

Transferability of decision trees for land cover classification in a heterogeneous area

Julie Verhulp¹, Adriaan Van Niekerk^{1,2}

¹Department of Geography and Environmental Studies, Stellenbosch University, Stellenbosch, South Africa,
Julie.Verhulp@drdlr.gov.za

²School of Plant Biology, University of Western Australia, Perth, Australia

DOI: <http://dx.doi.org/10.4314/sajg.v6i1.3>

Abstract

As the value of accurate land cover becomes more apparent, methods to decrease the costs associated with supervised land cover mapping are investigated. One such method is to use training data captured in one scene and apply it to a different scene through a process known as signature extension. This paper attempts to derive classification rules from training data of four Landsat-8 scenes by using the classification and regression tree (CART) implementation of the decision tree algorithm. The transferability of the ruleset was evaluated by classifying two adjacent scenes. The classification of the four mosaicked scenes achieved an overall accuracy of 80.6%, while the two adjacent scenes achieved 61.4% and 83.7% respectively. The low accuracy of the first adjacent scene can be ascribed to a misclassification of graminoids, urban and bare areas, attributed to the temporal changes of grasslands throughout the year. In an attempt to improve the results, a normalised difference vegetation index (NDVI) threshold was applied to each scene. This increased the accuracy of the first adjacent scene but decreased the accuracy of the second. We conclude that signature extension using CART is unreliable. However, simple rules can be added to improve the results.

1. Introduction

There are many applications of remotely sensed imagery, but the most common is undoubtedly land cover mapping (Gray and Song, 2013; Hu *et al.*, 2015). Moreover, the importance of accurate and up to date land cover and land use information is increasing as the significance of this information becomes recognised by the international scientific community (Rodriguez-Galiano and Chica-Olmo, 2012).

One method to produce land cover information is through the supervised classification of remotely sensed imagery. Generally, supervised classification takes place on a scene-by-scene basis (Gray and Song, 2013; Knorn *et al.*, 2009), which is problematic as the selection of training data can be time-consuming and expensive and requires the expertise of a skilled photo-interpreter (Knorn *et al.*, 2009; Richards and Jia, 2006). These factors could lower the quality and frequency of land cover products.

The costs associated with the collection of training data could be reduced if the training data are applied to a different area or time period (Giri, 2012). When confronted with a large area to be classified, one approach – initially developed in the mid-1970's – is signature extension, also

known as generalization (Pax-Lenney *et al.*, 2001). This method involves the use of spectral signatures created in one area and applied either to another scene, a different sensor or at a different time (Hu *et al.*, 2015; Knorn *et al.*, 2009). Spatial signature extension was initially found to be ineffectual, mainly due to poor radiometric calibration and normalisation between scenes (Olthof *et al.*, 2005). The success of signature extension was further hindered by spatial heterogeneity and phenological differences (McDermid *et al.*, 2005), specifically in a north-south direction, due to significant changes in vegetation (Laborte *et al.*, 2010).

Owing to variations in topography, phenology (Knorn *et al.*, 2009), the angle of the sun and atmospheric conditions (Hu *et al.*, 2015), accuracies of land cover maps generated through signature extension tend to decline by as much as 13% when classifying nearby scenes (Pax-Lenney *et al.*, 2001). This decline is, however, not yet fully understood as the patterns affecting it are complex (Pax-Lenney *et al.*, 2001). Using only two land cover classes as a method of monitoring forest change, Woodcock *et al.*, (2001) noted that spatial extension is possible, and in fact comparable to other methods, but only when used for nearby scenes. This observation supports the findings of other authors who noted that the success of land cover classification is hampered when an area is heterogeneous, as geographical complexity can have a negative effect on the spectral separability of classes (Rodriguez-Galiano and Chica-Olmo, 2012; Rodriguez-Galiano *et al.*, 2012) and can reduce classification accuracy (Okubo *et al.*, 2010). Spectral separability measures have been used to provide an indication of the potential accuracy of land cover classifications (Su *et al.*, 2010) in heterogeneous areas. For instance, Verhulp and Van Niekerk (2016) used spectral separability measures in four individual and mosaicked Landsat-8 scenes to evaluate the potential of signature extension for supervised land cover mapping. They noted that the overall spectral separability was substantially lower in highly heterogeneous scenes compared to less complex scenes. They also showed that the use of multi-temporal imagery can improve the separability of classes in heterogeneous areas as it better represents the phenological stages of vegetation (Rodriguez-Galiano *et al.*, 2012). The value of multi-temporal imagery for land cover classification was also demonstrated by Verhulp and Van Niekerk (2016) and Brown de Colstoun *et al.*, (2003). The inclusion of multi-temporal imagery could possibly improve the accuracies associated with signature extension.

An alternative to spatial classifier extension without the use of spectral signatures is to create a decision tree. A decision tree uses binary rules to classify pixels based on both spectral and ancillary information (Chuvieco and Huete, 2010). Each tree has a root node, a series of splits and terminal nodes known as leaves (Pal and Mather, 2001). The decision tree method is very popular in land cover classifications thanks to its flexibility, simplicity and ease of interpretation (Brown de Colstoun *et al.*, 2003). The accuracies obtained through decision trees are also similar to or better than other classification methods (Brown de Colstoun *et al.*, 2003; Zhai *et al.*, 2012).

Traditionally, the creation of the decision tree and the identification of the splitting criteria were based on expert interpreter's knowledge (known as expert rules) or on statistical approaches (Chuvieco and Huete, 2010). Recently algorithms have been developed to automatically generate

decision trees (Chuvieco and Huete, 2010). One such algorithm is classification and regression trees (CART). CART recursively splits the data into two nodes according to the independent variables (the spectral and ancillary data), until there is a consistency between the land cover classes. Since decision trees tend to overfit the data (i.e. produce trees with poor generalizability to other datasets), CART uses an independent dataset to test the classification and prune the tree to an optimal size, known as the best tree, which is a combination of predicted accuracy and complexity (Steinberg and Colla, 2001). Ideally, the best tree is “less complex yet has superior predictive capabilities” (Brown de Colstoun *et al.*, 2003: 317).

