General Disclaimer

One or more of the Following Statements may affect this Document

- This document has been reproduced from the best copy furnished by the organizational source. It is being released in the interest of making available as much information as possible.

- This document may contain data, which exceeds the sheet parameters. It was furnished in this condition by the organizational source and is the best copy available.

- This document may contain tone-on-tone or color graphs, charts and/or pictures, which have been reproduced in black and white.

- This document is paginated as submitted by the original source.

- Portions of this document are not fully legible due to the historical nature of some of the material. However, it is the best reproduction available from the original submission.

Produced by the NASA Center for Aerospace Information (CASI)

LANDSAT DATA FOR STATE PLANNING

ANNUAL TECHNICAL REPORT

by

N. L. Faust and G. W. Spann

(NASA-CR-144266) LANDSAT DATA FOR STATE PLANNING Annual Technical Report (Georgia Inst. of Tech.) 57 p HC $4.50 CSCL 08B

N76-22635

Unclas

G3/43 25264

December, 1975

Prepared for

George C. Marshall Space Flight Center
Marshall Space Flight Center, Alabama 35812

MAY 1976
RECEIVED
NASA STI FACILITY
INPUT BRANCH

Under Contract Number NAS8-30653

by
Engineering Experiment Station
Georgia Institute of Technology
Atlanta, Georgia 30332

LANDSAT DATA FOR STATE PLANNING

ANNUAL TECHNICAL REPORT

by

N. L. Faust and G. W. Spann

December, 1975

Prepared for

George C. Marshall Space Flight Center
Marshall Space Flight Center, Alabama 35812

Under Contract Number NAS8-30653

by
Engineering Experiment Station
Georgia Institute of Technology
Atlanta, Georgia 30332

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS	ii
ABSTRACT	iii
I. Introduction	1
II. Summary	2
III. Landsat Data as an Input to G DOT Transportation Planning Model.	4
Site Selection	4
Procedure for Classification	4
Results of Classification	5
Aggregation of Classified Data	7
Land-Use Adjustment	7
IV. Usefulness of Landsat Data	9
Advantages of Digital Processing of Landsat Data	10
Disadvantages	10
V. Cost-Effectiveness Calculations	11
Costs	11
VI. Geobotanical Indicators	18
VII. Summary and Conclusions	22
VIII. Continued Research	22
APPENDIX A - Unsupervised Classification of Landsat MSS Data	23
APPENDIX B - Supervised Classification of Landsat MSS Data	23
APPENDIX C - The Georgia Transportation Planning Land-Use Model	39
Employment and Population Submodel	41
Description of the Delphi Technique	42
Housing and Population Submodel	45
Relationships Between Land-Use Forecasting and Transportation Planning	47
Land Supply Submodel	49

ACKNOWLEDGEMENTS

The authors would like to thank Mr. Mike Hailperin and Mr. Hal Maggied of the Georgia Department of Transportation for their aid throughout this project, and for their evaluation of the usefulness of Landsat data for transportation planning. We would also like to express appreciation to Mr. Bruce Rado and Mr. Lawrie Jorden of the Georgia Department of Natural Resources and Mr. Arden Brey of the Office of Planning and Budget, State of Georgia, for their criticisms and suggestions as to how Landsat data could best be applied to State of Georgia problems.

We would like to especially thank Ms. Pat Sellers of the U. S. Corps of Engineers who provided valuable assistance to this project by collecting ground truth data around the Lake Lanier area.

ABSTRACT

This report presents the results of the second phase of an effort to generate and apply automated classification of Landsat digital data to State of Georgia problems. This phase of the project centers on an analysis of the usefulness of Landsat digital data to provide land-use data for transportation planning.

Hall County, Georgia was chosen as a test site for this phase of the project because it is part of a seventeen county area for which the Georgia Department of Transportation is currently designing a Transportation Planning Land-Use Simulation Model. The land-cover information derived from this study was compared to several other existing sources of land-use data for Hall County and input into this simulation.

The results of this study indicate that there is difficulty comparing Landsat derived land-cover information with previous land-use information since the Landsat data are acquired on an acre by acre grid basis while all previous land-use surveys for Hall County used land-use data on a parcel basis. An analysis of the expected future use of Landsat data for land-use planning, and a special section detailing the benefits and limitations of Landsat digital data are included in this report. In addition, land use categories such as residential, commercial and industrial were only accumulated within corporate boundaries in previous land use studies. An attempt at comparison of the Landsat derived landcover classification to the above land use categories was made by only accumulating Landsat cover categories associated with urban use within city boundaries. Areas outside city boundaries that were originally assigned to residential, commercial or industrial categories were reassigned to another category. These results provided a favorable comparison.

I. Introduction

In a series of tasks begun by the Engineering Experiment Station at Georgia Tech in 1972, several cooperative efforts have been carried out with federal, state, and local agencies. Previous work under this contract with Marshall Space Flight Center has involved cooperation with the Douglas County Planning agency, the Georgia Department of Natural Resources (DNR), Office of Planning and Budget (OPB), and others. In addition, early work on this effort was funded by the Georgia DNR in 1973.

This phase of the project was designed to extend capabilities for computer land-use/land-cover classification that were demonstrated earlier to another area in Georgia, and to interact with state and local agencies not previously involved with this project. The ability to classify the data, however, is only the beginning of its usefulness. Once land-cover information is available, the usefulness of many planning situations and models is enhanced. One such model is the transportation planning land-use simulation model presently being developed by the Georgia Department of Transportation (G DOT). This model is described in more detail in Section II. An evaluation by G DOT of the Landsat data as suitable input into the model is included in this report. A secondary effort was to search for geobotanical indicators relating to the Brevard Shear Zone in the study area.

II. Summary of "Georgia Transportation Planning Land-Use Model"

As indicated in the previous section, one of the primary purposes for this project was the evaluation of Landsat data as an input to the Georgia Transportation Planning Land-Use Model. This model is currently being developed by the College of Business Administration, Research and Services, University of Georgia, Athens, Georgia, under sponsorship of the Georgia Department of Transportation. An appreciation of the nature and scope of the model will aid in understanding the role to be played by the Landsat data as an input to the model.

The Georgia Transportation Planning Land-Use Model is designed primarily to forecast changes in employment, housing, population, and land-use as a result of growth associated with alternatives to selected transportation routes. The Model itself consists of four submodels: Transporation Submodel, (2) Employment Submodel, (3) Population-Housing Submodel, and (4) Land Supply Submodel.

Each submodel in turn utilizes a variety of techniques for evaluation of the effect of changes in the transportation patterns. Multiple evaluation techniques are used to avoid bias in the predictions for each submodel. Local, state, and national data sources are integrated and used as input for the employment and population-housing models. This integration provides consistency in projections with state and national forecasts.

As alternative transportation plans are considered, the results of the various submodels are considered by a Delphi panel (Appendix C) to access the overall effect of each transportation plan and to decide on the best plan to implement ignoring any environmental constraints. Information from the Land Supply submodel is then used to define any environmental constraints that exist which would affect the transportation plan. If these constraints are minimal, the Land Supply submodel will forecast the expected change in land-use if the plan is implemented. If the constraints are major, another iteration is made through the complete process including that plan.

The study area for the University of Georgia project is a seventeen county area in northeast Georgia. Data for all of the submodels are aggregated on a county by county basis, and the results are pertinent only at that scale.

For transportation planning, however, these results will allow Georgia DOT to consider various alternatives for major transportation routes and assess the general effect of each proposed plan.

For more information on the various submodels see Appendix C which was taken from Structuring the Georgia Transportation Planning Land-Use Model, working paper #1, Georgia Department of Transportation Land-Use Model, May, 1974. The authors of the report, Paul F. Wendt and Charles F. Floyd, are, respectively, Professor and Associate Professor of Real Estate and Urban Development, the University of Georgia.

The land-cover information obtained by processing Landsat data over Hall County is one source of data for the Land Supply Submodel of the above model. Even though data are available on a 1.05 acre grid, the submodel is only presently designed to utilize land cover data which has been aggregated over the whole county. The acreage of usable land that is desired for the model is estimated by subtracting the existing area of developed land, water and floodplains, and land unsuitable for development from the total county area. Information that is normally used to provide the usable land area comes from the Georgia Department of Revenue property tax records and previous land-use surveys. Land absorption coefficients for Hall County are then estimated subject to a set of national and state policy constraints. The absorption coefficients project the rate at which the area of various land use classes which are suitable for development will be absorbed by the commercial activity which accompanies the construction of a transportation facility.

---

*This information was taken from working paper # 1; however, several more recent working papers are available which deal with changes in the various submodels. For an overview of the operationalized model see "The Georgia Transportation Planning Land Use Model": Development of a Policy Sensitive Impact Analysis Planning Tool, M.P. Hailperin, H.S. Maggied, C.S. Floyd, presented at 57th Annual Conference, American Institute of Planners, San Antonio, Texas, October 27, 1975.

III. Landsat Data As An Input to G DOT Transportation Planning Model

Site Selection

Hall County, Georgia was selected as a test site for the application of computer processed Landsat data to Transportation Planning problems. Hall County is one of the seventeen counties that are being studied using the G DOT Transportation Planning Land-Use Simulation Model that was discussed previously. Hall County is presently extremely rural, but because of its proximity to Atlanta, and the presence of Lake Lanier, it is a rapidly changing area that needs an effective land-use policy to allocate its resources. The demand for new facilities will skyrocket with the rising population, and thus Hall County is an excellent test area for the application of such a dynamic data acquisition system as Landsat. In addition, Hall County has available information from two land-use studies that were compiled using windshield surveys and human interpretation of low altitude aerial photographs. The land-cover information derived from automatic classification of Landsat data will be compared with the previous land-use studies before the data are entered into the model. Finally, this information will also be used as one data source by the U. S. Corps of Engineers for a land-use survey of the Lake Lanier area in Hall County.

Procedure for Classification

After the Landsat digital tapes for the North Georgia area including Hall County were obtained, an initial clustering analysis (Appendix A) was accomplished on the area around Lake Lanier. The clustering results were output at a scale of 1:24000 so that the data could be compared to topographic maps of the area and some low altitude aerial photographs provided by the Georgia DOT. By overlaying the printouts and the aerial photographs on a light table, specific classes such as water, forests, commercial, etc. were identified. Subsequent field checking of these areas by Ms. Pat Sellers of the U. S. Corps of Engineers allowed us to verify the identity of the clusters. Once the clusters had been identified, the spectral signature of each class was computed and stored. Several iterations were made through the clustering

process until signatures were available for all the separable classes in the area. Next, the stored spectral signatures were used for a supervised classification (Appendix B) of the whole of Hall County. Software was not available to geographically reference each pixel and overlay the geographic co-ordinates of the county boundaries, so a manual method using topographic maps at a 1:24000 scale was used to define the county boundaries on the supervised classification results. Geographic rectification software will be implemented in the next phase of this project. Each supervised class was assigned a particular symbol on the scaled computer map. These symbols were later manually color coded to produce a land-cover map of Hall County (Fig. 1).

Results of Classification

Twelve land-cover classes were used for the supervised classification of Hall County. These classes are described below. The corresponding color on the Hall County land-cover map and the number of acres for each class are also given. The total acreage for Hall County is approximately 270,000 acres.

Class	Color	Description	Area (Acres)
1	Blue	Open water	12,627
2	White	Hard woods (deciduous)	116,118
3	Gold	Mixed-open + deciduous	31,812
4	Yellow-green	Conifers	24,260
5	Dark green	Open - I	13,599
6	Yellow-brown	Low density residential and secondary roads	28,358
7	Yellow	Residential	14,485
8	Blue-green	Sediment loaded water	6,743
9	Brown	Asphalt - commercial	8,536
10	Olive green	Open - II	7,134
11	Orange	Commercial (large buildings)	860
12	Red-brown	High density residential	5,705
		TOTAL	270,237

* These categories do not directly correspond tothe USGS/NASA land cover classification scheme. They were designed specifically to provide data needed for the Land Supply Submodel.

