Development of a Land-Cover Characteristics Database for the Conterminous U. S.

Thomas R. Loveland
USGS EROS Data Center, Sioux Falls, SD 57198
James W. Merchant
Center for Advanced Land Management Information Technologies, Conservation and Survey Division, Institute of Agriculture and Natural Resources, University of Nebraska-Lincoln, Lincoln, NE 68588-0517
Donald O. Ohlen
TGS Technology, Inc., USGS EROS Data Center, Sioux Falls, SD 57198
*Jesslyn F. Brown**
Center for Advanced Land Management Information Technologies, Conservation and Survey Division, Institute of Agriculture and Natural Resources, University of Nebraska-Lincoln, Lincoln, NE 68588-0517

ABSTRACT: Information regarding the characteristics and spatial distribution of the Earth's land cover is critical to global environmental research. A prototype land-cover database for the conterminous United States designed for use in a variety of global modeling, monitoring, mapping, and analytical endeavors has been created. Database development has involved (1) a stratification of vegetated and barren land, (2) an unsupervised classification of multitemporal "greenness" data derived from Advanced Very High Resolution Radiometer (AVHRR) imagery collected from March through October 1990, and (3) post-classification stratification of classes into homogeneous land-cover regions using ancillary data. Ancillary data sets included elevation, climate, ecoregions, and land resource areas. The resultant database contains multiple layers, including the source AVHRR data, the ancillary data layers, the land-cover regions defined by the research, and translation tables linking the regions to other land classification schema (for example, UNESCO, USGS Anderson System). The land-cover characteristics database can be analyzed, transformed, or aggregated by users to meet a broad spectrum of requirements. Future research plans include examination of impacts of interannual change, landscape and sensor interaction, development of improved analytical tools and methods, and appropriate modes for verification.

INTRODUCTION

INFORMATION REGARDING THE CHARACTERISTICS and spatial distribution of the Earth's land cover is critical to global change research. Capabilities to inventory and map land-cover conditions and to monitor change are required for, among other things, modeling the global carbon and hydrologic cycles, studying land surface-climate interactions, and establishing rates of tropical deforestation (Risser, 1985; Dale, 1990; International Geosphere-Biosphere Programme, 1990; Pinker, 1990; Pielke and Avissar, 1990; Dorman and Sellers, 1989). Global land process research heretofore has had to rely upon simple interpretations of gross land cover and surface properties, such as biomass, albedo, surface roughness, and canopy resistance, at low spatial resolution (Henderson-Sellers *et al.*, 1986). The Matthews land cover and natural vegetation (Matthews, 1983; 1984) and the Olson and Watts major world ecosystems (Olson and Watts, 1982) global databases are the most common sources of land-cover and surface parameter data. These data bases have, respectively, 1° by 1° and 0.5° by 0.5° spatial resolution. Higher resolution data with greater precision for classification are clearly required (International Geosphere-Biosphere Programme, 1990).

During the last decade, substantial progress has been made in using National Oceanic and Atmospheric Administration (NOAA) Advanced Very High Resolution Radiometer (AVHRR) data for land cover characterization (for example, Goward et al., 1985; Tucker *et al.*, 1985; Roller and Colwell, 1986; Townshend *et al.*, 1987; and Lloyd, 1990). AVHRR data have only moderate spatial resolution (1 km) when compared, for example, to Landsat's 80 m for multispectral scanner (MSS) and 30 m for thematic mapper (TM) or SPOT's 20 m for multispectral and 10 m for panchromatic data. AVHRR data are, however, collected more frequently, with virtually the entire globe imaged twice each day. The high frequency of coverage enhances the likelihood that cloud-free observations can be obtained for specific temporal windows, and makes it possible to monitor change in land cover conditions over short periods, such as a growing season (Miller *et al.*, 1988; Tappan and Moore, 1989; Justice *et al.*, 1985; Goward *et al.*, 1985). Moreover, the moderate resolution of the data makes it feasible to collect, store, and process continental or global data sets.

Research on applications of AVHRR data for land-cover inventory and monitoring has focused on analysis of vegetation "greenness." Greenness is most often measured using a vegetation index, commonly the Normalized Difference Vegetation Index (NDVI) (Goward *et al.*, 1985). A number of investigators have shown that changes in greenness during a growing season can be observed and often correlated with the spatial distribution of major biomes (Townshend *et al.*, 1987; Tucker *et al.*, 1985; Lloyd, 1990). Because of limitations in AVHRR data availability, almost all regional, continental, and global-scale analyses have used data that have been resampled to either 4- or 16-km pixels (Global Area Coverage [GAC] or Global Vegetation Index [GVI] data). Only recently have spatially extensive data sets at the highest nominal resolution (1.1 km) started to become available on a continuing basis for major land areas. (Note: 1-km AVHRR data are referred to as high resolution picture transmission [HRPT] for data collected directly by ground receiving stations, and as local area coverage [LAC] for data gathered using on-board satellite tape recorders.)

The U.S. Geological Survey, National Mapping Division's (USGS NMD) EROS Data Center (EDC) has a program to produce 1-km resolution AVHRR time series data sets for the conterminous U. S., Alaska, and Eurasia as products for applied research (Eidenshink *et al.*, 1991; Kelly and Hood, 1991; Sadowski, 1990). The EDC has direct reception capabilities for NOAA's TIROS

*Visiting scientist at the EROS Data Center.

PHOTOGRAMMETRIC ENGINEERING & REMOTE SENSING,
Vol. 57, No. 11, November 1991, pp. 1453–1463.
0099-1112/91/5711–1453$03.00/0

series of polar-orbiting satellites (AVHRR HRPT data) covering most of North America. The EDC also operates a Domestic Communications Satellite System (DOMSAT) downlink that facilitates near-real-time access to virtually all of the AVHRR LAC data collected globally. AVHRR data reception activities are integrated with georegistration, product generation (such as greenness maps and land-cover classifications) and archiving systems developed to insure that high quality data will be available to researchers and land managers.

Because spatially extensive 1-km data sets possessing high temporal resolution have been unavailable, capabilities to use such data for regional land-cover characterization have not been well explored. The initial results of research being conducted by the EDC with the Center for Advanced Land Management Information Technologies (CALMIT) of the University of Nebraska at Lincoln (UNL) focus on the design and evaluation of strategies for detailed land-cover characterization over continental-size areas. Central to the research is the conviction that there is synergism in the integration of data derived by remote sensing with Earth science data acquired from other sources.

RESEARCH OBJECTIVES

The principal objective of the research has been to define and evaluate the potential for using AVHRR 1-km digital imagery and multisource data (such as broad-scale climate, terrain, ecoregions) in concert to characterize global land cover. The investigation included numerous questions involving methodology, data, and product requirements. Initial work focused on development of a prototype 1-km resolution land characteristics database for the conterminous U. S. that is designed for scientists dealing with global and mesoscale climate modeling, land surface change, and biosphere-atmosphere-hydrosphere interactions.

SCIENTIFIC HERITAGE OF THIS RESEARCH

Attempts to characterize land cover over large areas (subcontinental, continental, or global) using AVHRR data extend back at least 15 years. Most studies have focused upon GAC (4-km) or GVI (16-km) data rather than on full-resolution 1-km imagery. Typically, data are transformed to a vegetation index, such as the NDVI, for analysis. Tucker *et al.*, (1985), for example, used NDVI derived from GAC data to map major biomes and observe phenological change over the African continent for a 19-month period in 1982 and 1983. Three-week maximum vegetation index composites and principal components analysis were used to define major ecosystems. The authors observed qualitative agreement between their results and published maps, but argued for further development of analytical techniques and examination of multiple years to determine effects of short-term climatic variations.

