

Article

Mapping Woodland Cover in the Miombo Ecosystem: A Comparison of Machine Learning Classifiers

Courage Kamusoko ^{1,*}, Jonah Gamba ² and Hitomi Murakami ³

¹ Asia Air Survey (AAS) Co., Ltd., Kanagawa 215-0004, Japan

² TOPS Systems Corp., Tsukuba 305-0032, Japan; E-Mail: jgamba@topscom.co.jp

³ Department of Computer and Information Science, Faculty of Science and Technology, Seikei University, Tokyo 180-8633, Japan; E-Mail: hi.murakami.g@gmail.com

* Author to whom correspondence should be addressed; E-Mail: cou.kamusoko@ajiko.co.jp; Tel.: +81-44-969-7510; Fax: +81-44-965-0917.

Received: 21 April 2014; in revised form: 13 June 2014 / Accepted: 13 June 2014 /

Published: 20 June 2014

Abstract: Miombo woodlands in Southern Africa are experiencing accelerated changes due to natural and anthropogenic disturbances. In order to formulate sustainable woodland management strategies in the Miombo ecosystem, timely and up-to-date land cover information is required. Recent advances in remote sensing technology have improved land cover mapping in tropical evergreen ecosystems. However, woodland cover mapping remains a challenge in the Miombo ecosystem. The objective of the study was to evaluate the performance of decision trees (DT), random forests (RF), and support vector machines (SVM) in the context of improving woodland and non-woodland cover mapping in the Miombo ecosystem in Zimbabwe. We used Multidate Landsat 8 spectral and spatial dependence (Moran's I) variables to map woodland and non-woodland cover. Results show that RF classifier outperformed the SVM and DT classifiers by 4% and 15%, respectively. The RF importance measures show that multidate Landsat 8 spectral and spatial variables had the greatest influence on class-separability in the study area. Therefore, the RF classifier has potential to improve woodland cover mapping in the Miombo ecosystem.

Keywords: Zimbabwe; Miombo woodlands; Landsat 8; decision trees; random forests; support vector machines

1. Introduction

Miombo woodlands are extensive in the Democratic Republic of Congo (DRC), Angola, Tanzania, Mozambique, Malawi, Zambia and Zimbabwe [1,2]. These broad-leaved deciduous woodlands—dominated by tree species, such as *Brachystegia*, *Julbernardia*, and *Isoberlinia*—provide important ecosystem, socioeconomic and cultural services in Central and Southern Africa [3]. In Zimbabwe, Miombo woodlands cover approximately 42% of the country [4]. The woodlands are the primary source of firewood, construction poles, medicine, and food in rural areas [5]. However, rapid population growth and tobacco farming by newly resettled farmers have increased deforestation and woodland degradation in the Miombo ecosystems [4]. As a result, the livelihoods of two-thirds of the rural population dependent on the Miombo ecosystem is under threat unless sustainable agro-forestry development policies are implemented [5].

Timely and up-to-date land cover information is required to formulate and implement effective sustainable agro-forestry development policies in the Miombo ecosystem. However, such land cover information is sparse or lacking given the high cost of conducting conventional land cover surveys [4]. Medium resolution satellite remote sensing data are relatively inexpensive sources for mapping land cover at a regional scale [6]. More recently, there has been an increased use of medium resolution satellite data, such as Landsat Thematic Mapper (TM) and Enhanced Landsat Thematic Mapper (ETM+) imagery, since the datasets are available for free [7]. Although Landsat data have improved forest cover mapping in tropical evergreen rainforest ecosystems [8,9], woodland cover mapping remains a challenge in the Miombo ecosystem [10]. This is mainly attributed to a number of environmental and anthropogenic factors. First, the Miombo woodlands exhibit high degrees of spatial heterogeneity (e.g., tree density and size), which is influenced by soil type, fire, herbivores, land use, *etc.* [11]. Consequently, it is difficult to quantify biophysical properties of the Miombo woodlands (e.g., canopy cover, structure), especially in areas where closed and open woodlands (characterized by wooded grassland, and bushland) alternates [12]. Second, it is difficult to discriminate closed and open woodlands at the resolution of Landsat 8 imagery because the two land cover types are close in space [13]. Third, woodland cover is influenced by seasonal changes. Therefore, spectral reflectance patterns of the woodland vary according to water availability during the short rainy season (November to March) and long dry season (April to October) [13]. Previous studies have used satellite imagery (e.g., Landsat and SPOT) to map land cover in the Miombo woodlands and dry forests [14–17]. However, these studies used single-date satellite imagery, which fail to capture dynamic vegetation changes in the Miombo ecosystem [18].

Spectral mixture analysis classifiers (which assume that a pixel's spectrum is a linear combination of spectral distinct endmembers) [19,20] have been recommended to map woodland cover in the dry Miombo ecosystems [10]. While spectral mixture analysis classifiers have been relatively successful for mapping land cover in the Brazilian Cerrado [10,21], the classifiers have some limitations [22]. First, spectral mixture analysis classifiers assume linear spectral mixing of land cover reflectance [23]. However, past studies have shown that land cover reflectance mix in non-linear trend, especially when multiple scattering effects from the background and canopy layers are taken into consideration [24,25]. Second, the number of endmembers (e.g., vegetation, high albedo, low albedo) must account for the number of classes in the pixel, and their spectral separability should be sufficient in order to avoid

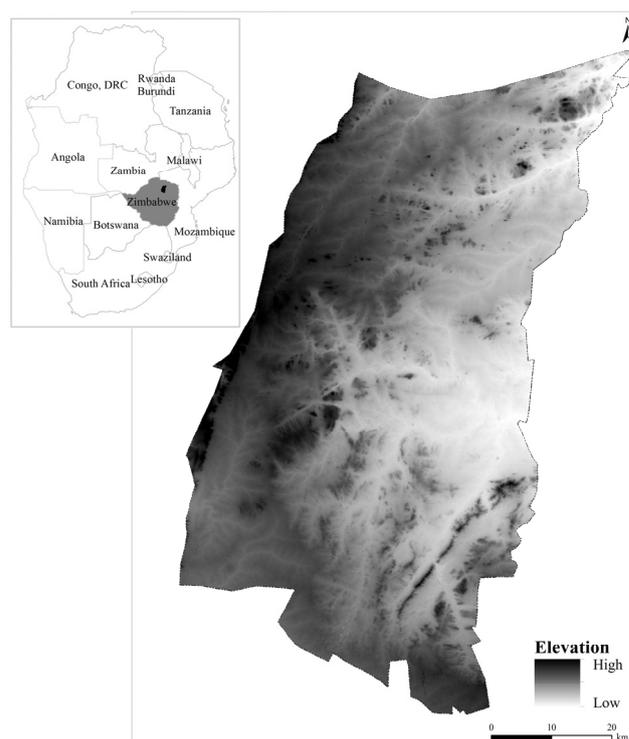
confusion [26]. Nonetheless, past studies have revealed that endmembers do not generally correspond to physical land cover components like tree canopy [22,23].