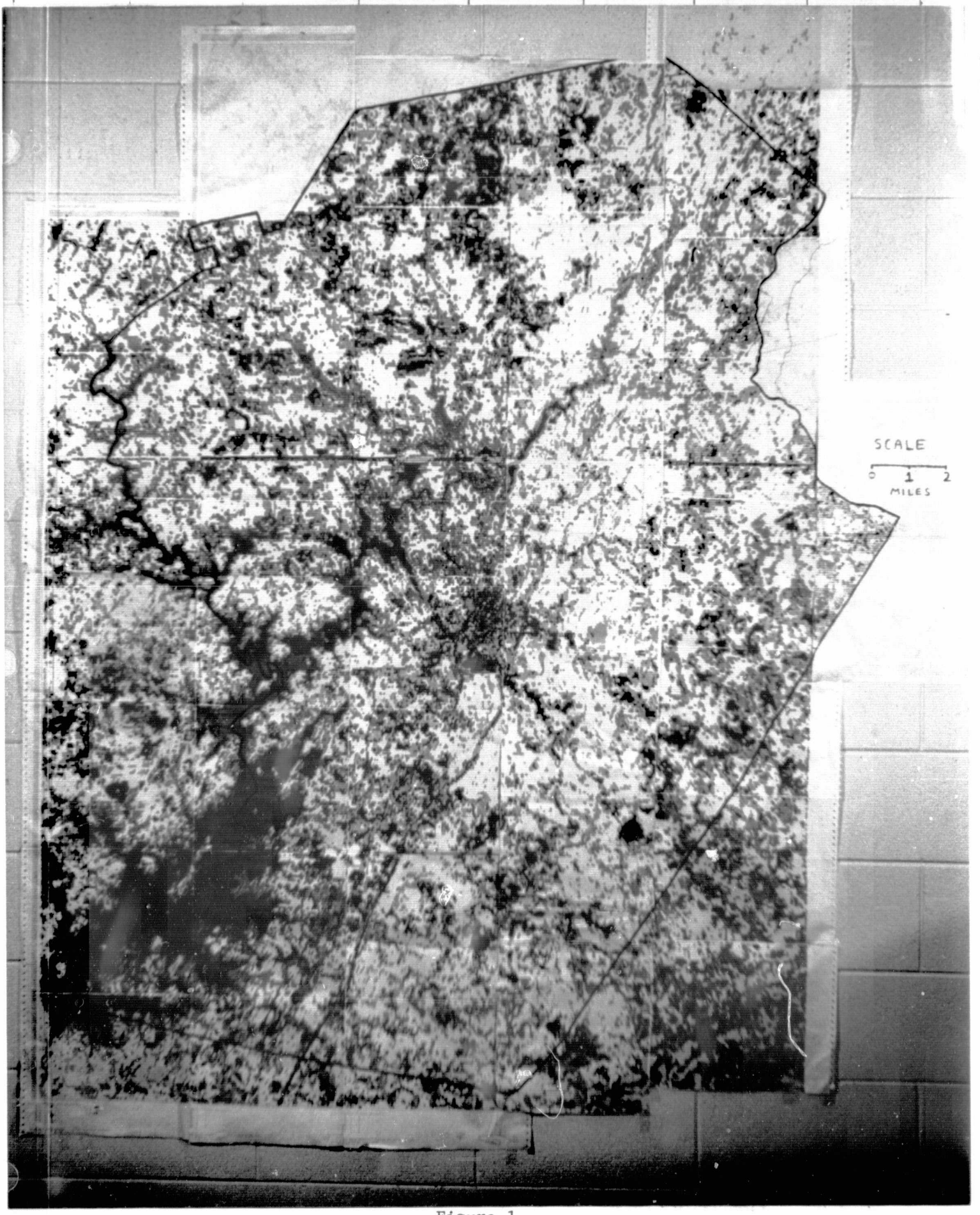


Figure 1.

Aggregation of Classified Data

The data required for input into the G DOT Land-Use Planning Simulation Model are the total acreages, county wide, of residential, commercial and industrial, water, and other. The "other" category is taken to be land suitable for development. No attention is given as to the spatial location of the classes within the county since, at present, the model is designed to attack only regional problems. The table below is an aggregation of the above 12 classes into the four required classes.

Class	Area (Acres)	Name
1	20,190	Residential
2	9,396	Commercial + Industrial
3	12,627	Water
4	228,023	Other

Land-Use Adjustment

Conventional land-use data are accrued on a parcel basis. The predominant land-use of a parcel is taken as the use of the whole parcel. Gridded data, such as Landsat data, are not referenced to parcel boundaries, and it is therefore impossible to aggregate land-use data by parcel until the parcel boundaries are computerized and overlaid onto the land-cover data in a data base management system. An example of where gridded data land-use results vary from parcel oriented results is in rural sections of Hall County. The conventional land-use technique classifies a whole parcel as farm land even though there may be chicken houses, barns, and residences on the property. Using Landsat data, the same area would be segmented into 1.05 acre units that could have the spectral signature of residential, commercial, or industrial classes depending on the type of structure imaged on the LANDSAT scanner.

Since, for the previous land use study there were assumed to be no commercial or residential areas outside corporate boundaries, an alternative would be to aggregate the residential and commercial pixels only within city boundaries and bordering major transportation routes. When these categories

occur outside corporate boundaries or transportation corridors the areas are assigned to the "other" category. These data correlate more closely to the previously gathered land-use data for Hall County. The new aggregations are given below. Note the county itself has not been reclassified.

Class		Area (Acres)
1	Residential	10,127
2	Commercial	1,875
3	Water	12,627
4	Other	245,607

IV. Usefulness Of Landsat Data

Several problems are inherent in the use of Landsat data for transportation planning. The problems are not so much directly related to this modeling process, however, as they are related to the problem of transferring a new technology. First, there must be a general realization of the capabilities and limitations of the data. Remote sensing is a tool, and as such, it must be used at the proper times and in the proper manner. As with all tools, there are some purposes for which it is ideally suited, and others for which it is totally useless. Failure to properly transfer its limitations, as well as its benefits, has probably hampered the general acceptance of remote sensing to date.

In line with the foregoing, it should also be stressed that remote sensing is a complementary technique to the existing methods of gathering land-use/land-cover data. It is not intended to replace the current techniques but only to enhance their utility, effectiveness, and efficiency.

A third problem exists with respect to the presentation of information derived from computer processing of digital remote sensing data. Data which are useful to the planning process are normally presented in pictorial form. Even statistics on population, housing, economics, and the like are generally presented in bar charts, pie charts, or similar graphics. Land-use data in particular are most often presented in the form of multi-colored maps. Thus through prior training and experience, most planners, geographers, and government officials relate best to pictorial information.

In many instances, however, a pictorial format is unnecessary and is the least efficient method for presenting the data. Computer processed Landsat data often fall into this category, certainly in the present instance. All that is needed for input to the transportation model is aggregate land-use statistics by county or sub-county units. Despite the current status of acceptance of remote sensing as a valid tool, it is difficult for many people to accept the "computer-to-computer" transfer of data.

This problem is not new. Business computers were viewed with much suspicion in the early years of their availability. Education of management has largely overcome these early fears of the business computers. Short courses,

serious demonstration projects, and other remote sensing technology transfer efforts are currently needed to lessen the uncertainty about the capabilities of remote sensing generally, and the computer processing of digital data in particular.

Advantages of Digital Processing of Landsat Data

* Landsat data are available regularly and cover the entire state contemporaneously.

* Landsat data are in gridded form and are well suited for inclusion into a data base.

* Area measurements are available automatically.

* Manual methods of making area measurements are slow and often inaccurate.

* Manual derivation of land-use data for the entire state would take an estimated 3.3 man-years; computer processing of the same area could be achieved in six to nine months.

* The computer would have a constant (if any) bias in classifying land-cover, whereas, individual interpreters would not.

* Landsat data include locational information for sub-county areas.

Disadvantages

* Photointerpretation and/or field checking may be more accurate than the computer processing technique.

* Existing land-use information is often in parcel form, therefore the comparison of Landsat information to existing information is difficult.

* Digital processing of Landsat data gives land-cover information, not land-use. Auxiliary information must be used to infer land-use.

* The definition of some categories in conventional land-use surveys differs from that obtained via Landsat classification or photographic interpretation, i.e. strip mining, chicken raising and farming are all in the same category in a conventional survey.

* The Landsat minimum mapping unit is one acre; therefore it is best suited for regional studies.

V. Cost-Effectiveness Calculations

Determination of an exact cost-effectiveness ratio for automatic versus normal methods of generating land-use data of the type needed by a transportation planning model is, at best, difficult. The calculation is complicated by numerous factors, including:

1. The tendency to use existing data of unknown accuracy when such data are available.
2. The conflicts between using land-use data and land-cover data.
3. The "quality" of the model in terms of the sophistication and aggregation of the land-use data needed.

Nevertheless, this section presents the results of a cost-effectiveness calculation based on inputs from the Georgia Department of Transportation, the University of Georgia and the Engineering Experiment Station, Georgia Institute of Technology. It applies specifically to the data needed for the Georgia Transportation Planning Land-Use Model.

Costs

Based on current limited land-use data gathering efforts sponsored by Georgia DOT at the University of Georgia, it is possible to derive cost estimates for providing state-wide land-use data for the Transportation Planning Land-Use Model. The best available estimates indicate that an average of five man-days per county are needed to produce land-use data manually. If an average salary of $15,000 is assumed along with a 100% overhead rate,† the cost per county becomes $625 or approximately $100,000 for the entire state of Georgia. This figure is approximately $1.70 per square mile.

In a previous survey conducted by the Engineering Experiment Station* on land-use mapping activities, reported land-use mapping costs varied from $1.08 to $149.00 per square mile with an average cost of $62.61 per square

---

* G. William Spann and N. L. Faust, "Study of USGS/NASA Land-Use Classification System," Engineering Experiment Station, Georgia Institute of Technology, for NASA/MSFC, March, 1975.

† The overhead rate given by Georgia DOT was 70%; this was adjusted to 100% as an estimate of a realistic overhead rate, if a private company had performed the study.

mile. Therefrore, the $1.70 figure calculated above should probably be considered a minimum cost for normal land-use mapping efforts.

Previous Landsat data processing efforts+ at EES indicated that land-cover mapping could be accomplished on the Georgia Tech U-1108 for about $1.00 per square mile. Georgia Tech no longer owns a U-1108, however, but now owns a CDC Cyber-74. Because this machine is faster and because of a change in the computer charge structure, it is now estimated that land-cover mapping from Landsat data would cost approximately $.60 to $.80 per square mile depending on the amount of analysis required. This cost estimate includes both computer costs and manpower costs.

Extrapolating the maximum figure to the entire State of Georgia indicates a land-cover mapping cost of approximately $48,000. This is less than half the cost of a manual effort. Furthermore, this mapping could be done in a matter of weeks or months instead of years.

Other studies have produced data consistent with the above cost estimates. While the figures for cost per square mile are not the same as determined in this project, the other studies resulted in data with the same relative magnitude. For example, Table 1 was prepared by ECON, Incorporated* under NASA contract NASW-2558. The data are in 1973 dollars. In order to make the data comparable to the cost derived above an inflation factor of 26% was added. Results of this calculation are shown in Table II.

These data, of course, show the collection of land-cover information to be much less costly via computer processing of LANDSAT data than by conventional techniques. While the costs for manual and automatic data collection are not the same as those found for this project, they bear the same relationship, i.e., manual methods of data collection are much more expensive than automatic methods.

---

* The Economic Value of Remote Sensing of Earth Resources from Space, An ERTS Overview and the Value of Continuity of Service, ECON, INC., Volume VI, Part II, page 3-17, NASW-2558, December, 1974.

\+ G. William Spann and N. L. Faust, "Study of USGS/NASA Land-Use Classification System," Engineering Experiment Station, Georgia Institute of Technology, for NASA/MSFC, March, 1975.

In another study of the Earth Resources Survey program prepared by Earth Satellite Corporation and Booze-Allen Applied Research Corporation* different costs were estimated for manual and automatic data collection; but, again, these costs have the same relationship as the costs derived for this project. Table III summarized the EARTHSAT/Booze-Allen estimates (adjusted for inflation). Their original data were presented in 1974 dollars so an inflation factor of 15% was assumed in preparing the table presented here. Again, using Landsat data to supply land-cover information is estimated to be much less costly than using other data sources.

TABLE 1

COST OF LAND COVER INFORMATION (1973 $/SQUARE MILE) AUTOMATIC

	Manual			Automatic		
	Satellite	Aircraft	Ground	Satellite	Aircraft	Ground
Level I	.14	1.13	11.0	.048	.80	11.0
Level II	NC	1.60	12.5	.194	.97	12.5
Level III	NC	NC	14.6	NC	1.42	14.6

NC = The sensor is incapable of providing required detail

TABLE 2

COSTS OF LAND COVER INFORMATION (1976 $/SQUARE MILE)

	Manual			Automatic		
	Satellite	Aircraft	Ground	Satellite	Aircraft	Ground
Level I	.18	1.42	13.86	.06	1.01	13.86
Level II	NC	2.02	15.75	.24	1.22	15.75
Level III	NC	NC	18.40	NC	1.79	18.40

*Earth Resources Survey Benefit-Cost Study, Earth Satellite Corporation and Booze-Allen Applied Research, Contract Number 135-19, USGS, Appendix 5, page iii, November, 1974.

