# Investigation of Possible Contributions NDVI's Have to Misclassification in AVHRR Data Analysis[1]

David L. Evans[2] and Raymond L. Czaplewski[3]

**Abstract.**---Numerous subcontinental-scale projects have placed significant emphasis on the use of Normalized Difference Vegetation Indices (NDVI's) derived from Advanced Very High Resolution Radiometer (AVHRR) satellite data for vegetation type recognition. In multi-season AVHRR data, overlap of NDVI ranges for vegetation classes may degrade overall classification performance. The preliminary work of this project is presented to illustrate potential differences in vegetative cover discrimination capabilities between original visible and near-infrared (NIR) channels of AVHRR data and NDVI's derived from these channels. Initial tests on AVHRR data have been performed for regions of known vegetative cover composition in Northern Mexico. The results are presented here to stimulate debate on the effectiveness of use of NDVI's for vegetation type discrimination.

## BACKGROUND

The U.S. Department of Agriculture, Forest Service, has developed an increasing interest in the use of Advanced Very High Resolution Radiometer (AVHRR) data for vegetation assessments of broad regions. The AVHRR sensor is mounted on National Oceanic and Atmospheric Administration (NOAA) polar-orbiting weather satellites.

Typical applications of AVHRR data for forest assessments include land cover characterization (Brown, et al. 1993; Loveland 1991), forest density analysis (Iverson, et al. 1989; Ripple 1994; Zhu 1994), and general forest distribution

[1] *Approved for publication as Proceedings Article No. FA-055-0296 of the Forest and Wildlife Research Center, Mississippi State University.*

[2] *Assistant Professor of Forestry, Mississippi State University, Starkville, MS.*

[3] *Mathematical Statistician, USDA Forest Service, Rocky Mountain Forest and Range Experiment Station, Fort Collins, CO.*

mapping (Zhu and Evans 1994; Evans, et al. 1992). The latter of these applications utilized both the original spectral data and Normalized Difference Vegetation Indices (NDVI's) in various interpretation schemes to characterize earth cover.

The accuracy of products derived from AVHRR data is difficult to assess at continental scales due to the comprehensive coverage of the data and the low resolution of each data element (1.1 km at nadir). Therefore, increased attention has been devoted to use of data analysis techniques that optimize chances for high classification success. One possible way to increase classification success is to identify the data combinations and transformations that can be used effectively in automated classification procedures. Exclusive use of some data transformations could deprive classification procedures of valuable information that is inherent in the original data. The concern identified in this paper is that NDVI's (data transformations) may forfeit useful information that could have been obtained from the original AVHRR spectral channels.

## DERIVATION AND USES OF NDVI's

The NDVI is a data transformation (Eq.1) used to characterize vegetation vigor or "greenness."

$$NDVI = (NIR - VIS / NIR + VIS) * 100 \qquad [1]$$

In Equation [1], NIR is the reflectance values for the near-infrared channel (channel 2 in AVHRR) and VIS is the reflectance values for the visible channel (channel 1 in AVHRR) from satellite data.

A common approach used to interpret AVHRR data is to acquire imagery over one year's time and calculate NDVI's to document vegetative phenology. This is done to identify different or distinctive phenological sequences that can be attributed to vegetation types, crop development, and regional climate conditions.

NDVI data have been used as input to classification algorithms to identify vegetative cover types (Brown, et al. 1993; Loveland, et al. 1991). With time as a critical component in classification, it is assuned that vegetative cover types have different phenological time frames. Given the low sensitivity of the AVHRR instrument to spectral variance in 1.1 km resolution elements, this assumption of separability may not hold in the classification process.

A complicating factor in vegetation identification is that the same vegetation may exhibit a wide range of latitudinal variation in phenology. There are numerous examples of wide-ranging eastern deciduous trees that occur from the deep south to northern states. Eastern cottonwood (*Populus deltoides*, Bartr.) ranges from southern Louisiana to central Minnesota and red maple (*Acer rubrum*, L.) ranges from central Florida to southern Canada. Many of the Oaks (*Quercus* spp. ) and hickories (*Carya* spp. ) also exhibit wide geographic ranges. These wide geographic ranges can easily mean that leaf fall or leaf flush could vary two to three months across the range for any given species. The timing of phenologic events for any given cover type must contribute to spectral reflectance pattern variation that can ultimately cause confusion in image classification.