Recently, researchers have shown that non-parametric machine learning classifiers such as decision trees (DT), support vector machines (SVM), and random forests (RF) improve land cover mapping [22,27–31]. For example, Rogriguez-Galiano *et al.* [32] and Grinand *et al.* [33] successfully applied machine learning classifiers for mapping land cover in the Mediterranean and evergreen tropical ecosystems. However, machine learning classifiers have not been tested for mapping land cover in general and woodland cover in particular in the Miombo ecosystem. The objective of this study is to evaluate the performance of DT, SVM, and RF classifiers in the context of improving woodland and non-woodland cover mapping in the Miombo ecosystem in Southern Africa. Taking Mazowe district in Zimbabwe as an example, we used multivariate Landsat 8 spectral and spatial (Moran's I) variables to map woodland and non-woodland cover. This study area was selected because of the complex nature of land use and management practices (e.g., a mixture of commercial and subsistence agriculture), which has resulted in a landscape mosaic that comprise closed and open woodland communities, grassland, and agriculture fields.

2. Study Area

Mazowe district is located in Mashonaland Central province of Zimbabwe (Figure 1). The study area covers an area of approximately 4662 km². The altitude varies from 1000 m to 1740 m above sea level (Figure 1). The highest temperatures usually occur in the second half of October or early November with an average maximum temperature in the range of 26 °C–35 °C. The study area receives a mean annual rainfall ranging from 700 mm to 1000 mm and is distributed from mid-October to April.

Figure 1. Location of Mazowe district, Zimbabwe.

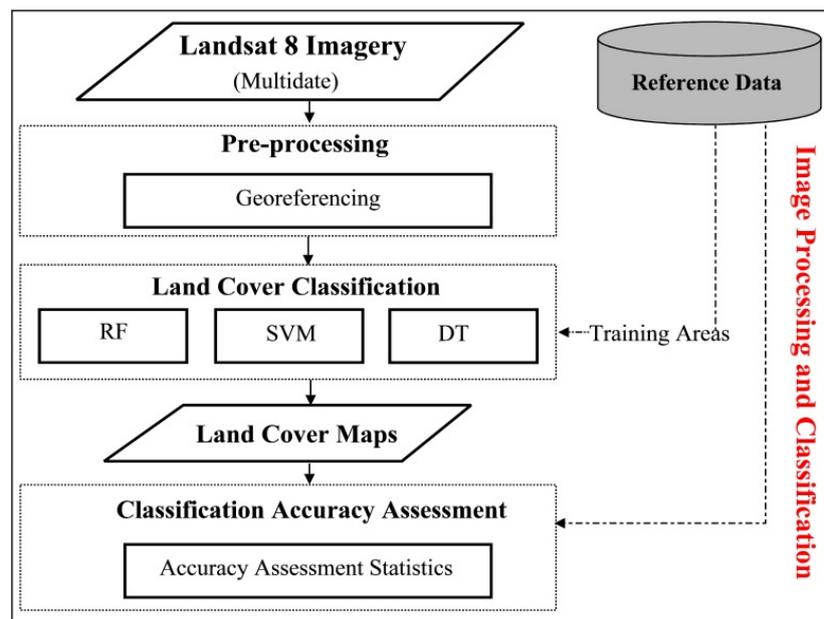


The study area is dominated by a complex landscape of closed and open Miombo woodlands, bushland, grassland and agriculture. Soil varies from ferralsols, luvisols, lithosols, and nitosols. The major economic activity in the study area is commercial and semi-subsistence agriculture, with major crops such as tobacco, cotton, maize, and groundnuts, as well as vegetables, in those areas with irrigation. However, production of major rainfed crops is usually affected by unreliable rainfall patterns, particularly the late onset of the rainy season. According to the 2012 population census, population increased from 198,319 in 1992 to 232,885 in 2012 [34].

3. Methodology

The methodology used in this study comprised four major components, namely data acquisition, pre-processing, land cover classification, and accuracy assessment (Figure 2). The following subsections describe data, classification scheme design, and land cover classification procedures.

Figure 2. Land cover classification flow diagram.



3.1. Data

We acquired four multidate Landsat 8 scenes (Table 1) for image processing and classification. Landsat 8 (originally called Landsat Data Continuity Mission) was launched on 11 February 2013, as the eighth satellite in the Landsat program [35,36]. The sensor consists of the Operational Land Imager (OLI) and the Thermal Infrared Sensor (TIRS) sensors, which provides images at a spatial resolution of 15 m (panchromatic), 30 m (visible, NIR, SWIR), and 100 m (thermal) [35,36]. All Landsat 8 image dates were selected from cloud-free scenes acquired during the post-rainy and dry seasons (Table 1) in order to account for seasonality or vegetation phenology in the classification. The four multidate Landsat 8 scenes were georeferenced to the Universal Transverse Mercator (UTM) map projection (zone 36 south). We did not perform atmospheric correction because the Landsat 8 multidate composite was classified as if it were a single image [20,33,37]. In addition, we used Moran's I to derive spatial dependence (autocorrelation) information from the Landsat 8 scene acquired on June 2013 because this

image captures healthy woodland canopy (leaf-on). Spatial dependence (autocorrelation) measures the degree to which spatial features and their data values are clustered in space (positive spatial autocorrelation) or dispersed (negative spatial autocorrelation) [38]. Past studies show that Moran's I improves land cover mapping, particularly in open canopy woodland areas [31,32]. A total of 27 features (that is 24 bands from Landsat 8 and three Moran's I images) were used for classification (Table 1).

Table 1. Summary of datasets used in the study.

Sensor	Path/Row	Acquisition Date	Remarks
Landsat 8	170/72	19 April 2013	Post-rainy season data
Landsat 8	170/72	6 June 2013	Post-rainy season data
Landsat 8	170/72	25 August 2013	Dry season data
Landsat 8	170/72	28 October 2013	Dry season data
Moran's I band 4		6 June 2013	Derived from band 4
Moran's I band 5		6 June 2013	Derived from band 5
Moran's I band 7		6 June 2013	Derived from band 7

Reference datasets were developed for classifier training (Table 2) and classification accuracy assessment for 2013. The primary reference data was obtained from very high-resolution images (e.g., Quickbird image) in Google Earth™ [39]. In addition, secondary reference data for 2013 was obtained from Global Positioning System (GPS) points collected in September 2012.

Table 2. Land cover classification scheme and the total number of training pixels used with the classifiers.