Despite the simple and transparent nature of decision trees, very little research has been conducted on the transferability of the resulting rules to different scenes. Wentz *et al.*, (2008) adapted expert rules, originally designed for the classification of Phoenix, Arizona, and applied them to Delhi, India. They achieved an overall accuracy (OC) of 80.0%, but noted that certain classes had been hardcoded by the expert system, resulting in 100% accuracy for those classes. Zhai *et al.*, (2012) used the C5.0 decision tree algorithm to classify 18 Landsat Enhanced Thematic Mapper (ETM) scenes, having used only a few of the scenes to develop rules. Using spectral data from one season, as well as normalised difference vegetation index (NDVI) and tasseled cap components, an accuracy of 78.87% was achieved. They concluded that it is possible to classify large areas using decision trees, and that sample selection in every scene is not necessary.

This paper aims to investigate the accuracy and robustness (transferability) of decision tree rules to classify a large, highly heterogeneous area into land cover classes. The potential to spatially transfer the rules to two adjacent Landsat-8 scenes is investigated and the results are interpreted in the context of finding cost-effective operational solutions for monitoring land cover in complex areas. The area selected for this study is particularly complex owing to the great variation in elevation, climate, environmental patterns and vegetation (Verhulp and Van Niekerk, 2016).

2. Study area

The study area is made up of six Landsat-8 scenes situated primarily in the Eastern Cape Province of South Africa (Figure 1). The scenes were acquired from the United States Geological Survey (USGS) archive and stretch from 22°20' E to 30° E, with four scenes positioned along the coastline and two positioned inland.

The western portion of the study area is separated by the east-west oriented Baviaanskloof Mountain range, with a maximum elevation of 2 130m. The eastern portion of the study area rises steadily from the coast to the Winterberge and Drakensberg, with a maximum elevation of 2 743m.

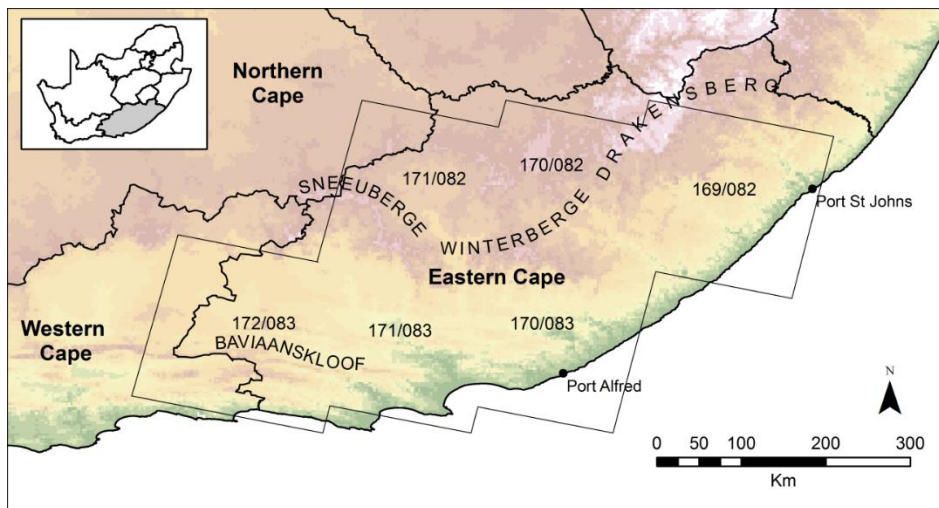


Figure 1. The location of the six Landsat-8 satellite scenes and their respective Path/Row reference number.

The climate of the study area varies greatly. The south-western coast up to Port Alfred experiences a maritime temperate climate and fairly even rainfall throughout the year. Further east towards Port St. Johns the climate becomes subtropical, with very high summer rainfall from October to March. An increase in elevation from Port St. Johns inland towards the Drakensberg Mountains is accompanied by a decrease in the average temperature and rainfall. The western interior is drier than the coast and experiences hot summers and cold winters (Schulze and Maharaj, 2006). In the central interior, north of the Winterberge Mountain range, the temperatures are much cooler, especially during winter.

The extent of the geographical area, along with the large variation in climate, has resulted in a highly diverse vegetation structure. According to Mucina and Rutherford (2006), the study area contains nine of the ten vegetation biomes found in South Africa. A biome is a “high-level hierarchical unit having similar vegetation structures exposed to similar microclimatic patterns, often linked to characteristic levels of disturbance such as grazing and fire” (Mucina and Rutherford, 2006: 32). The inland scenes are dominated by the grassland and nama-karoo biomes, while the coastal areas host a complex mixture of albaty thicket, fynbos, savanna, succulent karoo, azonal vegetation, Indian Ocean coastal belt and forests (Figure 2).

The study area was selected because the complex nature and heterogeneity of vegetation within the province would ensure that any methodologies developed should be applicable to less complex, more homogeneous areas.

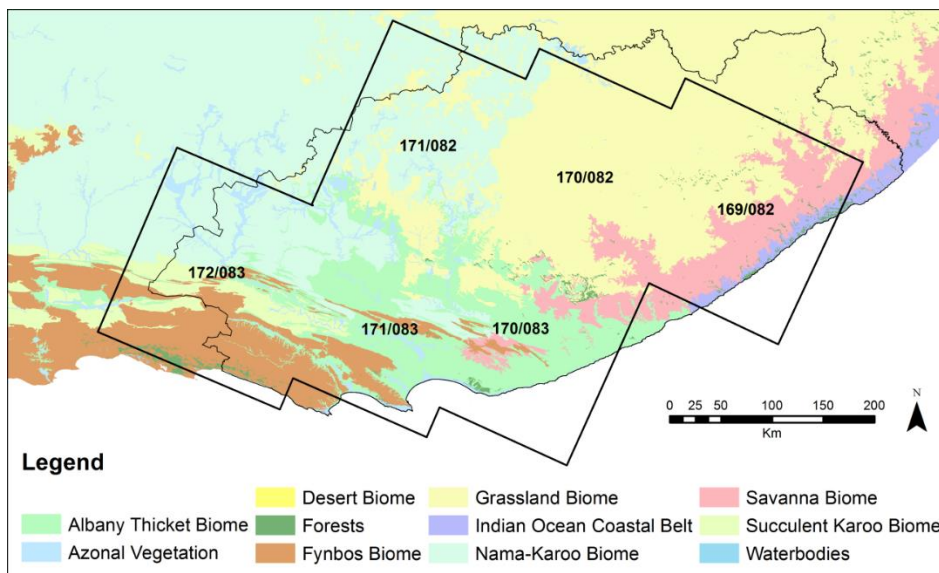


Figure 2. Distribution of biomes in the study area. Water bodies are not considered biomes.

