U.S. Forest Types and Predicted Percent Forest Cover from AVHRR Data

Zhiliang Zhu and David L. Evans

Abstract
Two forest-cover maps of national scale have been produced under the 1993 Forest and Rangeland Renewable Resources Planning Act (RPA) Assessment Update program. Both maps are based on 1-kilometre resolution Advanced Very High Resolution Radiometer (AVHRR) data, and use of multitemporal and multisource remote sensing data analyses. The forest-type groups map depicts the distribution patterns of 25 forest-cover types over the United States. Complementary to the forest-type groups map is the predicted percent forest-cover map of the conterminous United States which portrays the relative amount of forest cover per square kilometre. The two maps, in both digital and paper forms, provide current forest information for the U.S. Techniques and results of the 1993 RPA forest mapping project are presented, and apparent changes in forest cover since 1967 are also described.

Introduction
The U.S. Department of Agriculture Forest Service conducts nationwide assessments of forest resources, as mandated by the Forest and Rangeland Renewable Resources Planning Act (RPA) of 1974. The assessments are performed by the Forest Service Forest Inventory and Analysis (FIA) programs on a 10-year cycle, with updates every five years. In previous assessments, there has been no attempt to produce a forest-cover map of the whole country. The only available forest-cover type map at the national scale is the 1967 Major Forest Types map, compiled by the Forest Service and published in The National Atlas of the United States of America by the U.S. Geological Survey (U.S. Department of Agriculture, Forest Service, 1967). Information contained in the 1967 map was generalized at the regional level. Details in forest type distribution and overlapping of different forest types are not well represented, and the information for most locations is not current.

With advances in satellite remote sensing and geographic information systems (GIS), it has become feasible to characterize land cover and produce thematic maps for a very large area (Townshend et al., 1991). Developments in these technologies led to the decision to create a new U.S. forest type map for the 1993 RPA Assessment Update. The forest-cover mapping project was begun in early 1991 at the Forest Service, Southern Forest Experiment Station's Forest Inventory and Analysis (SO-FlA) unit. It was completed in late 1992.

The primary goal of this project was to create a new forest type map of the United States in support of the 1993 RPA Assessment Update. The map depicts the current spatial distribution of the forest type groups as reported in the RPA timber tables (Powell et al., 1993). In the process of achieving this goal, emphasis was placed on defining procedures that would utilize established remote sensing techniques and satellite data. Research results from this project will provide valuable inputs to SO-FlA's long-term plan of developing a forest inventory update model.

The project was also designed to update information in the 1967 Major Forest Types map. By depicting the same forest-cover types, the new map may be compared to the old to show locations of possible regional forest cover changes that have occurred since 1967. Such information is significant to the global change research.

Initial research in the first phase of a two-phase approach was focused on methodology development (Zhu and Evans, 1992) in seven southern states (Alabama, Arkansas, Louisiana, Mississippi, Oklahoma, Tennessee, and Texas), where SO-FlA routinely performs field inventories. The lower 48 states and Hawaii were mapped in the second phase based on the methodology developed in the first. Alaska forest mapping work was performed by the FIA scientists at the Pacific Northwest Forest and Range Experiment Station, and the results will be reported separately.

Background
RPA Forest-Type Groups
Forest surveys and RPA assessments performed by the FIA programs have used a cover-type classification scheme with similar forest types and forest-type groups (Table 1) as defined in Eyre (1980). The determination of a forest type is based on the plurality of tree stocking on a forest land. In absence of a stocking majority, other species are grouped within these types as associate species. Note the scheme only concerns forest cover. All other cover categories are grouped into two classes: nonforest and water.