Land Cover	Description	Training Pixels
Closed woodland:	All wooded areas with over 20% of the deciduous trees above 5 m in height. It also includes riverine vegetation with sparse grass cover, mainly of perennial species.	769
Open woodland:	Open deciduous or scattered trees with a canopy cover of about 5%–20% and height greater than 5 m. This class also includes a varying density of small shrubs and bush. The grass cover is well developed and continuous due to the low canopy cover.	584
Grassland:	Dominant grass cover areas with sparse or no shrubs and bush or trees.	721
Agriculture:	This class includes areas currently under crop, orchards, land under irrigation, cultivated land or land being prepared for cultivation.	423
Others:	Non-vegetated areas such as bare rocks, or areas with very little vegetation cover (excluding agricultural fields with no crop cover), where soil exposure is clearly apparent. This class also includes quarries, mine dumps and settlement areas.	799
Water:	Rivers and reservoirs.	253

3.2. Classification Scheme Design

A modified land cover classification scheme was used for image classification. The modified land cover classification scheme is based on the Forestry Commission (Zimbabwe) woody cover classes

and the author's *a priori* knowledge of the study area. The original land cover classes were modified with the aid of very high-resolution images (e.g., Quickbird image) from Google Earth™ [39] and fieldwork. In total, six land cover classes were considered in this study: (1) closed woodland; (2) open woodland; (3) grassland; (4) agriculture; (5) others; and (6) water. Detailed descriptions of these land cover classes are provided in Table 2.

3.3. Land Cover Classification

We used DT, RF, and SVM classifiers available in R [40] for land cover classification (Figure 2). R is a free and open source statistical and computer graphic software, which offers a wide range of machine learning classifiers.

Decision trees (DT) are non-parametric and hierarchical (top-down) splitting classifiers, which use a sequence of decisions to classify objects of interest [41,42]. Generally, DT classifiers are composed of a root node, a set of interior nodes and terminal nodes called “leaves” [41]. A multispectral remote sensing dataset is subdivided into categories based on a splitting mechanism, which chooses the best feature to split the dataset [43]. The CART (classification and regression trees) and C4.5 are the most commonly used decision trees [44,45]. The former is a binary classifier that uses the Gini impurity index to measure the impurity of a data partition, while the latter is a multiple classifier that uses information gain as a feature selection measure for node splitting [42]. The advantages of the DT classifiers are: (i) they can easily integrate numerical and categorical data; (ii) require less training time compared to artificial neural networks (ANN) and SVM while achieving similar accuracies [46]; and (iii) are free of normal distribution assumptions [47]. However, DT requires large training samples for tree classification and the stability of trees is affected by outliers or small changes in training data [47]. In this study, the CART algorithm available in the *rpart* package [48] was used to build a decision tree and classify the multivariate Landsat 8 spectral and Moran's I spatial dependence variables. The DT parameters were set as follows: the minimum split (Min Split) was specified as 20; maximum depth (Max Depth) was specified as 30; minimum bucket (Min Bucket) was specified as 7; and complexity parameter (cp) was specified as 0.01. The minimum split specifies the minimum number of observations that exist at a node before it is considered for splitting, while the minimum bucket size is the minimum number of observations in any leaf node [48]. The maximum depth is used to limit the depth of trees, whereas the complexity parameter is used to control the size of the decision trees and to select the optimal tree size (that is, for pruning the decision trees) [48].

Random Forests (RF) is an ensemble (collection) classifier, which uses bagging (bootstrap aggregated sampling) to build many individual decision trees for final classification [31]. The algorithm uses a random subset of predictor variables to split an observation data into homogenous subsets [31]. The node-splitting variable with the greatest increase in data purity (variance or Gini) is selected, which gives the overall model more generalization capacity before and after the split [23]. A majority voting procedure is used to produce the final labeling [43]. The RF classifier uses out-of-bag (OOB) sample data, which are derived from data that are not in the bootstrap sample to evaluate performance [43]. In addition, importance measures (mean decrease in accuracy or Gini index) are computed by comparing the proportion between misclassifications and OOB sample, which provides an unbiased estimation of the generalization error that is used for feature selection [32]. The

advantages of RF classifier are: (i) they can handle large database (e.g., thousands of input numerical and categorical variables); (ii) require less training time compared to other machine learning classifiers (e.g., ANN, SVM, boosting); (iii) are free of normal distribution assumptions; (iv) robust to outliers and noise; and (v) quantifies each input variable into importance measure [32,43]. We used the *randomForest* package [49] to classify all the Landsat 8 spectral and Moran's I spatial dependence variables. The *randomForest* package is based on the original Fortran code, which was developed by Breiman [50]. In this study, 500 trees were used to construct the RF model. The parameter *ntry*, which represents the number of variables to be considered at every node, was specified as 5. For this RF model, *ntry* is the square root of the total number of variables used for classification.

Support vector machines (SVM) are machine learning classifiers based on statistical learning theory [51]. The classifiers perform classification by constructing hyperplanes in a multidimensional space [52]. The SVM classifiers were introduced by Boser *et al.* [53] and Vapnik [54] to solve supervised classification and regression problems. In general, SVMs select the decision boundary from an infinite number of potential ones, leaving the greatest margin between the closest data points to the hyperplane, which are referred to as "support vectors" [29,55]. SVM employ a kernel function to transform the training data into higher dimensional feature space for non-linear classification problems [29]. In this regard, SVM are considered to be a kernel method since kernel functions are used to maximize the margin between classes. Therefore, the SVM have ability to delineate multi-modal classes in high dimensional feature spaces [56]. Previous studies have demonstrated the effectiveness of SVM for mapping land cover [57], especially in areas where training data is limited. However, SVM require more training time, especially if the dataset has many features. The SVM classifier available in the *e071* package [58] was used to classify the multivariate Landsat 8 spectral and Moran's I spatial dependence variables. We calibrated and fine-tuned the SVM classifier by changing the kernel functions (types) and regularization (penalty) parameter. In this study, the radial basis function was selected for classification since it had the best accuracy.

4. Results and Discussion

4.1. Classification Accuracy Assessment

We used reference pixels for accuracy assessment, which were independent from the training area pixels used for land cover classification. A total of 959 sample points were collected as reference data for each date based on a random sampling approach. Four measures of accuracy assessment namely, the producer's accuracy (accounting for errors of omission), user's accuracy (accounting for errors of commission), overall accuracy and overall kappa were computed to evaluate classification accuracy. The producer's accuracy or omission error show how well training set pixels were classified, while user's accuracy or commission error indicates the probability that a classified pixel actually represents that land cover class on the ground [20]. The overall accuracy gives the total number of correctly classified pixels divided by the total number of reference pixels, while the Kappa statistic incorporates the off diagonal elements of the error matrices and represents agreement obtained after removing the proportion of agreement that could be expected to occur by chance [59]. In addition, we used the accuracy assessment method proposed by Pontius and Millones [60] to assess classification accuracy.

This method divides the disagreements between classification and reference into quantity disagreement and allocation disagreement.

Table 3 shows the summary land cover accuracy assessment results for the RF, SVM, and DT classifiers. The overall classification accuracy for the RF, SVM, and DT classifiers were respectively, 80%, 76%, and 65% with Kappa statistics of 76%, 70%, and 58% (Table 3). In this study, the overall classification accuracy for the RF classifier is 4% and 15% higher than that of SVM and DT classifiers, respectively. These results are in agreement with other studies, which have noted relatively good accuracy from the RF classifier [31]. The individual land cover class accuracies are generally high for the RF classifier, with the exception of grassland that has a low user's accuracy of 53%. The SVM classifier have high individual land cover class accuracies for agriculture, others and water classes, while closed woodland, open woodland and grassland classes have lower individual class accuracies than the RF classifier (Table 3). Note that, individual class accuracy trend follows those of RF classifier, with the least accurate class being the grassland class, which exhibited low user's accuracy (Table 3). Although the DT classifier achieved the lowest overall classification accuracy, individual class accuracies, are relatively high for the closed woodland class. However, open woodland, grassland, agriculture and others classes have low class accuracies.

