AN EXAMINATION OF THE POTENTIAL APPLICATIONS OF AUTOMATIC CLASSIFICATION TECHNIQUES TO GEORGIA MANAGEMENT PROBLEMS

By Bruce Q. Rado, Department of Natural Resources,
Office of Planning and Research, Atlanta, Georgia 30334

ABSTRACT

Recent State and federal legislation, along with improved methods and techniques for data acquisition and manipulation, has created new opportunities for State, regional, and local planning.

One such opportunity included representatives from the State of Georgia and other southeastern states who participated in a workshop at the NASA/Earth Resources Laboratory in Bay St. Louis, Mississippi. The purpose of the workshop was to become more familiar with automatic classification techniques as they may relate to future statewide planning systems.

Aside from the lectures which the group was to receive, the opportunity was also offered to process on LANDSAT tape, provided we acquired the tape and the training samples.

During the three days at the NASA facility, the group also received an intensive briefing on how this system operates, the types of equipment and associated costs, the more scientific method of obtaining training samples, and a presentation of the assorted case studies as they pertain to different disciplines. The latter proved to be quite beneficial, because we were able to relate to specific issues and formulate potential applications which the automatic classification system might have.

The results of the classified tape were then presented to individuals in the various State, university, and regional planning agencies. Many of these groups will react favorably provided we can get the land cover classifications we desire to acceptable geographic reference standards. This is especially true if the system is flexible and meets the decision-maker's needs for data at a low cost. The new tape, once classified, can be reproduced at a variety of scales through electronic expansion without loss of information and detail. This group further decided that an investigation as to the system's full potential had merit. The opportunity for further investigation has been secured in the form of technology transfer between the State of Georgia and the Earth Resources Laboratory of NASA.

INTRODUCTION*

Recent State and federal legislation, along with empirical methods and techniques for data acquisition and manipulation, has created new opportunities for State, regional, and local planning.

Moreover, proposed federal and State land use legislation, which has resulted from the pressures of rapid growth and development of the environment, has caused an evaluation as to the responsibility of various governmental agencies. It is these responsibilities, along with a recognition of specific issues and possible data sources, which require further attention.

The objective of this paper is to present the applications of automatic classification techniques to Georgia management problems. Therefore, discussions will be presented as to some of the issues which need to be addressed in a statewide system, and how automatic classification techniques can be made compatible with such a system. The observations regarding the implementation of an automatic classification system in Georgia are based upon preliminary investigations with NASA and potential users throughout the State.

As appropriate, this paper presents information derived from the Georgia Resource Assessment Program and other statewide agency activities.

For the past two years, the Georgia Department of Natural Resources, Office of Planning and Research, has been engaged in a statewide Resource Assessment Program. This Program, which incorporates much of the data and results produced by the Department's engineers, geologists, biologists, and planners, as well as data from other agencies such as the USDA/Soil Conservation Service, has begun to serve as a data input for the Department of Natural Resources as well as other agencies of state government when engaging in various types of natural resource planning efforts.

The concept of the Resource Assessment is to provide data and standards for agencies to implement in their project design. In this way, these projects will be improved from a natural resource perspective. This should also result in a more efficient environmental impact statement review since we would be better able to support the alternatives which the implementing agency recommends.

Under its present structure, the Resource Assessment Program has four elements incorporated into its design.

- The first element, known as the Resource Index, is a publication entitled A Guide to the Natural Resource Information of Georgia Volumes I and II. This publication serves as a guide to all existing natural resource information contained in published reports which pertain to Georgia. Its purpose is to present professional individuals with a quick but thorough knowledge of all the available natural resource data within the State.

- The second element, known as the Resource Inventory, is the objective and reproducible mapping of soils, vegetation, watersheds, geology, and other important data on a statewide basis. This data is mapped at a scale of 1" = 1 mile and 1:250,000 (1" = approximately 4 miles) with minimum map units of about 40 acres.

- The third element, known as Application, deals with the recommended techniques for using the data base for the determination of relationships between natural resources and land disturbing activities.

