Concepts for Sensor Data Fusion to Detect Vegetation Stress and Implications on Ecosystem Health Following Hurricane Katrina

by Sam S. Jackson, George T. Raber, Jerry A. Griffith, and Mark R. Graves

PURPOSE: Forest ecosystems, in particular forest wetlands, are very dynamic and offer many ecological benefits because of their complex floral and faunal assemblages. It is important to understand these interactions, thus improving the ability to sustain this precious resource, and as stewards, pass it on. In addition, response to various natural influences, such as severe weather events, is also a vital part of understanding ecosystem health. It is important to quantify not only the obvious, visible damage but also the ambiguous stress these systems have undergone as a result of sustained wind damage. Satellite and airborne-based remote sensing (particularly imagery) are well-established methods for monitoring and assessing large-scale forest damage and are currently used to quantify visible damage. This research establishes proof of concept techniques for fusing sensor data from multiple remote sensing platforms to better understand the requirements needed to characterize subtle damage to forest environments impacted by hurricanes, in this case Hurricane Katrina. These advanced techniques may provide an indication of such vegetation stress before becoming visibly detectable, thus essentially predicting stress-induced mortality before it occurs. This information can be used in formulating mitigation practices in riparian areas and along streams to help reduce sediment intake due to erosion from loss of vegetation, thus improving water quality.

BACKGROUND: Apart from the direct impact of tree blow-down, wind stress to foliage is a more subtle form of damage that could make forested vegetation more vulnerable to disease or pest infestation. Extensive defoliation during periods of drought or other stressful conditions could amplify forest stress and lead to mortality, especially in coastal areas prone to repeated hurricane landfall. One form of remote sensing, hyperspectral imaging (HSI), uses instruments to collect data in hundreds of narrow wavelength bands and is specifically used for extracting much more detailed information than multispectral imaging. In fact, HSI has received increasing attention over the past decade as a tool to assess vegetation condition and mineralogical composition. Because of the very fine spectral, spatial, and radiometric resolution, these sensors extend the capability of remote sensing in a variety of applications.

Similarly, light detection and ranging (LiDAR) has also been used recently to extend the capabilities of remote sensing. LiDAR remote sensing is rapidly becoming the standard method for acquiring digital terrain information and producing digital elevation models. While this activity represents the majority of the application of airborne LiDAR data, a significant amount of research and commercial activity is also focused on deriving both qualitative and quantitative land cover information from the data.

In addition to airborne LiDAR systems, ground-based LiDAR (or terrestrial LiDAR), is another form of LiDAR imaging that can quantify structural characteristics from a horizontal perspective, which will greatly complement the vertical aspect of a forest canopy derived from the airborne LiDAR. Ground-based LiDAR will capture many portions of the forest sub-canopy structure (from a three-dimensional, horizontal perspective) at a much higher resolution and detail than airborne LiDAR. This characterization will accurately quantify the level of damage, although at a much smaller scale than airborne LiDAR, and provide the ability to correlate these quantifying metrics with early stress indicators captured by the hyperspectral imagery. This technology and ground-based LiDAR specifications are detailed under Experiment 3 on page 11.

Remote sensing approaches can be categorized by the characteristics of the sensor, and with respect to bio-physical applications, typically focuses on forest canopy cover, leaf area index (LAI), and fraction of photosynthetically active radiation (fPAR). LiDAR data allow for the quantification of tree crown dimension for habitat and biomass assessments in forest environments. Hyperspectral imaging and imaging spectroscopy provide information on structural, biochemical, and physiological properties of canopies including fractional material cover, nutrient concentration, pigment expression, and fPAR (Asner et al. 2005). This technical note reviews the hyperspectral and LiDAR technologies and examines the concept for combining these data sources for bio-physical condition assessment. It is anticipated that a better understanding of the data fusion concept for this utility will promote the technology for the stated purpose.

STUDY SITE: The principal study area is located in the DeSoto National Forest near Hattiesburg, Mississippi (Figure 1) and includes smaller sampling areas within the northern part of the forest. This inland site was readily accessible and appeared to be a suitable area to assess ambiguous stress with moderate to slightly damaged areas and limited severely damaged areas.

METHODS AND ANALYSIS: The original scope of work for this effort included the acquisition of hyperspectral imagery (Hyperion), airborne LiDAR, and selected locations of ground-based LiDAR. Individual statistics and band ratios were to be calculated on both the airborne LiDAR

Figure 1. Study area within the DeSoto National Forest near Hattiesburg, MS. Airborne LiDAR data (proxy) referenced later in the report are restricted to the coastal areas and are not represented on this map.

and Hyperion datasets. Various limitations restricted the collection of airborne LiDAR data that were spatially coincident with the ground-based LiDAR; therefore a proxy airborne LiDAR dataset south of the area of interest was used. Ideally, spatially coincident datasets could be utilized to grow decision tree models similar to those used by Ducic et al. (2006). To develop the experimental proof of concept and apply the methodology to the area of interest, other available resources had to be utilized. Three different experiments were conducted to demonstrate the potential for utilizing these datasets and develop the methodology for the stated purpose. Each experiment is summarized below.

Experiment 1. Validate the use of hyperspectral data for discriminating various areas of vegetation stress in the study area post Katrina.

Experiment 2. Evaluate the use of derived LiDAR statistics for identifying stressed areas using post-Katrina LiDAR data from Harrison County, Mississippi, south of the desired study area.

Experiment 3. Classify three ground-based LiDAR datasets using the classification system derived from experiment 2 and assess the effectiveness of characterizing impacted areas with high-density laser mapping technology.

Experiment 1: A goal of this study was to determine relationships between forest structure and vegetation indices. If significant relationships are found, this may provide a foundation to use the hyperspectral data alone to estimate forest condition. Examples of hyperspectral vegetation indices utilized in this study are found in Table 1 below.

