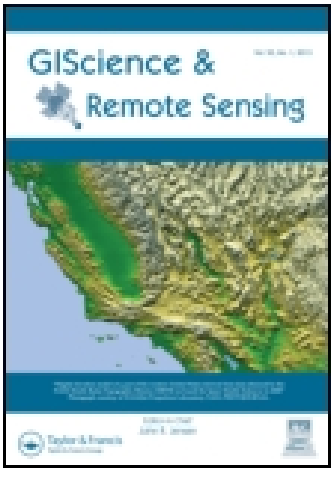

GIScience & Remote Sensing
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/tgrs20

Comparison of NAIP orthophotography and RapidEye satellite imagery for mapping of mining and mine reclamation

A.E. Maxwell[a], M.P. Strager[b], T.A. Warner[c], N.P. Zégre[d] & C.B. Yuill[b]

[a] College of Science, Technology, and Mathematics, Alderson Broaddus University, 101 College Hill Drive, Philippi, WV, USA

[b] Division of Resource Management, West Virginia University, 2018 Agricultural Sciences Building, PO Box 6108, Morgantown, WV 26506-6108, USA

[c] Geology and Geography Department, West Virginia University, 330 Brooks Hall, 98 Beechurst Ave, Morgantown, WV 26506-6300, USA

[d] Division of Forestry and Natural Resources, West Virginia University, PO Box 6125, Morgantown, WV 26506-6125, USA

Published online: 09 May 2014.

To cite this article: A.E. Maxwell, M.P. Strager, T.A. Warner, N.P. Zégre & C.B. Yuill (2014) Comparison of NAIP orthophotography and RapidEye satellite imagery for mapping of mining and mine reclamation, GIScience & Remote Sensing, 51:3, 301-320, DOI: 10.1080/15481603.2014.912874

To link to this article: http://dx.doi.org/10.1080/15481603.2014.912874

GIScience & Remote Sensing, 2014
Vol. 51, No. 3, 301–320, http://dx.doi.org/10.1080/15481603.2014.912874

Taylor & Francis
Taylor & Francis Group

Comparison of NAIP orthophotography and RapidEye satellite imagery for mapping of mining and mine reclamation

A.E. Maxwell[a]*, M.P. Strager[b], T.A. Warner[c], N.P. Zégre[d] and C.B. Yuill[b]

[a]*College of Science, Technology, and Mathematics, Alderson Broaddus University, 101 College Hill Drive, Philippi, WV, USA;* [b]*Division of Resource Management, West Virginia University, 2018 Agricultural Sciences Building, PO Box 6108, Morgantown, WV 26506-6108, USA;* [c]*Geology and Geography Department, West Virginia University, 330 Brooks Hall, 98 Beechurst Ave, Morgantown, WV 26506-6300, USA;* [d]*Division of Forestry and Natural Resources, West Virginia University, PO Box 6125, Morgantown, WV 26506-6125, USA*

(*Received 20 December 2013; accepted 26 March 2014*)

National Agriculture Imagery Program (NAIP) orthophotography is a potentially useful data source for land cover classification in the United States due to its nationwide and generally cloud-free coverage, low cost to the public, frequent update interval, and high spatial resolution. Nevertheless, there are challenges when working with NAIP imagery, especially regarding varying viewing geometry, radiometric normalization, and calibration. In this article, we compare NAIP orthophotography and RapidEye satellite imagery for high-resolution mapping of mining and mine reclamation within a mountaintop coal surface mine in the southern coalfields of West Virginia, USA. Two classification algorithms, support vector machines and random forests, were used to classify both data sets. Compared to the RapidEye classification, the NAIP classification resulted in lower overall accuracy and kappa and higher allocation disagreement and quantity disagreement. However, the accuracy of the NAIP classification was improved by reducing the number of classes mapped, using the near-infrared band, using textural measures and feature selection, and reducing the spatial resolution slightly by pixel aggregation or by applying a Gaussian low-pass filter. With such strategies, NAIP data can be a potential alternative to RapidEye satellite data for classification of surface mine land cover.

Keywords: National Agriculture Imagery Program; NAIP; RapidEye; mountaintop removal coal mining; surface coal mining; land cover classification; machine learning

Introduction

The United States Department of Agriculture (USDA) National Agriculture Imagery Program (NAIP) orthophotography offers high spatial resolution [1 m ground sampling distance (GSD)], up to four spectral bands, nearly cloud-free coverage of large areas (e.g., entire states) acquired during a single growing season and low cost to the user. For the state of West Virginia, it is available for the growing seasons of 2007, 2009, and 2011 (Aerial Photography Field Office 2012), potentially allowing multitemporal analysis. Because of these characteristics, it is compelling to consider these data for land cover mapping and monitoring. However, prior research has found that high spatial resolution, variable viewing geometry and illumination, and in some cases limited spectral resolution can complicate mapping tasks from aerial data (Cushnie 1987; Myeong et al. 2001;

*Corresponding author. Email: maxwellae@ab.edu

Bozheva, Petrov, and Sugumaran 2005; Guo et al. 2007; Baker et al. 2013). In comparison, high spatial resolution commercial satellite images can potentially offer large scenes with relatively consistent viewing and illumination geometry and four to eight well-characterized spectral bands (Toutin 2009).

This research compares the use of NAIP orthophotography and RapidEye satellite imagery for the mapping of mining and mine reclamation within an individual mine complex in the southern coalfields of the eastern United States. RapidEye satellite data offer five spectral bands and a 5 m cell size (Tyc et al. 2005). RapidEye data were used, as opposed to IKONOS or other high-resolution satellite data, because cloud-free coverage was available that was acquired at a time nearly coincident with the NAIP collection, because the RapidEye constellation of five satellites offers a frequent return time (5.5 days at nadir) (Tyc et al. 2005), and because the researchers have previously used this imagery for mapping surface mine land cover (Maxwell et al. 2014) with promising results. The following factors were investigated:

(1) Comparative classification accuracy for five land cover classes under various processing scenarios (at the original cell size of each data set, at a common cell size, at a common radiometric resolution, and at a common cell size and radiometric resolution).
(2) The relative importance of different image bands in the random forest (RF) classification.
(3) The relative importance of multiseasonal data.
(4) The relative importance of the near-infrared (NIR) band for NAIP classification.
(5) The relative improvement in classification accuracy as the number of classes is reduced.
(6) Improvement in classification accuracy of NAIP with the incorporation of texture measures from the gray level co-occurrence matrix (GLCM).
(7) Improvement in classification accuracy of NAIP resulting from image smoothing using a Gaussian low-pass filter.

Background

In recent years, data selection for fine-scale mapping has become more complex due to the wide range of choices of both aerial- and satellite-based high spatial resolution data. Determining which sensors offers the appropriate spatial, spectral, temporal, and radiometric scales to meet specific mapping needs may not be a simple task as multiple criteria must be considered and evaluated (Phinn 1998). Variations in the cost of data further complicate data selection (Tarnavsky et al. 2004; Warner, Nellis, and Foody 2009).

The complexity of land cover mapping from high-resolution aerial imagery has been previously explored. For example, Myeong et al. (2001) noted that shadows and similarity in spectral response between mapping classes complicated the mapping of urban cover from high-resolution (0.61 m pixel size) color-infrared imagery. There is also a growing literature on the use of NAIP orthophotography for land cover mapping. For example, Davies et al. (2010) used NAIP to map juniper cover in Idaho, USA, while Meneguzzo, Liknes, and Nelson (2013) attempted to classify isolated trees in Steele County, Minnesota, USA. Baker et al. (2013) compared pixel-based and object-based classification methods for mapping forest clearings associated with natural gas drilling operations in Greene County, Pennsylvania, USA. They noted the challenges of using NAIP data for image classification, including issues related to the range in time of day and season of

acquisition, which resulted in differences in digital number (DN) values between tiles, differences in shadow length and direction, and phenological differences. A search of the current literature suggests that there is little work published on comparison of NAIP and high-resolution satellite data or NAIP for the classification of mining and mine reclamation.

Mountaintop mining occurs in southern West Virginia, eastern Kentucky, eastern Tennessee, and southwestern Virginia, a region known as the southern coalfields of the eastern United States (USEPA 2005). In this region, mountaintop coal mining is the leading cause of land use/land cover change (Saylor 2008; Townsend et al. 2009; Drummond and Loveland 2010). Coal extraction results in forest clearing [e.g., 420,000 ha in Appalachia between 1973 and 2000 (USEPA 2005)] and fragmentation (Wickham et al. 2007), removal of top soil, and a recontouring of the landscape (Palmer et al. 2010; Bernhardt and Palmer 2011). Furthermore, mountaintop mining causes faster landscape alteration than more traditional mining methods such as auger, contour, and highwall mining (Fritz et al. 2010). After mining is complete, the disturbed areas are reclaimed, commonly to grasslands or shrublands (Simmons et al. 2008; Kazar and Warner 2013). More recently, there has been increased interest in reclamation to native forest species on the topographically altered terrain (Zipper et al. 2011).