3. Data collection and pre-processing

3.1 Satellite imagery

Landsat-8 imagery was collected for early spring of 2013 (August – September) and late summer of 2014 (February – April). The imagery was pre-processed by applying atmospheric and radiometric corrections. This is essential when multi-temporal and multi-scene images are utilised (Hu *et al.*, 2015), as it allows the user to compare the digital numbers across both space and time (Chuvieco and Huete, 2010). The corrections were implemented using the Atmospheric and Topographic Correction (ATCOR) procedure within the Interactive Data Language (IDL) environment. The 30m resolution imagery was pansharpended to 15m, while the two thermal bands were converted to a single surface temperature and resampled to 15m. The final result of the pre-processing workflow was a 12 bit, 15m TIFF image with seven bands: blue, green, red, near infrared (NIR), short wave infrared 1 (SWIR1), short wave infrared 2 (SWIR2) and surface temperature.

3.2 Training and reference data

The four coastal scenes were mosaicked and treated as a single entity. Training data for the coastal scenes were collected manually using a combination of Landsat-8, SPOT-5 and Google Earth imagery. A total of 1 464 polygons were collected for seven land cover classes (Table 1). The two inland scenes were kept separate and used to test classification via spatial extension of the developed ruleset. A total of 180 samples to be used as ground truthing (reference) data for scenes 170/082 and 171/082 were collected from SPOT-6 imagery (Table 1).

Table 1. The number of polygon samples used for each area.

| Class Name Number of Samples | Trees | Bushes and Shrubs | Forbs | Graminoids | Urban Built-up | Bare Areas | Water |
|------------------------------------|-------|----------------------|-------|------------|-------------------|---------------|-------|
| Coastal Scenes | 209 | 206 | 294 | 203 | 207 | 162 | 183 |
| 170/082 | 18 | 18 | 0 | 12 | 14 | 13 | 12 |
| 171/082 | 13 | 18 | 10 | 19 | 12 | 10 | 11 |

3.3 Auxiliary data

3.3.1 Principle component analysis and texture measures

Texture is characterised by the spatial variation of the spectral brightness within an image, and can be included to increase the classification accuracy (Rodriguez-Galiano and Chica-Olmo, 2012; Berberoglu *et al.*, 2007). A popular method used to determine texture is the grey-level co-occurrence matrix which evaluates the arrangement of grey values within a specified window in order to determine textural variation (Chuvieco and Huete, 2010; Berberoglu *et al.*, 2007; Coburn and Roberts, 2004). The use of textural features may, however, dramatically increase the dimension of the data as the calculation is applied to each image (Rodriguez-Galiano *et al.*, 2012). Additionally, many of these features may be redundant or highly correlated (Pacifci *et al.*, 2009).

Principal component analysis (PCA) is a feature selection procedure that results in the maximum amount of information for all bands condensed into a single band (Campbell and Wynne, 2012), made possible by the high correlation between bands. PCA is achieved through a linear transformation where the data axes are rotated in order to realign them with the maximum data variance (Giri, 2012). The first axis, which contains the maximum information in a single band, is known as the first principal component or PCA1. To avoid an increased complexity resulting from an increase in the number of bands and to reduce redundancy, texture is usually extracted from the PCA1 band (Rodriguez-Galiano *et al.*, 2012; Berberoglu *et al.*, 2007).

In this study, PCA1 was extracted from both the spring and summer bands, and contained 87% of the data variance. Six texture measures were applied to PCA1: homogeneity, second angular momentum, contrast, entropy, correlation and standard deviation. These six measures are generally accepted to be the most important measures for analysing images (Kayitakire *et al.*, 2006). A 3x3 window produces the largest classification accuracy as well as the highest Kappa value (Chet *et al.*, 2004), and was therefore used for calculating the texture measures.

3.3.2 Spectral indices

Spectral indices are formulas designed to extract quantitative information about each pixel (Chuvieco and Huete, 2010) and enhance latent or hidden information in the image data (Campbell and Wynne, 2012). Vegetation indices take advantage of the strong reflectance and absorption of chlorophyll in the NIR and red bands respectively (Chuvieco and Huete, 2010). The soil adjusted vegetation index (SAVI), second modified soil adjusted vegetation index (MSAVI-2) and the enhanced vegetation index (EVI) (Jensen, 2005) are all variations of the NDVI, and make use of this relationship. Water indices, such as the normalised difference water index (NDWI_MF)

proposed by McFeeters (1996) or the modified NDWI proposed by Xu (2006) (NDWI_XU), attempt to identify water and reduce shadow noise. Built-up and bare soil indices aim to emphasise non-vegetated features including urban areas, rock and bare soil. Examples of such indices include the enhanced built-up and bareness index (EBBI) (As-syakur, 2012), the index-based built-up index (IBI) (Xu, 2008), the soil index (Waqar, 2012), the normalised difference bareness index (NDBAI) (Waqar, 2012) and NDBAI-2 (Li *et al.*, 2014). These eleven indices were calculated for both the spring and summer image sets and included as additional input variables.

3.3.3 Ancillary data

The inclusion of topographic data as ancillary data in land cover mapping can improve classification accuracies (Ren, 2009). For this study, the 30m Shuttle Radar Topography Mission (SRTM) digital elevation model (DEM) covering the area of interest was obtained from the USGS. Slope gradient and aspect values were calculated and incorporated into the classification as additional features.

4. Data preparation and CART application

The four coastal scenes were mosaicked and treated as a single entity in order to produce a decision tree that incorporates scene-to-scene variations. No colour calibration, feathering or dodging parameters were selected during the mosaicking process.