Table 1 A Selection of Important Hyperspectral Vegetation Indices Used in this Study.	
Index	**Description**
NDVI	Narrow-band Normalized Difference Vegetation Index (can check all possible two-band combinations, and determine best band combinations through correlation plots); Relates to fPAR and LAI.
NDWI	Normalized difference Water Index. Relates to canopy water content.
PRI	Photochemical Reflectance Index; relates to light use efficiency and photosynthetic capacity.
CRI	Carotenoid Reflectance Index; indicator of plant stress and adaptation to resource conditions.
Red-edge	Examines indices on the derivative curve of the red-edge region of the EM spectrum.

Three adjacent Hyperion scenes were acquired in the fall of 2006 over a portion of the principle study area in the DeSoto National Forest. An archived post-Katrina Hyperion scene acquired immediately after the storm, which overlapped one of the new acquisitions, was obtained. This scene was analyzed in a similar method but unfortunately the data suffered from significantly greater atmospheric effects and could not be reliably utilized in the study.

Study sites (n = 17) were selected using a combination of high-resolution aerial photography and field assessment. Sites were chosen based on the following criteria: 1) They were within, or in close proximity to, the DeSoto National Forest, and 2) they could be assigned to one of the three

following categories, with respect to hurricane impact. These categories relate to increasing levels of damage.

1. Slightly affected
2. Moderately affected
3. Severely affected

Slightly affected areas are characterized by very little physical damage and exhibit only minor signs of impact from Katrina. Moderately affected areas are characterized by isolated incidences of whole-tree blowdown and perhaps a few leaning trees with slight canopy damage. Severely affected areas exhibit very apparent damage that is obvious to the untrained observer. It was expected that discrimination between the first two categories and the latter category would be possible in almost any dataset. The purpose of this experiment was to determine if it is possible to differentiate slightly impacted areas from moderately impacted areas using the hyperspectral dataset, since this discrimination was not always apparent to a standard observer and was very hard to extract from multispectral data.

Each of the sites was delineated and stored as an ESRI polygon shapefile. In addition to calculating the average spectral signature for each class, a number of band indices were calculated for each site. The band indices were:

- NDVI (Hyperion bands 45 & 33) (Figure 2)
- NDWI (Hyperion bands 51 & 109)
- PRI (Hyperion bands 18 & 22)
- CRI (Hyperion bands 16 & 20)
- Red-Edge position index (ratio of Hyperion bands 39 & 37)

These band indices were chosen because of their established success in amplifying subtle changes in vegetation characteristics.

An Analysis of Variance (ANOVA) test was performed to evaluate the effectiveness of each band index in separating the classes from each other. A box plot of each band index is displayed in Figure 3 along with the significance level of the ANOVA test. Those significant at the alpha 0.05 level are marked with an asterisk. The damage classes are labeled in plots 1 to 3, low to high.

Figure 2. An NDVI image of a portion of the study area.

Two of the five band indices are significant, indicating that the band index successfully discriminates between categories for these groups. NDVI and NDWI were very close to achieving a significant result, and were still particularly good at separating two of the three damage categories. A discriminant analysis was conducted to determine if a combination of the indices would be effective in completely separating the categories from each other in feature space.
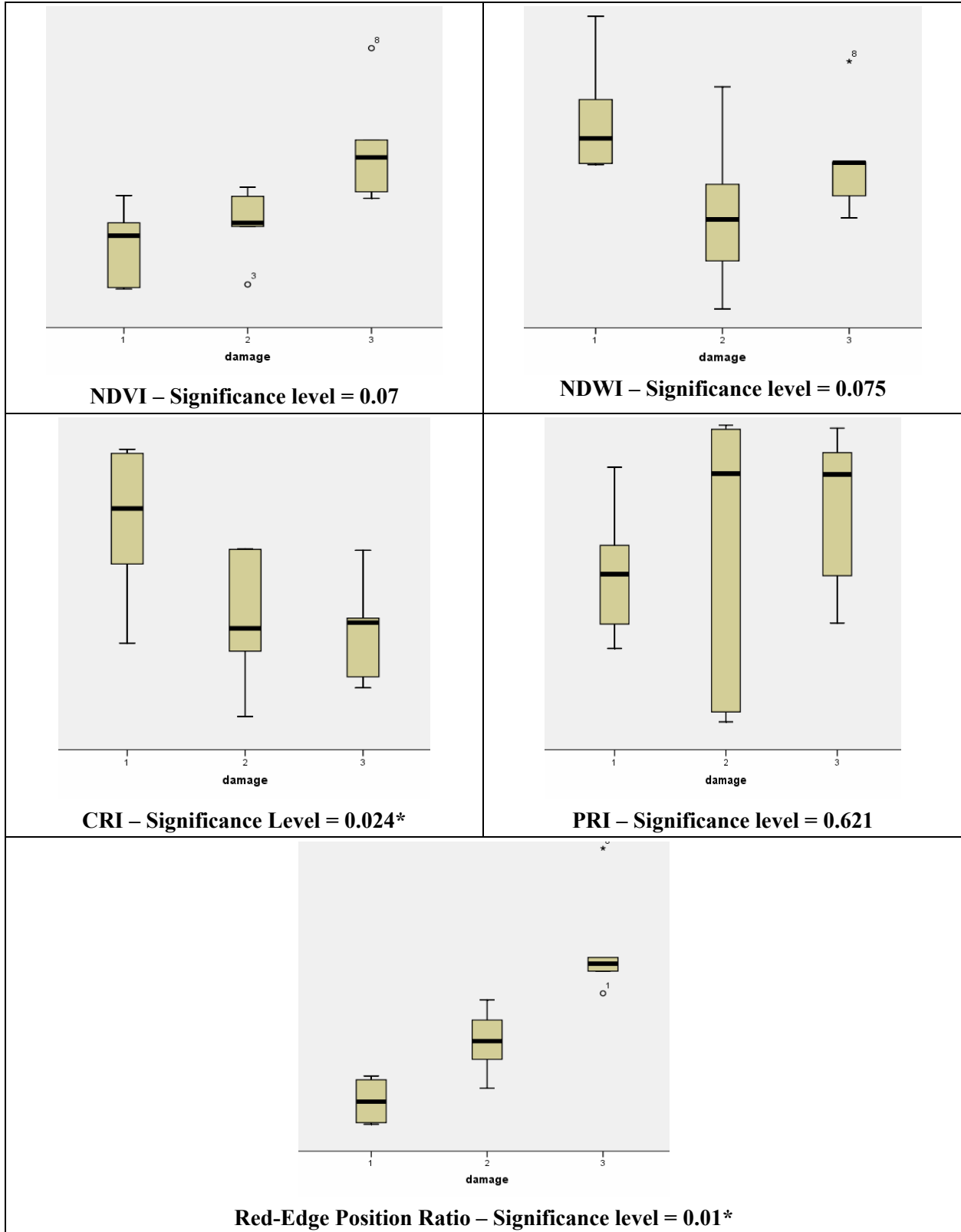
Figure 3.   Box plots demonstrating the ability of various hyperspectral indices to separate types of vegetation stress.

