Regional Scale Land Cover Characterization Using MODIS-NDVI 250 m Multi-Temporal Imagery: A Phenology-Based Approach

Joseph F. Knight[1] and Ross S. Lunetta
U.S. Environmental Protection Agency, National Exposure Research Laboratory, Research Triangle Park, North Carolina 27711

Jayantha Ediriwickrema
Computer Sciences Corporation, 2803 Slater Road, Morrisville, North Carolina 27560

Siamak Khorram
Center for Earth Observation, North Carolina State University, Raleigh, North Carolina 27695-7106

Abstract: Currently available land cover data sets for large geographic regions are produced on an intermittent basis and are often dated. Ideally, annually updated data would be available to support environmental status and trends assessments and ecosystem process modeling. This research examined the potential for vegetation phenology–based land cover classification over the 52,000 km^2 Albemarle-Pamlico estuarine system (APES) that could be performed annually. Traditional hyperspectral image classification techniques were applied using MODIS-NDVI 250 m 16-day composite data over calendar year 2001 to support the multi-temporal image analysis approach. A reference database was developed using archival aerial photography that provided detailed mixed pixel cover-type data for 31,322 sampling sites corresponding to MODIS 250 m pixels. Accuracy estimates for the classification indicated that the overall accuracy of the classification ranged from 73% for very heterogeneous pixels to 89% when only homogeneous pixels were examined. These accuracies are comparable to similar classifications using much higher spatial resolution data, which indicates that there is significant value added to relatively coarse resolution data though the addition of multi-temporal observations.

INTRODUCTION

Land cover (LC) classification methods for remotely sensed image data have, to date, made only limited use of multi-temporal remote sensor data. Typically, images

[1]Corresponding author; email: knight.joe@epa.gov. The U.S. Environmental Protection Agency funded and conducted the research described in this paper. It has been subject to the Agency's programmatic review and has been approved for publication. Mention of any trade names or commercial products does not constitute endorsement or recommendation for use.

GIScience & Remote Sensing, 2006, **43**, No. 1, p. 1-23.

are collected for the spring, summer, and fall seasons to support regional-scale LC classification efforts (Homer et al., 2004). Multi-temporal images are also used to support change detection. In change detection, images of a given area, separated by some time interval, are compared to identify pixels that changed during the interval (Lillesand and Kiefer, 2003). The images' acquisition dates are nominally chosen so that they fall as closely as possible to anniversary dates. This approach minimizes spectral differences between the images caused by factors such as seasonal vegetation phenology, sun angle and shading, cloud cover, and atmospheric particulate composition (Jensen, 2004). Using this traditional spectral analytical approach, vegetation phenology is essentially noise that can substantially contribute to poor classification accuracy.

Studies reporting the use of multi-temporal image data for classification often include relatively few dates, possibly due to a lack of cloud-free image availability, cost, and processing requirements. A basic multi-temporal approach is the use of leaf-on and leaf-off images, which provides greater vegetation phenology information than is available with only one image (Goetz et al., 1999; Varlyguin et al., 2001). Seasonal images have also been used in LC classification with some success (Loveland et al., 1995; Lambin, 1996; Roberts et al., 1997; Goetz et al., 2004). Vegetation phenology represents a potentially significant source of LC information (Reed et al., 1994; Loveland et al., 2000; Senay and Elliott, 2000). Because most of the landmass of the world is covered by vegetation, taking into account phenology when performing LC classification may yield more accurate maps. However, only a limited number of studies to date have explored the potential of employing a complete year of uninterrupted vegetation phenology data as the basis for an LC classification.

The objective of this study was to develop a phenology-based LC classification of the Albemarle-Pamlico Estuarine System (APES) using a time series of 23 MODIS 250 m Normalized Difference Vegetation Index (NDVI) composite images for calendar year 2001. The classification used a hybrid Anderson Level 1-2 schema (Anderson et al., 1976). To obtain optimal classification accuracy, the NDVI data were filtered to correct pixels containing low-quality or erroneous values, pixels containing water were removed from the image stack. The agriculture areas (pixels) were partitioned from non-agriculture areas. Non-agriculture image pixels were classified using hyperspectral image processing techniques, where the spectral data were substituted with NDVI data that tracked the annual vegetation phenology cycles.

STUDY AREA

The study area for this project was the 52,000 km^2 Albemarle-Pamlico estuary system (APES) drainage basin in North Carolina and Virginia (Fig. 1). The APES is the second-largest estuarine system in the United States (after the Chesapeake Bay) and represents a major resource base, through commercial and private fishing, tourism, recreation, agriculture, forestry, and mining (APNEP, 2005). The study area is composed of four physiographic provinces (Valley and Ridge, Blue Ridge, Piedmont, and Coastal Plain), and includes four major river basins (Chowan, Roanoke, Tar, and Neuse). The Neuse River Basin (NRB) is designated as a Near-Laboratory Long-Term Research Area (NLLTRA) by the U.S. Environmental Protection Agency (USEPA), and has been used extensively since 1998 for remote sensing methods

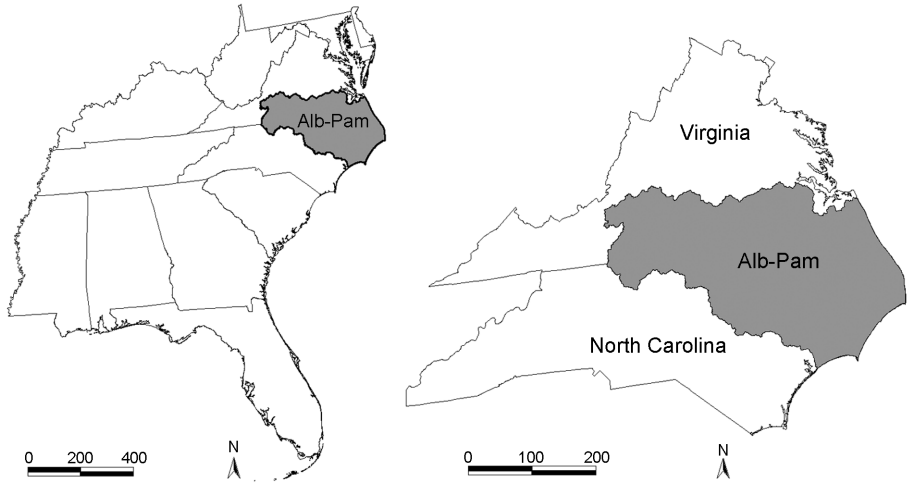

Fig. 1. The location of the Albemarle-Pamlico drainage basin.

development research (Lunetta et al., 2001, 2002b, 2003, 2004). Finally, the APES is one of 20 National Water Quality Assessment (NAWQA) Program Study Units designated by the United States Geological Survey (USGS). The NAWQA established the site in 1991 to assess and monitor water quality status and trends as part of a national monitoring program (Spruill et al., 1998). The APES was selected for this research because it contains a wide variety of cover types and land use activities, which are representative of the mid-Atlantic region and southeastern United States.

METHODS

The general approach in this project consisted of the following steps: reference data set creation, image preprocessing, water mask creation, agriculture mask creation, and land cover classification. The project flow for all preprocessing and classification steps is shown in Figure 2.

