GEOMATICS CANADA
OPEN FILE 17

Evaluation of Landsat Based Fractional Land Cover Mapping in the Alberta Oil Sands Region

D. Pouliot, W. Parkinson, and R. Latifovic

2015

Canada

Evaluation of Landsat Based Fractional Land Cover Mapping in the Alberta Oil Sands Region

D. Pouliot, W. Parkinson, and R. Latifovic

Natural Resources Canada, Earth Sciences Sector, Canada Centre for Mapping and Earth Observation, Canada Centre for Remote Sensing

2015

ABSTRACT

In the Alberta Oil Sand Region (AOSR) high spatial resolution (<5 m) remotely sensed multispectral time series are needed to capture the varying size and rates of change that occur. However, limited spatial-temporal coverage and cost of current high resolution sensors make such an approach impractical for retrieval of historical information and long term monitoring. Using moderate spatial resolution (~30m) time series such as that available with the Landsat series of sensors provides an alternative. In this research the potential to derived sub-pixel information on land cover types was evaluated for the AOSR using Landsat time series. Sub-pixel land cover fractions were trained for Landsat using high resolution (2 m) Geoeye data classified into basic land cover types. Cover types evaluated included conifer forest, broadleaf forest, shrub, low vegetation cover, bare, and water. The point spread function of Landsat was modeled to ensure that the reflectance properties measured were coincident with the training footprint in the higher spatial resolution Geoeye scenes. Decision tree classifiers were used for the fractional modeling. Results showed that land cover fractions could be estimated over the region with an average absolute error ranging from 7-17%. Sampling exerted a significant effect where validation using a holdout Geoeye scene preformed inferior to sampling from all available scenes as expected. Water and bare covers had limited sampling for fractions between 25-75% and therefore the results for these covers are uncertain in this range. Better controlling for spectral variability, fractional training and Landsat data quality in site specific analysis suggests significant improvement in accuracy compared to the regional analysis. The improvement for the site specific analysis ranged from 5-10%. Examination of forest fraction sensitivity to change revealed good agreement with forest harvesting and fire, but did not capture insect related damage well. These findings suggest there is potential for fractional land cover retrieval, but error is likely to remain moderate if training and remote sensing data are not carefully controlled for large regional applications.

Table of Contents

1. INTRODUCTION

Sub-pixel or fractional land cover represents the area covered by a given land cover type within a defined footprint on the land surface. For multispectral remote sensing systems the footprint typically corresponds to a mapping unit appropriate for the selected earth observation sensor. The advantage of fractional land cover is the ability to provide more detailed information on the spatial distribution of land cover and change in a region, but this depends on whether accurate fractional estimates can be retrieved. Numerous studies have been undertaken to better understand the potential to extract land cover fractions from different types of remotely sensed data. Spectral unmixing using hyperspectral image data has shown to be effective. However, without extensive and repeated coverage approaches based on hyperspectral data cannot currently be considered for large area monitoring needs. An alternative is machine learning algorithms and other statistical methods which can be trained to predict fractions within a region using multispectral satellite data records. These satellite data records provide historical measurements and are expected to be continued in the future to support ongoing monitoring. Examples of multispectral based fractional land cover in Canada have mostly made use of coarse spatial resolution data (> 250m) such as the MODIS continuous fields products (Hansen et al., 2003), boreal fractional assessment using SPOT\VEGETATION (Fernandes et al., 2004), and northern Canada fractions derived from SPOT\VEGETATION (Olthof et al., 2007). More recently moderate spatial resolution (30 m) Landsat data has been used for global forest fractions (Hansen et al., 2013). Much of the research on fractional retrieval has focused on snow, impervious or forest cover types. Other fractional covers such as water, shrub, and herb have received less attention.

Understanding change dynamics in the AOSR requires consideration of the cumulative effects of small and large changes occurring at rapid and moderate rates from a variety of causes including mining development, mining exploration, fire, forest harvesting, infrastructure development, hydrological alteration, and insect related mortality (Latifovic and Pouliot et al., 2013). The ability to derive land cover fractions provides detailed information that can be used to investigate the effect of land cover conversions on air quality, water storage, habitat use, carbon sequestration, and other environmental processes in the region. Thus the analysis undertaken here was designed to determine the accuracy for the estimation of different land cover fractions from Landsat data and identify fractions for which unacceptable accuracy was observed to guide future research. Landsat was selected as it provides the longest record of moderate spatial resolution remotely sensed observations and will be continued in the future with Landsat 8. The Sentinel constellation could also be integrated with past Landsat measurements for ongoing monitoring and trend analysis. All Landsat bands were used in the fraction model development. Training was based on high spatial resolution (2 m) Geoeye remotely sensed images that were classified into basic land cover types using an integrated object and pixel classification approach. Fraction results were validated using either holdout samples or entire high resolution scenes in a variety of combinations to assess sensitivity to sampling.

