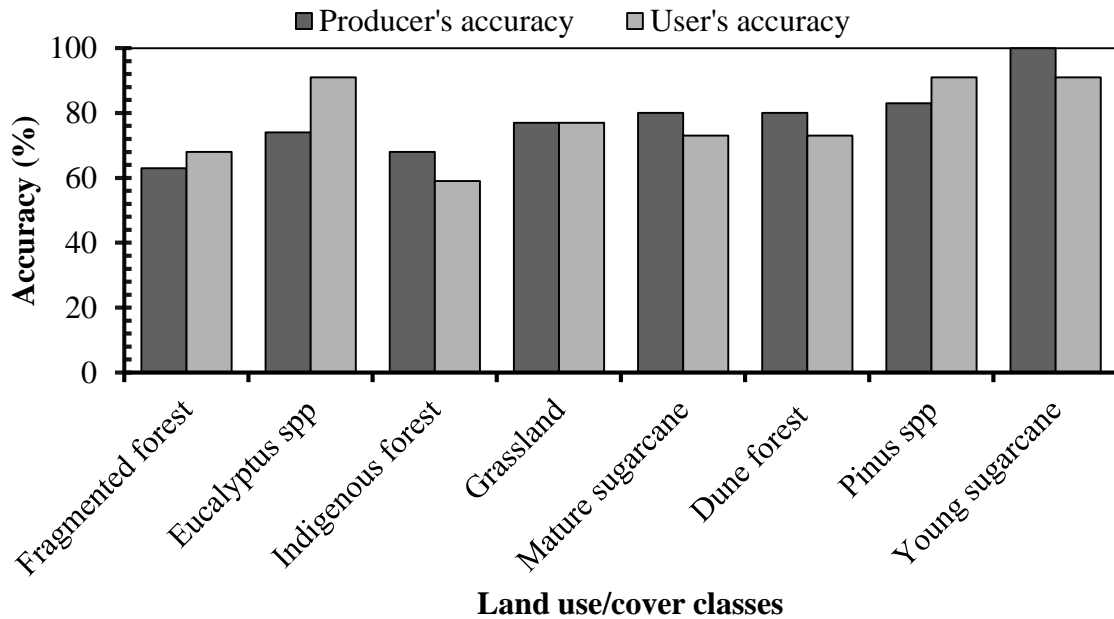
Figure 3. Producer's accuracy (%) and user's accuracy (%) of the studied eight LULC classes using all eight bands subset (8B) and support vector machines classifier for the 30% test data sets
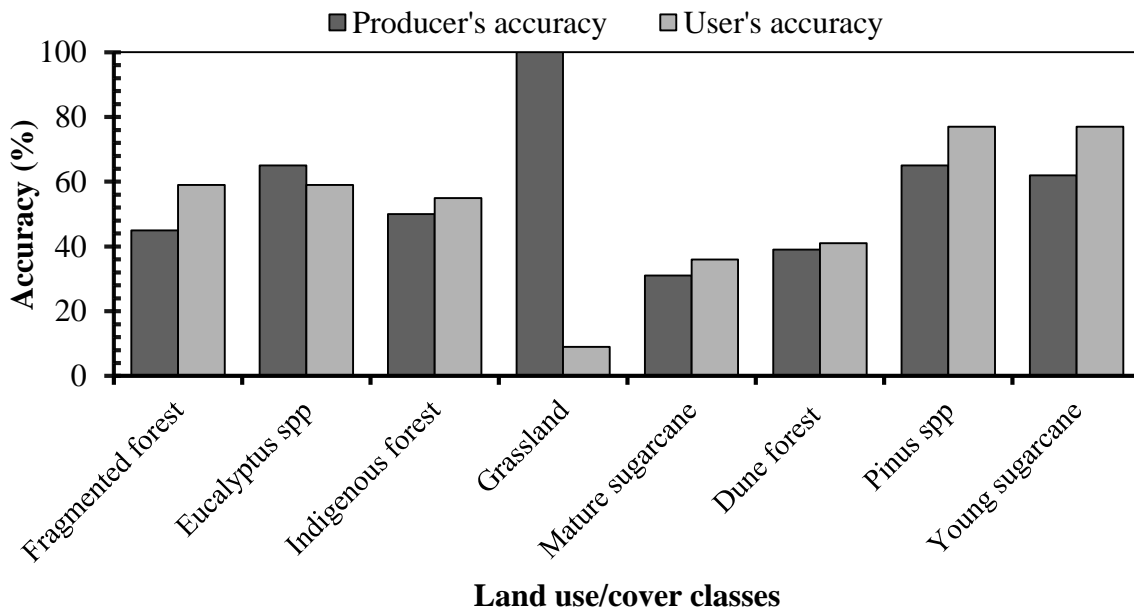

