

# GISTRAN

## Version 1.0

**A Geographic Information System  
for Modelling Forest Products Transportation<sup>1</sup>**

**User's Guide**

**D. C. Kapple and H. M. Hoganson<sup>2</sup>**

**September 1991**

**STAFF PAPER SERIES NO. 83**

---

<sup>1</sup> Research supported by the University of Minnesota Center for Transportation Studies. Published as paper #19,760 of the Minnesota Agricultural Experiment Station.

<sup>2</sup> The authors are research assistant, Department of Forest Resources, College of Natural Resources, University of Minnesota, St. Paul, MN 55108 and assistant professor Department of Forest Resources, College of Natural Resources, University of Minnesota stationed at the North Central Experiment Station, Grand Rapids, Minnesota.

## TABLE OF CONTENTS

ABOUT THE MANUAL .....	1
SOFTWARE DISCLAIMER .....	2
I. INTRODUCTION .....	3
II. MODEL REQUIREMENTS .....	3
III. LINKAGES WITH OTHER MODELS .....	3
IV. GEOGRAPHIC DATA SOURCES AND PREPARATION .....	4
A. GISTRAN TRANSPORTATION NETWORK .....	4
B. GISTRAN BOUNDARIES AND AREAS .....	5
V. GISTRAN PROGRAMS .....	5
A. MILLPATH .....	5
1. Description .....	5
2. Usage notes. ....	6
B. FINDROAD .....	6
1. Description. ....	6
2. Usage Notes .....	7
C. LISTSTAN .....	7
1. Description .....	7
2. Usage Notes .....	8
D. TRANMAP .....	8
1. Description .....	8
2. Using TRANMAP .....	9
VI. REFERENCES .....	16
Appendix 1 - GISTRAN Network Databases and Files .....	17
Appendix 2 - Other GISTRAN Files and Data Structures .....	20

## ABOUT THE MANUAL

This manual is designed to facilitate the use of GISTRAN, a geographic information system for managing political and transportation spatial data in forest management planning models. The authors assume that users of this software have some knowledge of both geographic information systems and forest management planning. The authors also assume that users have experience using personal computers running under the DOS operating system.

GISTRAN was developed as part of a forest management scheduling system that includes two other subsystems, RxWRITE and DTRAN. Although the methodology and algorithms can be incorporated into other systems, GISTRAN in its present form is linked with these other models. The authors assume that GISTRAN users will also be using RxWRITE and DTRAN and have access to their documentation.

The manual is broken down into distinct sections that the user may consult as needed. Sections I through IV provide an introduction to GISTRAN system. Section V provides descriptions of the programs and their use.

The appendices explain GISTRAN file formats and data structures.

**SOFTWARE DISCLAIMER**

The software on the GISTRAN diskette has been extensively tested and checked for accuracy. To the best of our knowledge, it contains no errors. However, neither the University of Minnesota, the Department of Forest Resources, nor the authors provide any guarantees and are not responsible for errors that may arise during the use of this software. The authors request that any errors discovered by users be brought to their attention in order to incorporate appropriate changes into future versions.

For permission to use or copy GISTRAN and for revised versions or updates to existing installations, contact:

Douglas C. Kapple  
University of Minnesota  
Department of Forest Resources  
115 Green Hall  
1530 North Cleveland Avenue  
St. Paul MN 55108

## **I. INTRODUCTION**

The cost of transporting wood to market is an extremely important component of forest products production costs. For most timber harvested in Minnesota, the cost of transporting wood to the mill is greater than the cost of purchasing the standing timber. However, traditional forest planning models, used to help make strategic decisions on forest management and forestry based economic development, give inadequate consideration to transportation factors. The objective of this project was to develop a computer model featuring a method for solving forest management scheduling problems that considers transportation costs in significantly greater detail.

Obviously, such a model requires a large amount of spatial data, including information about the transportation network and the location of forest areas. GISTRAN is a system of computer programs designed to perform the critical functions of managing, processing, and presenting this geographic information.

In its present implementation, GISTRAN uses Indexed Sequential Access Method (ISAM) databases, random-access binary files, and ASCII text files. It is worth emphasizing, however, that the methods and algorithms developed do not depend upon the data format and the programs can be modified for computer systems where a particular format is unavailable.

## **II. MODEL REQUIREMENTS**

Several requirements influenced the selection of the spatial data and the design and development of GISTRAN.

- A. Information associating each analysis area with the nearest point in the road network was needed to compute wood transportation costs and to determine the procurement zone in which each analysis area is located.
- B. The shortest or lowest cost route from each node in the road network to each market location was also required to compute transportation costs.
- C. An ability to display maps showing state and county boundaries, roads, and procurement zones was desired.
- D. GISTRAN was required to be compatible with the other components of a forest management modelling system.

## **III. LINKAGES WITH OTHER MODELS**

GISTRAN is specifically designed to be linked with other models developed by the University of Minnesota Department of Forest Resources.

RxWRITE is a new program that serves as a "prescription writer" or "management alternative generator" (McDill 1991). This program uses the standard U.S. Forest Service statewide forest

survey ground plot information to develop possible management alternatives. RxWRITE uses the same growth projection routine as used by the STEMS growth projection system (USDA Forest Service 1979) to predict yields in future periods. RxWRITE is extremely flexible in its ability to define possible management alternatives, harvest systems, associated harvesting costs, product utilization standards, and species groups. It is a highly detailed and user-friendly system.

DTRAN is a computer model that helps analyze a forest's ability to supply multiple product flows over time to different market locations (Hoganson and Kapple 1991). DTRAN extends the basic solution strategy of the DUALPLAN forest management scheduling model (Hoganson and Rose 1984, Hoganson and Rose 1989) to recognize important transportation considerations. It is designed to examine differences in timber supply within a large region where specific product requirements at specific market locations are important. It can be used to help examine a variety of state-wide or region-wide concerns ranging from mill expansion opportunities at a particular location to broad forest-wide policies influencing forest management practices.

#### **IV. GEOGRAPHIC DATA SOURCES AND PREPARATION**

Digital Line Graphs (DLGs) produced by the U.S. Geological Survey from 1:2,000,000 maps furnished the road network and political boundary data (U.S. Geological Survey 1990). A DLG is a digital map data set in vector form. The data derived from the 1:2,000,000-scale maps are in several data sets in three categories: boundaries, hydrography, and transportation. GISTRAN data was taken from the political boundaries and the roads and trails data sets. The DLGs provided complete coverage of Minnesota in a format from which the data required for GISTRAN could be conveniently extracted or computed. The DLG road network information was last updated in 1980 (U.S. Geological Survey 1990).