Figure 3 shows the analysis of classification errors in terms of quantity and allocation disagreements. The majority of classification errors for all classifiers are derived from allocation disagreement, which ranges from 14% to 19%. However, quantity disagreement for the DT classifier is very high (18%) compared to 5% and 6%, respectively, for SVM and RF classifiers (Figure 3). The quantity disagreement for the RF classifier is slightly higher than the SVM classifier despite the higher overall accuracy for the former (Table 3 and Figure 3). This is because the RF classifier for the grassland class is not stable as shown by a producer's accuracy of 70% *versus* a user's accuracy of 53% (Table 3). Although the quantity disagreement for SVM classifier is low (only 5%), the allocation disagreement is relatively high (19%) (Figure 3). This is attributed to low individual accuracies in open woodlands and grassland, which exhibits high commission error. For the DT classifier, both the quantity and allocation disagreements are high (Figure 3) because of low individual class accuracies, particularly for open woodland and grassland classes. While the single DT classifier performed poorly compared to RF and SVM classifiers, bagging DT algorithm which is used to improve DT was not tested in this study.

Figure 4 shows that RF, SVM and DT classifiers identified relatively small closed woodland areas. However, conspicuous differences are observed in the open woodland and non-woodland areas. Generally, RF classifier produced a relatively modest classification of the open woodland areas, while the SVM classifier overestimated the open woodland areas. The DT classifier on the other hand underestimated the open woodland areas. To visualize the differences in the classified areas more clearly, we extracted six subset images that show Landsat 8 images acquired on 6 June 25 August and 28 October 2013, as well as RF, SVM, and DT classified land cover maps at location A (Figures 4a and 5). Location A is a typically open woodland, subsistence agriculture and settlement area. As can be observed in the subset images of location A, the RF classifier extracted the open woodland, agriculture and other areas correctly, whereas SVM classifier overestimated the open woodland areas. Although the DT classifier managed to extract the closed woodland area, it however extremely underestimated the open woodland areas. Moreover, the DT classifier mislabeled open woodland and agriculture areas

as “others” class (Figure 5). The lower classification accuracy for the DT classifier, particularly for the open woodland and grassland areas is due to the small training pixel sample size used to construct the trees [47]. The DT and SVM classifiers results in lower classification accuracy for the open woodland and grassland areas because these classifiers fail to deal with inter-class variability differences caused by phenological changes (Figure 5). Furthermore, it should be noted that grassland areas had low accuracy compared to agriculture areas despite the fact that the two land cover classes are spectrally similar. This is because grassland areas are composed of small and fragmented patches that are difficult to discriminate, while agriculture areas are composed of large and homogenous patches that are relatively easy to classify.

Table 3. Summary of land cover classification accuracies (%): random forests (RF), support vector machines (SVM), and decision trees (DT).

Class	RF		SVM		DT	
	Producer’s Accuracy	User’s Accuracy	Producer’s Accuracy	User’s Accuracy	Producer’s Accuracy	User’s Accuracy
Closed woodland	81	83	64	77	74	93
Open woodland	70	78	63	67	65	43
Grassland	70	53	62	50	45	59
Agriculture	88	78	94	82	75	38
Others	85	97	91	93	54	86
Water	100	100	100	100	100	100
Overall Accuracy	80		76		65	
Overall Kappa	76		70		58	

Figure 3. Quantity and allocation disagreements for RF, SVM, and DT classifiers.

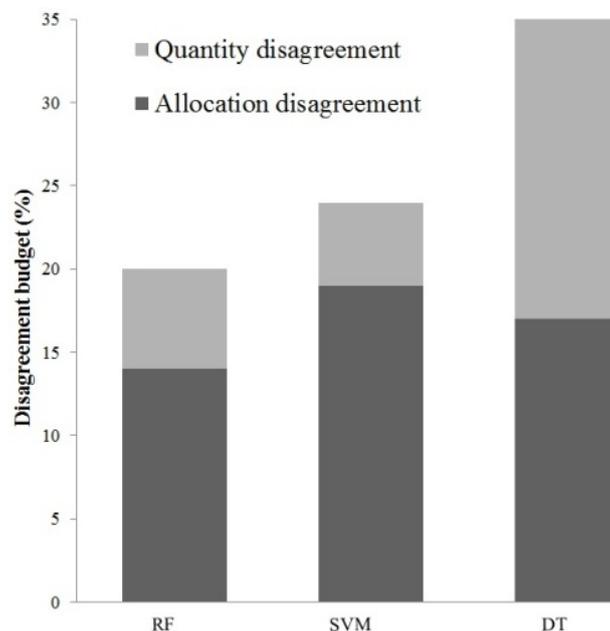


Figure 4. (a) Landsat 8 image in bands 6, 5, 4 (R,G,B) acquired on 6 June 2013 (note the square inset shows location A, while the red and black points shows validation and training areas, respectively); and land cover maps produced using (b) RF; (c) SVM; and (d) DT classifiers.