This test indicates the degree of separation present in the potential classes. It is similar in nature to running a reverse classification. Only those indices that were significant in the first series of ANOVA tests (CRI and Red-Edge Position Ratio) were included in the discriminant analysis function.

The results of the discriminant analysis indicate that the first derived function was significant and successfully discriminated between the three damage categories. The weights for this function indicate that the red-edge position index contributed more to the solution than CRI. The first two functions are plotted in Figure 4. The coefficients of the first function are -0.661 for CRI and 0.999 for red-edge position index. The correlation values of the first function to each of the input variables are -0.332 for CRI and 0.781 for the red-edge position index.
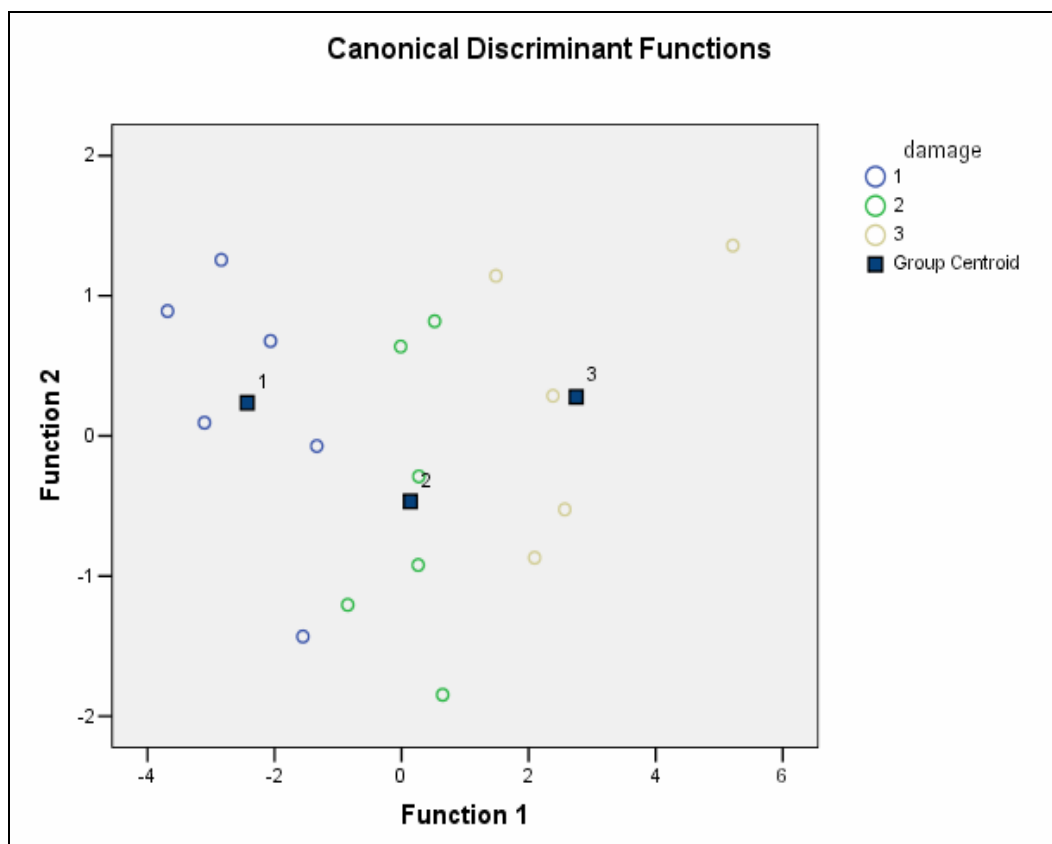

Figure 4.   Plot of the first two functions of the canonical discriminant analysis.