Townshend *et al.*, (1987) employed GAC and GVI data in examining three approaches to classification of land cover in South America. They compared a principal-components transformation of 13 dates, a multidate greenness curve-matching methodology, and a maximum-likelihood classification approach. The last method yielded the best outcome. The optimal result was achieved when 13 dates of coverage (rather than fewer) were used. Available ground reference material allowed only qualitative judgment that the outcome of their classification was successful.

Goward *et al.*, (1985) examined GVI data for North America. They analyzed three-week composite maximum greenness (NDVI) images from April through November 1982 to map regions of net primary productivity. They showed that seasonal NDVI patterns could be associated with major land-cover regions, and that multidate greenness images could be used to observe patterns of vegetation growth and senecence. The authors recommended research on interannual change and further technique development. In later work Goward *et al.* (1987) compared the vegetation characteristics of North and South American biomes by analyzing GVI data using methods developed in Goward's 1985 research. They found that the differential timing and longer duration of the South American growing season was well captured. Biome distributions appeared, qualitatively, well-associated with published maps. Lloyd (1990) used a supervised binary decision tree classification approach to map world biomes with multidate GVI data. Although the spatial distributions appeared reasonable, no quantitative verification was possible. Gallo and Brown (1990) used biweekly GVI composites to examine global phytoclimatological conditions. They concluded that biweekly histograms of greenness change could be used to indicate general climatic conditions and associated vegetation distributions.

One-kilometre AVHRR data have been used less often than GAC or GVI data because they have not been generally available. Tucker et al., (1984), however, employed 1-km data to monitor vegetation conditions in the Nile delta. No attempt was made to classify land cover, but changes in greenness conditions from May to October 1981 were observed to correspond to known phenological circumstances and agricultural practices. Gervin *et al.* (1985) compared 1-km data acquired over the Washington, D. C. area to Landsat MSS data. They performed unsupervised classification of single-date images collected in July 1981 to identify Anderson Level I land cover and land use. The first four channels of the AVHRR were used rather than a vegetation index. Accuracies of classification were similar for predominant land-use and land-cover classes, but the MSS classification had higher accuracy on classes that were spatially heterogeneous or of limited spatial extent. Overall accuracy was 71.9 percent for the AVHRR and 76.8 percent for the MSS. The authors concluded that additional work on AVHRR data classification was warranted.

DESIGN CONSIDERATIONS

Discussions of land-cover mapping often lead to debate over classification schemes, assignment of class descriptors and labels, and product specifications. Most classification schemes are designed to be useful for a rather narrow range of applications; conversely, no single classification scheme can satisfy all, or even most, applications. The International Geosphere-Biosphere Programme (IGBP), following a year-long discussion of appropriate land-cover products for global change applications, concluded that "...the varied requirements for the IGBP cannot be satisfied by a single map of one set of attributes..." (International Geosphere-Biosphere Programme, 1990).

A number of studies have indicated that it is possible to produce databases of land characteristics that can satisfy a wide range of applications assuming proper methodological design (Loveland, 1984; Fitzpatrick-Lins *et al.*, 1987). This study is directed to the matter of appropriate design. Five major principles were established to guide the land characterization research (Loveland and Ohlen, 1991). Data analysis strategies and methods developed had to be

- Applicable and repeatable over continental and larger areas;
- Capable of discerning significant seasonal, ecological, and cultural variations in land cover;
- Applicable to very large data sets;
- Able to deal with data varying in quality; and,
- Capable of producing results applicable to various studies.

In keeping with these principles, the initial conceptual strategy, through the use of geographic information system (GIS)-based tools, allows examination of relationships between spatial data sets to characterize land cover, yet relies upon relatively simple methods for image segmentation (Figure 1).

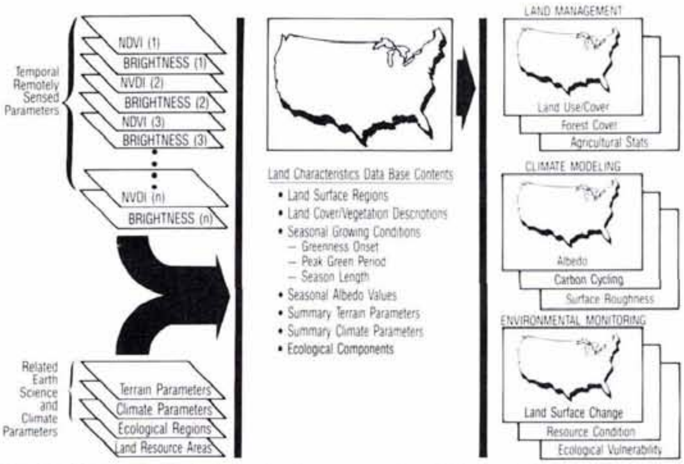


FIG. 1. Conceptual strategy for large-area land characterization includes use of remote sensing and multisource data to create a spatial database that includes seasonally distinct land-cover regions and associated attributes that can be tailored to a number of disparate applications.

Very large data sets present unique image analysis problems. Continental areas typically exhibit greater variations in climate, terrain, and vegetation than are encountered in analyses of single scenes. In addition, data set sizes tend to be extraordinarily large, posing computational and data handling difficulties.

Such problems can be dealt with in two ways. First, the study area can be partitioned into smaller data sets based on, for example, climatic or ecological regions. This would serve to limit the environmental diversity, but would likely create significant post-classification mosaicking and interregional class correlation problems. An alternative solution would be to treat the data set as a single unit. The latter approach was used in this research.

Previous studies suggest that multitemporal, multisource image classification techniques are required for large-area land-cover characterization. Single-date analyses, especially using AVHRR data, are frequently inadequate for discriminating land-cover types because disparate cover types can share spectral reflectance characteristics. The problems are compounded when one deals with large areas exhibiting great climatic, topographic, and ecological diversity. Classification of multitemporal AVHRR-NDVI data should have advantages over single-date observations, though some cover parameters required for global analyses are still likely to be imperfectly characterized. Ancillary data, such as elevation, climate variables, ecological regions are, therefore, considered critical in land cover description.

The prototype land characteristics database has several components:

- "Seasonally distinct" land-cover regions defined employing analysis of AVHRR and ancillary data. These regions exhibit unique phenological characteristics, such as time of onset, magnitude of peak, and seasonal duration of greenness, and possess relatively homogeneous vegetative associations.
- Attributes (or spreadsheets) that describe the characteristics of the landscape regions. Attributes contained in the U.S. prototype are (1) descriptions of vegetation composition and physiognomy; (2) quantitative seasonal characteristics including mean monthly NDVI (March-October 1990) and seasonal parameters (time of onset, magnitude of peak, duration of greenness, and total greenness); (3) site characteristics including, for every pixel, elevation, climate, and ecoregion and Major Land Resources Area (MLRA) membership; (4) translation tables linking the regions to common land-cover classification schemes such as UNESCO, USGS Anderson System, and the vegetation types used in the Simple Biosphere Model and the Biosphere Atmospheric Transfer Scheme (Dickinson *et al.*, 1986); and (5) summary data on climate, terrain, land use and land cover derived from publications describing U.S. ecological regions and MLRA's, and from sampled digital USGS land-use and land-cover data. The strategy of this approach is to give researchers a capability to compute new parameters, derive new classifications and aggregations of the data to suit specific needs, and develop custom products. This provides the flexibility that may allow the land characteristics database to be used in many models without extensive modification of inputs.
- All source data used to produce the first two components are also included to provide further customization.

As research evolves, other attributes will be added to the land characteristics data base. For example, measurements of surface albedo, primary production estimates, and other surface properties associated with canopy resistance may be added as concensus methods for their calculations are established.

DATA SOURCES

AVHRR DATA

Daily observations of NOAA-11 1-km data were calibrated to reflectance measurements, scaled to byte data, and georegistered to a Lambert Azimuthal Equal Area map projection (Kelly and Hood, 1991; Holben, 1986). The resulting data set dimensions are 2889 rows by 4587 columns.