Most of the forest-type groups are regionally distributed, such as longleaf-slash pine in the southeast and redwood along the Pacific coast. A few type groups consist of species that are trans-continental. The fir-spruce type group represents both fir species in the western states and white spruce in Alaska. Likewise, aspen-birch occurs both in the...
Experience indicated that the difficulty with the RPA forest-type mapping project was not the large area that was to be mapped, but the detail of the map defined by the classification scheme which was in greater detail than an Anderson et al. (1976) level-2 classification. Spectral variations caused by physiographic factors (e.g., terrain changes) are far greater in data sets covering a large area than a small study area, such as a Landsat scene. Stratification of a large area into physiographic regions has been suggested to reduce the effect of physiographic variations in spectral data (Loveland et al., 1991). Spectral variations due to non-vegetation factors are reduced when data are segregated and analyzed for separate regions and the classification is more focused on local vegetation types.

Many studies have attempted to synthesize the country's land characteristics at the regional scale. Geomorphic, climatic, and vegetation factors were employed to describe divisions of ecoregions or land surface forms. Study of Hammond's Land-Surface Form map (Hammond, 1964) and work by Bailey (1980), Fenneman and Johnson (1946), and Omernik (1987) provided a basis for stratifying the conterminous U.S. into 15 physiographic regions. Boundaries of the 15 regions were drawn on Hammond's map, then digitized into polygon files (Figure 1).

### AVHRR Data
The primary data used in this project were from the Advanced Very High Resolution Radiometer (AVHRR) on National Oceanic and Atmospheric Administration (NOAA) polar-orbiting weather satellites. Biweekly composite images and a normalized difference vegetation index (NDVI) were prepared from 1-kilometre AVHRR data (Gallo and Eiden- shink, 1988) by the U.S. Geological Survey EROS Data Center at Sioux Falls, South Dakota. The AVHRR composites from the EROS Data Center are composed of picture elements (pixels) that have the maximum NDVI value from multiple dates of data (Holben, 1986). Work by Loveland et al. (1991) used AVHRR NDVI and other ancillary data to produce an Anderson level-2 classification (Anderson et al., 1976) database of the conterminous U.S. Research conducted at SO-FIA has also shown the usefulness of AVHRR data to support FIA state forest inventories (Teuber, 1990; Zhu, 1992).

### Methods
Nine 1991 AVHRR U.S. composite image data sets were acquired from the EROS Data Center. The time-series of the nine data sets, spanning the growing season of 1991 from early March to middle of November, were georegistered to a common Lambert azimuthal equal area map projection, then partitioned into 15 physiographic regions using the polygon files. Each of the nine AVHRR images were examined visually to note problems with residual clouds, low haze, and remaining seams between different dates of the composited data. Only clear image segments were used for the regional data processing.

### Regression Procedure
Statistical regression analysis with AVHRR and TM data was performed (1) to construct a map of predicted within-AVHRR-pixel percent forest-cover values, and (2) to provide a mechanism for selection of the optimal AVHRR dates and spectral channels for forest-type classification. The regression procedure was based on research on resolving forest and nonforest components in mixed pixels (Iverson et al., 1989). They used expensive ancillary data with fine ground resolution but a limited sample area (Landsat TM) to predict the percentage of forest cover within pixels of AVHRR data of a larger region.

---

**Figure 1. Physiographic regions used in the RPA mapping project. The bold line outlines the seven southern states used as the test areas for methodology development in the first phase of the project.**
However, a model derived for one region is likely to be inaccurate when applied to other regions with different vegetation and geophysical conditions. Therefore, it was deemed necessary to develop separate models for each of the 15 physiographic regions defined in this project. In a few regions, two models were developed to cover extended areas. These regression models were not necessarily the same form (linear or nonlinear) or based on the same variables (AVHRR spectral channels).

Nineteen Landsat TM scenes and quarter scenes were used for the AVHRR-TM modeling process. Each TM scene was classified as forest or nonforest. The TM classifications were then registered to the AVHRR spectral data, and regressed against the AVHRR spectral channel data for the area under the TM scene boundary. Tests for simple linear and polynomial models were performed for each of the 19 regressions. Values of adjusted R-squares ($R^2$) for the 19 chosen models ranged from 0.51 to 0.85. Details of this process are being reported elsewhere (Zhu, in press).

