USING SATELLITE IMAGERY TO CHARACTERIZE LOCATIONS, AGES AND WOODY CANOPY COVER OF RECLAIMED SURFACE MINES IN APPALACHIA, USA[1]

S. Sen[2], C. E. Zipper, R. H. Wynne, P.F. Donovan, J.W. Coulston

Abstract: The Appalachian region of USA hosts diverse forests and abundant high-quality coal reserves. Surface mining methods are often used for coal extraction. Because common reclamation methods in past years have not restored forest vegetation, surface mining has created a diverse land base. Although some mined lands have been placed into managed uses, most have not. Little is known about the extent and nature of the land resource base created by surface coal mining in Appalachia. Here, we report on development of methods for interpreting imagery acquired by the Landsat satellites since the early 1980s to identify surface-mine land disturbances by date of mining, and to estimate current woody canopy on those mined areas. We have conducted these analyses working within a study area in southwestern Virginia's coalfield. The mined-area identification algorithm, when applied to an independent dataset, was found to identify mined/non-mined areas correctly with an overall accuracy of 89.1%, with 87.4% of mined areas within the independent dataset were classified correctly as mines. Incorrectly classified mines were often areas with low levels of vegetative cover nested within correctly classified mine areas. Preliminary results show that woody canopy cover on mined and reclaimed areas can be estimated successfully using Landsat (0.80 R^2). Future work will further develop these procedures and apply them over a test area.

Additional Key Words: Landsat, Coal Mine Reclamation, Mine Reforestation, Ecosystem Restoration.

---

[1] Paper was presented at the 2011 National Meeting of the American Society of Mining and Reclamation, Bismarck, North Dakota ***Reclamation: Sciences Leading to Success*** June 11 - 16, 2011. R.I. Barnhisel (Ed.) Published by ASMR, 3134 Montavesta Rd., Lexington, KY 40502.

[2] Susmita Sen is Graduate Research Assistant, Geospatial Environmental Analysis Program; Randolph H. Wynne is Professor, Forest Resources and Environmental Conservation; Carl E. Zipper is Associate Professor, and Patricia F. Donovan is Geospatial Analysis Manager, Department of Crop and Soil Environmental Sciences. Virginia Polytechnic Institute and State University, Blacksburg VA 24061. John W. Coulston is Supervisory Research Forester with the USDA Forest Service, Forest Inventory and Analysis Unit, in Knoxville, TN.

Proceedings America Society of Mining and Reclamation, 2011 pp 567-590
DOI: 10.21000/JASMR11010567
http://dx.doi.org/10.21000/JASMR11010567

Introduction

Coal mining has resulted in disturbance and reclamation of lands throughout the coal-bearing areas of eastern USA's Appalachian Mountains. The federal Surface Mining Control and Reclamation Act (SMCRA), passed in 1977, established minimum reclamation standards for coal mined lands throughout the USA, including Appalachia. US Office of Surface Mining data indicate that more than 600,000 hectares have been mined under SMCRA in the Appalachian region. Within the coalfields of southwestern Virginia, more than 40,000 hectares have been mined under SMCRA (US OSMRE) while over 20,000 hectares of land were affected by the pre-1977 mining activities (D'Appolonia, Inc. 1980).

In the years following SMCRA's implementation, Appalachian surface coal mines were often reclaimed with herbaceous vegetation which satisfied regulatory standards by establishing vegetative cover suitable for post-mining land use over the SMCRA-mandated bonding period, five years for active mines (Angel et al., 2005). Although some mined lands were reclaimed to support grazing and others have been reclaimed to support commercial or industrial uses, the majority are unmanaged. Reclamation with woody vegetation to support uses such as wildlife habitat and unmanaged forest have become more common in recent years (Angel et al., 2009; US GAO, 2009). On lands reclaimed using conventional SMCRA methods, mine soils were often reclaimed using practices that satisfied legal standards but hindered restoration of the hardwood forests that occur extensively as native vegetation throughout the area (Angel et al., 2005). While some of these lands have been planted with native trees such as eastern white pine (*Pinus strobus*) and black locust (*Robinia pseudoacacia*) and others may reforest with native trees naturally, many remain in predominantly herbaceous cover for extended periods (Zipper et al., 2007). It is possible for post-SMCRA mined lands to be converted to productive native woody vegetation, native forest trees and/or faster-growing biomass-producing species through application of cultural treatments (Evans et al., 2010; Fields-Johnson et al., 2008; Skousen et al., 2009), but such conversions require investment. Mined lands with soils that are not well suited for forest trees can become dominated by low-productivity woody vegetation including the invasive shrub autumn olive (*Elaeagnus umbellate)*, which acts as an obstacle to conversion by increasing the cost of necessary cultural treatments (Burger et al., 2011).

Although numerous research studies have been conducted on individual or small numbers of mine sites, there is no unified database that documents the status of post-SMCRA mined lands, collectively. Lack of knowledge about the extent and nature of the land resource base created by post-SMCRA coal mining is an obstacle to development of policies and strategies capable of improving the use and management of these lands. The status of these previously mined and reclaimed lands, including current uses and vegetative cover, and their capability to serve as renewable natural resources is a natural resource management concern.

Satellite-borne sensors are well suited to the task of characterizing the Appalachian mined land resource base. Satellite imagery is used commonly for detecting and monitoring changes on the earth's surface by providing consistent and repeatable measurements of land attributes. They can be applied over time and without the cost and difficulty of obtaining legal access and physically visiting mined properties. Images from the Landsat satellites have a spatial resolution adequate to locate mined lands and are multispectral – meaning that they detect surface-reflected radiation, both visible and non-visible, within several well-defined wavelength bands – which makes them well-suited for vegetation analysis. Landsat imagery is available from a 37-year public-domain archive for no cost and without copyright restrictions. We are developing Landsat data interpretation and analysis methods for potential use in characterizing the post-SMCRA mined land base in Appalachia and its vegetation.

Here, we report on studies intended to develop Landsat data interpretation methods that will:

1. Identify reclaimed mined areas, by time of mining / reclamation and their spatial extent.
2. Estimate woody canopy cover on these reclaimed mine areas.

Methods

Study area

The study area encompasses the coal-mining areas in 4 counties of southwestern Virginia: Wise, Dickenson, Buchanan and Russell (Fig. 1). These areas are heavily surface mined.

