Quantifying Forest Ground Flora Biomass Using Proximal Sensing

Paul F. Doruska* and **Robert C. Weih Jr.**
Arkansas Forest Resources Center
University of Arkansas School of Forest Resources
P.O. Box 3468
Monticello, AR 71656

Matthew D. Lane
Landscape Analysis Laboratory
University of the South
P.O. Box 222
Sewanee, TN 37375

Don C. Bragg
USDA Forest Service Southern Research Station
P.O. Box 3516
Monticello, AR 71656

*Corresponding Author

Abstract

Current focus on forest conservation and forest sustainability has increased the level of attention given to measures of ground flora in forest ecosystems. Traditionally, such data are collected via time- and resource-intensive methods of field identification, clipping, and weighing. With increased focus on community composition and structure measures of forest ground flora, the manner in which these data are collected must change. This project uses color and color infrared digital cameras to proximally sense forest ground flora and to develop regression models to predict green and dry biomass (g/m^2) from the proximally sensed data. Traditional vegetative indices such as the Normalized Difference Vegetative Index (NDVI) and the Average Visible Reflectance Index (AVR) explained 35-45% of the variation in forest ground flora biomass. Adding individual color band variables, especially the red and near infrared bands, to the regression model allowed the model to explain 66% and 58% of the variation in green and dry biomass, respectively, present.

Introduction

Many forest research projects estimate forest ground flora biomass via the labor-intensive technique of clipping, drying and weighing vegetative samples (Brower et al., 1990). When combined with species identification, such work is used to calculate a variety of community composition and structure measures (Magurran, 1988; Reed and Mroz, 1997; Elzinga et al., 1998) used in ecosystem studies and reporting. Such information is also used when assessing wildlife habitat (National Wildlife Research Center, 2000). The need to rapidly and preferably nondestructively determine such attributes has become apparent with the increased focus conservation and sustainability.

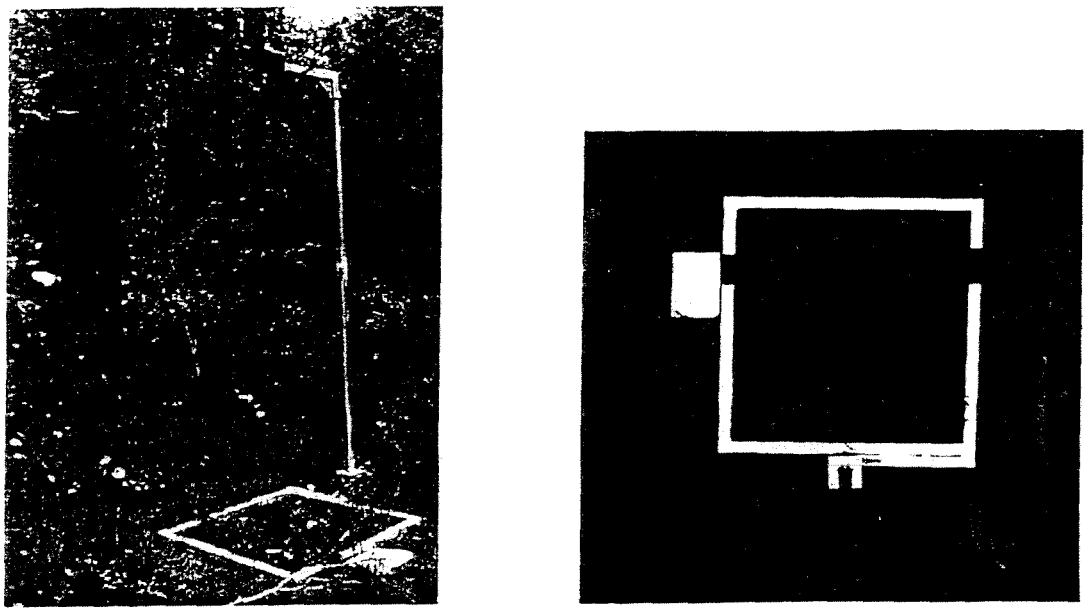
Satellite or airborne imagery combined with computer algorithms can be used to estimate forest biomass. (Baret et al., 1989; Ahern et al., 1991). However, such imagery cannot be used to estimate forest ground flora biomass for two primary reasons. First, the scale of the imagery is too coarse to adequately examine forest ground flora. Second, the presence of a forest canopy often prevents a satellite or airborne camera from capturing forest ground flora in its imagery. This project examines whether techniques used to estimate forest biomass from satellite or airborne imagery can be used to estimate forest ground flora biomass using proximally sensed data obtained from color and color infrared imagery.