Reference Data

An important step in any LC analysis is the creation of a reference data set with which to assess the accuracy of the final classification product (Congalton and Green, 1999). A complicating factor in assessing the accuracy of a classification derived from 250 m image data in the APES was the disparity between the image pixel size and the average patch size of the landscape. A previous classification of the NRB by Lunetta et al. (2003) using Landsat ETM+ and SPOT multi-spectral (XS) data had an effective minimum mapping unit (MMU) of 0.4 ha. Overlaying a MODIS 250 m pixel grid over this classification revealed that approximately 5% of MODIS pixels contained homogeneous cover types at that MMU. Because the APES is quite heterogeneous at the 250 m scale, with approximately 95% all of the pixels in the study area composed of a mix of cover class types, a reference data set was required to provide

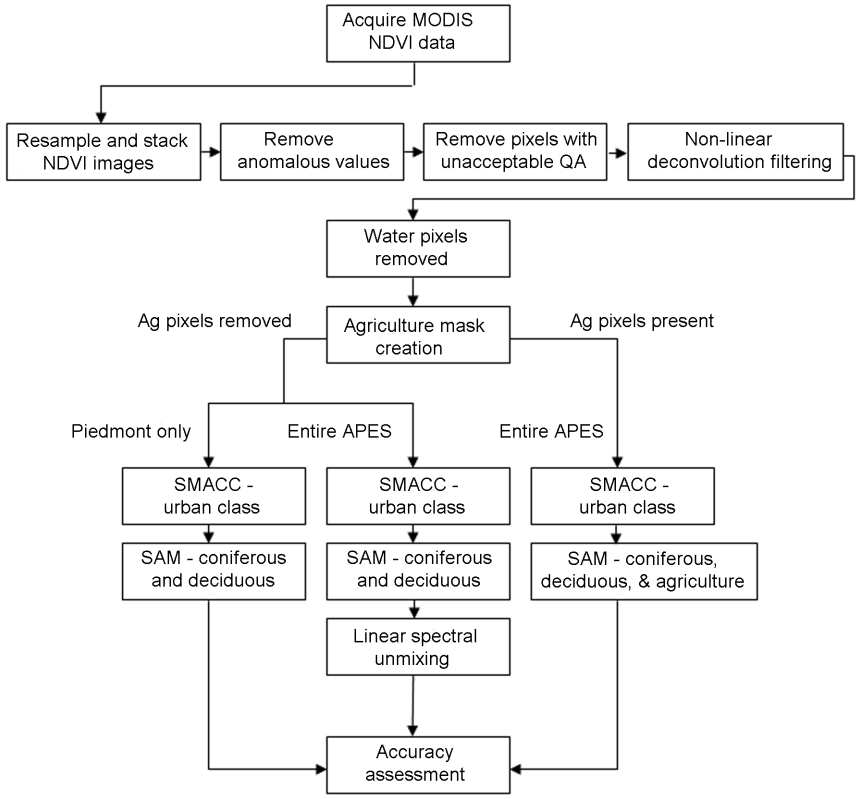

Fig. 2. Project methods flowchart.

sub-pixel LC information. USGS Digital Orthophoto Quadrangles (DOQs) were selected as the base data source for this sub-pixel reference information.

The APES contains elevations from mountains to coastal bottomland forests and beaches. To capture the variability of the study area, the study area was stratified by the 13 major Bailey eco-regions present (Table 1), and one USGS DOQs was randomly selected for each eco-region (Fig. 3, Table 2). The 13 DOQs were obtained from the North Carolina Center for Geographic Information and Analysis (NCCGIA) and the Virginia Economic Development Program (VEDP). Two quarter quads (one half) of the Spencer, NC DOQs were panchromatic, and the remaining 50 of the 52 total quarter quads were color infrared (IR). The acquisition dates of the DOQs ranged from 1998 to 1999, and represented the best contemporaneous data available for use with the 2001 MODIS data.

Each selected DOQ was overlaid with a 250 m fishnet corresponding to the pixels of the MODIS NDVI data (Fig. 4A). Each cell was assigned a unique numeric identifier (n = 31,322) and was then overlaid with a 100-point dot grid (Fig. 4B). The cells were examined by two trained aerial photography interpreters, who counted the numbers of dots that fell over the cover classes of interest (Table 3, Appendix). This

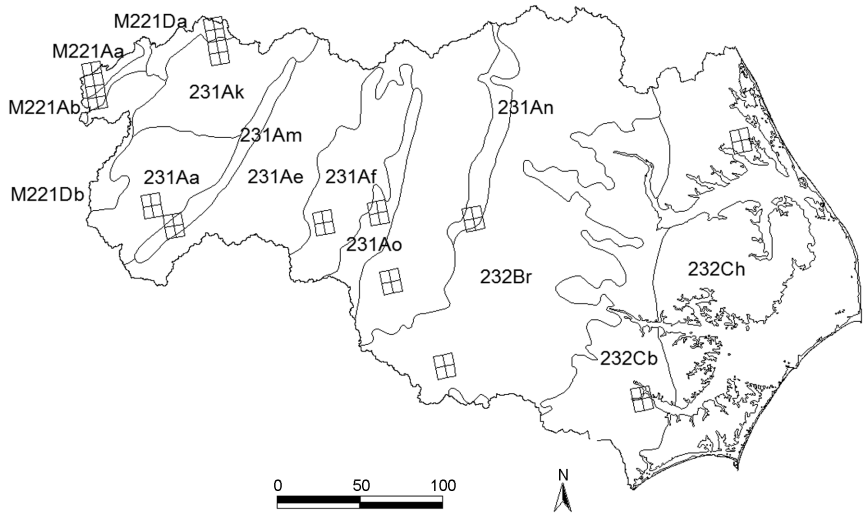

Fig. 3. Bailey eco-regions in the APES and reference data quadrangle locations.

Table 1. Bailey Eco-regions Present in the APES and Aggregate Eco-regions Used in Accuracy Assessment

Map code	Subsection name	Aggregate eco-region for accuracy assessment
231Aa	Midland Plateau Central Uplands	Mountains
231Ak	Lynchburg Belt	Mountains
M221Aa	Ridge and Valley	Mountains
M221Ab	Great Valley of Virginia	Mountains
M221Da	Northern Blue Ridge Mountains	Mountains
M221Db	Mestasedimentary Mountains	Mountains
231Ae	Charlotte Belt	Piedmont
231Af	Carolina Slate	Piedmont
231Am	Triassic Uplands	Piedmont
231Ao	Southern Triassic Uplands	Piedmont
231An	Western Coastal Plain-Piedmont Transition	Coastal plain
232Br	Atlantic Southern Loam Hills	Coastal plain
232Cb	Lower Terraces	Tidewater
232Ch	Tidal Area	Tidewater

procedure provided percentages of each cover class within the corresponding MODIS pixels. The interpreters were directed to provide Level 2 class percentages when possible. In situations where the correct Level 2 class label was difficult to ascertain, the interpreters provided only a Level 1 label. For example, in mixed forests where neither deciduous nor coniferous trees predominated, the interpreters chose the Level 1 label of "woody."

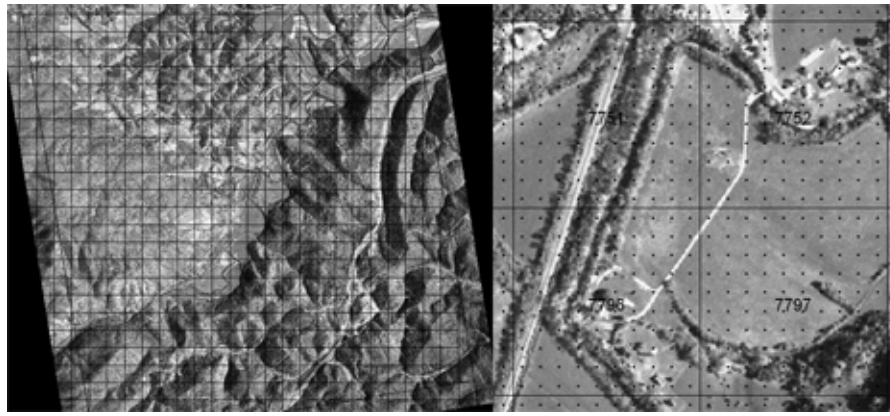

Fig. 4. Sample DOQ with (A) 250 m fishnet and (B) cells with 100-point dot grid overlays.