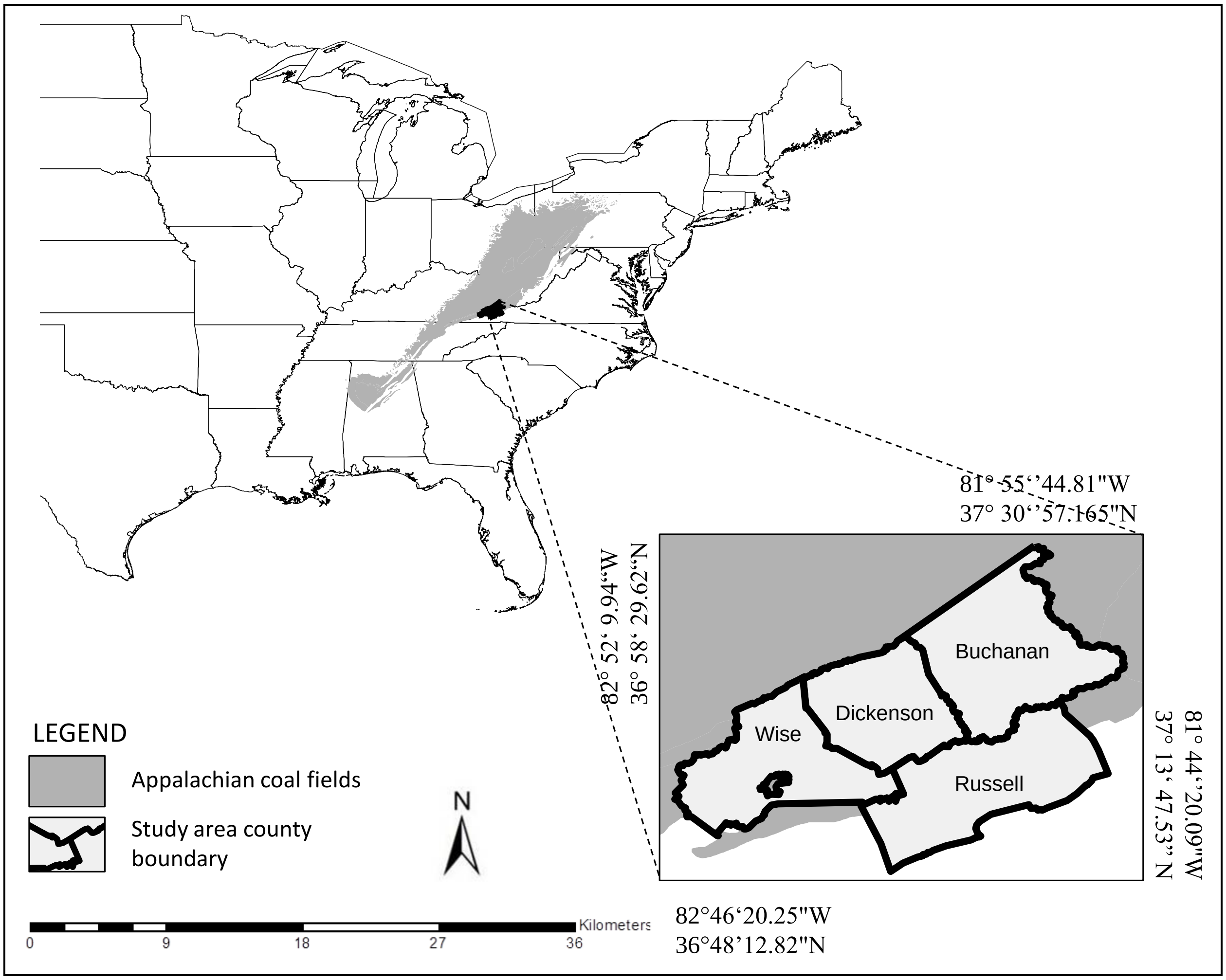

Figure 1. The study area encompasses four counties of southwestern Virginia where coal mining is conducted.

Objective 1: Identify reclaimed mined areas, by time of mining / reclamation and their spatial extent.

Image and Ancillary Data. The dataset used for analysis consisted of 23 Landsat images (Table 1), all from WRS-2 path 18 row 34. Best available leaf-on images for each year were chosen. Two years (1992 and 1996) are missing because suitable cloud-free leaf-on images were not available. All images used were acquired as level 1T product in a standard terrain-corrected form. Co-registration was verified and radiometric and atmospheric corrections were done using the Landsat Ecosystem Disturbance Adaptive System (LEDAPS) routine (Masek et al., 2006). The LEDAPS surface reflectance product was used. Clouds and cloud shadows were visually identified and eliminated from individual images by manual digitization. Additional geospatial data were acquired for training and validation (Table 2).

Table 1. Landsat image dates and assigned numbers in the chronosequence

Image Date	Chronosequence Number	Image date	Chronosequence Number
9/17/1984	1	5/19/1998	13
9/20/1985	2	9/3/1999	14
6/19/1986	3	6/9/2000	15
6/6/1987	4	8/15/2001	16
6/8/1988	5	5/22/2002	17
6/11/1989	6	6/2/2003	18
6/30/1990	7	9/24/2004	19
9/21/1991	8	9/11/2005	20
6/6/1993	9	7/12/2006	21
8/28/1994	10	9/17/2007	22
8/31/1995	11	9/3/2008	23
9/5/1997	12		

Note: all images from Landsat Thematic Mapper (™) except 5/22/2002, from the Enhanced Thematic Mapper Plus sensor.

Table2. Data sources used to develop training and validation points for objective 1.

Type	**Data**	**Source**	**Dates**
Aerial photos	Digital Ortho Quarter Quads (DOQQ)	U.S Geological Survey (USGS)	1996-1999
	National Agricultural Imagery Program (NAIP)	U.S Department of Agriculture (USDA)	2003, 2005, 2008
	Virginia Base Mapping Program (VBMP)	Virginia Geographic Information Network	2002, 2005, 2007
Geospatial data layers	Landuses, National Land Cover Dataset (NLCD)	Multi Resolution Land Characteristic Consortium (MRLC)	1992 and 2001
	Road data	Virginia Department of Transportation (VDOT)	2000

Developing a Reclaimed Mine Classification Model. The area's primary landcover is forest, but numerous non-mining forest disturbances also occur, including industrial, commercial, and residential development, and transportation infrastructure. We combined these non-forest non-mining landuse types as a single class called "urban" for study purposes. The study area contains little agriculture, and most of that which does occur is livestock grazing on reclaimed mines.

We developed a classification algorithm in FORTRAN 95 to discriminate mined from urban and forested areas. The algorithm operates by analyzing vegetation index (VI) data derived from the Landsat images for each discrete spatial unit as a multitemporal sequence. A VI is a dimensionless radiometric measure that functions as an indicator of relative abundance and activity of green biomass (Jensen, 2000). Mining disturbances are expected to have a multitemporal VI signature that represents a disturbance-recovery sequence (Fig., 2). We hypothesized that mines can be discriminated from forests and from urban disturbances by three diagnostic parameters derived from the disturbance-recovery sequence:

1. Disturbance Minimum (*Dmin*): Mining causes a sharp drop in VI due to vegetation removal. We expected the minimum VIs for mines to be lower than minima for less drastic forest disturbances such as fire and forest harvest; lower than the minima for many existing urban areas that retain vegetative cover (e.g. existing residential developments); and lower than for urban disturbances that did not fully remove vegetative cover over large, contiguous areas.

2. Recovery slope (*Rslope*): On mine sites reclaimed under SMCRA, the VI was expected to exhibit a rapid increase after reclamation; this occurs in response to the SMCRA-mandated revegetation standards that require rapid establishment of vigorous herbaceous cover. A mining recovery VI is expected to have steeper recovery slopes than most urban development because urban disturbance areas affected are usually not fully revegetated, and/or their vegetation is managed after establishment.

3. Recovery Maximum (*Rmax*), the maximum VI value within a recovery period: Because vegetation typically develops rapidly, without management such as cutting or trimming, and over entire reclaimed mine areas, a diagnostic maximum value is reached or exceeded within a defined recovery period. Urban development VIs are expected to remain generally at lower

levels, even after all construction and revegetation are completed, because only a portion of such areas are often revegetated and/or revegetated areas are managed.