Photoplots have been used in ecological research for change detection (Schwegman, 1986; Windas 1986). This project combines the idea of photoplots, utilizing both color and color infrared digital images of forest ground flora, with vegetative indices typically calculated from satellite images to estimate forest ground flora aboveground biomass in g/m^2 via regression models.

Materials and Methods

Equipment and Software.--Two digital cameras were used with this project. A Kodak DCS760 camera with a Nikon F5 body was used to take color digital images at a 6 million pixel (3038 x 2028) resolution. A Kodak DCS420CIR camera with a Nikon F90 body camera operating at a 1.5 million pixel (1524 x 1020) resolution was used to take color infrared images. 20 mm auto-focus lenses were used on both cameras and an Omega Optical band pass filter (500-900 nm) was used with the DCS420CIR camera. A GER2600 spectroradiometer was used to assist in image standardization.

An aluminum stand was constructed to mount the



(A) (B)

Fig. 1. The frame (A) used to photograph the plots and an image (B) taken with the color camera (grayscale version shown here).

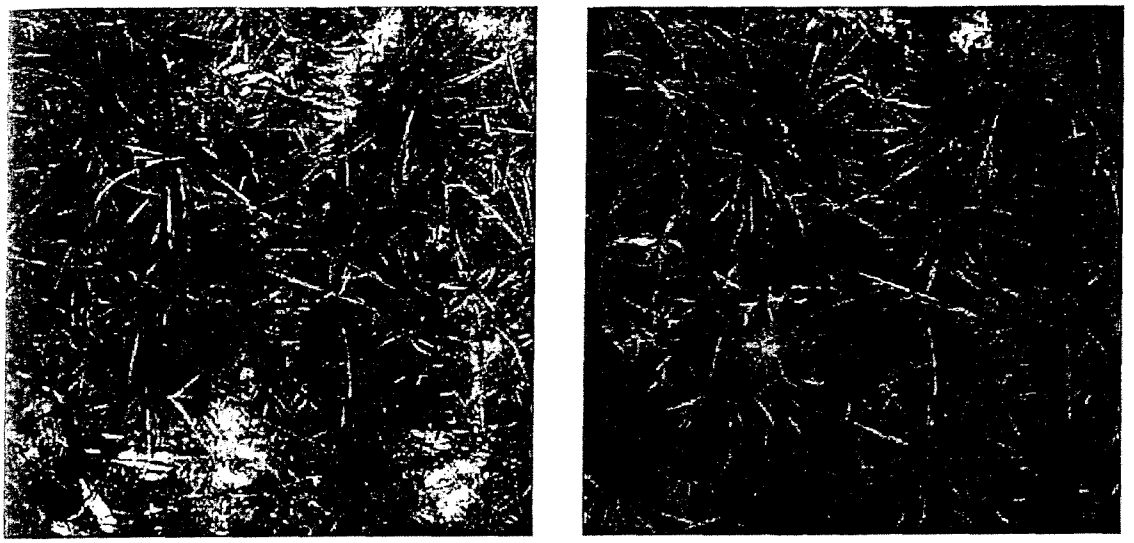
cameras and frame plots 1 m^2 in size. The actual frame size was 0.966 m^2 but hereafter it will be referred to as 1 m^2. The cell size for the imagery was 0.1015 cm. Black and white bands were painted onto the frame to calibrate image values from 0 to 255 to take into account variations in illuminations. Cross hairs (or tic marks) were drawn onto the frame in order to develop a local coordinate system for image comparisons. ESRI ArcGIS™ 8.x and ERDAS Imagine® 8.5 software were used to process the imagery and calculate the vegetative indices used herein.

Image Acquisition and Vegetation Collection.--A total of 75 1-m^2 plots was randomly established throughout the summer and early fall of 2002 in a variety of forest stands in the University Forest at the Univ. of Arkansas-Monticello, Monticello, AR (Drew County), near the Crossett Experimental Forest in Crossett, AR (Ashley County), and in established research areas in the Ouachita Mountains (Perry County). Young pine plantations, mature pine plantations, mixed pine-hardwood forests, and hardwood forests were visited. Once a plot location was established, the aluminum camera stand was set up (Fig. 1A), and any vegetation overlapping or extending beyond the border of the frame was removed to ensure only vegetation within the plot would appear in the images. For the purposes of this study, forest ground flora was defined as all vegetation less than 1 meter in height.