2. METHODS AND RESULTS

2.1 Fractional Calibration and Validation Data
Training and validation data for fractional land cover models were derived from high spatial resolution (2 m) Geoeye images classified into basic land cover types. The scenes used are shown in Figure 1. Geoeye includes 4 multispecral bands ranging from the green to the near-infrared at 2 m spatial resolution and a panchromatic band at 0.5 m spatial resolution.
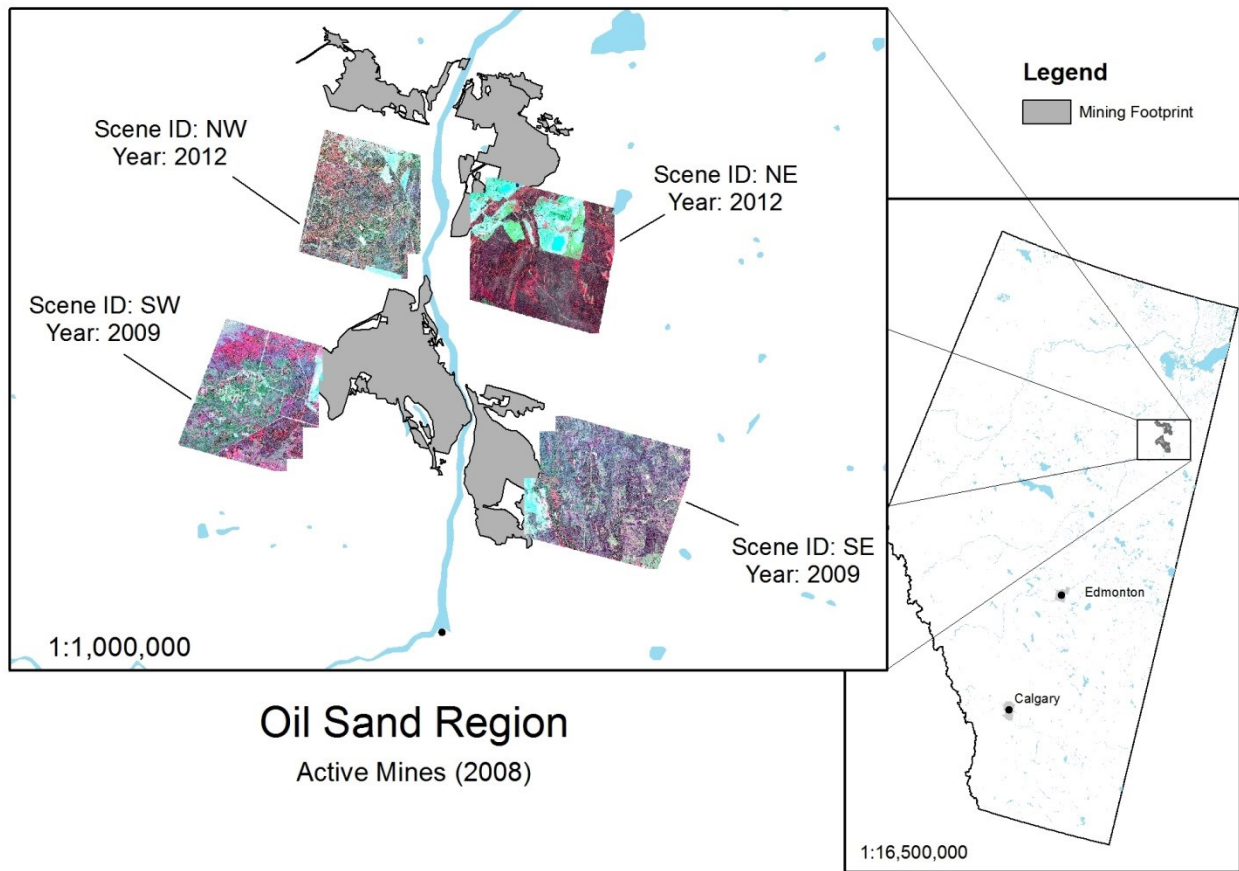
Figure 1: Location and dates of Geoeye scenes used to derive reference and land cover fractions in the AOS.

To improve classification accuracy of the reference data (i.e. land cover maps derived from Geoeye scenes) a pixel and object based classification approach was used. The pixel based classification was performed using the random forest algorithm implemented in R (Breiman, 2001). All four multispectral bands were used. Classes defined for the pixel and object classifications are given in the column and row header of Table 1 respectively. For the object classification, images were segmented using eCognition Developer 8.8 and objects classified using the embedded decision tree classifier. Object features included the object mean and standard deviation for the 4 multispectral bands and the Normalized Difference Vegetation Index (NDVI).

The object and pixel classification results were combined using a fuzzy expert rule set defined in Table 1, where the membership for the pixel based classification was multiplied by the value in the table for the associated object class. The class with the resulting maximum membership was taken as the adjusted class output. In this case, the purpose of the object based classification was to generalize the pixel based classification to achieve an overall more accurate result. Initially the values in Table 1 were given an equal probability across all classes. These values were then adjusted in a trial and error process to correct for clearly identifiable errors in the pixel based classification. For comparison to Landsat the final high resolution classification result was up-scaled to 30 m resolution based on the Landsat Point Spread Function defined in section 2.4.

Table 1: Classes for the pixel and object classification and membership adjustment rule set.

		PIXEL CLASS							
		Water	Broadleaf Forest	Insect Mortality	Shrub	Conifer Forest	Bare	Shadow	Low Vegetation
OBJECT CLASSES	Bare	0.1	0.1	0.001	0.1	0.1	0.399	0.1	0.1
	Broadleaf Forest	0.1	0.4	0.05	0.05	0.1	0.1	0.1	0.1
	Open Conifer Forest	0.1	0.05	0.01	0.1	0.27	0.27	0.1	0.1
	Shrub	0.1	0.15	0.001	0.249	0.1	0.15	0.1	0.15
	Conifer Forest	0.1	0.1	0.1	0.1	0.3	0.1	0.1	0.1
	Water	0.5	0.05	0.05	0.05	0.05	0.05	0.2	0.05
	Low Vegetation	0.1	0.1	0.001	0.1	0.1	0.1	0.1	0.399
	Unclassified	0.125	0.125	0.125	0.125	0.125	0.125	0.125	0.125

Training and validation data for the pixel and object land cover classifications were derived from field photos and interpretation of Pansharped Geoeye images. For each sample location an object containing 3-9 pixels was manually interpreted in the pansharpened imagery and a class label assigned. These data were split into a training set using 70% of the data and a validation set using the remaining 30%. Overall accuracy was measured for each Geoeye scene for the pixel and object-pixel adjusted classification results.