Figure 4. Producer's accuracy (%) and user's accuracy (%) of the studied eight LULC classes using standard bands subset (SB) and support vector machines classifier for the 30% test data sets
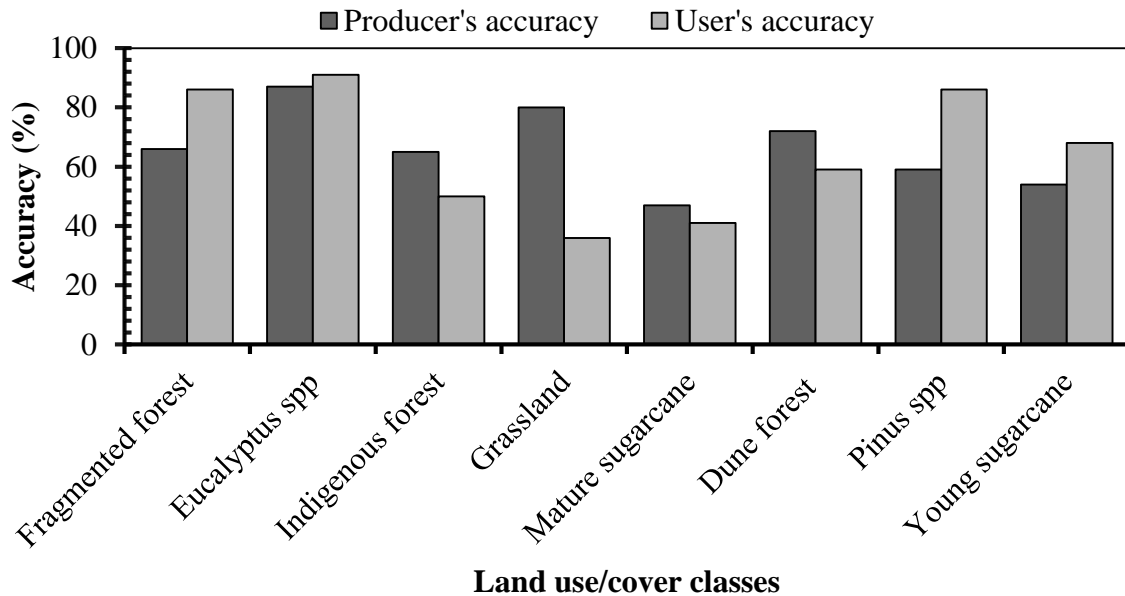
Figure 5. Producer's accuracy (%) and user's accuracy (%) of the studied eight LULC classes using additional bands subset (AB) and support vector machines classifier for the 30% test data sets

Table 4. McNemar's test result for comparing classification confusion matrices obtained using WV-2 8B, SB, and AB

		SB		
		Correctly Classified	**Misclassified**	**Total**
	Correctly Classified	98	12	110
8B	Misclassified	11	55	66
	Total	109	67	176
	Z Score	5.42		
		AB		
		Correctly Classified	Misclassified	Total
	Correctly Classified	96	14	110
8B	Misclassified	12	54	66
	Total	108	68	176
	Z Score	5.17		
		AB		
		Correctly Classified	Misclassified	Total
	Correctly Classified	93	16	109

SB	Misclassified	15	52	67
	Total	108	68	176
	Z Score	4.52		

8B = all WorldView-2 eight bands subset, SB = four WorldView-2 standard bands subset and AB = four WorldView-2 additional bands subset.

