Multi-temporal Analysis of Landsat Data to Determine Forest Age Classes for the Mississippi Statewide Forest Inventory–Preliminary Results

Curtis A. Collins, David W. Wilkinson, and David L. Evans
Forest and Wildlife Research Center, Mississippi State University
Box 9681
Mississippi State, Mississippi 39762
ccollins@sitl.cfr.msstate.edu

Abstract— **The use of Landsat data to aid in forest sampling stratification, area estimation, and future resource assessment through growth models is currently being investigated for the state of Mississippi with the goal of better understanding present and future wood resources. In such analyses, and as a part of this investigation, change detection techniques are being exploited to help determine these forest stand ages in approximate five year intervals. This preliminary report looks at post classification comparisons and temporal image differencing as two means to find these dates. The results find the post classification comparisons techniques, in an unrefined use, to work moderately well (overall accuracy = 0.6157, KHAT = 0.5386) and temporal image differencing with NDVI and tasseled cap transformations to disagree with each other in predicted age class sizes with no assessment data to validate accuracy at this time.**

Keywords-change detection; forest change; forest age; forest classification; forest; hardwood; pine; multitemporal; Landsat

I. Introduction

The Landsat program has yielded a plethora of geographically repeated time-series datasets that not only provide for a great deal of information extraction from an individual location's spectral and spatial component, but also through the differing degrees of temporal resolution afforded by the time-lapse of the data acquisition process. With the history of the program spanning three decades with scenes being revisited around the globe every 18 (Landsats 1 through 3) to 16 (Landsats 4 through 7) days, a wealth of temporal resolutions dating back to the early 1970's can be very useful in analyzing more lengthy landcover changes, including those related to forest age derivation. In combination with the broad forest type mapping available through single-year pre- (leaf-on) and post-leaf senescence (leaf-off) datasets, timber types can be determined and aged for sampling design purposes (such as strata definition) and growth and yield estimation. This not only provides government entities with a basis to judge resource status, but it also provides a means to encourage wise resource utilization by portions of the private sector spurring economic development and its many byproducts. This more detailed classification could also become useful in a variety of other environmental analyses leading to habitat mapping for a variety of forest species and other similar applications.

In order for this more detailed classification to be derived, past use of multi-image change detection was investigated. Being defined as the process of identifying differences, specifically radiance differences, by a remote sensor at different times [10], change detection analyses can be performed through a variety of methods. References [3] and [12] provide examples of long term, large scale change detection projects for natural resource management. References [5], [8], and [10] also provide general background to a wide list of these methods. Reference [1], in a similar project, uses Landsat-based area estimates and classes derived from single and multi-temporal datasets in inventorying the forest resources of Minnesota. With these methods in mind, however, [7] describes the basic change detection procedures, while not exhaustive, as: post classification comparisons, where each date's imagery is separately classed and compared to each other; classification of multi-temporal datasets, where a single classification is performed on a combined multi-date dataset; principal components analysis, where uncorrelated principal components layers from a combined multi-date dataset can be related to change; and temporal differencing or ratioing, where ratios or differences are taken from a multi-date dataset and observed to locate areas of change.

This study's first test method uses post classification comparisons as it looks at archived leaf-on images and their use in the classification of forest and non-forest areas. By independently classifying each year's set of leaf-on images, a series of forest/non-forest thematics can be used to tag each forested pixel from the most recent forest thematic layer with a date of origin, if one is found. Likewise, forest pixels in the next most recent dataset that were found to be non-forest and non-water in the most recent dataset can be assumed in some cases to be remaining in forest use, so the target classification can at this point be updated with a new class representing areas of forest regeneration. Prior to applying this method to the entire state, it will be tested against a land ownership GIS (in vector polygon format) of 73,767 hectares of intensively managed timberlands in southeast Louisiana attributed with stand age, among other things.

Another method explored the use of the time-series dataset's six leaf-on Landsat scenes covering a four county area in east Mississippi, which was to be analyzed using a temporal differencing method. This analysis was performed using two image transformations, Normalized Difference Vegetation Index (NDVI) and tasseled cap, resulting in a set of NDVI and tasseled cap images including brightness, greenness, and wetness (for TM and ETM+ scenes only) bands. For each date interval the pair or triplicate of tasseled cap or NDVI images were differenced, masked, stacked, and analyzed by a supervised classification procedure, similar to procedures performed in [2]. This resulted in two final land cover change images, one for both differenced transformations. At this time, preliminary results from this procedure were judged logically as well as by comparison to each other in order to observe agreement between these methods as well as, eventually, against the post classification comparisons method.

