Draft Final Report

Research Project GC8286, Task 42
State Level Transportation Applications

STATE LEVEL TRANSPORTATION APPLICATIONS USING TIGER

by

Edward McCormack
Research Engineer
Washington State
Transportation Center

Lawrence W. Blain
Transportation Engineer
Puget Sound Council
of Governments

Washington State Transportation Center (TRAC)
University of Washington
The Corbet Building, Suite 204
4507 University Way N.E.
Seattle, Washington 98105

Washington State Department of Transportation
Technical Monitor
Ron Cihon
Mapping Supervisor

Prepared for

Washington State Transportation Commission
Department of Transportation
and in cooperation with
U.S. Department of Transportation
Federal Highway Administration

March 21, 1990

TECHNICAL REPORT STANDARD TITLE PAGE

1. REPORT NO.	2. GOVERNMENT ACCESSION NO.	3. RECIPIENT'S CATALOG NO.
WA-RD 209.1		

4. TITLE AND SUBTITLE	5. REPORT DATE
STATE LEVEL TRANSPORTATION APPLICATIONS USING TIGER	March 1990
	6. PERFORMING ORGANIZATION CODE
7. AUTHOR(S) Edward W. McCormack Lawrence W. Blaine	8. PERFORMING ORGANIZATION REPORT NO.
9. PERFORMING ORGANIZATION NAME AND ADDRESS Washington State Transportation Center (TRAC) University of Washington, JE-10 The Corbet Building, Suite 204; 4507 University Way N.E. Seattle, Washington 98105	10. WORK UNIT NO.
	11. CONTRACT OR GRANT NO. GC8286, Task 42
12. SPONSORING AGENCY NAME AND ADDRESS Washington State Department of Transportation Transportation Building, KF-10 Olympia, Washington 98504	13. TYPE OF REPORT AND PERIOD COVERED Draft Final Report
	14. SPONSORING AGENCY CODE

15. SUPPLEMENTARY NOTES

This study was conducted in cooperation with the U.S. Department of Transportation, Federal Highway Administration.

16. ABSTRACT

The TIGER/Line Files are geographic databases developed by the Census Bureau as part of the 1990 Census to produce local maps for census takers. The TIGER (Topographically Integrated Geographic Encoding and Referencing) files contain the geographic information for every road, railroad, stream, and political boundary in the United States. This research explored state level transportation planning applications of the TIGER/Line files by building a link-node highway microcomputer modeling network from the TIGER data.

The use of TIGER resulted in a network with links and nodes that were accurately located by latitude and longitude and included link lengths. A TIGER-derived network potentially could serve as a powerful basis for incorporating census and, to a lesser extent, non-census based data into a transportation model. The disadvantages of using TIGER to create a model network were found to be roadway detail unnecessary on a state level, a time consuming downloading process, lack of appropriate TIGER manipulation software, and the need for network sorting, editing, and preparation. In spite of these drawbacks, TIGER offers considerable promise as a transportation planning tool, especially for smaller state and state level studies needing a detailed network.

17. KEY WORDS	18. DISTRIBUTION STATEMENT
TIGER/Line files, networks, state level transportation model	No restrictions. This document is available to the public through the National Technical Information Service, Springfield, VA 22616

19. SECURITY CLASSIF. (of this report)	20. SECURITY CLASSIF. (of this page)	21. NO. OF PAGES	22. PRICE
None	None	45	

DOT FORM 310-022 (11/86) -1216-

DISCLAIMER

TABLE OF CONTENTS

LIST OF FIGURES

EXECUTIVE SUMMARY

The TIGER/Line Files are geographic databases developed by the Census Bureau as part of the 1990 census to produce local maps for census takers. The TIGER (Topographically Integrated Geographic Encoding and Referencing) files contains the geographic information for every road, railroad, stream, and political boundary in the United States. This research explored state level transportation planning applications of the TIGER/Line files by attempting to build a link-node highway microcomputer modeling network from the TIGER data.

The initial step of the project involved exploration the relevant literature and examination of software capable of manipulating the TIGER data. Since the TIGER files require large amounts of computer file space, the next step of the project was to develop procedures to extract only the portion of the TIGER data concerned with roadways. During the extraction process, the data from the source computer tape were also converted to a format usable on a microcomputer. Once on a microcomputer, the research team attempted to extract, from all roads in the TIGER data, federal and state highways for the model network according to the roadway classification scheme built into the TIGER data. However, the scheme in use complicated the extraction of the state highway data necessary for this research, and the researchers had to fill in numerous missing links and make many corrections. As a result of this situation and because of software limitations, a simplified, partial state level network was developed. After the network had been finished, the research team investigated the utility of linking a TIGER derived model network with census and non census databases.

In summary, the research results demonstrated that a simple microcomputer state level link-node network can be created with the TIGER/Line files. The use of TIGER resulted in a network with links and nodes that were accurately located by latitude and longitude and included link lengths. In addition, a TIGER derived

network potentially could serve as a powerful basis for incorporating census and, to a lesser extent, non-census based data into a transportation model.

The disadvantages of using TIGER to create a model network include roadway detail unnecessary on a state level, a time consuming downloading process, lack of appropriate TIGER manipulation software, and the need for network sorting, editing and preparation. In spite of these drawbacks, TIGER offers considerable promise as a transportation planning tool, especially for smaller states and state level studies needing a detailed network.

CONCLUSIONS AND RECOMMENDATIONS

CONCLUSIONS

This project attempted to construct a link-node network of Washington's federal and state highways from the Bureau of the Census's TIGER database. The network was designed for use with a state level microcomputer transportation modeling effort. During the project, a number of conclusions related to the use of TIGER data on a state level were developed. They are listed below:

- A survey of software found only one package that appeared to be suitable for use with TIGER/Line files and with a state level transportation modeling effort. This recently developed package, TransCAD, appeared to have the capability to manipulate TIGER data and had some standard transportation network operation features. Additional software will probably be developed as the 1990 census data becomes generally available.
- The TIGER data files supplied on tape by the Census Bureau were sizable and tended to overwhelm the data storage capacity of a microcomputer (all the TIGER/Line files for the state of Washington required about 350 MB). However, the research team determined that many of the data in the TIGER/Line files were unnecessary for a state level transportation network, and many of the records represented features other than roads. By extracting only road segments and by compressing those records to include only essential data, the overall size was reduced to a level manageable on a microcomputer.
- The actual process of downloading the TIGER data into a microcomputer format was time consuming and fairly complex. A

number of in-house mainframe and microcomputer programs were required to extract the necessary data and to avoid overloading the microcomputer data storage capacity. The Bureau plans to make each state's files available on CD ROM when its workload permits. One CD ROM would have held all of Washington's files and would have been readily usable with a microcomputer.

- The research team had difficulty in extracting just the state highway segments from the TIGER/Line files. The main difficulty concerned the accuracy of the Feature Name and CFCC (Census Feature Classification Code) fields in the TIGER/Line files. Because of inconsistent highway labeling and classification, a number of highways in the state network that was extracted had missing or misplaced segments. Network construction included adding missing links (missed by the extraction process -- not necessarily missing from the TIGER/Line files) and removing unwanted links.
- The shape points between the end points of the TIGER roadway were not useful at a state level. They contained unnecessary detail and greatly increased the data storage requirements.
- The linking of a TIGER derived transportation network with census and non-census data structures could potentially enhance a state level transportation modeling process. However, the integration of any data set with a TIGER derived network would involve considerable work. Since the TIGER data were designed as a census product, the integration of census data should be the most easily completed.

RECOMMENDATIONS

TIGER has potential as an enhancement to a state level transportation modeling process. If the transportation modeler can acquire or develop appropriate

software, the TIGER/line files may be used to build an extremely accurate node-link model network with latitude and longitude coordinates and link distances. Since TIGER includes all roads, a network can be as comprehensive as needed. However, because of the large number of data in TIGER and problems with roadway classification, the state level modeler attempting to use the TIGER/Line files for network construction should be aware that considerable data sorting, editing, and preparation are required.