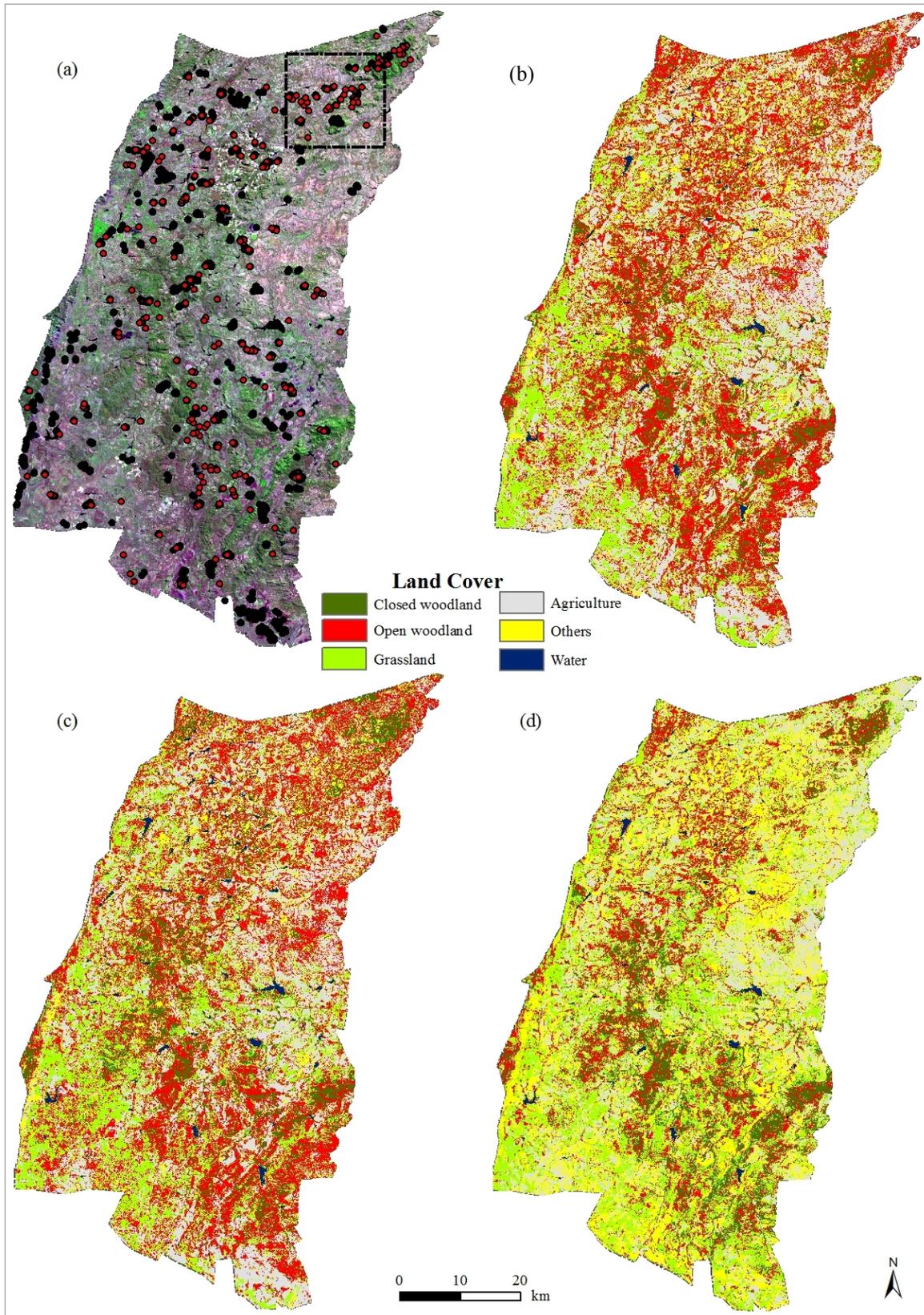
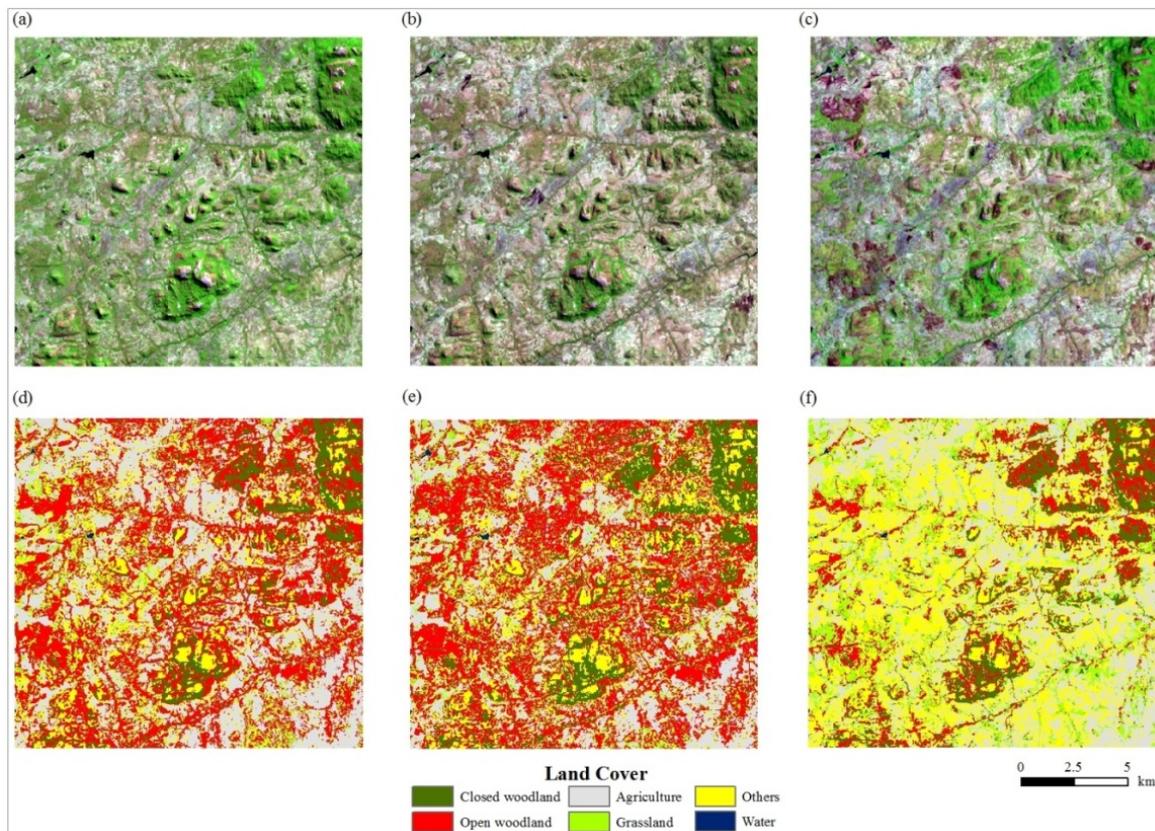


Figure 5. Subsets of Landsat 8 imagery in bands 6, 5, 4 (R,G,B) for location A acquired on (a) 6 June; (b) 25 August; and (c) 28 October; and subsets of land cover maps produced from (d) RF; (e) SVM; and (f) DT classifiers.



4.2. Performance of RF Classifier for Woodland Cover Classification

In order to gain deeper insights into closed and open woodland classification, we analyzed the relative importance of the contribution of the ten most important variables (Figures 6 and 7). The greatest contributions (with a mean decrease accuracy above 20) for the closed woodland class are derived from bands 4 (acquired on 6 June), 5 (acquired on 25 August), 4 (acquired on 19 April), and 5 (acquired on 28 October). This shows that multi-seasonal bands used in this study improved woodland cover classification given the complex seasonal behavior of vegetation in the study area. Note that when only post-rainy season imagery is used, closed and open woodland classes as well as the grassland class have the same spectral reflectance (Figure 5a). This is because the increase in greenness during the rainy and post-rainy season is associated with woodland canopy and grassland cover, which are at the peak of their phenological cycle [61]. In contrast, grassland areas appear as bare ground during the dry season peak in August (Figure 5b). However, the increase in greenness is attributed to woodland cover canopy leaf-on during the early growing season in late October (Figure 5c), while grassland cover is still in senesced state [61]. Therefore, the use of late dry season imagery, especially band 5 (acquired on 28 October) improved closed woodland cover mapping. This is supported by Yang and Prince [14] who stated that vegetation classification (e.g., scrubland and Miombo woodlands) is more effective if Landsat data is acquired when trees are in leaf and the grass layer is senescing. Furthermore, band 5 (acquired on 19 April), band 7 (acquired on 6 June), Moran's I

band 4 (6 June), band 5 (acquired on 6 June), band 4 (acquired on 28 October), and band 7 (acquired on 28 October) improved classification accuracy given that it had a relative importance contribution with a mean decrease accuracy between 15 and 20. The red (4), near infrared (5), and shortwave infrared (7) bands were the most important variables. In addition, Moran’s I band 4 also provided significant differentiation of the closed woodland and other land cover classes.