II. Study Areas

A. East Mississippi Study Area

The first area of interest in this project consists of 4,939 km^2 in east-central Mississippi, which includes the counties of Choctaw, Clay, Oktibbeha, and Winston. This area includes a variety of spectral responses that encompasses a small urban area around Starkville and Mississippi State University, agricultural areas along the black-belt prairie region located in the northeastern part of the study area along with pine, hardwood, and mixed pine-hardwood forested areas in central and southern portions of the area. Ownership types vary from small less-active owners to large active forest product-minded owners to the more diversified land management indicative of federal and state government ownership.

B. Southeast Louisiana Study Area

The second area of interest in this project consists of 3,947 km^2 in southeast Louisiana, which includes the parishes of St. Tammany and Washington. This area also includes a variety of spectral responses that encompasses urban and suburban areas in the greater New Orleans metropolitan area along with scattered agricultural areas, marsh along the Gulf of Mexico and Lake Pontchartrain, hardwood swamps in several riverine systems along with pine, hardwood, and mixed pine-hardwood forested areas in central and northern portions of the area. Ownership types also vary from small less-active owners to large active forest product-minded owners to the more diversified land management indicative of federal and state government ownership.

III. Methods

A. Data Description and Preprocessing

After choosing Landsat's TM (i.e., ETM+) and MSS sensors with their spatial (TM resolution is approximately 30 m and MSS resolution is approximately 60 m) and spectral (TM bands 2, 3, 4, and 5 and MSS bands 1-4) resolutions, the purpose of defining forest age relied on determining an optimal temporal resolution. This resolution was chosen to be approximately five years. These dates were also influenced by the acquisition of North American Landscape Characterization (NALC) data which were delivered as a decade-based leaf-on Landsat dataset. With the onset of this effort beginning in 2003, this meant that target dates were set at 2003, 1996, 1991, 1986, 1980, and 1973 after considering these and other data availability issues. Individual scenes were then selected and purchased from Earth Resources Observation Systems (EROS) to fill the gaps so that an archive of Landsat scenes covering the entire state existed for the specified years from both leaf-on and leaf-off seasons.

Next the data were georegistered to mosaicked county-level Digital Orthophoto Quarter Quadrangle (DOQQ) imagery which were registered by USGS National Map accuracy standards with a spatial resolution of 3 m. This was done in two stages with the panchromatic (from ETM+) layers from the leaf-off 2003 dataset being registered first with all subsequent scenes being registered to these single-band base-line images. Care was taken so that Root Mean Squared (RMS) errors from the georegistering process, performed in Leica's Erdas Imagine 8.7 using either first or second order models, never exceeded one pixel. This subpixel accuracy results in measures of less than 30 m for the TM data and less than 60 m for the MSS data. This accuracy is pertinent as image-to-image matching is essential in change detection analyses as demonstrated in the suggestion of one-quarter to one-half pixel RMS errors [7].

B. Present Classification

With the preprocessing done, the actual processing of the data were undertaken with the initial classification of the most recent, 2003 dataset in order to isolate 5 landcover classes: water, non-forested land, hardwood forestland, pine forestland, and mixed hardwood-pine forestland. This was done in two stages involving the two seasonal datasets. The first stage used the leaf-on dataset to identify regions of a scene that were occupied by water, forest, and non-forest cover types. This was done by isodata clustering in Imagine 8.7 using 100 or more clusters in order to reduce cluster-to-class confusion. Each cluster was in turn interpreted against the original image in order to recode this 100-plus class thematic into a three-class (water, non-forest, forest) thematic.