- The fourth element, known as the Resource Research, involves researching and analyzing the experiences of other state and federal groups relating to manual and computer systems design and remote sensing.

These applications of computer and manual techniques can take several forms, ranging from the more qualitative overlay to the highly technical simulation model having several objectives such as vulnerability, attractiveness, suitability, and carrying capacity. The user will ultimately be responsible for selecting those techniques which best serve his needs. In addition, User Guides to Soils and Vegetation Information have been prepared which describe the procedures involved in the data collection phase of the Resource Assessment along with possible methods of application for each of the data categories.

To date, soils and vegetation information have been mapped by county to a scale of 1" = 1 mile. Soils are mapped by associations which are derived from either the modern or general soils maps. The classifications were grouped according to the soil's texture and physical properties as determined by the USDA/SCS. Vegetation information was mapped by associations from the 1" = 1 mile USDA/ASCS aerial photographic mosaics which varied from one to approximately ten years in age.

The purpose of the Resource Assessment is to provide data in a common format for use by State, regional, and local groups when determining areas of natural resources vulnerability, determining the location of potential areas for sanitary landfill operations, and use as a data source in the A-95 and environmental impact statement process. At the interdepartmental level, the data has served as a natural resource base to the State's Coastal Zone Management Program and the Department of Transportation's Land Use Allocation Model. At the regional and local levels, the information has been in great demand by individuals working for public agencies and consulting firms engaged in a myriad of applications ranging from federally contracted projects to those of more local significance. One example is the use of our 1" = 1 mile data, augmented with other sources, by the Georgia Mountains Area Planning and Development Commission for determining the impacts of proposed projects within its 13 counties.

Given the common data base system requirements which several federal, state, regional and local governments, and consulting firms have for certain data, it can be anticipated that a need exists for an expanded data base which can supply existing programs with better quality data on a more timely and iterative basis. An automatic classification system utilizing LANDSAT data should allow us the flexibility of developing certain classification systems on an individual project basis, should provide us with objective and iterative data for a relatively inexpensive cost and should be geographically referenced to acceptable standards compatible with other data in both grid and polygon form. Many of these issues are fairly straightforward, and if given proper consideration, can be adequately addressed. There is, however, one technical issue which could possibly prove to be the greatest limiting factor in the design of a statewide system, that is the ability to economically collect and update the data bank.

Until recently, most states engaging in such a system have had few alternatives as to how natural resource data is collected. The standard procedure, which was also used for our vegetation mapping, has been to assemble university students and low altitude photographs and begin the rigorous interpretive process. The obvious disadvantages are that such a process is expensive to update and is quite difficult to ensure compatible results.

Presently, there are several methods which collect and manipulate data for natural resource decisions. In an attempt to look for new and better ways of improving and synthesizing data aquisition, storage and retrieval, a task force from the Southern Growth Policy Board sub-committee on Land Use and Carrying Capacity Studies was invited to a three day workshop at the Earth Resources Laboratory of NASA in Bay St. Louis, Mississippi. The purpose of the workshop was to examine automatic classification techniques as they may relate to future statewide information and natural resource planning systems.

During the three days at the NASA facility, the group received an intensive briefing on how this system operates, the types of equipment and the associated costs, the more scientific method of obtaining training samples, and a presentation of the assorted case studies as they pertained to application by different disciplines. The latter proved to be quite beneficial since we were able to relate to specific issues and formulate potential applications which the automatic classification system might have.

Aside from the three days of lectures which we were to receive, the opportunity was also offered to process one LANDSAT tape (approximately 100 miles by 25 nautical miles), providing we did the training samples and acquired the computer compatible tape. An effort was then launched which would give us access to a computer tape and teach us the art of collecting training samples.