Table 2. DOQs Selected for Accuracy Assessment

DOQ name	State
Essex	NC
Four Oaks	NC
Mayodan	NC
New Bern	NC
Oxford	NC
Shiloh	NC
Timberlake	NC
Wake Forest	NC
Goode	VA
Ironton	VA
McDonald's Mill	VA
Sedalia	VA
Spencer	VA

NDVI Data Preprocessing

The multi-temporal images used in this research were acquired from the NASA Terra (AM-1) satellite's Moderate Resolution Imaging Spectroradiometer (MODIS) sensor. The MODIS 250 m NDVI product (MOD13Q1) provided the needed vegetation phenology data. Although MODIS NDVI scenes (16-day composites) were acquired for calendar years 2000 through 2004, only the 2001 data (n = 23) were directly used for the LC analysis. The images were resampled using a nearest neighbor operator from their native sinusoidal projection to the Albers equal-area conic projection. The individual scenes were subset to the APES study boundary and stacked into one image. Separate image stacks were developed for the NDVI data and

Table 3. Reference Data and Classified Map Classes

Reference	Classified
Herbaceous	
Agriculture	Agriculture
Maintained	
Pasture	
Herbaceous Aquatic	
Impervious	
Urban	Urban
Bare	
Woody	
Deciduous	Deciduous
Coniferous	Coniferous
Shrub	
Woody Aquatic	
Water	Water
Other	

the Quality Assessment Science Data Set (QASDS), which provided pixel-wise data quality ratings. The NDVI data stack was then subjected to three filtering and cleaning steps. First, anomalous high and low values were identified using a fixed threshold value. Anomalous pixel values were deleted and flagged for further processing. Second, NDVI pixels that fell below the acceptable QASDS quality level were deleted and flagged. Third, the flagged pixel values were estimated using nonlinear deconvolution. The final preprocessing step was to remove (mask) those pixels containing water or agriculture from the data stack.

First, anomalous data values were filtered to remove gross errors in the NDVI data. Anomalous values were eliminated by deleting and flagging pixels when their NDVI data values suddenly dropped or increased and then returned to near the previous NDVI value. The threshold for removal of data spikes was determined through empirical comparisons of original and filtered spectral profiles at various threshold values. The NDVI data stack was filtered using threshold values ranging from 0.10 to 0.25 NDVI. A high threshold (0.25) retained many of the anomalous features, while a low threshold (0.10) eliminated what was judged to be valid data. The optimal threshold value was determined to be 0.15. This threshold value eliminated most of the obvious anomalous values while preserving seasonal and local variations. Clearly, choosing a threshold for this type of filtering was subjective; however, this step was necessary, as the QASDS data did not flag as low quality many of the pixels that exhibited clearly anomalous values. The threshold was carefully chosen to be as conservative as possible, while still eliminating the gross errors.

Next, QASDS filtering was performed to delete and flag low-quality NDVI pixels. The QASDS provides a large amount of information related to potential quality problems of individual NDVI pixels, including atmospheric aerosol quantity,

presence of cloud cover, presence of snow and ice cover, likelihood of shadow, and whether BRDF correction was applied. This information is summarized in the Usefulness Index (UI), which ranges from "quality too low to be useful" to "perfect quality." After examining the spectral profiles of pixels with UI values of fair, acceptable, good, high, and perfect, the "acceptable" quality level was chosen as the minimum. Pixels with values below "acceptable" were deleted and flagged for further processing.

Last, a non-linear deconvolution was incorporated that transformed the filtered data into the frequency domain by Fourier transformation. The data stack was processed using a nonlinear deconvolution approach described by Roberts et al. (1987). This procedure separated the noise spectrum from the signal spectrum of the data. The noise spectrum was discarded, and an inverse Fourier transformation of the signal spectrum back to the time domain provided a cleaned NDVI spectrum for all pixels in the stack. The previously deleted NDVI data points were assigned values from this cleaned spectrum. Thus, a complete filtered and cleaned NDVI spectrum was created for all pixels in the 2000–2005 data stack (Fig. 5).

Water Mask

Differences in water levels over time create significant temporal variation in NDVI at water-land boundaries. Clear, deep water absorbs nearly all incident near-IR energy, but shallow or turbid water's infrared absorption can vary greatly depending on depth and water constituents (Wetzel, 1975). Water's visible and infrared reflectances in general differ substantially from those of the bare soil or vegetation that may be exposed at times of low water depth. Because IR and red reflectances are the two components of the NDVI, tidal and seasonal depth changes present a significant source of NDVI contamination. The APES has a long coastline on the Atlantic Ocean, as well as numerous large lakes and other smaller bodies of water such as hog lagoons and ponds. Upon examination of the cleaned NDVI stack prior to classification, it was apparent that water-containing and water edge pixels had to be removed from the image.

A water mask was created using Landsat ETM+ 30 m data using image dates that varied for each World Reference System-2 (WRS-2) Path/Row due to cloud cover and data quality problems. However, at a minimum, winter and summer images were available for at least two years between 2000 and 2003 for each Path/Row. The nine Path/Row combinations in the APES were separately processed using the following approach.

- The images were georectified and the ETM+ bands 4 (0.76–0.90 μm) and 5 (1.55–1.75 μm) were subset and stacked in chronological order.

- In each band, reflectance values were sampled in water areas, and a reflectance threshold for each band was empirically determined.

- Binary water/non-water masks were created from the individual band thresholds.

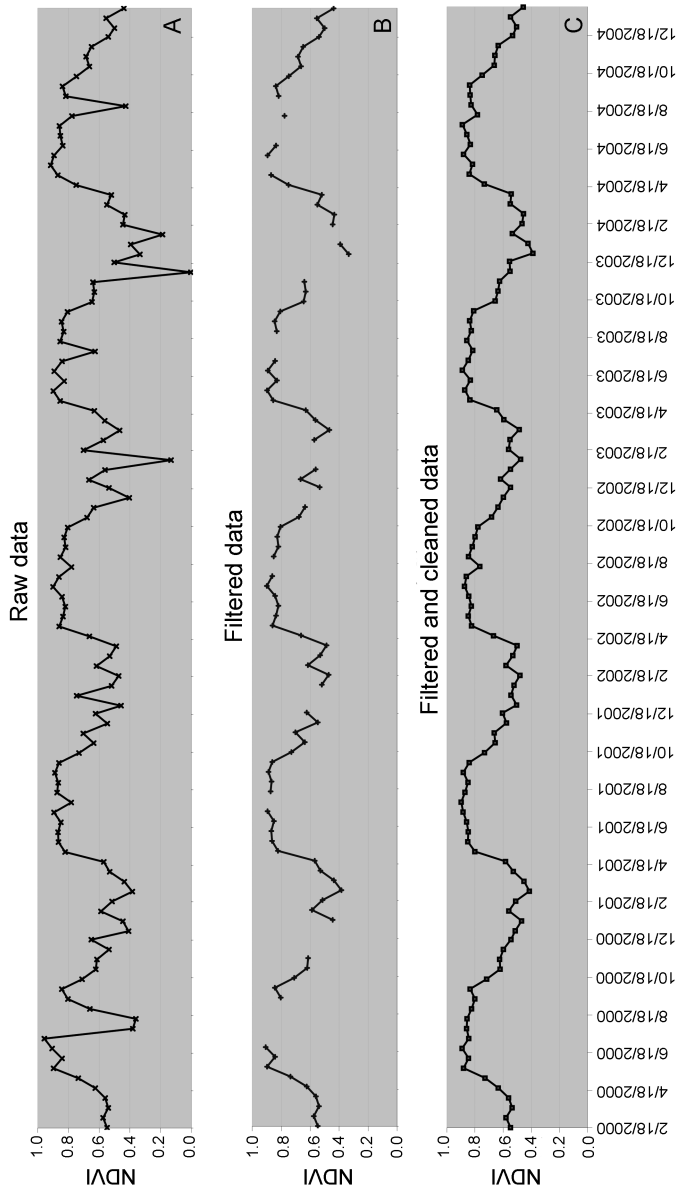

Fig. 5. Sample NDVI profiles showing errors fixed during filtering and cleaning.