The resulting three-class thematic layer was in turn used to mask the leaf-off dataset from the same period with the forest class in order to differentiate pine, hardwood, and mixed portions of the forest class's area. Training areas defined through the classification of DOQQs made the use of maximum likelihood classification possible using the field definition of pine-hardwood mixed areas. This definition stated that an 80% majority coverage among overstory hardwood or pine crowns resulted in a Landsat training pixel being coded as either a pure hardwood or pure pine training pixel. If less than 80% of one of the two pure classes existed, then the training pixel was labeled mixed. This too was accomplished by using isodata clustering with 100 or more clusters. The clusters were recoded, using expert knowledge, into pine, hardwood, and “other” classes and a zonal operation was performed, making sure to ignore “other” pixels, using the resolution and origin of the Landsat scene to be classified so that pine, hardwood, and mixed pixels could be identified for use in the maximum likelihood classifier in Imagine 8.7. Four or more predominantly forested DOQQs were used in training set definitions and signature development for each scene classified in this manner. The result from this maximum likelihood procedure was a most recent thematic layer from which various multi-temporal analytical techniques were to be tested in order to tag forest origin to presently forested pixels and to determine candidate forest regeneration areas.

C. Post Classification Comparisons

Using the concept of post classification comparisons, multi-temporal analyses were first performed in two manners with both involving independently classified datasets from each date interval of Landsat imagery over the southeast Louisiana test area (scene details are given in Table I). One used whole time-series scenes while the other used a masking protocol.

The first process used the most recent thematic layer created from DOQQ training data and a maximum likelihood classification along with subsequent leaf-on time interval imagery. Each of these whole scenes from previous periods were clustered, again using 100 or more clusters, with each scene being interpreted and recoded into a forest/non-forest dataset with classes: water, non-forest, and forest. When all dates were classed in this manner over the area of interest in southeast Louisiana, a model was constructed in Imagine 8.7's spatial modeler that tagged each forested pixel in the present five-class thematic with an approximate year of origin, or regeneration date, based on the most recent corresponding forest/non-forest pixel coded as non-forest. In the event where a location remained forested in the entire time series, it was noted as not having a known date of regeneration. This method accounted for all regeneration dates except the most recent. These areas of regeneration were approximated by cross-referencing the areas classed as forested from the 1996 dataset with non-forested areas in the present classification.

The second process involved the same principles as the first process, with the exception of the extent at which the data were clustered and recoded in the later portions of the time-series dataset. Instead of performing independent clustering tasks on each whole time-series scene of interest, each scene was classified in order from most recent with each date being masked to include forested areas whose origin had not yet been determined. This stepwise reduction of classification data was intended to focus on well-defined cluster differences between forest and recently harvested or regenerated areas, eliminating fallow and other areas that might cause classification confusion via fuzzy cluster boundaries in image feature space. Similar to the previous methods, upon clustering and recoding the 1970's era data, which were the final time-series data, all bands were cross-referenced in Imagine 8.7's spatial modeler to output a final forest age thematic for the southeast Louisiana area. Likewise, in order to account for the regions in this area presently in a regenerative state, the non-forest land class from the present classification's thematic was used to mask the 1996 dataset's image upon which clustering and recoding revealed the forest area at this earlier time that is at present in non-forest. Again these locations were classed as forest regeneration.

TABLE I. LANDSAT IMAGE DATA USED ON THE SOUTHEAST LOUISIANA TEST AREA

Approximate Year (ID)	Path	Row	Sensor	Season	Actual Acquisition Date
2003	22	39	ETM+	Leaf-off	12/28/02
2003	22	39	ETM+	Leaf-on	08/03/01
1996	22	39	TM	Leaf-on	09/28/95
1991	22	39	MSS	Leaf-on	10/05/92
1986	22	39	MSS	Leaf-on	08/31/85
1980	23[a]	39	MSS	Leaf-on	09/10/81
1973	22[b]	39	MSS	Leaf-on	10/08/74

a. Differs from previous scenes, Landsats 1-3 used a different path-row sequence than newer missions.

b. These data are from NALC which uses the newer Landsat (post-Landsat 3) path-row sequence.

D. Temporal Image Differencing

A series of six leaf-on Landsat images (Table II), with temporal resolutions ranging from three to eight years, were used in the temporal image differencing change detection procedures representing the approximate time-series interval of interest over the east Mississippi study area. Image normalization was not used because it was believed that the degree of change isolated in the differencing methods were to be greater than any radiometric noise that would occur between the images themselves. Band four of the 1981 image received from EROS contained approximately six bad lines that crossed the southern portion of the study area. An 11x11 distance-weighted (linearly) smoothing kernel was used to correct these bad lines. From this point a series of image transformations were tested using this image differencing technique.

