ANALYSIS OF BIOMASS DYNAMICS DERIVED FROM SPECTRAL TRAJECTORIES

Powell, S.L., Research Ecologist
scottpowell@fs.fed.us
W.B. Cohen, Research Forester
wcohen@fs.fed.us
R.E. Kennedy, Research Ecologist
robertkennedy@fs.fed.us
U.S.F.S. Pacific Northwest Research Station
3200 SW Jefferson Way, Corvallis, OR 97331

ABSTRACT

Quantification of the rates and variability of forest disturbance and regrowth at continental scales remains a critical challenge. We have undertaken an effort to address these issues, through funding from NASA, by coordinating research efforts between the USDA Forest Service Forest Inventory and Analysis and the North American Carbon Program. We present preliminary results of our approach to analyze 20+ years of biennial Landsat data, and translate change trajectories from spectral space into biomass change trajectories. We have implemented an empirical modeling approach to interpret spectral change as biomass change. To demonstrate our approach, we compare disturbance and regrowth trajectories in two distinct regions of the U.S. We anticipate that the outcome of this study will prove beneficial to both the FIA program for tracking spatially explicit estimates of forest disturbance and regrowth, and to the NACP for quantification of the consequences of these rates for the carbon cycle.

INTRODUCTION

This paper details a piece of a much broader study that was funded by NASA Carbon Cycle Science, titled, “North American Forest Disturbance and Regrowth Since 1972: Empirical Assessment with Field Measurements and Satellite Remotely Sensed Observations.” The principal investigator on this study is Samuel Goward at the University of Maryland, and the co-investigators are Jeff Masek at the NASA Goddard Space Flight Center, Warren Cohen and Robert Kennedy from the USFS Pacific Northwest Research Station, Gretchen Moisen from the USFS Rocky Mountain Research Station, and Andy Lister and Mike Hoppus from the USFS Northeast Research Station. There are currently a number of other collaborators on this project, including researchers from the University of Maryland and researchers from the USFS Forest Inventory and Analysis.

The overall goals of this project are to satisfy objectives of both the North American Carbon Program (NACP) and the Forest Inventory and Analysis (FIA) program. For the NACP, the objectives are to integrate satellite observations with field data to reduce uncertainty in carbon cycle models by improving estimates of disturbance and regrowth processes. For FIA, the objectives are to improve their ability to quantify disturbance and regrowth processes by enhancing their satellite change-detection methods. In doing so, this project will enable FIA to focus on specific regional disturbance issues that are relevant to each of the FIA regions.

To accomplish these objectives, we have selected a study design that integrates time-series of Landsat satellite data with FIA ground data. The benefit of this approach is that we will leverage the strengths of both Landsat time-series data and FIA plot data, and enhance our ability to use FIA plot data in a more spatially explicit context. We have selected to use Landsat data because of its moderate spatial grain and ability to span moderate to large spatial extents. Previous studies of forest disturbance and recovery have been carried out using repeat aerial photography, but they are often limited to sample locations or smaller spatial extents (Kennedy and Spies, 2004). Conversely, coarse grained data such as AVHRR have been used to investigate forest disturbance and regrowth at continental spatial extents (Potter et al., 2005). The spatial grain of these data, however, precludes their ability to discern finer grained forest disturbance and regrowth processes. For these reasons, we chose to use Landsat data for its moderate resolution and regional applicability. Landsat data have been effectively used to analyze forest disturbance and regrowth processes at regional scales (Cohen et al., 2002; Lawrence and Ripple, 1999). To boost the spatial extent of our study, however, we are implementing a continental-scale sampling approach for selecting Landsat data frames. It would be beyond our capabilities to process and analyze all Landsat scenes across North America,
therefore, we opted to sample approximately 20-30 Landsat scenes that capture the full range of forest conditions. This approach will enable two distinct approaches to estimating disturbance and regrowth processes. First, design-based estimates will stem from the sample of Landsat frames representing the range of variability in forest types across the United States. Second, model-based estimates will derive from the sample locations as reference data for model building using national-level data sets.