Figure 6. RF variable importance measures for the closed woodland class based on mean decrease accuracy.

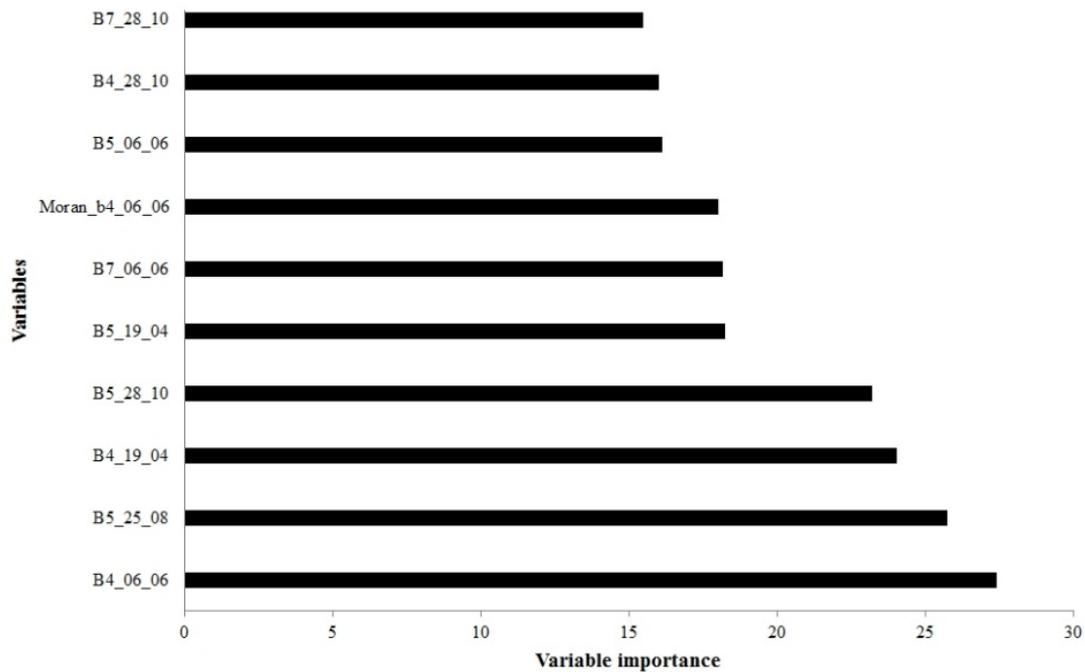
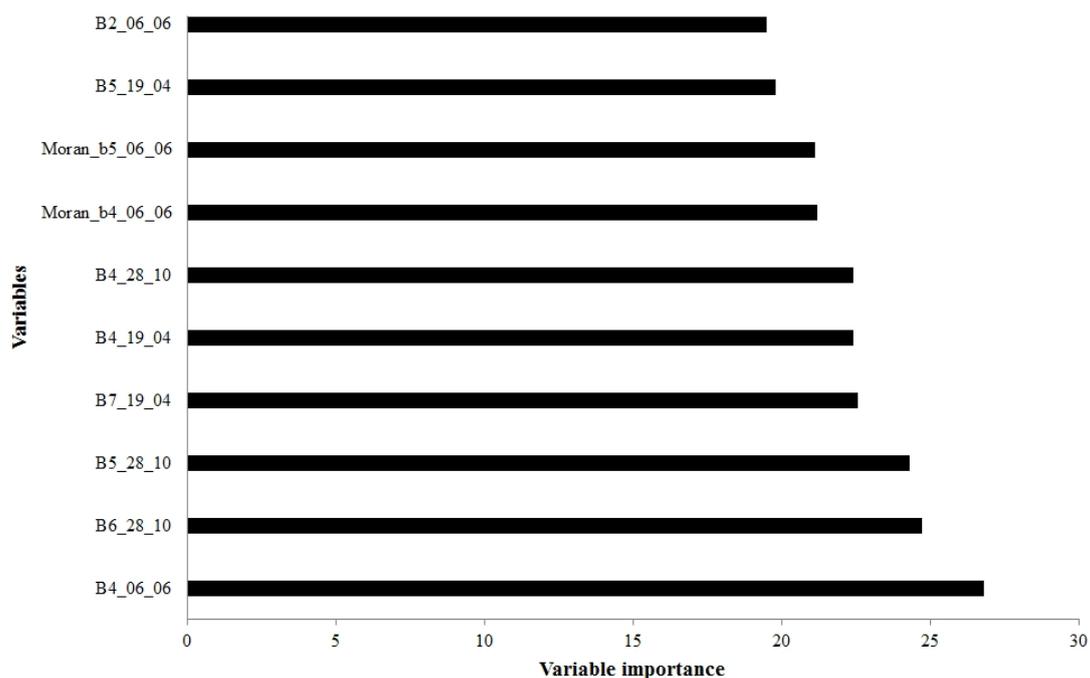


Figure 7. RF variable importance measures for the open woodland class based on mean decrease accuracy.



For the open woodland class (Figure 7), the greatest contributions (with a mean decrease accuracy above 20) are derived from bands 4 (6 June), 6 (28 October), 5 (28 October), 7 (19 April), 4 (19 April), 4 (28 October), Moran's I band 4 (6 June), and Moran's I band 5 (6 June). However, band 5 (19 April), band 2 (6 June) have a relative importance contribution with a mean decrease accuracy between 19 and 20. The red (4), near infrared (5) and shortwave infrared (6 and 7) bands were the most important variables. As observed in the classification of the closed woodland, multi-seasonal bands also significantly improved the open woodland classification. For example, the combination of multi-seasonal bands and the post-rainy season Moran's I bands 4 and 5 also provided significant contribution for the classification of the open woodland class. The results show the effectiveness of the RF classifier since one can evaluate the contribution of different spectral and spatial variables during the classification.

5. Conclusions

The objective of this study was to evaluate the performance of RF, SVM and DT classifiers for the classification of woodland and non-woodland cover in the Miombo ecosystem in Zimbabwe. We used multivariate Landsat 8 spectral and spatial dependence variables for classification. The results show that the RF classifier had a classification accuracy of 80% (with a Kappa statistic of 76%), while SVM and DT classifiers had 76% and 65% (with Kappa statistics of 70% and 58%). The RF classifier significantly increased the classification accuracy of both the SVM and DT classifiers by 4% and 15%, respectively.