Initially, relational database operations cannot be performed conveniently on DLG files because they are flat sequential files. However, DLG files are topologically structured, meaning that the relationships inherent in the map are preserved. Political boundary and road network databases were developed from the DLG files. The original files contained information for Michigan, Wisconsin, and Minnesota. To reduce computer storage and processing time requirements, data were extracted to form databases for Minnesota alone. Appendix 1 summarizes the structure and contents of these databases.

In the development and use of GISTRAN to date, forest survey plot locations have been used to represent analysis area locations. However, the ground coordinate information from the DLG files and the forest survey data were not in the same coordinate system. The forest survey data contained the Universal Transverse Mercator (UTM) coordinates of the forest survey plot locations, but points in the DLGs had Albers Equal-Area Conic projection coordinates. Also, forest survey plots were located in three UTM zones. The coordinate system for GISTRAN is UTM zone 15, where most of Minnesota is located. All of the coordinates of the Minnesota boundaries and roads and the coordinates of the forest survey plots in zones 14 and 16 were converted to UTM zone 15 coordinates.

##### **A. GISTRAN TRANSPORTATION NETWORK**

A network of arcs and nodes, taken from the DLG data, models the road system in Minnesota. At the small scale of 1:2,000,000, many secondary highways are not included, but the scale was adequate for development purposes and was deemed sufficient for studies in which broad regional analysis is the major interest. Each arc represents a section of road with one starting point, one ending point, and no branches. Every arc is identified by a unique integer and has a starting node, an ending node, a road classification, and an arc length. Each arc also has a list of coordinates of points along that road segment. Every node in the road network has an identification number, an Easting (x coordinate), a Northing (y coordinate), and a list of the identification numbers of the arcs incident upon it. It is important to distinguish between nodes and points. Nodes define the ends of the arcs and occur only at road intersections or ends. Points are coordinate pairs that describe the location of the road segment represented by an arc. The coordinates of the first point along an arc are the coordinates of its starting node and the coordinates of the last point are the coordinates of the ending node.

## **B. GISTRAN BOUNDARIES AND AREAS**

Nodes and arcs also describe the county and state boundaries. The nodes in the boundary network are similar to the nodes in the road network and the same information is associated with them. Each arc in the boundary network also has an identification number, a starting node, an ending node and a list of coordinates of points along that road segment. However, instead of a road classification and arc length, a right area and a left area are associated with each boundary arc. The right area identifies the area on the right side of the boundary line as one walks from the starting node to the ending node. Each area has an identification number and code numbers identifying the county and state it represents. Also associated with each area is a list of the arcs that form its boundary and a list of the road network arcs that lie within the area.

## **V. GISTRAN PROGRAMS**

GISTRAN programs perform three major tasks within the forest management scheduling model. MILLPATH finds the shortest route from each node in the road network to each market location and computes its distance. FINDROAD finds the nearest point on the nearest arc in the road network for analysis area locations. TRANMAP displays maps of the road network with wood procurement zones from DTRAN results.

### **A. MILLPATH**

#### **1. Description**

The MILLPATH program processes spatial information to generate several data files that are used in the modelling system. The data files generated can either cover the entire state of Minnesota or a user-specified subset of counties defined as a region. If a subset of the network is sufficient for a particular study, significant reductions in computer storage and processing time requirements are achieved by considering only that part of the network.

The identification numbers in the DLG databases are unique, but are not necessarily consecutive. For example, there are 914 arcs in the road network database with identification numbers between

1 and 3004. However, identification numbers are not used in DTRAN, where arcs and nodes are identified by the indices of the arrays in which they are stored. MILLPATH forms lists that map each node and arc identification number to its position in an ordered list. For this purpose, MILLPATH creates an arc list file, which contains four lists (columns) ordered by arc identification number: (1) arc identification number, (2) index number of the starting node, (3) index number of the ending node, and (4) arc length.

MILLPATH also generates a distance adjacency matrix and an arc identification adjacency matrix. The element in the *i*th row and *j*th column of a distance adjacency matrix is the distance along the arc between node *i* and node *j* in the network. The element in the *i*th row and *j*th column of an arc identification adjacency matrix is the identification number of the arc between node *i* and node *j*. Because they are sparse, these adjacency matrices are stored as three matrices in compressed matrix mode. Compressed matrix mode is described in Appendix 2.

MILLPATH also writes a file containing lists of the node coordinates, ordered according to node identification number. The program then writes a file with four items of information for each market: (1) The market identifier, (2) The identification number of the node where the market is located, (3) the node index number, and (4) the index of the market location row in the path matrices. Where the number of mills and the number of market locations are equal, items 1 and 4 will be identical.

For each node in the road network, MILLPATH determines the shortest route to each market location and computes its distance. The subroutine that performs these computations implements the Dijkstra algorithm (Horowitz and Sahni 1978). Two matrices are created. For every node to market combination, the shortest route path matrix identifies the nodes along the shortest route. The distance matrix indicates the distance along the shortest route. The transposes of these matrices are written in ASCII files.

## 2. Usage notes.

MILLPATH is designed as a one-time batch job to be run after the network databases are set up and a study region is defined. Subsequently, the program will be required only if region definitions change or a MILLPATH output file is destroyed. MILLPATH assumes that the boundary and roads databases exist as described in Appendix 1. To set up a MILLPATH job, create or edit the control file. The control file provides 12 items of information to the program. There is an example control file in Appendix 2.

To run the program, verify that MILLPATH.EXE is in the current directory and enter "millpath" from the DOS prompt. Enter the name of the control file when requested. No further user interaction is required. The program displays messages as processing progresses.

## B. FINDROAD

### 1. Description.

Given a pair of UTM coordinates, for example, the coordinates of a forest survey plot representing an analysis area, FINDROAD finds the nearest arc in the road network to the analysis area. The



program also finds the point on that arc nearest to the analysis area, in terms of its distance from the arc's starting node. This distance is scaled so that it is between zero and UNITSPERROAD, as defined for the associated DTRAN model. Let A be the length of the arc and D be the actual distance from the starting node to the nearest point to the analysis area. Then the distance is expressed as  $\text{UNITSPERROAD} * (D / A)$ . FINDROAD also computes and records the distance from the analysis area location to the nearest point. This value is not scaled.

Finding which road is closest to a particular point is essentially a searching process and a simple task for a person looking at a map. Conducting the search efficiently with a computer program is more complicated, and the following procedure was developed.