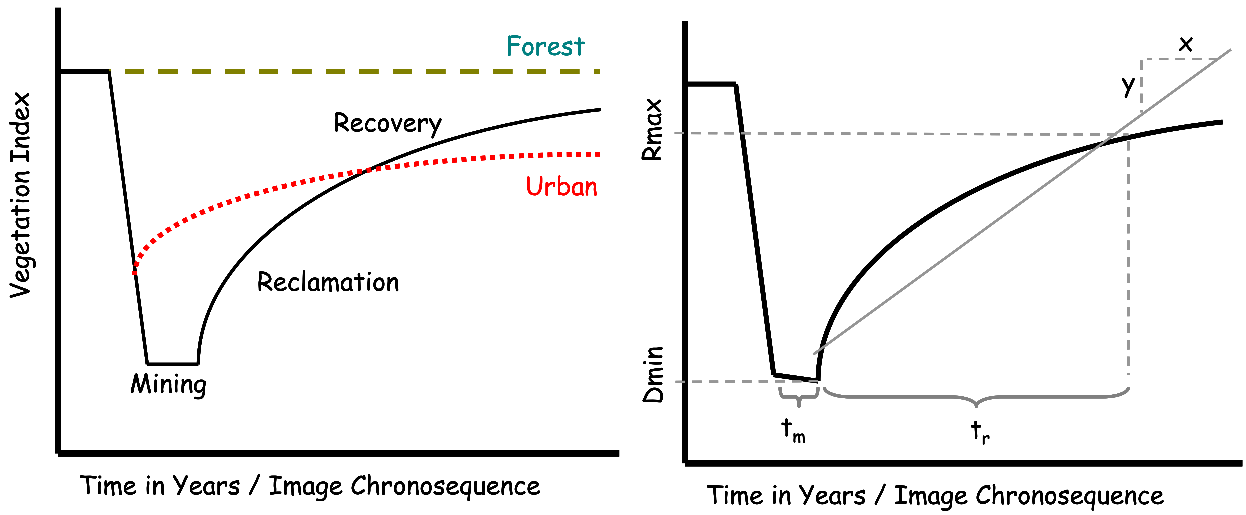

Figure 2. (Left) Hypothetical disturbance-recovery sequence produced by mining and reclamation, contrasted to an urban disturbance and undisturbed forest; (Right) illustration of the diagnostic parameters used for automated discrimination of mining from other disturbances for the mining-reclamation-recovery process, with t_m = period of mining disturbance and t_r = diagnostic recovery period (7 images in our analysis): *Dmin* = disturbance minimum, *Rmax* = recovery maximum, and *Rslope* calculated through ordinary least squares regression of VI change (y/x) over the diagnostic recovery period. *Rmax* is the maximum VI that occurs during period t_r. It may be the final value, as depicted here, or it may be an intermediate value if vegetation index peaks during an intermediate year.

The algorithm detects the disturbance minimum and also the year in which the minimum VI value was reached. This is considered to be the time of mining and is the point in time when the 7-year recovery period is initiated.

Testing the Classification Model. We executed several procedures to (1) determine which of the many vegetation indices that have been developed by various remote sensing studies over the years is best suited to our purposes, and (2) determine if the classification model is best applied on an individual pixel basis, or if a more effective procedure is to create and apply the classification algorithm to objects (i.e., groups of pixels that exhibit spectral similarities). These studies, as detailed in Sen et al. (2011), found that (1) the tasseled cap greenness-brightness index (TC G/B) (Crist and Cicone, 1984; Powell et al., 2010) is an effective vegetation index for our purposes, and (2) creation and application of the classification model using various VIs with objects produced more consistent and accurate classifications than did applying the model to individual pixels. Therefore, we used the multiresolution segmentation algorithm within Definien's Professional software (v 5.0, Definiens AG, München, Germany) to segment the Landsat images comprising our chronosequence into objects; we calculated an average TC G/B VI for each object within each image; and we tested the classification model's capability to discriminate mines from urban and forested areas using the TC G/B values computed over the multitemporal image sequence for each object.

To test the classification model, we developed training and validation data sets. Using ancillary data (Table 2), a stratified random sampling procedure was used to select 1262 points that represent landcover classes within the cloud-free-image portions of our study area. Mined areas were identified on DOQQ images and digitized to produce polygons; random points were generated within these polygons. For urban and forest classes, random points were developed on the landcover base layer from NLCD 1992 and 2001. Additional points representing roadways were generated from the GIS road layer acquired from Virginia Department of Transportation (VDOT); these were merged with the urban NLCD points to produce 'urban' training and validation points. The training dataset was comprised of 650 points, while 612 points comprised the validation dataset.

We analyzed the training dataset to develop diagnostic threshold for *Dmin* using the CART (Classification and Regression Trees; Salford Systems Inc., v.6.0) software and identified as "disturbed" points with VIs that dropped below the threshold. A CART classification tree (parameters: Gini splitting rule, en-fold cross validation) was prepared to classify the "disturbed" class as either mined or urban, using all three disturbance/recovery parameters. Then, we applied the classification model to the cloud-free-image sections of the study area using those threshold values. Because the classification model required a 7-image recovery period, areas with post-2001 disturbances that fell below *Dmin* were classified as "post-2001 disturbances." Using these procedures, areas that were cloud free over the entire chronosequence were classified as mined, urban, forest, or post-2001 disturbance. Classifications for objects containing each validation point were then compared to those points' actual landcover class and the classification model's ability to correctly identify mined areas was assessed.

Objective 2: Estimate woody canopy cover on reclaimed mine areas.

Objective 2 was pursued in cooperation with the USDA Forest Service (USFS). Our method is based on a protocol being developed by USFS for estimating woody canopy cover at the national level. It is the USFS intent that this protocol would be applied in production of the next National Land Cover Database (NLCD, 2011). Our study was a preliminary application of this protocol, testing its applicability over areas where mined lands constitute a significant fraction of the land base. The protocol seeks to model woody canopy cover as a function of both spectral and non-spectral landscape variables.

Model Development: Response Variable: The response variable was the percent woody canopy, which was developed by photo-interpretation of samples obtained using the same spatially randomized procedure that is being applied in other study areas of the NLCD 2011 effort. The samples were obtained from a 4x intensified FIA sampling grid (Bechtold and Patterson, 2005), which is developed from the EMAP (Environmental Monitoring and Assessment Program) sampling framework (White et al., 1992). Using this procedure, we located 547 photo-plots within the study area. Each of these photo-plots had 105 points, spread over a 3x3 Landsat pixel window covering approximately 90m x 90m on the ground (Fig. 3). These points were photo-interpreted using the leaf-on National Agricultural Image Program (NAIP; USDA 2010) image from 2008. An ArcMap™ (V 9.3.1, ESRI, Redlands, California, USA) extension "Canopy cover" was used to label each of the 57,435 points (i.e., 547 photo plots x 105 points per plot) as either "Canopy" or "Not Canopy". Each point was also labeled as "mine" or "not-mine". Based on these individual point classifications, each photo-plot was labeled as "mine" (if 100% mined), "non-mine" (if 0% mined) or "split-mine" (if partially mined). In addition to the NAIP image, Digital Orthophoto Quarter Quads (DOQQ) from mid to late 1990s were used to cross-check the "mine" vs. "not- mine" designations, checking for older mine areas which may have become obscured on the recent photos by post-mining vegetation development. For each photo-plot a percent canopy was computed using canopy labels from the 105 points. These percent canopy data were used as the response variable in the modeling procedure.