1) Image Transformations

NDVI and tasseled cap transformation layers were created for the Landsat imagery of interest. These transformations can provide information about the current state of the vegetation represented in a pixel, which can be used to determine if a pixel has changed from one date to another. For the NDVI difference images, a decrease in the NDVI value can, and in many cases does, indicate a loss of vegetation. While for the tasseled cap differenced images an increase in brightness values along with a decrease in greenness and wetness (where applicable) will usually indicate vegetation loss.

In order for tasseled cap transformations to be done on the images, the 1996 and 2003 datasets were processed using USGS-developed procedures to convert them from digital number images to at-satellite reflectance value images [11]. Tasseled cap transformations were then run using MSS and TM

TABLE II. LANDSAT IMAGE DATA USED ON THE EAST MISSISSIPPI TEST AREA

Approximate Year (ID)	Path	Row	Sensor	Season	Actual Acquisition Date
2003	22	37	ETM+	Leaf-off	01/26/02
2003	22	37	ETM+	Leaf-on	11/07/01
1996	22	37	TM	Leaf-on	09/28/95
1991	22	37	MSS	Leaf-on	10/05/92
1986	22	37	MSS	Leaf-on	10/21/86
1980	23[a]	37	MSS	Leaf-on	09/10/81
1973	22[b]	37	MSS	Leaf-on	09/12/72

a. Differs from previous scenes, Landsats 1-3 used a different path-row sequence than newer missions.

b. These data are from NALC which uses the newer Landsat (post-Landsat 3) path-row sequence.

data with values from [9], while the TM images were transformed using the at-satellite reflectance procedure in [6].

2) Simultaneous Image Differencing

The tasseled cap and NDVI images for each date were then differenced, by time interval, to produce 16 individual change images (Table III). Each change image was then put through a series of data masks. The first mask involved masking out non-forested pixels which was done by creating a mask from forest/non-forest classifications from the 1996 and 2003 datasets. The assumption at this point was that if a pixel was classified as non-forest for more than two scenes, going backwards in time from the most recent image, then the land use had not changed. With that said, pixels meeting this criteria were deemed as not being used for forestry applications in the 2003 dataset and, thus, would be classified as non-forest.

The second data mask was applied using a threshold derived from image statistics and expert interpretation. A mean and standard deviation were derived from the forest masked difference images. All pixels that were located within plus or minus 1.2816 standard deviations of the mean would be assumed to be no change pixels and removed from the image. The remaining pixels were then presumed to be areas where large amounts of change had occurred.

The twice masked layers were next stacked into two images based on transformation procedures (NDVI or tasseled cap), thus creating an 11- and five-layer change image. These images were then classified using training areas outlined by the user. For each of the six time intervals a minimum of 4,000 pixels per interval were used in creating signatures. The signatures were then used in a maximum-likelihood classification scheme in Imagine 8.7. The resulting classified images represented areas where forest change had occurred over the past 28 years by the time intervals represented. Using a data overlay operation the classified images were superimposed onto the 2003 dataset's five-class thematic (which was derived earlier) so that pixels that were unchanged over the 28 year time span could be determined. This resulted in the final product of a forest age map with six different age classes.

IV. Results and Discussion

A. Post Classification Comparisons

Accuracy assessment among single images can sometimes be a challenge, but in considering the assessment of change detection procedures using multiple stacked scenes the task becomes even more challenging [4]. Part of this problem is reduced, however, if the accuracy assessment treats the final ages output from the southeast Louisiana study area as a single scene's classification. In focusing within the forested area, which was done throughout this work as a high level of user confidence was maintained for the present forest/non-forest thematic, we are simply noting accuracies among predicted and known ages. In order to do this, however, either the GIS data needed to be categorized into classes corresponding to the classified data, or the classified data needed to be recoded so that instead of reflecting the date of the scene where the change was detected, they needed to display the value of the midpoint date between the two scenes where the change occurred.