Of the LANDSAT tapes we had access to, it was decided to pick the coastal frame which includes practically all of Georgia's islands and marshes, while extending inland to include the new I-95, the cities of Savannah and Brunswick, plus the river swamps and areas of upland vegetation. Before the data classifications were chosen, an attempt was made to determine the types of data which would be useful for the various State agencies. The following categories seemed appropriate: sand and spoil areas, salt water marsh grass, brackish marsh grass, lowland river swamp vegetation, upland grass areas, different densities of urban cover and upland vegetation by associations (i.e. upland hardwood, pine mixed age, pine even age, mixed hardwoods and pines). The training samples, which numbered approximately 75, were then aggregated until we had approximately five for each classification category containing a minimum of 15-25 acres and a maximum of several hundred acres. The total amount of time for collecting these training samples was about two days.

During our second evening at the facility, the training samples were located on the unclassified data tape which was displayed on the image display system. We were on the system for approximately three hours which enabled us to take all the 75 pre-selected training samples plus some additional new ones. The following morning the statistical information was ready for review and analysis. Each training sample was then reviewed for any bi-model characteristics while the divergence statistics were checked to determine if further training samples were needed and the number of groups which could be categorized. Following a review of the statistics, the classifications were grouped and the data was classified using spectral pattern recognition programs. The actual printing of the unclassified display (see figure 1) and the classified product (see figure 2) to a scale of 1:250,000 (1" = approximately 4 miles) was then performed on the film recorder. The classified final product displayed the following categories: low density urban, (light blue), higher density urban (cyan), beach and spoil areas (white,), upland grass areas (yellow), salt water marsh grasses (red), brackish marsh grasses (magenta), water bodies (blue), and upland vegetation (green). There was an extremely small area which was displayed as unclassified (black).

The results of this demonstration tape have proved impressive, although there were categories which we were unable to separate due to the time of year the tape was derived. Since we were using an October 26, 1973 tape, the specific signatures of the upland vegetation were difficult to separate. The problem of separating the vegetation associations was further extenuated because

of the quality of our training samples. Our review of NASA and USGS research products has indicated that given proper time of year tapes, that vegetation associations can be separated. Therefore, reaching this objective will depend upon the successful grouping of the various vegetation associations which we were unable to obtain from this tape. Our NASA work should provide us with the opportunity to separate vegetation associations provided we obtain better training samples, become more familiar with our training sites, and are able to obtain updates on these sites.

In arriving at any conclusions about an automatic classification system, one might decide that a system which is flexible, rapid, and inexpensive from a personnel and operating cost standpoint could be more desirable than a system lacking these advantages but able to obtain a greater degree of accuracy. An important characteristic of this computer implemented technique is that it is iterative and the newly acquired data can possibly be formatted for direct input into a system for analysis without digitizing or resorting to hand-drawn maps. Another advantage of this automatic classification technique is that it will operate on any multispectral scanner data whether it be derived from medium or high flight aircraft or from a satellite. One issue of concern is a comparison of the scanner capabilities of LANDSAT versus Skylab and aircraft data (see figure 3). The recent work by Coker et al. appears to indicate that if a Skylab or a aircraft-type scanner could be designed for operational LANDSAT application, then products of greater utility to Georgia, and possibly other states, could be derived.

The classified tape was presented to individuals in the various State, university, and regional planning agencies. Many of these groups will react favorably provided we can get the land cover classifications we desire to acceptable geographic reference standards. This is especially true if the system is flexible and meets the decision-maker's needs for data at a low cost. The new tape, once classified, can be reproduced at a variety of scales through electronic expansion without loss of information and detail. This group further decided that an investigation as to the system's full potential had merit. The opportunity for further investigation has been secured in the form of technology transfer between the State of Georgia and the Earth Resources Laboratory of NASA.