Once a network has been completed, a TIGER based data set offers potential as a base for integrating a model network with census and, to a lesser extent, non-census data sets. Ideally, the integration of socio-economic data with a TIGER derived model network could result in more accurate zoning structures and better data analysis.

BACKGROUND

In preparation for the 1990 decennial census, the U.S. Bureau of the Census has digitized every street and census block face in the United States. Since these data will be used to produce local maps for census takers, many other physical features such as railroad, lakes, streams and power lines are also encoded for reference. These data sets are contained in the Topographically Integrated Geographic Encoding and Referencing (TIGER) data structure. The entire TIGER data structure used internally by the Bureau of the Census consists of cross-referenced points, lines, and polygons with explicit topology. The TIGER/Line files distributed by the Bureau contain just TIGER line. The endpoints, adjoining polygons, and topology are defined implicitly.

These geometric and topological descriptions of features and the ability to explicitly relate these features to collection units of census and attribute data make the TIGER/Line files appealing for a variety of transportation planning and analysis functions. This research investigated the potential for using the TIGER/Line files in conjunction with state level transportation modeling efforts.

The research team attempted to integrate the TIGER data into a state level modeling process by using the TIGER/Line file to construct a link-node network of Washington state compatible with a microcomputer transportation planning model. The project network was to include all state highways under the jurisdiction of the Washington State Department of Transportation. The TIGER segments to be extracted from the 39 counties in Washington included all highways with the following classifications:

- Federal Interstates (I-5, I-405),
- U.S. Highways (U.S 2, U.S 101), and
- State Routes (SR 6, SR 506).

These three classes of roads had to be separated from the numerous county, municipal, Forest Service, and Park Service road segments found in the TIGER/ Line files. Once extracted, these segment were be placed in a computer file appropriate for a microcomputer.

REVIEW OF THE LITERATURE AND SOFTWARE

LITERATURE

Since this research involved prototype TIGER/line files, little literature concerning previous use of TIGER for state level transportation planning was found. Simkowitz (1988a) provided a good introduction to the use of TIGER/Line files for transportation planning. On a more fundamental level, several articles discussed the use of geographic systems and transportation planning; both Nyerges and Dueker (1988), and Simkowitz (1988b) provided insight into the issues of using geographic information with transportation modeling and planning.

Several other documents explained the purpose and structure of the TIGER/Line files. Notably among these was an article by the Census Bureau's Data User Service Division (1988) and information on the TIGER demonstration project in Boone County, Missouri (Fleck, 1989).

SOFTWARE

The research team searched for microcomputer software that was capable of manipulating and examining the Census's TIGER data files. In particular, they wanted software that was both transportation-oriented and readily compatible with the TIGER/Line files. The research team contacted a number of vendors who developed and distributed Geographic Informations Systems (GIS) and mapping software.

Two software products were found to be useful for examining entire TIGER/Line files, Geographic Data Technology's Safari and the Bureau's own Tiny Tiger. But neither product permitted analysis or manipulation of the TIGER data. The software that appeared to offer the most flexibility in processing TIGER data and building transportation models was TransCAD (the Caliper Corporation). TransCAD was billed as "a comprehensive Geographic Information System software

package for planning, management, operation, and analysis of transportation systems and facilities." This software was acquired.

PROCEDURES

The research procedure involved the following steps:

- extracting the TIGER data from a computer tape,
- transferring the extracted data to a microcomputer,
- constructing a model network, and
- exploring the integration of the TIGER derived network with other data sets.

The following sections discuss these steps.

EXTRACTING THE DATA

The TIGER/Line files were supplied by the Census Bureau encoded in straight ASCII, without record separators, on nine-track tape at 1,600 or 6,250 bpi. At the higher density, the files for Washington state required three tapes, although they were sent on four. The files had to be downloaded from tapes to a mainframe and then to a microcomputer, while extraction and compression occurred along the way.

The TIGER/Line files were loaded to a mainframe computer one county at a time. A batch file displayed the initial parts of selected files for confirmation. Another batch file processed the county's files with a FORTRAN program, which extracted and compressed the state highway and county boundary segments. The segments were appended to two growing state files, which were to be transferred to a microcomputer.

The critical element was the choice of extraction criteria. For the county boundary segments the process was obvious: choose the segments whose left-side county did not match the right-side county. In fact, since each TIGER/Line file

referenced nothing outside a single county, either the left side or the right side county field of county boundary segments was empty.

For the TIGER highway segments data, there were two fields in each record on which to base a choice: the Census Feature Classification Code (CFCC) and the Feature Identification, which consisted of direction prefix, feature name, feature type (ramp, road, lane, etc), and direction suffix.

For all road features the first character in the three-character CFCC was the letter "A". The second character was a digit. According to the technical documentation, the code for this digit was 0 for unclassified, 1 through 5 for U.S. Geological Survey (USGS) road classes, and 6 for special features such as traffic circles and interchanges. The third character, also a digit, indicated whether the road was divided, was in a tunnel, was underpassing, or had a rail line in the center. For example, the CFCC for a USGS class 1 road, divided, was "A15".

The USGS showed five classes of roads on its 7.5- and 15-minute maps. They were defined as follows:

Class 1	All officially designated interstate and federal routes (including alternates), primary state routes, and other routes of equivalent importance.
Class 2	Secondary state routes, primary county routes, and other comparable through roads important to statewide or county travel.
Class 3	Paved roads not included in a higher class and loose-surfaced improved roads passable in all kinds of weather and used mostly for local traffic.
Class 4	Unsurfaced and unimproved roads passable in fair or dry weather.
Class 5	Roads passable only with a 4-wheel-drive vehicle.

The technical documentation confirmed that Class 5 roads were "jeep trails." If the TIGER/Line files had consistently followed this classification scheme, then extracting interstate, federal, and state numbered highways would have been a

simple matter of choosing the class 1 and 2 roads. However, preliminary investigation showed that this scheme was reasonably reliable only in rural areas, where coding originated with the USGS 1:100,000-scale maps, which in turn were derived from the 7.5-minute maps. In urban areas, where the coding was carried over from the Bureau's older DIME files (which had no classification), the majority of roads were labeled Class 4 -- up to and including multilane divided state highways.

After consulting with Census Bureau staff, the research team found that in urban areas the TIGER/Line file classifications had been updated only for major highways, that records that had originated in DIME files did not contain alternative names, and that the Bureau had used a different classification scheme in which Class 4 was any local road and thus was appropriate for most city streets.

In an effort to capture segments of numbered highways that had "incorrect" CFCCs, the feature names were examined to find the string "USHWY", "STHWY", or "I" followed closely by a "5" or a "90" (I-90 and I-5 are the two principal interstates in Washington). Alternative names were also sought with the same strings and traced back to the originating records. To avoid capturing all the segments at every interchange, segments with the feature type "RAMP" were not extracted, even if they had met one or more of the above criteria.

In its final form the extraction program developed by the research team searched TIGER for desired names, built a list of the records with desired alternative names, and extracted these and all other desired road and county boundary segments. The program then embedded the coordinates of corresponding intermediate points between the start and end points. Next the program wrote compressed records, including a field containing the total number of segment shape points, to either or both of the county road and county boundary files. At the conclusion of processing for each county, another program appended the two county files to the state road and boundary files.

TRANSFERRING THE DATA TO A MICROCOMPUTER

After they had been extracted, the compressed records had to be transferred to the microcomputer for additional preprocessing before they would be ready for building TransCAD databases.

The number of road records extracted was 35,513, with a total file size of 3.5 MB. The number of county boundary records extracted was 13,576, with a total file size of 1.6 MB. The PC on which TransCAD was running was connected to the mainframe by a 2,400 baud modem. The estimated transfer time was 7 hours. The chosen alternative was to transfer the files from the VAX to another PC over an Ethernet connection (less than 15 minutes), back up the files within DOS, and restore them to the TransCAD machine.