The AVHRR channels (variables), identified by each regression model, were used to create an image data file for each region. These data represented optimum dates and spectral channels for spectral characterization of forest cover by region. The image files were used for forest-type classification and for creating a predicted percent forest-cover map for each region.

A map of percent forest cover was created by applying the regression coefficients from each model to the corresponding regional AVHRR image file formed with selected spectral channels. Pixel values for the percent forest-cover map ranged from 0 to 100. The values represented the predicted percent forest cover, as determined by number of TM forest pixels, within AVHRR pixels. The 15 regional maps were combined to form one predicted percent forest-cover map file for the entire lower 48 states.

Forest-Type Classification

Various reference materials were collected to assist in the forest-type classification. FIA research programs provided survey plot data with field locations for many parts of the country. The FIA data served as the basis for labeling AVHRR spectral classes in each of the physiographic regions. Other reference materials included regional forest type maps, regional and local vegetation studies, FIA publications, and Landsat image prints.

Classifications were performed for each physiographic region using the same AVHRR image data used for the regression process. In regions with high local relief (i.e., Appalachians in the east, most of western states), an elevation channel was included during the image classifications to help identify forest types that display altitudinal zonation.

Unsupervised classification algorithms were used in a process that frequently involved several iterations of classification, spectral signature evaluation, masking (selectively identifying certain pixels), and recoding. The percent forest-cover map for each region was used to separate nonforest from forest-cover classes. All pixels with a corresponding percent forest-cover value lower than or equal to 25 percent were assumed to be primarily nonforest, and were labeled accordingly. The basis for this cut-off value included examination of cross-tabulations between percent forest-cover values and the FIA field survey data, and precedents from other studies (for example, USDA Soil Conservation Service uses 25 percent canopy cover as a forest land definition).

After nonforest and water classes were identified, EPA forest types were assigned to other classes based on reference data, knowledge of regional vegetation, and class statistics (e.g., spectral scatter-grams, class histograms). Distribution of forest types in the western United States often results from interactions between climatic effects, topographic effects, and tree biology. Use of ancillary data, such as elevation and length of growing season, provided effective separation of many western type groups. On the other hand, areas with complex local terrain conditions sometimes resulted in highly mixed spectral classes. In these areas, notably northwestern Montana and northern Idaho, detailed forest maps and aerial photos were used to label the classes.

Compared to the west, eastern forests were less influenced by topographic and climatic factors. This led to larger transitional zones between forest types in the east. In these areas, current FIA field survey plot data provided a key reference to the class labeling process.

Feedbacks from FIA units, Forest Service regional offices, and other cooperators that reviewed draft copies of the initial classification, indicated that forest land was overestimated in areas either dominated by nonforest land uses, or with a high degree of forest land fragmentation. The overestimation seemed to be caused by assigning forest/nonforest mixture pixels in favor of forest rather than nonforest during classification. Visual examination of the relationship between predicted percent forest-cover values and the forest-type classification suggested that the AVHRR classifications masked by a predicted percent forest-cover value at 40 percent matched closely with most of the ground reference data. This value was subsequently used in a GIS model to reunify forest and nonforest pixels in areas of overestimation. It should be noted that this value was not universally applied. Relationships between percent forest cover and forest/nonforest classification varied between regions or different land-cover patterns.

Hawaii

Persistent cloud cover in many parts of the Hawaiian Islands precluded use of 1991 data for the Hawaii portion of the RPA map. Three AVHRR images, one acquired in 1992 and the other two in 1989, were obtained with acceptable cloud cover. There were no other recent Landsat or French SPOT satellite data available with minimal cloud cover. The Hawaii forest type classification was thus based primarily on the single 1992 AVHRR image and the same unsupervised classification techniques as used in the lower 48 states. Areas covered by remaining clouds were filled in by using classifications from the two older AVHRR images, and by manual editing.