- (1) Find the 16 closest nodes to the analysis area location.
- (2) Make a list of the arcs incident on those nodes.
- (3) For each of these arcs, find the 8 digitized points nearest the analysis area.
- (4) Find the place on the arc that is nearest to the analysis area. This point may be between two of the digitized points that define the arc.
- (5) If the place on the current arc is closer to the analysis area than the shortest distance from any previously scanned arcs, then the current arc becomes the nearest arc to the analysis area.

FINDROAD presently creates a table in the stands databases of the RxWRITE.

## 2. Usage Notes

FINDROAD is fully integrated into the RxWRITE system and can also be run as a stand-alone program. For input, the program requires the stands databases from RxWRITE (McDill 1991), along with the GISTRAN roads database and associated files. When run as a part of RxWRITE, the RxWRITE county and codes databases are also needed. If a network table exists in the stands database(s) to be processed, it must be deleted or emptied before FINDROAD can be executed. FINDROAD also uses the index adjacency matrix file, the arc identification matrix file, and the node coordinates file generated by MILLPATH.

To run the FINDROAD, verify that FINDROAD.EXE is in the current directory and enter "findroad" from the DOS prompt. You can select a single stands database or multiple stands databases for processing via a sequence of user-friendly menus. The program displays messages as processing progresses.

## C. LISTSTAN

### 1. Description

LISTSTAN is a simple program that extracts analysis area location information from three RxWRITE stands databases. LISTSTAN creates a woods-to-network data file and an associated pointer file. The data file contains the analysis area identification, the UTM coordinates, and the information computed by FINDROAD. The pointer file indicates the location in the data file for each county.

## 2. Usage Notes

LISTSTAN requires control file that specifies the required parameters and files. Appendix 2 contains an example of a control file. To run the program, verify that MILLPATH.EXE is in the current directory and enter "liststan" from the DOS prompt. Enter the name of the control file when requested. No further user interaction is required. The program displays messages as processing progresses.

## D. TRANMAP

### 1. Description

TRANMAP is a set of subroutines that enables a DTRAN user to view maps showing political boundaries, the road network, and the wood procurement zones at selected stages of the DTRAN solution process. TRANMAP is integrated into DTRAN so that displaying procurement zone maps is an option whenever the iteration process is stopped. TRANMAP can also be run with a stand-alone driver program using a DTRAN model formulation file as input.

Mapping the procurement zones is a sequence of five processes.

- (1) Determining the map scaling constants. The area to be displayed by TRANMAP is automatically scaled to approximately fill a square display window. To display a map, the UTM coordinates of the map data must be converted to screen coordinates. This is accomplished using five scaling constants: The UTM coordinates are multiplied by a *scale factor* to convert the ground dimensions to window dimensions. An *x shift value* is added to the x coordinates and a *y shift value* is added to the y coordinates to place the map in the window. In the Northern Hemisphere, UTM Northings, or y coordinates, increase from south to north, but y coordinates on a computer monitor increase from top to bottom. For this reason, a *minimum y value* and a *maximum y value* are needed to "flip" the y coordinates. These scaling constants are determined by scanning the coordinates of the nodes on the boundary of the area to be displayed for the maximum and minimum coordinate values.
- (2) Finding the procurement zone boundaries. Each analysis area is associated with an arc in the road network. Thus, the first step in finding the procurement zone for the analysis areas is to find the procurement zone boundaries on the arcs. For each arc, the procurement zones for the starting node and ending node are computed using the shortest route distance matrix generated by MILLPATH with the transportation cost and product price information computed by DTRAN. If the starting and ending nodes are not in the same zone, then the location of the zone boundary on the arc is computed.
- (3) Finding the procurement zones for the analysis areas. Having computed the procurement zone boundaries and the nearest points in the road network to the analysis areas, each analysis area is assigned to a procurement zone. The analysis area coordinates are then scaled to the graphics window, and a circle is drawn in the color representing its procurement zone.

- (4) Drawing the boundaries. The arcs forming the county boundaries of the map area are scaled and drawn.
- (5) Drawing the road network. The arcs forming the road network in the map area are scaled and drawn.

## 2. Using TRANMAP.

TRANMAP is started by selecting the "A. DISPLAY MAP" option from the DTRAN menu displayed when the solution process is stopped or by running the ZONEMAP program. Several databases and files are required. Appendix 1 describes the format of these files.

- The roads database and associated files.
- The boundary database and associated files.
- The market database.
- A DTRAN model formulation file (Hoganson and Kapple 1991).
- A market location data file (created by MILLPATH).
- An arc list file (created by MILLPATH).
- A shortest route distance matrix file (created by MILLPATH).
- The woods-to-network files created by LISTSTAN.

Upon selecting the "A. DISPLAY MAP" option from the DTRAN menu (or entering the model formulation file name, if you are running ZONEMAP), the menu shown in Figure 1 is displayed.