Seventeen biweekly maximum NDVI composites were generated for the period of March through October 1990. This process involved the creation of a composite image in which the pixel having the maximum NDVI for each composite period was retained (Eidenshink *et al.*, 1991). By selecting for maximum NDVI, nearly cloud-free data sets usually result. An image-to-image registration process was used to assure accuracy within a root-mean-square error of 1 pixel (Kelly and Hood, 1991).

Initial experiments using 1989 biweekly NDVI composites of the western United States suggested that the use of monthly composites would both minimize data volume problems and computational demands without unduly affecting results. Consequently, the 1990 biweekly composites were reduced to eight monthly composites of maximum NDVI. The original biweekly data were, however, retained for use in region characterization. Data quality was improved by the monthly compositing through the elimination of much of the remnant atmospheric, cloud, and off-nadir contamination in the biweekly composites. Although previous studies by other investigators with GAC or GVI data have documented improved classification results as more frequent observations are used (Townshend *et al.*, 1987), practical considerations argue for dimensionality reduction in continental studies using 1-km data.

TERRAIN DATA

Digital elevation data incorporated in the database were originally derived by the Defense Mapping Agency from 1- by 2-degree topographic maps, and were later refined by the National Telecommunications and Information Administration. These data are now distributed by the NOAA National Geophysical Data Center in Boulder, Colorado. The elevation values are rounded estimates to the nearest 20 feet for every 30 seconds of latitude and longitude.

CLIMATE DATA

Climate data layers, including length-of-frost-free-period, average annual precipitation, average monthly precipitation, and monthly mean temperature, were digitized from climate atlas maps (NOAA, 1979). All of the maps were based on long-term means of temperature and precipitation (for example, monthly precipitation from 1931 to 1960). The scales of these maps varied from approximately 1:7,000,000 to 1:18,000,000. Digitized isoline data were subsequently interpolated to a gridded surface. Because of the generalized nature of the source maps, these data relate to continental climate conditions and do not represent local or microclimate conditions.

Ecoregions

Ecoregion maps from the U.S. Environmental Protection Agency (Omernik, 1987; Omernik and Gallant, 1990) were digitized and attributes of the regions (land surface form, major soils, land use and potential natural vegetation) were summarized for use in characterization.

Major Land Resource Areas

Major land resource area (MLRA) regional boundaries were digitized from a 1:7,500,000-scale map published by the USDA Soil Conservation Service (USDA SCS, 1981). MLRA region attributes include soils, terrain, climate, potential natural vegetation, and land use.

Land-Use and Land-Cover Data

Land-use and land-cover (LULC) data were sampled from digital land-use and land-cover files obtained from the USGS (U.S. Geological Survey, 1986). These data, classified at Anderson Level II (Anderson et al., 1976), have been developed by the USGS over the past 20 years from visual analyses of aerial photography. The data are keyed to 1:250,000-scale USGS 1- by 2-degree quadrangles. Fifty-one quadrangle-based data sets were converted to a 1-km grid for use in the research. The quads, selected to sample major ecosystems, cover approximately 12 percent of the U.S.

Political Boundaries

State and county political boundaries from the USGS 1:2,000,000-scale digital line graph national data base were used as reference during the investigation (U.S. Geological Survey, 1990).

Water Mask

Surface water bodies were separated using Channel 2 data from daily AVHRR scenes. Cloud-free scenes were selected through a visual quality assessment of imagery. After a threshold between land and water values was identified, a binary mask was computed and the water bodies data set was added to a land characteristics database. Approximately 50 AVHRR scenes were used to create the mask.

Other Data

Many other supporting materials, including state, regional and national land use and land cover maps, vegetation maps, atlases, agricultural statistics, and crop calendars, were used.

ANALYTIC METHODS

The strategy developed to characterize U.S. land cover employed both AVHRR and ancillary data in a carefully structured manner (Figure 2). Analytic procedures involved overlaying, exploring, and interrelating the disparate spatial data and attributes.

Preliminary Experiments

The image analysis methodology used in the development of the 1990 conterminous U.S. land characteristics database evolved from a series of classification experiments conducted using 1989 AVHRR NDVI data covering the western U.S. These tests indicated that (1) an initial vegetated/barren land stratification would be required, (2) a minimum of 50 spectral-temporal classes would be required to define important land cover types, (3) unsupervised classification was suitable, and (4) the use of monthly rather than biweekly NDVI composites would yield acceptable results.

The vegetated/barren land stratification was used to ensure that classes exhibiting high intraclass variance, such as, water, bare soil, clouds, and snow/ice, would not dominate the clustering process. Masking of these classes optimizes the spectral discrimination of the classes directly associated with vegetation. Separation and characterization of non-vegetated areas is not, in any case, reliable with NDVI data because of insensitivity of this transformation to low-biomass conditions. Plans call for the characterization of non-vegetated areas using a brightness measure.

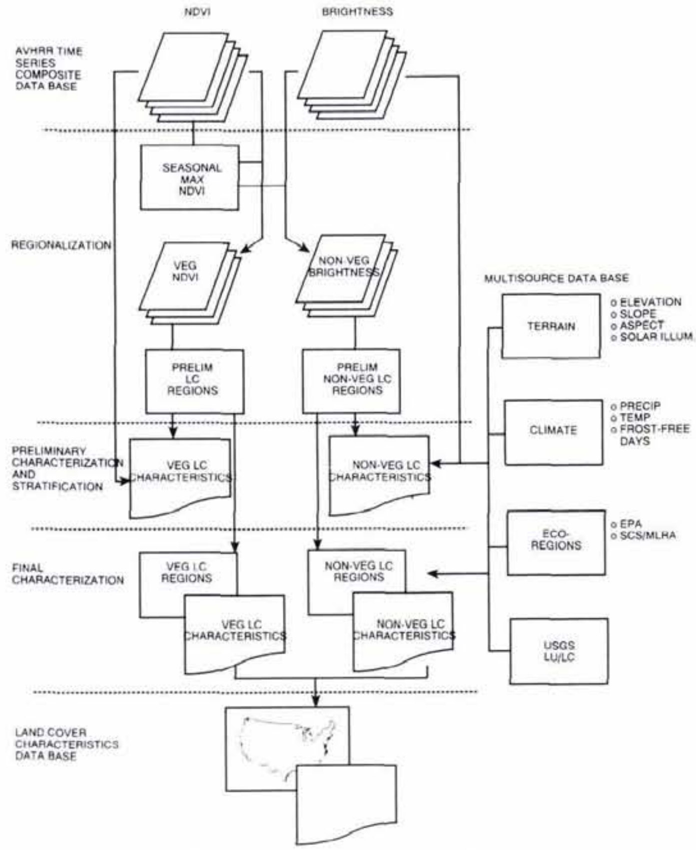


FIG. 2. Processing flow for the development of the prototype land characteristics database. Note that the analysis of brightness data is planned but not yet completed.

Image Classification

Vegetated and non-vegetated land were stratified by analyzing a maximum NDVI composite spanning the March to October 1990 period. Through visual interpretation, an NDVI threshold of 0.09 was selected to separate vegetated and non-vegetated lands. The threshold was determined by comparison of the strata to available maps and imagery, and published data on NDVI-land cover relationships.

An unsupervised clustering algorithm (ISOCLASS) and minimum-distance-to-mean classification methodology was used to define 70 spectral-temporal ("seasonally distinct") classes within the vegetated stratum (Plate 1). A 20 percent systematic sample of the eight monthly composites was employed to derive cluster statistics.

Land-Cover Characterization

Initial evaluation, labeling, and characterization of the 70 classes was based on a combination of graphic, statistical, and visual tools and techniques. For example, graphs portraying the variation of mean NDVI over the 8-month analysis period yielded a profile of the phenology of each class (Figure 3). The shapes of NDVI multitemporal curves often could be used to identify land cover, and comparisons between curves helped in identifying related classes when analyzed in concert with a display of the spatial distribution of each class. Maps, atlases, agricultural statistics, and Landsat image maps were also used in interpretation of classification results.