TABLE III.

UNIT COST ESTIMATES
LAND COVER INFORMATION (1976 $/SQUARE MILE)

Information Granularity	Unit Cost ($/square mile)		
	Without ERS	With ERS Manual	With ERS Automated
I Coarse	5.68	.35	.91
II Medium	6.16	1.86	1.98
III Fine	16.85	N/A	N/A

Table IV summarizes the different cost estimates for producing land-cover information at a categorization equivalent to Level II of the USGS/NASA land-use description system. Column 1 presents the "best" estimate for computer processing of Landsat data to obtain the land-cover information. Column 2 presents the "best " estimate for collection of the land-cover data by other, nonautomated means. These data support the contention that computer processing of Landsat data is a less costly method of obtaining land-cover information than more traditional methods, and in general can be accomplished in much less total time.

The only other considerations in this cost-effectiveness calculation of the land-cover data and the appropriateness of the categories for the Georgia Transportation Planning Model. These topics are discussed in detail below.

TABLE IV.

SUMMARY OF COST ESTIMATES FOR GATHERING LAND COVER INFORMATION

Source	Automatic Processing of Landsat Data ($/square mile)	Non Automated Data Collection ($/square mile)
Current project	$.60-$.80	$1.70
March 1975 EES report	$1.00	$62.61 (ave.)
ECON	$.24	$15.75
EARTHSAT/ Booze-Allen	$1.98	$6.16
Jayroe (1)	$.13#	*
Joyce (2)	$.58-$1.30	*

No estimate given, computer time only.

Since the Georgia Transportation Planning Land-Use Model was not planned with Landsat in mind, an exact comparison of the effectiveness of Landsat and conventional data sources is not possible. An example of the problems encountered is the discrepancy in the data classification provided by the normal methods adopted and by the Landsat data. Consider, for example, a 160 acre farm which might consist of a farm house, several commercial broiler houses and a small country store. Under the methodology currently used in data gathering for the model, all 160 acres would be classified rural agricultural. In fact no commercial activity outside corporate limits is recognized.

Using Landsat data, however, there would be at least three classifications for this particular geographical area: commercial, residential, and cropland/pasture. While there is little question that Landsat data could serve the needs of the model for land-use data, some conceptual revisions would be needed. This possibility is currently being explored.

In summary we can say that the use of Landsat data in the Georgia Transportation Planning Land-Use Model is cost-effective. The data are of sufficient accuracy on an acre to acre basis, are less costly to obtain, and are more timely than conventionally derived land-use data. The only remaining question is whether, at the current stage of development, the Georgia Transportation Planning Land-Use Model should be conceptually revised to take advantage of Landsat technology.

VI. Geobotanical Indicators

In phase one of this project a correlation was found in the Douglasville, Georgia area between one of the tree categories classified from digital Landsat data and a particular soil type derived from a mica schist. This mica schist was found to be an extremely metamorphosed rock unit associated with the Brevard Fault zone in Georgia. Thus, an indirect delineation of a fault zone was accomplished by the classification of Landsat digital data. This unexpected result led to an effort in this phase of the project to determine if a similar type of situation occurred in Hall County. Figure 2, a geologic map of northeast Georgia, shows the Brevard Fault zone extending from the Douglasville area northeast into Hall County. The Brevard Fault zone is a major fault zone and cuts many different rock units. Since the individual rock units metamorphosed along the Brevard Fault are not necessarily of the same composition as those in Douglas County, the identical soil units as found in Douglas County would not necessarily be found all along the fault; however, it was expected that a similar elongated trend in vegetation types in Hall County would be found, and that this trend could also be traced to a rock unit associated with the fault zone.

Landsat data from two seasons were studied in an attempt to detect consistent elongated vegetative trends. Figure 3 shows a clustered April Landsat scene next to a black and white print of the classified data, and Figure 1 shows the October classified data for the whole of Hall County. First, it should be noticed that the lake itself trends parallel to the fault zone. Possible fractures parallel to the fault and cross-fractures perpendicular to the fault are indicated by sharp, linear lake inlets.

Figure 1 shows a definite elongated vegetation group (yellow-green) striking northeast. It is most pronounced southeast and northeast of Gainesville. The yellow-green class has been identified with ground truth as loblolly pine. This is the same tree type that was observed to exhibit an elongated pattern corresponding to the Brevard Fault zone in the Douglasville, Georgia area. By overlaying a geologic map on the classified data, one can see the correlation of the loblolly pine vegetation group with a geologic unit, the Brevard Schist. Figure 3 shows in black the same vegetation group

ORIGINAL PAGE IS
OF POOR QUALITY

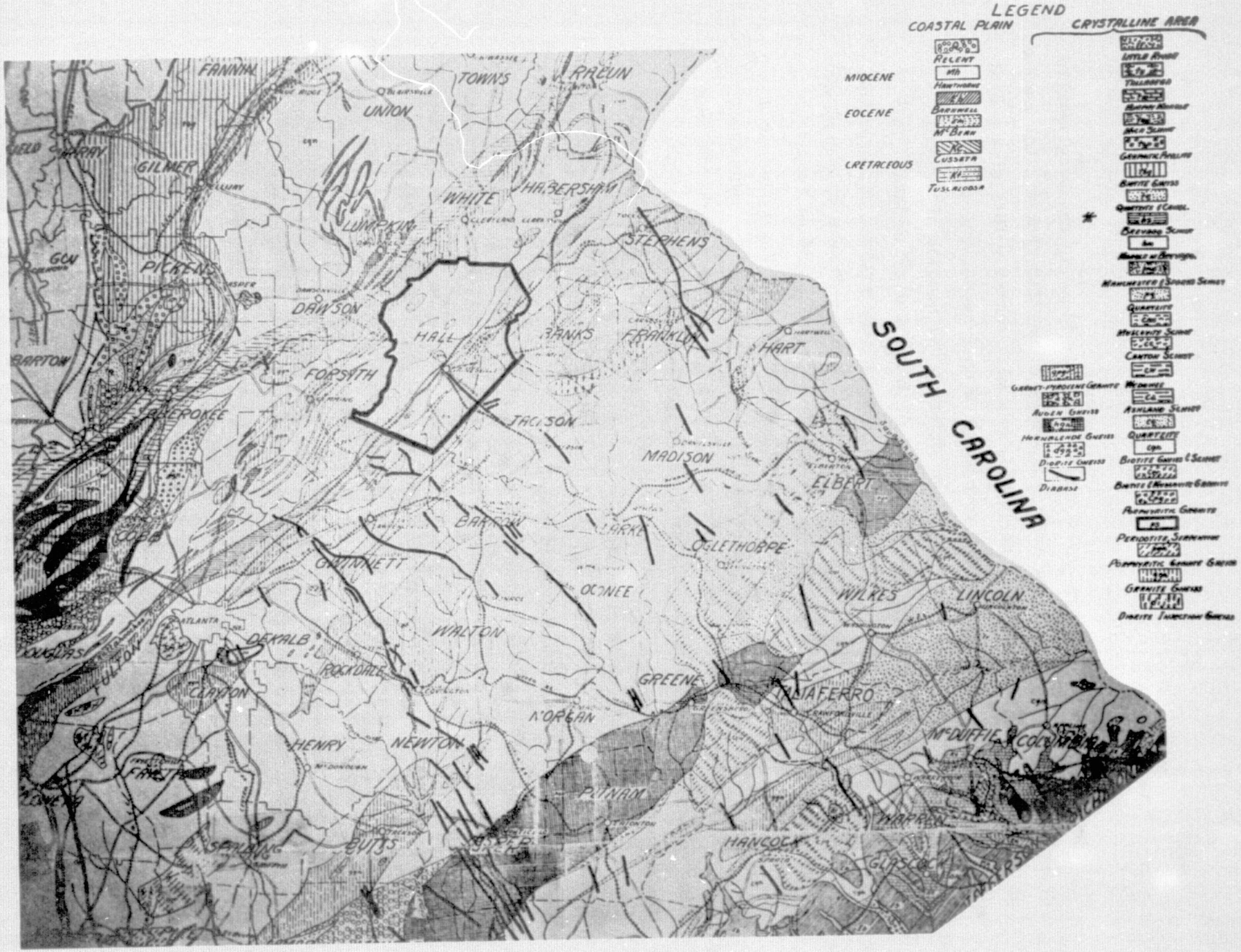
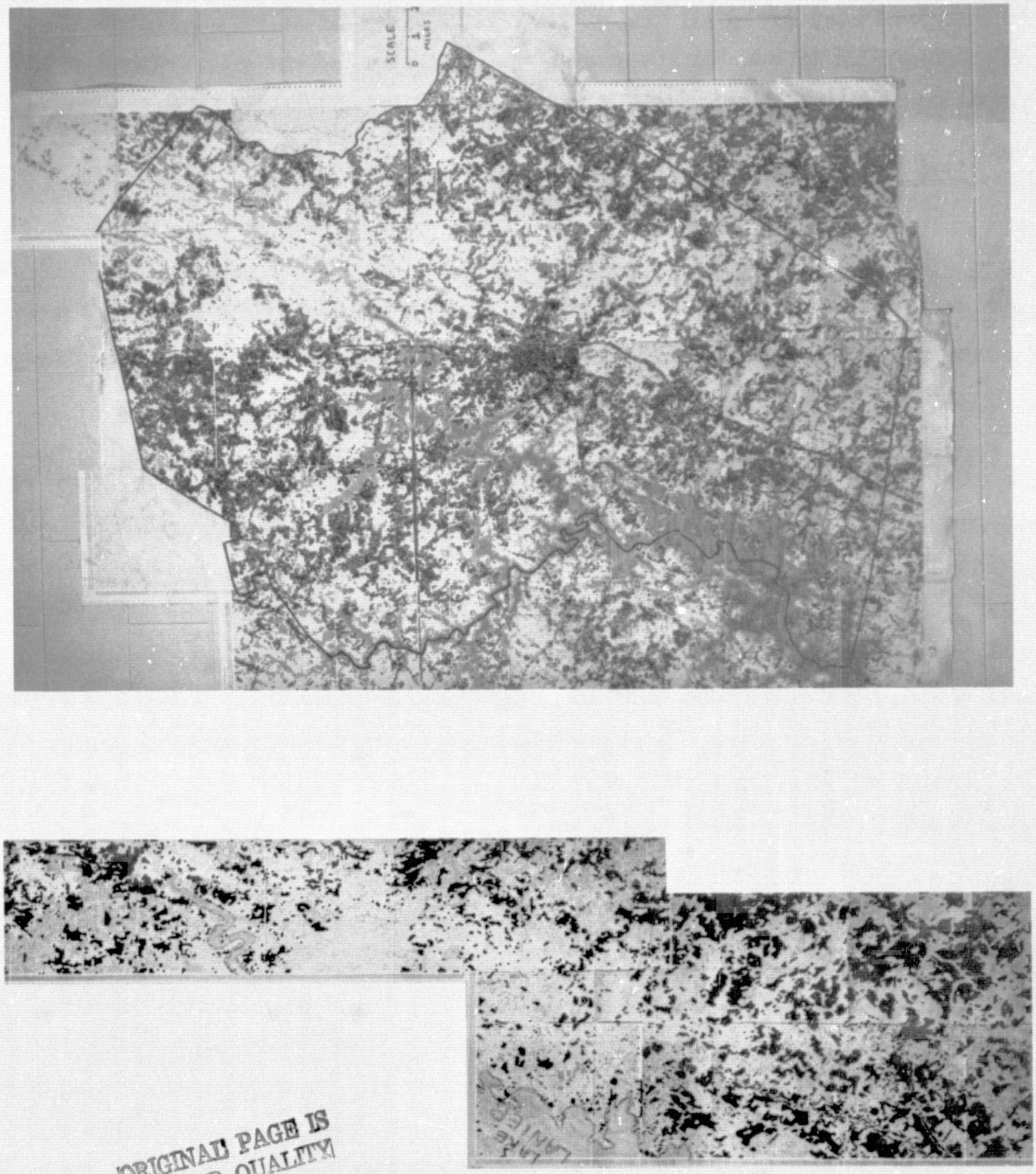


Figure 2.