Table 5. Area of each LULC class in the study area obtained from all WorldView-2 eight bands, four WorldView-2 standard bands and four WorldView-2 additional bands subsets based on support vector machines classification algorithm

Land use/cover class	Area (ha)		
	All WorldView-2 eight bands subset (8B)	**Four WorldView-2 standard bands subset (SB)**	**Four WorldView-2 additional bands subset (AB)**
Fragmented forest	1918.52	1931.66	2131.63
Dune forest	2146.92	1945.23	1572.41
Eucalyptus spp	1579.09	2636.70	1520.79
Indigenous forest	4292.60	2814.09	4569.94
Grassland	2762.34	4388.52	2697.19
Mature sugarcane	2348.33	2106.46	2444.29
Pinus spp	1848.63	1818.57	1700.74
Young sugarcane	2991.09	2246.29	3250.52

4. Discussion

The main finding of the present study was that WV-2 8B significantly outperformed SB and AB subsets in mapping LULC classes in a fragmented ecosystem. Our finding is in conformity with other studies (Elsharkawy *et al.*, 2012, Pu and Landry, 2012, Peerbhay *et al.*, 2014) that demonstrated the utility of WV-2 8B in mapping LULC class in intact landscapes. There are two reasons that may have led to high accuracy. Firstly, the LULC class in Dukuduku forest consists of vegetation and WV-2 strategically located bands are effective in differentiating vegetated surfaces (Marchisio *et al.*, 2010, Yang, 2011, Alsubaie, 2012). Secondly, the SVM classification algorithm is useful for LULC mapping of fragmented ecosystem because SVM reduces classification error on test data points without a prior assumption about their distribution (Mountrakis *et al.*, 2011, Ghosh and Joshi, 2014).

Moreover, SVM is a known versatile classifier that constructs models based on a small

data from different classes (Cortes and Vapnik, 1995) maximizing the margin between the support vectors and the hyperplane. The classification error is therefore significantly minimized. In the present study, we used a non-linear kernel function to perform SVM classification. A non-linear kernel is an efficient method to solve inseparability problems that may be found in the LULC classes. The relatively good performance of SVM classifier obtained in this study is consistent with the findings of Huang *et al.* (2002), Kavzoglu and Colkesen (2009) and Petropoulos *et al.* (2012) who utilized a kernel functions analysis of SVM for classifying remotely-sensed data and concluded that the classifier leads to improved classification accuracy. A number of authors have found that SVM was the best classification technique for mapping LULC using high spatial resolution imagery such as WV-2 (Pal, 2006, Chen, 2011, Pu and Landry, 2012). To our knowledge, WV-2 subsets (8B, SB and AB) have never been compared for mapping LULC classes in areas with small dataset samples such as the fragmented ecosystem in the Dukuduku area of KwaZulu-Natal, South Africa. Our study showed that SVM classifier was unable to fully deal with the high spectral variation inherent in some LULC classes like mature sugarcane and grassland which obtained relatively lower UA and PA (see Figs 4 and 5).This is a common problem when classifying heterogeneous landscapes using high spatial resolution (WV-2 image) based on per-pixel classification techniques (Lu and Weng, 2007).

Although the eight LULC classes could be separated accurately using only the four standard bands, the use of the WV-2 strategically located bands (coastal blue, yellow, red edge and NIR-2) led to a considerable improvement in the classification accuracy. That is expected when advanced machine learning algorithm is used with WV-2 data. The additional wavebands are expected to provide an increase of up to 30% in classification accuracy (Zhou *et al.*, 2012). The low UA for the GL and MS classes indicate that there is a probability that pixels classified as GL and MS may not actually exist on the ground. That is expected since the physiological age of a mature sugarcane crop could be similar to densely vegetated grassland as a result of some confounding factors such as weeds and abiotic stressors (Abdel-Rahman *et al.*, 2013) and hence similar spectral characteristics.

In summary, our findings are promising for accurate mapping of LULC in fragmented areas as it demonstrates the possibility of mapping LULC classes using WV-2 data and SVM classifier. Moreover, the relatively accurate classification result obtained in this study provides reliable information on LULC classes in the Dukuduku area. That could be used for the design of management plans policies as a basis for assessing and monitoring natural resources, ecological fragmentation and the ecosystem function. This information is therefore critical in the management of one of the most valuable landscapes in South Africa. Furthermore, mapping sugarcane ages is quite useful information for the Southern African sugar industry for making

informed decisions with regard to sugarcane harvesting and milling. Further research is needed to widen the use of WV-2 imagery in identifying the rare forest species within the indigenous forest in the north-western part of the study area.