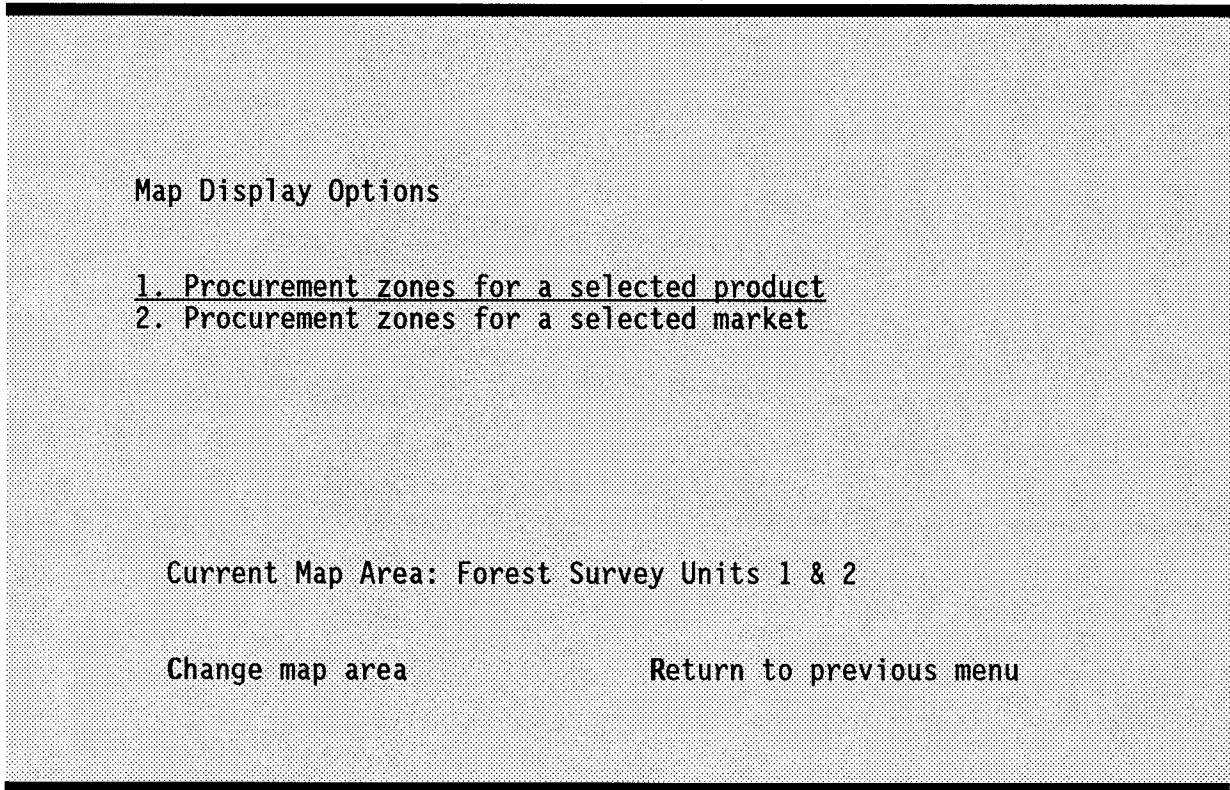


Figure 1 -- Initial Menu

On an actual screen, the menu is white on blue. In the example shown here, the current selection is underlined, but on the screen it is in reverse video (blue on white). To change the current selection, use the arrow keys or press the number of the selection. Press the enter key or the space bar to choose the current selection. Press the highlighted key ( C or R ) to change the map area or return to the previous menu.

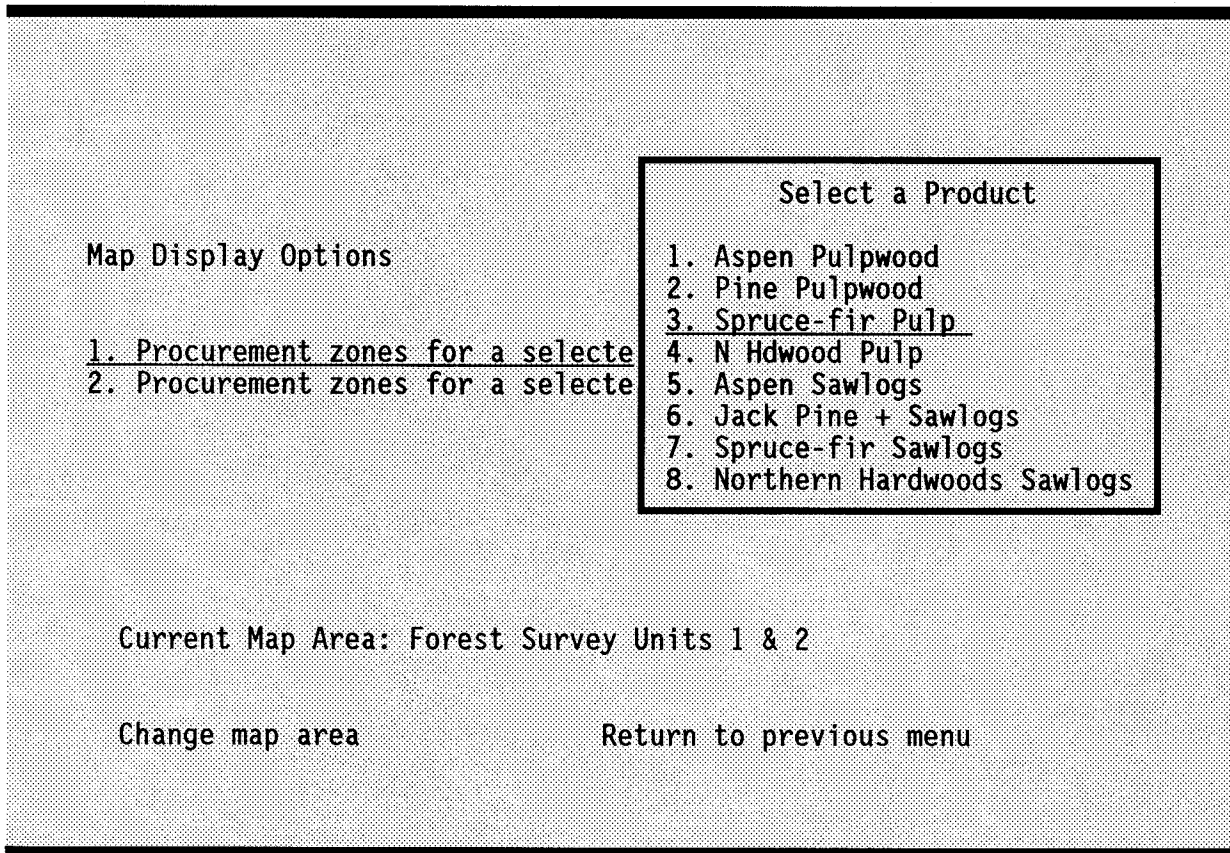


Figure 2 -- Product Selection Menu

When you choose the first display option, you need to select a product from the menu shown in Figure 2.

After you select a product, the procurement zones for the product are determined and a map like the one in Figure 3, but in full color, is displayed.