Both forest disturbance and regrowth processes are characterized by inherent variability at multiple scales, from local to regional to continental. Disturbance regimes are often characterized in terms of frequency, extent, and magnitude (Morgan et al., 2001), while regrowth is often characterized in terms of delay, rate, and extent (Lawrence and Ripple, 1999; Yang et al., 2005). Previous research at the PNW Research Station documented that disturbance is followed by significant variability in successional regrowth patterns in Oregon’s Coastal and Cascade Range provinces (Yang et al., 2005). Yang interpreted 50-year time-series of aerial photographs to quantify changes in percent conifer cover following clear-cut logging. His results demonstrated both within and between province variability in regrowth. While this method was effective at monitoring regrowth at sample locations, our study builds upon this approach by incorporating time-series of Landsat data. Trajectories of spectral change derived from Landsat time-series can effectively track disturbance and regrowth dynamics (Figure 1), and enable characterization of these processes by linking to FIA ground data. Our challenge is to characterize disturbance and recovery processes in forest structural terms, at multiple scales, from local to regional to continental.

Within the broader context of the overall project, the specific objectives of this paper are to: 1) Explore the empirical relationships between forest structure (specifically biomass) and spectral reflectance for two distinct study sites, one located in the western U.S. and one in the eastern U.S.; and 2) Explore techniques for characterizing forest disturbance and regrowth processes at these sites.

METHODS

We analyzed data from two sample locations, specifically chosen because they represent distinctly different forest types and disturbance regimes. The first location was northern Arizona (Landsat path37/row35), containing the Grand Canyon, the Kaibab Plateau, and the San Francisco Peaks (Figure 2). The forested area within this scene is primarily coniferous, dominated by pinyon pine/juniper at lower elevations and ponderosa pine at higher elevations, and interspersed with smaller patches of deciduous aspen forest.
Figure 2. Study area location: Landsat path 37/row 35, northern Arizona.

The second location was coastal South Carolina (Landsat path 16/row 37), containing Lake Moultrie and the city of Charleston along the eastern seaboard (Figure 3). The forested area within this scene is a mixture of coniferous and deciduous, dominated by young, managed pine plantations and older oak, maple, and cypress forests.

Figure 3. Study area location: Landsat path 16/row 337, coastal South Carolina.

There were three major methodological components to this study: Landsat processing methods, FIA plot methods, and integrated empirical modeling methods. Landsat imagery was processed for each of the sample locations by the Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS) (Masek et al., 2006). For the Arizona scene, we used 1991 and 2001 images, and for the South Carolina scene, we used a biennial stack of imagery between the years 1984 and 2005. All images underwent geometric correction and absolute radiometric correction to surface reflectance. After data acquisition from LEDAPS, we performed relative radiometric normalization on each image using a single reference image and the Multivariate Alteration Detection (MAD) method (Canty et al., 2004; Schroeder et al., in press).

We used the most recent annual inventory FIA data (cycle 3 in both AZ and SC) to construct plot-level estimates of biomass. FIA annual inventory data is based on the standard 4 subplot fixed-radius plot configuration. We scaled tree-level observations of species and diameter at breast height (DBH) up to plot-level biomass (lbs/acre).
using national-level Jenkins allometric equations (Jenkins et al., 2003). For each sample scene, we isolated the population of single-condition, forested plots to ensure more homogenous reference data. At each of the FIA plot locations, we extracted the Landsat spectral values from a 55-m radius circle, matching the Landsat image year with the field measurement year. We then integrated the raw Landsat spectral data for each reference plot into a single spectral index using Canonical Correlation Analysis (CCA), which maximized the correlation between the reference data (spectral bands) and the response variable (plot-level biomass). We performed Reduced Major Axis (RMA) regression, a form of orthogonal regression, between the spectral index from CCA and biomass. RMA regression minimizes error in both the X and Y direction, and therefore better maintains the data variance structure compared to ordinary least squares (OLS) regression (Cohen et al., 2003). Based on the RMA regression equations, we spatially extrapolated the biomass models across the sample scenes, and temporally extrapolated biomass across years of the image stack. We validated each model with “hold-back” data by comparing model predictions with observed values.