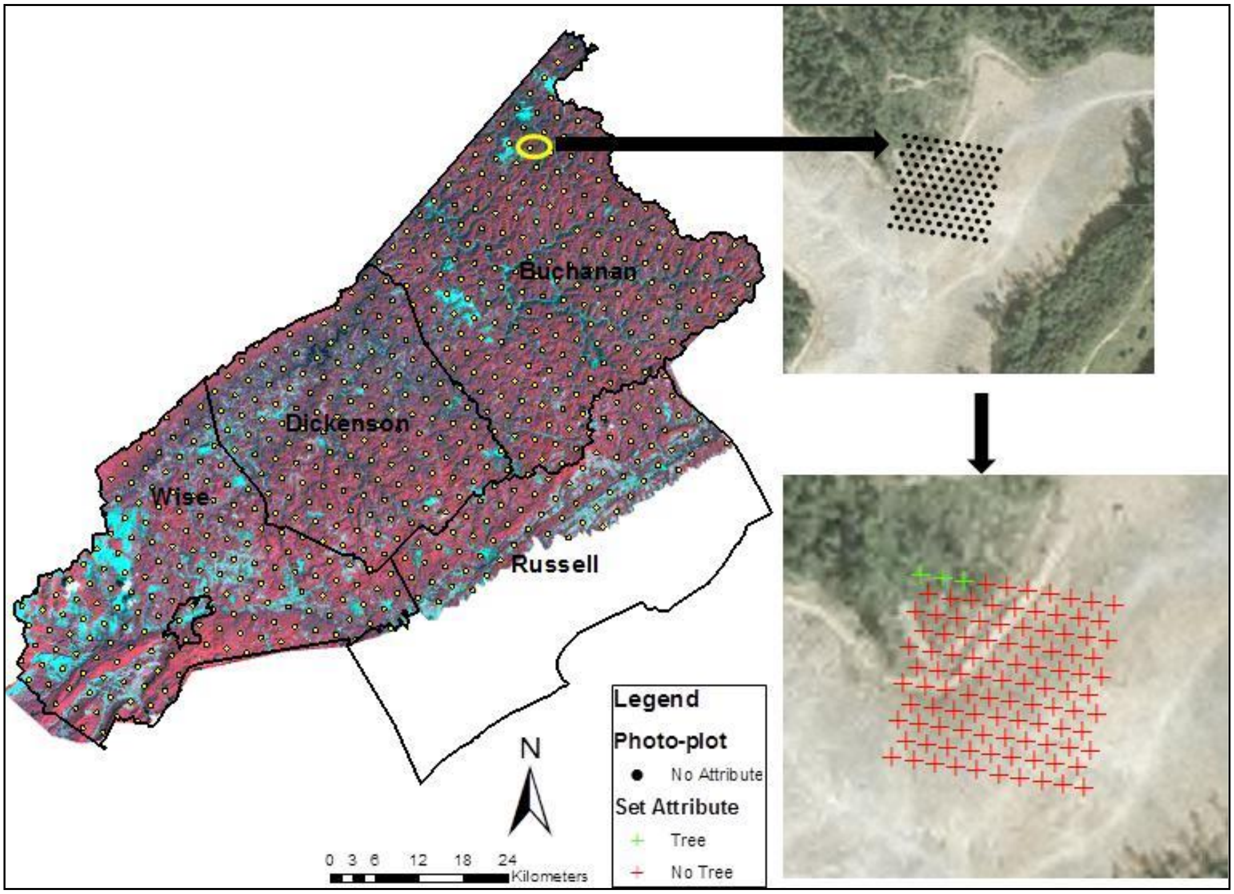

Figure 3. Left: distribution of the photo-plots over the study area. Inset upper right: A partially mined photo-plot with 105 photo-points, before labeling. Inset lower right: Points within the photo-plot as labeled by "canopy" vs. "no canopy." A point is labeled "tree canopy" (meaning woody canopy in this usage) only if the precise point location falls directly on a piece of ground that is visibly covered by tree or shrub canopy in the aerial photo.

<u>Explanatory Variables.</u> Fifty-five landscape variables were developed as potential explanatory variables for use in the canopy-cover estimation models (Table 3). Most variables were specified for the entire study area in a separate geospatial data layer. The photo-plot sampling points were then overlaid on each data layer, and the corresponding variable was specified for each photo plot using the "extract values to points" in the spatial analyst ArcMap™ extension. Values were specified as both focal mean (the mean value for all raster grid-points within the photo plot, *fcm*) and focal standard deviation (the standard deviation for all raster grid-points within the photo plot, *fst*). Data layers used to develop the explanatory variables included two recent Landsat images selected for image quality over the study area from among those

available, one during the leaf-on season and the other as leaf-off. These images were downloaded in a level 1T terrain-corrected form, and were radiometrically and atmospherically corrected using the LEDAPS routine (Masek et al., 2000). The resulting image is converted to surface reflectance and has substantial haze reduction. Bands 1, 2, 3, 4, 5 and 7 were spectrally subsetted from this surface reflectance image. Band 6, temperature, is obtained separately as a LEDAPS output in a brightness-temperature corrected form. Vegetation indices – the Normalized Difference Vegetation Index (NDVI; Rouse et al., 1973) and the three tasseled cap (TC) indices (band 1 - "brightness", band 2 - "greenness", and band 3 - "wetness"); (Crist and Cicone, 1984), were calculated for both leaf-on and leaf-off images.

Table 3. Detailed list of variables used in the model development process for Objective 2.

Variable classes	Variables	Focal mean (*fcm*)	Focal standard deviation (*fst*)	Count
DEM and DEM derivatives	Elevation	√	√	4 x 2 =8
	Aspect	√	√	
	Cosine aspect	√	√	
	Sine aspect	√	√	
Landsat leaf-on (9/3/2008)	Spectral bands 1,2,3,4,5,7	√ (6)	√ (6)	11 x 2 = 22
	Temperature band	√	√	
	NDVI	√	√	
	TC band 1 (Brightness)	√	√	
	TC band 2 (greenness)	√	√	
	TC band 3 (wetness)	√	√	
Landsat leaf-off (1/25/2009)	Same as for leaf -on			11 x 2 = 22
Non-image variables	NLCD 2001 canopy cover			1
	NLCD 2001 landcover			1
	Number of mined points			1
TOTAL				55

Terrain data were obtained from Virginia Department of Mines, Minerals and Energy (VDMME) as vector (polyline) shapefiles representing elevation contours. VDMME personnel had derived the data from Virginia Base Mapping Program 2007 Orthophotography (VITA, 2010). These data are obtained and developed by Virginia DMME for the purpose of representing current contours of mined areas, which may differ from the pre-mining contours that are commonly represented by publicly available terrain data. The polyline shapefiles were converted into a 30-m raster data layer. Slope and aspect were calculated using the DEM raster, expressed as degrees. Aspect values were transformed to radians; the sine and cosine aspect layers were computed from the transformed aspect.