To date, several meetings have been held with interested agency personnel to assist us in relating potential applications to the automatic classification technology. These meetings have addressed the uses and data needs which the various agency representatives plan to investigate. Below is a preliminary list of the potential uses which our study will initially address:

- agriculture

    crop acreage and yield predictions
    disease infestation studies

- forestry

    crop inventories and yield predictions
    percent crown density
    age of species
    disease infestation studies
    undesirable species studies
    wildlife habitat areas

water resources

> salinity rates
> flood plain studies
> estuary and river swamp studies
> marshland protection (monitoring change)

environmental planning

> land use classification systems
> environmental impact statements and A-95 reviews
> monitoring urban growth
> septic tank feasibility studies
> natural areas and ecosystem studies
> heritage trust sites
> resource assessment program

geology

> beach erosion and deposition of rivers
> monitor channel morphology
> land reclamation studies
> delineation of shoal areas

The Department of Natural Resources' Resource Assessment Program is serving as a natural resource planning element within our State. It is an issue-oriented program which relies upon data from other sources whenever available and develops its own data only when it is feasible. Its objective is to assist in resolving resource management problems regarding statewide concern such as: selecting feasible alternatives for highway corridors, selecting new airport sites, developing shoreline management programs for the Corps of Engineers, developing regional land-use programs and developing an analysis of natural resource impacts for projects such as reservoir construction, small watershed projects and the location of major industry.

To date there has been a substantial use of the data in its present form. But the problem with the existing data is that as the users become more sophisticated, they will prefer, and probably be required, to quantify rather than qualify their results. If we can obtain the desired classifications and solve any geographical control problems to acceptable minimum map accuracy standards, then we expect that automatic classification techniques can provide us with a more acceptable product.

There is, however, a limited regional scale of planning which can be addressed at scales such as 1" = 1 mile or 1:250,000. We must realize that the type of decisions at this scale will always be too general for site-specific areas. Although we never designed or anticipated that our system would provide us with site-specific data, it is anticipated that an automatic classification system which is flexible, iterative, digitized, and with a smaller minimum unit than our current hand-drawn maps, should be of great utility.

One example which the Resource Assessment Program can compare is the existing 1" = 1 mile vegetation maps prepared for each of the State's 159 counties.

The cost of this project was approximately $30,000, which did not include administrative costs to DNR, or about $190.00 per county. Each map was prepared on clear mylar while color was applied to delineate the appropriate vegetation associations. This exercise has proven quite valuable since it not only provides the State with one fairly consistent mapping effort, but also since it provides good indicator information for large portions of the State which, until now, had little land cover data. These hand-produced maps still have problems since they were drawn by different individuals, interpreted from the "best available photography" and the cost of updating would be expensive. These products would be more acceptable if improved sensors were able to provide iterative coverage for particular areas having site-specific problems.

These products should not only enable us to do more efficient work within our own departmental programs, but also enable us to provide better technical data and assistance to other public and private concerns.

In the early stages of this project we fully realized the need to meet with potential users and to gain some insight into the types of applications which could be anticipated. Although these uses serve a valuable function, it is probably more important to be able to identify those goals, objectives, target dates and issues which are needed to assist in making a final decision on whether such a statewide system is warranted.

The Georgia Department of Transportation is presently developing with the University of Georgia a transportation planning land-use simulation model. This model will enable the DOT to determine the impact of transportation corridors on a statewide basis upon land-use, population, housing and employment.

The model is an iterative land-use simulation model and contains a series of self-contained, interrelated submodels. The submodels to be developed are employment, population, housing, land-use, and transportation. The model's structure will enable it to run with alternative forecasting techniques and policy assumptions.

Presently, there is an inadequate supply of reliable land-cover data. Therefore, the only data which has been collected are gross classifications of available information such as soils, slope, vegetation, and a differentiation between urban and rural areas. These gross classifications indicate that a problem exists in obtaining and managing data for the modeling process.

The implementation of an automatic classification system could aid the modeling effort by increasing the amount and detail of land-cover data. This information would be able to provide a land-cover classification system delineating spatial and statistical data on a pixel basis.

The possible advantages of an iterative automatic classification system appear numerous, especially when comparing it to the current methods of manual acquisition. Those advantages include a savings of money and time, better detail of data, and a greater opportunity for sophistication and quantification.