All of the polygon training samples were converted to a series of vector points at a 15m sample distance, with an attribute representing the reference land cover class. This resulted in 625 939 sample points. An equally proportioned random subset of 89 238 sample points was created to produce the decision tree as CART performs best with an equal ratio (Campbell and Wynne, 2012).

For each point, the underlying pixel value of the Landsat-8 image features as well as the ancillary data was extracted. The image features consisted of the seven bands and eleven indices described in Section 3.3.2, as well as the six texture variables discussed in Section 3.3.1. The ancillary data consisted of slope gradient and aspect. The attribute data was exported for input into CART. Half of the points were used to build the initial tree, while the remaining points were used for pruning and obtaining the predicted classification accuracy.

CART often generates complex trees containing a large number of nodes that are not easily programmable or transferable. It is consequently common practice to limit the depth of the tree or the maximum number of nodes. However, such limitations generally have a negative effect on the resulting tree's predictive accuracy (Steinberg and Colla, 2001). Another approach is to manually prune the tree to the desired complexity. In this study, two tree complexity reduction methods were implemented and evaluated according to their resulting predicted accuracy. In the first scenario (Scenario 1), the number of nodes was limited during the tree-building phase. In the second scenario (Scenario 2), no limits were specified during the tree-building phase, but the tree was manually pruned. The predictive accuracies of these two scenarios were recorded for each tree-sized instance (from 900 to 20 nodes). The impact of merging different land cover classes was also

investigated by repeating the tree complexity reduction scenarios on different sets of classification schemes.

The classification rules derived from the decision tree with the smallest number of terminal nodes and the highest predictive accuracy was implemented using ERDAS Imagine's Knowledge Engineer Classifier. The ruleset was used to produce three land cover maps. The first map covered the four coastal scenes from which the rules had been derived, while the second and third maps were generated by implementing the ruleset on the two inland scenes. The purpose of the latter two maps was to test the ability of image extension as no training samples were collected in these areas.

An independent set of reference samples was used to determine the accuracy of the resulting maps. The points were randomly selected from the ground truthing polygons discussed in Section 3.2. As there were many clouds in scene 170/082, which could not be masked out, a cloud mask was created. No points inside the mask were used for the accuracy assessment, as this could negatively affect the result. A confusion matrix was used for the accuracy assessment. The user's and producer's accuracy, OC and the kappa index of agreement (KIA) coefficient were calculated from the confusion matrix.

5. Results

CART produced an initial decision tree with 975 terminal nodes. Allowing the tree to grow and then manually pruning it (Scenario 2) achieved a higher predicted accuracy than when a maximum number of nodes were imposed prior to tree-building (Scenario 1). However, even with manual pruning, a relatively large number of terminal nodes (42) were required to achieve a predicted accuracy above 80%. Pruning the tree further resulted in a sharp drop in predictive accuracy, mainly because of a poor distinction between trees and bushes on the one hand and the urban and bare classes on the other.

It is known that the accuracy of classifying urban features with the spatial resolution of Landsat-8 is low (Moran, 2010; Kahya *et al.*, 2010) and that urban and barren land features are easily confused due to the similarity in their spectral signature (Zhang, 2014). A third scenario (Scenario 3) was consequently tested in which the bare and urban classes were merged into a single class and its predictive accuracies tested (Figure 3). When manual pruning was applied to this simplified classification scheme, a predicted accuracy of 80.77% was achieved with only 21 terminal nodes. Limiting the number of nodes was to 21 during the tree building phase produced a predictive accuracy of 75.45%.

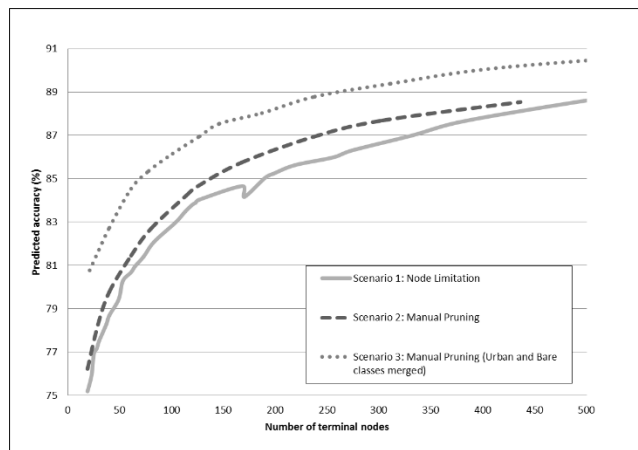


Figure 3. The predicted accuracy compared to the number of terminal nodes when the maximum number of nodes is specified prior to tree-building (Scenario 1), manual pruning is applied (Scenario 2), and when manual pruning was applied after the urban and bare class was combined (Scenario 3).

Under Scenario 3 the pruning process discarded the attributes that were not beneficial to the classification and only eleven attributes were retained. The attributes that remained included the blue, NIR, SWIR1 and thermal bands, as well as both water indices from the spring season; the blue and thermal bands and NDVI and EBBI indices from the summer season, and contrast as a texture measure. All other attributes were deemed unnecessary to achieve an accuracy over 80%. The decision tree in Figure 4 was applied to both the scenes from which the rules had been derived and two independent scenes to produce land cover maps. The accuracy of these maps is described in the following subsections.