Graphical summaries of elevation and frost-free period statistics for each cluster (Figure 4) enabled association of the spatial distribution of each class with site characteristics. Ecoregion and MLRA boundaries were overlaid on the 70-class data set; spatial interrelationships between the data sets were computed; and tables depicting the associations were constructed (Tables 1 and 2). Similar summaries were developed showing the association among the 70 classes and sampled USGS data (Table 3).

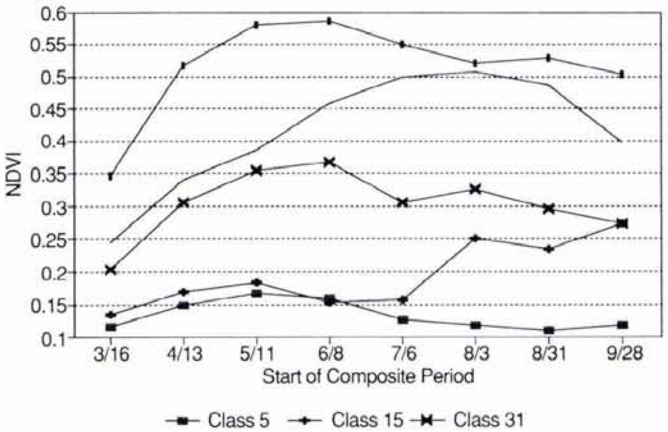
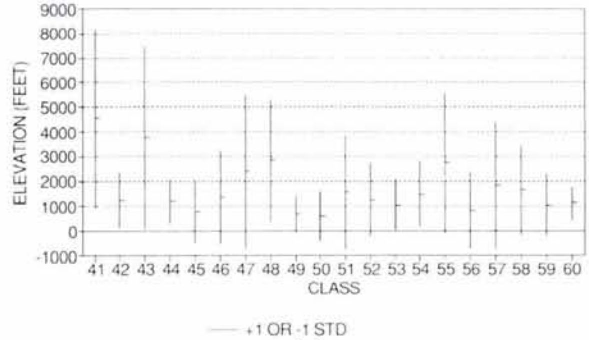


FIG. 3. Example of cluster class NDVI mean values for selected classes.



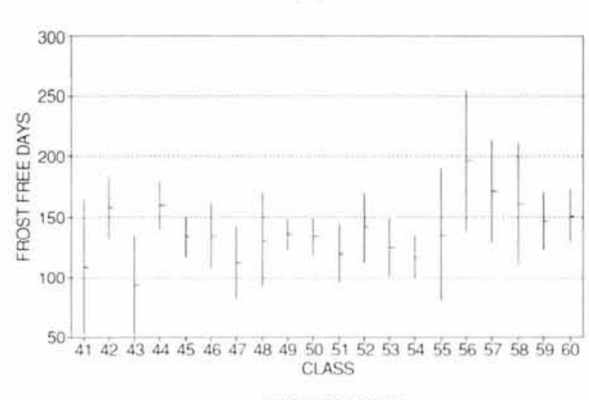


FIG. 4. Statistical relationships between vegetation greenness classes and (a) elevation and (b) frost-free period.

The tables indicate the percentage of each of the 70 "seasonally distinct" AVHRR-derived classes falling within MLRA and ecoregion classes, and associated data. Attributes of the ecoregion, MLRA, and LULC data were not considered "ground truth," but were used as aids in understanding site factors, describing land cover, and identifying instances of confusion in the classification. They also provided an opportunity for comparison of alternative methods of landscape regionalization and characterization.

Finally, interpretive maps were developed to portray, respectively, (1) the month in which the NDVI first rose above a threshold value (onset of greenness), (2) the month in which maximum NDVI occurred (peak of greenness), (3) the number of days when the NDVI reached or exceeded a threshold value (duration of greenness), and (4) the cumulative value of the NDVI (total NDVI) for March through October (Plate 2). These maps were derived through analysis of individual class NDVI statistics produced from the original 17 biweekly NDVI composites. Interpretation of temporal NDVI means led to the identification of the four interpretive maps (Figure 5). The four factors are strongly related to the phenologic cycle of vegetation. The month in which the NDVI increases dramatically corresponds to the time of emergence of green vegetation at the beginning of the growing season. The month of maximum NDVI reflects the time of maximum photosynthetic activity (Lloyd, 1990). The time that the NDVI exceeds a certain threshold value is similar to the length of the growing season (Lloyd, 1990; Brown, 1990). The cumulative NDVI through the growing season generally reflects total photosynthetic activity or net primary productivity (Goward *et al.*, 1987; Brown, 1990).

POSTCLASSIFICATION REFINEMENT

As expected, a number of instances of classification confusion were observed. These were instances in which the 70 classes were not uniquely associated with a single cover type. They provided considerable insight about phenological patterns of the U.S. through the process of observing the types and distri-

TABLE 1. ECOLOGICAL REGIONS AND THEIR CHARACTERISTICS THAT CORRESPOND TO CLASS 53

Class: 53 Percent: 75.4	Name	Northern lakes and forest
	Landform	Smooth to irregular plains, plains with hills
	PNV	Great Lakes Spruce/Fir, Pine/ Northern hardwood
	Landuse	Forest and woodland mostly ungrazed
	Soils	Podzolic
Class: 53 Percent: 5.2	Name	North central hardwood forest
	Landform	Irregular plains
	PNV	Maple/Basswood, Northern hardwoods
	Landuse	Cropland with pasture, woodland, and forests
	Soils	Podzolic
Class: 53 Percent: 4.3	Name	Northeastern highlands
	Landform	Low mountains, open low mountains
	PNV	Northern hardwoods/Spruce
	Landuse	Forest and woodland mostly ungrazed
	Soils	Spodosols

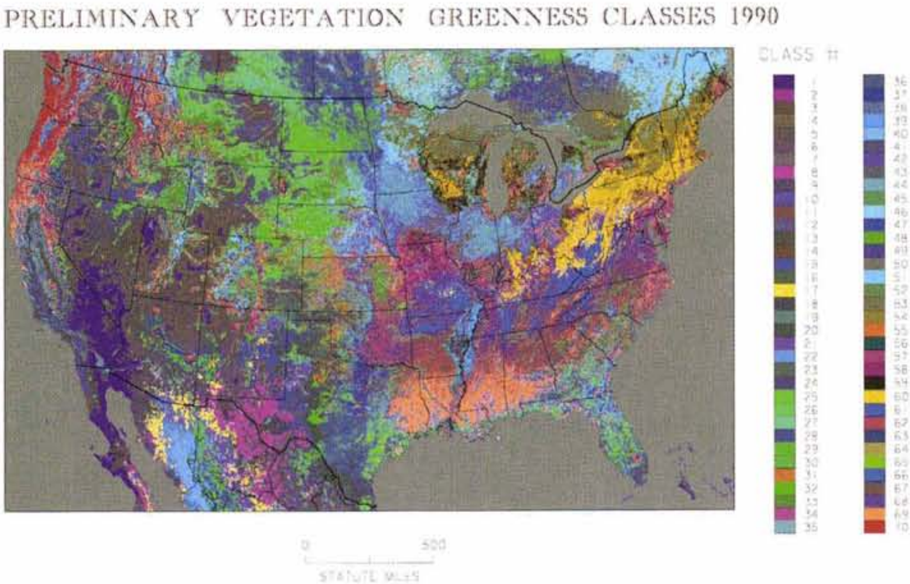


PLATE 1. Preliminary 1990 vegetation greenness classes derived from unsupervised classification of March-October monthly AVHRR NDVI composites.