Overall accuracy for the pixel based classification of the Geoeye scenes was reasonable ranging from 84 to 89% (Table 2). The incorporation of the object classification improved accuracy in the range of 2 to 5%, with an average improvement of 3.5%. The largest source of confusion was between shrub and broadleaf forest and this class generally benefited the most from the object adjustment. A visual example is provided in Figure 2.

Table 2: Overall classification accuracies for the Geoeye reference data

	Per-Pixel Classification Accuracy (%)	Object Adjusted Classification Accuracy (%)	Change (%)
NW	83.76	85.90	+2.14
NE	88.36	93.33	+4.97
SE	85.56	90.52	+4.96
SW	89.36	91.18	+1.82

In the example shown in Figure 2, along the river edge there is a broadleaf patch which in the pixel based classification has been confused with shrub. The object based adjustment has reduced this problem, but some shrub/broadleaf forest confusion remained.
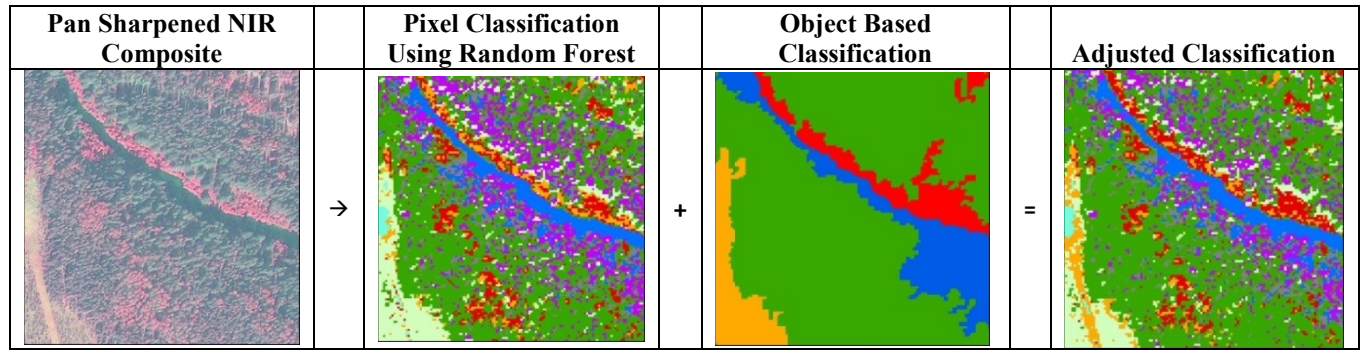
Pan Sharpened NIR Composite		Pixel Classification Using Random Forest		Object Based Classification		Adjusted Classification
	→		+		=	

Figure 2: Example results of the Geoeye classification and the object based adjustment. For the classified images, green = conifer, red = broadleaf, orange = shrub, blue = water, and purple = defoliation.

2.2 Landsat Data

Landsat observations from seven scenes covering the AOS were selected for the analysis (Table 3). All spectral bands and the brightness temperature measurements were used as well as NDVI, the Normalized Difference Moisture Index (NDMI), and a proxy for atmospheric transparency calculated as the difference between the red and blue bands, divided by the shortwave band at 1500 nm. Top of atmosphere reflectance was used and cloud and cloud shadow screening was implemented using software developed based on the FMASK algorithm (Zhu and Woodcock, 2012), but with specific modifications to improve performance over Canada. Scenes were further checked and haze or missed clouds were manually removed.

Table 3: List of Landsat scenes used in the analysis.

Path	Row	Date	Sensor
41	20	2009-07-21	TM
42	20	2009-07-28	TM
42	20	2009-08-29	TM
43	20	2009-08-28	ETM+
42	20	2011-07-02	TM
42	20	2011-09-04	TM
42	20	2012-07-28	ETM+

2.3 Point Spread Function Analysis

To provide a ground measure that is more consistent with the sensor observations the Geoeye scenes were up-scaled to a 30 m grid cell using the Landsat point spread function (PSF). The point spread function is known to vary by band, season, year, sensor, and atmosphere condition (Kavzoglu, 2010). For the purpose of this analysis we sought to define a generic PSF as resampling, imperfect atmosphere correction, and geolocation errors would contribute much greater error than the small deviations of the PSF from the generic version developed. Here we considered PSF estimates from McGillen and Yu (1983), Markham (1985), Solomonson et al. (1988), Kavzoglu (2004), and Kaiser and Schneider (2008). Estimates of PSF full width half maximum (FWHM) ranged from 31-50 m (Table 4). We applied different line PSFs for the scan and track direction to be consistent with

laboratory measurements of the PSF, where in the scan direction 3 m was added to the FWHM in the track direction.

Table 4: PSF Estimates for Landsat sensors.