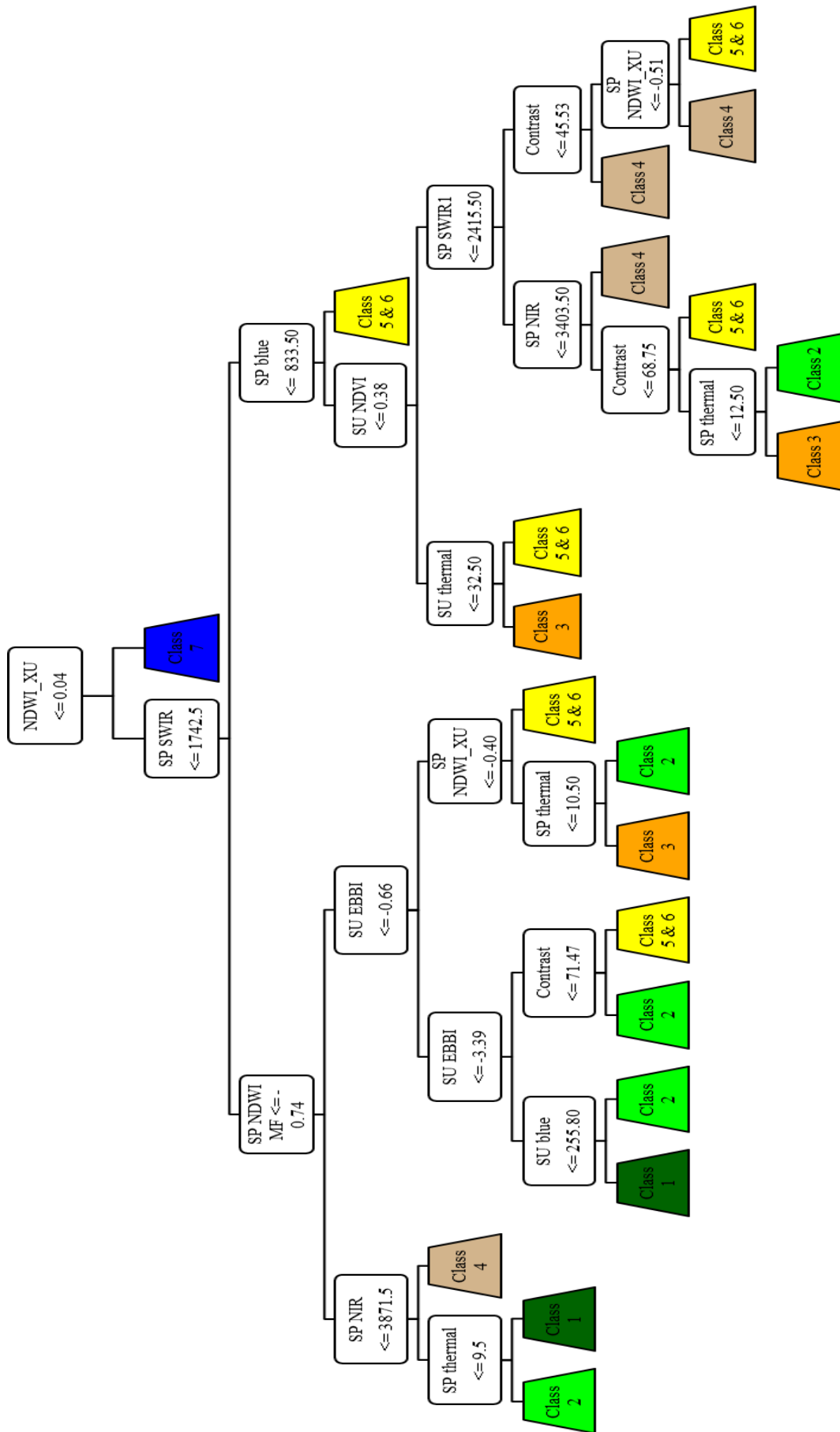


Figure 4. Decision tree with 21 nodes. Refer to Section 3.3.2 for an explanation of the class splitting abbreviations

The land cover classification of the mosaicked coastal scenes (Figure 5) achieved an OC of 80.58% with a KIA of 0.76. Table 2 shows the confusion matrix along with the user’s and producer’s accuracy of the resulting map. Bushes and trees could not be clearly distinguished from one another (66.9% and 64.3% producer’s and user’s accuracy respectively). Urban and bare areas were slightly over-classified, with 1 915 pixels (15%) of bushes, forbs and graminoids samples being incorrectly classified as urban and bare.

Table 2. Confusion matrix and the user’s and producer’s accuracy for the classification of the coastal scenes.

| Reference Data Classified Data | Trees | Bushes | Forbs | Graminoids | Urban and Bare | Water |
|-----------------------------------|-------|--------|-------|------------|----------------|-------|
| Trees | 4861 | 1225 | 14 | 80 | 22 | 0 |
| Bushes | 1175 | 4248 | 548 | 299 | 310 | 24 |
| Forbs | 23 | 440 | 4308 | 181 | 653 | 0 |
| Graminoids | 65 | 233 | 1193 | 5314 | 952 | 0 |
| Urban and Bare | 37 | 204 | 336 | 291 | 10353 | 93 |
| Water | 0 | 0 | 0 | 0 | 95 | 6145 |
| Prod Accuracy | 78.90 | 66.90 | 67.32 | 86.20 | 83.59 | 98.13 |
| User’s Accuracy | 78.38 | 64.32 | 76.86 | 68.51 | 91.51 | 98.48 |

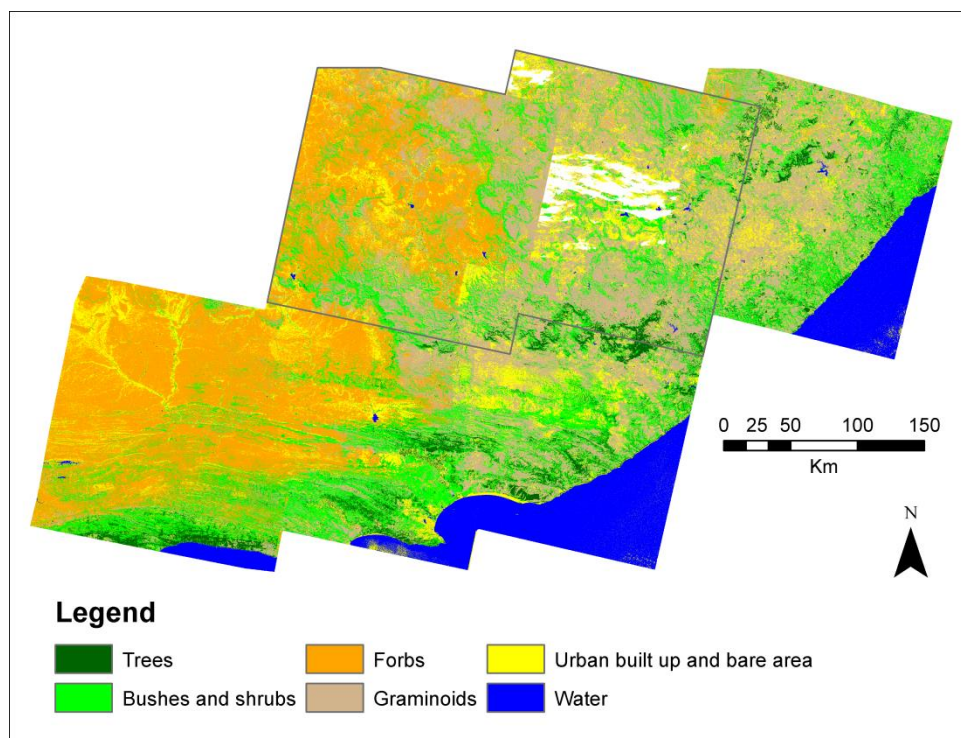


Figure 5. The land cover classification for the coastal scenes, as well as scenes 170/028 and 171/082 with the clouds masked out in white.