Other DNR applications should be in addressing environmental impact statements and A-95 reviews. One recent example, which may have had application to automatic classification techniques, is a 60,000 acre small watershed project proposed by

the USDA/Soil Conservation Service for the purposes of watershed protection and flood prevention. This coastal watershed has approximately 242,000 acres containing three primary vegetation habitats: lowland hardwoods, mixed hardwoods with pines and pine flatwoods. The area to be drained is about 90% mixed hardwoods with pines and 10% lowland hardwoods. If an automatic classification system were operational it could have possibly provided us with the following spatial data and statistical acreage information: location of drainage ways containing lowland hardwoods, areas inundated by water at various times of year, a delineation of wildlife habitats, and a determination of the salt and fresh water interfaces. This data would have enabled us to make a better quantitative determination of the project's effects as required by the National Environmental Policy Act of 1969. This system should also allow us to monitor the impacts of increased peak flows on the estuary and low flow conditions on the river swamps and tributaries.

The Department has also indicated its intention to increase the total acreage of land for wildlife management from 1.2 to 3.0 million acres. A major portion of this increase could come through leases from land owners associated with the wood products industry. Accurate and current land-cover data would greatly facilitate the identification of suitable lands for possible acquisition or lease. The types of data required include vegetation associations and age of stands, the proximity of any potential area to developed lands, and the location of surface water bodies. If this data can be obtained on a timely and cost-effective basis, then our management program could be substantially improved.

In the coastal marshlands, the automatic classification system should assist us in determining seasonal patterns and changes in the relationship between upland development and the marsh. It should also assist in monitoring and evaluating the effects of federal disposal of dredge and fill on the marshes and coastal waters. Another opportunity is to spatially locate and survey where the wetlands are and how the management practices of adjacent states effect Georgia's marshes. One disadvantage of this system using LANDSAT data is the difficulty in locating small alterations which are beyond the satellite's present resolution and registration capabilities.

The agricultural community within our State is interested in examining automatic classification techniques since it appears to offer the opportunity to continually collect agricultural resource data resulting in more efficient inventory and management procedures. Research results to date indicate that remote sensing, either from aircraft or satellites, is beginning to play a significant role in agricultural operations provided high standards of accuracy can be achieved and if the cost to the users can be efficient. Listed below are a few of the issues which need to be addressed if LANDSAT data is to be applied for agricultural purposes in our state:

- higher spectral and spatial ($<$ 10-30 meters) resolution
- sensors capable of cloud penetration
- repetitive and dependable coverage at periods of better than 18 days
- thermal scanner capability
- better real-time data availability and processing capabilities

Currently, we are attempting to identify those issues which will need future attention. Until recently most departments of state government have been operating

within their own limited boundaries. Now, with the appearance of programs such as the Department of Transportation's Land Use Allocation Model, the Department of Natural Resources' Resource Assessment Program, the State Office of Planning and Budget's Coastal Zone Management Program, and legislation such as the National Environmental Policy Act of 1969, a new era of interdepartmental problem-solving is emerging. By nature, these projects tend to be more comprehensive in scope and demanding in terms of quality and quantity of data. Preliminary indications from these programs appear to be pointing out that traditional ways of collecting, disseminating, and manipulating data are rapidly becoming inadequate. It is therefore imperative that new options and a determination of their potentials be explored.

A first step in the investigation of any statewide system or subcomponents of that system, such as the NASA automatic classified system, is to answer the types of questions which Calkins discusses in this paper. These questions include such obvious, but all too often overlooked, issues such as "does the system meet the needs of the user"; "is the user constrained from using the system"; and "does the user have any incentive to use such a system?" If the answer to these questions is no, then we have an example which could indicate a design problem. Perhaps the system as originally implemented was never useful, or, the needs of the user changed while the system did not or could not change. Whatever the reason, these examples point to a greater need for participation and communication between the groups charged with financing, designing, and ultimate application of such a system. It is therefore imperative that a more three-dimensional approach to systems design be observed. This would include the above objectives while also taking into consideration issues such as: how to obtain user confidence, how to reduce duplication of efforts, how to input changing conditions during a study, and how more reliable data and systematic processing techniques can prove to make planning and its results more accountable.