	FWHW (m)				
Sensor	Scan	Track	Directional Invariant	Stdev	Reference
TM			38.8	16.5	McGillen, 1983
TM			49.9	21.2	Solmanson et al., 1998
ETM+	35	31	33.0	14.0	Kavzoglu, 2004
TM		25	24.7	10.5	Kaiser & Schneider, 2008
TM	34	31	32.5	13.8	Markham, 1985

To determine the optimal FWHM two analyses were undertaken. The first up-scaled the Geoeye near-infrared reflectance using different FWHM estimates and compared this to the Landsat near-infrared reflectance. Pearson correlation was used as the measure of agreement between the up-scaled Geoeye and Landsat data as relative instead of absolute difference was considered the most important due to differences in spectral response functions and radiometric calibration between sensors. Two Geoeye scenes were upscaled for this purpose the NE and NW scenes as identified in Figure 1. For the second analysis, the Geoeye land cover was up-scaled to create fractions used for training and validating Landsat based fraction models. Models were computed using different PSFs and the Nash Sutcliffe Efficiency calculated with a hold out sample for validation. Nash-Sutcliff Efficiency (NSE) is a measure of model deviation from the 1:1 line between observed and predicted model outputs. NSE was defined as:

$$NSE = \frac{\Sigma_i(y_i - y_p)}{\Sigma_i(y_i - \bar{y})} \tag{1}$$

Where y_i is the observed value or reference value, y_p was the result from the model estimate and $\bar{y}$ was the mean of the observed data. NSE can be interpreted similar to that of correlation, but negative values are possible when models do not perform well.

Results showed the PSF parameter of 36 m (standard deviation of 15.2 m = FWHM/2.355) to be reasonable in both sets of analysis (Figure 3). For the selected PSF, approximately 47 % of the received signal from the surface is assumed to come from within the 30 m grid cell and 90% within a 60 m grid cell. Figure 4 shows the PSF selected for this study with an intersecting plane representing a 30 m grid cell in the x and y directions. Applying too large a PSF reduces the variation in the data essentially acting as a smoothing operator. This is undesirable as it reduces the predicted fractional cover range. For the PSF selected here the reduction in the up-scaled land cover fractions was less than 1%.
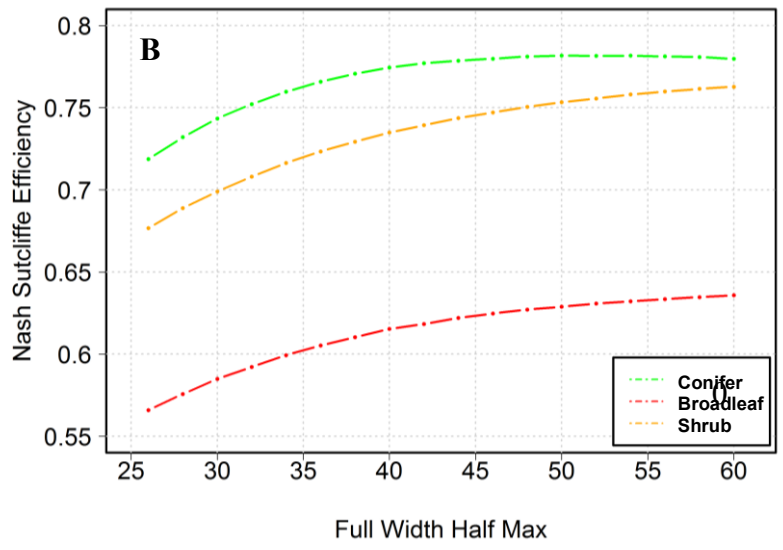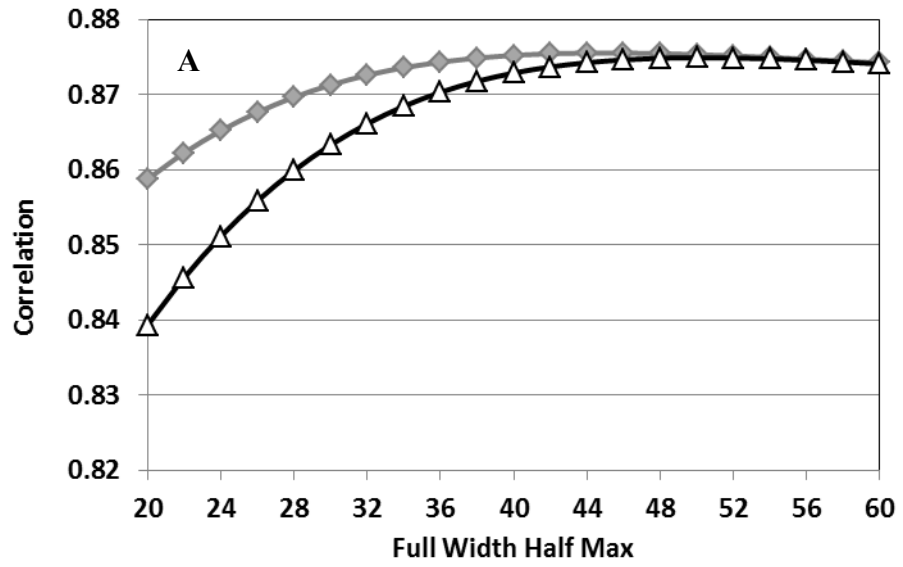
Figure 3: A) Effect of PSF on upscaling Geoeye data to Landsat (NE grey and NW black) and B) fractional land cover model prediction for conifer, broadleaf and shrub.
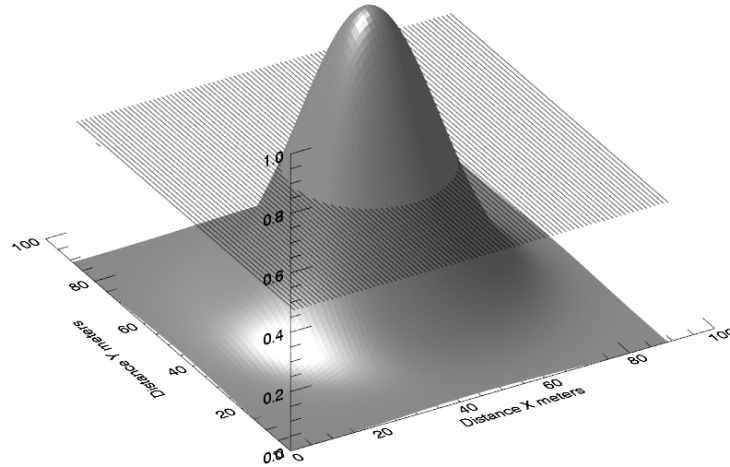
Figure 4: PSF selected for this study with a standard deviation of 15 m. The PSF is centered at 45 m. The intersecting plane identifies the 30 m grid cell in x and y.