Remote sensing has been investigated as a means to map landscape change due to mining and mine reclamation using data from the Landsat Multispectral Scanner, Thematic Mapper, Enhanced Thematic Mapper Plus, and SPOT (Rathore and Wright 1993; Townsend et al. 2009; Sen et al. 2012; Zégre, Maxwell, and Lamont 2013). Mapping mine reclamation is of specific concern as this is commonly the legacy imprint of mining on the landscape (Negley 2002), although Rathore and Wright (1993) found that mine reclamation was mapped with lower accuracy than active mining. Recently, there has been an interest in multitemporal imagery for separating mining and non-mining as a way to overcome the spectral similarities of reclaimed and undisturbed landscapes such as pastureland or herbaceous cover (Townsend et al. 2009; Sen et al. 2012). Textural measures have been investigated as a means to increase classification accuracy using single-date data (Warner 2011), and such variables have been shown to improve land cover classification in multiple studies (e.g., Chica-Olmo and Abarca-Hernández 2000; Agüera, Aguilar, and Aguilar 2008; Ghimire, Rogan, and Miller 2010; Rodríguez-Galiano et al. 2012a).

Machine learning algorithms such as support vector machines (SVMs) and RF offer particular advantages for classifying high spatial resolution imagery of mined landscapes. Machine learning algorithms have been shown generally to be more accurate and efficient than parametric classifiers, which make assumptions regarding the data distribution (Hansen, Dubayah, and DeFries 1996; Huang, Davis, and Townshend 2002; Rogan et al. 2003; Pal 2005; Ghimire et al. 2012; Loosvelt et al. 2012). Pal (2005) suggests that SVM and RF can provide comparable map accuracies, though RF is simpler to apply. Previous work by the authors (Maxwell et al. 2014) suggests that SVM provided a more accurate classification than RF for classification of surface mine land cover from high-resolution satellite imagery. In addition, optimizing the SVM algorithm, often the most time-consuming step in executing the algorithm, did not statistically increase the accuracy of the classification, suggesting that the method is relatively robust (Maxwell et al. 2014), although other studies have suggested that optimization may be of importance (Pal 2005).

SVMs make use of structural risk minimization to find the hyperplane that separates two classes with the maximum margin; it is an implementation of statistical learning theory and optimization algorithms to locate decision boundaries between data classes

(Vapnik 1995; Burges 1998; Pal and Mather 2005; Pal 2005; Su and Huang 2009). The points that lie closet to the hyperplane, termed "support vectors," define the margin. Because SVM algorithms are designed for two-class problems only, strategies have been developed incorporating multiple SVM algorithms to generate a multiclass classification (Vapnik 1995; Pal and Mather 2005; Pal 2005). The e1071 package in R, which is the implementation of SVM used in this research, uses a "one-against-one" approach for multiclass classification in which binary classifiers are trained and the appropriate class is found by a voting scheme (Meyer et al. 2012).

RF, introduced by Breiman (2001), uses multiple decision trees as a means to improve the accuracy and consistency of single-tree classifications. RF functions as an ensemble of decision trees and is a nonparametric learning algorithm. Bagging, which comprises bootstrap sampling with replacement (Breiman 1996), is used to generate a subset of the data to train each tree instead of using the entire training data set. The remaining data, termed out-of-bag (oob) data, are withheld and can be used for accuracy assessment. Instead of using all predictor variables in each tree, RF uses a random subset of the predictor variables (the number of which is defined by the user) to grow each tree of the ensemble. Although the use of a subset of variables decreases the classification accuracy of any one tree, correlation between trees is also reduced, resulting in a reduction in generalization error, which is the strength of RF. The Gini index, which provides a measure of impurity of a given class with respect to the rest of the classes, is used to select the best predictor among the randomly selected predictor variables available at each node (Breiman 2001; Ghimire, Rogan, and Miller 2010; Rodríguez-Galiano et al. 2011; Ghimire et al. 2012; Rodríguez-Galiano et al. 2012a, 2012b).

Study area

The study area is a single large mine complex, the Hobet-21 mountaintop mine in Boone and Lincoln counties, West Virginia, USA (Figure 1). It is the largest surface mine complex in the Appalachian region (Keene and Skousen 2010), with a wide variety in age of disturbance, vegetation, and land cover. Historical imagery shows that some of the mining disturbance predates 1987, while portions of the mine were still active at the time of this study. Multiple remotely sensed data sets are available for this site, including NAIP orthophotography, light detection and ranging (LiDAR) data, multiple RapidEye scenes, and high-resolution imagery from Pictometry International Corporation (Rochester, New York). Multiple data sources allowed the collection of accurate training and assessment data, making this an excellent location to study surface mine mapping. A total area of 5500 ha was classified within the extent of the Hobet-21 mine.

Methods

Data

The characteristics of the NAIP and RapidEye data used are described in Table 1. The RapidEye data were provided by the supplier as an orthorectified product (termed 3A), which has a nominal error of potentially less than 1 pixel under ideal circumstances (e.g., nadir view and flat terrain) (RapidEye 2009). The primary RapidEye data over the study site were acquired at a view angle of 6.72° (i.e., close to nadir), but the topography of the study site is complex; therefore, the error may be greater than 1 pixel. Nevertheless, a visual inspection of an overlay of the two image data sets did not show noticeable

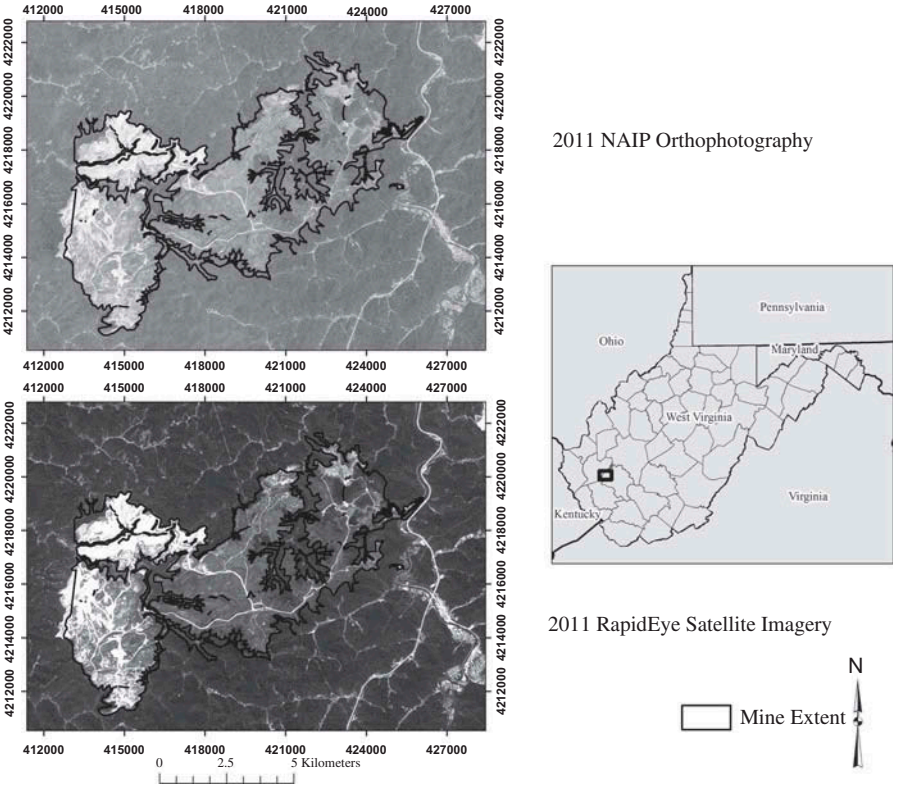

Figure 1. Hobet-21 surface mine complex. Base image for the top image is the NAIP orthophotography acquired on 14 July 2011. Base image for the bottom image is the RapidEye satellite imagery (© (2013) BlackBridge S.àr.l. All rights reserved) acquired on 1 August 2011. The depicted mine extent is based on the surface mining permit obtained from WVDEP. The map is projected in NAD83 UTM Zone 17 N.

Table 1. Characteristics of NAIP orthophotography and RapidEye satellite imagery used.