In our NASA technology transfer project, it will be essential that potential users of the system document their particular areas of concern and that criteria for desired results be determined, it will then be necessary to hold workshops between the users and the individuals transferring the technology. This will result in discussions of potential applications and use of the system. The next step might be to determine what the results of the various disciplines have in common and how, if at all, a statewide system could be used to address the issues of single and multi-disciplinary projects.

Following a determination of the applications of the automatic classification system, another dimension needs to be analyzed. That is, how can other programs, which include either data and/or analytical techniques, be integrated with the automatic classification techniques to enable the various agency-collected data to compliment each other. One example might be, how can the information from the NASA automatic classification system be made compatible with that of the LUDA program being developed within the USGS? From a user perspective, it will be quite useful if the two systems are compatible. The most obvious reason being that the apparent advantage of the NASA automatic classification system is its ability to produce land-cover information, plus associated statistics, on a pixel basis for large regions with iterative capabilities. Although the land-cover classification may not be sufficient for certain applications in urban areas, it should prove adequate and flexible in its classification in regions of natural land-cover. When dealing with urban areas, the LUDA program could prove more

effective since it will map level II land uses* to 10 acre accuracy in urban areas and 40 acre accuracy in rural areas. In any event, there is a tremendous amount of work necessary to evaluate individual components and then their relationships to each other. Comparative analysis will be made in Georgia as to pros and cons of each of these systems for both urban and natural resource application.

In closing, there are possibly more questions to be asked than products to present. Listed below are a few questions which surface when considering the implementation of a statewide automatic classification system.

- If we decide that an automatic classification system will do the job for us, what guarantee do we have that LANDSAT will become operational?

- Do we need operational aircraft with an MSS scanner?

- Will anyone assist us in some of the equipment costs?

- Is private industry able to assist government in this effort at a price we can afford?

- How can we assist an individual such as a farmer, who has a problem on a given day if there is no satellite overpass?

- Will it be more appropriate to rent or purchase equipment such as a portable image display system (PIDS) and film writer?

- What type of output and format do we need?

- How do we set up a user system in terms of communicating and disseminating data?

- Is it more appropriate to have 1 system or several systems within the State?

- How do we effectively incorporate equipment deliveries and budget cycles into the operation?

- Will the state need to begin a LANDSAT tape library, and if so, how much will it cost on a yearly basis?

- How can we obtain tapes on a more timely basis from Goddard?

- What is the continuing need and appropriate role of manual interpretation?

---

* Although the LUDA program claims to be mapping land use, probably the more appropriate term should be land-cover. This is suggested since the LUDA program will interpret from high altitude photographs without the needed element of field checking or other data sources.

There are several states which are either in the process of dealing with the issues in this paper or have already dealt with them. It is our hope that while dealing with these issues within our own State, that a forum be established for discussing and resolving these and other issues which are of mutual concern to all states.

BIBLIOGRAPHY

Calkins, Hugh W., "The Proposed Washington State Land Information Service", Bellevue, Washington, 1973.

Coker, A. E. et al., "Automatic Categorization of Land-Water Cover Types of the Green Types of the Green Swamp, Florida, Using Skylab Multispectral Scanner (S-192) Data", U.S. Geological Survey, Water Resources Division, August, 1974.

Colvocoresses, Alden P., "Recommended Specifications for a Follow on ERTS (ERTS-C) Satellite", U.S. Geological Survey National Center, Reston, Virginia, March, 1975.

Dueker, Kenneth J. et al., "Land Resource Information System: A Development and Demonstration for the State of Iowa", Institute of Urban and Regional Research, University of Iowa, Iowa City, Iowa, December, 1974.

Halperin, Mike, "Use of ERTS Land Use Information for Transportation Planning", Georgia Department of Transportation, Engineering Experiment Station, Georgia Institute of Technology, Atlanta, Georgia, December, 1974.

Joyce, T., Armond T., and Derbonne, J.D., "A Computer Implemented Land Use Classification Technique applied with ERTS Digital Data Acquired Over Southern Missouri", Lyndon B. Johnson Space Center, April, 1975.