2.4 Regional Fractional Modeling

Two decision tree classifiers were tested to derive fractional cover mapping algorithms, Random Forest implemented in R and the RuleQuest SEE5 software. Initial tests showed the two algorithms performed similarly, but with a very slight advantage to Random Forest. Thus Random Forest was used in subsequent analysis. Fractions tested included conifer forest, broadleaf forest, shrub, low vegetation cover, bare, and water. Dead trees due to insect mortality were considered part of the bare fraction. Samples with greater than 15% shadow fraction were removed from the analysis. With the dataset different approaches to training and validation were possible. The different training\validation sampling strategies included using all data or training from 3 scenes and validating on the fourth. Using all data was sub-divided into using scenes only for the north verses using only data for the south. The objective of the separation was to investigate the effect of Landsat data from the two different years North (2012) and South (2009). Metrics used to evaluate performance included the mean absolute error, mean error, error standard deviation, and the Nash-Sutcliff Efficiency (NSE).

Results for the different fractions and sampling/validation schemes are shown in Table 5. The lowest mean absolute errors were observed for water and bare covers. This was due to sampling limitations where fractions between 25% and 75% were insufficient. This caused the model to be optimized for predicting low and high water fractions with high accuracy and lower accuracy for mid-range fractions. For fractions where sampling was representative of the full range the mean absolute error rates were in the range of 6-11% for the all data validation tests. Shrub was the most variable between validation tests with a standard deviation of 5% followed by low vegetation, and broadleaf forest with a standard deviation of 4%. Conifer forest was the most stable with a standard deviation between validation tests of 2%. Conifer also had the best model performance as determined by Nash-Sutcliff efficiency. As expected the sampling and validation using samples from all scenes produced the best results compared to the holdout scene tests. No consistent difference was observed between validation tests for using all data, only North, or only South. Example scatterplots for the NE holdout validation test are shown in Figure 5.

Table 5: Fractional error metrics for the different fraction covers tested.