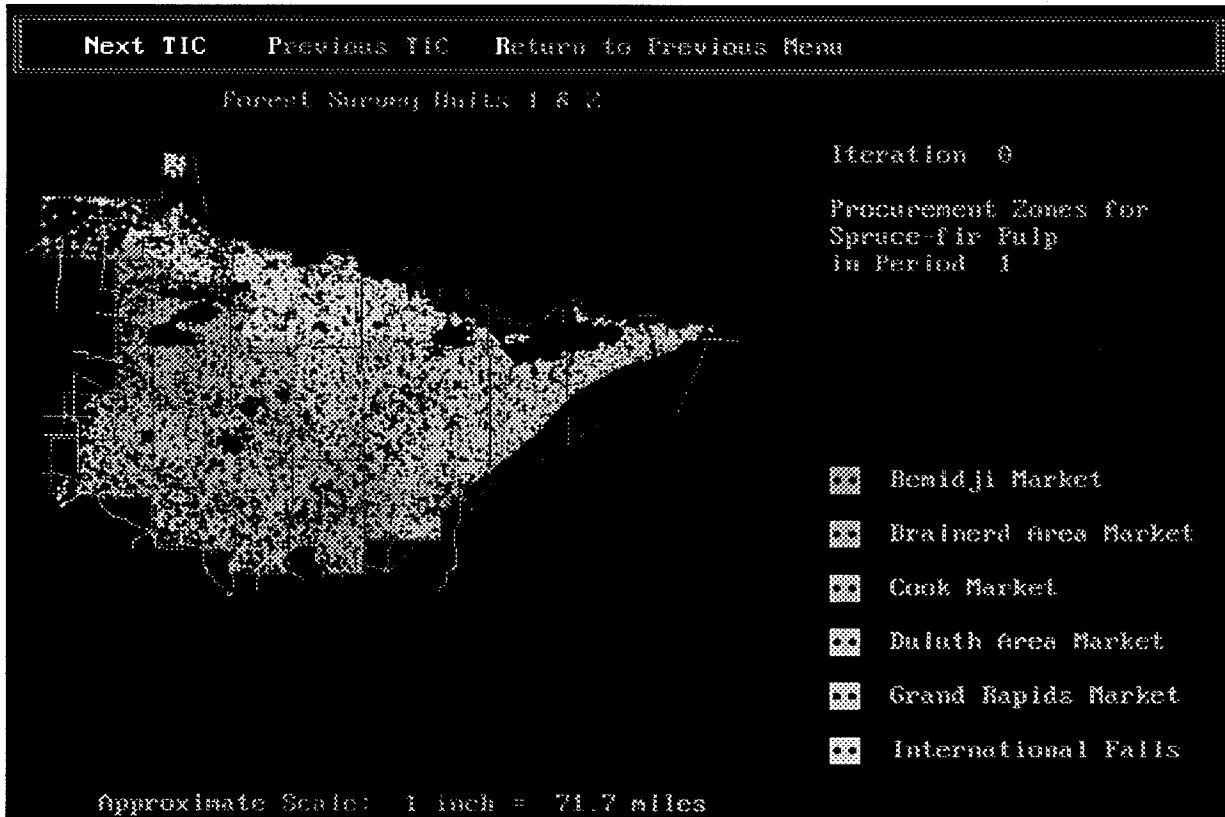


Figure 3

The current selection of the top line menu is highlighted. Use the arrow keys to change the current selection. Press the enter key or the space bar to choose the current selection or press a highlighted letter to make that selection immediately. When you select "Next TIC" or "Previous TIC", the program recomputes the procurement zones and recolors the analysis areas that change zones. Note that "Iteration 0" is shown near the upper right of the figure. For this example, TRANMAP was run with the ZONEMAP program, so the iterations were not counted. Actually, the map shows the results from a model formulation file generated after 70 DTRAN iterations.

If you select the second display option, to display procurement zones by market, you must choose a market from a market selection menu, and then choose a product. This option displays a map showing the procurement zone of one market for one product. You can display the procurement zone for a different product for this market without returning to the initial menu.

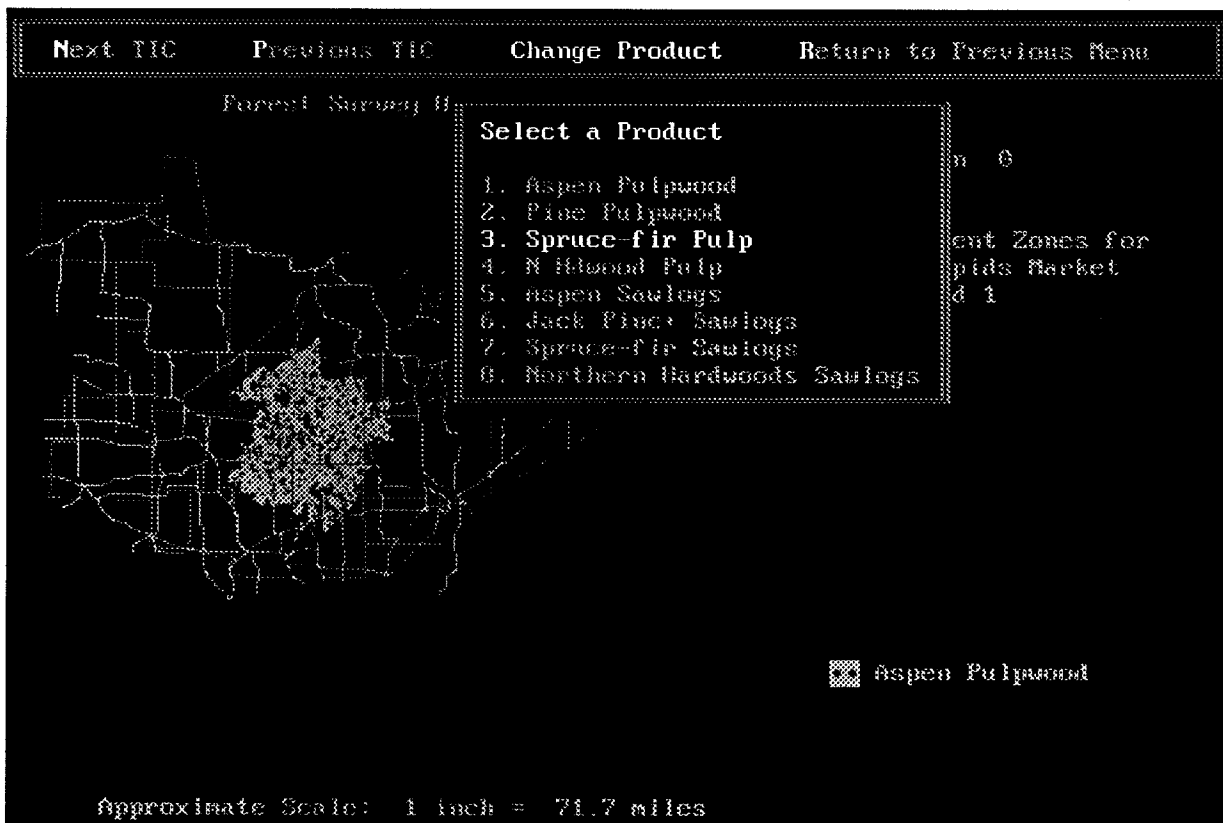


Figure 4

Figure 4 shows a map with the product selection menu.