ORIGINAL PAGE IS
OF POOR QUALITY

Figure 3.

in the April 1973 scene. Even though the geobotanical indicators are not as pronounced in the Hall County areas as in the Douglasville area, a definite correlation exists between the vegetation and the area geology. This same techniques should be used in other areas for future verification. A visual geologic interpretation study using the same premise of geobotanical indicators in being pursued at Georgia Southwestern College by Arden and Westra.*

* Personal Communication

VII. Summary and Conclusions

Results of this study indicate that digitally processed Landsat data can be a valuable future source of land-use/land-cover data for input into the G DOT Transportation Planning Land-Use Simulation Model if (1) the data is used along with other supplementary information and (2) if several conceptual problems are solved. Gridded data such as Landsat data are well suited for inclusion in a data base management scheme whereby several sources of information about a particular area may be overlaid and analyzed. The use of Landsat data was found to be a cost effective method for obtaining land-use/land-cover information for Hall County. In addition, the geobotanical investigation in Hall County led to a rough delineation of the Brevard Fault zone in the area. An elongated vegetation class representing loblolly pine was observed to parallel the trend of the Brevard. This is the same vegetation group that was observed to parallel the Brevard in Douglas County during the first phase of this project.

VIII. Continued Research

Work that will be done in the next phase of this project includes the transfer of various software techniques for Landsat rectification, geometric referencing and classification via Table Lookup to State of Georgia computers. Detailed interfacing with state user agencies will be required in this effort. An investigation of the available software for data base management will also be accomplished in this phase with effort toward the implementation of the most applicable technique on State of Georgia computers. The selected technique/techniques will possess the capability to incorporate land cover information generated from Landsat digital tapes.

APPENDIX A

UNSUPERVISED CLASSIFICATION OF LANDSAT MSS DATA

As discussed in Section 2 of this report, each resolution element for the Landsat scanner system represents an area on the ground of approximately 1.05 acres. Each resolution element in turn has a set of four measurements associated with it. These four measurements are the intensitites of light received by the detectors on the spacecraft in each of four spectral bands and may be considered a four-dimensional vector associated with each plot of ground. We would like to have some intuitive feeling for where the tip of this vector is located in four-space. Unfortunately, four dimensions are difficult to visualize so for an example we will take a three-dimensional vector. This might represent measurements in three regions of the spectrum instead of four. Now if we let each axis of a coordinate system represent intensities in one spectral region, we can visualize the location of each vector in three-space. For example, let us have three measurements (Vector A) normalized between 0 and 256:

Reading	Axis	Spectral Region
222	x	.5 - .6 micrometers
250	y	.6 - .7 micrometers
210	z	.7 - .8 micrometers

Figure A-1 shows the location of this vector in three-space. If we have another data vector B associated with a different area:

Reading	Axis	Spectral Region
234	x	.5 - .6 micrometers
220	y	.6 - .7 micrometers
230	z	.7 - .8 micrometers

Figure A-2 shows the location of vectors A and B in three-space. Now, we

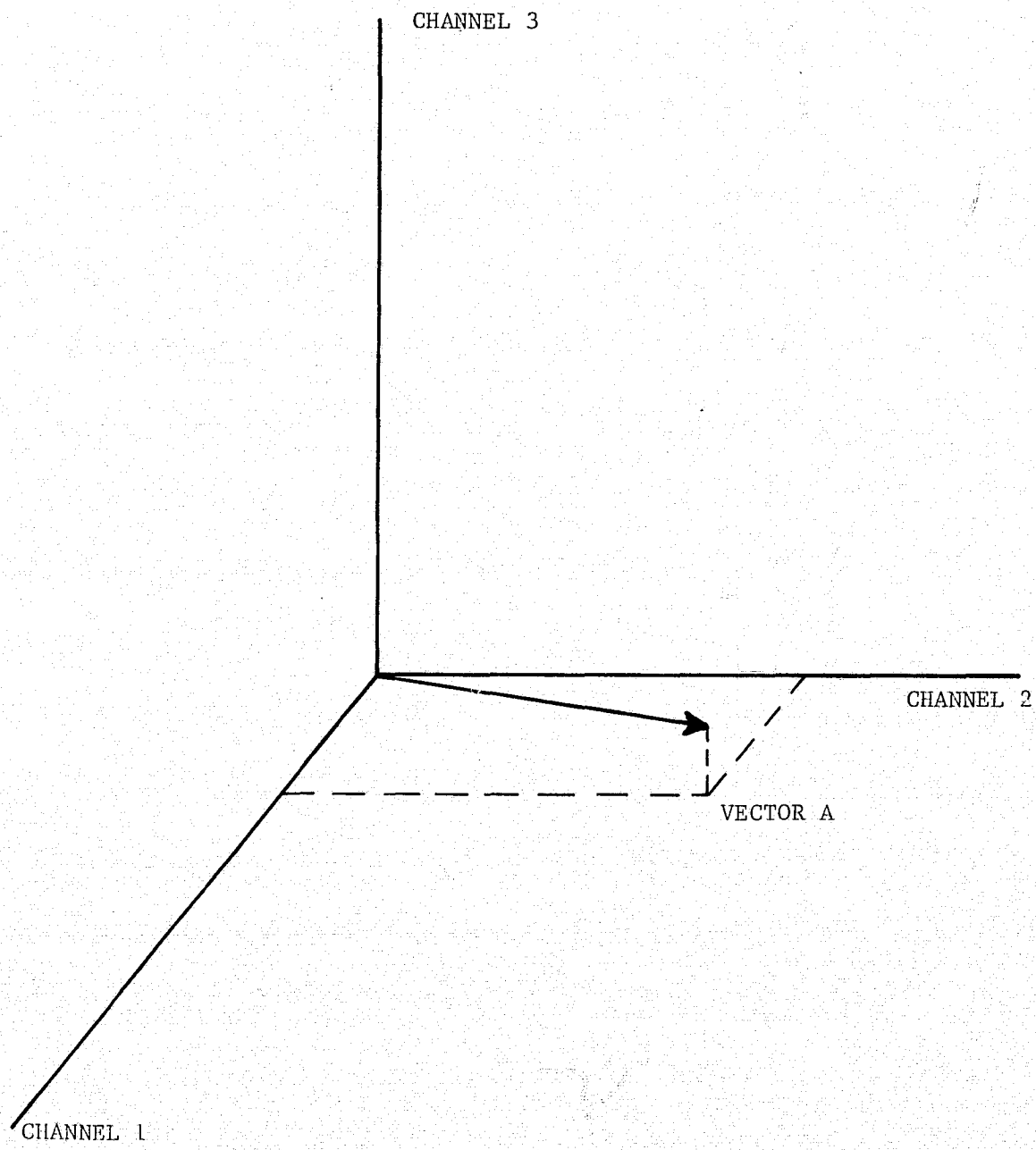


Figure (A-1)

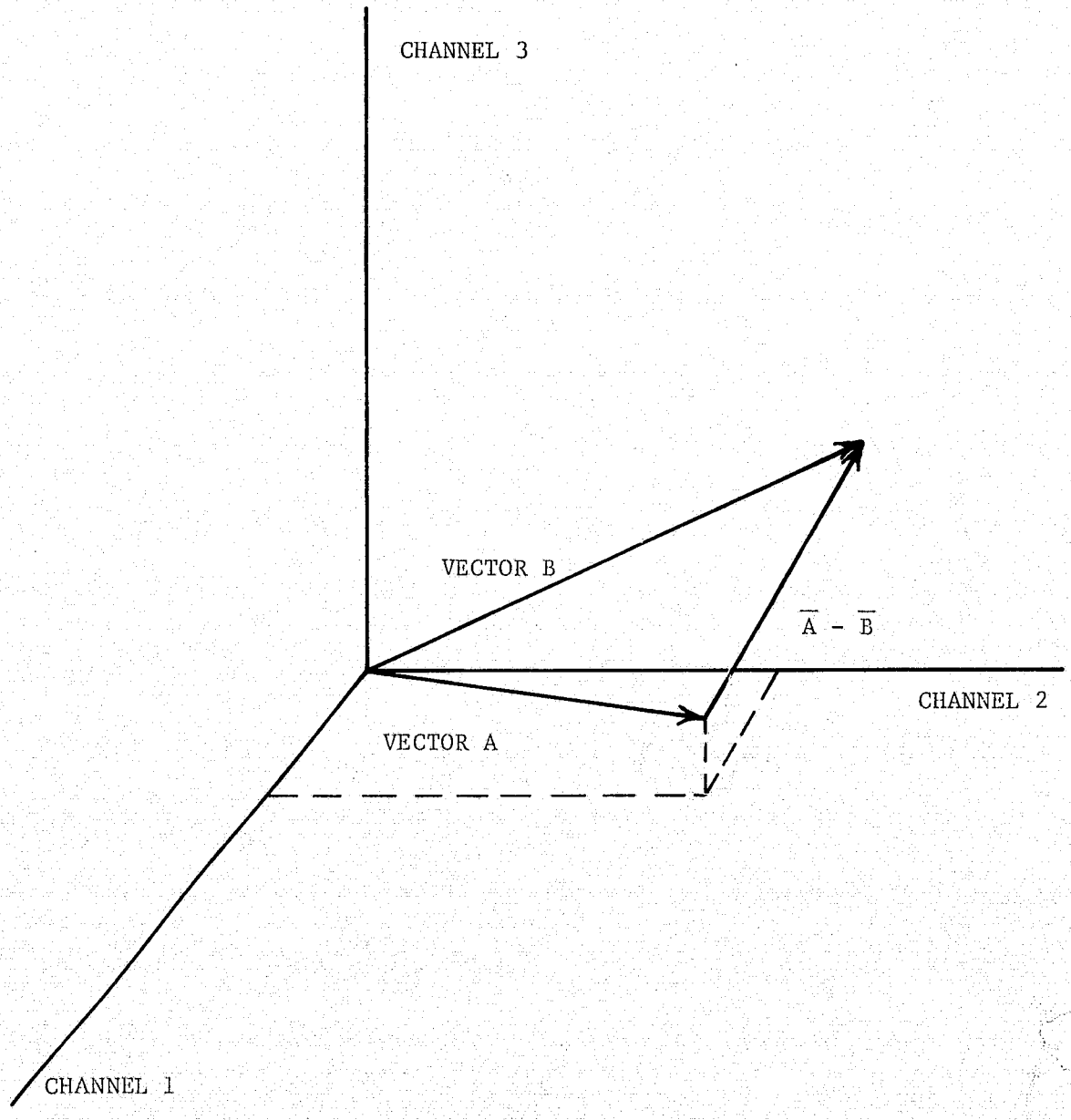

CHANNEL 3
VECTOR B
$\overline{A} - \overline{B}$
CHANNEL 2
VECTOR A
CHANNEL 1

Figure (A-2)

would like to have some measure of the difference between measurement vector A and B. The most logical choice for a difference measure is the distance between the two vector tips. This distance is given by

$$d = |\overline{A} - \overline{B}|$$

where $|\ |$ indicates absolute value of a vector and $\overline{A}$ means that A is a vector quantity. Expanding to evaluate d, we have

$$d = (a_1 - b_1)^2 + (a_2 - b_2)^2 + (a_3 - b_3)^2$$

where a_1 and b_1 are the first components of the vectors $\overline{A}$ and $\overline{B}$. The angle between A and B may also be calculated by

$$\Theta = \cos^{-1}\left(\frac{\overline{A} \cdot \overline{B}}{|\overline{A}|\ |\overline{B}|}\right)$$

where $A \cdot B$ is the inner product of A and B; i.e.,

$$A \cdot B = a_1b_1 + a_2b_2 + a_3b_3$$

Therefore,

$$\Theta = \cos^{-1}\left[\frac{(a_1b_1 + a_2b_2 + a_3b_3)}{(a_1^2 + a_2^2 + a_3^2)^{1/2}\ (b_1^2 + b_2^2 + b_3^2)^{1/2}}\right]$$

These equations will be used later. Another quantity that we would like to define is the mean vector. This vector is essentially the average vector associated with a set of N vectors. It is calculated by

$$\overline{M} = \frac{1}{N} \sum_{i+1}^{N} \overline{A}_i$$

where A_i is the ith individual vector. In terms of four components, we have

$$M_1 = \frac{1}{N}\sum_{i=1}^{N} (a_1)_i = \frac{1}{N} (a_{11} + a_{12} + a_{13} + \ldots + a_{1N})$$

$$M_2 = \frac{1}{N}\sum_{i=1}^{N} (a_2)_i = \frac{1}{N} (a_{21} + a_{22} + a_{23} + \ldots + a_{2N})$$

$$\vdots$$

$$M_4 = \frac{1}{N}\sum_{i=1}^{N} (a_4)_i = \frac{1}{N} (a_{41} + a_{42} + a_{43} + \ldots + a_{4N})$$

where a_{21} is the second component of the first vector considered and a_{23} is the second component of the third vector considered.

Now consider the situation in Figure A-3. The multispectral scanner scans a region normal to the flight path of the spacecraft. At any instant in time the rotating mirror displays an image representing approximately one acre on the ground and measurements in four regions of the spectrum are taken. The spacecraft velocity and the scanner rotation speed are such that after one scan line of data is taken, the spacecraft has moved forward enough so that the next scan line is contiguous to the first.