- If a pixel was identified as water in all bands, it was assigned to the water category. Otherwise, if ≥2 bands identified the pixel as water, the pixel was assigned to a transitional category. Remaining pixels were assumed to be non-water.

- The transitional pixels were separated using the ISODATA algorithm into 15 clusters. These clusters were then assigned to one of four groups: seasonal water, submerged agriculture areas, shadows, and unclear. Seasonal water pixels were added to the water sample. Submerged areas and shadows were assumed to be non-water.

- Unclear pixels from the previous step were separated using ISODATA into 15 further clusters. These clusters were then divided into the same four groups as above. Again, seasonal water was added to the water sample, and submerged areas and shadows were assumed to be non-water. There were no unclear pixels remaining.

- To generate the final 30 m water masks, pixels in the water sample were assigned a value of one and all others a value of zero.

- Finally, the water mask was scaled up to 250 m by overlaying the MODIS pixel grid over the 30 m water mask. All 250 m pixels that contained any 30 m pixels identified as water were removed from the NDVI image stack and were later classified as water.

This approach, while complex and tedious, produced a high-quality water mask. A qualitative third party review of the mask was performed by comparing the water mask image to the base ETM+ data. No significant discrepancies between the two data sets were found.

Land Cover Classification

Hyperspectral image classification techniques operate on the hypothesis that different image features of interest exhibit reflectance spectra that vary by wavelength in predictable and differentiable ways (Boardman et al., 1995; Green et al., 1998). Temporal image profile features can also exhibit predictable variation. For example, NDVI temporal profiles of healthy deciduous forests show a spring green-up as new leaves are formed, a sustained period of high NDVI in the summer months, a gradual decrease in NDVI in the autumn as senescence occurs and leaves are shed, and sustained low NDVI in the winter. In contrast, the NDVI "spectra" of healthy coniferous forests are much flatter, showing only a small green-up in the warmer months. Figure 6 shows sample spectra from the APES for 2001. The final output class list for the image stack is shown in Table 3. The data stack used for classification consisted of the filtered MODIS-NDVI data for 2001.

A complicating factor that arose during initial attempts to classify the NDVI stack using hyperspectral analysis techniques was that the temporal spectra of agriculture pixels were found to be very similar to other cover classes of interest. For

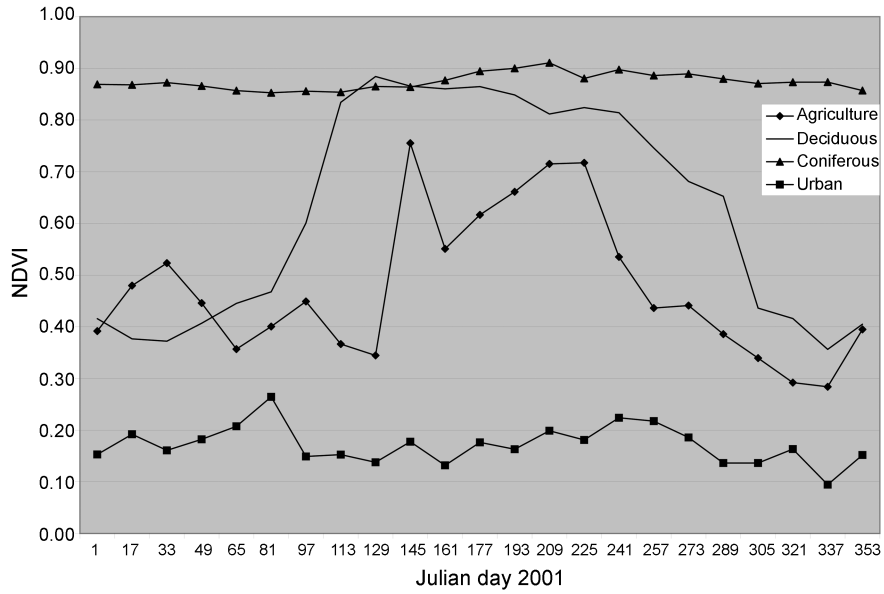

Fig. 6. Sample NDVI temporal spectra for land cover classes of interest.

example, row crop pixels often exhibited temporal profiles that were nearly identical to deciduous forest pixels. Similarly, fallow agriculture pixels were similar to maintained lawns and some coniferous forest pixels. These confusions could not be satisfactorily resolved, and thus the agriculture pixels were separated from non-agriculture using a mask created from Landsat ETM+ 30 m mosaic from March 2000. Pixels identified as agriculture in this mask were assigned that label as their final classified value.

The non-agriculture image stack was processed using the Sequential Maximum Angle Convex Cone (SMACC) endmember model (Gruninger et al., 2004; Research Systems Inc., 2005). The SMACC tool identifies spectral image endmembers and determines abundances through a sequential partitioning of bright and dark pixels. SMACC identifies the brightest pixel in the image and then finds the pixel most different from that bright pixel. Then, SMACC finds the pixel that is most different from the previous two. This process is iterated until a pixel is found that is already a member of a previous group or until the requested number of endmembers is found. The SMACC tool was designed to find endmembers from calibrated hyperspectral data, and was not intended to process multi-temporal imagery. However, the second endmember returned by the SMACC tool was found to correspond well to urban areas. Thresholding this endmember to select values from 0.2 to 0.7 provided a delineation of urban areas in the APES. These areas were subset from the image stack, and the remaining pixels were processed using other hyperspectral techniques, as the SMACC tool did not prove useful for further classification.

Image-based phenology spectra representative of the remaining classes of interest were chosen from homogeneous pixels throughout the study area (identified from the reference data). These included three deciduous forest, three coniferous forest,

and two woody wetland spectra. These phenology spectra were supplied as inputs, along with the image stack, to the Spectral Angle Mapper (SAM) algorithm. The SAM uses an *n*-dimensional angle vector to match pixels to provided spectra. The SAM determines the similarity of two spectra by computing the angle between the spectra (Kruse et al., 1993; Research Systems Inc., 2005). While performing well in separating deciduous and coniferous forest, the SAM was not able to discriminate between deciduous forest and woody wetland pixels, as the temporal profiles for those areas in the APES were very similar. As a result, only coniferous and deciduous classes were output from the SAM. The vast majority of known woody wetland areas were classified as deciduous forest.

The four major aggregated Bailey eco-regions in the APES—mountains, piedmont, coastal plain, and tidewater—differ substantially in topography and contain different vegetation types. To determine whether the classification results would also differ if processed separately by eco-regions, the piedmont region was subset from the APES stack and processed using the same steps as above.

As mentioned above, confusion related to the similarity of agriculture temporal profiles with other cover classes required agriculture pixels to be separated from non-agriculture pixels. To demonstrate the necessity of this step, the above classification steps were repeated with agriculture pixels present in the image stack. The only departure from the steps outlined above is that the SAM was provided with two agriculture spectra in addition to those previously listed.

To determine whether sub-pixel class abundances could be extracted using temporal profiles, the deciduous and coniferous pixels classified by the SAM were analyzed using Linear Spectral Unmixing (LSU). The LSU method assumes that each pixel is a linear combination of its constituent endmembers (Research Systems Inc., 2005). In this case, the endmembers were the coniferous and deciduous forest spectra that were provided to the SAM. The output of the LSU was a two-band image giving the abundances of the two endmembers. The LSU class abundances were correlated with the reference class abundances for all pixels in the reference data set that contained deciduous or coniferous forest. The correlation was measured using the Pearson correlation coefficient.

Accuracy Assessment

The full reference data set for this project consisted of 31,322 samples; however, the final classified map contained only a subset of the classes present in the reference data. The woody aquatic, shrub, bare, herbaceous aquatic, maintained, and pasture could not be differentiated during classification. Thus, reference data pixels containing greater than 50% coverage of these classes were removed from consideration. In addition, since some Level 2 classes were classified, reference data pixels containing a majority of the Level 1–only samples in those categories were also removed from the reference data. For example, since the classification discriminated between coniferous and deciduous forest, the general Level 1 label of "woody" was ignored. Allowing the removed samples to remain in the reference data set would have unfairly lowered the accuracy estimate because all of the samples would have been counted as incorrect. Removing the samples, however, did not undeservedly raise the accuracy

estimate, since the samples were not counted as correct—they simply were not considered. The final number of samples in the reference data set was 17,511 pixels.

