Status of Forests — Baker Creek, BC, Canada:

Classification of Land Cover Types Using Remotely Sensed Imagery

Simon A. Campbell

34982165
Department of Geographical Sciences
The University of British Columbia, Vancouver, BC, Canada

GEOB 373 - Introductory Remote Sensing
Final Project
April 7th, 2019

Abstract:

The British Columbia Ministry of Forests, Lands and Natural Resource Operations is interested in assessing the status of forests in an area near Baker Creek, British Columbia, Canada. This report is an effort to quantify the area associated with various land cover types — those being recent clear-cuts with no timber standing, older clear-cuts with vegetation re-growth and remaining forest using Landsat 8 OLI imagery provided by the United States Geological Survey (USGS). My methodology will consist of an (unsupervised) ISO cluster classification, followed by a (supervised) maximum likelihood classification (MLC) using the ESRI ArcGIS 10.6.1 platform. I will then conduct an accuracy assessment which provides information regarding the reliability and efficacy of my classification results.

My unsupervised classification results indicate that ~33% of the area is remaining forest, ~29% is clearcut, and ~33% is older clearcuts. Supervised classification results indicate that ~25%, 41%, and ~29% respectively. The accuracy assessment shows a kappa coefficient value of 0.72, indicating that our classification efforts were a moderate success.

Table of Contents:

1. Introduction:

The site under investigation covers an area of 3354 square kilometers, and is located about 90 kilometers North-West of Williams Lake, west of highway 97. The image occupies an area surrounding Baker Creek BC, with land cover primarily being forested areas under the influence of logging activities, small streams and various lakes. Dominant land cover types include recent clearcuts with little to no vegetation regrowth, older clearcuts with moderate to significant vegetation regrowth, remaining forests unaltered by logging activities, and various water bodies.

The image used in our analysis was captured by the LANDSAT 8 OLI satellite, bands 1-7, scene path/row: 48/23, taken November 14th, 2016. The Operational Land Imager (OLI) measures in the visible, near infrared, and short wave infrared radiation from the Earth's surface. The Landsat 8 mission — a collaboration between NASA and the U.S. Geological Survey, is working to monitor the changes in land cover and use are having profound consequences on resource management and the national and global economy (Roy et al., 2014). The image is a level 1 collection with a GeoTiff output format and was delivered as a 16-bit image. For the purposes of my analysis, I used a 8-bit unsigned resolution. A colour infrared (432;RGB) band combination was used to help identify and discern between land cover with different vegetation density. This band combination highlighted the contrast between areas with different vegetation density (displayed in red), and also between non-vegetated areas (displayed in grey/light blue). Water is easily distinguished from other land cover types (appears black). A histogram stretch (percent clip min/max = 0.5/0.5, nearest neighbour interpolation) was used to improve image quality.

To gain an idea of the general land cover types in the area, the satellite imagery was compared to ESRI base-map world imagery with boundaries and labels. Although the ESRI base-map was not as up to date as the Landsat imagery, it still aided greatly in the interpretation of clear cut areas, water bodies, roads and other features that were easier to identify with a higher resolution image. The difference in capture dates between these two images was noticed by recent clearcuts only showing in the Landsat image.

2. Analysis:

2.1. Unsupervised classification

Land Cover Type	Area (hectares)	Area (%)
Water Bodies	512	48.0
Remaining Forest	3517	32.9
Clearcuts	3109	29.1
Older Clearcuts	3531	33.0

Table 1. Land cover types and their respective areas. Values were determined from the *unsupervised* classification results. Area determined from pixel count for each class.

The unsupervised classification was able to clearly distinguish between waterbodies, clearcuts and remaining forests. However, older clearcuts with moderate to considerable vegetation regrowth were difficult to classify (fig. 2). This may be due to variation in the amount of vegetation in these areas. There were also some recent clear cut areas with minor vegetation regrowth that were classified as older clear cuts. In general, I noticed that areas with distinct patters of vegetation density/height were best represented in the unsupervised classification, due to distinct spectral signatures. It should be noted that shadows cast by hills and clouds were classified as water bodies, due to the 'dark' spectral signature that both these land cover types are associated with in the image. Results of the unsupervised classification show that roughly two thirds of the area covered by the image has been affected by logging activities, with ~29% of there area being recent clearcuts (table 1).