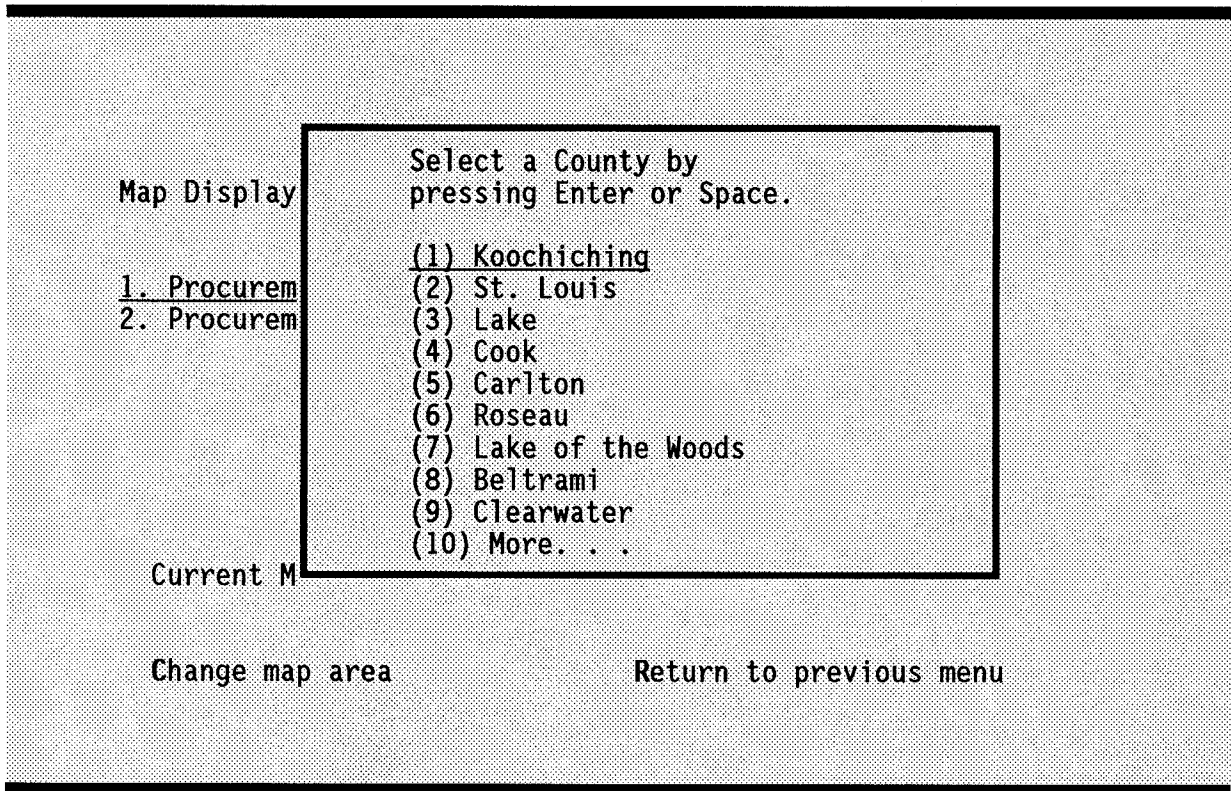


Figure 5 -- County Selection Menu

If you want to change the map area to be drawn, press "C" from the initial menu. Then choose whether you want to display a region or a single county. If you want a map of one county, then choose from the county selection menu, shown in Figure 5. The region selection menu is similar.



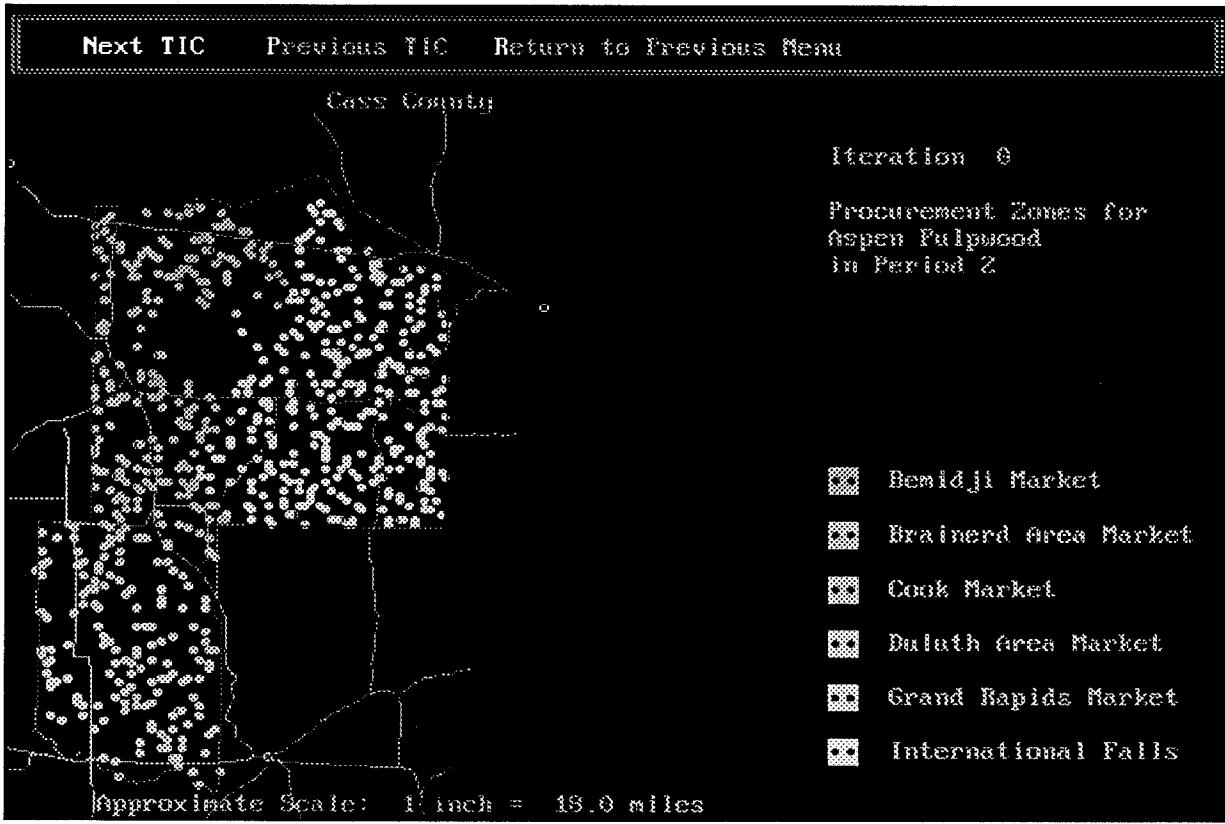


Figure 6

Figure 6 is a map of aspen pulpwood procurement zones in Cass County, Minnesota.