The RF importance measures showed that multivariate spectral and spatial variables provide the greatest influence on class-separability in the study area. In addition, the red (4), near infrared (5) and shortwave infrared (7) bands were important for the classification of closed and open woodland classes. The RF classifier discriminated closed woodland, open woodland, and non-woodland classes better than other classifiers. While the results show great promise of machine learning classifiers for classifying woodland cover in the study area, more studies are needed in other Miombo ecosystem in order to improve classification accuracy. For example, other spatial variables, such as digital elevation model (DEM), precipitation, and fire occurrence, can be included to improve woodland cover mapping in the Miombo ecosystem.

In this study, freely available Landsat 8 images acquired in 2013 were used to map woodland cover in the Miombo ecosystem. This is important given the lack of woodland cover information in the region. Furthermore, this study has shown that multivariate Landsat 8 images can be used to improve woodland cover mapping in the Miombo ecosystem. Last but not least, Landsat 8 sensor and other upcoming satellite sensors such as Sentinel 2 are opening a new era for mapping and monitoring woodland cover changes at the landscape scale.

Acknowledgments

We would like to thank the three anonymous reviewers who provided useful comments, which were incorporated into this manuscript. The reviewers significantly improved the quality of this paper.

Author Contributions

Courage Kamusoko was responsible for designing and conducting the study. He evaluated the machine learning classifiers and drafted major sections of the article. Jonah Gamba was responsible for overseeing accuracy assessment and revising the manuscript. He also contributed to the major sections of the article. Hitomi Murakami was responsible for providing advice for the overall study.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Frost, P.G.H. *Coupling Land Use, Land-Cover Changes and Ecosystem Processes in Miombo Woodlands*; Report for the Period April 1999–November 1999 on UVA Subcontract under NASA Grant No. NAG5–6384; Institute of Environmental Studies, University of Zimbabwe: Harare, Zimbabwe, 1999.
2. Walker, S.M.; Desanker, P.V. The impact of land use on soil carbon in Miombo woodlands of Malawi. *For. Ecol. Manag.* **2004**, *203*, 345–360.
3. Chidumayo, E. Using natural fertilizers in Miombo woodlands. *Issues Afr. Biodivers.* **1999**, *2*, 1–8.
4. Forestry Commission. *State of Forest Genetic Resources in Zimbabwe 2002–2011: A Country Report*; Ministry of Environment and Natural Resources Management: Harare, Zimbabwe, 2011.
5. Gambiza, J.; Bond, W.; Frost, P.G.H.; Higgins, S. Special section: Land use options in dry tropical woodland ecosystem in Zimbabwe. A simulation model of Miombo woodland dynamics under different management regimes. *Ecol. Econ.* **2000**, *33*, 353–368.
6. Weng, Q. Land use change analysis in the Zhujiang Delta of China using satellite remote sensing, GIS and stochastic modelling. *J. Environ. Manag.* **2002**, *64*, 273–284.
7. Woodcock, C.E.; Allen, R.; Anderson, M.; Belward, A.; Bindschadler, R.; Cohen, W.; Gao, F.; Goward, S.N.; Helder, D.; Helmer, E.; *et al.* Free access to Landsat imagery. *Science* **2008**, *320*, 1011.
8. Foody, G.M.; Hill, R.A. Classification of tropical forest classes from Landsat TM data. *Int. J. Remote Sens.* **1996**, *17*, 2353–2367.
9. Lu, D.; Mausel, P.; Brondizio, E.; Moran E. Classification of successional forest stages in the Brazilian Amazon basin. *For. Ecol. Manag.* **2003**, *181*, 301–312.
10. Ribeiro, N.; Cumbana, M.; Mamuky, F.; Chauque, A. Remote sensing of biomass in the Miombo woodlands of Southern Africa: Opportunities and limitations for research. In *Remote Sensing of Biomass: Principles and Applications*; Fatoyinbo, L., Ed.; InTech: Rijeka, Croatia, 2012; pp. 77–98.
11. Chidumayo, E.; Okali, D.; Kowero, G.; Larwanou, M. Climate Change and African Forest and Wildlife Resources. African Forest Forum, 2011. Available online: http://www.afforum.org/index.php?option=com_docman&task=doc_download&gid=46&Itemid (accessed on 3 April 2014).

12. Frost, P. The ecology of Miombo woodlands. In *The Miombo in Transition: Woodlands and Welfare in Africa*; Campbell, B., Ed.; Center for International Forestry Research (CIFOR): Bogor, Indonesia, 1996; pp. 11–57.
13. Sedano, F.; Gong, P.; Ferrao, M. Land cover assessment with MODIS imagery in southern African Miombo ecosystems. *Remote Sens. Environ.* **2005**, *98*, 429–441.
14. Yang, J.; Prince, S.D. Remote sensing of savanna vegetation changes in Eastern Zambia 1972–1989. *Int. J. Remote Sens.* **2000**, *21*, 301–322.
15. Petit, C.; Scudder, T.; Lambin, E. Quantifying processes of land-cover change by remote sensing: Resettlement and rapid land-cover changes in southern-eastern Zambia. *Int. J. Remote Sens.* **2001**, *22*, 3435–3456.
16. Kamusoko, C.; Aniya, M. Hybrid classification of Landsat data and GIS for land use/cover change analysis of the Bindura district, Zimbabwe. *Int. J. Remote Sens.* **2009**, *30*, 97–115.
17. Koller, R.; Samimi, C. Deforestation in the Miombo woodlands: A pixel-based semi-automated change detection method. *Int. J. Remote Sens.* **2011**, *32*, 7631–7649.
18. Chidumayo, E.N.; Gumbo, D.J. *The Dry Forests and Woodlands of Africa: Managing for Products and Services*; Earthscan: London, UK, 2010.
19. Settle, J.J.; Drake, N.A. Linear mixing and the estimation of ground cover proportions. *Int. J. Remote Sens.* **1993**, *14*, 1159–1177.
20. Jensen, J.R. *Remote Sensing of the Environment: An Earth Resource Perspective*; Pearson Prentice Hall: Upper Saddle River, NJ, USA, 2007.
21. Asner, G.P.; Knapp, D.E.; Cooper, A.N.; Bustamante, M.; Olander, L.P. Ecosystem structure throughout the Brazilian Amazon from Landsat observations and automated spectral unmixing. *Earth Interact.* **2005**, *9*, 1–30.
22. Huang, C.; Townshend, J.R.G. A stepwise regression tree for nonlinear approximation: Applications to estimating subpixel land cover. *Int. J. Remote Sens.* **2003**, *24*, 75–90.
23. Walton, J.T. Subpixel urban land cover estimation: Comparing cubist, random forests, and support vector regression. *Photogramm. Eng. Remote Sens.* **2008**, *74*, 1213–1222.
24. Borel, C.C.; Gerstl, S.A. Nonlinear spectral mixing models for vegetative and soil surfaces. *Remote Sens. Environ.* **1994**, *47*, 403–416.
25. Ray, T.W.; Murray, B.C. Nonlinear spectral mixing in desert vegetation. *Remote Sens. Environ.* **1996**, *55*, 59–64.
26. Theseira, M.A.; Thomas, G.; Sannier, C.A.D. An evaluation of spectral mixture modelling applied to a semi-arid environment. *Int. J. Remote Sens.* **2002**, *23*, 687–700.
27. Lu, D.; Weng, Q. A survey of image classification methods and techniques for improving classification performance. *Int. J. Remote Sens.* **2007**, *28*, 823–870.
28. Seto, K.C.; Liu, W. Comparing ARTMAP neural network with the maximum-likelihood classifier for detecting urban change. *Photogramm. Eng. Remote Sens.* **2003**, *69*, 981–990.
29. Griffiths, P.; Hostert, P.; Gruebner, O.; van der Linden, S. Mapping megacity growth with multi-sensor data. *Remote Sens. Environ.* **2010**, *114*, 426–439.
30. Knorn, J.; Rabe, A.; Radeloff, V.C.; Kuemmerle, T.; Kozak, J.; Hostert P. Land cover mapping of large areas using chain classification of neighboring Landsat satellite images. *Remote Sens. Environ.* **2009**, *113*, 957–964.