Figure 5 is a representation of the spectral signature of each category depicting only those bands that were utilized in the calculation of the indices (with the exception of PRI). There are a few items of interest. First, the degree of separation is significantly smaller between categories in the visible portion of the electromagnetic spectrum than the infrared portion. Second, the degree of separation between slightly affected areas and moderately affected areas in the visible portion of the spectrum is least apparent in the green band. This appears to indicate that even after one full growing season following Katrina there are differences in vegetation water content (near infrared differences) and chlorophyll absorption (red and blue differences) between the categories, despite these changes being subtle and not as visually apparent in the field.
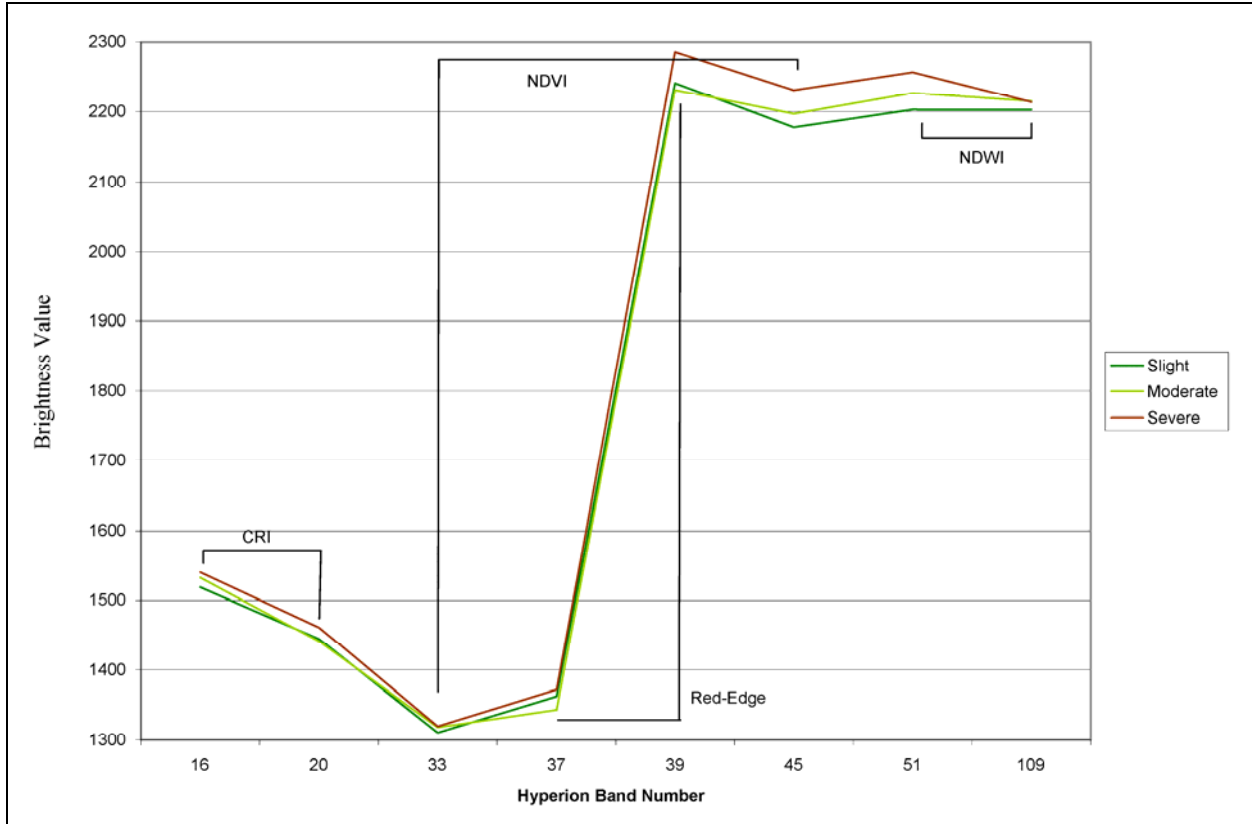
6

Figure 5.   A plot of selected hyperspectral bands with the indices shown.

Experiment 2: This experiment utilized airborne LiDAR data collected about one month after Hurricane Katrina and south of the desired area of interest. Unfortunately, the data are not spatially coincident with the ground-based LiDAR data and were not collected during the same time frame; therefore fusion of these two data types was not possible. These data were collected using a traditional small-footprint airborne LiDAR system. A unique method was utilized, which was developed to identify vegetation characteristics based on information derived exclusively from airborne LiDAR data. Some of the statistics calculated on the airborne LiDAR data seek to mimic or reconstruct a sampling of what a full waveform large-footprint LiDAR system might create over the same area, which is more ideal for a comprehensive vegetation condition assessment. In addition to evaluating the airborne LiDAR dataset alone, a portion of the derived LiDAR statistics were fused with spectral data in an attempt to demonstrate the sensor fusion proof of concept for utilization of forest vegetation condition assessment. However, suitable post-Katrina Hyperion data were not spatially coincident with the airborne LiDAR data; therefore a Landsat multispectral satellite image was used to demonstrate feasibility.

The study area for this experiment is located in western Hancock County, northeast of Diamond-head, MS. Approximately 40 sites were selected from a high-resolution photograph. These sites were classified into areas of light stress, moderate stress, and damaged areas. Each site was subsequently field verified. The categories and methodology are similar to those described in Experiment 1.

The methods in this experiment consist of constructing a simulated waveform to extract vegetation information from the airborne LiDAR data. Small-footprint LiDAR data collected in forested environments are often characterized by point distributions throughout the top or main canopy, mid-canopy, understory vegetation, and finally the terrain surface (Figure 6a). Rendering these points graphically yields a LiDAR "point cloud" that depicts the vertical structure from forest canopy to forest floor (Figure 6b). Integrating the point cloud over a geographic area yields a histogram that reveals the frequency of laser returns along the *z*-axis (Figure 6c). This histogram is similar to a waveform because it changes based on the vertical vegetation structure present within each geographic area. Extracting vegetation information using only small-footprint airborne LiDAR data is based on the assumption that a derived vertical histogram changes with point distribution and is dictated by the vegetation characteristics. For example, a homogeneous forest and the terrain surface (Figure 6a) are discernible as the two modes of the distribution in the vertical histogram. When mixed forests of uneven age are measured, however, the bimodal histogram tends to be dampened when compared to homogeneous forest stands. To characterize these distributions over small areas, a software program was written to analyze and display statistics generated from the LiDAR data. The program calculates aggregate statistics based on the vertical distribution of intercepted points within a specified polygon and stores the value in a shapefile attribute table. Initially, only four distribution descriptors were calculated, thus creating four additional attributes. These descriptors were mean, standard deviation, range, and skew. The results were only moderately successful.