## VI. REFERENCES

- Hoganson, H. M. and D. W. Rose. 1984. A simulation approach for optimal timber management scheduling. *For.Sci.* 30(1):220-238.
- Hoganson, H. M. and D. W. Rose. 1989. DUALPLAN version 1.0 user's manual. University of Minnesota Department of Forest Resources Staff Series Paper No. 73. 48 pp.
- Hoganson, H. M. and D. C. Kapple. 1991. DTRAN version 1.0 user's guide. University of Minnesota Department of Forest Resources Staff Series Paper No. \*\*. (in progress).
- Horowitz, E. and S. Sahni. 1978. Fundamentals of computer algorithms. Computer Science Press, Inc. Rockville, MD. 626 pp.
- McDill, M. 1991. RxWRITE users guide. University of Minnesota Department of Forest Resources Staff Series Paper No. \*\*. (in progress).
- USDA Forest Service. 1979. A generalized forest growth projection system applied to the Lake States Region. USDA Gen. Tech. Rep. NC-49. 96 pp.
- U. S. Geological Survey. 1990. Digital Line Graphs from 1:2,000,000-scale maps. National Mapping Program Technical Instructions Data Users Guide 3. U. S. Geological Survey. Reston, VA. 70 pp.

## Appendix 1 - GISTRAN Network Databases and Files

GISTRAN programs are written in Microsoft BASIC 7.1 and use the PROISAM feature which is part of that program development system. The present implementation of GISTRAN includes two Indexed Sequential Access Method (ISAM) databases with associated files that are built from Digital Line Graph (DLG) data sets.

The Political Boundary database contains five tables. In addition, there are four associated files that are separate from the database file.

- The node table contains a record for each node. This table is extracted from the DLG political boundary file. Records in the node table are made up of the following fields:
  - IDNum is a unique integer identification number.
  - Easting is the X coordinate of the node in UTM zone 15.
  - Northing is the Y coordinate of the node in UTM zone 15.
  - nLines is the number arcs incident on the node or the number of adjacent nodes.
  - LineLoc is the position where the list of the identification numbers of these arcs begins in a random-access binary file, in number of bytes from the beginning of the file.
- The area data table contains a record for each area. The area table is also extracted from the DLG file. Area table records contain the following fields:
  - IDNum is a unique integer identification number.
  - nLines is the number arcs that form the area boundary.
  - LineLoc is the position where the list of the identification numbers of these arcs begins in a random-access binary file.
  - nArcs is the number of arcs in the roads database that are in the area. If any point along an arc is within the area, that arc is considered in the area.
  - ArcLoc is the position where the list of the identification numbers of these arcs begins in a random-access binary file.
  - State is the FIPS (Federal Information Processing Standards) code for the state.
  - County is the FIPS code for the county.
- The boundary line (arc) table is also extracted directly from the DLG file and has records containing 7 fields:
  - IDNum is a unique identification number.
  - StartNode is the identification number of the starting node.
  - EndNode is the identification number of the ending node.
  - LeftArea is the identification number of the area on the left side of the arc as one walks from the starting node to the ending node.
  - RightArea is the identification number of the area on the right side of the arc.
  - nPoints is the number of digitized points along the arc.

- PointLoc is the position where the list of the coordinates of these points begins in a random-access binary file.
- A county table contains records with four fields. County names are not part of the original DLG data. The process of ordering the counties and creating this table is not fully automated.
  - FIPS is the FIPS code for the county.
  - Name is the county name, a 20-character string.
  - DrawOrder is the position of this county in the sequence in which the counties are drawn in GISTRAN maps. This ordering is also important in defining the regions.
  - InRegion is the region, or set of counties, to which the county belongs.
- A Region table defines regions and sets of regions for mapping. In GISTRAN, regions are disjoint sets of counties. This information is not in the original DLG data. Regions are based upon Forest Survey Unit boundaries with implementation-specific modifications. Each record contains four fields.
  - IDNum is a unique identifier.
  - Name is the name of the region or set of regions.
  - FirstCounty is the position of the first county where the counties are listed in ascending DrawOrder.
  - nCounties is the number of counties in the region or set of regions.
- The line list file is a random-access binary file which contains lists of boundary arc identification numbers. For each node, there is a list of arcs incident upon it. For each area, there is a list of arcs forming its boundary.
- The county arc list file is a random-access binary file containing lists of road arc identification numbers. For each area (county), there is a list of road arcs in that area.
- An arc point file contains lists of coordinates ( x followed by y ) for each arc in the boundary arc table. This is also a random-access binary file.
- The county list file is an ASCII text file listing the FIPS codes of the counties in ascending DrawOrder, followed by (FirstCounty, nCounties) pairs as defined for the region table above.

The roads database contains three tables and there are two files separate from the database.

- The road arc table is also extracted directly from the DLG roads and trails file and has records containing 7 fields:
  - IDNum is a unique identification number.
  - StartNode is the identification number of the starting node.
  - EndNode is the identification number of the ending node.
  - nPoints is the number of digitized points along the arc.

- PointLoc is the position where the list of the coordinates of these points begins in a random-access binary file.
  - RoadClass is the road attribute from the DLG data.
  - Distance is the length of the arc in meters.
- The node table contains a record for each node. This table is also directly from the DLG roads and trails file. Records in the node table are made up of the following fields:
    - IDNum is a unique integer identification number.
    - Easting is the X coordinate of the node in UTM zone 15.
    - Northing is the Y coordinate of the node in UTM zone 15.
    - nLines is the number arcs incident on the node or the number of adjacent nodes.
    - LineLoc is the position where the list of the identification numbers of these arcs begins in a random-access binary file, in number of bytes from the beginning of the file.
- The city table associates city names with selected nodes in the network. City name information is not in the original DLG files. The two fields in this table are:
    - NodeID is the identification number of the node where the city is located.
    - CityName is the name of the city.
- The road arc list file is a random-access binary file which contains lists of road arc identification numbers. For each node, there is a list of arcs incident upon it.
- An arc point file contains lists of coordinates ( x followed by y ) for each arc in the road arc table. This is also a random-access binary file.

## Appendix 2 - Other GISTRAN Files and Data Structures

Below is an example of a MILLPATH control file. Each indented line specifies a variable required by the program. The preceding line describes the item.

Directory path to network files and databases  
c:\dlg\network\  
Market database (input)  
c:\dlg\network\market6.MDB  
File name for county list for region definitions (input)  
c:\dlg\network\ctylist.DAT  
Region index in region definition file specified above  
7  
File name for the index adjacency matrix (output)  
c:\dlg\network\roadnode.906  
File name for the distance adjacency matrix (output)  
c:\dlg\network\roaddist.906  
File name for the Arc ID adjacency matrix (output)  
c:\dlg\network\arcid.906  
File name for the coordinates of road network nodes (output)  
c:\dlg\network\nodexy.906  
File name for the Arc IDs, starting nodes, ending nodes, distances  
c:\dlg\network\arclist.906  
File name for the Market location data  
c:\dlg\network\market6.906  
File name for the Shortest route distance matrix  
c:\dlg\network\pathdist.906  
File name for the Shortest route path matrix  
c:\dlg\network\pathfind.906

- The market database contains several tables but only one of these is required by GISTRAN programs.
  - The mill data table contains a record for each market. Each record has four fields.
    - IDNum is a unique integer identification number.
    - Name is a string variable containing the market name.
    - AtNode is the identification number of the node in the road network database where the market is located.
    - Kind indicates the kind of market. (This field is not accessed by GISTRAN.)
- The county list for region definitions is the one associated with the political boundary database described above.