Verifications

Much progress has been made on the accuracy assessment of land-cover classifications by remotely sensed data (Congalton, 1991), but applications of these techniques have been limited in studies using AVHRR. It is impractical to obtain a reference data set for testing positional accuracy of an AVHRR classification largely because (1) it is often difficult to identify a unique cover type in one square kilometre pixels, and (2) very large study areas are typical of AVHRR applications.

The strategy to assess the quality of the RPA forest mapping project was to adopt a combination of tests using available reference data. In a related project, National High Altitude Photography (NHAP) and National Aerial Photography Program (NAPP) aerial photography (nominal scales 1:56,200 and 1:40,000 respectively) for three Louisiana par-
ishes were scan digitized and interpreted. Polygons from the photointerpretation were used to compare to the AVHRR forest-type groups map; results of this test are being reported separately.

Another test compared forest area data between the RPA map and the database developed for the current RPA assessment. The database contains 1991 state-level forest variables summarized from FIA field surveys in all 50 states. FIA state forest area estimates were retrieved from the database and were subtracted from those of the RPA map to yield percent differences for the 50 states and associated statistics.

An accuracy assessment was performed on a random selection of nine of the 19 TM scenes used for the AVHRR/TM models. The purpose was to further validate the credibility of the forest percentage models. Recent (1 to 9 years old) NHAP and NAPP aerial photos were used. Fifteen to 20 sample points were randomly located on each aerial photo in either forested or non-forested areas that were at least two hectares in size. The point locations were transferred to topographic map sheets, digitized, and coded as being forested or non-forested. A GIS buffer algorithm was used to delineate a polygon of approximately two hectares around each point. These verification data and the TM classifications were then compared to generate an estimate of the TM classification accuracy.

Results and Discussion
Forest-Type Groups Map
The new forest type groups map (Plate 1) supports assessment results reported in the current RPA Assessment Update document (Powell et al., 1993). Among the 25 forest-type groups, oak-hickory covers the largest area and occurs in a broad range from southern Texas to New England, including a few continuous areas such as West Virginia and the Ozark Plateaus. The most forested region is New England, while the least is the plains states.

Conventional large-scale cover maps, such as the 1967 Major Forest Types map, are often produced by interpolation and generalization of point data into polygons. Transitional zones among forest types and between forest and nonforest lands are better represented in the new RPA forest-type groups map than in the 1967 map. The general distribution of forest lands in the western states has remained the same; areas forested many years ago are still forested today.

Changes in forest-land cover have taken place only in relatively small areas. In contrast, distribution of forest lands has changed in many of the eastern states except the northeast. Impacts on eastern forest lands have been greater in recent years than on those of the west (Alig et al., 1990), as the result of a high degree of human modification. In many areas across the midwest, forest lands are sparsely distributed relative to dominance of nonforest land cover. The 1967 map, by contrast, depicts more uniform and broad coverage of forest lands.

Examples of changes in forest land cover and forest types can be found in the southeast and south-central regions. A 14 percent decline in forest area was reported for the southwest region of Georgia between 1961 and 1988 (Knight and McClure, 1974; Thompson, 1989). This finding is supported by the RPA map which depicts primarily nonforest land cover in this area. Similar land diversions in southeast Alabama indicate a large agricultural base in that region. In Alabama, a shift from pine to upland hardwoods is reflected in two recent SO-FIA surveys which found that the oak-hickory acreage increased by 26 percent from 1972 to 1990 (Rudis et al., 1984; McWilliams, 1992). This change in forest types is illustrated in the new map, where much of the forested area in central Alabama is now a large transitional zone between the loblolly-shortleaf type and the oak-hickory and oak-pine types.

Predicted Percent Forest-Cover Map
Values (ranging between 0 and 100) in the predicted percent forest-cover map of the conterminous U.S. (Figure 2) represent the amount of forest as determined by number of forested TM cells within each AVHRR pixel.