Sensor and data attributes	National Agriculture Imagery Program (NAIP) orthophotography	RapidEye satellite imagery
Collection Date	14 July 2011	1 August 2011
Platform	Aircraft	Satellite
Sensor type	Digital camera (Intergraph Z/I Imaging Digital Mapping Camera)	Multispectral push broom imager
Spatial resolution	1 m GSD	6.5 m GSD (resampled to 5 m)
Spectral resolution	Blue (400–580 nm) Green (500–650 nm) Red (590–675 nm) NIR (675–850 nm)	Blue (440–510 nm) Green (520–590 nm) Red (630–685 nm) Red edge (690–730 nm) NIR (760–850 nm)
Radiometric resolution	12-bit (provided at 8 bit)	12-bit (provided at 16 bit)
Temporal resolution	Potentially every other growing season	Potentially daily (5.5 days at nadir)
Seasonality	Growing season (leaf-on)	Potentially multiseasonal

differences, suggesting that the registration was adequate for the analysis and not likely to affect the classification. However, as a precaution in case of co-registration problems or potential changes in land cover between the data set acquisitions, each pixel in the training and assessment data sets was checked to ensure that the assigned land cover class for that location was the same in each data set. Any sample that failed this test (less than 5% of all samples) was excluded from the analysis. In this study, single-date imagery in which the training, assessment, and image data were on the same scale was used; as a result, as suggested by Song et al. (2001), DN values were used as delivered, not converted to reflectance or with atmospheric corrections applied.

The NAIP orthophotography was collected with an Integraph Z/I Imaging Digital Mapping Camera (DMC) on 14 July 2011, and thus leaf-on conditions during the growing season were captured. This imagery has a 1 m GSD and four spectral bands (blue, green, red, and NIR) as described in Table 1. The data were provided by the Aerial Photograph Field Office of the USDA Farm Service Agency as uncompressed quarter quadrangles. These tiles were mosaicked to produce a single image covering the mine.

The RapidEye data were collected on 1 August 2011. The scene has an average center azimuth angle of 279.6°, sun azimuth of 165.1°, and sun elevation of 69.6°. The satellites have sensors with five spectral bands, as described in Table 1 (Tyc et al. 2005). The GSD of the system is 6.5 m. The 3A product used in this study has radiometric, sensor, and geometric corrections applied and is orthorectified to a 5 m grid. The Hobet-21 mine complex is covered by a single image tile; so there was no need to mosaic multiple tiles.

A second RapidEye image was also used in the analysis in order to assess the benefit of multiseasonal data. The image was collected on 1 April 2010, prior to spring leaf out. The scene has an average center image view angle of –2.82°, azimuth angle of 110.2°, sun azimuth of 171.2° and sun elevation of 56.5°. Although the multiseasonal data were combined to assess the potential benefits of multiseasonal data, which is not possible with NAIP, the primary objective of this study was to compare leaf-on, single-date imagery.

Image processing

Resampling was necessary in order to compare the images at a common cell size. As the RapidEye data are collected at a 6.5 m resolution and then subsequently resampled to 5 m, the coarsest resolution of the two data sets, 6.5 m, was used as the common cell size for analysis. Due to the much smaller original cell size of the NAIP imagery (1 m), it was resampled using pixel aggregation to 6.5 m, while the RapidEye data were resampled using cubic convolution. As cubic convolution resampling may smooth the image, nearest neighbor resampling was also tested, and no statistical difference in classification was noted between the resulting classifications using the two resampling methods. As a result, cubic convolution was used in the subsequent analyses. Also, each image was converted to an 8-bit stretch using a linear rescaling in order to allow for comparison of the images at a common radiometric scale.

GLCM texture measures (Haralick, Shanmugam, and Dinstein 1973) were calculated from the NAIP orthophotography. To facilitate comparison between the data sets, the texture calculations were applied to the imagery at the common 6.5 m pixel size. The following textural bands were calculated: variance, homogeneity, contrast, dissimilarity, entropy, second moment, and correlation. The mean was not included as the impact of smoothing the data was assessed separately. The textural measures were calculated from the red and green bands and also the first and second principle component bands. In order to assess the texture at different spatial scales, three different offsets, 3, 2, and 1 pixels,

were tested. The kernel size, which determines the amount of smoothing (Warner 2011), was set to be close to the minimum possible given the offset distance. The kernel size was set at 11 × 11 for the 3 pixel offset, 7 × 7 for the 2 pixel offset, and 5 × 5 for the offset of 1. Texture was calculated in the vertical, horizontal, and diagonal directions and then averaged in order to produce a composite, non-directionally sensitive value (Warner 2011). A total of 84 textural measures were produced.

A visual comparison of the 6.5 m NAIP and RapidEye data suggested that the latter image has a smoother texture (i.e., more autocorrelation), even though both images had the same nominal pixel size. Therefore, as an experiment, the 6.5 m NAIP was smoothed with a Gaussian low-pass filter with kernel sizes of 3 × 3, 5 × 5, and 7 × 7 pixels. Although applying a simple Gaussian low-pass filter does not replicate the modulation transfer function (Myneni et al. 1995) of the RapidEye sensor, this method provides a simple means to assess whether or not blurring the image (i.e., reducing the effective spatial resolution) would improve the classification accuracy of the NAIP data.

In summary, the NAIP imagery provided four bands, or predictor variables, for the classification. These four bands were used in the classification algorithms at the original cell size (1 m), resampled to 6.5 m, and also smoothed using the Gaussian filter. After calculating the textural measures, 88 predictor variables were available from NAIP (the four image bands and 84 textural measures). RapidEye provided five image bands, which were used in the classification algorithms at the original cell size, resampled to 6.5 m, and reduced to an 8-bit radiometric range. Combining the primary, leaf-on RapidEye data with the leaf-off scene resulted in a total of 10 predictor variables available to classify land cover.

Land cover classes

Five land cover classes were mapped within the surface mine: forested (not disturbed), reclaimed-herbaceous vegetation, reclaimed-woody vegetation, barren, and water (Table 2). In order to assess potential increases in classification accuracy when the number of classes is reduced, the data were also classified with the forested and reclaimed-woody vegetation classes combined, resulting in a four-class problem: woody vegetation, herbaceous vegetation, barren, and water. In addition, the data were also classified with the vegetation classes combined, resulting in just three classes: vegetated, barren, and water.

Training data

Training data polygons were delineated by manual interpretation of multiple sources including the following: NAIP orthophotography, RapidEye imagery (both leaf-on and

Table 2. Land cover class definitions.

Class	Description
Forested (not disturbed)	Land dominated by mature, woody vegetation that has not been directly disturbed by surface mining; mature forest that generally represents pre-mining conditions of the slopes
Reclaimed-herbaceous vegetation	Reclaimed areas dominated by herbaceous/non-woody vegetation
Reclaimed-woody vegetation	Reclaimed areas dominated by clumped or clustered woody plants that include shrubs and immature trees
Barren	Barren land lacking vegetation, manmade structures, haul roads, active quarries, lands disturbed by mining
Water	Water, including retention ponds, streams, and standing water

Table 3. Number of training polygons, areas (ha) of training polygons, and number of randomly selected training pixels manually interpreted for each land cover class based on manual interpretation of multiple sources.

Land cover class	Number of training polygons	Area of training polygons (ha)	Number of training pixels randomly selected from within the training polygons
Forested (not disturbed)	81	77.0	1000
Reclaimed-herbaceous vegetation	85	25.5	1000
Reclaimed-woody vegetation	55	8.7	1000
Barren	64	21.7	1000
Water	30	1.7	386

leaf-off), 1 m resolution Pictometry orthophotography, LiDAR-derived topographic slope, first return intensity, and normalized digital surface model raster grids. As these data span the period from March 2010 (Pictometry orthophotography) to August 2011 (most recent RapidEye image), there was potential for landscape change during the data collection period due to mine expansion or reclamation activities. Therefore, the training data were screened for potential change, and only samples where the land cover appeared to be the same in all data sets were used. The number of training polygons collected is summarized in Table 3.

The software tool Geospatial Modeling Environment (Beyer 2012) was used to draw a random sample of training pixels from the polygons; 1000 pixel examples of each class were randomly selected. Because water is not a common cover type in the mine complex, it was not possible to collected 1000 examples for that class. In total, 4386 pixels were utilized to train the models as summarized in Table 3.

Classification methods

In order to generalize the results of this study, two different per-pixel classifiers were utilized, SVM and RF, as implemented within the statistical software tool R (R Development Core Team 2012), using the e1071 package (Meyer et al. 2012) and the randomForest package (Liaw and Wiener 2002), respectively.

Optimization of the parameters for the SVM and RF algorithms was carried out using a 10-fold cross-validation grid search approach using the e1017 package in R (Meyer et al. 2012). This tuning function partitions the data into 10 disjoint training sets, using a random assignment. The classifier is then trained 10 times, and each time the remaining 90% of the data not used for training in that instance is used for validation to test for optimal parameter settings. Overall accuracy is calculated as the average across the 10 cross-validations. This function offers a means to test parameter combinations relative to how well withheld training data are classified. Once a coarse grid search using a wide range of parameter values was performed, a second grid search of values centered on the initial optimal value, and with a narrow range, was performed in order to further refine the estimate. Optimization was performed for each classification separately to ensure optimal parameters for the particular set of predictor variables tested. For SVM, a radial basis function kernel was used, and it was necessary to optimize the cost (C) and gamma (γ) parameters using the tuning function. For RF, it was necessary to optimize the number of predictor variables randomly sampled as candidates at each node (m) and also the number of trees to grow (k).