Miller, Robert H., "The Significance of Remote Sensing Techniques for Agricultural, Forestry, and Rangeland Management", in American Society of Photogrammetry, proceedings, 41st Annual Meeting, March 9-14, 1975, pp. 119-206.

Smith, John E., "New Direction for Land-Related Planning Information Systems: A State Perspective", New York State Office of Planning Services, Albany, New York, October, 1974.

"Texas Remote Sensing Plan", Prepared by the Task Force on Remote Sensing for Submission to the Interagency Council on Natural Resources and the Environment, December, 1974.
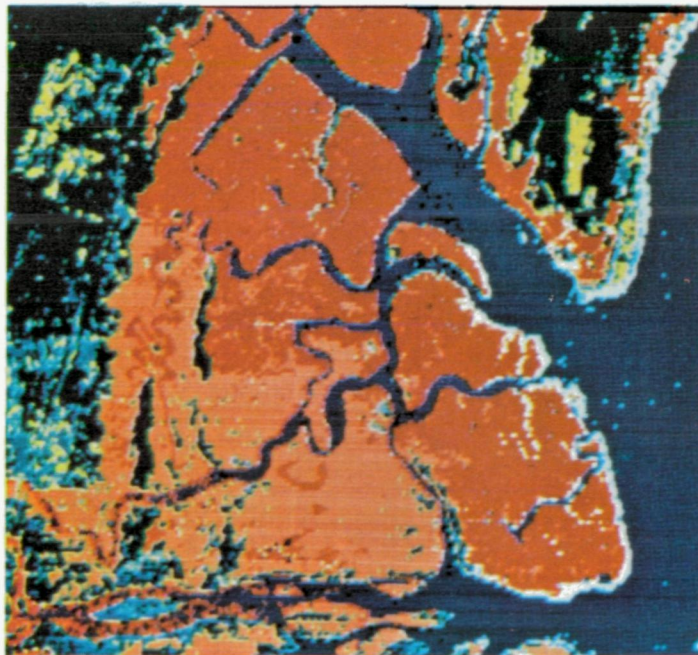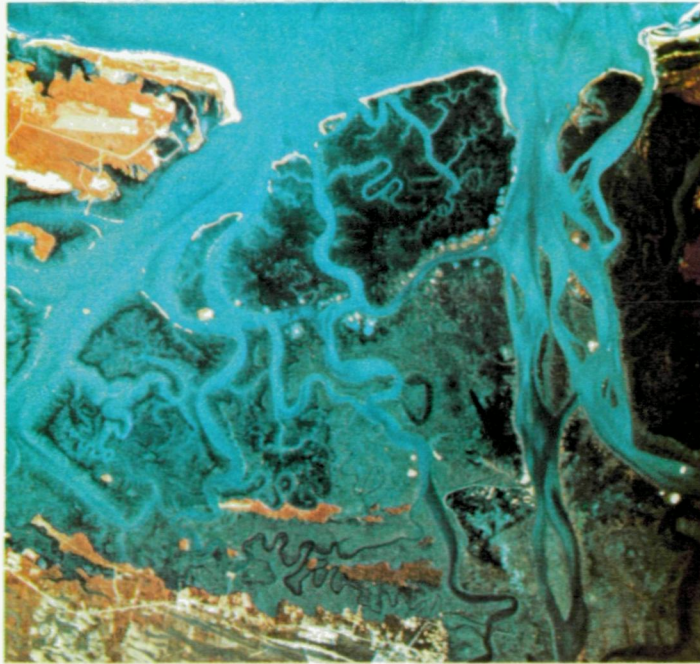
Figure 1

LANDSAT unclassified display of the Savannah area. It was from
a video display of this tape that the following figure was classified.