To improve the LiDAR-based method for characterizing vegetation, the program was modified to describe the entire vertical distribution using the following approach. One point was
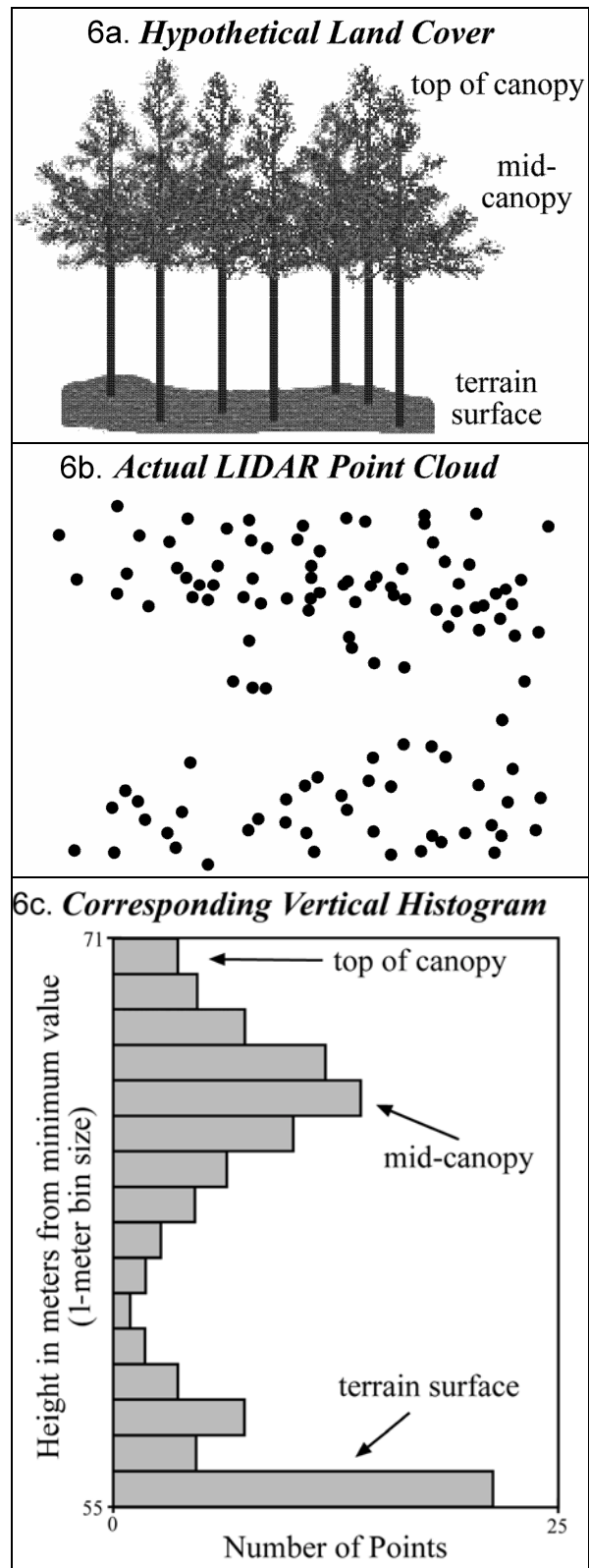

Figure 6. Histogram development representing the vertical distribution of LiDAR points.

sampled at each 10th percentile of height from the minimum value. The z-value of each of these points above the minimum point (relative z-value) became the value for 10 new attributes in the output shapefile. The mathematical explanation of the process is based on Equation 1 below:

$$V_k = P_{k*10} (Z_{poly})$$
(1)

where:

V_k = value of one single attribute in the output shapefile data table
P_k = mathematical notation for a percentile
Z_{poly} = set of relative height values for the LiDAR points within a polygon of interest

The output is a cumulative version of the vertical histogram displayed in Figure 6c. This technique was applied on the shapefile representing the damaged areas and yielded an additional 10 attributes. These 10 attributes were then normalized using two separate techniques. The first normalization simply subtracted each percentile value by the height. The second normalization repeated this and then divided the number by the range and multiplied by 100. Thus the first normalization yielded numbers from 0 to the maximum height, and the second yielded a number between 0 and 100. Figures 7 and 8 depict the curves associated with the three damage types for both of these normalized LiDAR statistics. In effect, these curves represent a LiDAR-based signature for these damage classes.

Note the differences in the curves in Figures 7 and 8. In the first normalization (Figure 7), the slight damage curve begins to taper off (at the 8th percentile at about 11m) and does not increase at the same consistent rate. This may indicate a healthier, more intact canopy structure, rather than bare or isolated branches sticking up above the main canopy. In the second normalization (Figure 8), the slight and moderate categories appear very close together, with only slight separation apparent in the lower percentiles. These differences could be explained by additional ground clutter, although the differences appear minor and are consistent with visual observations.

To experiment with the fusion of airborne LiDAR data with spectral information, a Landsat dataset was used instead of the desired Hyperion data. The fusion was performed by conducting an automated decision tree analysis called the See5 routine. For an initial proof of concept study, the structure of the decision tree rules is often more interesting than the accuracy of the classification produced. The See5 algorithm is the commercial version of a software algorithm originally pioneered by J. Ross Quinlan (Quinlan 1993), and has been used in many multi-sensor projects.