Each camera was then mounted to the frame separately and raised to the appropriate level to digitally capture an image of the plot. Three pictures were taken per camera to be sure at least one usable image was captured for each camera on each plot (Fig. 1B). After the images were taken, the vegetation in the plots was identified, clipped at ground level, sorted by species, placed into labeled plastic bags, and sealed for laboratory analysis. No protected species were encountered so all vegetation could be clipped for subsequent mass determination.

Mass Determination.--The green mass of the contents of each bag was determined immediately upon return to the laboratory. The contents of each plastic bag were then transferred to labeled paper sacks and placed in a drying oven at 60°C for three to four days. Upon completion of the drying process, the dry mass of the contents of each bag was determined and summed to obtain plot-level values.

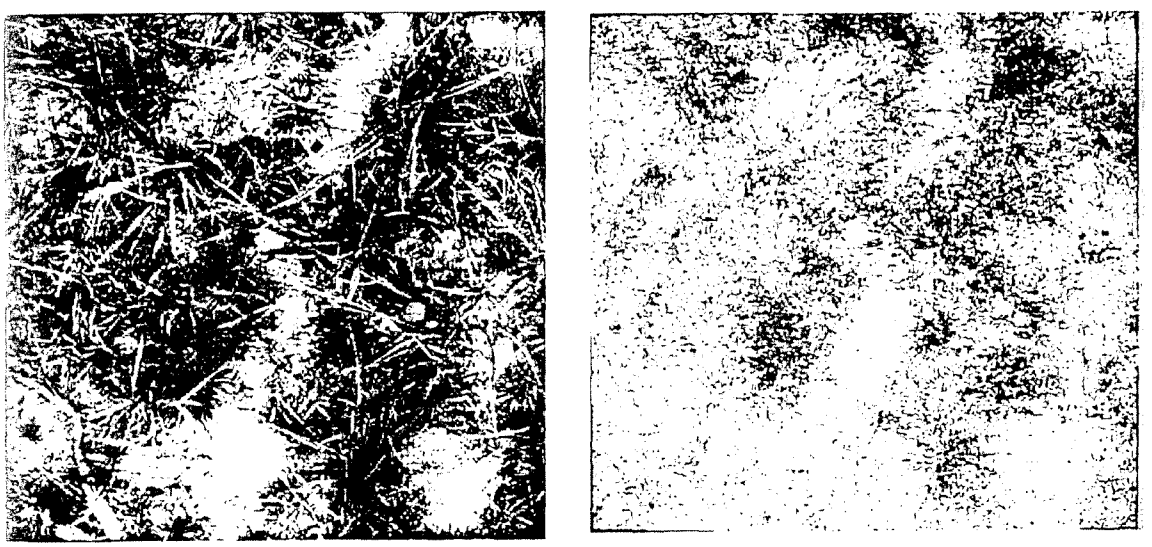
Image Registration and Standardization.--The camera stand used in this study had a set of seven tick marks on its frame. These tick marks were measured to within 0.025 cm and placed in a shapefile to represent a local coordinate system for the camera stand. Each collected image was then registered to that coordinate system within ArcGIS™ ArcMap® using the georeferencing extension. The referencing was accomplished by aligning the measured tick marks with the marks seen on the image. Doing so insured that all images would align exactly with each other and could be compared.



(A) (B)

Fig. 2. Color (A) and color infrared (B) images of an area of interest (grayscale versions shown here).

The camera stand also had black and white painted regions on it so that the illumination of images could be standardized, as the amount of solar energy incident on the plots changed. A GER2600 spectroradiometer was used to determine the reflectance of the painted regions for 4 bands (Near Infrared [NIR], Red [R], Green [G], and Blue [B]) and represented the extremes of the range of colors present in any image for any band. A simple linear regression was created per band per image to convert the range of values present within a given band/image combination to the range defined via the spectroradiometer. The regressions were used to calibrate each image. Upon completion of the



(A) (B)

Fig. 3. AVR (A) and NDVI (B) images created for the area of interest shown in Fig. 2 (a grayscale version of the NDVI image is shown here).