Delineation of classes was achieved using a dendrogram analysis, which was used to reduce statistical misclassification. Using this technique, similar classes were merged and then reduced to only the four major land cover types. The study area was not very heterogeneous, therefore the dendrogram aided in merging classes that were statistically close together.

2.2. Supervised classification

Results of the supervised classification show that water bodies were better able to be classified properly. This is most likely due the separation of shadows as a distinct training class, and

Land Cover Type	Area (hectares)	Area (%)
Shadows	160	1.5
Water Bodies	363	3.4
Remaining Forest	2625	24.6
Clearcuts	4406	41.3
Older Clearcuts	3115	29.2

Table 1. Land cover types and their respective areas. Values were determined from the *supervised* classification results.

therefore the maximum likelihood classification was able to distinguish shadows from water bodies. It is also evident that the supervised classification shows more of the image area classified as clearcut than the unsupervised classification.

Training samples were carefully delineated and combined to accurately capture the proper range of spectral signatures associated with each class. However, the histogram plots for the training samples used (fig. 4), shows relatively high covariance and standard deviation values for the 'Clearcuts' training sample. It should be noted that these were the lowest values achieved through multiple training sample choices, and this seemed to display the best results using the 'interactive supervised classification tool,' despite the unfavourable statistics for this class. We can see that the 'Clearcuts' class has a very broad distribution on the histogram plots (fig. 4) , which tells us that there is high variability in the spectral signatures for this class. The colour composite image (fig. 1) shows clearcut areas as somewhere between a light blue to a light pink shade due to different amounts of vegetation regrowth, which helps to explain the variability we see in the statistics and histogram plots. It should be noted that significant overlap existed in the histogram plots, therefore our training samples may not have been perfectly accurate. This source of error will be further discussed in section 4.

On the other hand, statistical information tells us that every other class was able to be classified well, as relatively low covariance and standard deviation values tell us that these land cover types had more distinct spectral signatures (fig. 4). Upon visual inspection of our supervised classification results (fig. 5), in comparison with the colour composite image (fig.1), show that remaining forests, water bodies, and older clear cuts were accurately classified. Proportional area calculations show that roughly one quarter of the area is covered by remaining forest, clearcuts cover ~41% and older clearcuts cover ~29% of the area.

3. Accuracy Assessment

	Water Bodies	Clear Cuts	Older Clear Cuts	Remaining Forest	Total	U_Accuracy	Kappa
Water Bodies	17.00	2.00	5.00	0.00	24.00	0.71	0.00
Clear Cuts	23.00	339.00	9.00	67.00	438.00	0.77	0.00
Older Clear Cuts	0.00	30.00	241.00	36.00	307.00	0.79	0.00
Remaining Forest	2.00	72.00	2.00	299.00	373.00	0.80	0.00
Total	42.00	443.00	257.00	402.00	1142.00	0.00	
P_Accuracy	0.40	0.77	0.94	0.74	0.00	0.74	0.00
Kappa	0.00	0.00	0.00	0.00	0.00	0.00	0.72

Table 2. Confusion matrix: pseudo accuracy assessment of classes, comparing the (unsupervised) ISO cluster classification to the (supervised) MLC 'ground truth' classification. User accuracy results show that waterbodies and clearcut areas were the most poorly classified, while remaining forest and older clearcut areas were the most accurately classified. Kappa coefficient of 0.72 tells us that our classification was moderately successful.

A (pseudo) accuracy assessment was conducted to examine the reliability of our classification of land cover types. This involved computing a confusion matrix that compares the unsupervised ISO cluster classification with the ground truth (i.e, the supervised MLC classification), based on 1500 randomly selected points. Our results show that water bodies, clearcuts, older clearcuts and remaining forest classes have percent correctly classified values of 71, 77, 79, and 80 % respectively (based on user accuracy values). This tells us that waterbodies and clear cuts classes were the most poorly classified, while remaining forest and older clearcut classes were the most accurately classified. It was also noted that the producers accuracy for water bodies was significantly lower than other classes, telling us that these areas were the most difficult to map. As mentioned above, this is most likely due to the inability of our unsupervised classification to differentiate waterbodies from shadows.