Error assessment

Classification accuracy was assessed through an independent data set, which did not use any pixel from the training data or from within the training polygons from which they were selected. A stratified sample for the accuracy assessment was chosen because reclaimed-woody vegetation and water classes are not common cover types in the study area and thus would typically be only rarely selected in a purely random approach. A total of 1300 pixels were randomly selected, 260 in each of the five land cover classes.

For each of the validation pixels, the dominant land cover category over the 6.5 × 6.5 m nominal area of the pixel was visually interpreted from the Pictometry, RapidEye, NAIP, and LiDAR data. As one goal of this work was to compare the statistical difference in map output, it was necessary to use the same validation data for all predictor variable combinations. Also, only validation points that were interpreted as the same cover type in all of the reference data sets were used in the analysis. This may have resulted in a slight inflation of map accuracy as mixed pixels may have been excluded.

From the randomly selected accuracy assessment data, error matrices and kappa statistics were produced. Since stratified random sampling was implemented, it was necessary to calculate the overall accuracy by weighting the class accuracies by the proportion of land cover within the study area (Stehman and Foody 2009). Following Pontius and Millones (2011), quantity disagreement and allocation disagreement were also calculated. Quantity disagreement provides a measure of error in the proportions of the categories, while allocation disagreement provides a measure of error in the spatial allocation of the categories (Pontius and Millones 2011). Allocation and quantity disagreement sum to overall error. Although the value of kappa is questioned (Pontius and Millones 2011), the statistic was nevertheless provided for potential comparisons to other studies.

In order to evaluate the statistical significance of any differences in the classifications, the results were compared on a pairwise basis using McNemar's test (Dietterich 1998; Foody 2004). First, 2-by-2 confusion matrices were produced that summarized the number of pixels correctly classified by both of the methods being compared, the number incorrectly classified by both methods, and the number classified correctly by one attempt but not the other. McNemar's test is a test of statistical difference that generates a z-score from this 2-by-2 matrix and is based on a chi-square test that compares the distribution of error expected under the null hypothesis (that the classifications have the same error rates) and the observed error. A z-score larger than 1.645 indicates a 95% confidence of statistical significance for the one-directional test of whether one classification is more accurate than the other (Bradley 1968; Dietterich 1998; Foody 2004; Agresti 2007).

Importance in RF model

The RF algorithm generates an estimate of predictor variable importance during training by excluding each variable sequentially and recording the resulting increased oob error (Breiman 2001; Rodríguez-Galiano et al. 2012a, 2012b). This feature of RF was used as a means to assess the importance of predictor variables, including raw image bands and textural measures. For example, the importance of the NAIP and RapidEye spectral bands were assessed in a model using both image data sets.

Results and discussion

Table 4 compares the NAIP and RapidEye data at the original cell sizes of each data set (i.e., 1 m for NAIP, 5 m for RapidEye), a common cell size (6.5 m), a common

Table 4. Comparison of map accuracy for NAIP and RapidEye.

Method	Parameter	NAIP	RapidEye	McNemar *z*-score
Comparisons at original cell size (1 m NAIP, 5 m RapidEye)				
SVM	OA (%)	81.2	90.8	8.062*
	K (%)	74.5	87.4	
	AD (%)	10.1	2.9	
	QD (%)	8.7	6.3	
RF	OA (%)	80.0	90.6	8.322*
	K (%)	73.0	87.1	
	AD (%)	10.2	3.4	
	QD (%)	9.8	6.0	
Comparisons at common cell size (6.5 m NAIP and RapidEye)				
SVM	OA (%)	83.6	90.3	5.852*
	K (%)	77.8	86.6	
	AD (%)	7.0	3.5	
	QD (%)	9.4	6.2	
RF	OA (%)	83.4	90.9	6.041*
	K (%)	77.4	87.4	
	AD (%)	8.8	4.3	
	QD (%)	7.8	4.8	
Comparison at same radiometric scale (8 bit) and original cell size (1 m NAIP, 5 m RapidEye)				
SVM	OA (%)	81.2	90.5	7.673*
	K (%)	74.5	87.1	
	AD (%)	10.1	3.3	
	QD (%)	8.7	6.2	
RF	OA (%)	80.0	91.1	8.806*
	K (%)	73.0	87.8	
	AD (%)	10.2	3.0	
	QD (%)	9.8	5.9	
Comparison at common cell size (6.5 m) and same radiometric scale (8 bit)				
SVM	OA (%)	83.6	90.2	5.826*
	K (%)	77.8	86.5	
	AD (%)	7.0	3.2	
	QD (%)	9.4	6.6	
RF	OA (%)	83.4	91.1	6.652*
	K (%)	77.4	87.7	
	AD (%)	8.8	3.3	
	QD (%)	7.8	5.6	

Notes: OA = Overall accuracy, K = Kappa statistic, AD = Allocation disagreement, QD = Quantity disagreement. A *z*-score for McNemar's test larger than 1.645 (*) indicates a 95% confidence interval of statistical significance for the one-directional test of whether one classification is more accurate than the other.

radiometric resolution (8 bit) and the original cell sizes, and a common cell size and radiometric resolution (6.5 m, 8 bit). At the original cell size, the RapidEye imagery was found to provide an overall accuracy for mapping the five classes 9.6% greater than the NAIP imagery using the SVM algorithm and 10.6% greater using the RF algorithm, and both these differences were found to be statistically significant using McNemar's test [*z*-score (SVM) = 8.062 and *z*-score (RF) = 8.322]. This higher accuracy for RapidEye data compared to NAIP is also reflected in kappa, allocation disagreement, and quantity disagreement; the classification of the RapidEye data had lower allocation and quantity disagreement and a higher kappa.

With a coarsening of the NAIP data to a 6.5 m pixel, accuracy increased by 2.3% for the SVM algorithm and by 3.4% for the RF algorithm. Both increases were shown to be

statistically significant [z-score (SVM) = 2.668 and z-score (RF) = 3.621]. Reduced spatial resolution associated with an increase in the cell size or with smoothing may result in a less complex and heterogeneous signatures for the map classes, producing increased classification accuracy (Latty et al. 1985; Warner, Nellis, and Foody 2009). However, the NAIP 6.5 m data classification was still found to be less accurate than the RapidEye 6.5 m classification (Table 4) [z-score (SVM) = 5.852 and z-score (RF) = 6.041], though the z-scores were smaller than those for the comparison at the original cell sizes. Allocation and quantity disagreement were also reduced. This suggests that coarsening the NAIP data increased classification accuracy, but not to the level obtained from the RapidEye data. Baker et al. (2013) also found that coarsening NAIP improved classification accuracy for a pixel-based classification.

Table 4 suggests that radiometric resolution had little impact on the classification accuracy. For example, reducing the RapidEye data to an 8-bit radiometric range only decreased the accuracy by 0.3% using the SVM algorithm, which was not a statistically significant difference [z-score (SVM) = 1.347 and z-score (RF) = 0.714]. Also, the RapidEye data on an 8-bit scale were still found to be statistically more accurate than the NAIP data at the original cell size and at a common cell size. This finding suggests that decreasing the radiometric resolution, perhaps to reduce the file size, may not have a large impact on classification accuracy for this specific classification task.

Figure 2 shows the relative importance of predictor variables as estimated by the oob mean decrease in accuracy for a RF model using both the RapidEye and NAIP spectral bands at the original cell sizes. The analysis suggests that all of the RapidEye bands were more useful in the model than the NAIP bands. The greater importance of the RapidEye bands is likely due to a more consistent viewing geometry and illumination conditions for satellite data in comparison to an aerial image mosaic, perhaps resulting in a more consistent class signature throughout the extent of the mapped area. Figure 3 provides the same comparison of relative importance, except for the data at a common cell size (6.5 m). Once the NAIP data were coarsened, the NAIP bands take on a greater importance in the model, particularly the NAIP blue and NAIP NIR bands. This supports the finding discussed above: coarsening the data improves classification accuracy of NAIP orthophotography for this mapping task. As an improvement in classification was found when the NAIP data were coarsened to 6.5 m, the 6.5 m data were used for the rest of the analysis.