Focal mean and focal standard deviation were calculated for all geospatially derived data layers, using a 3x3 pixel window (~90m x ~90m). The 2001 NLCD canopy cover estimates and 2001 classified landcover data were obtained from MRLC. Anderson II classification scheme had been followed for the land cover classification. The remaining variable, number of mined points within the photo plots, was recorded during the photo-interpretation process.

Model Development. The 55 potential explanatory variables were reduced in number by assessing their inter-correlations. A non-parametric correlation coefficient (Spearman's rho) was computed for each variable pair. For pairs with rho > 0.80, each variable of the pair was assessed for correlation to the response variable (percent canopy cover) and the less-correlated variable was removed from the variable list. This process resulted in a set of 29 landscape variables, a subset of the original 55 variables (Table 3), for use in developing a woody canopy cover estimation model.

A best subsets regression procedure was carried out in Minitab, as an effort to estimate canopy cover (the response variable) as a function of the landscape variables. Using model selection criteria of high R^2, high adjusted R^2, low Mallow's Cp statistic, and low standard deviation of prediction error, a preliminary canopy-cover estimation model with 14 explanatory variables was chosen as a basis for further work.

The 14-variable model was further refined. Using this model, an analysis was conducted to identify extreme outliers in the sample. The standard deviation (sd) of the residuals obtained from the best subset regression was 0.163. Using 2sd = 0.326 as the cut-off threshold, eight of 547 photoplot samples were identified as extreme outliers. These outlying data points were excluded from further modeling procedures, reducing the sample size to 539. Also, a decision was made to eliminate the NLCD 2001 canopy cover variable from further consideration because of its dependence on a result generated during an earlier time period. This decision was made considering the dynamic nature of reclaimed mine landscapes, and our desire to ensure that rapid change in canopy cover, when it does occur, can be assessed by our procedures

The remaining 13 explanatory variables and 539 sample points were used to produce a more parsimonious model by applying standard least-squares regression procedures in JMP (v. 9.0.0, SAS Institute Inc., NC, USA), using all 13 explanatory variables. From the regression results, the six variables with most significant p-values were selected to construct the parsimonious model.

Using those six variables only as explanatory variables and the 539 sample points, another multiple regression procedure was employed to evaluate the six variables' significance and their coefficients. That model's potential applicability for use in estimating canopy cover for mined areas only was assessed by fitting the predicted percent-canopy values against the measured estimates for the mined points only. Based on this outcome, a final model with six explanatory variables was selected.

Results and Discussion

Objective 1

Results: The three diagnostic parameters behaved as expected, with *Dmin* being generally lower, and *Rslope* and *Rmax* being generally higher for the mines than the urban category (Fig. 4). Although the mines' mean and median Dmin values were lower than the urban disturbance values, the lower quantiles and minima were at comparable levels, demonstrating that some urban disturbances (large-scale land development and highway construction, for example) create significant surface disturbances similar to those created by mining, with all surface soils and vegetation removed, and therefore create a similar spectral signature during the maximum disturbance phase; but mines produce this high level of surface disturbance more frequently (Sen et al., 2011).

Rmax and *Rslope* were generally higher for mines than for urban disturbances, a finding that we attribute to the extensive and rapid revegetation that is commonly required under SMCRA, but *Rslope* was found to give better discrimination than *Rmax*. We attribute this latter finding to the high level of image-to-image variance in the VIs spectral signatures; sources for such variance include seasonal differences among the image acquisition dates, year-to-year differences in moisture availability during the growing season, and image quality differences due to factors such as atmospheric conditions, solar and camera angles, and the like. Because *Rslope* is computed across 7 consecutive images, it is less influenced by image-to-image variability than *Rmax*. This interpretation is analogous to findings by Kennedy et al. (2007), whose studies revealed that analysis of data patterns derived from multitemporal image sequences can increase the accuracy of Landsat interpretations, compared to dual image comparisons, because multitemporal analyses are influenced less by image-to-image variance. Of the three diagnostic parameters, no single parameter proved adequate when used alone to provide a high level of correct classifications, and all three parameters contributed to the final classification model.

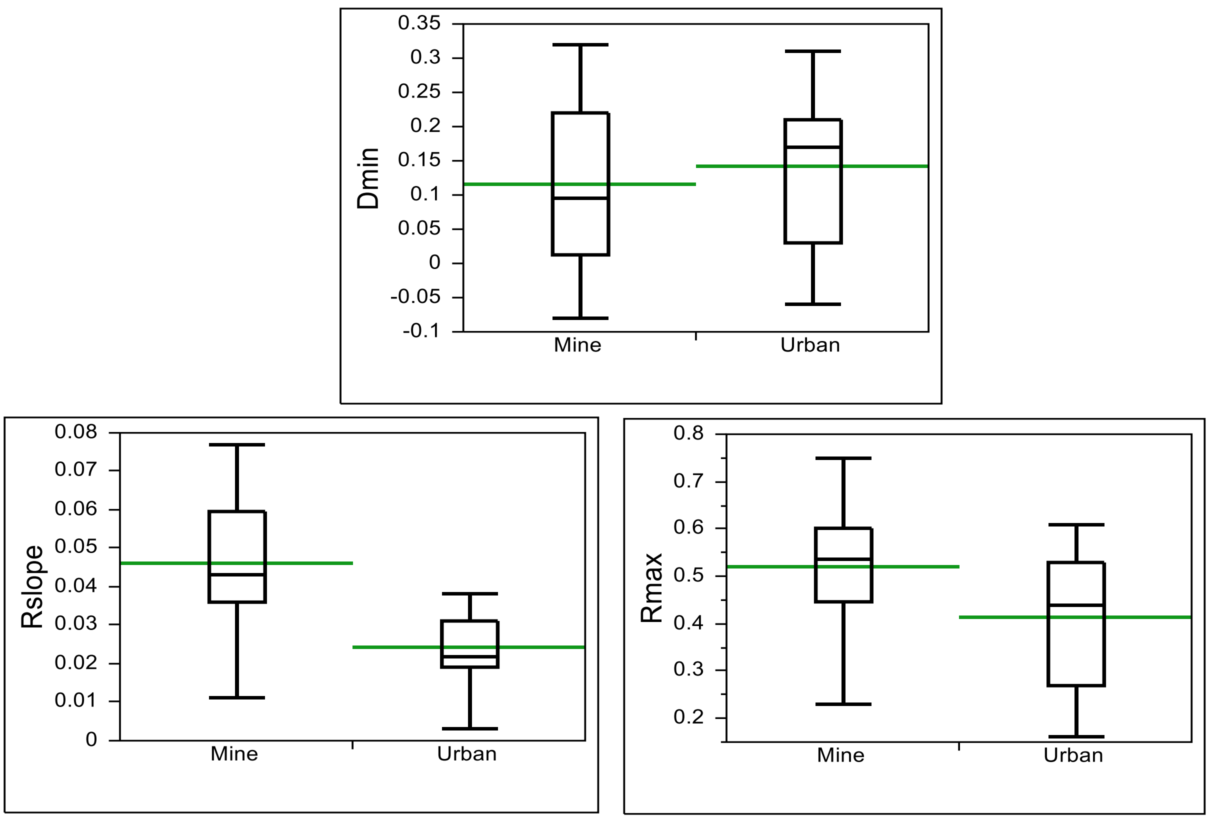

Figure 4. Training data distributions of the 3 diagnostic parameters for mined and urban classes. Box plots represent the 25th, 50th, and 75th percentiles of each distribution; the whiskers represent the 5^{th} and 95^{th} percentiles; and the horizontal lines represent means.