TransCAD stores the geometric and attribute data for each type of entity in a database, which may be composed of one or more layers. A database for polygons (2-cell) has two layers, the boundary segments and the enclosed regions. A segment (1-cell) database has a node layer and the segment layer. A point (0-cell) database has only one layer. TransCAD comes with a versatile utility (TCBuild) that builds the appropriate databases from external data files. During the building process TCBuild reserves user-specified fields in each record for attributes that may be added when the database is built or created later.

In order for TCBuild to accept coordinate data, the longitude and latitude must be expressed in fields of the same width. TIGER uses the format F9.6 for the longitude and F8.6 for the latitude. The research team had to develop a program to add a space before each latitude to obtain a usable format.

TCBuild had difficulty with segments that had a large number of shape points. Since that level of detail was not needed for the state highway model, a microcomputer utility was written to eliminate the intermediate points, leaving each segment defined by only start and end points. The county boundary file was reduced to 0.7 MB in size and the state highway file to 2.6 MB.

The final TransCAD database sizes were approximately 1.7 MB for the counties and 18 MB for the roads.

CONSTRUCTING THE NETWORK

The network construction step originally was to involve the identification and attempted extraction of the Interstates, U.S. highways, and state route highways from TIGER data that had been downloaded to a microcomputer. Because of gaps in the coverage and in order to limit the size of the project network, the research team decided to construct a selected network from just Interstate and U.S. highways.

The primary network construction activity was to select from the 35,513 segments in the extracted TIGER database a continuous network of segments to represent the selected network. The first step in the selection process was to use TransCAD's "select by condition." The condition was defined as "any segment whose name starts with "US" or "I" but whose CFCC is not "A63" (interchange). This condition failed to select segments of federal highways that were named "MAIN" or other names, and it selected in error segments that were misnamed or that were Class 2 roads with names such as "ICICLE RIVER ROAD." The entire network was examined and corrected in TransCAD, which allowed the researchers to add or remove segments.

Thirty seven gaps appeared in the network where the database lacked a segment completely. The missing segments were in the original TIGER/Line files but had not been extracted. These segments were added to the database and selected. Altogether, 9,784 road segments were selected, including the 37 added to the database.

TransCAD used one data structure to represent the geometry of a database and another to represent the topology of the network. The topological data structure was used to construct a model network. However, TransCAD's network building procedure was unable to process a large topologically correct network that

included all state level highways because of the large number of nodes (between 9,000 and 10,000). Therefore, a simplified network had to abstracted from the selected network.

Since the forecasting procedures of a state level transportation model could logically be based on county level zones, the simplified network was abstracted to show straight links between nodes established at intersections and county border crossings. TransCAD was unable to select the appropriate nodes automatically, so the research team selected the nodes by zooming in and pointing on the screen. Once the nodes had been selected, their identifiers and coordinates were sent to an external file where they could be used to build the simplified database.

The links were penciled on to a series of maps on which the nodes had been printed and labeled with their identifiers. Each link was then hand coded with start-node, end-node, and name. These files were used to build a new TransCAD database. The result was a simplified state network consisting of 94 links connecting the 86 selected nodes. The network included accurate node latitude/longitude and the possibility of calculating link lengths.

INTEGRATING THE DATA

After the network was constructed, the project team assessed the potential for linking census and non-census data structures with a TIGER derived transportation modeling network. Since the 1990 census had yet to be conducted and because an investigation of non-census data structures was beyond the scope of this project, this step did not involve an actual application.

DISCUSSION AND RESULTS

The result of the above process was proof that it is possible build a simple, state level link/node network from TIGER/Line files. With only one exception, the county (polygon) and road (segment) databases were built with no errors. TCBuild was able to close all the county polygons except Lewis County. By building a segment database of just the Lewis County boundary segments, the research team could find two places where segments were missing in the boundary. After these two segments had been added to the data, the county database was built with no further errors.

Figure 1 shows the county boundaries and all the road segments extracted from the TIGER/Line files. When this figure compared with the WSDOT digitized map of all numbered highways (Figure 2), the detail available in TIGER/Line files, even without shape points, is evident. There are apparent mismatches in adjacent county boundaries (e.g., between Ferry and Lincoln), but these mismatches are simply the artifacts of omitted shape points. When shape points are used, the boundaries match exactly.

By extracting all class 1 and 2 roads, the research team captured additional county roads in central Washington. However, in the urban areas the coverage of state highways was quite poor. This problem was especially evident in King County. Figure 3 shows the numbered highways in the urban part of King County, while Figure 4 shows the segments extracted from the TIGER/Line file. Notable omissions included Highway 99 (six-lane, limited access) and the Evergreen Point Bridge (a four-lane freeway). Examination of the original TIGER/Line file showed that both of these highways are coded as class 4, with feature names "AURORA AVENUE" and "EVERGREEN POINT BRIDGE", respectively, with no alternative names.