TABLE 2. MAJOR LAND RESOURCE AREAS AND THEIR CHARACTERISTICS THAT CORRESPOND TO CLASS 53

Class	53	Name	Superior Stony and Rocky Loamy Plains
Percent	25.9	Landuse	Forest (80%); crops/pasture (10%); cranberries
		Elev.	300-600M
		Topo.	Undulating to rolling glacial drift plains
		AAP	750mm; majority during growing season
		AAT	2-6 degrees C
		AFFP	80-140 days
		Soils	Orthods
		PNV	Northern hardwood/Pine; Spruce/Larch/Sphagnum Bogs
Class	53	Name	Central Wisconsin and Minnesota Thin Loess and Till
Percent	14.4	Landuse	Forage/feed grains (25%); pasture (15%); forests (60%)
		Elev.	300-500m
		Topo.	Level to rolling till plains mantled with Loess
		AAP	625-750mm; majority in growing season
		AAT	4-7 degrees C
		AFFP	120-140 days
		Soils	Boralfs
		PNV	Mixed northern hardwoods (Oak/Maple/Ash/Elm/Basswood)
Class	53	Name	North Michigan and Wisconsin Sandy Drift
Percent	9.5	Landuse	Forest/lumber; forage/feed crops; cranberries
		Elev.	200-500m
		Topo.	Morainic hills and glacial drift plains
		AAP	675-850mm; minimum in winter
		AAT	4-7 degrees C
		AFFP	120-140 days
		Soils	Orthods or Saprists
		PNV	Decid (Sugar Maple/Birch/Beech/Hemlock)/Jack/Red Pine
Class	53	Name	Northern Minnesota Gray Drift
Percent	6.8	Landuse	forest/timber (50%); forage/feed grains (50%)
		Elev.	300-500m
		Topo.	Rolling glacial Moraine and outwash
		AAP	525-675mm; majority during growing season
		AAT	3-6 degrees C
		AFFP	100-120 days
		Soils	Boralfs, Aqualfs, Fibrists
		PNV	Forest (Aspen, Northern hardwoods, White Spruce, Blackberry)

butions of confusion. Such information will be useful in future attempts to refine and improve the classification strategy.

Examples of confused land cover are warm season desert grasslands and alpine meadows (class 9). The late "greenup," moderate peak greenness, and short duration of greenness exhibited by grasslands in arid regions receiving limited mid-summer precipitation mimics the phenology of alpine meadows occuring at high elevations. In these instances, elevation and frost-free period data were used for stratification.

In another instance, classes were observed to occur both in areas of the southern Great Plains dominated by winter wheat, and in coastal California where they were associated with cool season grasslands (class 35). Consideration of regional variables led to explanation of this confusion. In the southern Plains,

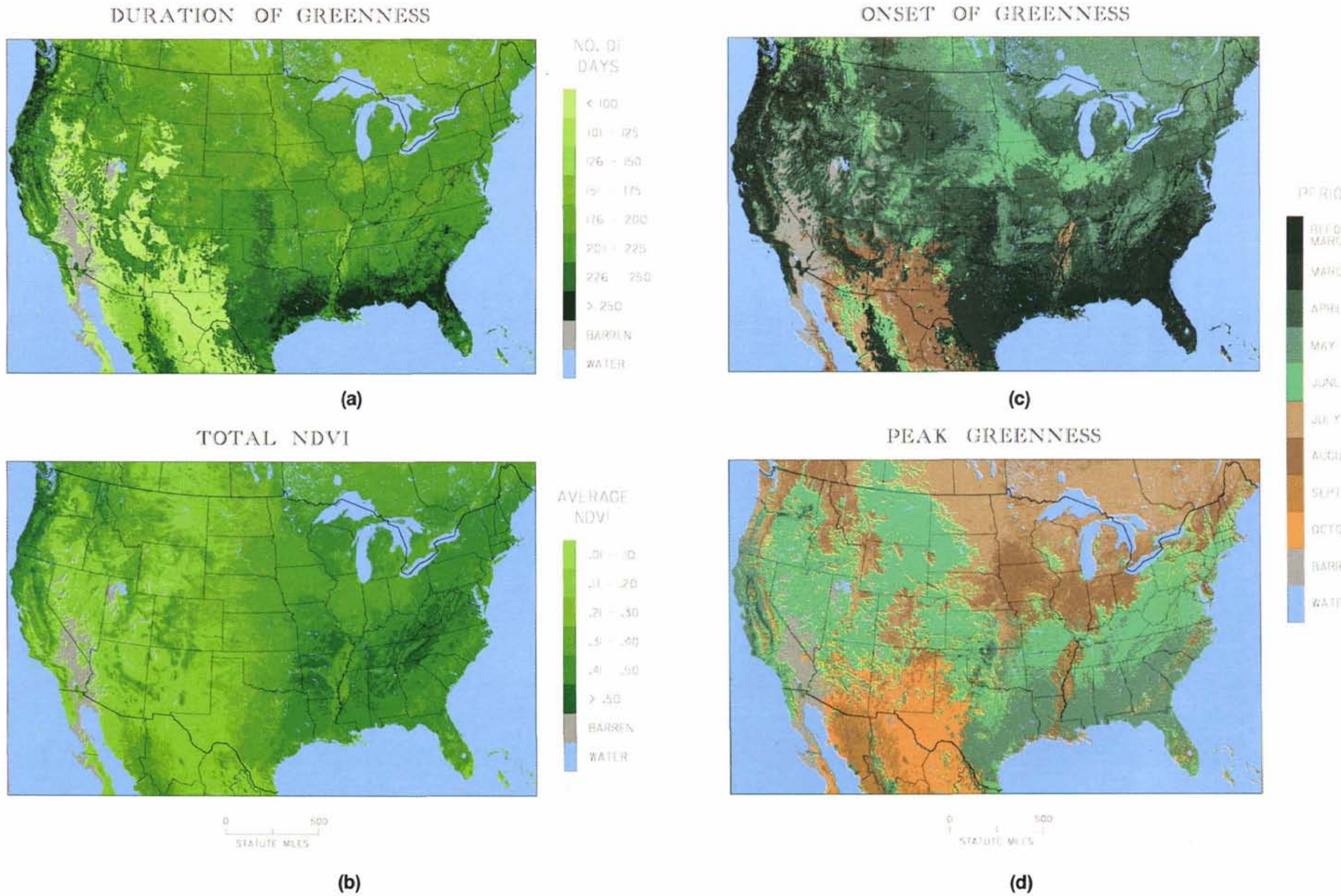


PLATE 2. Seasonal parameters calculated for preliminary vegetation greenness classes include (a) onset of greenness, (b) period of peak greenness, (c) duration of green period, and (d) total NDVI.

TABLE 3. USGS LAND-USE AND LAND-COVER CATEGORIES AND PROPORTIONS FOUND IN CLASS 53

Class	Anderson Code	Percent	Level II Category
53	43	46.43	Mixed forest land
53	41	21.78	Deciduous forest land
53	61	14.93	Forested wetlands
53	21	6.55	Cropland and pasture

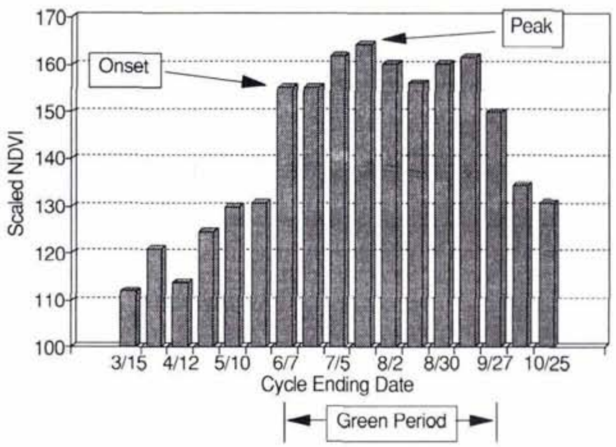


FIG. 5. Example of relationship between NDVI class temporal means (class 53) and selected seasonal parameters.

winter wheat fields "greenup" quickly in April and May, senesce, and are harvested in June. On the west coast, the unique timing of precipitation – winter maximum – influences a similar phenologic pattern in grasses, a type of vegetation similar to wheat in physiognomy and biomass. In this case, the unique ecological characteristics of the two cover types led to the use of ecoregions for resolving confusion.