The massive amount of data that is taken for one Landsat scene of 100 nmi x 100 nmi can be analyzed digitally using unsupervised classification and the quantities described above. Each resolution element's radiance values are represented in four-space, and we would like to decide which resolution elements resemble others in an Landsat scene. A typical situation in three-space is shown in Figure A-4. It can be seen that there are several groupings of data points which probably represent radiance values from the same or similar

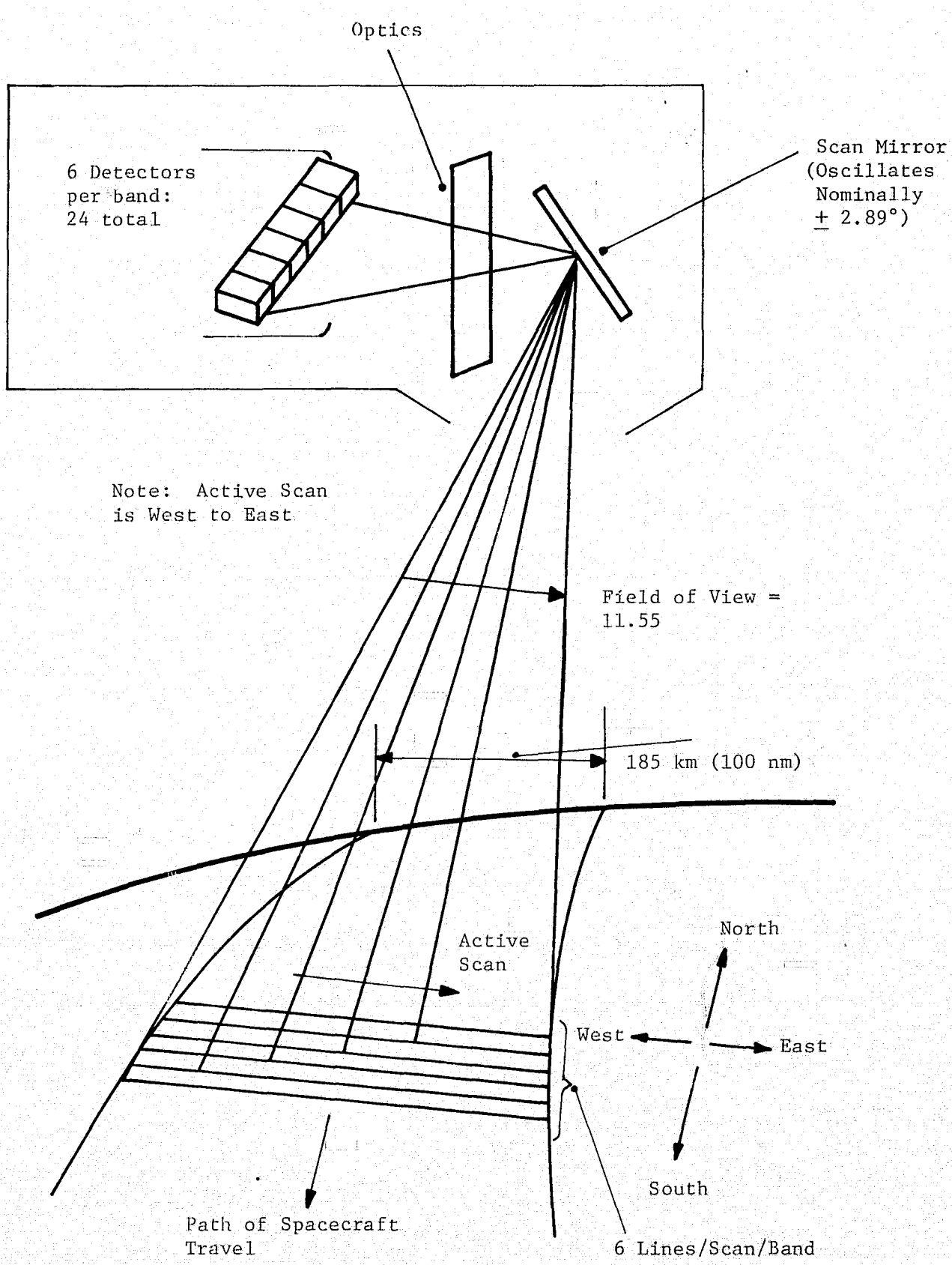

Optics
6 Detectors
per band:
24 total
Scan Mirror
(Oscillates
Nominally
± 2.89°)
Note: Active Scan
is West to East
Field of View =
11.55
185 km (100 nm)
Active
Scan
North
West
East
South
Path of Spacecraft
Travel
6 Lines/Scan/Band

Figure (A-3)

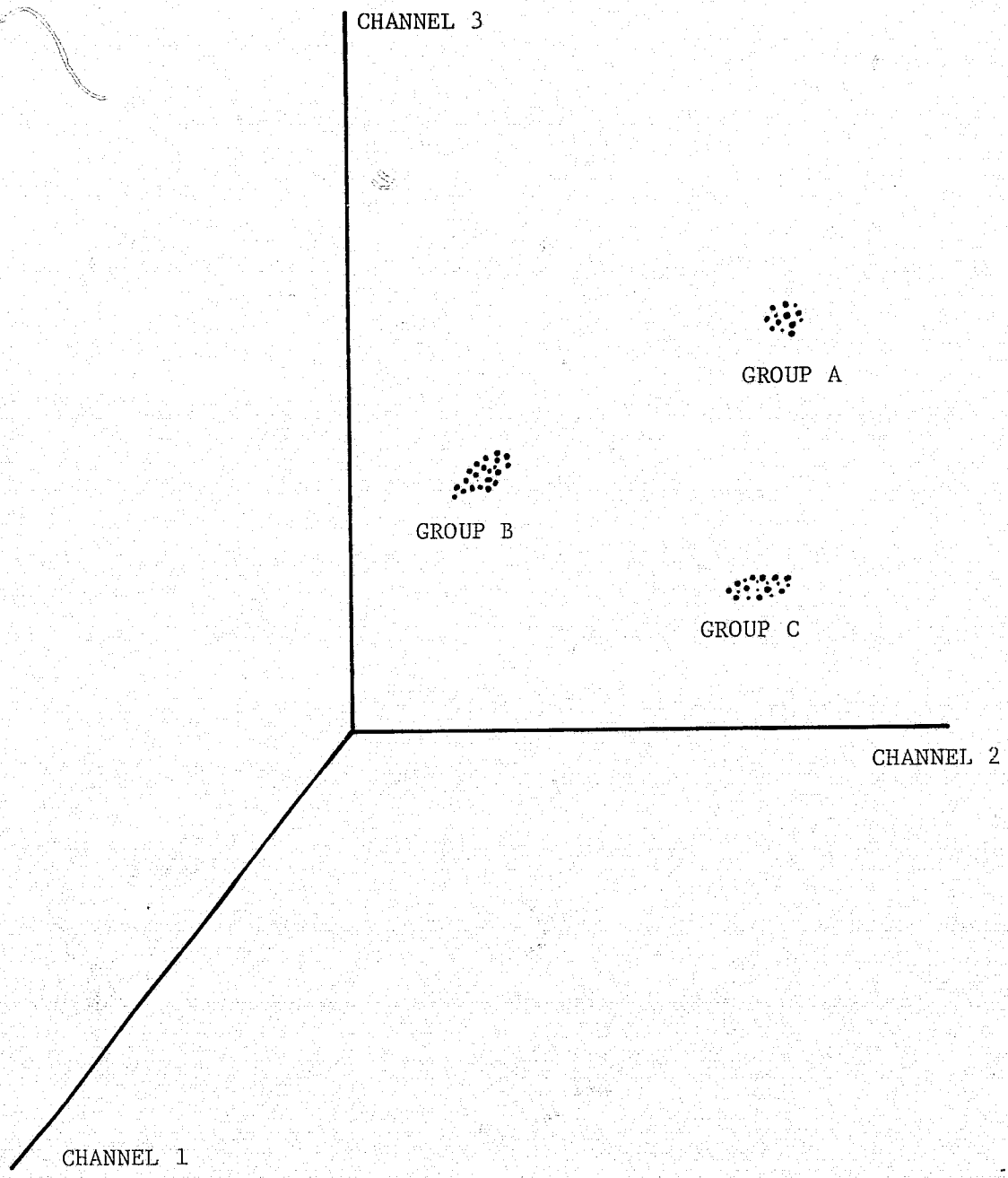


Figure (A-4)

ORIGINAL PAGE IS
OF POOR QUALITY

objects. For example, group A might be radiance values from trees, Group B from buildings, and Group C from water. Using the techniques developed above we may crudely represent each group or cluster by a mean vector and a chosen radius in three-space (Figure A-5). Any radiance vector that falls within this radius of the Group A mean is assigned to Group A. This follows similarly with other groups. If a vector does not fall within the prescribed radius of any of the previously defined clusters, a new cluster is generated using that vector as the first point. The data are usually considered sequentially considering one resolution element at a time for a whole area. One obvious disadvantage is that if the radii are chosen too small only a few points are allowed in a cluster and many additional clusters will have to be formed. The selection of radius values is essentially a trial and error procedure. As the number of clusters increases so does computer time and storage. This limits the number of clusters that may be considered. The present limit for our computer program is 20 clusters. If the program determines that a 21st cluster should be formed, then a statistical method considering the number of points in each cluster is used to decide which of the original clusters to eliminate. Actually a user of the program may set the maximum number of clusters to any number he likes up to 20.

The ASTEP (Algorithm Simulation Test and Evaluation Program) utilizes a sequential clustering as described above with minor modifications. Two iterations are made through the entire data set. The first iteration considers each measurement vector separately; i.e., the first vector is the first cluster; the second vector, if it is not within the specified radius of the first cluster, forms a second cluster and so on. If it is, the two vectors are averaged to form the cluster mean. It can be seen that this method may be biased due to the starting point in the data set. To eliminate this bias, a second iteration is made not allowing the mean vectors to be updated sequentially. The final product is a set of less than 20 groups of objects or things that look similar. These groups may often be associated with different objects on the ground such as water, rock, etc. These programs require a great deal of experience to determine radius values that will separate natural objects on the ground. A computer printout may be generated

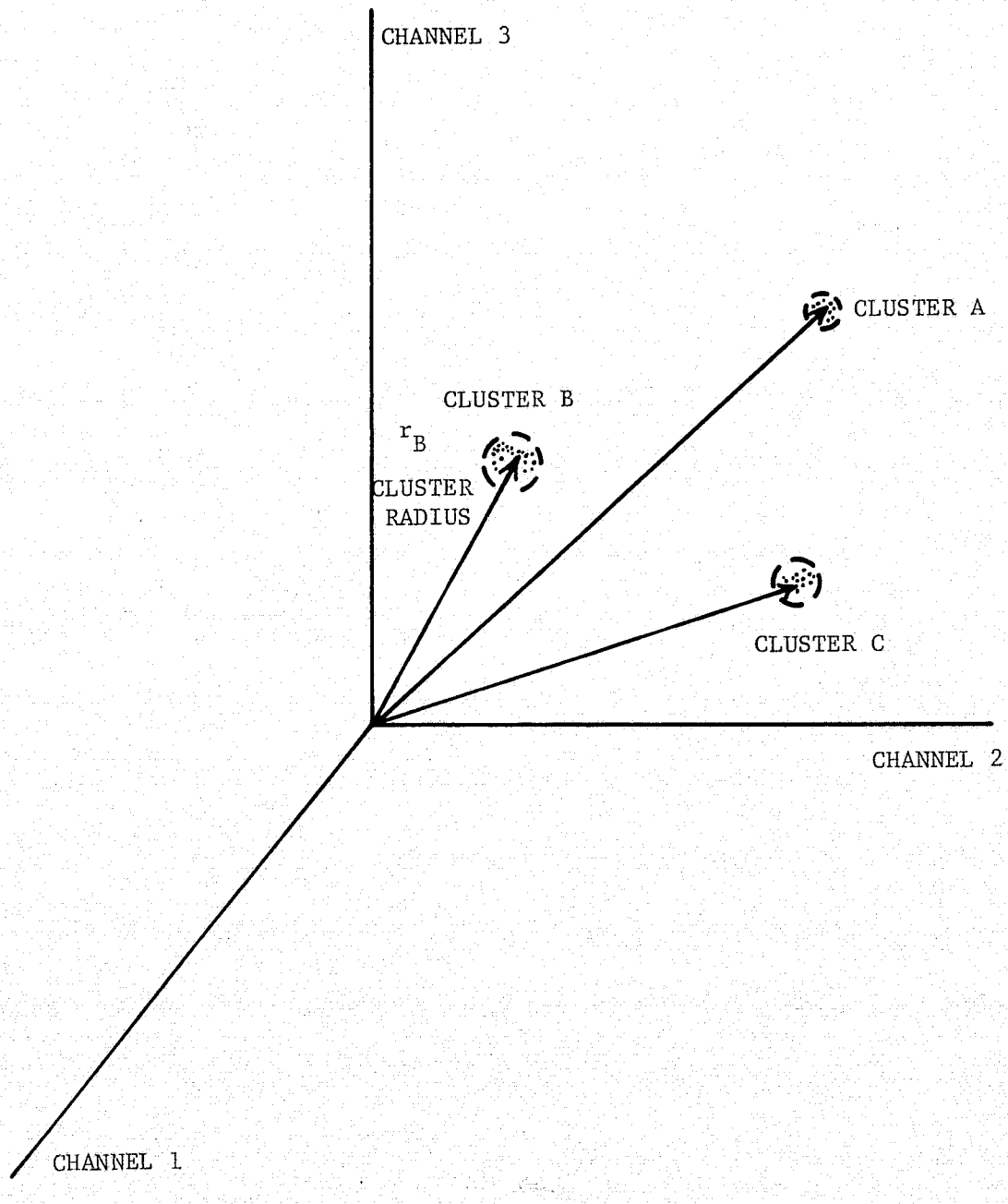


Figure (A-5)

ORIGINAL PAGE IS
OF POOR QUALITY

that represents the area that the satellite has imaged. Each character on the printout is associated with one of the previously determined clusters. Thus, one can see the spatial location of similar and dissimilar things on the earth's surface. With some checking with maps and aerial photos, these clusters may be used to represent major housing and development trends within a city as well as many other uses including geological.