We have mapped eight course LULC classes using remotely-sensed data of fine pixel size. The intra-classes variability could have exceeded the fine pixel size of WV-2 image. Hence, remotely-sensed data of medium spatial resolution (e.g., 10 or 20 m) could yield relatively better classification results. Furthermore, we did not consider the scale of the fragmentation which should match the image pixel size to accurately map the LULC classes, since our study area is a fragmented landscape (Cho *et al.*, 2013). Future studies should look at the scale of the fragmentation using fragstats metrics extracted from multi-temporal satellite data. Scale of fragmentation should then be matched with an appropriate image resolution to derive more accurate LULC classification maps.

5. Conclusions

The present study shows a successful application of multispectral WV-2 data and the machine learning SVM classifier in mapping eight LULC classes in a fragmented ecosystem. The results show that WV-2 8B significantly outperforms both SB and AB subsets in mapping LULC classes, achieving an overall accuracy of 78%. On the other hand, WV-2 AB subset yielded significantly higher classification accuracy than SB subset. The results further demonstrate that the classification error for mature sugarcane and grassland was relatively higher. Our study provides LULC maps that could be used as essential information for decision-making regarding land management and policy making strategies in the fragmented Dukuduku area. It is recommended that further studies should look at identifying threatened (rare) tree species within the indigenous Dukuduku forest using a sub pixel classification algorithm.

6. Acknowledgments

7. References

Abdel-Rahman, E. M., Mutanga, O., Adam, E. & Ismail, R. 2014. Detecting Sirex noctilio grey-attacked and lightning-struck pine trees using airborne hyperspectral data, random forest and support vector machines classifiers. *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 88, pp. 48-59.

Abdel-Rahman, E. M., Way, M., Ahmed, F., Ismail, R. & Adam, E. 2013. Estimation of thrips (Fulmekiola serrata Kobus) density in sugarcane using leaf-level hyperspectral data. *South African Journal of Plant and Soil*, vol. 30, no.2, pp. 91-96.

Adam, E., Mutanga, O., Odindi, J. & Abdel-Rahman, E. M. 2014. Land-use/cover classification in a heterogeneous coastal landscape using RapidEye imagery: evaluating the performance of random forest and support vector machines classifiers. *International Journal of Remote Sensing*, vol. 35, no.10, pp. 3440-3458.

Alsubaie, N. M. 2012. '*The Potential of Using Worldview-2 Imagery for Shallow Water Depth Mapping'*. Master degree. University of Calgary.

Benitez-Malvido, J. 1998. Impact of forest fragmentation on seedling abundance in a tropical rain forest. *Conservation Biology*, vol. 12, no.2, pp. 380-389.

Burges, C. J. 1998. A tutorial on support vector machines for pattern recognition. *Data mining and knowledge discovery*, vol. 2, no.2, pp. 121-167.

Chen, Q. 2011. Comparison of Worldview-2 and IKONOS-2 Imagery for Identifying Tree Species in the Habitat of an Endangered Bird Species in Hawaii; 8-Band Research Challenge; DigitalGlobe: Longmont, CO, USA.

Cho, M., Naidoo, L., Mathieu, R. & Asner, G. 'Mapping savanna tree species using Carnegie Airborne Observatory hyperspectral data resampled to WorldView-2 multispectral configuration'. Proc. 34th International Symposium on Remote Sensing of Environment, Sydney, Australia, 2011.10-15.

Cho, M. A., Debba, P., Mutanga, O., Dudeni-Tlhone, N., Magadla, T. & Khuluse, S. A. 2012. Potential utility of the spectral red-edge region of SumbandilaSat imagery for assessing indigenous forest structure and health. *International Journal of Applied Earth Observation and Geoinformation*, vol. 16, pp. 85-93.

Cho, M. A., Ramoelo, A., Debba, P., Mutanga, O., Mathieu, R., van Deventer, H. & Ndlovu, N. 2013. Assessing the effects of subtropical forest fragmentation on leaf nitrogen distribution using remote sensing data. *Landscape ecology*, vol. 28, no.8, pp. 1479-1491.

Cingolani, A. M., Renison, D., Zak, M. R. & Cabido, M. R. 2004. Mapping vegetation in a heterogeneous mountain rangeland using Landsat data: an alternative method to define and classify land-cover units. *Remote Sensing of Environment*, vol. 92, no.1, pp. 84-97.