Figure 2
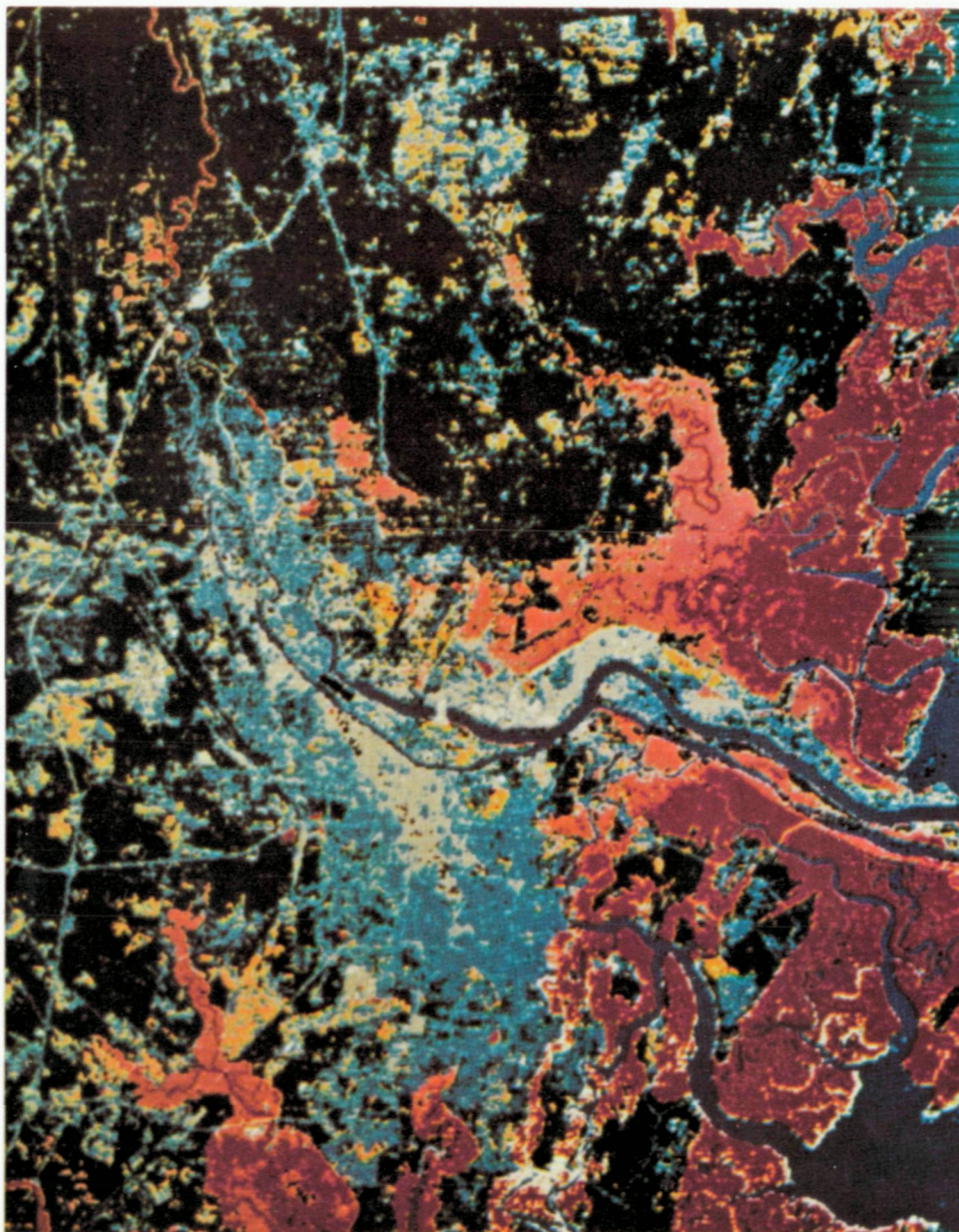
LANDSAT classified tape of the Savannah area. These classifications were derived from training samples obtained from the unclassified display tape.

Figure 3

A comparison of a color infrared highflight photograph (TOP) and a
classified LANDSAT tape (BOTTOM) both taken on October 26, 1973.
The city of Darien is to the west, Sapelo Island to the north, and
the Altamaha River to the south.