The postclassification refinement criteria were developed using interactive spatial and graphical comparison techniques. The methodology involved spatial display of each of the 70 classes with histograms of class relationships to ancillary variables such as elevation, ecosystem, and frost-free period. Through interactive selection of minimum and maximum threshold values of the ancillary data, the affected pixels within each class display would be alarmed. Thus, the pixels displayed in specific classes were highlighted in real time, reflecting the effects of selecting a particular threshold value.

Through analytical processes such as those set forth above, 75 percent of the original 70 preliminary vegetation greenness classes were subdivided into 171 seasonally distinct land-cover regions. The final characterization of the 171 classes was then completed, with the development of the descriptive and quantitative attributes of each region.

VERIFICATION

Classification accuracy is a complex issue. The coarse resolution of AVHRR data leads to the development of classes based commonly on land cover mosaics rather than on homogenous landscape regions. The accessibility of consistent site data for verification is also a limitation. An additional complication is the fact that the land characteristics data base is not based on well-defined categories. As a result, verification was limited to comparisons with other data sets such as ecoregions, MLRA's, and LULC.

LINKAGES TO OTHER CLASSIFICATION SYSTEMS

The final step in the prototype effort was to link the AVHRR-based classification and data to other commonly used classification systems (Table 4). Efforts are underway to develop relationships with the USGS (Anderson, 1976), the UNESCO vegetation classification system, and the vegetation types used in the Simple Biosphere Model (SiB) and the Biosphere Atmosphere Transfer Scheme (BATS).

RESULTS AND DISCUSSION

In general, homogeneous land-cover regions were well identified if they comprised relatively large, regular landscape patches. In spatially complex areas, such as the eastern U.S., seasonally distinct land-cover regions were more often correlated with mosaics of land cover having variable physiognomic and vegetative characteristics.

Rangeland classes tended to have seasonal minimum NDVI values of approximately 0.10, and seasonal maxima near 0.30, usually not exceeding 0.40. These regions have low percentage cover and low standing biomass. Many rangeland or grassland classes displayed a dispersed non-contiguous pattern in locations adjacent to and interspersed with forest land cover such as subalpine zones, or agricultural classes. For example, Class 4 has extensive coverage west of the Rocky Mountains, including some contiguous regions, especially in southwestern Wyoming and northeastern Arizona. There are also, however, widespread areas of scattered small occurrences throughout Nevada. These appear to correspond to sagebrush steppe cover in basins between forest-covered mountains.

This biome tends to be confused with eastern urban areas and coastal mixed pixels. In the Southwest, some confusion occurs with mid-elevation open stands of pinyon/juniper woodland having a grass understory. Two differing grassland phenologies, cool season and warm season, contribute to some confusion. For example, Class 14 encompasses both cool-season grasses in California and winter wheat in both Oklahoma and Oregon (Figure 6). Alpine meadows tend to be grouped with other warm-season rangeland.

Regionally distinct patterns representing forest lands were well identified in the classification. Class 54, for example, represents mixed forest land cover (maple/birch/beech with spruce/fir species) of the northeastern mountains and foothills. Class 53 is also primarily northern forests, but, in this case, corresponds to Great Lakes deciduous hardwoods (maple/birch). Class 61 represents a deciduous forest cover of oak/hickory within the Ozark-Boston Mountains and southern Appalachians. The unique hemlock/Douglas fir forests of the northwestern United States are represented by class 70. Figure 7 provides monthly NDVI characteristics for these four classes.

Major agricultural regions are clearly identifiable, and NDVI profiles for agricultural classes reveal much about phenology and crop types (Figure 8). For example, winter wheat regions (class 35) in the southern Great Plains are clearly distinguished from spring wheat (class 30) in the northern Great Plains by the different period of greenness onset. Class 44 corresponds to the corn and soybeans regions of the Midwest (Iowa, Illinois, Indiana). Class 43 also is distributed throughout the Midwest, but represents a more mixed landscape with oats, woodlands, and pasture land cover interspersed with corn and soybeans. The NDVI curves for these two classes differ slightly, with class 43 displaying a lower peak green level. It also displays a less rapid greenup rate, which is likely caused by the earlier green-up of the non-corn and soybean elements of the landscape.

The preliminary evaluation indicates that the procedures used are, for the most, part acceptable. However, the research has illuminated many issues that remain to be addressed. For example, the outcome of the NDVI-based classification was clearly

TABLE 4. RELATIONSHIP BETWEEN SELECTED SEASONALLY DISTINCT LAND-COVER REGIONS AND OTHER CLASSIFICATION LEGENDS. PRELIMINARY CLASSIFICATION — CONTERMINOUS U.S.

Class	Typical Vegetation	Anderson LevelII	UNESCO Vegetation
3	Saltbrush, Great Basin Sage, Greasewood	32 Shrub and Brush Rangeland	Extremely Xeromorphic Shrubland
6	Grama/Galleta, Bur Sage, Greasewoods	32 Shrub and Brush Rangeland	Mainly Deciduous Scrub
10	Saltbrush, Great Basin Sage, Greasewoods	32 Shrub and Brush Rangeland	Mainly Deciduous Scrub
15	Grama/Buffalo Grass	33 Mixed Rangeland	Short Grassland
16	Pinyon-Juniper, Blackbrush	42 Evergreen Forest Land	Mainly Evergreen Woodland
22	Subalpine Spruce/Fir, Alpine Meadows	42 Evergreen Forest Land	Mainly Evergreen Scrub
26	Wheatgrass/Needlegrass	31 Herbaceous Rangeland	Medium Tall Grassland
37	Ponderosa Pine, Pinyon-Juniper	42 Evergreen Forest Land	Mainly Evergreen Woodland
41	Cedar/Hemlock/Spruce/Fir/Pine	42 Evergreen Forest Land	Mainly Evergreen Forest
53	Maple, Beech, Birch	41 Deciduous Forest Land	Mainly Deciduous Forest
54	Aspen	41 Deciduous Forest Land	Mainly Deciduous Forest
63	Ponderosa Pine	42 Evergreen Forest Land	Mainly Evergreen Forest

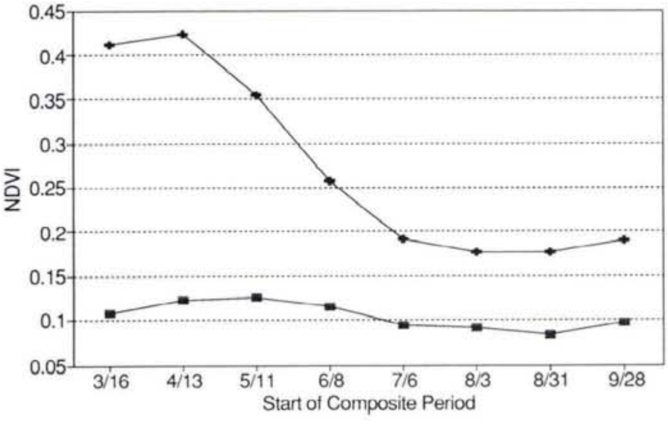


FIG. 6. Monthly NDVI means for selected rangeland categories.