31. Mellor, A.; Haywood, A.; Stone, C.; Jones, S. The performance of random forests in an operational setting for large area sclerophyll forest classification. *Remote Sens.* **2013**, *5*, 2838–2856.
32. Rodriguez-Galiano, V.F.; Chica-Olmo, M.; Abarca-Hernandez, F.; Atkinson, P.M.; Jeganathan, C. Random Forest classification of Mediterranean land cover using multi-seasonal imagery and multi-seasonal texture. *Remote Sens. Environ.* **2012**, *121*, 93–107.
33. Grinand, C.; Rakotomalala, F.; Gond, V.; Vaudry, R.; Bernoux, M.; Vieilledent, G. Estimating deforestation in tropical humid and dry forests in Madagascar from 2000 to 2010 using multi-date Landsat satellite images and the random forests classifier. *Remote Sens. Environ.* **2013**, *139*, 68–80.
34. ZimStats (Zimbabwe National Statistics Agency). *Census 2012: Preliminary Report*; ZimStats: Harare, Zimbabwe, 2012.
35. NASA. Landsat 8 Overview; 2013. Available online: http://www.nasa.gov/mission_pages/landsat/overview/index.html (accessed on 25 May 2013).
36. USGS. Landsat 8 Data Product Information, 2013. Available online: http://landsat.usgs.gov/LDCM_DataProduct.php (accessed on 25 May 2013).
37. Song, C.; Woodcock, C.E.; Seto, K.C.; Lenney, M.P.; Macomber, S.A. Classification and change detection using Landsat TM data: When and how to correct atmospheric effects? *Remote Sens. Environ.* **2001**, *75*, 230–244.
38. Chang, K. *Introduction to Geographic Information Systems*; McGraw-Hill: Boston, MA, USA, 2010.
39. Google Earth 7.1.2.2041. Available online: <http://www.earth.google.com> (accessed on 2 November 2013).
40. R Development Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. 2005. Available online: http://r-project.kr/sites/default/files/2%EA%B0%95%EA%B0%95%EC%A2%8C%EC%86%8C%EA%B0%9C_%EC%8B%A0%EC%A2%85%ED%99%94.pdf (accessed on 3 April 2014).
41. Tso, B.; Mather, P. *Classification Methods for Remotely Sensed Data*; Taylor and Francis: London, UK, 2001.
42. Miao, X.; Heaton, J.S.; Zheng, S.; Charlet, D.A.; Liu, H. Applying tree-based ensemble algorithms to the classification of ecological zones using multi-temporal multi-source remote-sensing data. *Int. J. Remote Sens.* **2012**, *33*, 1823–1849.
43. Mather, P.M.; Koch, M. *Computer Processing of Remotely-Sensed Images: An Introduction*; Wiley-Blackwell: Chichester, UK, 2011.
44. Breiman, L. *Classification and Regression Trees*; Chapman & Hall: New York, NY, USA, 1993.
45. Quinlan, J.R. *C4.5: Programs for Machine Learning*; Morgan Kaufmann Publishers: San Mateo, CA, USA, 1993.
46. Pal, M.; Mather, P.M. An assessment of the effectiveness of decision tree methods for land cover classification. *Remote Sens. Environ.* **2003**, *86*, 554–565.
47. Sesnie, S.E.; Gessler, P.E.; Finegan, B.; Thessler, S. Integrating Landsat TM and SRTM-DEM derived variables with decision trees for habitat classification and change detection in complex neotropical environments. *Remote Sens. Environ.* **2008**, *112*, 2145–2159.
48. Therneau, T.M.; Atkinson, B.; Ripley, B. Rpart: Recursive Partitioning. R Package Version 2010. Available online: <http://cran.r-project.org/package=rpart> (accessed on 3 April 2013).

49. Liaw, A.; Wiener, M. Classification and regression by randomForest. *R News* **2002**, *2*, 18–22.
50. Breiman, L. Random forests. *Mach Learn.* **2001**, *45*, 5–32.
51. Cortes, C.; Vapnik, V. Support-vector networks. *Mach Learn.* **1995**, *20*, 273–297.
52. Huang, Z.; Lees, B.G. Combining non-parametric models for multisource predictive forest mapping. *Photogramm. Eng. Remote Sens.* **2004**, *70*, 415–426.
53. Boser, B.E.; Guyon, I.M.; Vapnik, V.N. A training algorithm for optimal margin classifiers. In *Proceedings of the Fifth Annual Workshop on Computational Learning Theory*; Association for Computing Machinery (ACM): New York, NY, USA; 1992. pp. 144–152. Available online: <http://dl.acm.org/citation.cfm?id=130401> (accessed on 22 February 2014).
54. Vapnik, V.N. *The Nature of Statistical Learning Theory*; Springer: New York, NY, USA, 2000.
55. Pal, M.; Mather, P.M. Support vector machines for classification in remote sensing. *Int. J. Remote Sens.* **2005**, *26*, 1007–1011.
56. Zhang, R.; Ma, J. An improved SVM method P-SVM for classification of remotely sensed data. *Int. J. Remote Sens.* **2008**, *29*, 6029–6036.
57. Nemmour, H.; Chibani, Y. Multiple support vector machines for land cover change detection: An application for mapping urban extensions. *ISPRS J. Photogramm. Remote Sens.* **2006**, *61*, 125–133.
58. Hornik, K.; Meyer, D.; Karatzoglou, A. Support vector machines in R. *J. Stat. Softw.* **2006**, *15*, 1–28.
59. Congalton, R.G.; Green, K. *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*; CRC Press/Taylor & Francis: Boca Raton, FL, USA, 2009.
60. Pontius, R.G., Jr.; Millones, M. Death to Kappa: Birth of quantity disagreement and allocation disagreement for accuracy assessment. *Int. J. Remote Sens.* **2011**, *32*, 4407–4429.
61. Chidumayo, E. Climate and phenology of savanna vegetation in Southern Africa. *J. Veg. Sci.* **2001**, *12*, 347–354.

© 2014 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).