A random sample of 40 002 points was used to test the accuracy of this scene 170/082. A KIA of 0.55 and an OC of 64.1% were obtained for the scene. Table 3 shows the confusion matrix and user’s and producer’s accuracy for the scene. Trees have a high user’s accuracy (98.4%), but a large portion (22%) of trees was classified as bush. This scene did not contain any forbs, which meant that any classification thereof was inherently incorrect. Nearly 1 000 (2%) samples were incorrectly

classified as forbs, with 48% of them verified as being urban and bare, while 40% of them were meant to represent graminoids.

The urban and bare class was substantially over-classified, with 46% of pixels classified as such, when they were graminoids according to the reference data. Graminoids were also confused with bush, with a further 25% being classified as such.

In an attempt to reduce the misclassification of vegetation as bare areas and vice versa, an additional rule was manually added. The rule considers all pixels classified as urban and bare and applies a threshold to reclassify pixels with NDVI values of higher than 0.2 as graminoids. This improved the overall classification accuracy to 70.4% and the KIA to 0.63. The user's accuracy of the bare and urban class increased considerably (from 48.8% to 86.9%).

Table 3. Confusion matrix and the user's and producer's accuracy for the classification of scene 170/082, as well as user's and producer's accuracy for the classification before and after the addition of an NDVI threshold

| Reference Data | Trees | Bushes | Forbs | Graminoids | Urban and Bare | Water |
|--------------------------------|-------|--------|-------|------------|----------------|-------|
| Classified Data | | | | | | |
| Trees | 4820 | 57 | 0 | 6 | 7 | 7 |
| Bushes | 1436 | 4450 | 0 | 3388 | 68 | 44 |
| Forbs | 97 | 5 | 0 | 377 | 451 | 2 |
| Graminoids | 221 | 1835 | 0 | 5238 | 1530 | 0 |
| Urban and Bare | 92 | 320 | 0 | 4325 | 4577 | 57 |
| Water | 1 | 0 | 0 | 0 | 34 | 6557 |
| Original Classification | | | | | | |
| Prod Accuracy | 72.3 | 66.8 | | 39.3 | 68.7 | 98.4 |
| User's Accuracy | 98.4 | 47.4 | | 59.4 | 48.8 | 99.5 |
| NDVI threshold 0.20 | | | | | | |
| Prod Accuracy | 72.3 | 69.3 | | 77.4 | 52.3 | 98.5 |
| User's Accuracy | 99.0 | 64.1 | | 44.2 | 86.9 | 99.3 |

The classification of scene 171/082 was distinctly better than that of scene 170/082 and achieved an OC of 83.7% with a KIA of 0.80. Table 4 shows the confusion matrix and user's and producer's accuracy for the classification of the scene. Trees, bushes, urban and bare, and water classes all had accuracies above 80%. Graminoids had a very low producer's accuracy (30.2%), with 69.6% of graminoids classified as either forbs or urban and bare. The inclusion of an NDVI mask as in scene 170/082 only served to reduce the OC of the classification.

Table 4. Confusion matrix and the user's and producer's accuracy for the classification of scene 171/082

| Reference Data | Trees | Bushes | Forbs | Graminoids | Urban and Bare | Water |
|-----------------|-------|--------|-------|------------|----------------|-------|
| Classified Data | | | | | | |
| Trees | 5308 | 96 | 0 | 0 | 1 | 0 |
| Bushes | 691 | 5365 | 0 | 13 | 65 | 79 |
| Forbs | 0 | 88 | 5947 | 2127 | 739 | 0 |
| Graminoids | 0 | 351 | 21 | 1811 | 34 | 0 |
| Urban and Bare | 1 | 100 | 32 | 2049 | 9799 | 7 |
| Water | 0 | 0 | 0 | 0 | 147 | 5914 |
| Prod Accuracy | 88.4 | 89.4 | 99.1 | 30.2 | 90.7 | 98.6 |
| User's Accuracy | 98.2 | 86.4 | 66.8 | 81.7 | 81.7 | 97.6 |

6. Discussion

The classification accuracy achieved for the coastal scenes was very similar to the accuracy predicated during the creation of the decision tree (0.19% lower). Scene 171/082 performed distinctly better than scene 170/082, with an overall classification accuracy of 83.7% compared to 64.1%. The user's accuracy differences were noticeable for bushes (86.4% compared to 47.4%), graminoids (81.7% compared to 59.4%) and urban and bare (81.7% compared to 48.8%). Confusion between bushes and trees is common when only using spectral information (Geerling, 2007). The inclusion of height data (such as LiDAR data) could possibly assist with discriminating between the two, as they have different structures (Geerling, 2007). Research on the use of LiDAR data for discriminating these and similar land covers is recommended.

The inaccurate classification of pixels as forbs in scene 170/082 (none were present), is a known limitation of supervised classification, as spectral classes are forced to be classified in terms of operator defined classes (Campbell and Wynne, 2012). A possible solution is to amend the decision tree so that all forbs are classified as either grasslands or bare, however, operational issues must be considered. Furthermore, the analyst may not be aware of the absence of a certain class within a specific scene.

Scene 170/082 contains predominantly grassland, while scene 171/082 is primarily made up of nama-karoo vegetation, as well as grassland and Albany thicket. This large proportion of grassland in scene 170/082 may be the cause of its reduced accuracy. The producer's accuracy of graminoids was poor in both scenes (39.3% and 30.2% for 170/082 and 171/082 respectively), with over 30% being classified as urban and bare in each case.