APPENDIX B

SUPERVISED CLASSIFICATION OF LANDSAT MSS DATA

Landsat supervised classification is different from unsupervised classification in that instead of having a digital technique find separate clusters of measurement vectors in four space, we require a method which will classify each measurement vector into one of several classes whose position in four space has been previously computed. Each class in the supervised method represents a particular physical characteristic of the area imaged by the Landsat multispectral scanner system. For example, supervised classes may be defined as water bodies, commercial areas, cleared land, etc. To completely define a class we need more information than was used in unsupervised classification. Instead of a mean vector and a radius around it describing a class, we now use a method which allows us to describe the shape of the enevelope surrounding all points in one class. For example, in clustering we assumed that the points were symmetrical about the mean vector. Much statistical work has been done that indicates that most natural phenomena may be adequately described by a mean vector with a normal distribution of points around it, and not by a mean vector with an envelope equidistant in all spectral channels. In three-space a normal distribution resembles an ellipsoid about the mean (Figure B-1). Thus, if we wanted to describe an ellipsoid in three-space we would need to calculate the mean and the direction and length of the semi-minor and major axes. This may be done in three-space and extended into n-space by the calculation of the variance of the data from the mean. The variance, denoted by σ^2, is a measure of the elongation of the data in a particular direction. It may be calculated by standard statistical methods. An intuitive feeling for σ is found by the following equation. In 95% of the cases considered a random data value x will fall in the region defined by $|x - \mu| \leqq 2$ where μ is the mean value. Figure B-2 shows the region for one dimension. σ may be considered to be a difference in spectral response in one channel from the mean value. This may be extended to N channels of data. Since we are dealing with data randomly distributed within a normal distribution, we

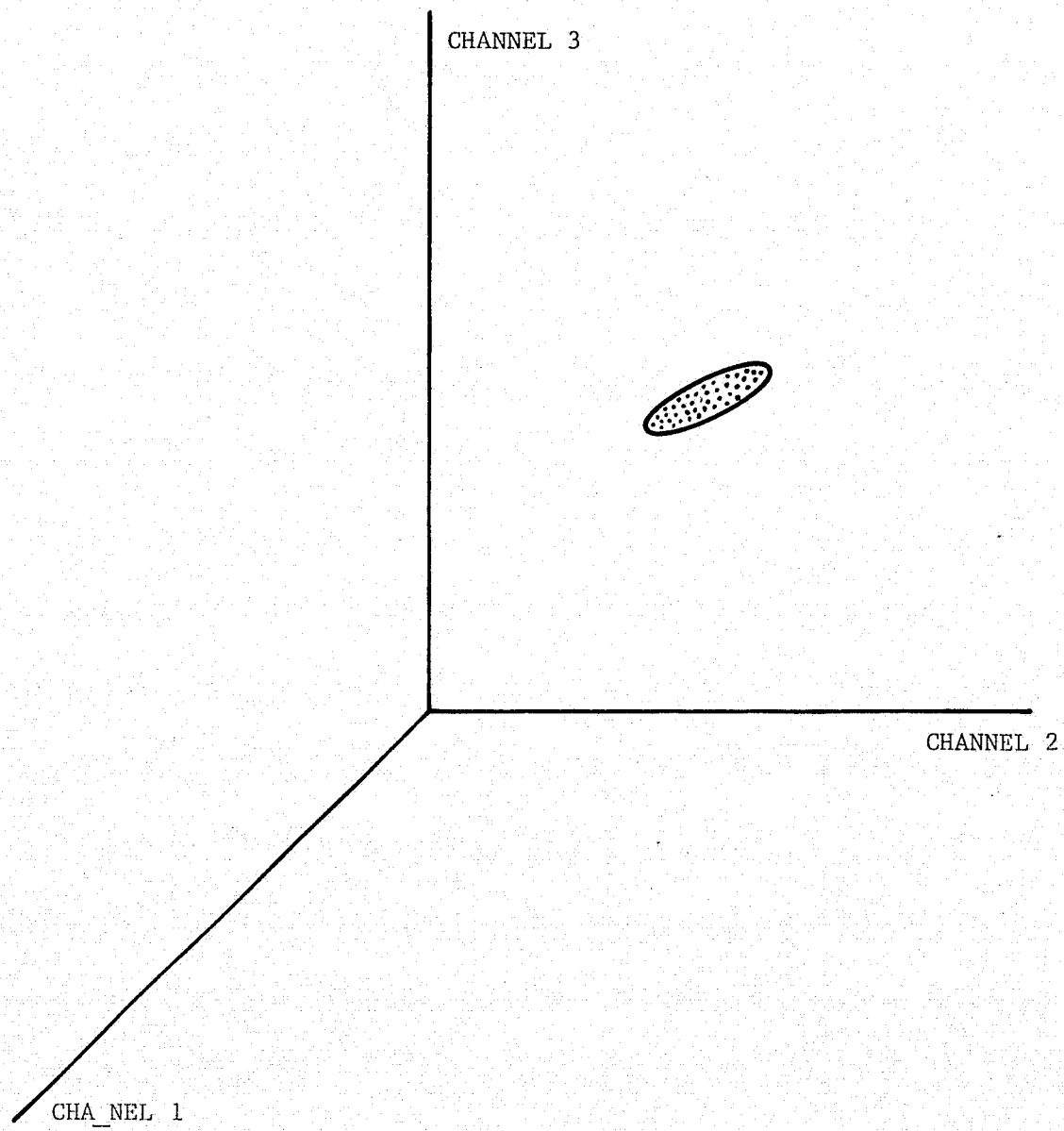


Figure (B-1)

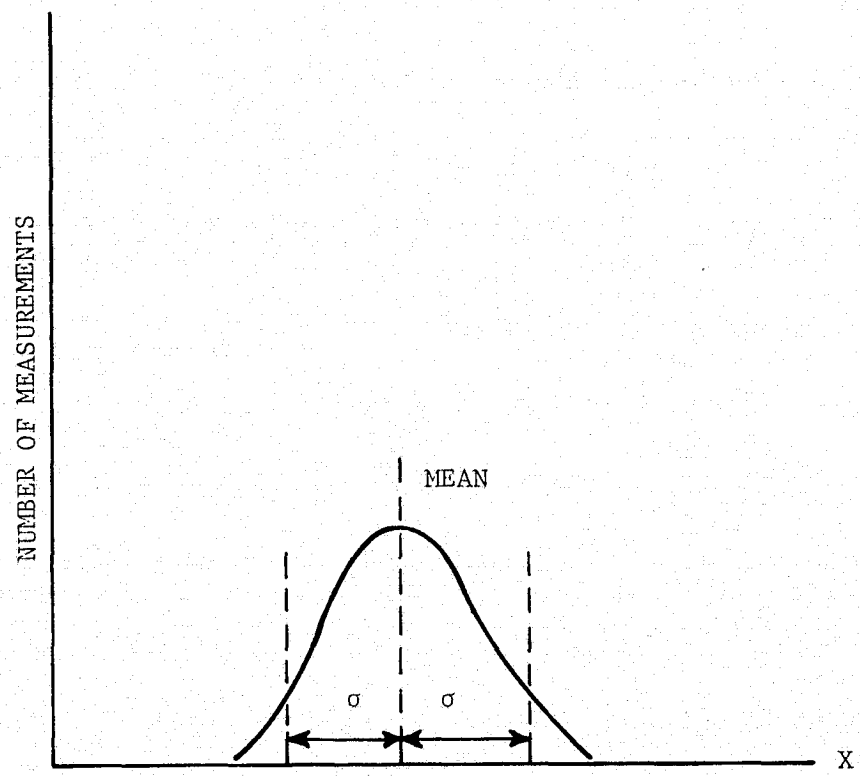


Figure (B-2)

ORIGINAL PAGE IS
OF POOR QUALITY

can only estimate the values for the mean and the variance associated with a particular class. In general, if a large number of samples are considered to calculate the mean value, the mean will approximate the true mean. If only a small number is considered there may be significant error in the calculation of the mean for a particular class. In multivariate analysis, the variances in each of the spectral regions are not the only considerations. If data values in some channels depend on data values in other channels, there will be a covariance between the two channels of data. For N channels this may be represented in an N by N matrix (the covariance matrix). If there is no interdependence, the channels are said to be independent and the covariance is zero. The best estimate for the mean and covariance matrix is given below.

$$\hat{\mu} = \frac{1}{N} \sum_{k=1}^{N} \overline{X}_k$$

where X is a single data vector and

$$\hat{\Sigma} = \frac{1}{N} \sum_{k=1}^{N} (\overline{X}_k - \hat{\mu})(\overline{X}_k - \hat{\mu})^t$$

where the t indicates the second matrix is transposed. If a sufficient number of samples are used to define the above population, the diagonal elements of the covariance matrix will be the variances squared for each channel and the off-diagonal elements describe the interaction between channels of data. A sample case for three channels is shown below.

$$\hat{\Sigma} = \begin{matrix} \sigma_1^2 & \sigma_1\sigma_2 & \sigma_1\sigma_3 \\ \sigma_2\sigma_1 & \sigma_2^2 & \sigma_2\sigma_3 \\ \sigma_3\sigma_1 & \sigma_3\sigma_2 & \sigma_3^2 \end{matrix}$$

If the channels of data were independent then

$$\hat{\Sigma} = \begin{matrix} \sigma_1^2 & 0 & 0 \\ 0 & \sigma_2^2 & 0 \\ 0 & 0 & \sigma_3^2 \end{matrix}$$

Thus, given a sufficient number of radiance vectors that are identifiable with one class of natural phenomena, an estimated mean and a covariance may be computed for the total population of that phenomena. By comparing each data vector to these estimates, we may decide if that data vector in fact represents a certain class of material, i.e., water. This will be discussed further below.

Discriminant functions are developed in classification theory for special distributions of data. These discriminant functions are the criteria by which a radiance vector may be assigned to a particular class. Since the normal density function is very often used to represent reality, the discriminant function for it has been known for some time. The discriminant function for a radiance vector $\overline{X}$ to be in the ith class is

$$g_i(\overline{X}) = -1/2\ (\overline{X} - \overline{\mu})^t \Sigma_i^{-1}\ (\overline{X} - \overline{\mu}) - \frac{d}{2} \log 2\pi - 1/2 \log |\Sigma_i| + \log P(w_i)$$

where $\overline{\mu}$ is the mean vector and Σ_i^{-1} is the inverse of the ith class covariance matrix. In general the $\frac{d}{2} \log 2\pi$ term is only additive and is not a function of which class is considered. Thus it may be ignored. By replacing $g_i(X)$ by $f[g_i(\overline{X})]$ where f is a monotonically increasing function, the resulting classification is unchanged (Ref. 1). Thus if we take the exponential of $g_i(\overline{X})$

$$Q_i = f[g_i(\overline{X})] = \frac{e^{-1/2(\overline{X} - \overline{\mu})^t\ \Sigma_i^{-1}(\overline{X} - \overline{\mu})}}{|\Sigma_i|^{1/2}}$$

Now for every radiance vector $\overline{X}$ a Q is calculated for each class previously defined. The vector is then assigned to the class that has the largest value of the discriminant function Q. This proceeds until all the radiance vectors for the imaged area are processed. One pitfall of this method is that a vector is always assigned to one of the classes even though it actually may not be similar to any of the classes. This problem may be attacked by a thresholding approach.