The classification model was able to classify 89.1% of the validation points correctly, including 87.4% of the mined validation points (Table 4). Using those classifications, a map of mined and urban areas was produced for a portion of the study area (Fig. 5, upper right). Years of mining were verified by a visual check using historical aerial photos and Landsat images and were found to be correctly identified in all the cases. The area was also classified by the year of mining and a map of year of mining is represented for the same area in Fig. 5 (lower left).

Table 4. Classification accuracies for 612 validation points.

		Reference (validation) data			
		Mine	Non-Mine	Reference Total	User's accuracy
Classified data	Mine	160	23	183	87.4%
	Non-Mine	44	385	429	89.7%
	Classified Total	204	408	612	89.1%
Producer's accuracy		78.4%	94.4%	89.1%	

Discussion: Analysis of the incorrect classifications revealed some confusion between the urban and mined categories which upon investigation using the DOQQ and NAIP leaf-on aerial photos revealed that most mined areas incorrectly classified as urban were areas that were not well revegetated. These included refuse piles that remain active, mined areas that remain in use for structures and/or for vehicular maintenance and for parking, and reclaimed areas that were not as well vegetated as other reclaimed areas (Fig. 6). Some of these latter areas appeared to have near-level configurations, suggesting soil compaction created by vehicular activity as a possible cause for lack of vigorous revegetation. This post-classification study of mis-classified mine areas also revealed that most (91%) were within or bordering other mined areas that had been classified correctly, suggesting that manual review of such mined-urban adjacencies could be employed as a means of increasing the accuracy of classification. Similarly, most of the urban areas incorrectly classified as mines were "nested" within other correctly classified urban areas, suggesting that a manual review and error-checking of results should focus on such adjacencies, should this classification algorithm be placed into operational use.

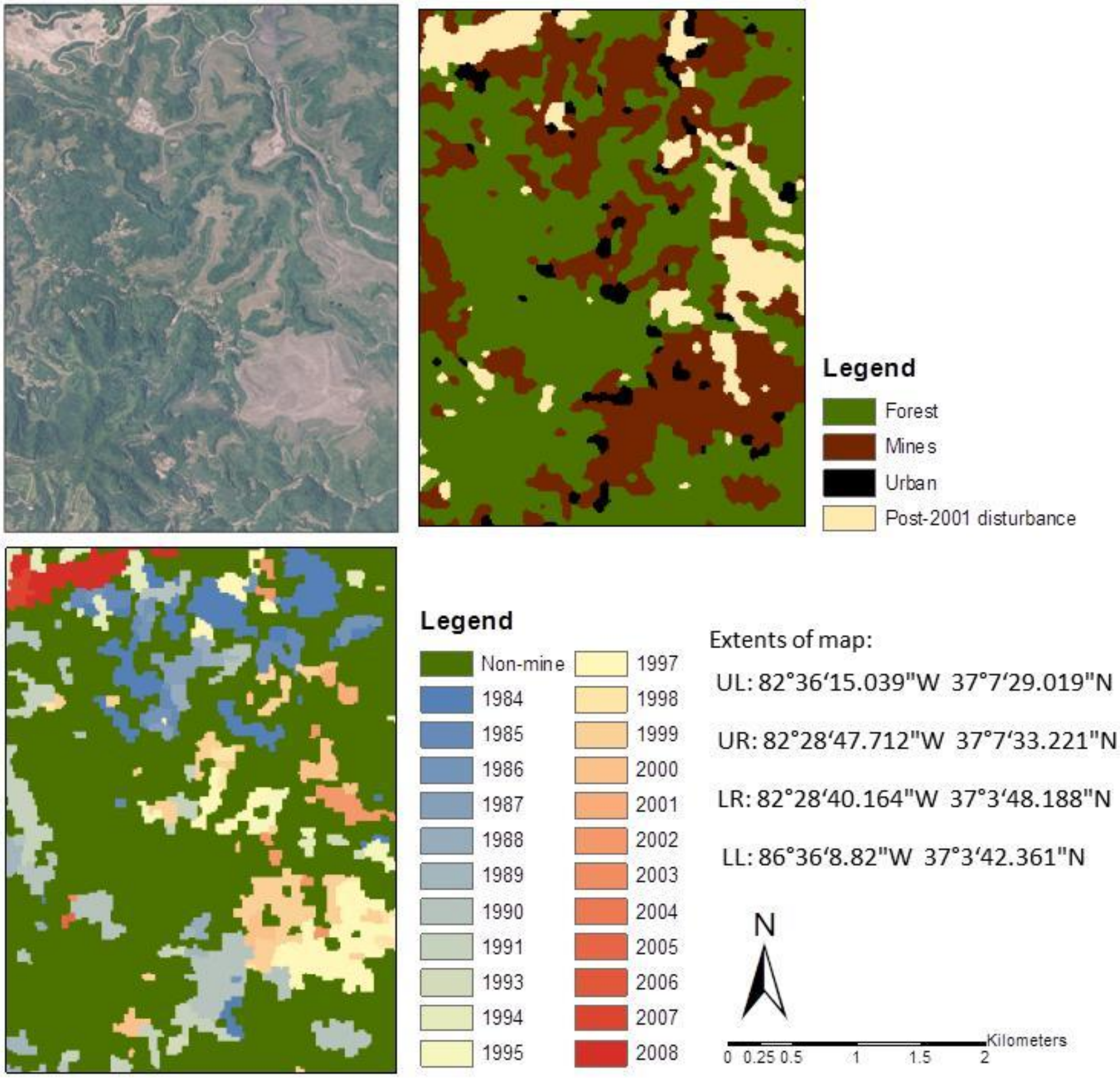

Figure 5. (Left) an aerial photo (NAIP-2008) of a portion of the study area, located in Wise County, Virginia, (Upper-right) the classification map of that area produced by the classification model. Note that several “pockets” of misclassified “urban” occurred within the extensive reclaimed mines for reasons described in the text. (Lower-left) classification map of mines categorized by date of mining, and of post-2001 disturbances by date of VI minimum.

Figure 6. Examples of mined areas that were not correctly classified due to lower Rslope and/or Rmax values than those which are characteristic of most reclaimed and revegetated coal surface mines. Left - a mined area used for mine-related activity (building, parking, etc.); Right - an area that is not as well vegetated (mine refuse). The black lines form the image-objects and the points are those used for validation purposes.

Objective 2

Results: The 14-variable model selected from the best-subsets regression results (Table 5) had an R^2 of 0.63, adjusted R^2 of 0.63, and Mallows Cp statistic of 3.8. This model was selected from the 29 potential models generated by best-subsets regression because it had among the highest R^2 and adjusted R^2 values, but with Mallows Cp and number of variables that were among the lowest, while having fewer significant explanatory variables than other potential models yielded by this procedure.