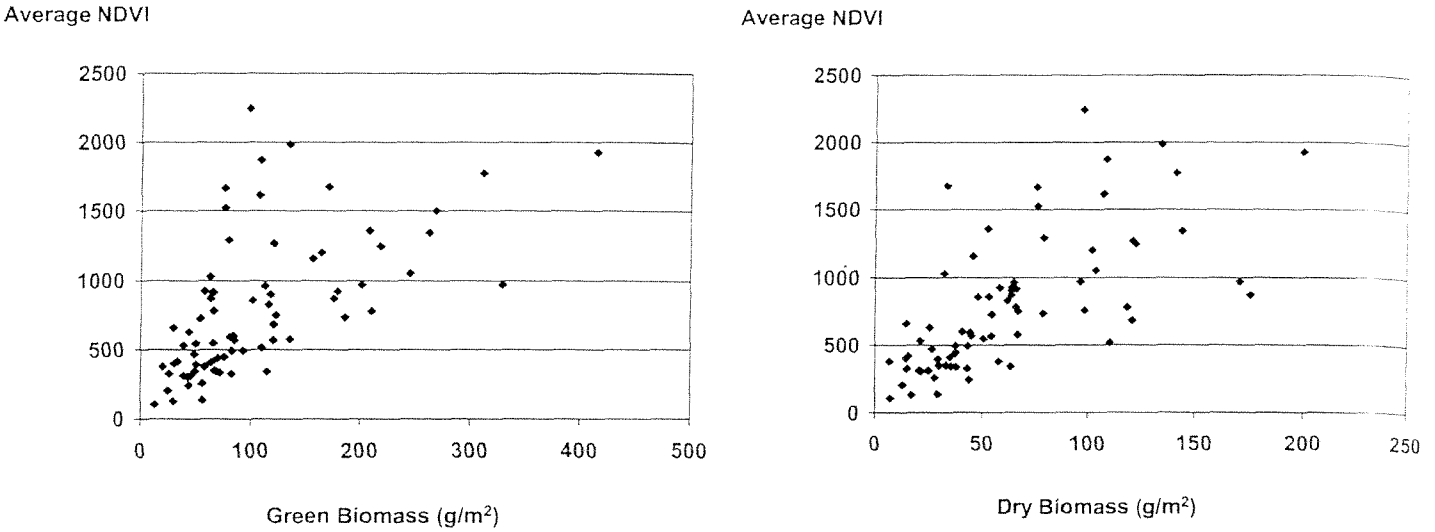


Fig. 4. Average NDVI values from each of the 75 forest ground flora plots versus (A) green biomass and (B) dry biomass in g/m^2.

calibration, the digital values were standardized and represented the same colors from one image to the next.

Images were then cropped to areas of interest (AOI's) created manually in Imagine®. The AOI's contained only that portion of each image that was inside the frame of the camera stand, and only that was used for all subsequent analyses (Fig. 2). When applying the regression models to the areas of interest for each respective band/image combination, any values in the output grid that were less than 0 were reset to equal 0 (negative values can disrupt calculation of certain vegetation indices). A vegetation mask was applied to better distinguish vegetation from non-vegetation in the images. The final output grid was a 6-band image consisting of standardized NIR, red and green bands from the color infrared image, and the red, green, and blue bands from the color image.

Vegetation Indices and Regression Modeling.--The CIR camera images were used to calculate the Normalized Difference Vegetation Index(NDVI), and the color camera images were used to calculate the Average Visible Reflectance(AVR) index for each pixel in each image.

$$NDVI = (NIR-R)/(NIR+R) \quad (1)$$
$$AVR = (G+R+B)/3 \quad (2)$$

The output image from this step was a two band grid (NDVI and AVR) with a cell size of 0.1015 cm by 0.1015 cm (Fig. 3). Once the NDVI and AVR values for the images were calculated, they were summed and averaged for use as potential independent variables in regression equations to predict green or dry biomass. Typical regression diagnostics (Myers, 1990) were used to construct models to predict green and dry biomass (g/m^2) of the forest ground flora from independent variables derived from the imagery.