The over-classification of bare and urban areas, prevalent in both inland scenes, could be resulting from the point sampling process. Urban areas often contain large quantities of vegetation (Zhang *et al.*, 2014), and this contamination of the training areas could have affected the classification result. Another reason for the extent of the misclassification of grassland and bare is that sensitive areas may be bare during the dry season, but contain vegetation during the wet season. This temporal complexity would then confuse the classifier when using dual season imagery. The grassland biome is particularly seasonal, with strong summer rainfall and droughts in winter (Mucina and Rutherford, 2006). Figure 6 shows the difference in vegetation between the wet and dry season, specifically the transformation from grassland to bare areas.



Figure 6. The substantial difference between the wet season (Parts (a) and (b)) (lush and green vegetation) and the dry season (Parts (c) and (d)).

The application of a masked NDVI or similar vegetation index to the urban and bare area at a specific season could reduce this misclassification and clarify the actual land cover class. A second mask, outlining the urban edge, could then separate the urban and bare class into two. The addition of other ancillary data, such as a biome or vegetation map, can also improve the classification, but may complicate the decision tree development.

7. Conclusion

This study evaluated the transferability of decision tree rules for land cover classification. A sample of 89 238 points was used to develop a decision tree ruleset. The information attributed to each point included the spectral information of the Landsat-8 images from two seasons, various indices, as well as elevation and texture information. The decision tree was pruned so as to reduce the complexity of the ruleset, while maintaining a predicted accuracy of above 80%. The ruleset was then applied to two adjacent scenes to test the transferability.

The results of this study provided new insight into the extent to which a decision tree ruleset can be transferred to adjacent scenes. The accuracy of scene 171/082 was 83.7%, while scene 170/082 only achieved an accuracy of 64.1%. The inclusion of an NDVI mask, however, improved the classification accuracy of scene 170/082 to 70.4%. Although the use of a decision tree via image extension for classification is possible, more insight into factors affecting the accuracy is needed, especially when complex, heterogeneous areas are involved. As noted by Pax-Lenney *et al.*, (2001), a single classification test of spatial extension is insufficient to draw concrete conclusions. This study showed that it is possible to transfer decision rules in complex areas, but that the accuracy varies depending on vegetation and distance from the original scene. Further research on the transferability of decision tree rules in complex, heterogeneous areas is needed; specifically on improving class specific accuracies and determining the optimal distance over which rules can be transferred.

References

- As-syakur, A.R., Adnyana, I.W.S., Arthana, I.W. and Nuarsa, I.W. 2012, 'Enhanced built-up and bareness index (EBBI) for mapping built-up and bare land in an urban area', *Remote Sensing*, vol. 4, pp. 2957-2970, viewed 29 May 2014, <<http://www.mdpi.com/2072-4292/4/10/2957>>
- Berberoglu, S., Curran, P.J., Lloyd, C.D. and Atkinson, P.M. 2007, 'Texture classification of Mediterranean land cover', *International Journal of Applied Earth Observation and Geoinformation*, vol. 9, no. 3, pp. 322-334.
- Brown de Colstoun, E.C., Story, M.H., Thompson, C., Commisso, K., Smith, T.G. and Irons, J.R. 2003, 'National park vegetation mapping using multitemporal Landsat 7 data and a decision tree classifier', *Remote Sensing of Environment*, vol. 85, pp. 316-327.
- Campbell, J.B. and Wynne, R.H. 2012, *Introduction to remote sensing*, 5th ed. New York: Guilford.
- Chen, D., Stow, D.A. and Gong, P. 2004, 'Examining the effect of spatial resolution and texture window size on classification accuracy: an urban environment case', *International Journal of Remote Sensing*, vol. 25, no. 11, pp. 2177-2192.
- Chuvieco, E. and Huete, A. 2010, *Fundamentals of satellite remote sensing*, Boca Raton: Taylor & Francis.
- Coburn, C.A. and Roberts, A.C.B. 2004, 'A multiscale texture analysis procedure for improved forest stand classification', *International Journal of Remote Sensing*, vol. 25, no. 20, pp. 4287-4308.
- Geerling, G.W. 2007, 'Classification of floodplain vegetation by data fusion of spectral (CASI) and LiDAR data', *International Journal of Remote Sensing*, vol. 28, no. 19, pp. 4263-4284.
- Giri, C.P. (ed) 2012, *Remote sensing of land use and land cover: Principles and applications*. Boca Raton: Taylor & Francis.
- Gray, J., and Song, C. 2013, 'Consistent classification of image time series with automatic adaptive signature generalization', *Remote Sensing of Environment*, vol 134, pp. 333-341.
- Hu, Y., Liu L., Caccetta P., and Juao Q. 2015, 'Landsat time-series land cover mapping with spectral signature extension method', *Journal of Remote Sensing*, vol. 19 no. 4, pp. 639-656.
- Jensen, J.R. 2005, *Introductory digital image processing: A remote sensing perspective*. 3rd ed. Upper Saddle River: Pearson Prentice Hall.
- Kahya, O., Bayram, B. and Reis, S. 2010, 'Land cover classification with an expert system approach using Landsat ETM imagery: a case study of Trabzon', *Environmental Monitoring and Assessment*, vol. 160, pp 431-438.
- Kayitakire, F., Hamel, C. and Defourny, P. 2006, 'Retrieving forest structure variables based on image texture analysis and IKONOS-2 imagery', *Remote Sensing of Environment*, vol. 102, pp. 390-401.
- Knorn, J., Rabe, A., Radeloff, V.C., Kuemmerle, T., Kozak, J., and Hostert P. 2009, 'Land cover mapping of large areas using chain classification of neighboring Landsat satellite images', *Remote Sensing of Environment*, vol. 113, no. 5, pp. 957-964.
- Laborte, A.G., Maunahan, A.A. and Hijmans, R.J. 2010, 'Spectral signature generalization and expansion can improve the accuracy of satellite image classification', *PLoS ONE*, vol. 5, no. 5, pp 1-9.
- Li, S. and Chen, X. 2014, 'A new bare-soil index for rapid mapping developing areas using Landsat 8 data', *Proceedings of international archives of the photogrammetry, remote sensing and spatial information sciences*, 4-16 May 2014, Suzhou.
- McDermid, G.J., Franklin, S.E. and LeDrew, E.F. 2005, 'Remote sensing for large-area habitat mapping', *Progress in Physical Geography*, vol. 29, no. 4, pp. 449-474.
- McFeeters, S.K. 1996, 'The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features', *International Journal of Remote Sensing*, vol. 17, pp. 1425-1432.
- Moran, E.F. 2010, 'Land cover classification in a complex urban-rural landscape with Quickbird imagery', *Photogrammetric Engineering and Remote Sensing*, vol. 76, no. 10, pp. 1159-1168.
- Mucina, L. and Rutherford, M.C. 2006. *The vegetation of South Africa, Lesotho and Swaziland*. Pretoria: South African National Biodiversity Institute.