A standard least-squares-regression canopy-cover prediction equation was generated using 6 explanatory variables (Table 5). The resulting model was fitted over the mined points, and the predicted vs. actual fit generated an $R^2 = 0.74$ (Table 6). However, the resulting plot showed that the model was predicting a wide range of canopy covers, including negative values, for mined photo points where actual canopy cover had been recorded as zero. Thus, all negative canopy-cover prediction values were manually adjusted to zero. This manual adjustment improved the predicted vs. actual fit to $R^2 = 0.80$ (Table 6). The six-variable model (Table 5), applied with manual adjustment of negative canopy-cover estimates to zero, is considered to be the final model (Fig. 7).

Table 5. Lists of variables selected by best subsets regression for the 14-variables, and coefficients and coefficient significance for the final woody canopy cover estimation model.

Variables in 14-variable Model	Coefficient in Final Model†	p-value in Final Model
Cosine of aspect, *fcm**	- 0.03	0.0001
Sine of aspect, *fcm**	- 0.01	0.0001
2009 Landsat band 1, *fcm*		
2009 NDVI, *fcm**	+ 0.5	0.0001
2008 TC band 2, *fcm*		
2008 TC band 3, *fcm**	+ 0.0005	0.0001
2008 Landsat band 2, *fst*		
2008 Landsat band 3, *fcm*		
2008 Landsat temp, *fcm*		
2008 NDVI, *fcm**	+ 0.52	0.0008
2008 TC Band 3, *fst**	- 0.0006	0.0021
NLCD 2001 canopy cover		
Number of mined points		
2009 TC band 2, *fst*		

† The final canopy-cover estimation model's intercept is +0.15. It estimates woody canopy cover as a percent of total area. The final model's application requires zero-adjustment of negative canopy-cover estimates, as explained in text. * indicates variables chosen for a final parsimonious model.

Table 6. Results of final model (Table 5) development.

	R^2	Adj R^2	*RMSE*
Regression on all samples (539)	0.57	0.56	0.17
Model prediction fit vs. actual on "Mine" samples (64)	0.74	0.73	0.18
Model prediction fit vs. actual on "Mine" samples (64) after zero adjustment.	0.80	0.79	0.16
Model prediction fit vs. actual on "Non-mines" samples (64) after zero adjustment.	0.44	0.43	0.18
Model prediction fit vs. actual on "Split-Mine" samples (64) after zero adjustment.	0.41	0.40	0.16

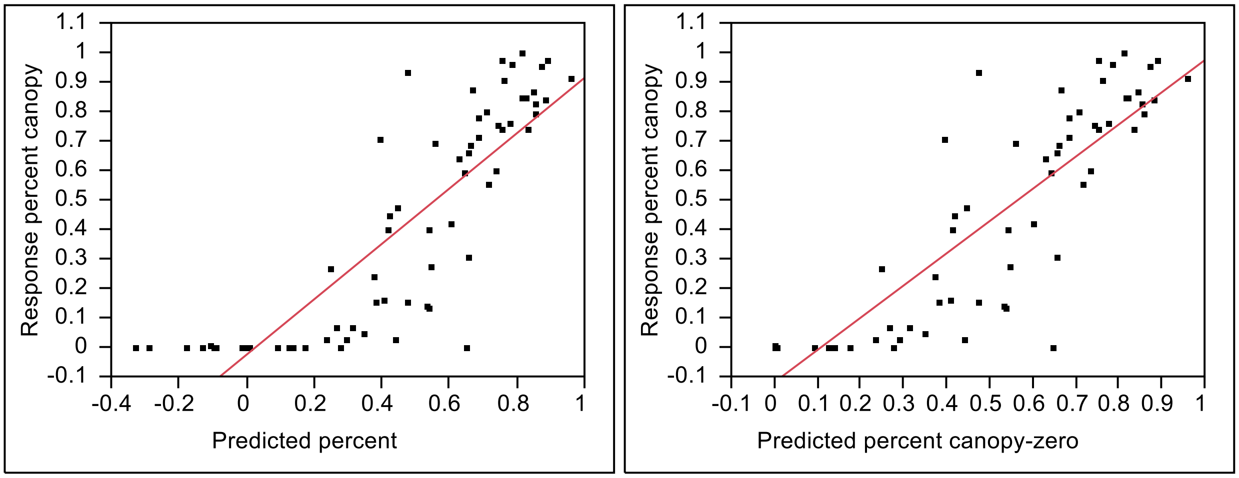

Figure 7. Model predicted canopy cover vs. actual before (left) and after (right) resetting all negative predictions to zero ("zero adjustment"), mined points only (64).

Discussion: The procedures described above were conducted with the intent of generating a canopy-cover estimation model that can be applied to both mined- and non-mined areas within our study region. The outcome, however, did not support the potential for combined application, as the final model proved far more accurate when applied to mined areas only than when applied to a full data set comprised predominantly of un-mined areas (Table 6). In a general sense, this result indicates that mined- and non-mined landscapes within the study region differ in their spectral and topographic characteristics, and/or the interaction of those characteristics. The greater capability of the final model to discriminate canopy from non-canopy on mined areas may indicate that spectral differences between these two cover types are more distinct on the mined areas.

The final model includes 6 explanatory terms that appear to us as potentially reflective of physical relationships with canopy cover, although documented and verified explanations for those physical relationships are not yet clear. Two of these explanatory variables – the cosine and sine of aspect – reflect terrain. The cosine transformation of the aspect is a measure of the northness and the sine transformation of the eastness. It is generally well known that woody vegetation characteristics (e.g., dominant species and/or vigor) are affected by aspect, and that north and east aspects are generally favorable.

Four of the explanatory variables are spectral and all are derived from vegetation indices rather than raw spectral bands. Both the NDVI and the TC band 3 (wetness) indices contributed one explanatory variable from both the leaf-on (2008) and the leaf-off (2009) image. It is logical that leaf-on vegetation indices would contribute to woody canopy cover estimation. It is possible that the vegetation indices are producing greater distinctions between woody canopy and herbaceous vegetation for mined areas because the herbaceous-vegetation spectral signal includes greater influence by underlying soils and/or dead and dying leaf area tissue. It is possible that the herbaceous vegetation and/or underlying soils of mined areas differ spectrally from those of unmined areas. NDVI contrasts the difference between visible and near infrared reflectance with their sum, and therefore highlight areas of high green biomass. It is possible that NDVI distinguishes well between herbaceous mined areas from the woody-canopy covered mines, since the background soil reflection is possibly greater in herbaceous vegetation. TC band 3 or the wetness band has been shown to be sensitive to soil and plant moisture (Crist and

Cicone, 1984), and to be more responsive to the interaction of water content and the structure of canopy (Cohen et al., 1995).

Two of the spectral explanatory variables are derived from leaf-off vegetation indices. It is possible that these variables contributed to canopy-cover estimation because some of the woody canopy is coniferous, and thus influences the vegetation indices via contrast with other landscape features during the winter months. It is possible that coniferous species constitute a larger fraction of total woody vegetation on mined than on unmined areas, and that this difference contributes to the improved performance by the 6-variable model on mined areas. These potential explanations are described as "possible" because, although suggested by the authors' considering our general knowledge of these systems, they have not checked or verified through analysis of field or photographic data.