Results

The green biomass data collected from the plots averaged 106.06 g/m^2, possessed a standard deviation of 77.95 g/m^2, and ranged from 12.90 to 413.60 g/m^2. The corresponding dry biomass observations averaged 53.68 g/m^2, possessed a standard deviation of 39.82 g/m^2, and ranged from 6.90 to 199.50 g/m^2. The green and dry biomass observations served as the dependent variables in the regression models constructed. The scatter plots of average NDVI versus green and dry biomass are depicted in Fig. 4, while Fig. 5 depicts the relationship between average AVR and green and dry biomass. Initial regression models used average NDVI and average AVR in an attempt to predict observed biomass.

Sole use of either of these averaged vegetative indices as independent variables did not result in very strong simple linear regressions, as the R^2 from such models were in the 25% to 35% range. Even after the non constant variance issues present in Figs. 4 and 5 were addressed by using a natural log transformation of the respective dependent variables (the biomass measures), the R^2, or the proportion of variation in biomass explained by the model (the respective averaged vegetative indices) were well under 50%. Clearly, alternate model forms or use of additional

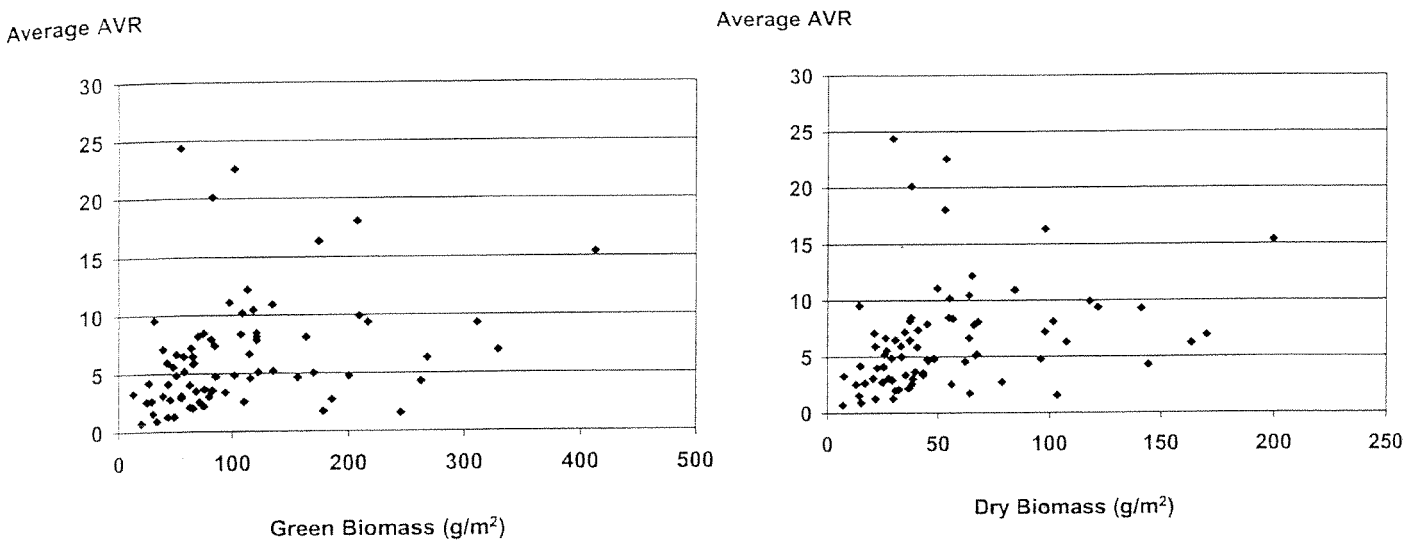


Fig. 5. Average AVR values from each of the 75 forest ground flora plots versus (A) green biomass and (B) dry biomass in g/m^2.

independent variables were needed to improve model performance.

Several nonlinear model forms (Sit and Poulin-Costello, 1994) were examined in order to improve the fit of the regression model. None of the nonlinear forms, which used just average NDVI or average AVR as independent variables, outperformed the simple linear regression models already developed.

In addition to the average NDVI value and average AVR value for each image, additional attributes were available for use as potential independent variables. The following attributes were available for each image taken using the color infrared camera: average of the NIR band, average of the G band, and average of the B band. Similarly, averages of the R, G, and B bands were available for each image taken with the color camera. These attributes were used in conjunction with the average NDVI and AVR attributes previously described to build a regression model via a stepwise regression procedure.