A kappa coefficient of 0.72 tells us that our classification of the Landsat image was successful, as the overall as an overall assessment of the accuracy results showed relatively high agreement (ArcGIS).

4. Conclusion:

4.1. Success of Classification Efforts

The results of our classification efforts show that although the area of interest was classified successfully, there were still some land cover types that showed misclassification. These land cover types were water bodies and areas with variable vegetation regrowth such as older and more recent clearcuts. Statistics and histogram plots of our training samples used in the supervised classification (fig. 4), as well the confusion matrix assessing overall accuracy (table 2), give us relevant information regarding the overall efficacy and reliability of our results. We have shown that our classification efforts have been able to successfully classify remaining forest areas, and although older and recent clearcuts were less so, we are still able to conclude that the overall image classification was a success based on the relatively high kappa coefficient given in our confusion matrix (table 2).

Based on visual inspection of the maps created, as well as cross referencing with ground truth imagery (fig. 1), land cover that displayed uniformity (i.e, remaining forest unaffected by logging activities) was able to be identified clearly. Land cover that did not display uniformity (in terms of spectral characteristics), such as areas with variable and uneven distribution vegetation growth (i.e, clearcuts), were not identified as clearly. Although waterbodies were highly uniform across the area of interest, it was difficult discern between them and shadows. Only the supervised classification was able to make this classification.

4.2. Recommendations

The results of our classification efforts allow us to speculate on possible sources of error, and therefore we are able to provide a few recommendations to avoid them. First of all, it should be stated that the supervised classification routine showed the best results in terms of ability to differentiate between land cover types, especially waterbodies and shadows. The use of 'ground truthed' training samples delineated by the map producer, allows for more accurate classification of land cover types. However, the unsupervised routine was useful in that it highlights areas of interest and limits bias from the map producer. This is critical information that helps improve the accuracy of the supervised routine. Our histogram plots (mentioned in section 2.2), show ed significant overlap, and therefore did not best represent classes of pixels with distinct spectral characteristics. Ideally, our histogram plots should show distinct classes with minimal overlap.

In addition to error associated with producer bias (which can be improved with training / practice) land cover identification could be improved by limiting sources of error associated training sample delineation. Images with a higher spatial resolution facilitates higher confidence in the differentiation between classes by the map producer, and allows for more a more accurate assessment of land cover types. Furthermore, having high resolution images captured at different timescales allows the map producer to cross reference between images, and see changes in land cover over time. This will also help with preliminary classification of land cover types before the supervised and unsupervised routines, and also with the delineation of training samples thereafter. It is also important increase the temporal resolution of images used in classification efforts. Increasing the frequency of images taken will allow the map producer to choose images with limited cloud cover and shadows cast by the sun. Another method for limiting error associated with training sample delineation would be field calibration and site visits to validate 'ground truth' classes. This could help avoid issues in discerning between water bodies and shadows, which was a significant source of error in my analysis. I recommend that the BC ministry of Forests, Lands and Natural Resources consider these methods to improve the analysis presented.

5. References:

Roy, D. P., Wulder, M. A., Loveland, T. R., Woodcock, C. E., Allen, R. G., Anderson, M. C., ... & Scambos, T. A. (2014). Landsat-8: Science and product vision for terrestrial global change research. *Remote sensing of Environment*, *145*, 154-172.

ESRI. (2019). ArcGIS Desktop: Tools, Tool Reference, Spatial Analyst toolbox, Segmentation and Classification toolset; *Compute Confusion Matrix.* Retrieved from: https://desktop.arcgis.com/en/arcmap/latest/tools/spatial-analyst-toolbox/compute-confusion-matrix.htm on April 6, 2019.

NASA, USGS. (2019) Landsat Science: *Landsat 8.* Retrieved from: https://landsat.gsfc.nasa.gov/landsat-data-continuity-mission/ on April 6, 2019.