Error matrices were constructed to assess the accuracy of each classified map. These included overall and class accuracies when agriculture was removed before classification versus when it was included in the classification procedure, accuracies of the broader aggregate eco-regions (Table 1) when partitioned from the overall classification, and accuracy of the Piedmont aggregate eco-region when it was classified separately. Accuracies were further broken out by class pixel composition threshold (i.e., the area percentage of the majority class in the pixel). This procedure showed how the classification performed with varying degrees of pixel heterogeneity.

RESULTS

The final LC product for the APES is illustrated in Figure 7. This map was derived from the steps described above in which agriculture pixels were removed prior to classification. Table 4 shows accuracy estimates for this classification broken out by pixel composition threshold (PCT). For a high PCT, i.e., 90% or 100%, the classification accuracy estimate was very good—88% and 89%, respectively. As the pixels became more mixed, however, classification accuracy decreased, to a low of 73% at a PCT of 50%. This result was not unexpected, as one would predict that mixed pixels would be more challenging to classify than homogeneous pixels. Standard *z*-tests, based on the Kappa statistic and its variance, showed that, with the exception of the 90% PCT versus 100% PCT matrices, each matrix in Table 4 is significantly different from the others at an alpha level of 0.05.

The poorest performing class was urban, which exhibited widely variable class accuracy estimates over the six PCTs. Much of the variability in urban class accuracy resulted from the fact that there were relatively few urban areas in the 13 DOQs selected as the reference data source. For example, there were only five urban reference pixels at a PCT of 100%. The greatest number of urban reference pixels (n = 266) occurred at the 50% PCT. However, that number represented only 1.9% of the 13,871 total reference pixels at that PCT. Given the small urban sample size, accuracy estimates for the class should be viewed with skepticism. Visual accuracy assessment of urban areas, through overlay of the classification on ETM+ images, indicated that the estimates were not representative of the actual accuracy of the class, which was judged to be satisfactory.

Accuracy Assessment

Overall accuracy assessment results, including estimates with agriculture included in the classification procedure, are illustrated in Figure 8 and summarized Table 5. In addition, Figure 8 and companion Table 5 show the percentages of the reference pixels that fell into each PCT. For example, only 6.0% of the APES was composed of homogeneous 250 m pixels. This result agreed quite well with the 5.0% estimate that was previously derived from the NRB 1998–1999 classification using ETM+ data (Lunetta et al., 2003).

The difference in the accuracies of the "agriculture-excluded" versus "agriculture-included" classifications varied considerably over the PCTs. The largest difference,

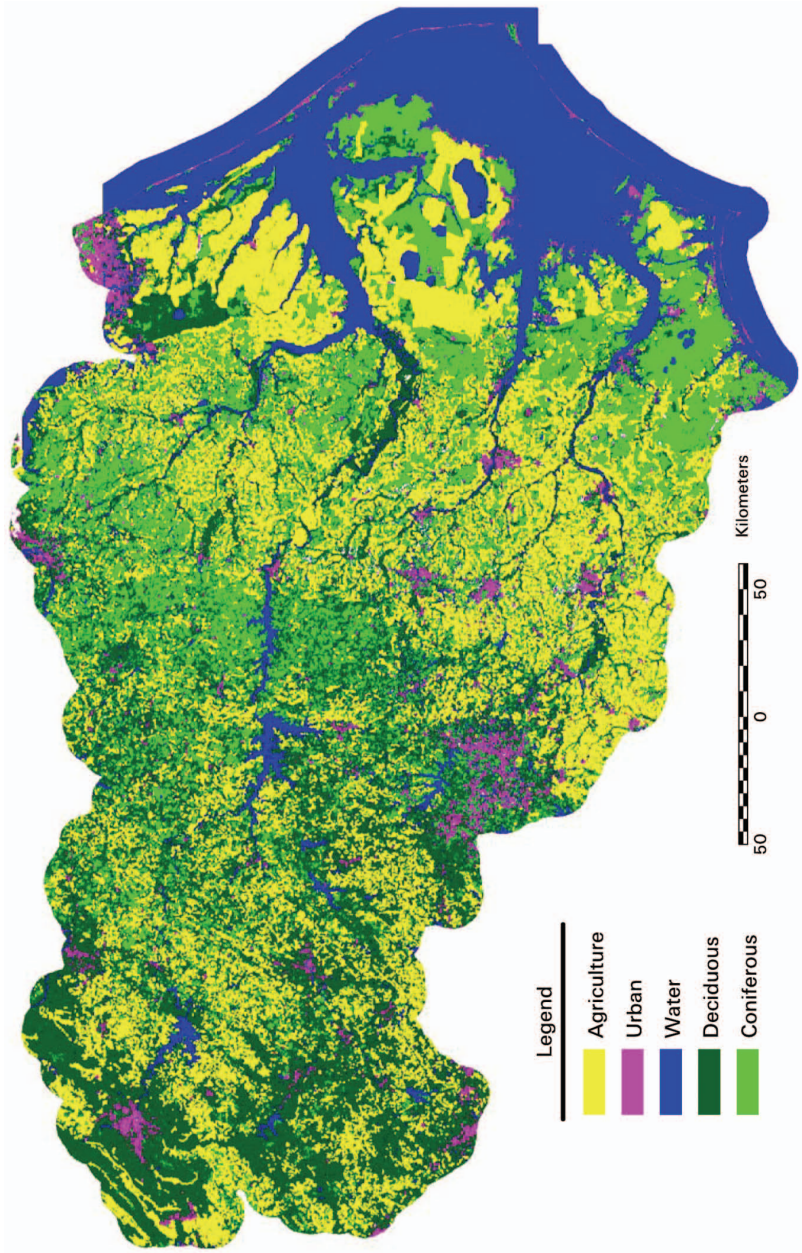

Fig. 7. The final APES land cover classification.

over 20%, occurred at the 50% PCT. *Z-test* results indicated that the difference between the two accuracies was significant at all PCTs. Clearly, the separation of agriculture pixels improved the accuracy of the classification considerably, and was a necessary step in achieving acceptable results. The creation of the agriculture mask was difficult, involving approximately 80 man hours of work, and is an obvious drawback to using such a classification method in larger study areas. However, in parts of the world where agriculture is not as prevalent, or, like the U.S. Midwest, occurs in larger and more regular areas, this approach is highly feasible for larger studies.

Figure 9 illustrates the accuracies of the "agriculture-included" and "agriculture-excluded" classifications broken out by aggregate eco-region and by PCT. The pattern displayed in Figure 8, in which accuracies were higher when agriculture was excluded, held true for every eco-region except the Tidewater, where the accuracy was slightly higher when agriculture was included in the classification. However, *z*-tests showed that none of the Tidewater accuracy differences were statistically significant. All other "agriculture-included" versus "agriculture-excluded" *z*-tests indicated a significant difference except the Coastal Plain 100% PCT test. The likely reason for the Tidewater's deviation from the previous pattern is that the Tidewater in the APES contains several very large collective farms, which are much more homogenous than other smaller farms in the study area. As a result, the spectral classifier was able to identify these areas when agriculture was included. In the Mountains eco-region, the disparity between the two accuracy estimators was quite large, and the accuracy of the "agriculture-included" classification decreased as the PCT increased. This result likely occurred because the agriculture areas in the mountains tended to be linear in shape, and thus increased the number of heterogeneous edge pixels, which the classifier was less able to correctly categorize.

Also shown in Figure 9 are the accuracy estimates for the Piedmont eco-region when it was classified separately from the rest of the study area. At every PCT, the Piedmont-only classifier performed marginally worse, but *z*-test results showed a significant difference only at the 90% PCT. The Piedmont is the most heterogeneous part of the APES, so it is possible that the accuracy estimates were lower because the classifier was not able to take advantage of the purer pixels present in other eco-regions (i.e., Mountains and Tidewater).