NAIP orthophotography is not always provided with a NIR band. For example, the 2009 collection for the state of West Virginia was only a true color collection. Therefore, the usefulness of the NIR band was tested here. Removing the NIR band from the analysis resulted in a 5.7% decrease in classification accuracy for the SVM algorithm and a 6.0%

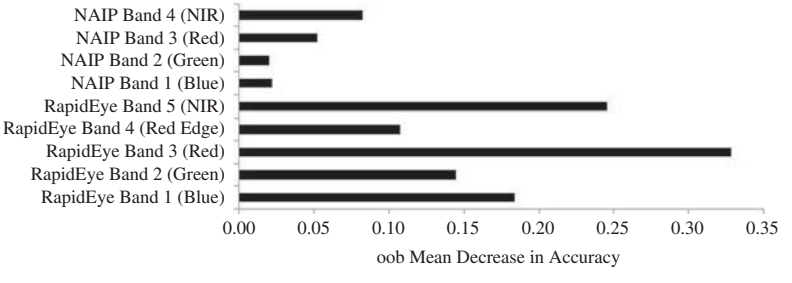

Figure 2. Relative importance of predictor variables for NAIP and RapidEye at provided cell size as estimated by the oob mean decrease in accuracy by RF.

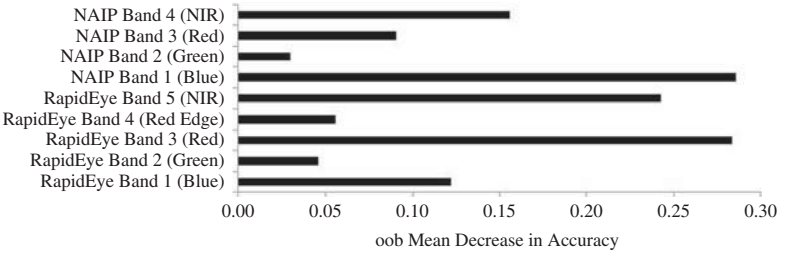

Figure 3. Relative importance of predictor variables for NAIP and RapidEye at common cell size (6.5 m) as estimated by the oob mean decrease in accuracy by RF.

decrease in accuracy for the RF algorithm. The overall accuracy dropped below 80% for both the SVM and RF models, and both allocation and quantity disagreement increased. Removing the NIR band resulted in a statistically significant decrease in accuracy for both classifiers [z-score (SVM) = 7.993 and z-score (RF) = 8.459]. This confirms the expected finding that there is merit in using the NIR band if NAIP data are to be used for classification tasks and NIR data are available.

As NAIP is acquired during the growing season, leaf-off data are not available, unlike satellite data where phenology can potentially be captured through multitemporal imagery. To assess the merit of multiseasonal data, leaf-on and leaf-off RapidEye data were compared, and then combined for a classification using leaf-on and leaf-off bands. Leaf-on data provided a statistically more accurate classification than the leaf-off data using both the SVM and RF algorithms [z-score (SVM) = 3.273 and z-score (RF) = 4.020]; however, the accuracy difference is much less than the difference between classification of NAIP and leaf-on RapidEye data. In comparison to the leaf-off data, classification of the leaf-on data resulted in only a 0.9% increase in accuracy using SVM and a 3.2% increase using RF; however, the differences were in both cases statistically significant [z-score (SVM) = 4.020 and z-score (RF) = 3.273]. This suggests that leaf-on data are marginally more useful for this classification task. When the data were combined in a multiseason layer stack, the accuracy statistically improved compared to the leaf-on and leaf-off models, resulting in a 3.7% increase over the leaf-on data using the SVM algorithm and a 3.3% increase using the RF algorithm. Figure 4 indicates that both leaf-on and leaf-off bands were important to the RF model. In summary, multiseasonal data

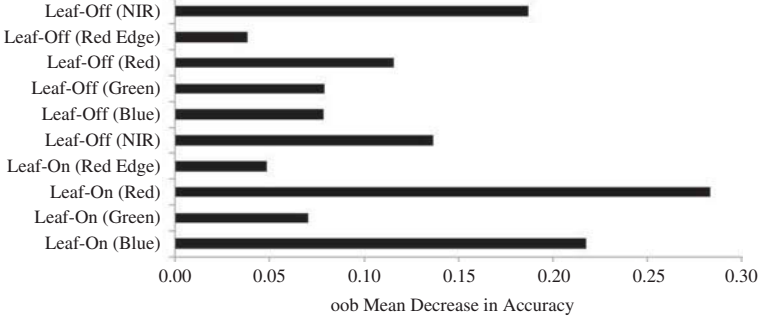

Figure 4. Relative importance of predictor variables for leaf-on and leaf-off RapidEye bands as estimated by the oob mean decrease in accuracy by RF.

Table 5. Error matrix for NAIP (6.5 m cell size, 8-bit) classification using RF.

		Reference data					
		Forested (not disturbed)	Reclaimed-herbaceous vegetation	Reclaimed-woody vegetation	Barren	Water	User's accuracy
Classification	Forested (not disturbed)	238	0	67	0	6	76.5%
	Reclaimed-herbaceous vegetation	10	232	11	8	12	84.0%
	Reclaimed-woody vegetation	12	21	182	0	15	79.1%
	Barren	0	7	0	241	22	89.3%
	Water	0	0	0	11	205	94.9%
	Producer's accuracy	91.5%	89.2%	70.0%	92.7%	78.8%	

Note: The overall accuracy is 83.4%.

increased map accuracy for this classification task. This is of concern for NAIP, as multiseasonal data are not available.

The error matrix for the NAIP (6.5 m) classification using RF is shown in Table 5, and the error matrix for the RapidEye classification using RF is shown in Table 6. Four classification results are shown in Figure 5, using both the RF and SVM classifiers. The error trends are very similar for both the SVM and RF models, with the reclaimed-woody vegetation class having the lowest producer's accuracy for both the NAIP and RapidEye classifications. This can be attributed to confusion with forest cover and somewhat to confusion with reclaimed-herbaceous cover. This is expected, as reclaimed-woody

Table 6. Error matrix for leaf-on RapidEye (6.5 m cell size, 12-bit) classification using RF.

		Reference data					
		Forested (not disturbed)	Reclaimed-herbaceous vegetation	Reclaimed-woody vegetation	Barren	Water	User's accuracy
Classification	Forested (not disturbed)	245	0	29	0	0	89.4%
	Reclaimed-herbaceous vegetation	0	238	13	3	2	93.0%
	Reclaimed-woody vegetation	15	4	218	0	0	92.0%
	Barren	0	18	0	251	19	87.2%
	Water	0	0	0	6	239	97.6%
	Producer's accuracy	94.2%	91.5%	83.8%	96.5%	91.9%	

Note: The overall accuracy is 90.9%.

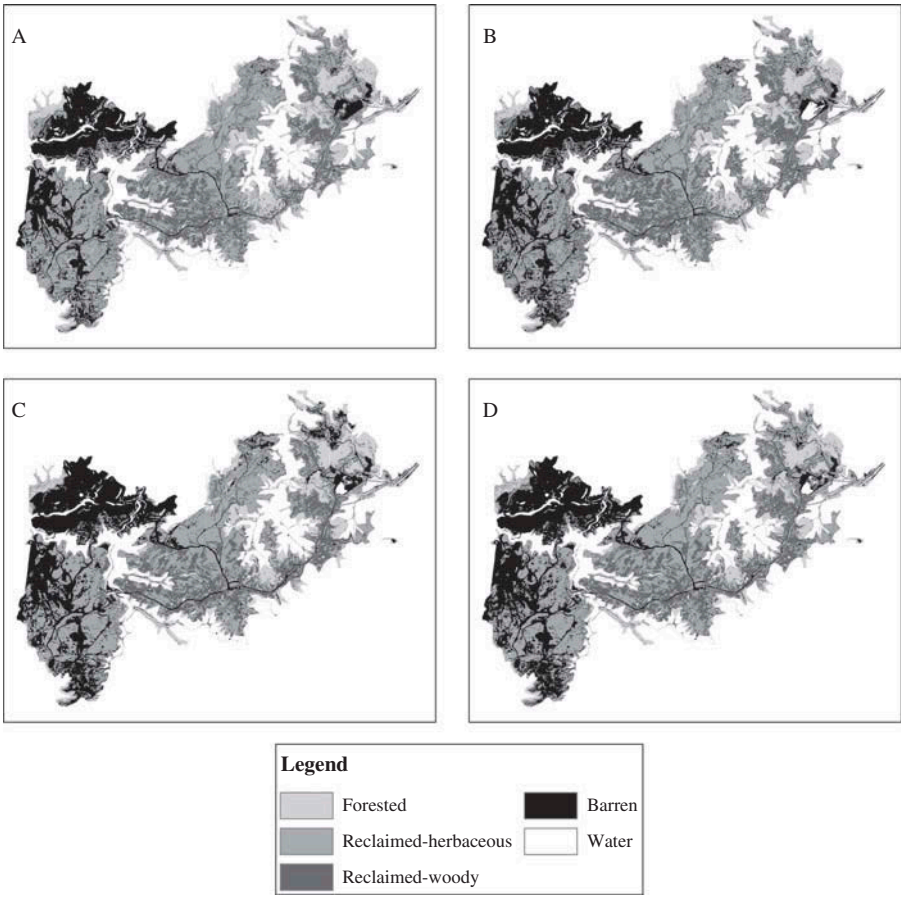

Figure 5. Land cover classification using (A) NAIP orthophotography (6.5 m) and SVM, (B) NAIP orthophotography (6.5 m) and RF, (C) RapidEye (6.5 m) and SVM, (D) RapidEye (6.5 m) and RF.

vegetation has similar spectral characteristics to both of these classes. For both image data sets, reclaimed-herbaceous vegetation was confused with barren cover; this confusion may be attributed to the patchy, heterogeneous nature of the vegetation in some reclaimed areas, which likely resulted in a mixed pixel problem. In the early stage of reclamation, in which vegetation is just beginning to regrow on the barren landscape, herbaceous cover is often present but can be sparse. In active surface mines where a wide range of disturbance and reclamation classes exist on a steep and heterogeneous landscape, the separation of these two classes is complex, and the boundary between them may be gradational as barren cover transitions to herbaceous cover.