Since the limiting data type in this case was categorical (discrete), the more continuous data presented in the GIS were abandoned in their raw form and recoded to correspond to the change detected classes. Since the change detection method recognizes areas of regeneration and harvest, the date of regeneration from the GIS could also be biased when compared to change detected dates. For example, an area which was harvested the month before Landsat acquisition will probably not be regenerated until the following year (the usual minimum wait is nine months prior to planting), so if change was noted in this scenario in 1981 and the GIS depicts a regeneration date of 1982, then a one-to-one recode is not needed. Instead, the GIS-based date of regeneration should have this bias corrected for by subtracting one, then recoding these smaller dates into corresponding change categories. For the category containing the image from 1974 (representing the 1973 dataset) we moved the interval minimum to 1969 making the interval width 5 years (our initial temporal resolution target). The resulting test dataset's classes ranged in size from 1,700+ ha to 16,000+ ha. The resulting producer's and user's accuracies are shown in Tables IV and V and the overall accuracies and estimated kappa statistic (or KHAT) were found to be 0.6002 and 0.5139, respectively, for the whole scene trial and 0.6157 and 0.5386, respectively, for the stepwise scene masking trial .

Among the masked and unmasked examinations of post classification comparisons, there appeared to be no tangible differences as overall accuracy proportions and KHAT values were so close to each other in both trials. With regard to overall performance of post classification comparisons techniques the results appeared promising. Some reasons that might explain the lower than expected assessment values include the fact that the test data (GIS dataset) required some assumptions to be made on the part of the user due to the fact

TABLE III. Transformed Change Images to be Used in Simultaneous Image Differencing

Approximate Year (ID) Intervals	Actual Year Intervals	Transformed Layers
2003-1996	2001-1995	NDVI, Brightness[a], Greenness[a], Wetness[a]
1996-1991	1995-1992	NDVI, Brightness[a], Greenness[a]
1991-1986	1992-1986	NDVI, Brightness[a], Greenness[a]
1986-1980	1986-1981	NDVI, Brightness[a], Greenness[a]
1980-1973	1981-1972	NDVI, Brightness[a], Greenness[a]

a. Derived from tasseled cap transformations.

TABLE IV. User's and Producer's Accuracy Values for the Whole Scene Post Classification Comparisons Procedure

Approximate Year (ID) Intervals	Actual Year Intervals	Producer's Accuracy Statistics	User's Accuracy Statistics
2003 (present)	2001	0.7091	0.7675
2003-1996	2001-1995	0.8847	0.5041
1996-1991	1995-1992	0.5511	0.6635
1991-1986	1992-1985	0.6884	0.5200
1986-1980	1985-1981	0.2334	0.7558
1980-1973	1981-1974	0.4966	0.2571

TABLE V. USER'S AND PRODUCER'S ACCURACY VALUES FOR THE STEPWISE MASKED POST CLASSIFICATION COMPARISONS PROCEDURE

Approximate Year (ID) Intervals	Actual Year Intervals	Producer's Accuracy Statistics	User's Accuracy Statistics
2003 (present)	2001	0.7000	0.7728
2003-1996	2001-1995	0.9066	0.4957
1996-1991	1995-1992	0.5330	0.6700
1991-1986	1992-1985	0.6828	0.564
1986-1980	1985-1981	0.3695	0.7503
1980-1973	1981-1974	0.4644	0.2438

that the GIS attribute of interest involves the date of regeneration, which is not necessarily linked to harvest date, which is what the change detection process is set to determine. The other inconsistency that could be causing problems involves the georectification of the data, which was done completely in Mississippi and, even though the test data were located very close to the state boundary, there might be some spatial misalignment leading to poorer than expected classification results.

B. Temporal Image Differencing

With ground-based and photo interpretative accuracy assessments planned to validate the statewide classification, including the four counties in the east Mississippi study area, there will at a future point be more definitive results from the temporal image differencing methods initiated in this paper. For the moment, however, these definitive results are not available. The only available means of determining results from this process at this point include logical observations and comparisons between the use of NDVI and tasseled cap data, which are noted in Table VII.

While the area estimates appear somewhat realistic in distribution when the NDVI layers were used, since the counts are nearly uniform in nature across the time-series of interest, the tasseled cap column illustrates a more volatile pattern which does not match the NDVI trend. This inconsistent pattern is so extreme that it also appears counterintuitive as, although not constant due to market and weather factors from season-to-season and even year-to-year, over five year intervals timber harvesting is usually more consistent in the southeast U.S. than this trend appears.