Civco, D. L. 1993. Artificial neural networks for land-cover classification and mapping. *International Journal of Geographical Information Science*, vol. 7, no.2, pp. 173-186.

Clark, M. L., Roberts, D. A. & Clark, D. B. 2005. Hyperspectral discrimination of tropical rain forest tree species at leaf to crown scales. *Remote sensing of environment*, vol. 96, no.3, pp. 375-398.

Cortes, C. & Vapnik, V. 1995. Support-vector networks. *Machine learning*, vol. 20, no.3, pp. 273-297.

Dalponte, M., Bruzzone, L., Vescovo, L. & Gianelle, D. 2009. The role of spectral resolution and classifier complexity in the analysis of hyperspectral images of forest areas. *Remote Sensing of Environment*, vol. 113, no.11, pp. 2345-2355.

Eldeen, I. M. S. 2005. Pharmacological Investigation of some trees used in South African Traditional Medicine PhD Thesis. Kwazulu-Natal University. South Africa.

Elsharkawy, A., Elhabiby, M. & El-Sheimy, N. 'Improvement in the detection of land cover class using the Worldview-2 imagery'. *ASPRS, International Scientific Conference* in Sacramento, California, March 19-23, 2012.

ENVI 2006. Environment for Visualising Images, ITT Industries Inc., Boulder, Colorado, ITT Industries Inc, USA.

Foley, J. A., DeFries, R., Asner, G. P., Barford, C., Bonan, G., Carpenter, S. R., Chapin, F. S., Coe, M. T., Daily, G. C. & Gibbs, H. K. 2005. Global consequences of land use. *science*, vol. 309, no.5734, pp. 570-574.

Foody, G. M. 2002. Status of land cover classification accuracy assessment. *Remote sensing of environment*, vol. 80, no.1, pp. 185-201.

Foody, G. M. 2004. Thematic Map Comparison. *Photogrammetric Engineering & Remote Sensing*, vol. 70, no.5, pp. 627-633.

Ghosh, A. & Joshi, P. 2014. A comparison of selected classification algorithms for mapping bamboo patches in lower Gangetic plains using very high resolution WorldView 2 imagery. *International Journal of Applied Earth Observation and Geoinformation*, vol. 26, pp. 298-311.

Heumann, B. W. 2011. An object-based classification of mangroves using a hybrid decision tree—Support vector machine approach. *Remote Sensing*, vol. 3, no.11, pp. 2440-2460.

Huang, C., Davis, L. & Townshend, J. 2002. An assessment of support vector machines for land cover classification. *International Journal of Remote Sensing*, vol. 23, no.4, pp. 725-749.

Kavzoglu, T. & Colkesen, I. 2009. A kernel functions analysis for support vector machines for land cover classification. *International Journal of Applied Earth Observation and Geoinformation*, vol. 11, no.5, pp. 352-359.

Kavzoglu, T. & Mather, P. 2003. The use of backpropagating artificial neural networks in land cover classification. *International Journal of Remote Sensing*, vol. 24, no.23, pp. 4907-4938.

Larsen, M. 2007. Single tree species classification with a hypothetical multi-spectral satellite. *Remote Sensing of Environment*, vol. 110, no.4, pp. 523-532.

Leeuw, J. D., Jia, H., Yang, L., Liu, X., Schmidt, K. & Skidmore, A. 2006. Comparing accuracy assessments to infer superiority of image classification methods. *International Journal of Remote Sensing*, vol. 27, no.1, pp. 223-232.

Lu, D. & Weng, Q. 2007. A survey of image classification methods and techniques for improving classification performance. *International journal of Remote sensing*, vol. 28, no.5, pp. 823-870.

Marchisio, G., Pacifici, F. & Padwick, C. 'On the relative predictive value of the new spectral bands in the WorldWiew-2 sensor'. *IEEE International Geoscience and Remote Sensing Symposium* (IGARSS), 2010, pp. 2723-2726.

McCarthy, M. J. & Halls, J. N. 2014. Habitat Mapping and Change Assessment of Coastal Environments: An Examination of WorldView-2, QuickBird, and IKONOS Satellite Imagery and Airborne LiDAR for Mapping Barrier Island Habitats. *ISPRS International Journal of Geo-Information*, vol. 3, no.1, pp. 297-325.