For the NAIP classification, producer's accuracy for water was also generally low, likely as a result of the variability in water reflectance on mine sites. Spectral reflectance may vary between small settlement ponds, water retention ponds, and treatment ponds that may have a wide variety of chemical precipitates, dissolved chemicals, and sediment levels. This complicates the classification of water in surface mines.

In comparison to the NAIP classification, producer's accuracy was greater for all classes for the RapidEye classification using either the SVM or RF algorithm, and the user's accuracy was higher in comparison to NAIP for all classes except barren for the RF

model. This generally suggests that the increase in accuracy associated with RapidEye cannot be attributed to alleviating a single misclassification problem; RapidEye allows for enhanced separation of most classes in comparison to NAIP.

Since the reclaimed-woody vegetation class showed confusion with forest cover, it was combined with the forest class to produce a woody vegetation class and, in doing so, simplifying the classification to a four-class problem. This increased the classification accuracy of the NAIP (6.5 m) data by 5.4% using the SVM algorithm and 5.2% using the RF algorithm. Although the reduction in the number of classes resulted in a substantial improvement in classification accuracy for NAIP, these results were still significantly less accurate than the RapidEye classification with four classes [z-score (SVM) = 4.469 and z-score (RF) = 4.833]. Combining all vegetation classes to create a three-class problem increased the accuracy by 10.7% in comparison to the five-class classification for the NAIP classification using the SVM algorithm and 10.7% using the RF algorithm. Nevertheless, the RapidEye results with just three classes were still more accurate than the NAIP classification [z-score (SVM) = 4.181 and z-score (RF) = 3.579]. Even though RapidEye provided a higher classification accuracy, the NAIP classification accuracy was nevertheless high (SVM = 94.3% and RF = 96.7%), suggesting that NAIP may be appropriate if a relatively small number of classes are to be separated.

Although many studies suggest an improvement in classification with the inclusion of textural measures (e.g., Ghimire, Rogan, and Miller 2010; Rodríguez-Galiano et al. 2012a), in this study, incorporating all 84 textural measures derived from the NAIP data did not increase accuracy to that of the RapidEye spectral data alone. Improvements were observed for the NAIP spectral and texture bands compared to classification using only the image bands from NAIP, but only in some cases was this improvement statistically significant. On the other hand, in some cases, the addition of the textural bands decreased the overall accuracy, especially for the SVM algorithm. For all bands used (red, green, first principle component, and second principle component), the textural measures were found to be much less important than the spectral bands, suggesting they provide little additional information for the classification.

It has been suggested that increasing the number of predictor variables can cause a decrease in classification accuracy, contrary to what might be expected, a phenomenon known as the Hughes effect or "curse of dimensionality" (Hughes 1968; Rodríguez-Galiano et al. 2012a). For the RF algorithm, this decrease in performance with additional predictor variables was noted by Evans et al. (2011) and Hastie, Tibshirani, and Friedman (2009). As a result, feature selection was tested as a means to increase the classification accuracy when using measures of texture. The feature selection method implemented is after Murphy, Evans, and Storfer (2010) and uses a model selection approach that uses variable importance measures to select a model with reduced dimensionality and potentially increased classification accuracy. Models using the top 10%, 20%, 50%, 70%, 90%, and all the variables were tested. The greatest accuracy was found using the top 10% of the variables, which included all the image bands and a subset of the spectral bands. Even though an increase in accuracy was observed using the top 10% of the predictor variables, all of the image bands were nevertheless found to be of greater importance than the selected textural measures. Using the top 10% of the bands resulted in a statistically significant increase in accuracy relative to using all bands and textural measures (z-score = 2.090) and a statistically significant increase relative to using just the spectral bands (z-score = 4.016) for the RF algorithm. However, it was still less accurate than the RapidEye classification (z-score = 3.491).

Given that producing these texture bands is time-consuming, deciding upon which measures to use, appropriate kernel sizes and offset distances, and which spectral bands to calculate them from, can be subjective, and feature selection may be required to reap the full benefit of these measures, it is questionable whether there is merit in calculating these measures as the potentially slight classification improvement is outweighed by complexity and the time-consuming nature of the task.

The NAIP data were also smoothed using a Gaussian low-pass filter with kernel sizes of 3 × 3, 5 × 5, and 7 × 7 pixels. Smoothing was found to produce a greater increase in accuracy than adding texture bands. Smoothing with a 3 × 3 pixel kernel did not statistically increase the classification accuracy with the SVM algorithm (z-score = 0.686); however, it did for RF (z-score = 2.546). At kernel sizes of 5 × 5 and 7 × 7 pixels, a statistically significant increase was observed, and the highest classification accuracy for the five-class problem using the NAIP imagery was obtained using the RF algorithm with the image resampled to 6.5 m and a Gaussian low-pass filter with a 7 × 7 pixel kernel size applied. An overall accuracy of 88.1% was obtained (a 4.7% increase in accuracy relative to the NAIP 6.5 m classification without smoothing and a 8.1% increase in accuracy relative to the NAIP 1 m classification without smoothing); however, this classification was still found to be statistically less accurate than the classification of the RapidEye data without any smoothing (z-score = 2.965). The model using the top 10% of the image bands and textural measures was found to be statistically comparable to this output (z-score = 0.825).

Comparing the results using the textural measures and the Gaussian low-pass filter indicates that smoothing the data yielded a higher increase in classification accuracy than utilizing texture unless feature selection was implemented. Perhaps this is a result of reduced heterogeneity within class signatures. Minimizing this heterogeneity using smoothing may be an alternative to trying to characterize the heterogeneity using textural calculations.

Conclusions

In this study, the accuracy of classification with NAIP was statistically less accurate than that of RapidEye satellite imagery for classification of five classes within a surface mine permit. This suggests that RapidEye satellite data are more suited for this mapping task. Nevertheless, NAIP orthophotography has many characteristics that make it useful for surface mine mapping and monitoring, including its availability for multiple years, a general lack of cloud cover, contiguous coverage of large areas, availability, and low cost to the user.

Although the accuracy of classification with NAIP was not found to be comparable to that of RapidEye, the NAIP imagery provided a high classification accuracy when the number of classes was reduced to four or fewer classes. If the aim of a mapping project is only to differentiate vegetated and non-vegetated cover, or to differentiate woody vegetation from non-woody vegetation, NAIP may be adequate. Reducing the spatial detail, either by simply increasing the pixel size or by smoothing the data using a Gaussian low-pass filter, also increased classification accuracy. Textural measures from the GLCM were of value, but feature selection was necessary to select amongst the large number of derived texture variables. If NAIP data are to be used for classification, the NIR band is of value and should be acquired and used, if available.

In summary, NAIP orthophotography offers a wealth of data and should not be ignored as a potential data set for classification. Although perhaps not as useful as satellite imagery for image classification, NAIP does offer many favorable characteristics for

mapping changing terrains, such as the mountaintop mining region of the eastern United States.

Acknowledgements

The views, opinions, and recommendations expressed herein are solely those of the authors and do not imply any endorsement by ARIES employees, other ARIES-affiliated researchers, or industrial members. The contents are solely the responsibility of the authors and do not necessarily represent the official views of the USGS. Last, we would like to thank two anonymous reviewers for their helpful comments that greatly improved the manuscript.

Funding

Funding support for this study was provided by West Virginia View, Alderson Broaddus University, and the Appalachian Research Initiative for Environmental Science (ARIES). The project described in this publication was also supported in part by Grant/Cooperative Agreement Number [08HQGR0157] from the United States Geological Survey via a subaward from AmericaView.

This work was supported by the Appalachian Research Initiative for Environment Science (ARIES) and United States Geologic Survey under Grant [number 08HQGR0157].