The decision tree that was generated by the algorithm for this study has only one branch with seven rules. The rules based on "L" variables are derived from the spectral data and the remaining three rules are based on LiDAR-derived statistical information. The P10 variable is the first variable in the first normalization LiDAR statistic and the S50 variable is the middle variable in the second normalization LiDAR statistic. Interestingly, there appear to be greater differences between these particular values in Figures 7 and 8.
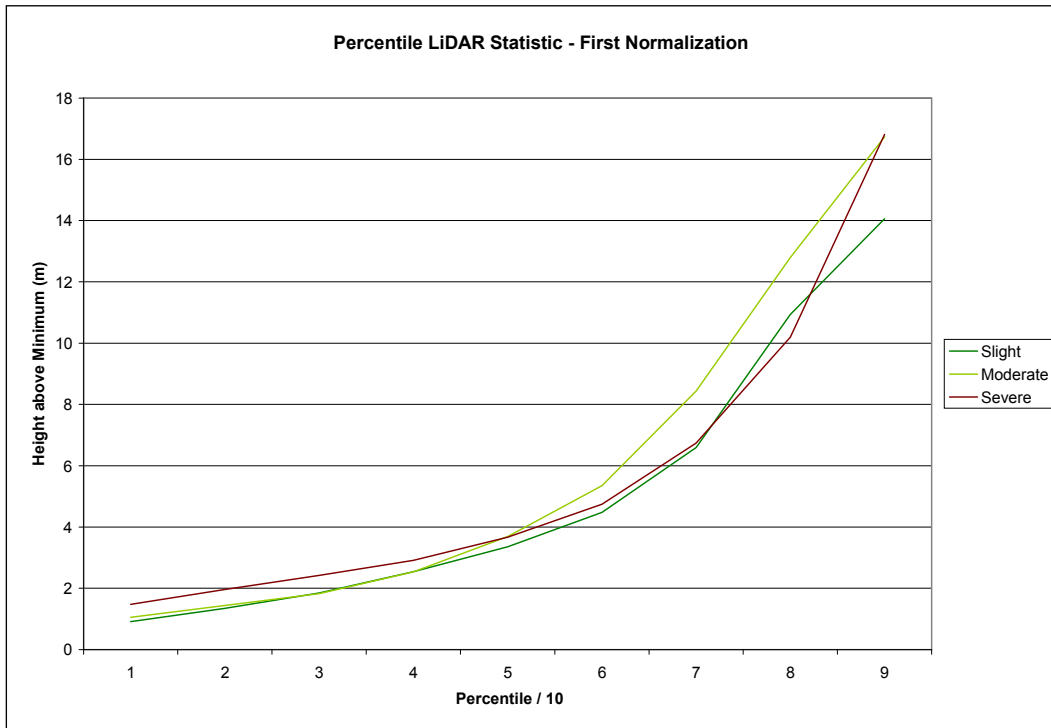
Percentile LiDAR Statistic - First Normalization

Figure 7.   Percentile LiDAR Statistic (First Normalization).

Percentile LiDAR Statistic - Second Normalization
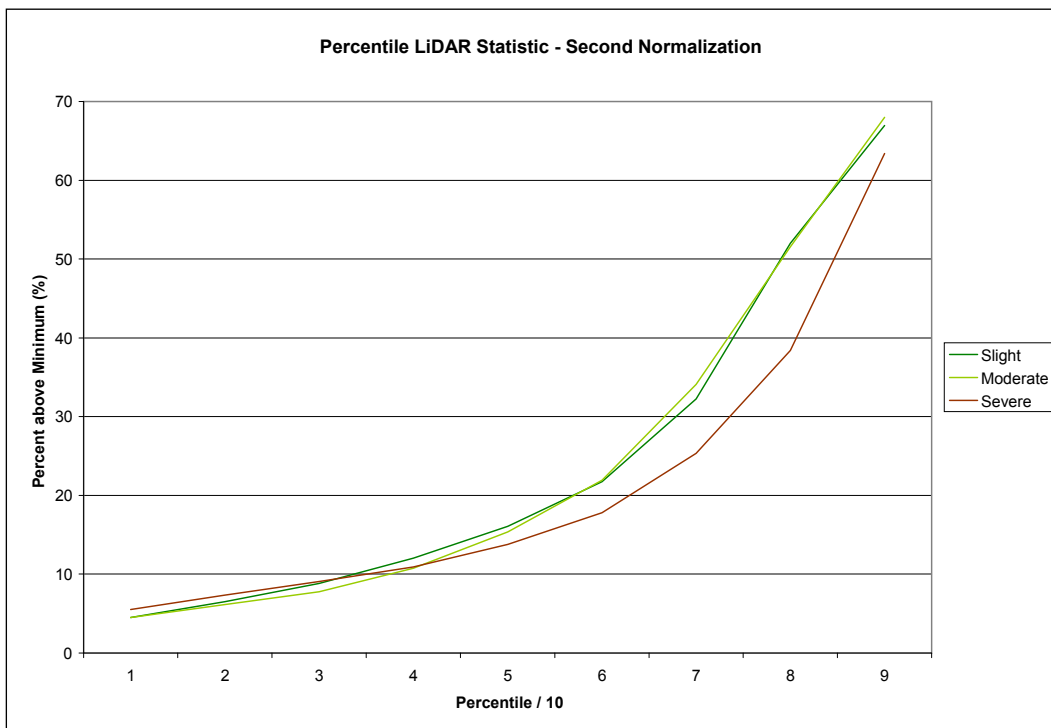
Figure 8.   Percentile LiDAR Statistic (Second Normalization).

Experiment 3: The final experiment utilized data collected from a high-density, very accurate Leica HDS3000 3D ground-based LiDAR system (Figure 9). Leica HDS (high-definition surveying) laser scanners use a rapid-firing, pulsing green (523 nm) laser (class 3R rating) for very detailed data acquisition. The highly accurate and dense point data captured by the terrestrial 3D laser scanner permits the development of robust datasets for GIS modeling and dynamic surface characterization, visualization, and quantification.

The scanner utilized for this study has a 360° horizontal by 270° vertical field of view that delivers positional, range, and angular (vertical and horizontal) single point accuracies (at 50 m range) of 6 mm, 4 mm, and 60 micro-radians, respectively. Each returned point coordinate is relative to the scanner's position and has an

Figure 9.   The Leica HDS3000 (ground-based) 3D LiDAR unit utilized for this study.

aggregate expected accuracy of 6 mm (one sigma standard deviation). The pulse rate is 1000 points/second with an optimal effective range up to 100 m that can produce a maximum point cloud density of 1.2 mm[1]. The laser also measures the intensity value of each point. The intensity value is a measure of the color and texture of the objects from which the laser reflects. Laser scanning, in general, is a rapid non-invasive form of data acquisition that is suitable for characterizing areas with restricted or limited access or where environmental conditions limit the ability to physically access the area.