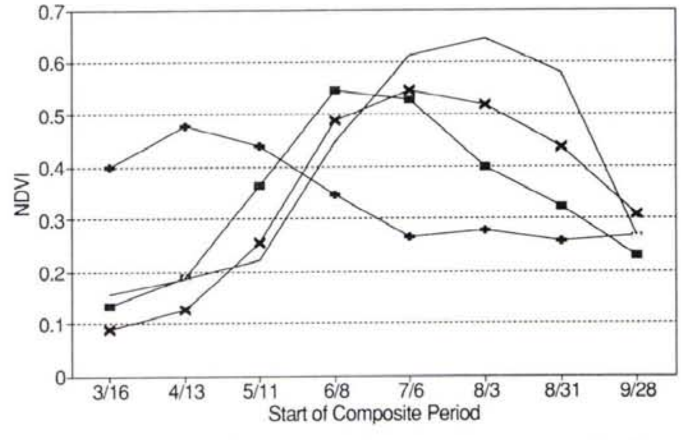


FIG. 8. Monthly NDVI means for selected agricultural categories.

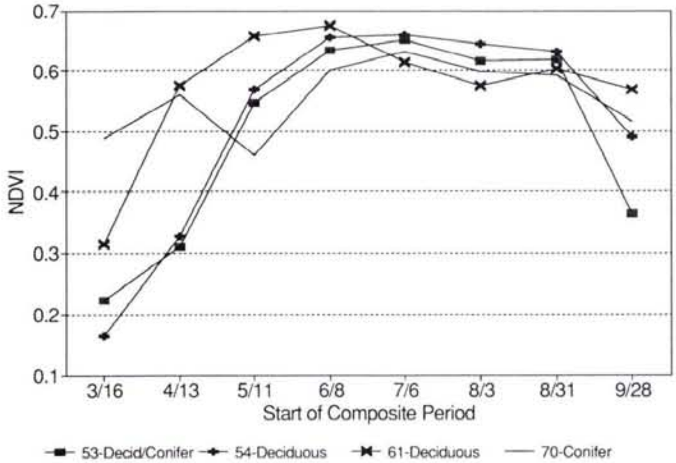


FIG. 7. Monthly NDVI means for selected forest categories.

influenced by the weather during 1990. The preliminary work carried out using the 1989 western United States data set resulted in classes corresponding to drought conditions in the Great Plains during that year. California experienced similar conditions during 1990, which undoubtedly has affected the 1990 classification. The specific effects of climatic anomalies on classification of land cover are, however, uncertain and remain to be investigated.

It is also likely that the classification was affected by the availability of AVHRR data for 1990. The fact that there were no seasonal observations from late-November to March very likely reduced the ability to discern some important cover types in the southeastern United States. The addition of winter composites must be part of future work. It is also not clear whether eight monthly observations are required to characterize land cover. Similar results may be derived from analysis of fewer composites selected at phenologically critical times.

Some cover types cannot be adequately identified using NDVI data. Barren lands, snow and ice, and water bodies have similar NDVI characteristics because of the absence of photosynthetically active plant material. The coarse resolution of AVHRR data in relation to the typically small landscape patches that comprise cover types of limited areal extents, such as wetlands, was limiting. Urban areas could not be uniquely identified because of the complex mixtures of surface conditions within 1-km urban pixels.

The strategy to employ ancillary data in postclassification stratification of the 70 preliminary vegetation greenness classes served to identify some important problems in working with data sets covering large areas. One such problem is exemplified by the case of warm-season desert grassland and alpine cover confusion discussed earlier. Although the initial supposition was that stratification was possible using elevation thresholds, in practice this worked only in local circumstances. Because of the related effects of altitude and latitude on vegetation phenology, the elevation threshold needed to split these classes had to be continually lowered, moving north to south, to achieve acceptable results. In other words, the elevation threshold actually proved to be very difficult to apply. Instead, climate variables such as length of frost-free period, were found to be more effective in postclassification refinement.

Verification of classification results has also presented problems, though these are not unique to this research. In fact, little has been reported on quantitative accuracy assessment of land-cover products derived from analysis of AVHRR data. Tucker *et al.* (1985) note that such work is hampered by the dearth of suitable "ground truth" and lack of agreement between the few extant land cover maps covering the continents. Townshend *et al.* (1987) assert that because existing maps of land cover and land use have been developed differently from AVHRR land-cover databases, they may not even be acceptable standards of reference where available.

Land characteristics databases derived from classification of AVHRR data may produce unconventional regional definitions that do not match classifications used in existing maps, but may be useful. Experience from this study suggests that often the spatial resolution, and probably also the classification precision, of the AVHRR-derived data are higher than those of existing maps. Adequate methods to verify 1-km resolution land-cover classifications conducted over continental-sized areas do not exist. Standards of reference, when they exist at appropriate resolution, are frequently old or have incompatible classifications or other problems (Matthews, 1983). Research is necessary to re-examine conventional image classification accuracy assessment. To employ methods such as those reported, accuracy requirements for global land-cover inventory and monitoring must be established along with a definition of procedures for gauging the quality of land-cover data from coarse-resolution satellite data.

FUTURE RESEARCH DIRECTIONS

The research results reported here are preliminary. Although these findings represent milestones, problems that limit current efforts to characterize continental land cover using multitemporal AVHRR 1-km imagery and ancillary data require work. Some of the important areas in which research is needed include

- Assessment of the effects of seasonal and annual variations on identification and characterization of land-cover regions; year-to-year effects of weather and climate on the development of seasonally distinct land-cover regions, and single-year variation of vegetation and its impacts on determination of the appropriate sample period (biweekly, monthly, or seasonal) for temporally-based classification.
- Identification of influences of landscape-sensor interaction on the definition and characterization of land-cover regions.
- Refinement of data analysis methods, including integration of data from other sensors, use of brightness measures in characterizing unvegetated areas, and potential use of AVHRR thermal channels in land-cover classification.
- Development of verification strategies appropriate for continental-scale land-cover data.

The research suggests that 1-km resolution multitemporal AVHRR/NDVI data employed in concert with ancillary data can be used to characterize land cover over very large areas. Successful land-cover characterization and database development alone, however, are insufficient. The databases must be useful to the global change community and others. Therefore, an important component of future work must be to address specific needs for products.

ACKNOWLEDGMENTS

Most financial support for the participation of Dr. Merchant and Ms. Brown in this research was provided by the Conservation and Survey Division (CSD), University of Nebraska-Lincoln. Special thanks are extended to Dr. Perry B. Wigley, CSD Director. A portion of Dr. Merchant's and Ms. Brown's participation was supported by a U.S. Environmental Protection Agency (EPA) grant-number X007526-01.

REFERENCES

Anderson, J. R., E. E. Hardy, J. T. Roach, and R. E. Witmer, 1976. *A Land Use and Land Cover Classification System for Use with Remote Sensor Data*: U.S. Geological Survey Professional Paper 964, 28 p.

Brown, J., 1990. *Vegetation Discrimination Using AVHRR Global Vegetation Index Data*: M.A. Thesis, University of Nebraska-Lincoln.

Dale, V. H., 1990. *Report of a Workshop on Using Remote Sensing to Estimate Land Use Change*: ORNL Environmental Sciences Division Publication No. 3397, Oak Ridge National Laboratory, Oak Ridge, Tennessee.

Dickinson, R. E., A. Henderson-Sellers, P. J. Kennedy, and M. F. Wilson, 1986. *Biosphere-Atmosphere Transfer Scheme (BATS) for the NCAR Community Climate Model*: NCAR Technical Note 275+STR, National Center for Atmospheric Research, Boulder, Colorado.

Dorman, J. L. and P. J. Sellers, 1989. A Global Climatology of Albedo, Roughness Length and Stomatal Resistance for Atmospheric General Circulation Models as Represented by the Simple Biosphere Model (SiB): *Journal of Applied Meteorology*, Vol. 28, pp. 833–855.

Eidenshink, J. C., R. E. Burgan, and R. H. Haas, 1991. Monitoring Fire Fuels Conditions by Using Time-Series Composites of Advanced Very High Resolution Radiometer Data, *Proceedings, Resource Technology International Symposium on Advanced Technology in Natural Resource Management*, Second, Washington, D.C., November 1990, pp. 68–82.