References

Aerial Photography Field Office (APFO). 2012. "Imagery Programs: National Agriculture Imagery Program (NAIP)." http://fsa.usda.gov/FSA/apfoapp?area= home&subject= prog&topic=naip

Agresti, A. 2007. *An Introduction to Categorical Data Analysis*. 2nd ed. Hoboken, NJ: Wiley-Interscience.

Agüera, F., F. J. Aguilar, and M. A. Aguilar. 2008. "Using Texture Analysis to Improve Per-Pixel Classification of Very High Resolution Images for Mapping Plastic Greenhouses." *ISPRS Journal of Photogrammetry and Remote Sensing* 63 (6): 635–646. doi:10.1016/j.isprsjprs.2008.03.003.

Baker, B. A., T. A. Warner, J. F. Conley, and B. E. McNeil. 2013. "Does Spatial Resolution Matter? A Multi-Scale Comparison of Object-Based and Pixel-Based Methods for Detecting Change Associated with Gas Well Drilling Operations." *International Journal of Remote Sensing* 34 (5): 1633–1651. doi:10.1080/01431161.2012.724540.

Bernhardt, E. S., and M. A. Palmer. 2011. "The Environmental Costs of Mountaintop Mining Valley Fill Operations for Aquatic Ecosystems of the Central Appalachians." *Annals of the New York Academy of Sciences* 1223 (1): 39–57. doi:10.1111/j.1749–6632.2011.05986.x.

Beyer, H. L. 2012. *Geospatial Modeling Environment (Version 0.6.0.0)*. http://www.spatialecology.com/gme

Bozheva, A. M., A. N. Petrov, and R. Sugumaran. 2005. "The Effect of Spatial Resolution of Remotely Sensed Data in Dasymetric Mapping of Residential Areas." *Giscience & Remote Sensing* 42 (2): 113–130. doi:10.2747/1548–1603.42.2.113.

Bradley, J. V. 1968. *Distribution-Free Statistical Tests*. Englewood Cliffs, NJ: Prentice Hall.

Breiman, L. 1996. "Bagging Predictors." *Machine Learning* 24 (2): 123–140. doi:10.1007/BF00058655.

Breiman, L. 2001. "Random Forests." *Machine Learning* 45 (1): 5–32. doi:10.1023/A:1010933404324.

Burges, C. J. C. 1998. "A Tutorial on Support Vector Machines for Pattern Recognition." *Data Mining and Knowledge Discovery* 2 (2): 121–167. doi:10.1023/A:1009715923555.

Chica-Olmo, M., and F. Abarca-Hernández. 2000. "Computing Geostatistical Image Texture for Remotely Sensed Data Classification." *Computers & Geosciences* 26 (4): 373–383. doi:10.1016/S0098-3004(99)00118-1.

Cushnie, J. L. 1987. "The Interactive Effect of Spatial Resolution and Degree of Internal Variability within Land-Cover Types on Classification Accuracies." *International Journal of Remote Sensing* 8 (1): 15–29. doi:10.1080/01431168708948612.

Davies, K. W., S. L. Petersen, D. D. Johnson, D. B. Davis, M. D. Madsen, D. L. Zvirzdin, and J. D. Bates. 2010. "Estimating Juniper Cover from National Agriculture Imagery Program (NAIP) Imagery and Evaluating Relationships between Potential Cover and Environmental Variables." *Rangeland Ecology & Management* 63 (6): 630–637. doi:10.2111/REM-D-09-00129.1.

Dietterich, T. G. 1998. "Approximate Statistical Tests for Comparing Supervised Classification Learning Algorithms." *Neural Computation* 10 (7): 1895–1923. doi:10.1162/089976698300017197.

Drummond, M. A., and T. R. Loveland. 2010. "Land-Use Pressure and a Transition to Forest-Cover Loss in the Eastern United States." *Bioscience* 60 (4): 286–298. doi:10.1525/bio.2010.60.4.7.

Evans, J. S., M. A. Murphy, Z. A. Holden, and S. A. Cushman. 2011. "Modeling Species Distribution and Change Using Random Forest." In *Landscape Ecology: Concepts and Application*, edited by C. A. Drew, Y. F. Wiersma, and F. Huettman. New York: Springer.

Foody, G. M. 2004. "Thematic Map Comparison: Evaluating the Statistical Significance of Differences in Classification Accuracy." *Photogrammetric Engineering & Remote Sensing* 70 (5): 627–633. doi:10.14358/PERS.70.5.627.

Fritz, K. M., S. Fulton, B. R. Johnson, C. D. Barton, J. D. Jack, D. A. Word, and R. A. Burke. 2010. "Structural and Functional Characteristics of Natural and Constructed Channels Draining a Reclaimed Mountaintop Removal and Valley Fill Coal Mine." *Journal of the North American Benthological Society* 29 (2): 673–689. doi:10.1899/09-060.1.

Ghimire, B., J. Rogan, and J. Miller. 2010. "Contextual Land-Cover Classification: Incorporating Spatial Dependence in Land-Cover Classification Models Using Random Forests and the Getis Statistic." *Remote Sensing Letters* 1 (1): 45–54. doi:10.1080/01431160903252327.

Ghimire, B., J. Rogan, V. Rodríguez-Galiano, P. Panday, and N. Neeti. 2012. "An Evaluation of Bagging, Boosting, and Random Forests for Land-Cover Classification in Cape Cod, Massachusetts, USA." *Giscience & Remote Sensing* 49 (5): 623–643. doi:10.2747/1548-1603.49.5.623.

Guo, Q., M. Kelly, P. Gong, and D. Liu. 2007. "An Object-Based Classification Approach in Mapping Tree Mortality Using High Spatial Resolution Imagery." *Giscience & Remote Sensing* 44 (1): 24–47. doi:10.2747/1548-1603.44.1.24.

Hansen, M., R. Dubayah, and R. DeFries. 1996. "Classification Trees: An Alternative to Traditional Land Cover Classifiers." *International Journal of Remote Sensing* 17 (5): 1075–1081. doi:10.1080/01431169608949069.

Haralick, R. M., K. Shanmugam, and I. H. Dinstein. 1973. "Textural Features for Image Classification." *IEEE Transactions on Systems, Man and Cybernetics* 3: 610–621. doi:10.1109/TSMC.1973.4309314.

Hastie, T., R. Tibshirani, and J. Friedman. 2009. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. 2nd ed. New York: Springer.

Huang, C., L. S. Davis, and J. R. G. Townshend. 2002. "An Assessment of Support Vector Machines for Land Cover Classification." *International Journal of Remote Sensing* 23 (4): 725–749. doi:10.1080/01431160110040323.

Hughes, G. 1968. "On the Mean Accuracy of Statistical Pattern Recognizers." *IEEE Transactions on Information Theory* 14 (1): 55–63. doi:10.1109/TIT.1968.1054102.

Kazar, S. A., and T. A. Warner. 2013. "Assessment of Carbon Storage and Biomass on Minelands Reclaimed to Grassland Environments Using Landsat Spectral Indices." *Journal of Applied Remote Sensing* 7 (1): 073583. doi:10.1117/1.JRS.7.073583.

Keene, T., and J. Skousen, 2010, "Mine Spoil Reclamation with Switchgrass for Biofuel Production," In *2010 National Meeting of the American Society for Mining and Reclamation*, June 5–11, Pittsburgh, PA, 1–16.

Latty, R. S., R. Nelson, B. Markham, D. Williams, D. Toll, and J. Irons. 1985. "Performance Comparison between Information Extraction Techniques Using Variable Spatial Resolution Data." *Photogrammetric Engineering & Remote Sensing* 51 (9): 1459–1470.

Liaw, A., and M. Wiener. 2002. "Classification and Regression by Randomforest." *R News* 2 (3): 18–22.

Loosvelt, L., J. Peters, H. Skriver, H. Lievens, F. M. B. Van Coillie, B. De Baets, and N. E. C. Verhoest. 2012. "Random Forests as a Tool for Estimating Uncertainty at Pixel-Level in SAR Image Classification." *International Journal of Applied Earth Observation and Geoinformation* 19: 173–184. doi:10.1016/j.jag.2012.05.011.

Maxwell, A. E., T. A. Warner, M. P. Strager, and M. Pal. 2014. "Combining Rapideye Satellite Imagery and Lidar for Mapping of Mining and Mine Reclamation." *Photogrammetric Engineering & Remote Sensing* 80 (2): 179–189. doi:10.14358/PERS.80.2.179-189.

Meneguzzo, D. M., G. C. Liknes, and M. D. Nelson. 2013. "Mapping Trees outside of Forests Using High-Resolution Aerial Imagery: A Comparison of Pixel- and Object-Based Classification Approaches." *Environmental Monitoring and Assessment* 185: 6261–6275.