Experiment 3 had two components that used the ground-based LiDAR unit. One component assessed the ability to classify three ground-based LiDAR datasets applying the rules generated by the See5 algorithm implemented in Experiment 2. The second component evaluated data collected by the unit for the ultimate development of techniques to characterize storm-impacted areas. The three datasets were all located in the study area near or within the DeSoto National Forest. Each ground-based LiDAR dataset was processed in the same manner as the airborne LiDAR in experiment 2. The primary results are statistical data describing the complete distribution of LiDAR points. Spectral data were also derived for the area using the technique and dataset described in Experiment 2.

The three sites selected most closely related to the three levels of damage or hurricane impact and thus remained consistent with previous experiments. However, the sites differed significantly in species composition. The first site (Attix Road) can be described as a hardwood riparian area within a mixed pine stand (Figure 10). This site had the most physical damage with many wind-thrown, downed hardwood trees. Many understory trees and saplings were impacted from surrounding, mature trees that were uprooted. The second site (Black Creek) was categorized as moderately damaged but was similar to the Attix Road site in damage characteristics and

---

[1] Upgraded specifications are improved at 4000 pts/sec effective up to the 300 m range.

represents a more uniform mix of hardwood and pine trees (Figure 11). The third and final site (Ashe Nursery) is a longleaf pine stand (Figure 12) and had the least amount of damage. Some of the trees were slightly damaged with some larger broken branches, but no trees were snapped or downed as a result of Hurricane Katrina.

Figure 10.   The Attix Road Site. This site is characterized by mostly hardwood trees and exhibited the most damage (Severe).

Figure 11.   The Black Creek Site. This site is characterized as a pine-hardwood mix and had some damage (Moderate).

Figure 12.   The Ashe Nursery Site. This is a longleaf pine stand with mostly herbaceous understory and had the least amount of damage (Slight).

After calculating the statistics for each site, the dataset representative of the site was classified accordingly. Given the very small sample size, the utility of this experiment is still in doubt. It is interesting to note that although the sites differed significantly, they were correctly classified using the same spectral information, but were in fact a different LiDAR dataset (in this case ground-based). Additional experiments are needed to test the validity of this apparent pattern. Also, future considerations should focus on determining if more descriptive or perhaps quantitative measures of forest stress could be derived using similar techniques.

The other component of Experiment 3 for this proof of concept was to demonstrate the utility of ground-based LiDAR data to characterize or otherwise better describe compromised forest environments using 3D mapping technology as opposed to visual inspection or traditional field assessments. Documenting a site's physical condition with LiDAR data allows for the development of robust quantitative descriptors as opposed to qualitative observations or time-consuming measurements. Using ground-based LiDAR to derive forest metrics is new research currently in the experimental phase. Establishing the technology for this purpose will help promote the implementation of additional sensor fusion concepts (such as integration with airborne LiDAR), greatly improving the ability to identify and monitor impacted areas in need of mitigation. Figure 13 illustrates the amount of detail afforded by the ground-based LiDAR and the modeling capabilities that exist.
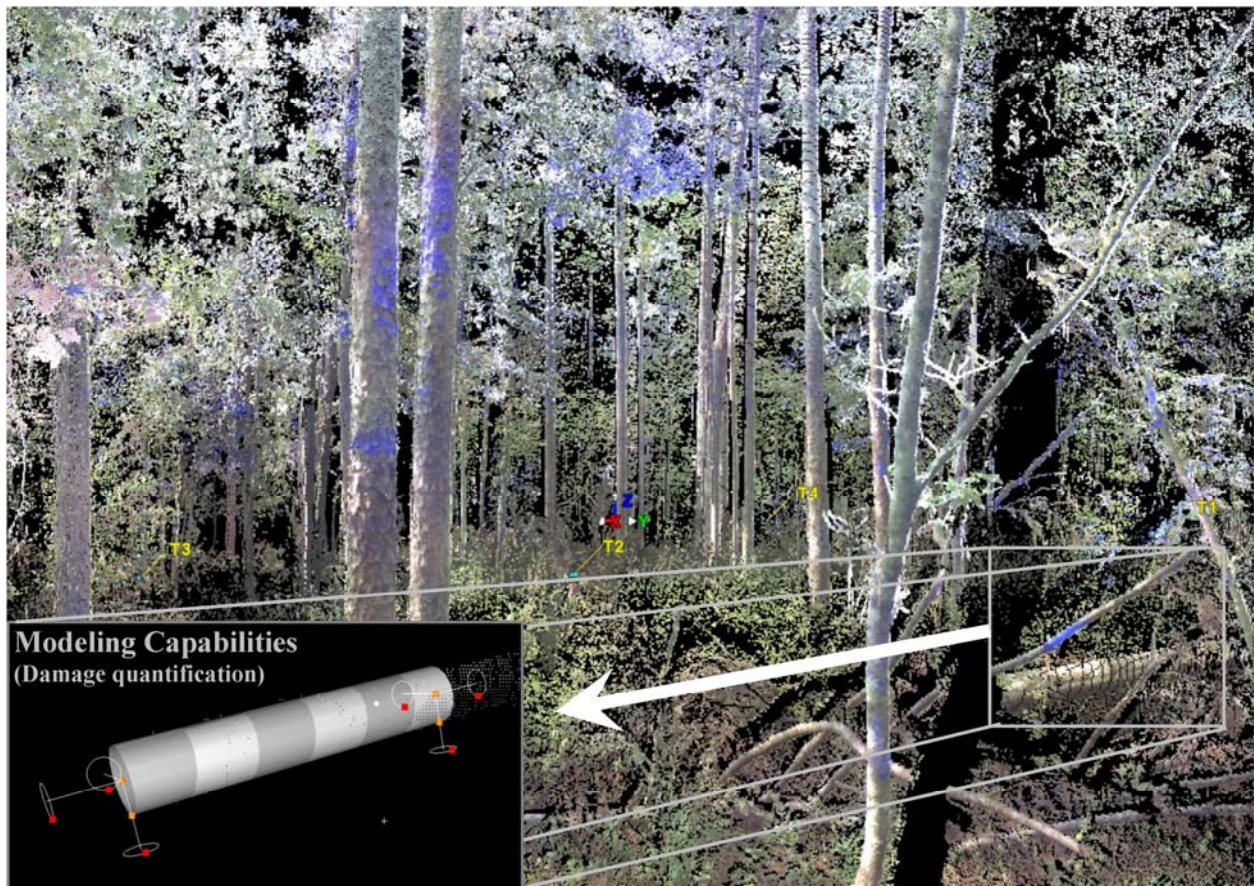