Mountrakis, G., Im, J. & Ogole, C. 2011. Support vector machines in remote sensing: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 66, no.3, pp. 247-259.

Otukei, J. & Blaschke, T. 2010. Land cover change assessment using decision trees, support vector machines and maximum likelihood classification algorithms. International *Journal of Applied Earth Observation and Geoinformation*, vol. 12, pp. S27-S31.

Ozdemir, I. & 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*, vol. 13, no.5, pp. 701-710.

Pal, M. 'Random forests for land cover classification'. *IEEE International Geoscience and Remote Sensing Symposium, IGARSS'03. Proceedings*. 2003, pp. 3510-3512.

Pal, M. 2006. Support vector machine-based feature selection for land cover classification: a case study with DAIS hyperspectral data. *International Journal of Remote Sensing*, vol. 27, no.14, pp. 2877-2894.

Peerbhay, K. Y., Mutanga, O. & Ismail, R. 2014. Investigating the Capability of Few Strategically Placed Worldview-2 Multispectral Bands to Discriminate Forest Species in KwaZulu-Natal, South Africa. *IEEE Journal of Selected Topics in Applied Earth Observation and Remote sensing*, vol. 7, no.1, pp. 307-316.

Penner, J. E. 1994. Atmospheric chemistry and air quality. *Changes in land use and land cover: a global perspective*, vol., pp. 175-209.

Petropoulos, G. P., Kalaitzidis, C. & Prasad Vadrevu, K. 2012. Support vector machines and object-based classification for obtaining land-use/cover cartography from Hyperion hyperspectral imagery. *Computers & Geosciences*, vol. 41, pp. 99-107.

Pignatti, S., Cavalli, R. M., Cuomo, V., Fusilli, L., Pascucci, S., Poscolieri, M. & Santini, F. 2009. Evaluating Hyperion capability for land cover mapping in a fragmented ecosystem: Pollino National Park, Italy. *Remote Sensing of Environment*, vol. 113, no.3, pp. 622-634.

Pontius, R. G. & Millones, M. 2011. Death to Kappa: birth of quantity disagreement and allocation disagreement for accuracy assessment. *International Journal of Remote Sensing*, vol. 32, no.15, pp. 4407-4429.

Pu, R. & Landry, S. 2012. A comparative analysis of high spatial resolution IKONOS and WorldView-2 imagery for mapping urban tree species. *Remote Sensing of Environment*, vol. 124, pp. 516-533.

R Development Core, T. 2012. 'R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, http://www.R-project.or, Vienna. (accessed 5.01.13)'. ISBN 3-900051-07-0.

Vaiphasa, C., Skidmore, A. K., de Boer, W. F. & Vaiphasa, T. 2007. A hyperspectral band selector for plant species discrimination. *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 62, no.3, pp. 225-235.

van Wyk, B.E. 2008. A broad review of commercially important southern African medicinal plants. *Journal of Ethnopharmacology*, vol. 119, no.3, pp. 342-355.

van Wyk, G. F., Everard, D. A., Midgley, J. J. & Gordon, I. G. 1996. Classification and dynamics of a southern African subtropical coastal lowland forest. *South African Journal of Botany*, vol. 62, no. 3, pp. 133–142.

Vapnik, V. 1998. Statistical learning theory. Support Vector Machines for Pattern Recognition. John Wiley & Sons, New York.

Von Maltitz, G., Mucina, L., Geldenhuys, C., Lawes, M., Eeley, H., Adie, H., Vink, D., Fleming, G. & Bailey, C. 2003. Classification system for South African indigenous forests: an objective

classification for the Department of Water Affairs and Forestry. Environmentek report. ENV-PC, vol. 17, pp. 1-28.

Yang, X. 2011. Parameterizing support vector machines for land cover classification. *Photogrammetric engineering and remote sensing*, vol. 77, no.1, pp. 27-37.

Zhang, Y., Gao, J. & Wang, J. 2007. Detailed mapping of a salt farm from Landsat TM imagery using neural network and maximum likelihood classifiers: a comparison. *International Journal of Remote Sensing*, vol. 28, no.10, pp. 2077-2089.

Zhou, X., Janscó, T., Chen, C. & Verőné, M. W. 'Urban land cover mapping based on object oriented classification using WorldView 2 satellite remote sensing images'. *International Scientific Conference on Sustainable Development & Ecological Footprint*, 2012.