The forest-type groups map and the percent forest-cover map represent complementary forest land information. Information on location and type of a forest is given by the forest-type groups map, while the relative amount of forest at the same location can be obtained from the percent forest-cover map. Central Oklahoma, northern Arkansas, and much of Tennessee are shown dominated by the same oak-hickory type group (Plate 1). However, hardwoods in central Oklahoma have lower percent forest cover values than those in the other two regions. Among the 22 RPA forest-type groups in the lower 48 states, pinyon-juniper and chaparral were consistently associated with percent forest-cover values lower than the other type groups — an expected condition in typical western woodland landscapes. Over the entire lower 48 states, the most contiguous forests occur in the northwest Coastal Ranges, the Cascades and Sierra Nevada (Douglas-fir being the major component), and the central Appalachian Mountains (mostly oak-hickory forests).

Because the percent forest-cover map is a predicted output, the values in many localities may not reflect the true forest-cover density. Improvements on future research may consider use of additional calibrations in areas of confusion, and a corresponding confidence layer to provide an uncertainty measure.

Verification Results
The mean and standard deviation of the difference variable between FIA and the RPA map estimations of 1991 state percent forests were found to be 2.02 and 1.48 percent, respectively. The minimum and maximum percent differences for the above variable were 0.07 (North Dakota) and 7.32 (Alaska). Figure 3 shows the differences, by state, for the conterminous 48 states.

The average agreement for the nine TM accuracy tests for forest and nonforest classes were 85 percent and 91 percent, based on average sample sizes of 1933 and 1421 TM pixels, respectively. The average of overall agreements was 89 percent.

Classification Scheme
The RPA scheme of forest-type groups is not always comparable with local land-cover patterns or with spectral and spatial resolutions of remotely sensed data. This was particularly true in areas where multiple RPA type groups together form distinctive ecosystems. In California's Sierra Nevada, Douglas-fir, ponderosa pine, and white fir are often co-dominant in mixed conifer forests, which are considered a successional ecosystem to pure stands by the component species (Tappeiner, 1980). It may be undesirable to allocate this type of forest to one of its component groups. Similarly, a mixed conifer or mixed forest-type group would have been preferable in parts of western Montana and northern Idaho (western larch, ponderosa pine, lodgepole pine), and in
Maine where spruce, maple, and beech can be intermixed within the one square kilometre resolution. However, the classification scheme identified in this work was predetermined based on forest-type groups used for the RPA forest resource assessment program.

A hierarchical format could be useful for the RPA classi-
This approach, following the Anderson system (Anderson et al., 1976), would provide an advantage for projects of different levels of detail to use the same system. For example, a mixed forest class may be organized at one level higher than the forest-type groups when individual forest-type groups in certain areas are indistinguishable in a particular remote sensing data type (e.g., AVHRR). In addition, under such a format, classification results may be more easily compared with, or integrated to, hierarchical ecoregion classifications.

Summary and Conclusions

A forest-type groups map of the United States and a predicted percent forest-cover map of the lower 48 states were produced in this project. Key to the project were methods which combined remote sensing data processing techniques with regression analysis of multitemporal and multisource data. The major methods used in this project include:

- use of multitemporal data sets in one growing season,
- physiographic stratification to reduce spectral variations due to abiotic factors,
- multisource remote sensing data and statistical analysis to derive percent forest-cover images (sub-pixel measurement), and
- relating percent forest-cover values in the forest type classification.
The objectives of the mapping project were such that the end products were not intended to be absolute or precise in terms of accuracy in minute detail. It is the regional perspective and analysis that are most important in using the maps.

It was concluded from this work that multitemporal AVHRR data can be used to produce fairly detailed forest-cover maps, provided that sufficient ancillary data are available for identifying of spectral classes. The percent forest-cover models were important in both defining input data for spectral classifications as well as in refining the results of those classifications.

Results from this project should be useful for other large-area land feature studies. A database can be constructed using the forest-type group and predicted percent forest-cover map files as the base layers. Along with other layers of information (e.g., elevation, growing season), this database can be used in ecosystem modeling research, or serve as an input to global change studies.

References