Figure 13.   High-resolution laser scan of Black Creek site. True-color laser postings are spaced at 3 m and generated from a Leica HDS3000 with an integrated digital camera. The inset image (in lower left corner) illustrates modeling capabilities of LiDAR-derived, wind-blown tree trunks.

Modeling forest damage characteristics will provide a quantifiable description of the compromised site. These areas can be spatially correlated with the spectral areas that exhibited definite indications of stress. An interesting approach would be to integrate the ground-based LiDAR data with spatially coincident airborne LiDAR data to determine any variability in the data models derived from both datasets, then fuse with high-resolution hyperspectral data to establish the relationships on a larger scale. Due to the density of the vegetation and ground clutter at each site, and with increasing distance from the unit, the ability to characterize compromised areas

becomes limited. In these situations, it may be necessary to conduct many individual, focused scans of these areas as a more effective approach. However, even with the limited data available for this study, it appears that modeling techniques are possible and can successfully quantify structural components of a damaged forest utilizing ground-based LiDAR technology. Future work should focus on applying the techniques to forest mensuration practices such as volume-loss estimation, canopy/branch loss, and perhaps implication to growth and yield estimates of areas impacted by significant storm events.

CONCLUSION: Using remote sensing to detect forest ecosystem stress is based upon the idea that stress causes changes in the spectral response of forest vegetation. Also, since LAI can be affected by stress, it is an important structural parameter for quantifying energy and mass exchange characteristics of terrestrial ecosystems (including photosynthesis, respiration, transpiration, carbon and nutrient cycling, and rainfall interception). These forest characteristics and parameters can all be used interactively to quantify and understand the effects hurricanes have on forested ecosystems; fortunately, most of these metrics can be derived and interpreted using remotely sensed data.

Based on this proof of concept using the limited data and analyses, there is evidence to suggest that both hyperspectral and LiDAR data can be used to successfully discriminate between categories of stressed vegetation. Even smaller stress differences were evident in the hyperspectral data, such as those in moderately versus slightly impacted areas. Furthermore, it was confirmed that techniques such as the automated decision tree analysis were applicable to this study. Even with a reduced sample size, the results appear portable to ground-based LiDAR systems. The ability to utilize these types of data, or perhaps even more ideal data sources such as full waveform LiDAR or high-resolution hyperspectral data, can only enhance the results and increase the significance of these findings.

Further research will be needed to establish a definitive correlation between quantifiable, stressed areas on the ground using LiDAR and hyperspectral data. Once this correlation has been made, a standardization to identify these forest stands could be possible and may be implemented on a larger scale, perhaps with hyperspectral imagery alone. Development of this knowledge may also help establish a timeline from the onset of damage to the onset of stress symptoms and may ultimately help predict mortality in compromised areas. Furthermore, the techniques described in this proof of concept will help identify stress-related factors resulting from subsequent wind damage (e.g., defoliation and foliage wilt) which indicate compromised tree vigor.

Dr. Beth C. Fleming, Chief of EE was Dr. David J. Tazik, and Chief of ESB was J. Mark Null. Dr. James R. Houston was Director of ERDC, and COL Richard B. Jenkins was Commander.

POINTS OF CONTACT: For additional information please contact Sam S. Jackson, U.S. Army Engineer Research and Development Center (ERDC), Waterways Experiment Station, Vicksburg, MS (601-634-3317, Sam.S.Jackson@usace.army.mil) or the SWWRP Program Manager Dr. Steven L. Ashby (601-634-2387, Steven.L.Ashby@usace.army.mil). This technical note was written by Sam S. Jackson, George T. Raber, Jerry A. Griffith, and Mark R. Graves. This document should be cited as follows:

> Jackson, S. S., G. T. Raber, J. A. Griffith, and M. R. Graves. 2008. *Concepts for sensor data fusion to detect vegetation stress and implications on ecosystem health following Hurricane Katrina*. SWWRP Technical Notes Collection, ERDC/EL TN-SWWRP-08-06. Vicksburg, MS: U.S. Army Engineer Research and Development Center. *http://el.erdc.usace.army.mil/*

REFERENCES

Asner, G. P., K. Carlson, and R. Martin. 2005. Substrate age and precipitation effects on Hawaiian forest canopies from spaceborne imaging spectroscopy. *Remote Sensing of Environment*, 457-467.

Ducic, V., M. Hollaus, A. Ullrich, W. Wagner, and T. Melzer. 2006. 3D vegetation mapping and classification using full-waveform laser scanning. Workshop on 3D Remote Sensing in Forestry, 14th-15th Feb. 2006, Vienna – Session 8a.

Quinlan, J. R. 1993. *C4.5 Programs for machine learning*. San Mateo, CA: Morgan Kaufman Publishers, 302.