| Class	Mean Absolute Error	Mean Error	Error Standard Deviation	Nash Sutcliff Efficiency	Analysis
Bare	3.00	0.12	0.01	0.87	All Data
Bare	1.78	0.03	0.03	0.93	North Only (2012)
Bare	3.89	0.19	0.03	0.70	South Only (2009)
Bare	3.90	2.05	0.13	0.92	Validation NE
Bare	2.87	1.60	0.03	0.37	Validation NW
Bare	3.82	-1.69	0.04	0.50	Validation SE
Bare	5.80	-0.61	0.14	0.47	Validation SW
Conifer	11.32	0.02	0.03	0.80	All Data
Conifer	9.96	0.03	0.03	0.85	North Only (2012)
Conifer	12.26	-0.08	0.07	0.75	South Only (2009)
Conifer	12.26	3.04	0.08	0.80	Validation NE
Conifer	12.63	4.24	0.07	0.72	Validation NW
Conifer	15.40	-6.68	0.12	0.53	Validation SE
Conifer	15.24	-3.35	0.31	0.59	Validation SW
Broadleaf	6.87	0.17	0.02	0.79	All Data
Broadleaf	8.58	0.12	0.06	0.77	North Only (2012)
Broadleaf	5.06	0.19	0.01	0.80	South Only (2009)
Broadleaf	9.07	3.31	0.08	0.57	Validation NE
Broadleaf	16.90	-14.44	0.10	0.47	Validation NW
Broadleaf	12.44	11.84	0.08	-1.82	Validation SE
Broadleaf	11.77	-1.66	0.04	0.71	Validation SW
Low Vegetation	7.33	0.26	0.03	0.66	All Data
Low Vegetation	3.49	0.20	0.01	0.60	North Only (2012)
Low Vegetation	10.39	0.28	0.06	0.62	South Only (2009)
Low Vegetation	9.05	6.54	0.04	-0.12	Validation NE
Low Vegetation	7.67	5.09	0.06	-0.07	Validation NW
Low Vegetation	13.43	-10.94	0.12	0.07	Validation SE
Low Vegetation	14.80	-8.77	0.40	0.36	Validation SW
Shrub	6.42	0.17	0.03	0.65	All Data
Shrub	8.48	0.24	0.03	0.66	North Only (2012)
Shrub	3.85	0.17	0.01	0.56	South Only (2009)
Shrub	16.06	-14.47	0.03	-0.07	Validation NE
Shrub	11.18	5.53	0.08	0.24	Validation NW
Shrub	10.07	7.42	0.11	-0.33	Validation SE
Shrub	18.37	17.57	0.24	-3.40	Validation SW
Water	2.77	2.67	0.01	-2.67	All Data
Water	3.41	3.27	0.05	-3.02	North Only (2012)
Water	1.88	1.81	0.02	-1.75	South Only (2009)
Water	2.78	2.51	0.11	-0.83	Validation NE
Water	3.74	3.27	0.27	-1.07	Validation NW
Water	7.04	6.98	0.28	-28.82	Validation SE
Water	4.53	4.31	0.09	-0.86	Validation SW
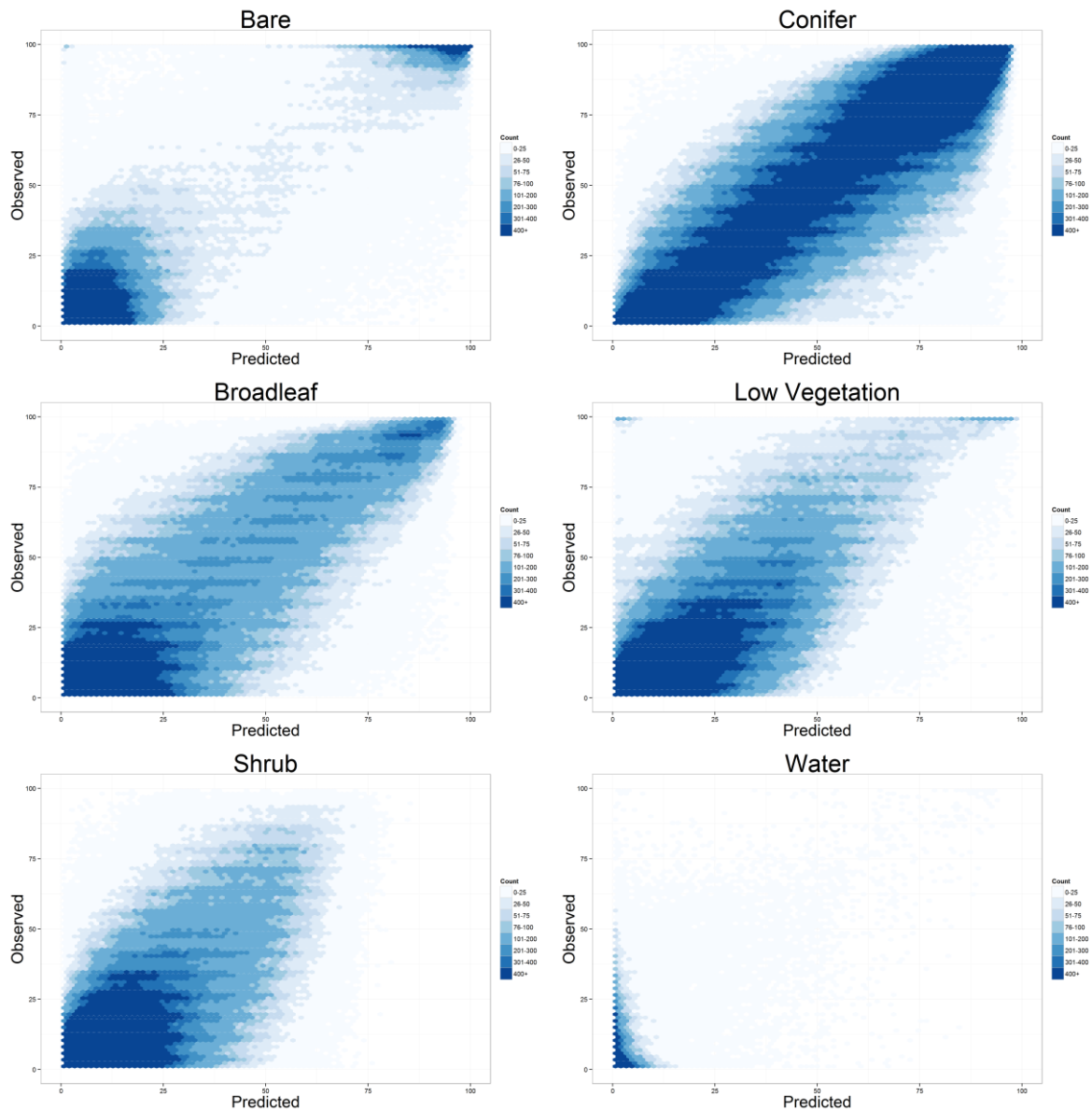
Figure 5: Scatterplots for the different fractional cover models at the regional scale. Axes indicate % fractional cover.