Meyer, D., E. Dimitriadou, K. Hornik, A. Weingessel, and F. Leisch. 2012. "E1071: Misc Functions of the Department of Statistics (E1071)." R Package Version 1.6-1. http://CRAN.R-project.org/package=e1071

Murphy, M. A., J. S. Evans, and A. S. Storfer. 2010. "Quantifying *Bufo Boreas* Connectivity in Yellowstone National Park with Landscape Genetics." *Ecology* 91: 252–261. doi:10.1890/08-0879.1.

Myeong, S., D. J. Nowak, P. F. Hopkins, and R. H. Brock. 2001. "Urban Cover Mapping Using Digital, High-Spatial Resolution Aerial Imagery." *Urban Ecosystems* 5: 243–256. doi:10.1023/A:1025687711588.

Myneni, R. B., S. Maggion, J. Iaquinta, J. L. Privette, N. Gobron, B. Pinty, D. S. Kimes, M. M. Verstraete, and D. L. Williams. 1995. "Optical Remote Sensing of Vegetation: Modeling, Caveats, and Algorithms." *Remote Sensing of Environment* 51: 169–188. doi:10.1016/0034-4257(94)00073-V.

Negley, T. L. 2002. "A Comparative Hydrologic Analysis of Surface-Mined and Forested Watersheds in Western Maryland." M.S. thesis, College Park, MD: University of Maryland.

Pal, M. 2005. "Random Forest Classifier for Remote Sensing Classification." *International Journal of Remote Sensing* 26(1): 217–222. doi:10.1080/01431160412331269698.

Pal, M., and P. M. Mather. 2005. "Support Vector Machines for Classification in Remote Sensing." *International Journal of Remote Sensing* 26 (5): 1007–1011. doi:10.1080/01431160512331314083.

Palmer, M. A., E. S. Bernhardt, W. H. Schlesinger, K. N. Eshleman, E. Foufoula-Georgiou, M. S. Hendryx, A. D. Lemly, et al. 2010. "Mountaintop Mining Consequences." *Science* 327: 148–149. doi:10.1126/science.1180543.

Phinn, S. R. 1998. "A Framework for Selecting Appropriate Remotely Sensed Data Dimensions for Environmental Monitoring and Management." *International Journal of Remote Sensing* 19 (17): 3457–3463. doi:10.1080/014311698214136.

Pontius, R. G. Jr., and M. Millones. 2011. "Death to Kappa: Birth of Quantity Disagreement and Allocation Disagreement for Accuracy Assessment." *International Journal of Remote Sensing* 32 (15): 4407–4429. doi:10.1080/01431161.2011.552923.

RapidEye. 2009. *Rapideye Standard Image Product Specifications*. Berlin: RapidEye AG.

Rathore, C. S., and R. Wright. 1993. "Monitoring Environmental Impacts of Surface Coal-Mining." *International Journal of Remote Sensing* 14: 1021–1042. doi:10.1080/01431169308904394.

R Development Core Team. 2012. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Rodríguez-Galiano, V. F., F. Abarca-Hernández, B. Ghimire, M. Chica-Olmo, P. M. Atkinson, and C. Jeganathan. 2011. "Incorporating Spatial Variability Measures in Land-Cover Classification Using Random Forest." *Procedia Environmental Sciences* 3: 44–49. doi:10.1016/j.proenv.2011.02.009.

Rodríguez-Galiano, V. F., M. Chica-Olmo, F. Abarca-Hernandez, P. M. Atkinson, and C. Jeganathan. 2012a. "Random Forest Classification of Mediterranean Land Cover Using Multi-Seasonal Imagery and Multi-Seasonal Texture." *Remote Sensing of Environment* 121: 93–107. doi:10.1016/j.rse.2011.12.003.

Rodríguez-Galiano, V. F., B. Ghimire, J. Rogan, M. Chica-Olmo, and J. P. Rigol-Sanchez. 2012b. "An Assessment of the Effectiveness of a Random Forest Classifier for Land-Cover Classification." *ISPRS Journal of Photogrammetry and Remote Sensing* 67: 93–104. doi:10.1016/j.isprsjprs.2011.11.002.

Rogan, J., J. Miller, D. Stow, J. Franklin, L. Levien, and C. Fischer. 2003. "Land-Cover Change Monitoring with Classification Trees Using Landsat TM and Ancillary Data." *Photogrammetric Engineering & Remote Sensing* 69 (7): 793–804. doi:10.14358/PERS.69.7.793.

Saylor, K. L. 2008. *Land Cover Trend Project: Central Appalachians*. Washington, DC: U.S. Department of the Interior, U.S. Geological Survey. http://landcovertrends. usgs.gov/east/eco69 Report.html

Sen, S., C. E. Zipper, R. H. Wynne, and P. F. Donovan. 2012. "Identifying Revegetated Mines as Disturbance/Recovery Trajectories Using an Interannual Landsat Chronosequence." *Photogrammetric Engineering & Remote Sensing* 78 (3): 223–235. doi:10.14358/PERS.78.3.223.

Simmons, J. A., W. S. Currie, K. N. Eshleman, K. Kuers, S. Monteleone, T. L. Negley, B. R. Pohlad, and C. L. Thomas. 2008. "Forest to Reclaimed Mine Land Use Change Leads to Altered Ecosystem Structure and Function." *Ecological Applications* 18 (1): 104–118. doi:10.1890/07-1117.1.

Song, C., C. E. Woodcock, K. C. Seto, M. P. Lenney, and S. A. Macomber. 2001. "Classification and Change Detection Using Landsat TM Data: When and How to Correct Atmospheric Effects?" *Remote Sensing of Environment* 75: 230–244.

Stehman, S. V., and G. M. Foody. 2009. "Accuracy Assessment." In *The SAGE Handbook of Remote Sensing*, edited by T. A. Warner, M. D. Nellis, and G. M. Foody. London: Sage Publications.

Su, L., and Y. Huang. 2009. "Support Vector Machine (SVM) Classification: Comparison of Linkage Techniques Using a Clustering-Based Method for Training Data Selection." *Giscience & Remote Sensing* 46 (4): 411–423. doi:10.2747/1548–1603.46.4.411.

Tarnavsky, E., D. Stow, L. Coulter, and A. Hope. 2004. "Spatial and Radiometric Fidelity of Airborne Multispectral Imagery in the Context of Land-Cover Change Analyses." *Giscience and Remote Sensing* 41 (1): 62–80. doi:10.2747/1548–1603.41.1.62.

Toutin, T. 2009. "Fine Spatial Resolution Optical Sensors." In *The SAGE Handbook of Remote Sensing*, edited by T. A. Warner, M. D. Nellis, and G. M. Foody. London: Sage Publications.

Townsend, P. A., D. P. Helmers, C. C. Kingdon, B. E. McNeil, K. M. de Beurs, and K. N. Eshleman. 2009. "Changes in the Extent of Surface Mining and Reclamation in the Central Appalachians Detected Using A 1976–2006 Landsat Time Series." *Remote Sensing of Environment* 113: 62–72. doi:10.1016/j.rse.2008.08.012.

Tyc, G., J. Tulip, D. Schulten, M. Krischke, and M. Oxfort. 2005. "The Rapideye Mission Design." *Acta Astronautica* 56: 213–219. doi:10.1016/j.actaastro.2004.09.029.

USEPA. 2005, *Final Programmatic Environmental Impact Statement (PEIS) on Mountaintop Mining/Valley Fills in Appalachia* (EPA 9–03-R-05002). Washington, DC: U.S. Environmental Protection Agency. http://www.epa.gov/region3.mtntop/eis2005.htm

Vapnik, V. N. 1995. *The Nature of Statistical Learning Theory*. New York: Springer-Verlag.

Warner, T. A. 2011. "Kernel-Based Texture in Remote Sensing Image Classification." *Geography Compass* 5 (10): 781–798. doi:10.1111/j.1749-8198.2011.00451.x.

Warner, T. A., M. D. Nellis, and G. M. Foody. 2009. "Remote Sensing Scale and Data Selection Issues." In *The SAGE Handbook of Remote Sensing*, edited by T. A. Warner, M. D. Nellis, and G. M. Foody. London: Sage Publications.

Wickham, J. D., K. H. Riitters, T. G. Wade, M. Coan, and C. Homer. 2007. "The Effect of Appalachian Mountaintop Mining on Interior Forest." *Landscape Ecology* 22: 179–187. doi:10.1007/s10980-006-9040-z.

Zégre, N. P., A. E. Maxwell, and S. Lamont. 2013. "Characterizing Streamflow Response of a Mountaintop-Mined Watershed to Changing Land Use." *Applied Geography* 39: 5–15. doi:10.1016/j.apgeog.2012.11.008.

Zipper, C. E., J. A. Burger, J. G. Skousen, P. N. Angel, C. D. Barton, V. Davis, and J. A. Franklin. 2011. "Restoring Forests and Associated Ecosystem Services on Appalachian Coal Surface Mines." *Environmental Management* 47 (5): 751–765. doi:10.1007/s00267-011-9670-z.