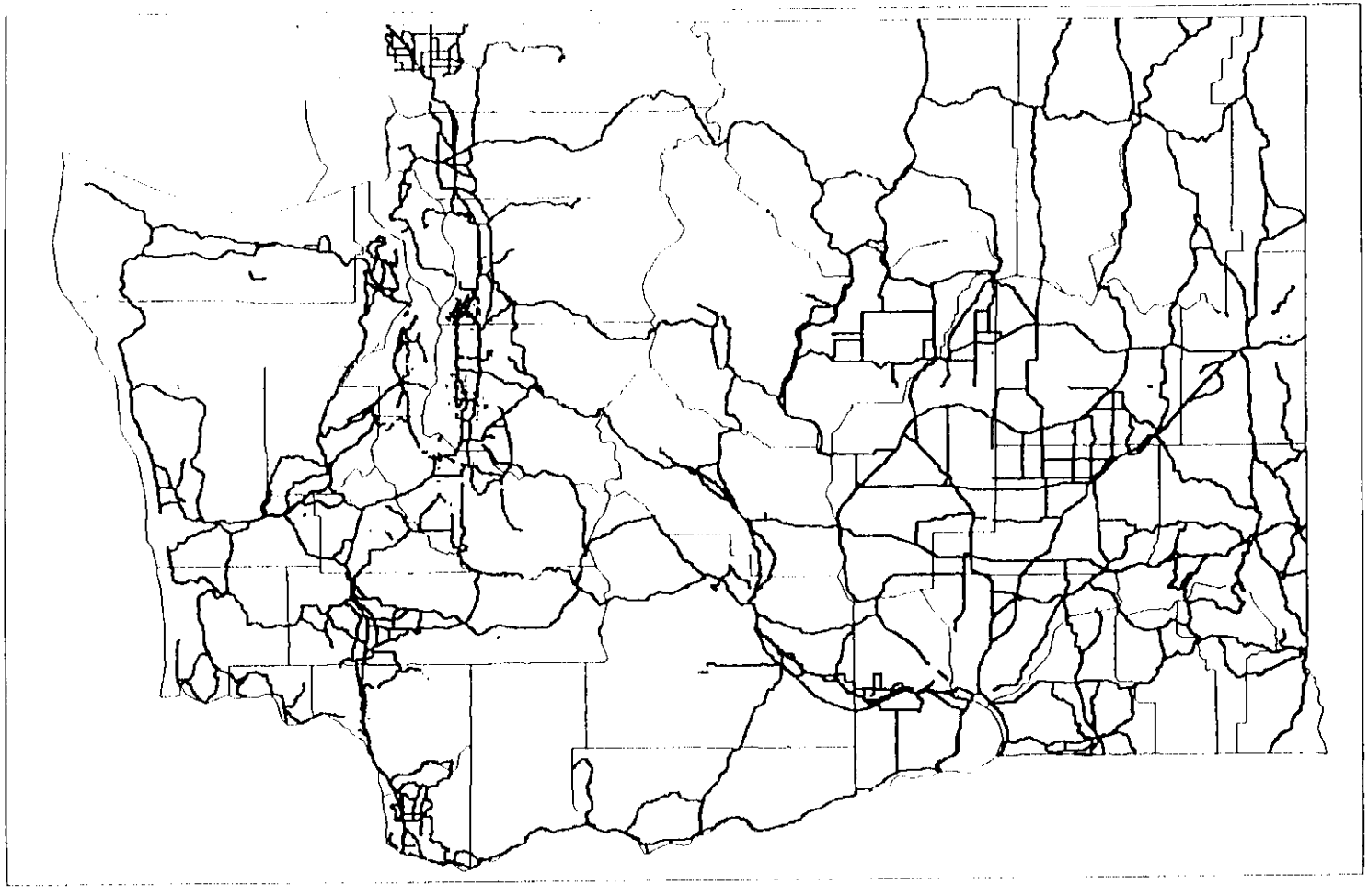

Figure 1. All Extracted Roads

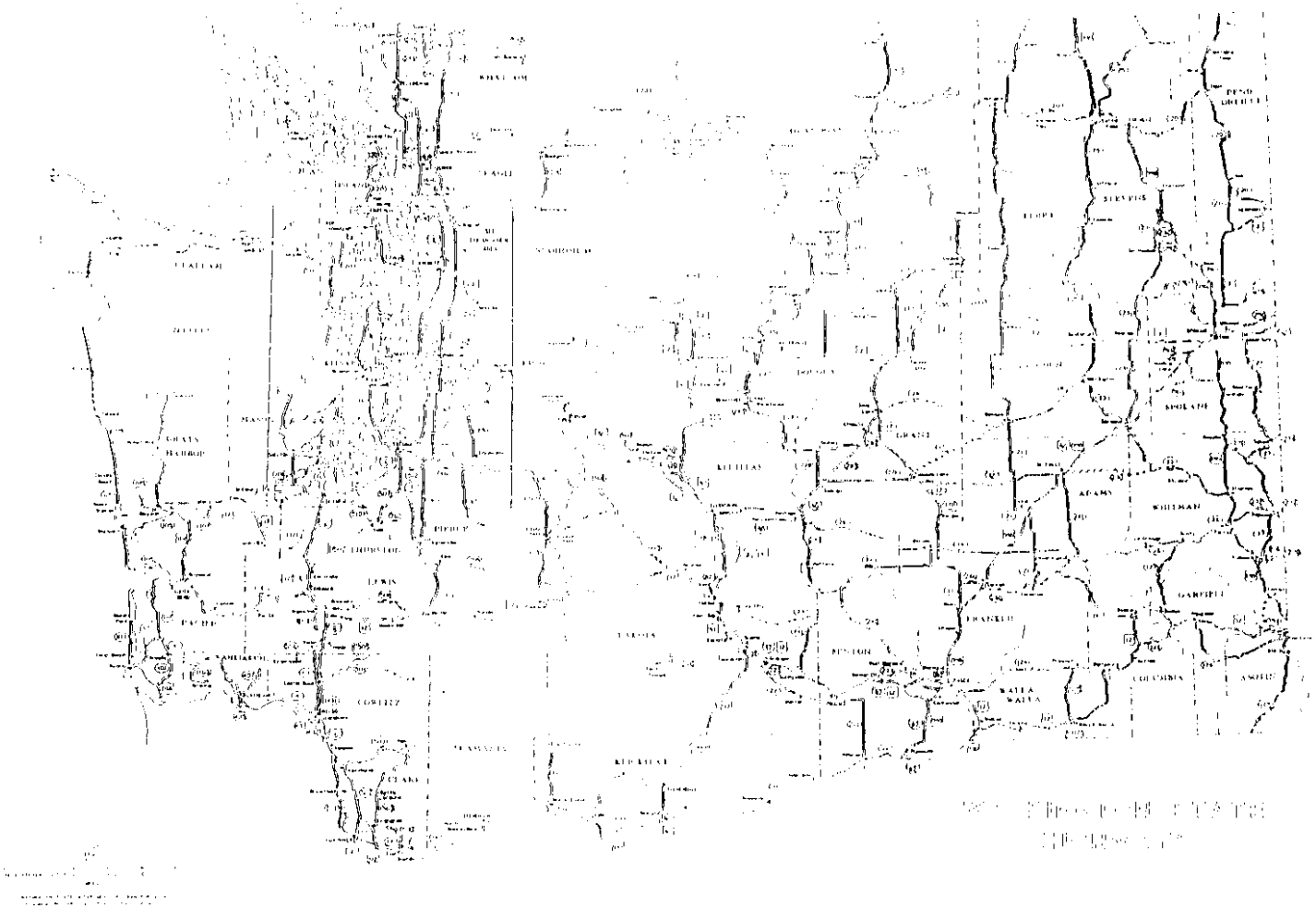

Figure 2. WSDOT Digitized Highway Network

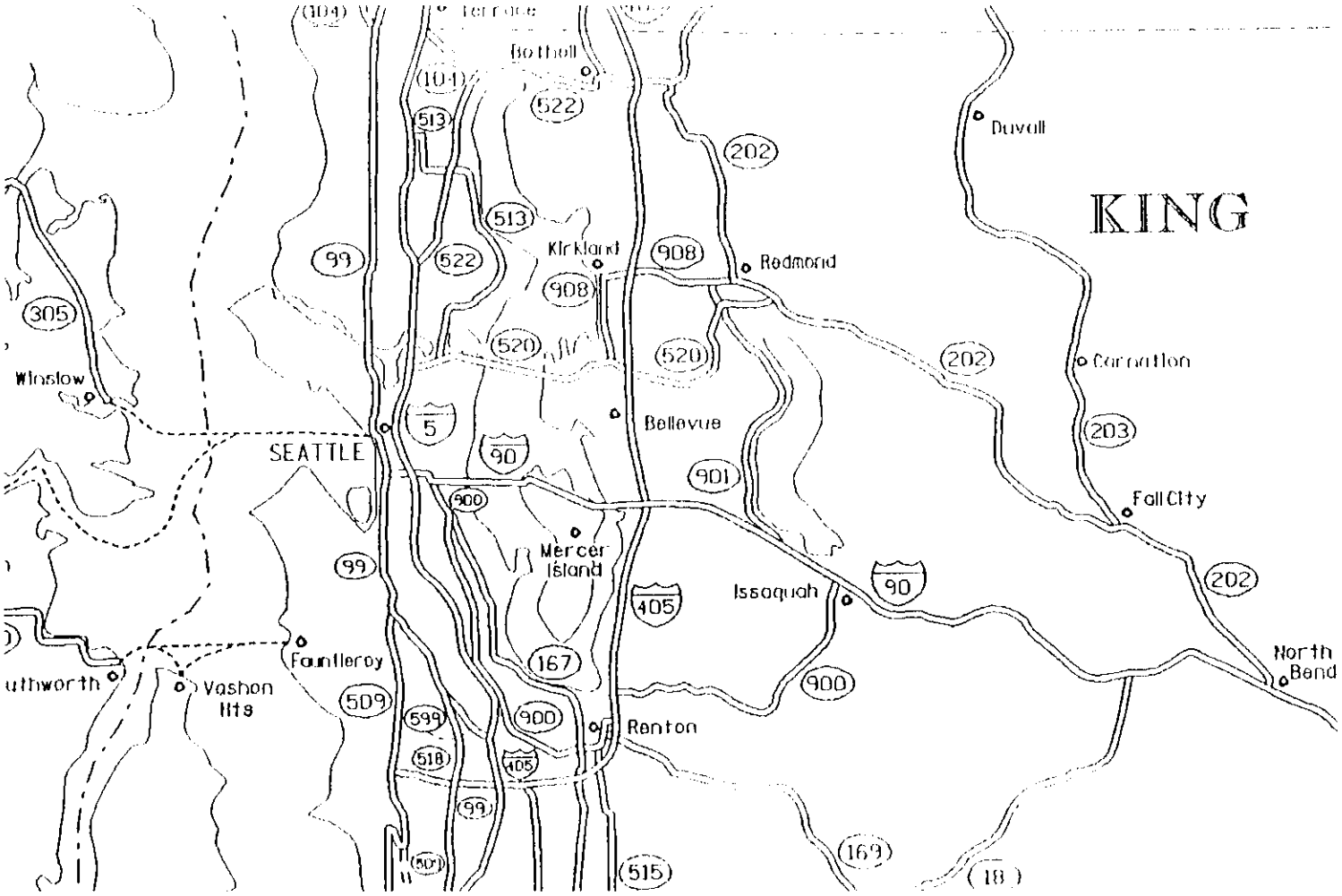

Figure 3. King County Numbered Highways

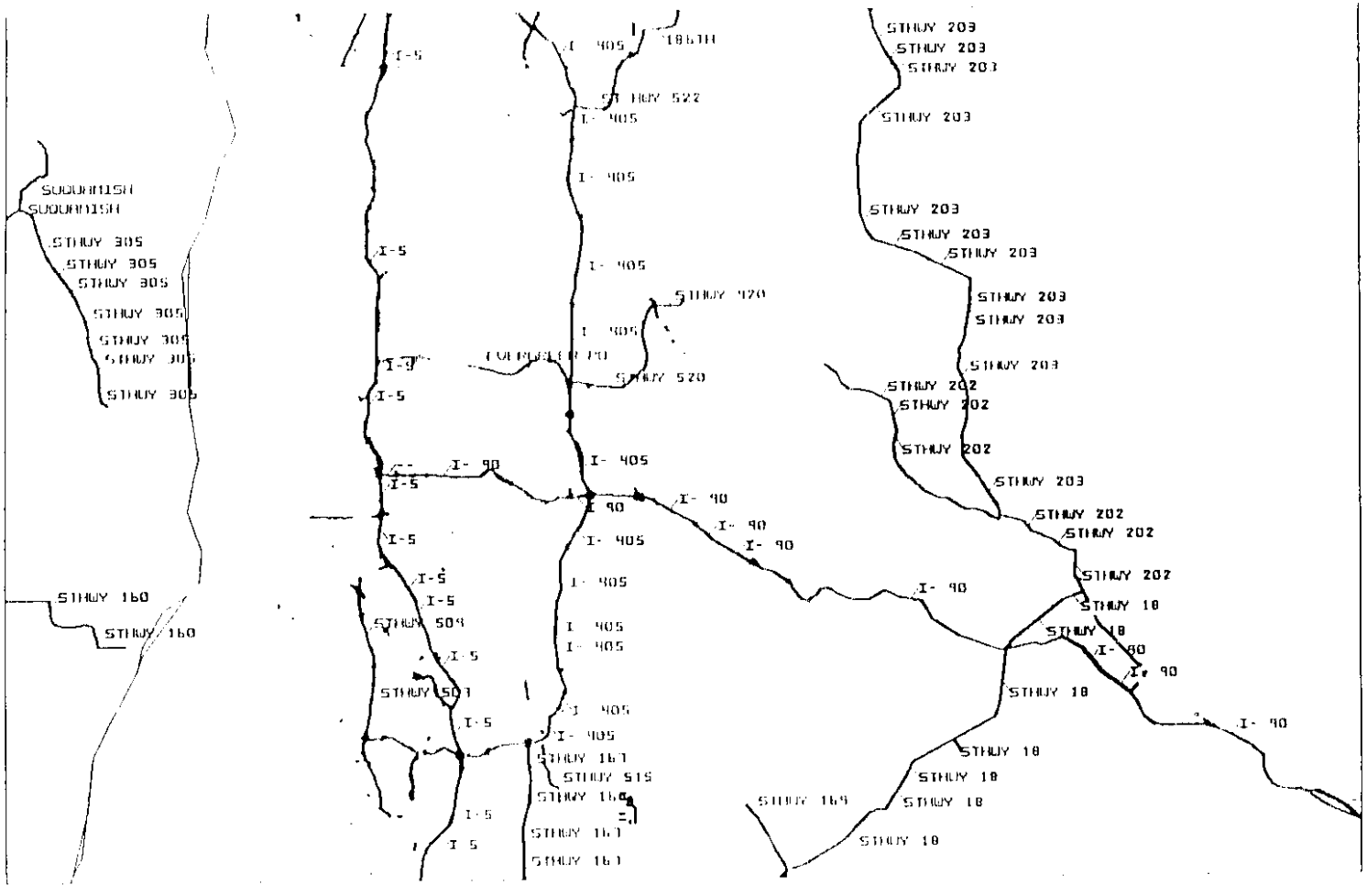

Figure 4. King County Extracted Highways

During the process of selecting the segments that represented the interstate and U.S. highway network, problems with the extracted data revealed other inconsistencies and omissions in the TIGER/Line data.

Inconsistent Form of Names

- More than one format for naming highways, making extraction by feature name more difficult. For example, the names given to segments of Interstate 5 north of Vancouver (Figure 5 -- arrows added outside TransCAD) were "I5", "I- 5", and "I 5". Between Olympia and Tacoma, Interstate 5 was also named "I-5" and "1-5" (Figure 6).

Rename without Alternate

- The U.S. highways were often given other names where they passed through cities and towns. For example, U.S. Highway 12 was called "MAIN" where it passed through Dayton and Pomeroy in southeast Washington (Figure 7). The designation "USHWY 12" was not given, even as an alternative name. The segments were extracted only because they were coded as class 2.
- U.S. Highway 195 south of Spokane (Figure 8) was named variously as "USHWY 195" (consistent with the naming of U.S. highways), "US 195", "SR 195" (an error), and "INLAND EMPIRE HWY". In some cases highway segments had no name, such as U.S. Highway 101 near Olympia; or main line segments were identified as feature type "RAMP", which led to their omission by the extraction program (Figure 9).

Misnamed Segments

- In southeast Snohomish County, many of the local class 4 roads were given the feature name "I 5" or "I 5 RAMP" in error. This resulted in the extraction of many segments that had no connection to the highway network (Figure 10).