Fitzpatrick-Lins, K., E. F. Doughty, M. B. Shasby, T. R. Loveland, and S. Benjamin, 1987. Producing Alaska Interim Land Cover Maps from Landsat Digital and Ancillary Data, *Proceedings, Pecora XI Symposium, Satellite Land Remote Sensing: Current Problems and a Look to the Future*, Sioux Falls, South Dakota, pp. 339–347.

Gallo, K. P. and J. F. Brown, 1990. Satellite-Derived Indices for Monitoring Global Phytoclimatology, *Proceedings of the International Geoscience and Remote Sensing Symposium*, Piscataway, New Jersey, pp. 261–264.

Gervin, J. C., A. G. Kerber, R. G. Witt, Y. C. Lu, and R. Sekhon, 1985. Comparison of Level 1 Land Cover Classification Accuracy for MSS and AVHRR Data: *International Journal of Remote Sensing*, Vol. 6, No. 1, pp. 47–57.

Goward, S. N., D. G. Dye, A. Kerber, and V. Kalb, 1987. Comparison of North and South American Biomes from AVHRR Observations: *Geocarto International*, Vol. 1, pp. 27–39.

Goward, S. N., C. J. Tucker, and D. G. Dye, 1985. North American Vegetation Patterns Observed with the NOAA-7 Advanced Very High Resolution Radiometer: *Vegetatio*, Vol. 64, pp. 3–14.

Henderson-Sellers, A., M. F. Wilson, G. Thomas, and R. E. Dickinson, 1986. *Current Global Land-Surface Data Sets for Use in Climate-Related Studies*: NCAR Technical Note 272+STR, Boulder, Colorado, 110 p.

Holben, B., 1986, Characteristics of Maximum Value Composite Images from Temporal AVHRR Data, *International Journal of Remote Sensing*, Vol. 7, pp. 1417–1434.

International Geosphere-Biosphere Programme, 1990. *Global Change*: Report No. 12, Stockholm, Sweden.

Justice, C. O., J. R. G. Townshend, B. N. Holben, and C. J. Tucker, 1985. Analysis of the Phenology of Global Vegetation using Meteorological Satellite Data: *International Journal of Remote Sensing*, Vol. 6, pp. 1271–1318.

Kelly, G., and J. Hood, 1991. AVHRR Conterminous United States Reference Data Set: *Vol. 3, Technical Papers, ACSM-ASPRS Annual Convention*, Baltimore, Maryland, pp. 232–239.

Lloyd, D., 1991. A Phenological Classification of Terrestrial Vegetation Using Shortwave Vegetation Index Imagery: *International Journal of Remote Sensing*, Vol. 11, No. 12, pp. 2269–2279.

Loveland, T. R., 1984. *Copper River, Alaska Terrain and Land Cover Project – Final Report*: EROS Data Center Technical Report, Sioux Falls, South Dakota, 110 p.

Loveland, T. R., and D. O. Ohlen, 1991. A Strategy for Large-Area Land Characterization - the Conterminous U.S. Example, *Proceedings, U.S. Geological Survey Global Change Research Forum*, Reston, Virginia (in press).

Matthews, E., 1983. Global Vegetation and Land Use: New High Resolution Data Bases for Limited Studies: *Journal of Climatology and Applied Meteorology*, Vol. 22, pp. 474–487.

———, 1985. *Atlas of Archived Vegetation, Land Use and Seasonal Albedo Data Sets*: NASA Technical Memorandum 86199, Washington, D.C.

Miller, W. A., S. M. Howard, and D. G. Moore, 1988. Use of AVHRR Data in an Information System for Fire Management in the Western United States, *Proceedings, International Symposium on Remote Sensing of Environment*, 20th, Nairobi, Kenya, December 1986, Vol. 1, pp. 67–79.

National Oceanic and Atmospheric Service, 1979. *Climatic Atlas of the United States*: U.S. Dept. of Commerce, National Oceanic and Atmospheric Administration, Environmental Data Services, Ashville, North Carolina, 80 p.

Olson, J. S., and J. A. Watts, 1982. *Major World Ecosystem Complexes Map, Scale=1:30,000,000*: Oak Ridge National Laboratory, Oak Ridge, Tennessee.

Omernik, J.M., 1987. Ecoregions of the Conterminous United States: *Annals of the Association of American Geographers*, Vol. 77, No. 1, pp. 118–125.

Omernik, J. M., and A. L. Gallant, 1990. Defining Regions for Evaluating Environmental Resources, *Proceedings of the Global Natural Resources Monitoring and Assessment Symposium*, Bethesda, Maryland, pp. 936–947.

Pielke, R. A., and R. Avissar, 1990. Influence of Landscape Structure on Local and Regional Climate, *Landscape Ecology*, Vol. 4, No. 2-3, pp. 133–155.

Pinker, R. T., 1990. Satellites and Our Understanding of the Surface Energy Balance: *Paleogeography, Paleoclimatology and Paleoecology (Global and Planetary Change Section)*, Vol. 82, pp. 321–342.

Risser, P. G., 1985. *Spatial and Temporal Variability of Biospheric and Geospheric Processes: Research Needed to Determine Interactions with Global Change*: International Council of Scientific Unions, Paris.

Roller, N. E. G., and J. E. Colwell, 1986. Coarse-Resolution Satellite Data for Ecological Surveys, *BioScience*, Vol. 36, No. 7, pp. 468–475.

Sadowski, F. G., 1990. Prototype Land Data Sets for Studies of Global Change, *Proceedings, International Geoscience and Remote Sensing Symposium*, 10th, Washington, D.C., May 1990, Vol. 2, p. 1235.

Tappan, G., and Moore, D.G., 1989. Seasonal Vegetation Monitoring with AVHRR Data for Grasshopper and Locust Control in West Africa, *Proceedings, International Symposium on Remote Sensing of Environment*, 22nd, Abidjan, Cote D'Ivoire, October 1988, Vol. 1, pp. 221–234.

Townshend, J. R. G., and C. J. Tucker, 1984. Objective Assessment of Advanced Very High Resolution Radiometer Data for Land Cover Mapping: *International Journal of Remote Sensing*, Vol. 5, No. 2, pp. 497–504.

Townshend, J. R. G. and C. O. Justice, 1988. Selecting the Spatial Resolution of Satellite Sensors Required for Global Monitoring of Land Transformations: *International Journal of Remote Sensing*, Vol. 9, No. 2, pp. 187–236.

Townshend, J. R. G., C. O. Justice, and V. Kalb, 1987. Characterization and Classification of South American Land Cover Types: *International Journal of Remote Sensing*, Vol. 8, No. 8, pp. 1189–1207.

Tucker, C. J., J. A. Gatlin, and S. R. Schneider, 1984. Monitoring Vegetation in the Nile Delta with NOAA-6 and NOAA-7 AVHRR Imagery: *Photogrammetric Engineering & Remote Sensing*, Vol. 50, No. 1, pp. 53–61.

Tucker, C. J., J. R. G. Townshend, and T. E. Goff, 1985. African Land-Cover Classification Using Satellite Data: *Science*, Vol. 227, No. 4685, pp. 369–375.

U.S. Department of Agriculture, Soil Conservation Service, 1981. *Land Resource Regions and Major Land Resource Areas of the United States*, Agriculture Handbook 296, Washington, D.C.

U.S. Geological Survey, 1986. *Land Use and Land Cover Digital Data from 1:250,000- and 1,100,000-Scale Maps*, U.S. Geological Survey Data User's Guide No. 4, Reston Virginia.

———, 1990. *Digital Line Graphs from 1:2,000,000-scale maps*, U.S. Geological Survey Data User's Guide No. 3, Reston Virginia.