Upon successful completion of the stepwise regression procedure, the following linear regression model performed best with respect to estimating either green or dry biomass.

$$\hat{Biomass}_i = b_0 + b_1(A\_NDVI_i) + b_2(A\_AVR_i) + b_3(A\_NIR420_i) + b_4(A\_G420_i) + b_5(A\_R760_i) \quad (3)$$

where $\hat{Biomass}_i$ = predicted biomass, green or dry (g/m^2) for plot i,

A_NDVI_i = average NDVI value for plot i,

A_AVR_i = average AVR value for plot i,

A_NIR420_i = average NIR band value from the color infrared image for plot i,

A_G420_i = average G band value from the color infrared image for plot i,

A_R760_i = average R band value from the color image for plot i, and

$b_0, b_1, b_2, b_3, b_4, b_5$ = parameters to be estimated.

Fit statistics for equation (3), when fit to the green and dry biomass data, respectively, can be found in Table 1. The standard error of the estimate was 48.00 g/m^2, and R^2 equaled 66% for the regression fit to the green biomass data, whereas those statistics for the dry biomass regression fit were 25.50 g/m^2 and 58%, respectively. All parameter estimates are significantly different than 0 at the 0.05 α-level. Clearly, both of these model fits perform considerably better than the models using just NDVI or AVR as the sole independent variable.

Variance Inflation Factors (VIFs) were used to assess the potential presence of multicollinearity between the independent variables in the fitted models. Since none of the VIFs in either model exceeded 10, multicollinearity issues are not present in either model (Myers, 1990). The presence of outliers and influential data points were addressed by examining studentized residuals, difference in fits (DIFFITs) values, and difference in betas (DFBETAs) values (Myers, 1990), and no outliers or influential points were detected. Therefore, the parameterizations of equation

Table 1. Fit statistics for equation (3) when fit to the green and dry biomass data, respectively.

Dependent Variable	Parameter	Estimate	Standard Error	p-value
Green Biomass	b_0	97.8778	30.0038	0.0017
	b_1	0.1338	0.0182	<0.0001
	b_2	4.9456	1.5056	0.0016
	b_3	-2.5581	0.7175	0.0007
	b_4	2.4904	1.1584	0.0351
	b_5	-2.0427	0.3893	<0.0001
Dry Biomass	b_0	48.5839	16.2102	0.0038
	b_1	0.0627	0.0099	<0.0001
	b_2	2.4722	0.8135	0.0034
	b_3	-1.3929	0.3877	0.0006
	b_4	1.2709	0.6258	0.0461
	b_5	-0.7762	0.2104	0.0004

(3) shown in Table (1) are the final regression models examined for these data.

Discussion and Conclusions

The ability to quantify forest ground flora attributes, such as the amount of biomass present is becoming increasingly important (Reed and Mroz, 1997). Such attributes are commonly estimated using the time- and labor-intensive method of field identification, clipping, drying, and mass determination in the lab. That technique poses three problems as the intensity of this type of research/reporting increases: the nature of destructive sampling, the amount of time involved to collect such data, and the necessity of field identification.

In certain areas destructive sampling is acceptable; in others it is not, such as in ecologically-sensitive areas or areas where vegetation is sparse (thus impact of removal would be great). This project attempts to address this issue. By using proximally-sensed data of subsequently clipped forest ground flora and building regression models to predict biomass from image characteristics, this work is setting the stage for future research and improvements in this data collection arena. While R^2 of 58% - 66% may at first seem a bit low, the authors are quite pleased with these results given the truly experimental nature of this work. The ability to explain variation in forest ground flora biomass, even to the degrees achieved herein, via a regression equation that requires no destructive sampling to apply is a very positive development. It should be noted that plots with much overlapping foliage were problematic because the images acquired only captured the top level of foliage. Future improvements in the data collection procedure (such as employing plots smaller than 1 m^2 in size) could positively impact results by reducing the amount of overlap that may be present in a given plot. More total plots would need to be employed to offset a reduction in plot size in any such sampling scenario.