Confusion between different cover types with the same NDVI's can be illustrated by a hypothetical example. Let us suppose that forest land in a region has spectral reflectance values of 20 and 40 for the visible and NIR channels in AVHRR data. Grassland in the same region may have values of 30 and 60 for the same two channels. Clearly these two cover types are distinguishable in the separate spectral channels but the NDVI's for each are identical (rounded value of 33). In this hypothetical but realistic case, the use of NDVI's inadvertently filters out useful information and causes confusion. This is an assumed single-date case but even if there were multiple dates of imagery, it is possible that some NDVI's would be the same for otherwise different cover classes and therefore, would degrade classification performance.

## RESEARCH OBJECTIVE

The basic objective of use of AVHRR data by the USDA Forest Service is to distinguish forest cover from other earth cover over broad continental-scale regions. This has been demonstrated by (Zhu and Evans 1994). This project utilized NDVI's and the original spectral data, but placed significant reliance on ancillary data and post-processing techniques to resolve conflicts in cover class identification. The concern for future uses of AVHRR NDVI's is the ability to assess earth cover in regions where little ancillary information is available. Lack of ancillary information dictates that the maximum amount of information has to be extracted from original image data. This project was initiated to evaluate information content in satellite data to better define appropriate interpretation techniques that will maximize success in cover type identification.

## INITIAL METHODOLOGY

The first phase of this project is to compare cover class separability with NDVI's versus the visible and NIR AVHRR channels for areas of known ground cover. Early work has concentrated on use of data from an on-going project to map the forest cover of Mexico and Central America (Lannom 1995). Landsat TM data of Northeastern Mexico were classified as agriculture (generally fallow fields), forest, and scrub land.

AVHRR data used in this initial investigation were near-cloud-free composites derived for North America by the maximum NDVI method described by Eidenshink (1992). These data sets were composed of information compiled over 10-day intervals. Those AVHRR data sets that did not exhibit clouds or substantial compositing boundaries (due to abrupt changes in reflectance values between image dates) were chosen for initial study. This yielded 9 data sets from those available over one year that were deemed useful for the study.

The TM data were co-registered to the AVHRR data. A linked display system implemented in PCI's Imageworks[4] was used to locate 3 areas in each TM cover class that covered 25 AVHRR cells (75 total cells). EASI/PACE software was then used to export the AVHRR visible, NIR, and NDVI data values to a text file for analysis with the Splus statistics software package. The discriminant analysis function in Splus was used to generate discriminant correlations (measure of the ability to identify the three cover classes) for the visible, NIR, visible and NIR combined, and NDVI channels of the AVHRR data set.

## RESULTS AND DISCUSSION

The results of the discriminant analysis work are presented in Table 1. The statistical significance of these findings has not been determined because these tests represent a limited area and should be greatly expanded to cover a wider range of conditions that are assumed to occur in large AVHRR data sets.

**Table 1.—Discriminant correlations and averages for three cover classes as represented by visible NIR and NDVI channels from selected AVHRR data sets.**

| AVHRR Composite | Channels VIS | NIR | VIS & NIR | NDVI |
|---|---|---|---|---|
| April1 | 0.9148 | 0.5656 | 0.9839 | 0.9586 |
| April2 | 0.6793 | 0.4311 | 0.9878 | 0.8900 |
| April3 | 0.9458 | 0.6711 | 0.9702 | 0.9323 |
| June2 | 0.9062 | 0.2806 | 0.9688 | 0.9617 |
| July3 | 0.9305 | 0.2411 | 0.9459 | 0.8468 |
| Sept1 | 0.9549 | 0.2290 | 0.9734 | 0.9445 |
| Sept2 | 0.9297 | 0.2883 | 0.9773 | 0.9187 |
| Oct2 | 0.8337 | 0.1872 | 0.9470 | 0.9175 |
| Nov1 | 0.8296 | 0.1802 | 0.9486 | 0.8721 |
| Average | 0.8805 | 0.3416 | 0.9670 | 0.9158 |

It was interesting to note that on average, the visible channel seemed to have better discriminating capabilities for the three subject cover classes than did the NIR. One would expect that the NIR channel, which is sensitive to vegetation response, would provide better discriminating capabilities. This area of Mexico (particularly west of the mountains) is dominated by fairly arid lands. The agriculture and scrub land categories undoubtedly had significant amounts of exposed soil that influenced the visible reflectance channel values. The visible

[4] *Mention of company or product names is for information only and does not constitute official endorsement by Mississippi State University or the USDA Forest Service.*

channel generally exhibits greater contrast and variance in spectral values between vegetated and non-vegetated cover. Thus, the greater contrast in visible channel pixels significantly influenced differentiation of forest from the other categories that had more bare ground exposed to the sensor.