There were a number of significant data quality issues that were addressed during this analysis. First, the spatial quality of both Landsat and FIA data were critical. The positional error of FIA plot locations renders working with these data in a spatially explicit context challenging. In conjunction with Landsat positional error, accurate georegistration and orthorectification among Landsat scenes was essential. Second, the radiometric quality of the Landsat data was critical for interpreting spectral trajectories. We tested the ability of relative radiometric normalization using MAD to reduce the interannual variability of surface reflectance across the image stacks. The results from sample coniferous and deciduous locations indicated reduced interannual variation in reflectance following relative radiometric normalization (Figure 4).

Third, the temporal data quality was critical for matching the appropriate Landsat image year with the FIA field measurement year. Figure 5 demonstrates the effect of a mis-match between image year and measurement year. Fourth, a “catch-all” category of spectral-physical data quality issues were critical to address. These included the effects of erroneous spectral values due to clouds, shadows, roads, water, or positional error. The solutions to these problems included focusing exclusively on single-condition forested plots, masking out undesired features such as non-forest, clouds, and shadows, and performing rigorous manual and statistical flagging of potential outliers.

**Figure 4.** Comparison of radiometric corrections for band 5 spectral trajectories: Solid lines are absolutely radiometrically corrected trajectories (pre-MAD) (red=hw=hardwood stand; blue=con=conifer stand), and dashed lines are absolutely plus relatively radiometrically corrected trajectories (MAD). Band 5 surface reflectance (on the Y-axis) was scaled between 0-10,000. Two dates of 2002 imagery were compared in this time-series (2002 and 2002b).

Third, the temporal data quality was critical for matching the appropriate Landsat image year with the FIA field measurement year. Figure 5 demonstrates the effect of a mis-match between image year and measurement year.
Figure 5. Correlation coefficients (r) between basal area and Landsat band 5 for subsets of FIA plot data corresponding to field measurement year. The Landsat band 5 data was from 2002, and showed the highest correlation with the 2002 field measurement year. Conversely, the 2001 field measurement year had the lowest correlation with the 2002 Landsat spectral data, potentially as a result of spectral-physical data quality issues.

RESULTS

For the northern Arizona study scene, we plotted the empirical relationship between biomass (transformed to square root) and the Landsat spectral index derived from CCA (Figure 6).

Figure 6. Empirical relationship between biomass (square root transformed) and the Landsat spectral index derived from CCA, for single-condition forested FIA plots in the northern Arizona (p37/r35) study scene.

These data showed a negative linear relationship, with a correlation coefficient of -0.72. The RMA regression model for these data had an R^2 of 0.52. We extrapolated this RMA regression model across the entire study scene for the year 2001 (Figure 7). From a visual standpoint, the 2001 biomass map conformed to our expectations, with higher biomass evident in the ponderosa pine forests of the Kaibab Plateau, relative to the pinyon-pine/juniper woodlands at lower elevations.
The correlation coefficient between observed and predicted biomass for the hold-back data was 0.78 and the RMSE was 5,847 lbs/acre of biomass (Figure 7). We extrapolated the biomass model to the 1991 image year and inspected known areas of disturbance between 1991 and 2001 to visually evaluate the results (Figure 8). Areas of stand-replacement forest harvest between 1991 and 2001 were appropriately labeled as lower biomass in 2001. Field reference data from an earlier date will be necessary to more fully evaluate these results.

For the South Carolina study scene, exploratory data analysis revealed distinctly different empirical relationships between biomass and Landsat spectral data for the coniferous versus deciduous FIA plots (Figure 9).
Figure 9. Empirical relationships between biomass (square root transformed) and the Landsat spectral index derived from CCA, for single-condition deciduous (hardwood) FIA plots (left) and single-condition coniferous FIA plots (right) in the South Carolina (p16r37) study scene.