Reference Data

A significant concern when using multiple interpreters is inter-interpreter bias. Some degree of bias will always exist; however, proper training and clear guidelines can minimize the problem. For assessment purposes, the two interpreters duplicated efforts on a randomly chosen quarter quad. The Level 1 class percentages assigned by the interpreters to the pixels in this test quarter quad were compared. Table 6 shows the results of this comparison, which indicate that the two interpreters provided very similar class percentages. For example, in 81.2% of the cells for which percentages were provided for the impervious class, the two interpreters gave identical values. At a minimum, across all of the classes, approximately 90% of the cells had a difference of 10% or less. Very few cells, 1.4% for the herbaceous class and 1.7% for the woody class, had differences in assigned class percentages that exceeded 50%. Some of those cases were the result of typographical errors that were subsequently corrected. These

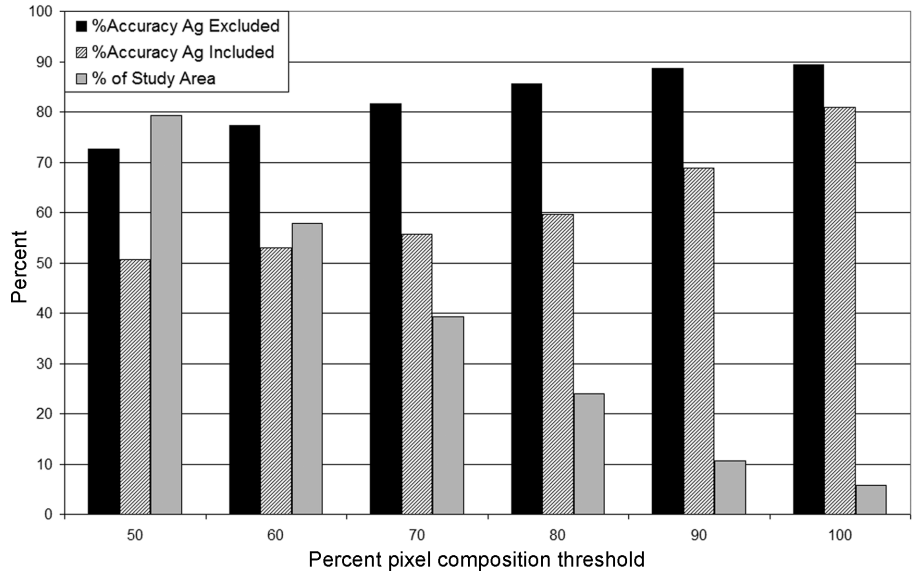

Fig. 8. Overall classification accuracy when agriculture was removed vs. included in classification procedure.

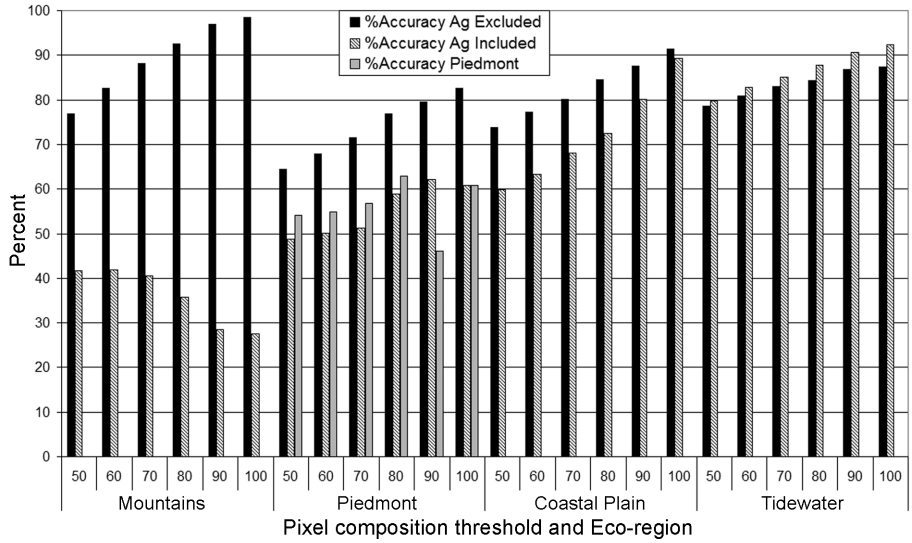

Fig. 9. Accuracies of the individual eco-regions of the main classification, and the accuracy of the Piedmont region when classified separately.

results show that the two interpreters provided very similar interpretations of the test DOQQ, and likely did so throughout the study area.

The reference data sample was not drawn from a completely randomized design. The DOQs used were selected at random, but the entire contents of each DOQ were used as reference data. As a result, spatial autocorrelation among reference pixels

Table 4. Confusion Matrices for Selected Pixel Composition Thresholds

Class	Reference data						Pct. correct	Pct. commission
	Agriculture	Urban	Water	Deciduous	Coniferous	Row Total		
A. Pixel composition threshold of 50%								
Agriculture	**2748**	65	71	478	130	3492	79	21
Urban	2	**23**	7	8	0	40	58	42
Water	48	3	**456**	9	4	520	88	12
Deciduous	933	40	97	**5006**	243	6319	79	21
Coniferous	243	135	96	1107	**1843**	3424	54	46
Column total	3974	266	727	6608	2220	n = 13,871		
Pct. correct	69	9	63	76	83			
Pct. omission	31	91	37	24	17	Overall accuracy = 73%		
						Kappa = 0.59		
B. Pixel composition threshold of 60%								
Agriculture	**2188**	38	41	298	83	2648	83	17
Urban	1	**16**	5	4	0	26	62	38
Water	46	2	**410**	8	1	467	88	12
Deciduous	608	19	51	**3725**	141	4544	82	18
Coniferous	139	86	51	620	**1486**	2382	62	38
Column total	2982	161	558	4655	1711	n = 1,012		
Pct. correct	73	10	73	80	87			
Pct. omission	17	90	27	20	13	Overall accuracy = 78%		
						Kappa = 0.66		
C. Pixel composition threshold of 70%								
Agriculture	**1645**	24	29	157	56	1911	86	14
Urban	1	**11**	4	3	0	19	58	42
Water	42	0	**362**	3	0	407	89	11
Deciduous	355	9	35	**2521**	68	2988	84	16
Coniferous	78	44	20	286	**1081**	1509	72	28
Column total	2121	88	450	2970	1205	n = 6,881		
Pct. correct	78	13	80	85	90			
Pct. omission	22	87	20	15	10	Overall accuracy = 82%		
						Kappa = 0.73		
D. Pixel composition threshold of 80%								
Agriculture	**1136**	12	19	65	28	1260	90	10
Urban	1	**3**	1	0	0	5	60	40
Water	41	0	**321**	0	0	362	89	11
Deciduous	195	3	18	**1450**	27	1693	86	14
Coniferous	36	24	3	101	**683**	847	81	19
Column total	1409	42	362	1616	738	n = 4,196		
Pct. correct	81	7	89	90	93			
Pct. omission	19	93	11	10	7	Overall accuracy = 86%		
						Kappa = 0.80		

(*table continues*)

Table 4. Continued

Class	Reference data						Pct. correct	Pct. commission
	Agriculture	Urban	Water	Deciduous	Coniferous	Row Total		
E. Pixel composition threshold of 90%								
Agriculture	**649**	5	7	25	11	697	93	7
Urban	0	**1**	0	0	0	1	100	0
Water	36	0	**277**	0	0	313	88	12
Deciduous	67	0	3	**473**	7	550	86	14
Coniferous	9	4	1	30	**249**	293	85	15
Column total	761	10	288	528	267	n = 1,862		
Pct. correct	85	10	96	90	93			
Pct. omission	15	90	4	10	7	Overall accuracy = 88%		
						Kappa = 0.84		
F. Pixel composition threshold of 100%								
Agriculture	**411**	3	1	11	5	431	95	5
Urban	0	**1**	0	0	0	1	100	0
Water	33	0	**236**	0	0	269	88	12
Deciduous	30	0	3	**153**	2	188	81	19
Coniferous	1	1	1	13	**109**	125	87	13
Column total	475	5	241	177	116	n = 1,019		
Pct. correct	87	20	98	86	94			
Pct. omission	13	80	2	14	6	Overall accuracy = 89%		
						Kappa = 0.85		