2.5 Site Specific Fractional Modeling

The potential to improve results was examined in site specific areas where the spectral range, training, and Landsat data quality could be better controlled. At the regional scale the model must account for greater confusion due to spectral overlap for a greater number of fractional land cover combinations. Further at a site specific scale the Geoeye training data can be more rigorously checked to remove/reduce error resulting from geolocation or classification. Cloud, haze, and shadow screening of Landsat scenes could also be enhanced leading to improved modeling. Figure 6 provides scatterplots for comparison with the regional analysis. For the areas tested the mean absolute errors were substantially smaller. On average the mean absolute error was reduced by 8%. This analysis severs as an indication as to the upper bound accuracy that could be achieved with Landsat TM and ETM+ data.
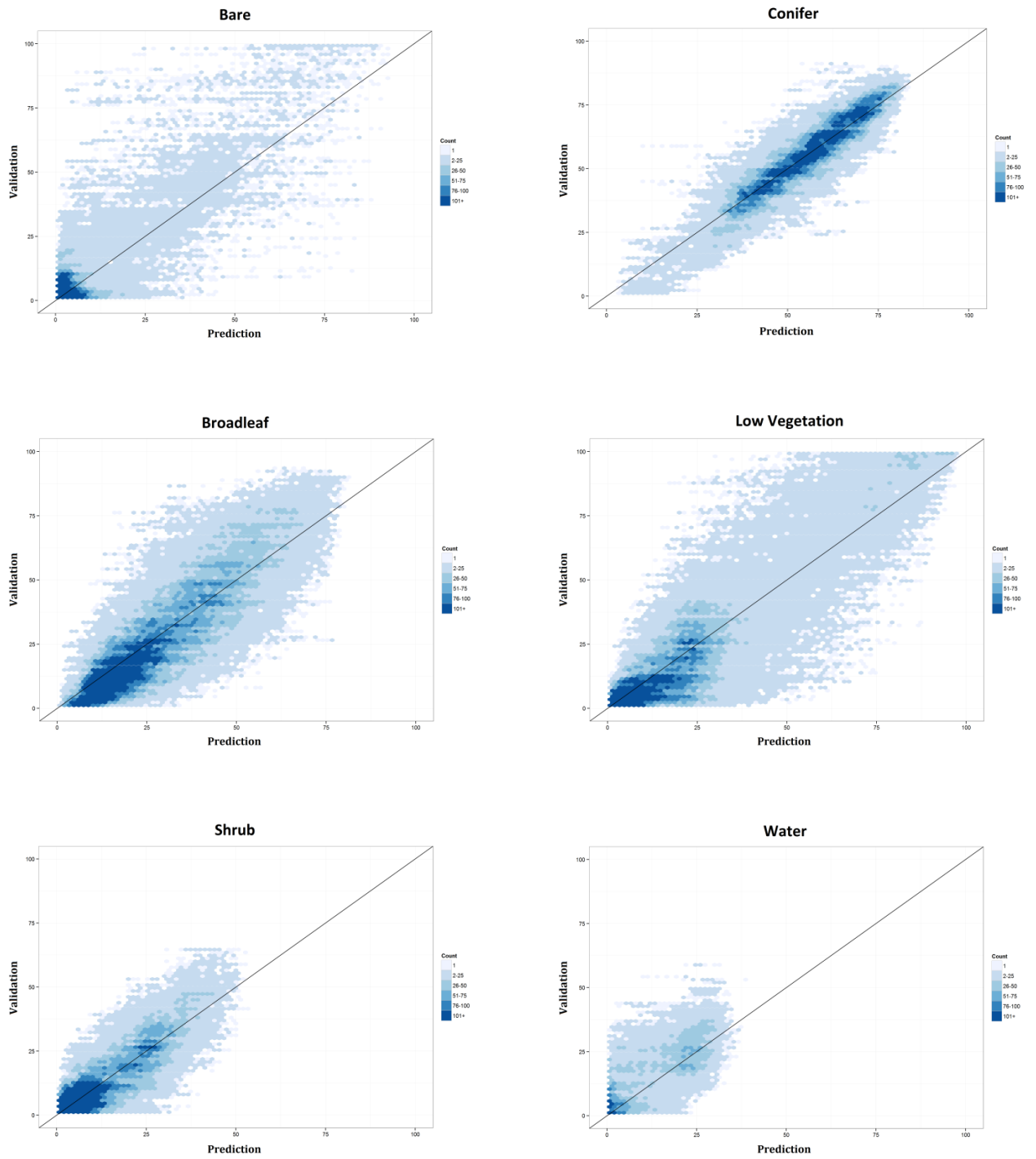
Figure 6: Scatterplots for the different fractional cover models at the site specific scale. Axes indicate % fractional cover.