For the coniferous data, there was a positive linear relationship between biomass and the spectral index, with a correlation coefficient of 0.78. For the deciduous data, the relationship was a negative linear one, with a correlation coefficient of -0.68. Because of these different relationships, we modeled the two data sets independently. For the coniferous RMA model, the $R^2$ was 0.61, and for the deciduous RMA model, the $R^2$ was 0.48. We extrapolated these RMA regression models across the entire sample scene for the year 2001 (Figure 10). Visually, the map conformed to our expectations, with higher biomass evident in the deciduous forests, especially the riparian areas.

Figure 10. 2001 Biomass map for the South Carolina study scene.

We validated each of these models with hold-back data (Figure 11). For the coniferous data, the correlation coefficient between observed and predicted biomass was 0.66 with an RMSE of 5,915 lbs/acre of biomass. For the deciduous data, the correlation coefficient was 0.27 with an RMSE of 24,823 lbs/acre of biomass.
Figure 11. Observed vs. predicted scatter-plot validations of the deciduous biomass model (left) and the coniferous biomass model (right) with hold-back data.

Further exploratory analysis of the age-structure of the deciduous data revealed much stronger empirical relationships between biomass and spectral data for young stands versus older ones, potentially indicating room for improvement over the original empirical model (Figure 12).

Figure 12. Empirical relationships between biomass (square root transformed) and the Landsat spectral index derived from CCA, for single-condition deciduous (hardwood) FIA plots < 10 years old (top-left), single-condition deciduous FIA plots < 20 years old (top-right), single-condition deciduous FIA plots > 20 years old (bottom-left), and single-condition deciduous FIA plots > 40 years old (bottom-right) in the South Carolina (p16/r37) study scene.

We spatially and temporally extrapolated the RMA regression models across all the images in the sample scene between 1984 and 2005. A visual examination of known areas of disturbance (forest harvest) revealed expected correspondence between spectral patterns and biomass patterns (figure 13).
DISCUSSION

Empirical relationships between Landsat spectral data and FIA plot-level data across the 2 study scenes support our approach for modeling biomass, and ultimately characterizing disturbance and regrowth dynamics. Many previous studies have exploited these empirical relationships to map forest structural characteristics such as biomass (Hall et al., 2006; Zheng et al., 2004). Fewer studies, however, have attempted to model changes in forest structural characteristics with time-series of satellite imagery (Dong et al., 2003; Healey et al., 2006). In the absence of data to more robustly validate biomass change, our approach is still developmental, though the ability of spectral data to track change is well accepted (Song et al., 2002; Healey et al., 2005).

In this paper, we specifically intended to contrast coniferous and deciduous forest systems, and our results suggest that at least for the 2 regions we examined, biomass is more highly correlated with Landsat spectral data in coniferous forests than deciduous forests. The biomass regression model for the Arizona study scene was the most accurate, followed closely by the South Carolina coniferous biomass model. The South Carolina deciduous biomass model was the least accurate. Despite a weaker biomass model for deciduous forests, we demonstrated the potential for exploiting a higher correlation between Landsat data and biomass data in younger, deciduous forests. This suggests a potential for at least focusing on modeling biomass trajectories in young, regrowing deciduous stands.

Temporal stacks of Landsat imagery contain a wealth of information that can aid in the characterization of disturbance and regrowth processes (Schroeder et al., in press). This information can be analyzed at multiple scales using a variety of approaches. At the scale of individual FIA plots, spectral and biomass trajectories reveal disturbance and regrowth dynamics. Traditionally, FIA plots have been used for statistical summarization at broad spatial scales (e.g. county-level). For these purposes, the actual spatial location for individual plots is irrelevant, but rather the combined properties of plot-level characteristics are integrated to draw inference about broad scale conditions. If multiple inventories have been carried out in a particular region, then inference can be drawn about trends in forest properties. For example, trends in forest disturbance have been reported at the state-level (Smith et al., 2003), or county-level (Shaw et al., 2005) but it remains unknown where precisely within the state or county changes occurred. The approach we have taken of integrating time-series of Landsat data with FIA ground data potentially addresses this issue by enabling more spatially explicit use of plot data. For example, Figure 14 depicts both a spectral and biomass trajectory between 1984 and 2005 for an FIA plot in South Carolina. The coniferous stand that contains the FIA plot was clear-cut after 1992 and regenerated until 2005. Both the spectral and biomass trajectories depict the disturbance and regrowth dynamics. At the time of the FIA field measurement in 1999, the
FIA field crew measured a stand biomass of 52,104 lbs/acre, which is consistent with the modeled biomass for that time period. As this example demonstrates, FIA plots can easily be summarized by their spectral and biomass trajectories. In the absence of multiple inventories or repeat measurements, this technique offers an opportunity to investigate trends at specific locations, including FIA plots.