Table 5. Overall Percent Accuracy by Pixel Composition Threshold

Threshold pct.	Ag. excluded	Ag. included	Pct. of study area
50	73	51	79
60	77	53	58
70	82	56	39
80	86	60	24
90	88	69	11
100	89	81	6

may have introduced bias into the sample (Congalton and Green, 1999). This bias was mitigated by assessing the accuracy of the classification at the different PCTs. As the pixels became more pure, they became less numerous, and so not as affected by spatial autocorrelation. For example, at the 100% PCT, there were only 1019 reference pixels throughout the APES. Finally, as is shown in Figure 8, only approximately 80% of the study was represented by the reference data. Areas such as wetlands were

Table 6. Inter-interpreter Comparison of Reference Data at Level 1

Difference	Percentage of cells				
	Impervious	Herbaceous	Woody	Water	Other
0	81.2	50.2	44.1	90.4	99.1
>0–10	18.3	39.7	45.2	9.6	0.1
>10–20	0.5	6.3	6.8	0	0.3
>20–50	0	2.4	2.2	0	0.3
>50	0	1.4	1.7	0	0.2

excluded from the reference data. As a result, accuracy estimates for these areas are not provided, and should not be inferred from the existing reference data.

Linear Spectral Unmixing

The Linear Spectral Unmixing algorithm that was applied to pixels containing deciduous and coniferous forests resulted in class abundances that correlated well with the pixel compositions quantified in the reference data. The correlation coefficients (r^2) between the sub-pixel LSU data and the reference data were 0.89 for the deciduous class, and 0.83 for the coniferous class. This result indicated that the unmixing algorithm was able to extract sub-pixel data from the coarse 250 meter pixels. The use of greater temporal resolution compensated for the lack of spatial resolution in the image data. This analysis was limited, in that it examined only two classes that had quite different phenologic responses; however it may be possible to obtain similar results using less distinct classes.

DISCUSSION

Spectral- and temporal-based classification techniques are employed to accomplish the same goal, accurate LC classification; however the methods by which they do so are very different. The choice of a spectral or temporal (or hybrid) classifier depends upon the objectives of the study. Spectral classifiers can be used to achieve goals that temporal classifiers cannot. For example, a temporal classifier would likely be less useful for differentiating species of co-located deciduous trees, as the spring green-up would be occurring nearly simultaneously among the species. In contrast, a temporal classifier may provide better results than a purely spectral classifier when mapping agricultural crops, as planting and harvest dates for different crop species may vary significantly.

When choosing a data source for temporal classification, sensor resolution is an important issue. High spatial resolution satellites have relatively long revisit times, and, given inevitable weather or technical problems, it can be difficult to obtain sufficient images to incorporate adequate phenology information for a large area such as the APES. To construct only spring and fall ETM+ mosaics for the APES would require 18 scenes. The 16-day revisit time of the Landsat satellites, and the high likelihood that clouds would obscure parts of one or more scenes, make it likely that

images could not be acquired during the desired times. In contrast, lower spatial resolution sensors such as MODIS provide daily coverage of the earth, and so weather events are much less of an obstacle.

For any multi-temporal analysis, positional accuracy of the data sets used is of paramount importance. This is particularly true for an approach such as that presented here, which requires near perfect co-registration of the images in the time series stack. MODIS project scientists have reported 50 m geolocation accuracy at nadir (Wolfe et al., 2002). An in-house assessment of MODIS geolocation accuracy in the APES that encompassed the entire MODIS swath (both nadir and off-nadir) resulted in an RMS error of approximately 113 meters. This result was well below one-half of a MODIS pixel width, and so the positional accuracy of the data appeared to be quite good.

The inability of the classifier to discriminate wetlands was a significant limitation of the approach applied here. However, classifying wetlands has long been problematic in remote sensing (Ozesmi and Bauer, 2002). In the first iteration of the Multi-Resolution Land Characteristics (MRLC) Consortium's National Land Cover Data (NLCD), wetlands were not classified spectrally. Rather, the U.S. Fish and Wildlife Service's National Wetlands Inventory (NWI) data were used as a mask to identify wetlands. An analogous approach could have been used in this study.

The temporal classification method presented here is expected to be applicable to areas other than the APES. The most onerous tasks would be creation of the water and agriculture masks. In this study, the water mask was developed from ETM+ data already in hand rather than from the MODIS data stack. Preliminary attempts to mask water from the MODIS data using NDVI-band thresholds produced usable results, but did not allow for sub-pixel identification of water. The decision was taken to use the ETM+ data to make the water mask as broad as possible to remove even small streams and ponds. Acceptable results may be obtained through the use of a less comprehensive water mask based on the MODIS data.

The agriculture mask was developed by heads-up digitizing ETM+ data, and required a considerable time investment. This step may be eliminated through the use of the forthcoming agriculture mask data set for the United States that is being developed by the United States Department of Agriculture (USDA), National Agriculture Statistical Service (NASS). The product is based on ETM+ data and will provide a high-resolution delineation of agricultural plots that may be substituted for a manually derived mask, thus substantially reducing the preparation time required for a project of this type (NASS, 2002). Using the NASS product in combination with a water mask created from the MODIS data, a classification of this type for an area the size of the APES could be completed relatively quickly and economically.

CONCLUSIONS

This study demonstrated successful phenology-based LC classification using established hyperspectral analysis techniques applied to temporal data. The value added to relatively coarse spatial resolution images by incorporating vegetation phenology information was significant; sub-pixel information was extracted from the image data through the use of multiple temporal observations per pixel. The approach described herein provides LC classification accuracies comparable to those of maps

derived from higher resolution data. Level 1 accuracy estimates of the 80 m Landsat MSS–derived North American Landscape Characterization (NALC) project for Mexico range from 60% to 67% and had very high standard deviations (±20%) (Lunetta et al., 2002a). Accuracy assessment of the 1992 MRLC NLCD data set yielded Level 1 estimates ranging from 70% to 85%, depending on which federal region was considered (Wickham et al., 2004). Given that the accuracy results are comparable, data availability is greater, costs are lower, and the approach is simpler than spectral techniques typically used in large projects, temporal classification may provide a viable alternative for regional or national classifications.

ACKNOWLEDGMENTS

The authors would like to thank Barry Hester and Robert Stallings for their DOQ interpretation support for this research.

REFERENCES

Anderson J. R., Hardy, E. F., Roach, J. T., and R. E. Witmer, 1976, *A Land Use and Land CoverClassification System for Use with Remote Sensor Data,* Washington, DC: United States Geological Survey, Prof. Paper 964.

APNEP, 2005, Albemarle-Pamlico National Estuary Program, North Carolina Department of Environment and Natural Resources [http://www.apnep.org].

Boardman, J. W., Kruse, F., and R. O. Green, 1995, "Mapping Target Signatures via Partial Unmixing of AVIRIS Data" in Pasadena, CA, *Summaries of the Fifth Annual JPL Airborne Earth Science Workshop, AVIRIS Workshop Proceedings*, Vol. 1, 23-26.

Congalton, R. and K. Green, 1999. *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*, Boca Raton, FL: CRC/Lewis Press, 137 p.

Cowardin, L. M., Carter, V., Golet, F. C., and E. T. LaRoe, 1979, *Classification of Wetlands and Deepwater Habitat of the United States*, Washington, DC: Fish and Wildlife Service, U.S. Department of the Interior.