Since the $|\Sigma_i|^{1/2}$ and the Σ_i^{-1} need only be calculated once for each class, the most time consuming part of the calculation for each data vector is the quadratic computation of $(\overline{X} - \overline{\mu})^t \Sigma_i^{-1} (\overline{X} - \overline{\mu})$.

Thus the supervised method of classification uses statistics generated by a large number of samples to describe each class of data that a vector may be assigned to. Once these statistics are calculated, the discriminant function must be calculated for each class for every data vector. The vector is then assigned to one of those classes by inspection of the discriminant functions.

The ASTEP program has the supervised classification scheme described above implemented as a classification module. Training sets of data are usually located by comparing clustering outputs as described above with aerial photos or maps. The homogeniety of each training set may be tested by histograms of the data. Next, the statistics for each training class are computed and saved on magnetic tape. When the supervised module is requested, these training set statistics provide the necessary information to be used to classify other multispectral data into the selected classes.

APPENDIX C*

THE GEORGIA TRANSPORTATION PLANNING LAND-USE MODEL

The Georgia Transportation Planning Land-Use Model, shown in Chart 1, has three distinguishing features:

1. Alternative approaches are used in the estimation of key exogenous variables as well as in the locational assignment algorithms employed.
2. Judgemental human intervention is explicitly provided for at key junctures in the modeling process.
3. The model user assumes an important participating role in model planning, development testing, and implementation.

The state employment and population forecasts are integrated with national forecasts through the use of the Shift and Share technique, as well as econometric and input-output models available for the State of Georgia. The outputs of these models will be judgementally compared with employment and population estimates made by U.S. Government and other agencies and by the Georgia Office of Planning and Budget.

The Shift and Share, Input-Output, and Delphi techniques will be supplemented by trend extrapolation and judgement in estimating employment growth by counties.

Three techniques are identified in Chart 1 for estimating population and households at the county level. These techniques, described in more detail below, are also integrated by judgemental comparison. Previous difficulties in estimating employment, population and housing growth at the sub-county level led to the tentative conclusion that such estimates should be derived primarily through the sampling of expert opinions.

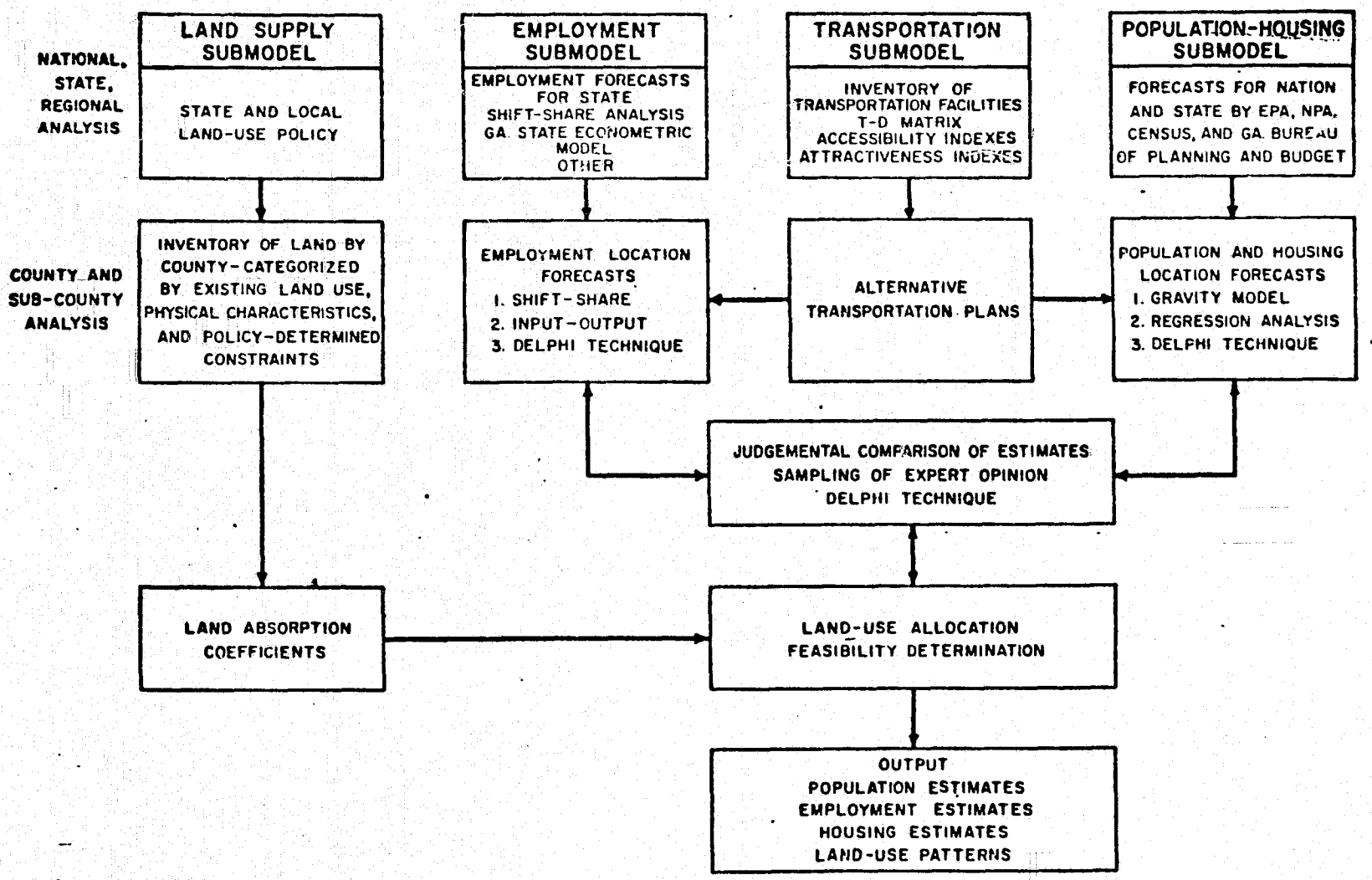
The interaction between these submodels and the land supply and transportation submodels in the framework shown in Chart 1 is discussed on the page following the chart.

* This appendix was copied directly and taken verbatim from a University of Georgia publication with their permission. A disscussion of revisions in the model is found in "The Georgia Transportation Planning Land Use Model: Development of a Policy Sensitive Impact Analysis Planning Tool", M.P. Hailperin, H.S. Maggied, C.S. Floyd, presented at 57th Annual Conference, American Institute of Planners, San Antonio, Texas, October 27, 1975.

ORIGINAL PAGE IS
OF POOR QUALITY

CHART I

GEORGIA TRANSPORTATION PLANNING LAND-USE MODEL FRAMEWORK



Employment and Population Submodel

A serious shortcoming of some small area employment forecasts is their lack of consistency with national, state, and major sub-state regional projections. The Georgia model will avoid this pitfall by allocating state and regional employment projections to multi-county and then to smaller areas.

National industry employment projections are prepared by several agencies, including the U. S. Department of Labor, the National Planning Association, and the U. S. Department of Commerce. These estimates will be used in combination with the output of the Georgia State Econometric Model to forecast future employment and population at the state level. An alternative forecast will be developed using the Shift and Share method of analysis, a technique which has been widely applied in the field. Its application in regional and small area employment projections has been the subject of recent discussion in the literature.[11]

The Shift and Share technique, combined with judgemental modifications of the regional share component, will also be employed to allocate projected Georgia state employment to sub-state multi-county regions. In view of the

---

11. For a description and evaluation of this technique, see Charles F. Floyd and C. F. Sirmans, "Shift and Share Projections Revised, "Journal of Regional Science, Vol. 13, No. 1, 1973, pp. 115-20; L. D. Ashby, Growth Patterns in Employment by County, 1940-1950 and 1950-1960, Vols. 1-8 (U. S. Department of Commerce, Office of Business Economics), Washington, D. C.: U. S. Government Printing Office, 1965; H. J. Brown, "Shift and Share Projections of Regional Economic Growth: An Empirical Test," Journal of Regional Science, Vol. 9, (1969), pp. 1-18; H. J. Brown, "The Stability of the Regional Share Component: Reply," Journal of Regional Science, Volume 11, (1971), pp. 113-114; C. F. Floyd, The Changing Structure of Employment and Income in the Regions of the United States, Vols. 1-6, Washington D. C.: U. S. Department of Commerce, Economic Development Administration, 1971; C. C. Paraskevopoulos, "The Stability of the Regional-Share Component: An Empirical Test," Journal of Regional Science, Volume 11, (1971), pp. 107-112.

apparent instability of the critical regional share component, multiple regression techniques will be employed in an effort to identify the important determinant variables affecting it. Forecasts based upon the Shift and Share technique will be compared to similar forecasts prepared by the bureau of Economic Analysis, U. S. Department of Commerce.[12]

Estimation of Employment by Counties: As in previous modeling formulations, the location of future employment becomes the determining influence upon the location of housing, commercial and public service development. The problems in forecasting employment by counties are of such dimension as to require two independent methods of estimation. The first technique, identified in Chart 1 as Shift-Share-Judgement, was discussed earlier.

The research team will supplement this method by use of the Delphi technique for obtaining a consensus of expert opinions. A research group in Regional Environmental Systems Analysis at the Oak Ridge National Laboratory has experimented successfully with the Delphi technique in forecasting the location of future development in a pilot project in the Knoxville, Tennessee, area.[13] On the basis of this experiment, the Delphi technique is being adapted to a 16-county Delphi land-use study at ORNL.[14]

Description of the Delphi Technique

The prospect of using the Delphi Technique in forecasting employment and residential location at the county and sub-county levels requires a careful review and consideration of the applicability and reliability of the

---

12. Bureau of Economic Analysis, U. S. Department of Commerce, OBERS Projections-Regional Economic Activity in the U. S., Washington, D. C.: U. S. Water Resources Council, 1972.

13. Osbin L. Ervin and Charles R. Meyers, Jr., The Utilization of Local Opinion in Land-Use Simulation Modeling, A Delphi Approach, Regional Environmental Systems Analysis, Memo Report #73-8, Oak Ridge, Tennessee, February, 1973.

14. Ibid., p. 57.

technique for these purposes.

The Delphi technique, a methodology for eliciting and refining expert or informed opinion, has gained increasing recognition in recent years as a result of experiments with its use in obtaining a consensus of expert views on a variety of topics at the Rand Corporation, the University of California, the University of Michigan, and elsewhere.[15] The technique involves the sequential administration of a questionnaire to a panel of experts in successive "rounds". The answers of the group are summarized at the conclusion of each round and the respondents are then asked to reconsider earlier responses in light of the summary of group answers. The theory of the technique is that successive iteration and feedback of group responses will result in a gradual approximation to the "true" prediction of "value" sought. The results are finally expressed in measures of central tendency and dispersion for the group as a whole. Theoretically, the anonymity of the respondents and the absence of face-to-face discussion eliminates the probable distorting effect of dominating personalities present in conventional group or committee discussions. The results of experiments comparing the performance of groups in face-to-face discussion with groups interacting through an anonymous questionnaire with controlled feedback have indicated that the estimates of the Delphi groups were more accurate than were those of the groups in face-to-face interaction.[16]

Experiments with the Delphi technique were conducted at the Oak Ridge National Laboratory in an 8 by 10 mile study area in North Anderson County, Tennessee. The experiments were designed to estimate the location of new factories anticipated in the area in the future. Participants from municipal government, utilities, banking, industrial development, real estate, and local planning agencies were selected for one panel because of the history of their

---

15. A citation to the principal published works in the field is found in the Notes and Selected Bibliography to Regional Environmental Systems Analysis, Memo Report #73-8.

16. Ibid., p. 10.

interest and involvement in the uses of land in the area. The respondents were first asked to identify, in order of importance, the principal factors determining future conversion of land to each category of use, to rate their individual knowledge of the general subject of the questionnaire, and to recommend others whom they considered to be knowledgeable in the area of expertise. In this experiment, it was found that each successive iteration of the questionnaires, accompanied by a feedback of the responses of the group as a whole, resulted in a gradual convergence of the responses about a median opinion, with a successively lower interquartile range.

The results of the manufacturing study conducted at Oak Ridge National Laboratory led to the decision to develop separate panels for each two-digit SIC category of manufacturing activity, as well as two panels for wholesaling activity and panels for predicting residential and retailing activity.