**Figure 14.** Sample South Carolina FIA plot spectral (U1) and biomass (lbs/acre) trajectories between 1984 and 2005. The FIA plot is depicted by the small red circle near the center of each image.

Spectral and biomass trajectories can also be analyzed at the landscape, or regional scale. Spatial extrapolation of spectral and biomass models enable inference of biomass dynamics for any pixel within a modeling area. For a subset of the South Carolina image, we analyzed the full temporal stack of spectral data for conifer pixels to characterize disturbance and regrowth trajectories (Figure 15). By performing an unsupervised classification on the full 20-year stack of imagery (each year was represented by the spectral index derived from CCA), we classified the subset according to one of three distinct disturbance and regrowth trajectories: areas disturbed pre-1984, areas disturbed during the 1980’s, and areas disturbed during the 1990’s. Each of the trajectories was graphed using the mean spectral values for each of the time periods. The starting points of the three trajectories were clearly different in spectral space. For the areas disturbed pre-1984, the spectral index value was lower, representing areas with lower biomass. The ending points for all three trajectories were similar, however, revealing that biomass regrowth had occurred to a similar degree by the end of the time-series. As evidenced by figure 14, and the linear relationship between the spectral index and biomass, these spectral trajectories can also be portrayed as biomass trajectories. Furthermore, any number of classes can be identified using this approach. For this example, we chose only three classes for simplicity, but more detail can be derived from a larger number of classified trajectories. In another example, with more classified trajectories, we extracted time-since-disturbance, roughly equating to a stand age map (Figure 16). This level of detail is invaluable for modeling the biomass of early successional stands. For example, we demonstrated that for young deciduous stands versus older stands in South Carolina, there was a stronger correlation between biomass and spectral data (Figure 12). Therefore, by mapping stand age, we can isolate young deciduous stands and develop improved biomass models for these pixels. Maps of disturbance and regrowth can be used to analyze the extent, rates, patterns, and mechanisms of these dynamics within and across regions. It will be possible to examine specific disturbance events (e.g. fires, hurricanes, windthrow, etc…) and characterize the resulting effect on forest biomass. Comparisons can be made across study areas of the extent, location, average size, and patterns of disturbance. Following disturbance events, the relative areas of faster versus slower regrowth can be mapped, facilitating development of models to aid in the interpretation of the mechanistic drivers of regrowth variability.

This paper describes preliminary results from only two locations. It is intended to provide background on the overall project as well as demonstrate how these data might be used. We are in the process of acquiring and processing data for all of the sample scenes across North America. To date, we have developed data for 8

Eleventh Forest Service Remote Sensing Applications Conference
Salt Lake City, Utah * April 24-28, 2006
"prototype" scenes, 2 of which we have shown examples from here. We are also in the process of evaluating other approaches for biomass modeling, including nearest neighbor imputation, and other approaches for characterizing spectral and biomass trajectories.

![Conifer Spectral Trajectory Classification](image)

**Figure 15.** Classified disturbance/regrowth trajectories for a subset of the South Carolina study scene between 1984 and 2005.

![Time-Since-Disturbance](image)

**Figure 16.** Time-since disturbance map for a subset of the South Carolina study scene.

**REFERENCES**