Goetz, S. J., Prince, S. D., Thawley, M. M., Smith, A. J., and R. Wright, 1999, "The Mid-Atlantic Regional Earth Science Applications Center (RESAC): An Overview," in *American Society of Photogrammetry and Remote Sensing, Conference Proceedings*, Washington, DC.

Goetz, S. J., Varlyguin, D., Smith, A. J., Wright, R. K., Prince, S. D., Mazzacato, M. E., Tringe, J., Jantz, C., and B. Melchoir, 2004, "Application of Multitemporal Landsat Data to Map and Monitor Land Cover and Land Use Change in the Chesapeake Bay Watershed," in *Analysis of Multi-temporal Remote Sensing Images*, Smits, P. and L. Bruzzone (Eds.), Singapore: World Scientific Publishers, 223-232.

Green, R. O., Eastwood, M. L., Sarture, C. M., Chrien, T. G., Aronsson, M., Chippendale, B. J., Faust, J. A., Pavri, B. E., Chouit, C. J., Solis, M., Olah, M. R., and O. Williams, 1998, "Imaging Spectroscopy and the Airborne Visible Infrared Imaging Spectrometer (AVIRIS)," *Remote Sensing of Environment*, 65(3):227-248.

Gruninger, J., Ratkowski, A. J., and M. L. Hoke, 2004, "The Sequential Maximum Angle Convex Cone (SMACC) Endmember Model," in *Proceedings SPIE, Algorithms for Multispectral and Hyper-spectral and Ultraspectral Imagery*, Vol. 5425-1, Orlando, FL.

Homer, C., Huang, C., Limin, Y., Wylie, B., and M. Coan, 2004, "Development of a 2001 National Land-Cover Database for the United States," *Photogrammetric Engineering & Remote Sensing*, 70(7):829-840.

Jensen, J. R., 2004, *Introductory Digital Image Processing: A Remote Sensing Perspective*, Upper Saddle River, NJ: Pearson Prentice Hall, 526 p.

Kruse, F. A., Lefkoff, A. B., Boardman, J. B., Heidebrecht, K. B., Shapiro, A. T., Barloon, P. J., and A. F. H. Goetz, 1993, "The Spectral Image Processing System (SIPS)—Interactive Visualization and Analysis of Imaging Spectrometer Data," *Remote Sensing of Environment*, 44:145-163.

Lambin, Eric F., 1996, "Change Detection at Multiple Temporal Scales—Seasonal and Annual Variations in Landscape Variables," *Photogrammetric Engineering & Remote Sensing*, 62(8):931-938.

Lillesand, T. W., Kiefer, R. W., and Chipman, J. W., 2003, *Remote Sensing and Image Interpretation*, 5th ed., New York, NY: John Wiley and Sons, Inc, 704 p.

Loveland, T. R., Merchant, J. W., Brown, J. F., Ohlen, D. O., Reed, B. C., Olson, P., and J. Hutchinson, 1995, "Seasonal Land-Cover Regions of the United States," *Annals of the Association of American Geographers*, 85(2):339-355.

Loveland, T. R., Reed, B. C., Brown, J. F., Ohlen, D. O., Zhu, Z., Yang, L. , and J. W. Merchant, 2000, "Development of a Global Land Cover Characteristics Database and IGBP Discover from 1 km AVHRR Data," *International Journal of Remote Sensing*, 21(6-7):1303-1330.

Lunetta, R. S., Alvarez, R., Edmonds, C. M., Lyon, J. G., Elvidge, C. D., Boniface, R., and C. Garcia, 2002, "NALC/Mexico Land-Cover Mapping Results: Implications for Assessing Landscape Condition," *International Journal of Remote Sensing*, 23(16):3129-3148.

Lunetta, R. S., Ediriwickrema, J., Iiames, J., Johnson, D. M., Lyon, J. G., McKerrow, A., and A. Pilant, 2003, "A Quantitative Assessment of a Combined Spectral and GIS Rule-Based Land-Cover Classification in the Neuse River Basin of North Carolina," *Photogrammetric Engineering & Remote Sensing*, 69(3):299-310.

Lunetta, R. S., Ediriwickrema, J., Johnson, D. M., Lyon, J. G., and A. McKerrow, 2002, "Impacts of Vegetation Dynamics on the Identification of Land-Cover Change in a Biologically Complex Community in North Carolina, USA.," *Remote Sensing of Environment*, 82:258-270.

Lunetta, R. S., Iiames, J. S., Knight, J. F., Congalton, R. G., and T. H. Mace, 2001, "An Assessment of Reference Data Variability Using a "Virtual Field Reference Database"," *Photogrammetric Engineering & Remote Sensing*, 63(6):707-715.

Lunetta, R. S., Johnson, D. M., Lyon, J. G., and J. Crotwell, 2004, "Impacts of Imagery Temporal Frequency on Land-Cover Change Detection Monitoring," *Remote Sensing of Environment*, 89(4):444-454.

NASS (National Agriculture Statistical Service), 2002, *Census of Agriculture: Volume 1. Geographic Area Series* [http://www.nass.usda.gov/census].

Ozesmi, S. L. and M. E. Bauer, 2002, "Satellite Remote Sensing of Wetlands," *Wetlands Ecology Management,* 10(5):381-402.

Reed, B. C., Brown, J. F., VanderZee, D., Loveland, T. L., Merchant, J. W., and D. O. Ohlen, 1994, "Measuring Phenological Variability from Satellite Imagery," *Journal of Vegetation Science*, 5:703-714.

Research Systems Inc., 2005, *ENVI 4.1 User's Guide*, Boulder, CO: Research Systems Inc.

Roberts, D. H., Lehar, J., and J. W. Dreher, 1987, "Time Series Analysis with CLEAN I Derivation of a Spectrum," *The Astronomical Journal*, 93(4):968-989.

Roberts, D. A., Green, R. O., and J. B. Adams, 1997, "Temporal and Spatial Patterns in Vegetation and Atmospheric Properties from AVIRIS," *Remote Sensing of Environment*, 62(3):223-240.

Senay, G. B. and R. L. Elliott, 2000, "Combining AVHRR-NDVI and Land Use Data to Describe Temporal and Spatial Dynamics of Vegetation," *Forest Ecology and Management*, 128:83-91.

Spruill, T. B., Harned, D. A., and G. McMahon, 1995, *The Albemarle-Pamlico Drainage National Water-Quality Assessment Study: Background and Design*, Reston, VA: United States Geological Survey, Open-File Report 95-164, 4 p.

Varlyguin, D., Wright, R., Goetz, S. J., and S. D. Prince, 2001, "Advances in Land Cover Classification for Applications Research: A Case Study for the mid-Atlantic," in *American Society of Photogrammetry, and Remote Sensing, Conference Proceedings*, St. Louis, MO.

Wetzel, R. G., 1975. *Limnology*, Philadelphia, PA: W. B Saunders, 743 p.

Wickham, J. D., Stehman, S. V., Smith, J. H., and L. Yang, 2004, "Thematic Accuracy of the 1992 National Land-Cover Data for the Western United States," *Remote Sensing of Environment*, 91:452-468.

Wolfe, R. E., Nishihama, M., Fleig, A. J., Kuyper, J. A., Roy, D. P., Storey, J. C., and F. S. Patt, 2002, "Achieving Sub-pixel Geolocation Accuracy in Support of MODIS Land Science," *Remote Sensing of Environment*, 83:31-49.

APPENDIX

Appendix Table 1. Land Cover Class Definitions

Class	Definition
Agriculture	Areas used for the production of crops or livestock grazing
Urban	Includes areas with a mixture of constructed materials and vegetation, highly developed areas, and infrastructure
Water	All areas of open water (streams/canals, lakes/ponds, reservoirs, bays, open marine)
Deciduous	Areas dominated by trees that shed foliage simultaneously in response to seasonal change
Coniferous	Areas dominated by trees that maintain their leaves all year; canopy is never without green foliage

Source: Adapted from Cowardin et al., 1979.