- The adjacency matrices are used only within GISTRAN in the MILLPATH and FINDROAD programs. Suppose there are  $m$  nodes in the network. Then the full adjacency matrix would have  $m$  rows and  $m$  columns. The element in row  $i$  and column  $j$  of an adjacency matrix is nonzero if and only if there is an arc linking node  $i$  and node  $j$ . Obviously, an adjacency matrix representing a road network will be sparse. Now suppose that  $n$  is the maximum degree of any node in the network, *i.e.* the maximum number of arcs incident on any node. Then all the relevant information can be stored in two  $m$  by  $n$  matrices. For GISTRAN, the first matrix is augmented with two columns and a third matrix is also formed. Suppose a network has 5 nodes and 5 arcs and that the highest degree of any node is 3. Then  $m = 5$  and  $n = 3$ . Let A, B, and C be the index, distance, and arc identification adjacency matrices for this network.

	A					B			C		
1	2	4	5	0	34.0	22.1	0.0	1	0	0	
2	1	3	0	0	18.8	0.0	0.0	2	0	0	
4	2	2	4	0	18.8	25.4	0.0	0	3	0	
5	3	1	3	5	34.0	25.4	28.1	0	0	4	
6	2	1	4	0	22.1	28.1	0.0	0	5	0	

A is a 5 by 5 ( $m$  by  $n+2$ ) matrix. For some row  $i$ ,  $A(i,1)$  is the identification number of the  $i$ th node. Note that the identification numbers are not consecutive.  $A(i,2)$  is the degree of the  $i$ th node. For example, the I.D. number of the third node in the list is 4 and its degree is 2. The remaining nonzero elements in the row are the row numbers of the adjacent nodes. For example, the third node is adjacent to the second and fourth nodes.  $B(i,j)$  is the length of the arc between node  $i$  and node  $A(i,j+2)$ . Thus the length of the arc linking nodes 1 and 4 is 34.0. The distance from node  $i$  to node  $j$  is the same as the distance from node  $j$  to node  $i$ , so B represents a symmetric matrix. If  $C(i,j)$  is nonzero, then  $C(i,j)$  is the database I.D. number of the arc starting at node  $i$  and ending at node  $A(i,j+2)$ . For example,  $C(4,3) = 4$ , so arc 4 begins at node 4 and ends node  $A(4,5)$  or node 5.

The adjacency matrices are filed as random-access binary files. The values of  $m$  and  $n$  are written in the file for matrix A before the matrix data.

- The following is a portion of an arc list file.

```

12,8,10,3305.02713222245
13,8,15,4078.82036073621
14,13,118,19030.2271937104
17,14,102,20980.4089799084
18,102,103,34177.5926690295
19,103,104,31910.6981614666

```

The arc list file has one line for each arc in the road network. The first value on the line is the arc identification number in the road network database. Note that the numbers are not consecutive. Next are the node index numbers (their positions in the adjacency and path

matrices) of the starting and ending nodes, respectively. The fourth value is the length of the arc in meters.

- An example of a market data file appears below.

```
6,6
1,68,52,1
2,234,168,2
3,113,97,3
4,180,127,4
5,175,122,5
6,164,111,6
```

The first line indicates that there are six markets and six market locations. The number of markets is always greater than or equal to the number of market locations. The four values on each of the remaining lines specify the market I.D. number, node identification number in the road network database, the node index number in the adjacency and shortest path matrices, and the market location number for each market.

- The files containing the shortest route matrices contain the transposes of the matrices as stored in the programs. If there are  $n$  nodes in the road network and  $m$  market locations, then the shortest route matrices are  $m$  by  $n$  as stored in the computer programs and  $n$  by  $m$  as stored in the files. Consider the  $n$  by  $m$  form, where each row represents a node in the network and each column represents a market location. Suppose  $D$  is the distance matrix and  $P$  is the path matrix.  $D(i,j)$  is the distance, in meters, of the shortest route between node  $i$  and market location  $j$ .  $P(i,j)$  is the index of the next node on the shortest path from node  $i$  to market location  $j$ . Suppose  $P(i,j) = k$ . Then  $P(k,j)$  is the index of the next node on the shortest path from node  $k$  to market location  $j$  and also the index of the next node on the shortest path from node  $i$  to market location  $j$ . The path from node  $i$  to market location  $j$  can be traced through the nodes by continuing this procedure.

The following is a sample LISTSTAN control file:

```
1000, 0.1,          "values of DTRAN variables UNITSPERROAD, ACCESSUNITS!"
"c:\dtran\output\testfile.dat",      "Woods-to-network data file name"
"c:\dtran\output\testfile.loc",      "Woods-to-network pointer file name"
"c:\dlg\stands\",      "directory containing the RxWRITE stands databases"
3, "number of distinct counties or subregions, database names listed below"
"Carl05.MDB"
"Cook05.MDB"
"U4NO05.MDB"
```

- Except for the database names, the rightmost string, enclosed in quotation marks, describes the preceding values and is not used by the program.



- The values of UNITSPERROAD and ACCESSUNITS! must be exactly the same as those values in the DTRAN model formulation file(s) with which the output files are being used.
- The county or subregion databases must be listed in the drawing order specified in the political boundary database and associated files.

The woods-to-network data file containing analysis area location information extracted from three tables in the RxWRITE stands databases. For each analysis area, six items of information are stored in a random-access binary file:

- A stand identification number.
- The UTM Easting ( x coordinate ).
- The UTM Northing ( y coordinate ).
- The identification number of the nearest arc in the road network.
- The distance from the analysis area location to the nearest point on that arc.
- The distance from the starting node of the arc to the point on the arc nearest the analysis area location.

This information is stored by county according to the drawing order, as defined above. The pointer file contains a set of pointers into the random-access data file. A portion of such a file is shown below:

1	1240
17361	2232
48609	684
58185	435
64275	291
68349	185
70939	351
75853	709

The number in the jth row and first column indicates the byte position of the beginning of the first data item for jth county in the random-access binary file. the number in the jth row and second column is the number of analysis areas in that county.