TABLE VI. SIMULTANEOUS IMAGE DIFFERENCING TIME-SERIES AREA ESTIMATES

Approx. Year (ID) Intervals	Actual Year Intervals	NDVI-Derived Area (ha)	Tasseled Cap-Derived Area (ha)	Difference (NDVI-T.C.)
2003-1996	2001-1995	24,982	30,268	-5,286
1996-1991	1995-1992	27,171	23,392	3,779
1991-1986	1992-1986	31,938	83,400	-51,462
1986-1980	1986-1981	31,383	12,706	18,678
1980-1973	1981-1972	25,477	56,310	-30,833
1973-	1972-	219,556	167,959	51,597

V. CONCLUSION AND SUMMARY

In conclusion, the techniques attempted so far are forming a rudimentary base for more complex, and possibly more accurate, change detection techniques. Possible transformations that may be of help in future analyses are also being noted, as in the case of NDVI bands in temporal image differencing. It is also of interest to note that the accuracy assessment methods used here are less definitive and crude in comparison to the more robust ground and photo interpretative techniques to be used in the future. Foremost among the areas that need improvement is the identification of the present regeneration class. It is understood that this class is capturing most of the "urban sprawl" development, albeit that Mississippi is not as prone to this as other states, but it is a situation that should and can be better handled. All-in-all it is a good start in aiding a complex procedure, monitoring the timber resources of Mississippi.

ACKNOWLEDGMENTS

A sincere debt of gratitude is owed to Weyerhaeuser company for allowing the use of their GIS data in examining the accuracy of the post classification comparisons process as well as various individuals from the Department of Forestry at Mississippi State University and the Mississippi Institute for Forest Inventory.

REFERENCES

[1] M. E. Bauer, et al., "Satellite inventory of Minnesota forest resources," Photogramm. Eng. Remote Sens., vol. 60, pp. 287-298, 1994.

[2] W. Cohen, M. Fiorella, J. Gray, E. Helmer, and K. Anderson, "An efficient and accurate method for mapping forest clearcuts in the pacific northwest using Landsat imagery," Photogramm. Eng. Remote Sens., vol. 64, pp. 293-300, 1998.

[3] W. Cohen, et al., "Characterizing 23 years (1972-95) of stand replacement disturbance in western Oregon forest with Landsat imagery," Ecosystems, vol. 5, pp. 122-137, 2002.

[4] R. G. Congalton and K. Green, Assessing the Accuracy of Remotely Sensed Data: Principles and Practices, Boca Raton, FL: Lewis Publishers, 1999.

[5] P. Coppin, I. Jonckheere, K. Nackaerts, B. Muys, and E. Lambin, "Digital change detection methods in ecosystem monitoring: a review," Int. J. Remote Sens., vol. 25, pp. 1565-1596, 2004.

[6] C. Huang, B. Wylie, L. Yang, C. Homer, and G. Zylstra, "Derivation of a tasseled cap transformation based on Landsat 7 at-satellite reflectance," Int. J. Remote Sens., vol. 23, pp. 1741-1748, 2002.

[7] T. M. Lillesand and R. W. Kiefer, Remote Sensing and Image Interpretation. 3rd ed., New York, NY: John Wiley and Sons, 2000.

[8] D. Lu, P. Mausel, E. Brondizio, and E. Moran, "Change detection techniques," Int. J. Remote Sens., vol. 25, pp. 2365-2407, 2004.

[9] S. Schrader and R. Pouncey, Erdas Field Guide. 4th ed., Atlanta, GA: ERDAS, 1997.

[10] A. Singh, "Digital change detection techniques using remotely-sensed data," Int. J. Remote Sens., vol. 10, pp. 989-1003, 1989.

[11] USGS, "MRLC 2001 image preprocessing procedure," [Online document] 2001, [November 2004], Available at HTTP: http://landcover.usgs.gov/pdf/image_preprocessing.pdf

[12] E. Yuan and C. Elvidge, "NALC land cover change detection pilot study: Washington D.C. area experiments," Remote Sens. Environ., vol. 66, pp. 166-178, 1998.