The NDVI channel outperformed the visible and NIR channels when used singly. It did not seem to discriminate the subject classes as well as the combined visible and NIR channels. This finding seems to support the original premise that NDVI may mask useful information found in the original channels that could be used to discriminate different vegetative cover. Additional and more rigorous testing is recommended to substantiate these findings.

## FUTURE WORK

Additional testing using these or modified techniques will be much broader based than this exploratory study. TM and AVHRR data sets are readily available for different locations in Central America, Mexico, and the U.S. Testing over this broad range of conditions will provide a better determination of the utility of different AVHHR channels in vegetation discrimination. The findings from this work will likely have broad application in the improvement of on-going and future studies that utilize AVHRR for global vegetation distribution assessments.

## LITERATURE CITED

Brown, Jesslyn F., Thomas R. Loveland, James W. Merchant, Bradley C. Reed, and Donald O. Ohlen. 1993. Using multisource data in global land-cover characterization: concepts, requirements, and methods. Photogramm. Engr. and Remote Sensing. 59(6):977-987.

Eidenshink, Jeffery C. 1992. The 1990 conterminous U.S. AVHRR data set. Photogramm. Engr. and Remote Sensing. 58(6):809-813.

Evans, David L., Zhiliang Zhu, Susan Eggen-McIntosh, Pedro Garcia Mayoral, and Jose Luis Ornelas de Anda. 1992. Mapping Mexico's forest lands with advanced very high resolution radiometer. USDA Forest Service, Southern Forest Experiment Station, Research Note SO-367, 4 p.

Iverson, L.R., E.A. Cook, and R. L. Graham. 1989. A technique for extrapolating and validating forest cover across large regions: calibrating AVHRR data with TM data. International J. Remote Sensing. 10(11):1805-1812.

Lannom, Keith B. 1995. Mapping forest distributions of Mexico and Central America. Geomatics Info Magazine. 9(5):33-37.

Loveland, Thomas R., James W. Merchant, Donald O. Ohlen, and Jesslyn F. Brown. 1991. Development of a land-cover characteristics database for the conterminous U.S. Photogramm. Engr. and Remote Sensing. 57(11):1453-1463.

Ripple, William J. 1994. Determining coniferous forest cover and forest fragmentation with NOAA-9 very high resolution radiometer data. Photogramm. Engr. and Remote Sensing. 60(5):533-540.

Zhu, Zhiliang. 1994. Forest density mapping in the lower 48 States: a regression procedure. USDA Forest Service, Southern Forest Experiment Station, Research Paper SO-280. 11 p.

Zhu, Zhiliang, and David L. Evans. 1994. U.S. forest types and predicted percent forest cover from AVHRR data. Photogramm. Engr. and Remote Sensing. 60(5):525-531.

## BIOGRAPHICAL SKETCHES

David L. Evans is an Assistant Professor of Forestry at Mississippi State University, Starkville, MS. He completed the Ph.D. in Forest Management at Louisiana State University in 1986. His teaching and research interests are in the use of remote sensing and geographic information systems technologies for natural resource assessments. Dr. Evans was formerly with the USDA Forest Service as a Supervisory Research Forester and Team Leader for remote sensing research in the Southern Research Station.

Ray Czaplewski has spent the past 13 years working on statistical design and analysis of regional monitoring systems for forests and other environmental features, primarily the Forest Service's nationwide Forest Inventory and Analysis Program. Other accomplishments include statistical design of the 1990 and 2000 global surveys of forests by the Food and Agricultural Organization of the United Nations, and design to monitor the effectiveness of the President's 1994 Forest Plan for the States of Oregon, Washington, and California. He has specialized in linking sample surveys to remotely sensed data and geographic information systems.