2.6 Assessment of Fraction Change Sensitivity and Temporal Stability

To determine the variability of fractions in time, the fractional model using all four input training Geoeye classifications was applied to a set of temporally smoothed Landsat Mosaics from 1984 to 2012. Forest fraction was derived by adding the broadleaf and conifer forest fractions. Temporal smoothing of the Landsat time series consisted of first detecting outliers using an iterative Lowess filter to reduce noise and then fitting linear piecewise segments using a modified version of the procedure developed by Kennedy et al. (2010). Areas of specific changes and where change was known not to have occurred were sampled. Example results are shown in Figure 7. Shrub and low vegetation covers were in a constant state of change in this area and thus temporal variability for these covers was not assessed. For fire and forest harvesting forest fraction showed good sensitivity to change, however for insect damage areas, confusion in the training data reduced sensitivity. For areas where change did not occur the average inter-annual difference was typically less than ±10%. Conifer tended to be the more temporally variable due to the more open nature of the canopy and the spectral variability associated with shrub and other non-forest ground covers.
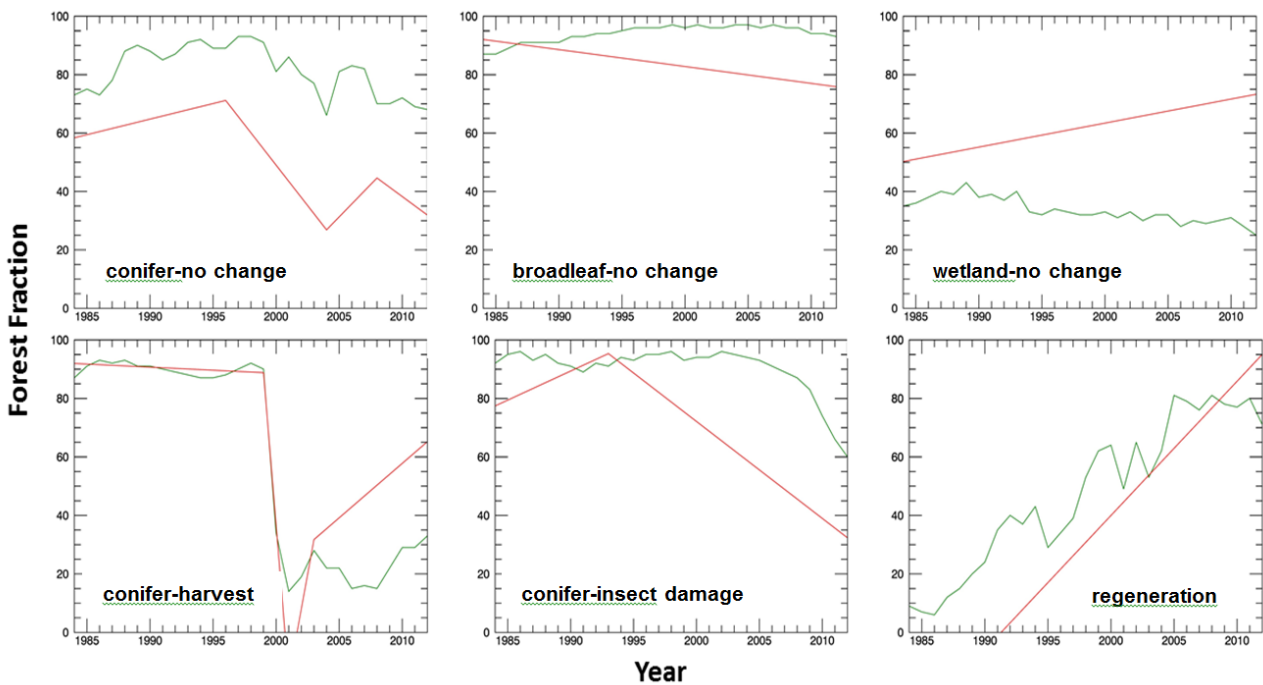

Figure 7: Examples of forest fraction sensitivity to change and temporal variability (Green line). Red line is the Normalized Difference Moisture Index (NIR-SWIR)/(NIR+SWIR) linearly scaled to be on fractional scale for comparison.

3. CONCLUSIONS

The accuracy for fractional land cover mapping in the AOSR was investigated. Results reveal that fractional models could be derived, but with error rates between 7and 17%. Significant improvement could be made with greater control on stratification, Landsat data processing and training data quality. Temporal stability appeared reasonable. Further research is required to improve training data development, Landsat data processing, and modeling for greater reliability in fractional model development.

ACKNOWLEDGEMENTS

REFERENCES

Breiman, L. (2001). Random forests. Machine learning, 45(1), 5-32.

Fernandes, R., Fraser, R., Latifovic, R., Cihlar, J., Beaubien, J., & Du, Y. (2004). Approaches to fractional land cover and continuous field mapping: A comparative assessment over the BOREAS study region. Remote Sensing of Environment, 89(2), 234-251.

Hansen, M. C., DeFries, R. S., Townshend, J. R. G., Sohlberg, R., Dimiceli, C., & Carroll, M. (2002). Towards an operational MODIS continuous field of percent tree cover algorithm: examples using AVHRR and MODIS data. Remote Sensing of Environment, 83(1), 303-319.

Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., & Townshend, J. R. G. (2013). High-resolution global maps of 21st-century forest cover change. Science, 342(6160), 850-853.

Kaiser, G., & Schneider, W. (2008). Estimation of sensor point spread function by spatial subpixel analysis. International Journal of Remote Sensing, 29(7), 2137-2155.

Kavzoglu, T. (2004). Simulating Landsat ETM+ imagery using DAIS 7915 hyperspectral scanner data. International journal of remote sensing, 25(22), 5049-5067.

Kennedy, R. E., Yang, Z., & Cohen, W. B. (2010). Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr—Temporal segmentation algorithms. Remote Sensing of Environment, 114(12), 2897-2910.

Latifovic, R., & Pouliot, D. (2014). Monitoring Cumulative Long-Term Vegetation Changes Over the Athabasca Oil Sands Region. Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of , 7(1).

Markham, B. L. (1985). The Landsat sensors' spatial responses. Geoscience and Remote Sensing, IEEE Transactions on, (6), 864-875.

McGillem, C. D., Anuta, P. E., Malaret, E., & Yu, K. B. (1983). Estimation of a remote sensing system point-spread function from measured imagery. LARS Technical Reports, 81.

Olthof, I., & Fraser, R. H. (2007). Mapping northern land cover fractions using Landsat ETM+. Remote Sensing of Environment, 107(3), 496-509.

Salomonson, V. V., Nickeson, J. E., Bodechtel, J., & Zilger, J. (1988). Comparative point-spread function calculations for the MOMS-1, thematic mapper and SPOT-HRV instruments. In Spectral Signatures of Objects in Remote Sensing (Vol. 287, p. 407).

Zhu, Z., & Woodcock, C. E. (2012). Object-based cloud and cloud shadow detection in Landsat imagery. Remote Sensing of Environment, 118, 83-94.