- Okubo, S., Parikesit, Muhamad, D., Harashina, K., Takeuchi, K. and Umezaki, M. 2010, 'Land use/cover classification of a complex agricultural landscape using single-dated very high spatial resolution satellite-sensed imagery', *Canadian Journal of Remote Sensing*, vol. 36, no. 6, pp. 722-736.
- Olthof, I., Butson, C. and Fraser, R. 2005, 'Signature extension through space for northern landcover classification: A comparison of radiometric correction methods', *Remote Sensing of Environment*, vol. 95, pp 290-302.
- Pacifici, F., Chini, M. and Emery, W.J. 2009, 'A neural network approach using multi-scale textural metrics from very high-resolution panchromatic imagery for urban land-use classification', *Remote Sensing of Environment*, vol. 113, no. 6, pp. 1276-1292.
- Pal, M. and Mather, P.M. 2001, 'Decision Tree Based Classification of Remotely Sensed Data', *Proceedings of the 22nd Asian Conference on Remote Sensing, Centre for remote imaging, sensing and processing*. 5-9 November, Singapore,
- Pax-Lenney, M., Woodcock, C.E., Macomber, S.A., Gopal, S. and Song, C. 2001, 'Forest mapping with a generalized classifier and Landsat TM data', *Remote Sensing of the Environment*, vol. 77, no. 3, pp. 241-250.
- Ren, G., Zhu, A., Wang, W., Xiao, W., Huang, Y., Li, G., Li, D. and Zhu, J. 2009, 'A hierarchical approach coupled with coarse DEM information for improving the efficiency and accuracy of forest mapping over very rugged terrains', *Forest Ecology and Management*, vol. 258, pp. 26-34.
- Richards, J.A. and Jia X. 2006, *Remote Sensing Digital Image Analysis*. 4th ed. Berlin: Springer.
- Rodriguez-Galiano, V.F. and Chica-Olmo, M. 2012, 'Land cover change analysis of a Mediterranean area in Spain using different sources of data: multi-seasonal Landsat images, land surface temperature, digital terrain models and texture', *Applied Geography*, vol. 30, pp. 208-218.
- Rodriguez-Galiano, V.F., Chica-Olmo, M., Abarca-Hernandez, F., Atkinson, P.M. and Jeganathan, C. 2012. 'Random Forest classification of Mediterranean land cover using multi-seasonal imagery and multi-seasonal texture', *Remote Sensing of Environment*, vol. 121, pp. 93-107.
- Schulze, R.E. and Maharaj, M. 2006, 'Daily Mean Temperatures', in Schulze, R.E. (ed), *The South African Atlas of Climatology and Agrohydrology*. Pretoria, Water Research Commission.
- Steinberg, D. and Colla, P. 1995, *CART: Tree-Structured Non-Parametric Data Analysis*, Salford Systems, San Diego, viewed 27 June 2016, <http://docs.salford-systems.com/CART_Main.pdf>
- Su, H., Kanemasu, E.T., Ransom, M.D. and Yang, S. 1990, 'Separability of soils in tallgrass prairie using SPOT and DEM data', *Remote Sensing of Environment* vol. 33, pp. 157-163.
- Verhulp, J. and Van Niekerk, A. 2016, 'Effect of inter-image spectral variation on land cover separability in heterogeneous areas', *International Journal of Remote Sensing*, vol. 37, no. 7, pp. 1639–1657.
- Waqar, M.M., Mizra, J.F., Mumtaz, R. and Hussain, E. 2012, 'Development of new indices for extraction of built-up area & bare soil from Landsat data', *Open Access Scientific Reports*, vol. 1, no. 1, viewed 29 May 2014, <<http://omicsonline.org/scientific-reports/JGRS-SR136.pdf>>
- Wentz, E.A., Nelson, D., Rahman, A., Stefanov, W.L. and Roy, S.S. 2008, 'Expert system classification of urban land use/cover for Delhi, India', *International Journal of Remote Sensing*, vol. 29, no. 15, pp. 4405-4427.
- Woodcock, C.E., Macomber, S.A., Pax-Lenney, M. and Cohen, W.B. 2001, 'Monitoring Large Areas for Forest Change Using Landsat : Generalization across Space, Time and Landsat Sensors', *Remote Sensing of Environment*, vol. 78, pp. 194–203
- Xu, H. 2006, 'Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery', *International Journal of Remote Sensing*, vol. 27, pp 3025-3033.
- Xu, H. 2008, 'A new index for delineating built-up land features in satellite imagery', *International Journal of Remote Sensing*, vol. 29, pp. 4269-4276.

- Yuan, F., Bauer, M.E., Heinert, N.J. and Holden, G.R. 2005, 'Multi-level land cover mapping of the twin cities (Minnesota) metropolitan area with multi-seasonal Landsat TM/ETM+ data', *Geocarto International, vol. 20, no. 2*, pp. 5-13.
- Zhai, L., Sun, J., Sang, H., Yang, G. and Jia, Y. 2012, 'Large Area Land Cover Classification with Landsat ETM+ Images Based on Decision Tree'. *Proceedings of the XXII ISPRS Congress, International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 25 August – 1 September, Melbourne.*
- Zhang, J., Li, P. and Wang, J. 2014, 'Urban built-up area extraction from Landsat TM/ETM+ images using spectral information and multivariate texture', *Remote Sensing, vol. 6, no. 8*, pp. 7339–7359.