The authors of the ORNL report indicate that "there is reason for optimism about the potential of a Delphi approach to weighting indices of land-use conversion and isolating growth areas," and that "the second round index weights are a better reflection of real world processes than those which might have been developed by other methods."[17]

Successful experimentation with the application of the Delphi method in predicting land-use in the ORNL test area has provided the University of Georgia research team with an operationally tested supplementary technique for estimating future land-uses for the Georgia prototype corridor study. Observation that the ORNL experiment and related workshops "established an atmosphere of broad, area-wide involvement and cooperation in the land-use simulation program" provides an added inducement for use of the Delphi method in the Georgia study.

---

17. Ibid., p. 41.

Housing and Population Submodel

State and county population estimates derived from the employment submodel will be compared to population estimates made by the Bureau of Economic Analysis, the National Planning Association, the Environmental Protection Agency, and the Georgia Bureau of Planning and Budget.[18] Because the relationship between the population and employment estimates is so critical, the two will be carefully reviewed for consistency.

Estimation of Housing by Counties: Virtually all previous land-use models have allocated households by some variation of the gravity-model approach. The gravity model had its origins in The Law of Retail Gravitation, by W. J. Reilly, who observed that the relative retail attraction of cities for retail trade would vary directly with population size and inversely with distance between population centers.[19]

The gravity-model approach to estimating future residential development assumes that most workers will seek residences convenient to their place of work with decreasing proportions of workers commuting long distances as measured in miles or in time.[20]

The principal criticism of the gravity-model approach to the estimation of the location of future residential development has been that often some of the parameters used in the estimating equations must be determined

---

18. Environmental Protection Agency, Population by County, Historic (1940-1970) and Projected (1980-2020), Region IV, Atlanta, Georgia, July. 1972; Office of Planning and Budget, State of Georgia, County Population Projections: Georgia 1980, Atlanta, Georgia, Office of Planning and Budget, September, 1972.

19. W. J. Reilly, The Law of Retail Gravitation, (New York: Knickerbocker Press, 1931).

20. Projective Land-Use Model, Volume I, Plan Making with a Computer Model, Chapter Four. For the equations used in estimating the work-to-home probability used in the Plum Model, see Projective Land-Use Model-Plum, Volume II - Theory and Application, pp. 80-81.

judgementally, due to the wide variations in commuting behavior among different regions, income classes, trip purposes and occupational groups. Preliminary studies of population growth and urban development in Georgia confirm the findings of earlier studies that distance alone (or time distance) from urban population centers provides an unsatisfactory explanation for commuting behavior and resulting land development patterns.

The recent publication of summary tables showing county population by place of work, from the U. S. Census of Population 1970, Fourth Count Summary Tapes, and supplementary census tapes, provides substantial information on commuting patterns in the United States. Experiments in analyzing commuting patterns in East Tennessee have indicated that factor analysis techniques can be used successfully in determining the similarities or differences among sub-areas with respect to commuting patterns.[21]

The Georgia Transportation Planning Land Use Model will take advantage of this and other recent research directed toward improvement in the techniques for measuring the all-important "Attractiveness" or "Accessibility" indices used in the application of the gravity model technique as well as the exponents for Distance or Time Distance used in the gravity-model equations.

The accepted shortcomings of the gravity-model approach in forecasting the location of future housing, commercial, and related development dictate that alternative approaches to such estimation should also be employed in the Georgia Transportation Planning Land-Use Model. Chart 1 identifies the second technique to be employed as Multiple Regression using national, state, and county variables. This widely used technique has been applied in one way or another in virtually every major land-use modeling effort.[22] The EMPIRIC Model, used in the Boston Area Transportation Study, estimated the

---

21. Charles R. Kerley, Labor Commuting Sub-Areas of the East Tennessee Development District, Regional Environmental Systems Analysis Program, ORNL, NSF-ER-60, November, 1973.

22. H. James Brown, et al., Op. Cit., pp. 22-23.

rate of growth of sub-areas by regressing the relationship between growth in a past period and a number of "locator" variables, including densities of land-use, zoning practices, the quality of water and sewerage services, automobile and transit accessibilities, and relative attractiveness indices for various locations.[23] Recent improvements in the quality of data available and in the techniques of analysis and interpretation identify multiple regression as a key approach to the estimation of housing location in the model.

The use of the Delphi technique as a supplementary method of forecasting future changes in county employment was discussed earlier. This successful experimentation suggests that this technique be used as an additional method of forecasting future housing location in the Georgia model. The Delphi technique has particular advantages for small area analysis, where data limitations are particularly severe and where the exercise of judgement requires extensive local experience and knowledge. For this reason, the Delphi technique is listed in Chart 1 as one of a number of alternatives in estimating future housing development at the county level but is identified as the major technique for estimating future development for sub-county areas. This technique may also provide a means of establishing, for county and sub-county areas, measures of attractiveness to be used in the multiple regression approaches.

Relationships Between Land-Use Forecasting and Transportation Planning

Traditionally, land-use forecasts have been used as inputs in transportation planning with only trivial feedbacks of transportation investments on future patterns of urban development. The principal emphasis in transportation studies has been upon the estimation of trips, model splits, and network assignments.[24] Major transportation study budgets have ranged in

---

23. Donald M. Hill, "A Growth Allocation Model for the Boston Region," Journal of the Americal Institute of Planners, XXXI, May, 1965, pp. 111-120.

24. H. James Brown, et al., Empirical Models of Urban Land Use: Suggestions on Research Objectives and Organization, op. cit., Chapter 1.

cost between $2 million and $5 million, with the dominant portion of expenditures used for gathering original data on travel behavior. Land-use modeling expenditures on the average have represented less than six percent of total costs.[25]

Chart 1 indicates that in the proposed framework for the Georgia Transportation Planning Land-Use Model, principal emphasis will be placed upon the interaction of proposed changes in transportation with employment and population growth, location, and land-use. The initial forecasts for the 17-county test area will assume the completion on schedule of presently planned transportation facilities. A time distance matrix for the major urban areas of the State will be developed from data available from the Georgia Department of Transportation. The models will then be "run" to test the impact of changes on the timing, location, or nature of alternative transportation improvements. Thus, an attempt will be made to assess the influence of transportation investment decisions and their impact on employment growth, development, and land-use.

Theoretically, and actually, the effect of a given change in transportation facilities will depend upon the nature and extent of other changes in transportation facilities in other areas competing for development. The effect, for example, of the opening of a major highway will depend upon whether or not it is assumed that other facilities are completed concurrently or soon thereafter. This of course adds a new dimension of difficulty to the analysis and suggests a possible further application of the Delphi technique. A panel of experts on employment location, for example, might be asked first to assess the effect of the completion of a single facility and then to separately assess the effect of the concurrent opening of that facility along with a competitive one. The alternative to using this technique for sampling expert opinion has proved to be exceedingly costly and of doubtful accuracy. Needless to say, historical analysis of the apparent influence of alterations in transportation facilities will be used in the design of the questionnaires and in checking the results judgementally.

---

25. Ibid, pp. 91-95.

Land Supply Submodel

The proposed land-use submodel consists of three major components: a continuing inventory of the usable land supply; the determination of land absorption coefficients; and a set of policy determined constraints. Georgia Department of Revenue property tax records will furnish the principal source of current land-use data.

Consistent with criteria of simplicity and minimum specification of model output, and mindful of the shortcomings of more ambitious approaches, the University of Georgia model research team proposes the following specifications for the land supply submodels.

1. The usable but unused land supply in each county will be estimated by subtracting from the total land supply all presently developed land, together with all land identified for public and semi-public use, existing and proposed water bodies and flood plains, and other land suitable for development, based upon slope and other characteristics.

2. Future land absorption will be estimated in the limited categories of residential, manufacturing and wholesaling, commercial (including retailing and service), agricultural, and public.[26]

The objectives of the University of Georgia research modeling team are to develop a simple, objective, transportation planning land-use model which is theoretically sound, and can be used and understood by State transportation planners to assess the impact of alternate transportation routes.

A review of previous modeling efforts led to the conclusion that the Georgia model should allow for human intervention and evaluation at several stages, and that the introduction of judgemental estimations should be made explicit in elaborate equation systems. The limited financial resources available for the project have reinforced this decision to trade-off

---

26. The BASS model, for example, forecast future land-uses for six classes of residential use, as well as for manufacturing and wholesaling, service employment, commercial, public and recreational, and for agriculture, mining and construction. See Jobs, People and Land, op. cit., Appendix tables.

theoretical elegance and mathematical and computational sophistication for simplicity, economy, comprehension, and feasibility.

The goal of the project team is to build a model which will retain the advantages of a computer-based, iterative approach, but which will avoid some of the demonstrated shortcomings of large econometric, land-use computer models. Supplemental use of expert opinion survey methods will assure local inputs to the modeling process and enhance community support.

The research framework outlined in Chart 1 will rely at every point upon the cumulative knowledge of the land-use modeling art developed by others working in the field. In this respect, it is particularly fortunate that the preliminary results of a multi-million dollar NSF RANN research contract for a Regional Environmental Systems Analysis Program at Oak Ridge, Tennessee, are becoming available at this time. Hopefully, much time and money can be saved in the present modeling effort through the opportunity to share in the current findings of the Oak Ridge and other on-going projects.

An overview of the state of knowledge and the art of land-use forecasting provides a sobering influence and reinforces the conclusion that a variety of approaches rather than a single forecasting technique should be employed. However, caution must be exercised to assure that the limited resources available are not dissipated over too broad an area of application.

Needless to say, it is impossible to predict at this point in time the weight which will be accorded the alternative estimating techniques identified in Chart 1. Hopefully, potential problems of cumulative error can be avoided by the use of independent techniques and human intervention.

The final processes of judgemental weighting will depend upon the quality and quantity of data available, statistical measures of reliability, the track record of both the researchers and the techniques employed, and most importantly, the dispersion among the estimates provided by the different approaches.

Previous experience has indicated that the principal value of land-use planning models lies in their use for measuring the impact of alternative private and public policy decisions. The responsiveness of the model output to assumed changes in key variables affecting future land-use will be of

extreme importance in evaluating the Georgia model.

Long-term land-use plans will become a virtual necessity for transportation and other planning agencies during this decade. The criteria set forth in the Federal guidelines for consideration of the economic, social, and environmental effects of highways specify the use of a systematic, interdisciplinary approach, alternative analytical procedures, and public involvement.[27] Hopefully the framework outlined for the Georgia Transportation Planning Land-Use Model will facilitate the weighing of transportation alternatives and the reaching of sound judgements.

---

27. Department of Transportation, Federal Highway Administration. Report to Congress on Section 109 (h), Title 23, United States Code - Guidelines Relating to the Economic, Social, and Environmental Effects of Highway Projects. Washington, D. C.: U. S. Government Printing Office, 1972.

Department of Transportation
State of Georgia
Office of Materials and Research
15 Kennedy Drive
Forest Park, Georgia 30050

Research and Development Bureau

March 22, 1976

Mr. Nick Faust
Research Scientist
Electro-Magnetics Laboratory
Engineering Experiment Station
Georgia Institute of Technology
Atlanta, Georgia 303322

Dear Nick:

I have reviewed your draft report: Study of USGS/NASA Land Use Classification System and found no major problems. Other than the suggested editorial changes that we discussed I suggest that you revise "Appendix C" to reflect the changes that have occurred with the GaTPLUM model development through Phase I. The enclosed paper "The Georgia Transportation Planning Land Use Model: Development of a Policy Sensitive Impact Analysis Planning Tool" documents the major elements of the model' s initial stages. As you will note substantial changes have occurred since the state-of-the-art paper to which you refer was authored. I believe that it would benefit your report substantially if you were to update the relevant section to comport with our project successes.

In reviewing the topics discussed during our meeting in mid-December, it becomes apparent that the evident discrepancies between the Landsat data and aerial photography results, not only from interpretation, but also (and equally important) from differences in definitional categories structured by various disciplines. The disparity seems to derive in part from the software system incompatibilities. Concomitant with that problem is a lack of a more complete methodology for collecting ground truth data. Also, photo-interpretation compounds the problem. Although adjustments have been made to compensate for these inherent problems, the images viewed and translated often do not relate to the real world. In order to obviate these problems, it necessitates development of a more discrete technique for resolving ground truth.

Concerning the utility to the Department of LANDSAT digital data for planning purposes, it does provide a viable mechanism for identifying changes in land-cover which will enhance macro transportation planning. The ability to automate classification of land-use by observing changes in land-cover affords transportation planners significant opportunities in their activities. The range that LANDSAT covers from satellite altitudes coupled with the detailed level offers the Department an exceptional tool to add to the developing cadre of policy analysis tools. We will be interested in reviewing further developments of your project.

If there are further questions, please contact me at 404:363-7583.

Sincerely,

Hal

Hal Maggied, AIP
Senior Research Scientist

HM:ddh

Attachment