The ability to better estimate green biomass than dry biomass as evidenced by the R^2 values is at first a somewhat surprising result because many biomass studies do far better at estimating dry biomass than green biomass. In this study, the vast majority of the forest ground flora sampled was herbaceous. Reflectance in the NIR band is particularly sensitive to moisture content fluctuations in the complex internal structure of foliage (Swain and Davis, 1978). Therefore, use of the NIR band may have led to better green biomass estimation than dry biomass estimation

because biomass in the green condition was captured in the images.

The time necessary to sample a plot is dramatically reduced if just images are taken and no vegetation is actually clipped. While times were not directly studied in this endeavor, a time versus accuracy comparative study has been suggested as an interesting follow-up project (Dr. Jim Guldin, USDA Forest Service, pers. comm.).

One aspect of community structure and composition not fully addressed by this project is species identification. Identification of the three most dominant species on a plot or determining the cover type (graminoid versus broadleaf forb versus shrub) of the plot is used to generate species-abundance or cover-type abundance curves. Species identification combined with biomass observations can be used to calculate similarity indices (Reed and Mroz, 1997). However, this project is providing valuable information that might be used to address this issue over time.

Over the course of this project, nearly 120 forest ground flora species were identified and digitally captured. Spectral attributes of some of these species have been identified. As work continues in this arena and more observations are taken throughout the entire growing season, computer algorithms and perhaps other equipment will be used in an attempt to better identify species thus addressing the third issue with collecting these data: species identification. With the traditional method of data collection, a botanist, or someone well-versed in forest ground flora species identification, must participate in the field identification process. In the future, perhaps some, but most likely not all, of this identification could be done from the images as well.

As interest in and reporting of herbaceous plant community and structure measures in forested ecosystems continues to increase, the need to more rapidly and nondestructively quantify such measures is apparent. This project has attempted to set the stage for using proximal sensing as the technique to satisfy that need, and has suggested means to pursue to improve model performance in the future.

ACKNOWLEDGMENTS.—The authors would like to thank the National Science Foundation Biodiversity and Ecosystem Informatics program for primary financial support of this project (EIA/0131801). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. The Arkansas Forest Resources Center is also hereby acknowledged for providing additional financial support. The assistance of Dr. Michael Shelton when identifying available study areas is also greatly appreciated. This manuscript is approved for publication by the Director of the Arkansas Agricultural Experiment Station.

Literature Cited

Ahern, F. J., T. Erdle, D. A. McLean, and **I. D. Kneppeck.** 1991. A quantitative relationship between forest growth rates and thematic mapper reflectance measurements. Int. J. Rem. Sens. 13:699-714.

Baret, F., G. Guyot, and **D. J. Major.** 1989. TSAVI: A vegetation index which minimizes soil brightness effects on LAI and APAR estimation. Pp. 1355 1358 *In:* 12th Canadian symposium on remote sensing and IGARSS'90, Vancouver, Canada.

Brower, J. E., J. H. Zar, and **C. N. von Ende.** 1990. Field and Laboratory Methods for General Ecology. William C. Brown Publishers, Dubuque, IA. 237 pp.

Elzinga, C. L., D. W. Salzer, and **J. W. Willoughby.** 1998. Measuring and Monitoring Plant Populations. Bureau of Land Management, US Government Printing Office, Washington, D.C. 477 pp.

Magurran, A. E. 1988. Ecological Diversity and its Measurement. Princeton University Press, Princeton, NJ. 179 pp.

National Wildlife Research Center. 2000. web; http://www.nwrc.gov/wdb/pub/hsi/hsiindex.html Accessed 2002 December 1.

Myers, R. H. 1990. Classical and Modern Regression with Applications, 2nd Edition. PWS-Kent Publishing, Boston, MA. 488 pp.

Reed D. D., and **G. D. Mroz.** 1997. Resource Assessment in Forested Landscapes. John Wiley & Sons, Inc., NY. 386 pp.

Schwegman, J. 1986. Two types of plots for monitoring individual herbaceous plants over time. Natural Areas J. 6:64-66.

Sit, V., and **M. Poulin-Costello.** 1994. Catalog of Curves for Curve Fitting. Biometrics Information Handbook No. 4. Province of British Columbia Ministry of Forests, Victoria B.C. 110 pp.

Swain, P. H., and **S. M. Davis** (eds). 1978. Remote Sensing: the Quantitative Approach. McGraw-Hill, Inc., NY. 396 pp.

Windas, J. L. 1986. Photo-quadrat and compass mapping tool. Natural Areas J. 6:66-67.