These results should be considered as preliminary. Even so, we interpret these results to indicate that there is potential for use of Landsat imagery in characterizing mined lands' woody canopy cover.

Conclusions

These results should be seen as the outcomes of early steps in a multi-year process. We are confident in the procedure for identifying and aging mine sites; future work will focus on development of algorithms for bridging cloud-obscured areas in multiple-image sequences; and for interpreting the rate of spectral recovery after reclamation. The preliminary woody canopy characterization model will be further refined and model development efforts will be validated against aerial photographs and/or independent datasets. We are developing these methods, anticipating eventual application over pilot-study areas and more broadly.

The post-SMCRA mined land resource constitutes on the order of 600,000 hectares in eastern USA's Appalachian region. Lack of knowledge about the extent and nature of the land resource base created by post-SMCRA coal mining is an obstacle to development of policies and strategies capable of improving its utilization. The Landsat satellite multispectral data series is well suited to use in characterizing these land resources, due to its spatial resolution, extensive archive, low cost, and temporal extent. Our studies have found that Landsat data can be used to identify post-SMCRA surface mines, by areal extent and by year of mining, using techniques with potential for application over broad areas. These results also show promise for Landsat interpretation techniques to characterize these lands' woody canopy cover, an important indicator of ecological status and future use potentials on reclaimed coal surface mines.

Acknowledgements

This research has been supported by the Geospatial and Environmental Analysis Program, and Department of Crop and Soil Environmental Sciences, at Virginia Tech. Sincere thanks to Steve Mullins and Daniel Kestner, Virginia Department of Mines, Minerals and Energy, for providing the digital elevation spatial data.

References

Angel, P., V. Davis, J. Burger, D. Graves and C.E. Zipper. 2005. The Appalachian regional reforestation initiative. U.S. Office of Surface Mining, Forest Reclamation Advisory No. 1. 2 p. http://arri.osmre.gov/

Angel, P., J.A. Burger, V.M. Davis, C.D. Barton, M. Bower, S.D. Eggerud, P. Rothman. 2009. The forestry reclamation approach and the measure of its success in Appalachia, Proceedings America Society of Mining and Reclamation, 20091 pp 18-36. http://dx.doi.org/10.21000/JASMR09010018.

Bechtold, W.A., Patterson P.L. (Editors). 2005. The enhanced Forest Inventory and Analysis program-national sampling design and estimation procedures. General Technical Report SRS-80. Asheville, NC: U.S. Department of Agriculture Forest Service, Southern Research Station.

Burger, J.A., C. Zipper, P. Angel, S. Eggerud. 2011. Reforestation guidelines for unused surface mined lands: Development, application and adoption. Proceedings America Society of Mining and Reclamation, 2011 pp 113-141 http://dx.doi.org/10.21000/JASMR11010113

Cohen, W.B., T.A. Spies and M. Fiorella. 1995. Estimating the age and structure of forests in a multi-ownership landscape of western Oregon, U.S.A., International Journal of Remote Sensing 16 (1995): 721-746. http://dx.doi.org/10.1080/01431169508954436.

Crist, E. P. and R. C. Cicone. 1984. A physically-based transformation of Thematic Mapper data - the TM tasseled cap, IEEE Transactions on Geoscience and Remote Sensing, 22 (3): 256-263. http://dx.doi.org/10.1109/TGRS.1984.350619.

D'Appolonia, Inc. 1980. Abandoned Coal Mine Land Inventory. Project # 78-411. Prepared for Virginia Division of Mine Land Reclamation.

Evans, D.M., C.E Zipper, J.A. Burger, C. Fields-Johnson. 2010. Tree species and density effects on woody biomass production on unused mined lands: Establishment and two-year results. Proceedings America Society of Mining and Reclamation, 2010 pp 276 -291. http://dx.doi.org/10.21000/JASMR10010276291.

Fields-Johnson C., T. Fox, J. Burger, and C. Zipper. 2008. Fourth-year tree response to three levels of silvicultural input on mined land. Proceedings America Society of Mining and Reclamation, 2008 pp. 389-404. http://dx.doi.org/10.21000/JASMR08010389.

Jensen, J.R., Remote Sensing of the Environment - An Earth Resource Perspective, Prentice Hall, Upper Saddle River, New Jersey.

Kennedy, R. E., W. B. Cohen and T. A. Schroeder. 2007. Trajectory-based change detection for automated characterization of forest disturbance dynamics. Remote Sensing of Environment, 110 (3): 370-386. http://dx.doi.org/10.1016/j.rse.2007.03.010.

Masek, J. G., E. F. Vermote, N. E. Saleous, R. Wolfe, F. G. Hall, K. F. Huemmrich, F. Gao, J. Kutler and T. K. Lim. 2006. A Landsat surface reflectance dataset for North America, 1990-2000. IEEE Geoscience and Remote Sensing Letters, 3 (1): 68-72. http://dx.doi.org/10.1109/LGRS.2005.857030.

Multi-Resolution Land Characteristics Consortium (MRLC). 2011 .National Land Cover Database. http://www.mrlc.gov/

Powell, S. L., W. B. Cohen, S. P. Healey, R. E. Kennedy, G. G. Moisen, K. B. Pierce and J. L. Ohmann. 2010. Quantification of live aboveground forest biomass dynamics with Landsat time-series and field inventory data: A comparison of empirical modeling approaches. Remote Sensing of Environment 114 (5): 1053-1068. http://dx.doi.org/10.1016/j.rse.2009.12.018.

Rouse, J. W., J. Haas, J. A. Schell and D. W. Deering. 1973. Monitoring vegetation systems in the Great Plains with ERTS. Third ERTS symposium, NASA SP-3511: 309-317.

Sen, S., C. Zipper, R.H. Wynne, and P.F. Donovan. 2011. Identifying revegetated mines as disturbance/recovery trajectories using an interannual landsat chronosequence. Photogrammetric Engineering & Remote Sensing (in review).

Skousen J., J. Gorman, E. Pena-Yewtukhiw, J. King, J. Stewart, P. Emerson, and C. DeLong. 2009. Hardwood tree survival in heavy ground cover on reclaimed land in West Virginia: Mowing and ripping effects. Journal of Environmental Quality 38:1400-1409. http://dx.doi.org/10.2134/jeq2008.0297.

US Department of Agriculture (USDA). 2010. Imagery Programs – NAIP Imagery. Farm ServiceAgency. http://www.fsa.usda.gov/FSA/

USGAO. 2009. Surface coal mining characteristics of mining in mountainous areas of Kentucky and West Virginia. US Government Accounting Office. GAO-10-21.82.

Virginia Information Technologies Agency (VITA). 2010. VBMP Orthophotography. http://www.vita.virginia.gov/isp/default.aspx?id=8412.

White, D., J. Kimerling, and W.S. Overton. 1992. Cartographic and geometric components of a global sampling design for environmental monitoring. Cartography and Geographic Information Systems 1(19): 5-22. http://dx.doi.org/10.1559/152304092783786636