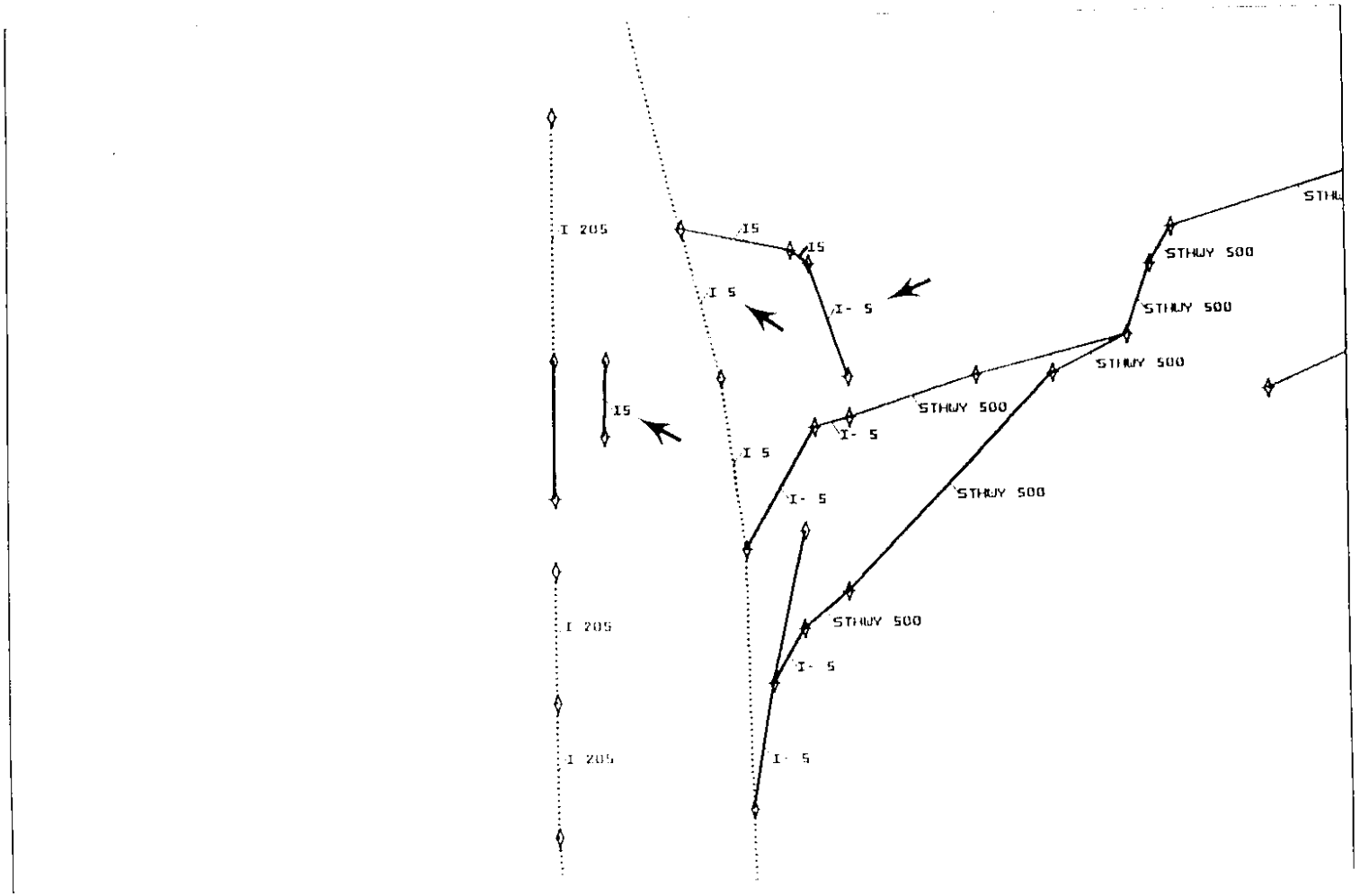

Figure 5. Interstate 5 Names

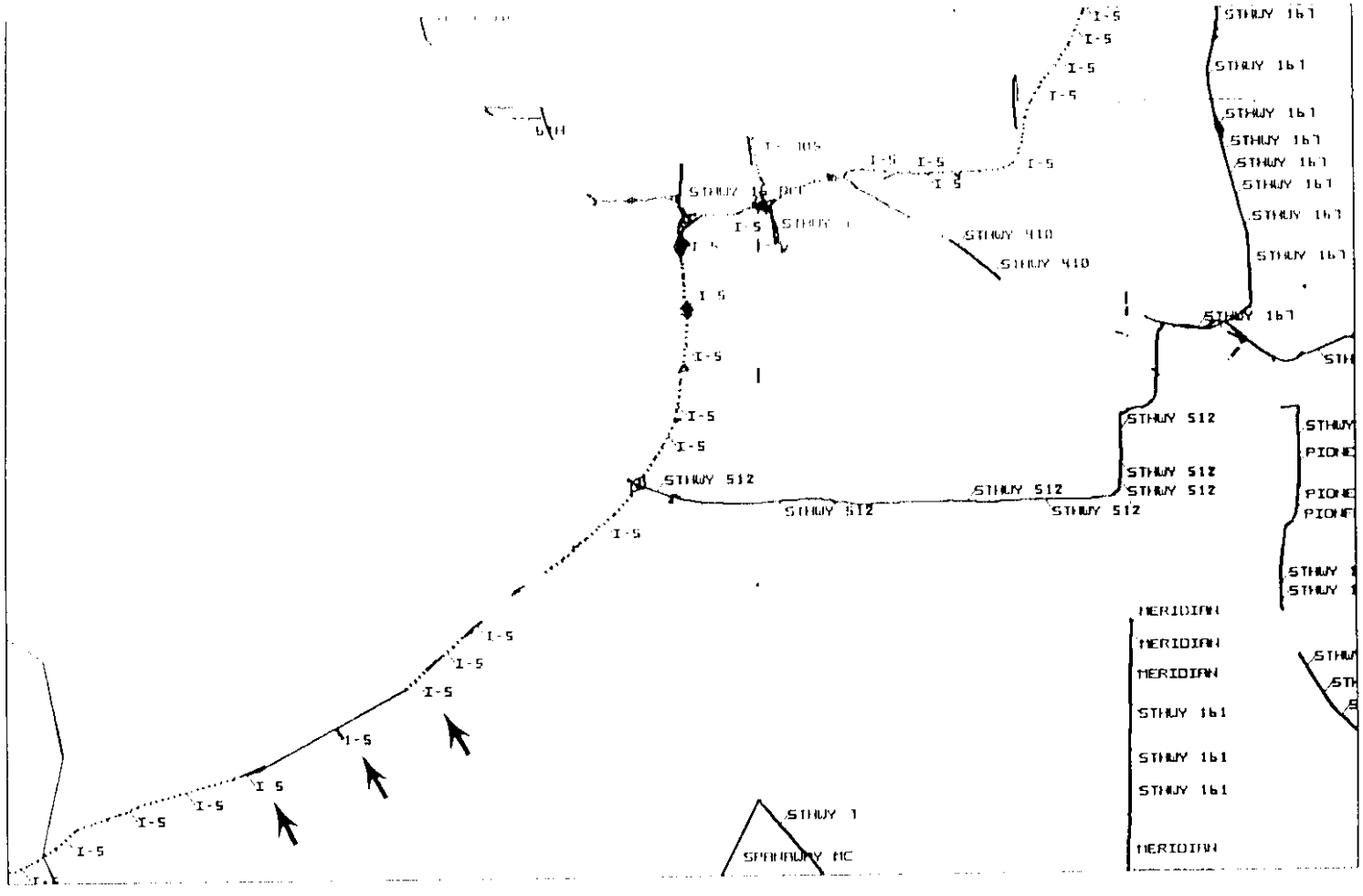

Figure 6. More Interstate 5 Names

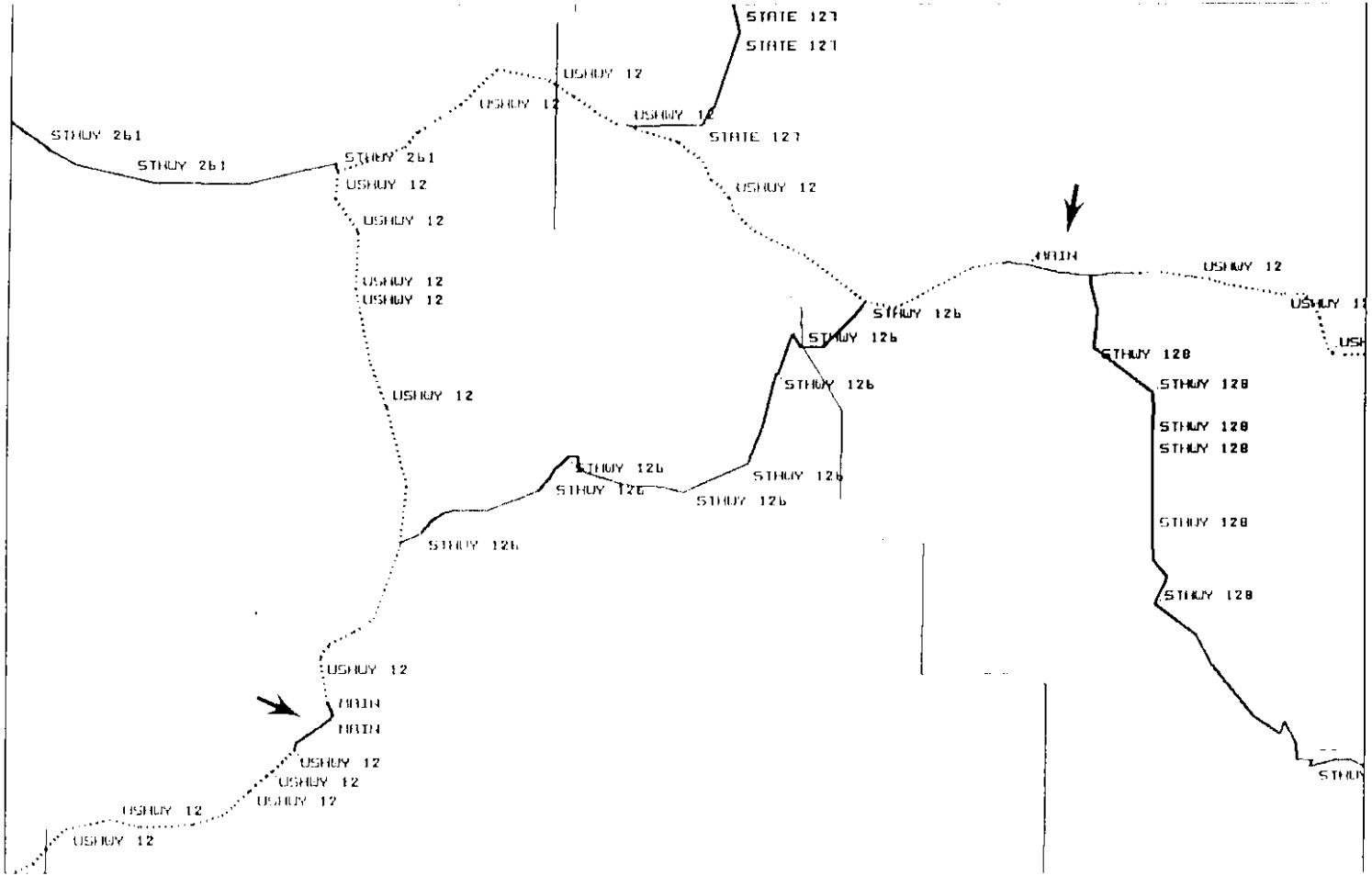

Figure 7. U.S. Highway 12 called MAIN

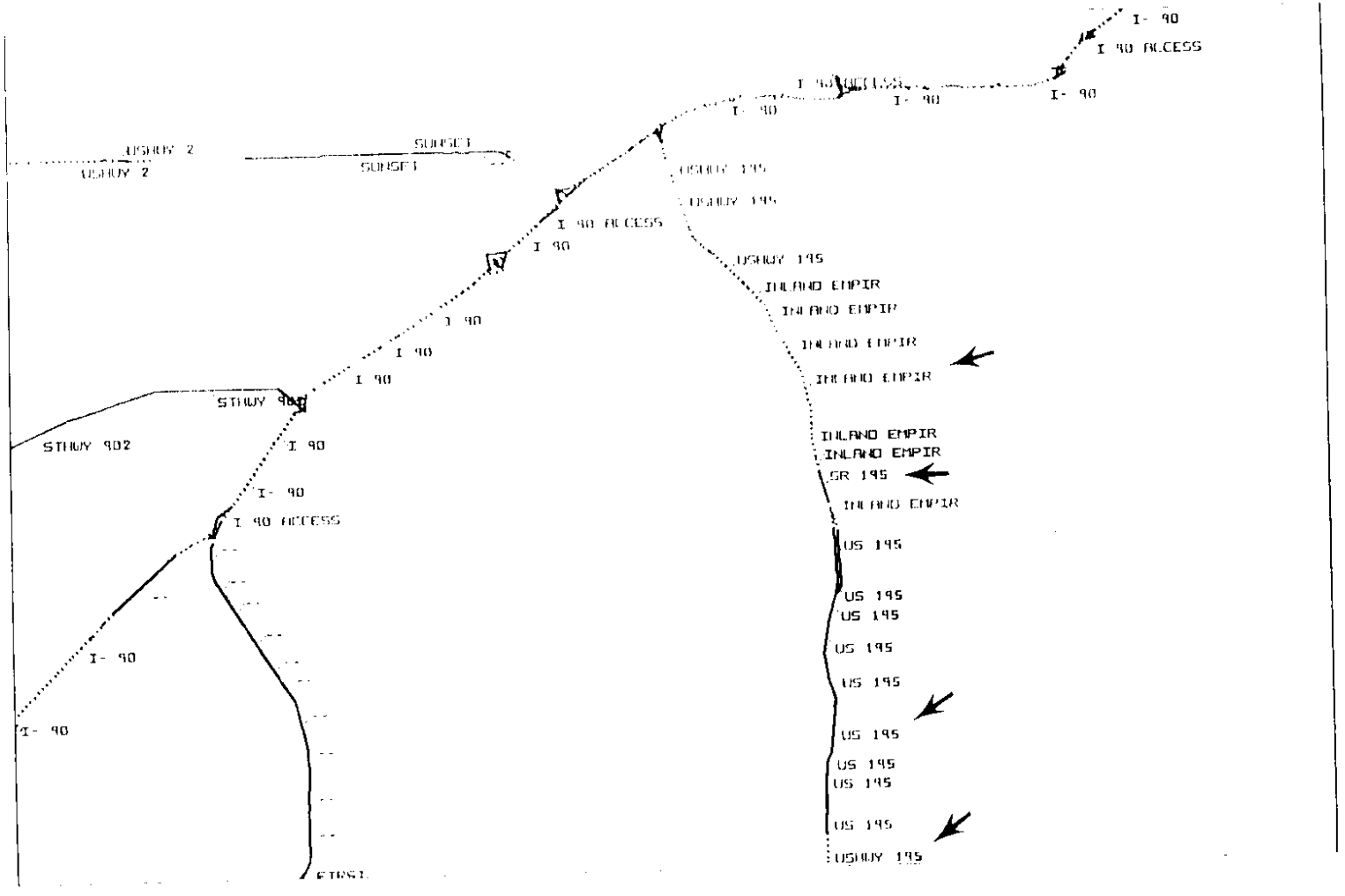

Figure 8. U.S. Highway 195 Names

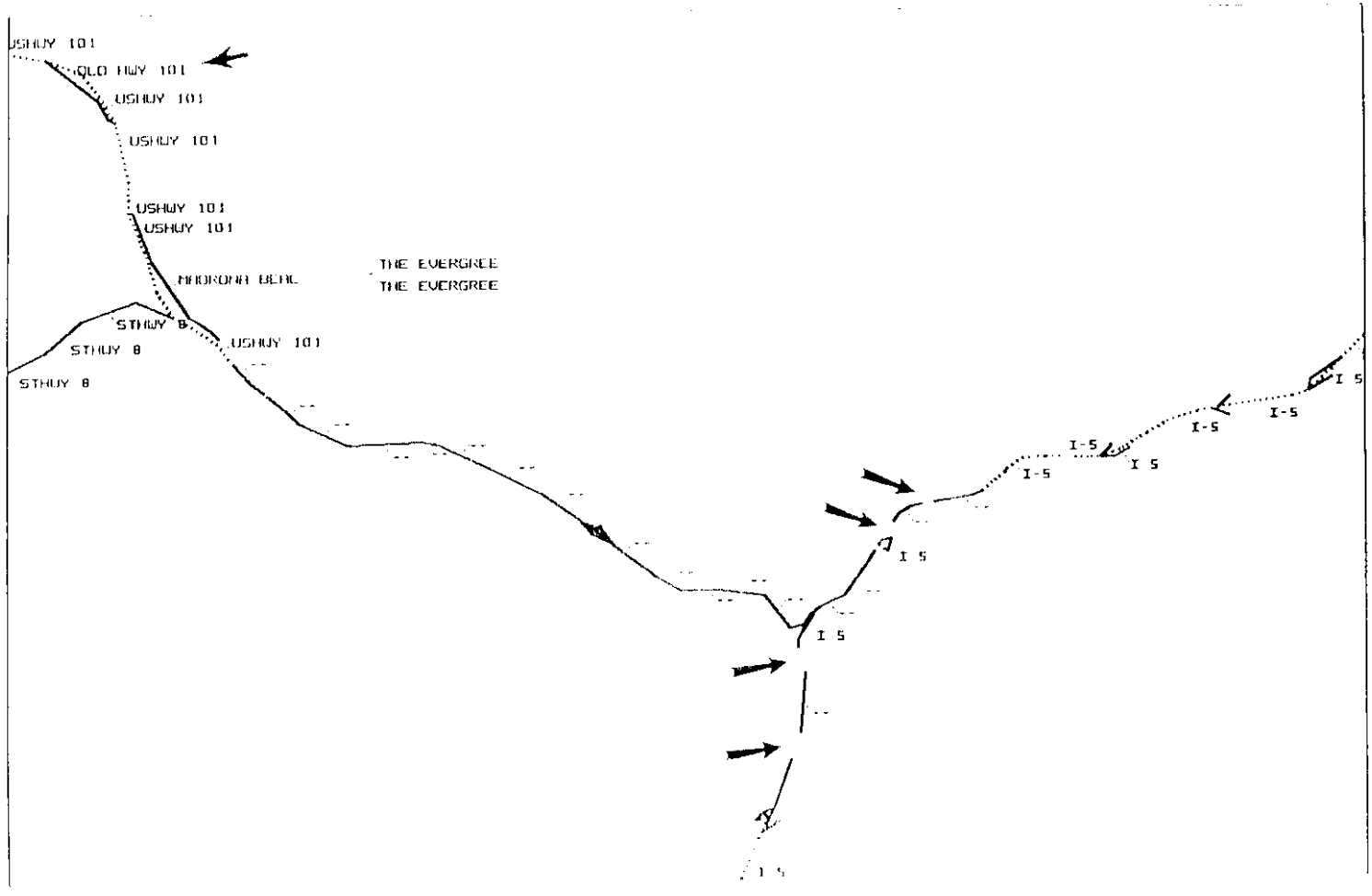

Figure 9. Missing Names and Segments

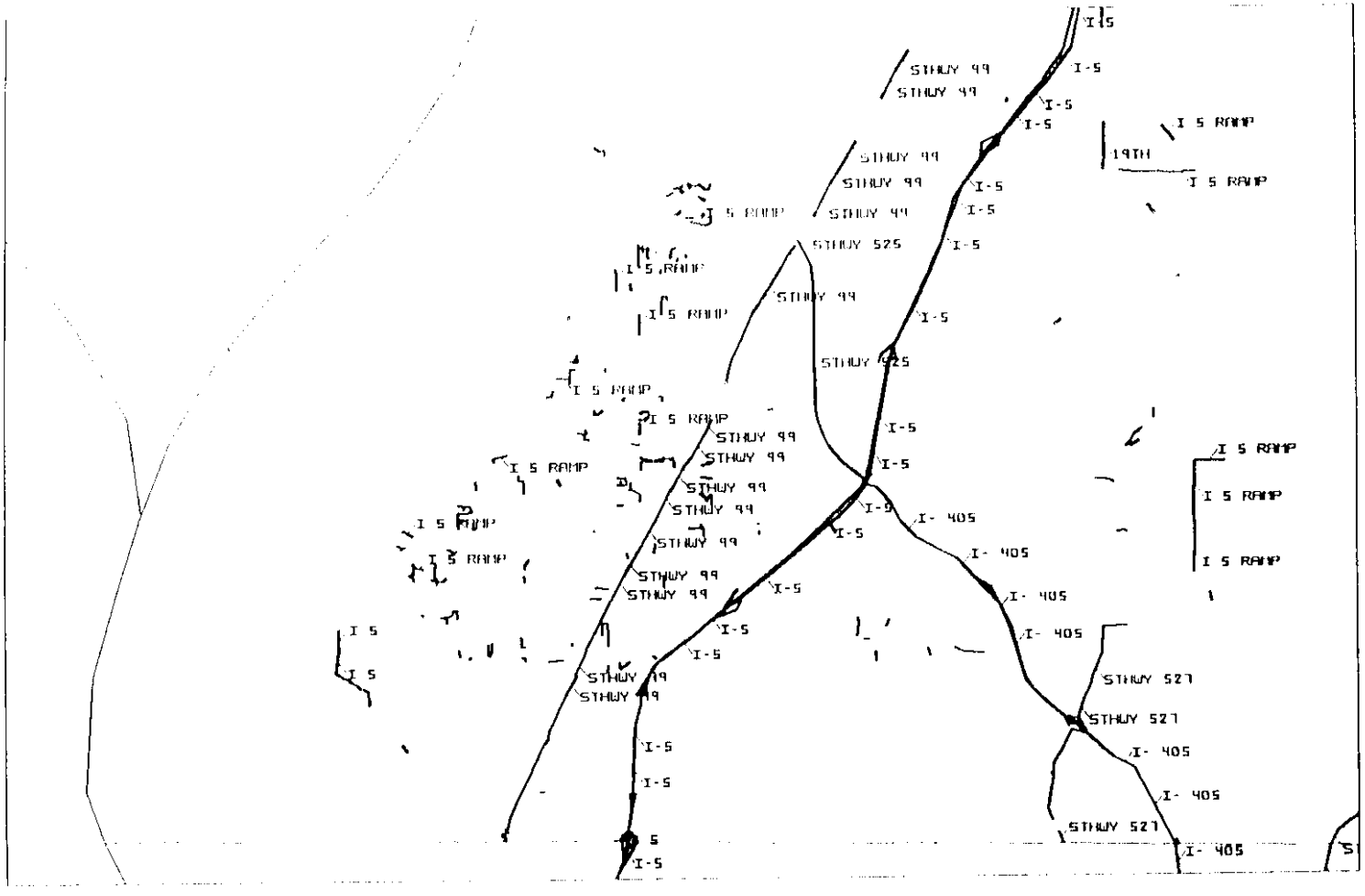

Figure 10. I 5 Ramps

- In 37 places in the network the database lacked a segment completely. The missing segments were in the original TIGER/Line files but had not been extracted. These segments were added to the database and selected.

THE MODEL NETWORK

Figure 11 shows all the roads in the original TIGER/Line file for one county (Snohomish). Figure 12 shows the TIGER segments that were extracted and built into the TransCAD database. Figure 13 show the segments selected within TransCAD by any name that started with "US" or "I" and was not an interchange. Figure 14 shows the clean selected network for Snohomish County after the research team had manually removed unnecessary segments and added missing links. The clean selected interstate and U.S. numbered highway network for the whole state are shown in Figure 15.

Figure 16 compares the selected Interstate and U.S. highway network with the simplified abstracted network based on county level zones. Figure 17 is the simplified network with a node at each county boundary and roadway intersection. This is output that could be translated into a microcomputer transportation planning package.

INTEGRATION OF A TIGER NETWORK

Once a node-link model network has been developed from TIGER, census based and non-census data structures could be integrated with the network. The research team assessed this situation for both census and non-census data.

Census Data

Since TIGER was developed primarily to delineate census enumerator boundaries, the 1990 census data would be compatible with a TIGER derived model network. The integration of census data would require software (that would probably use GIS technology) to sort the data and identify the census tract (or

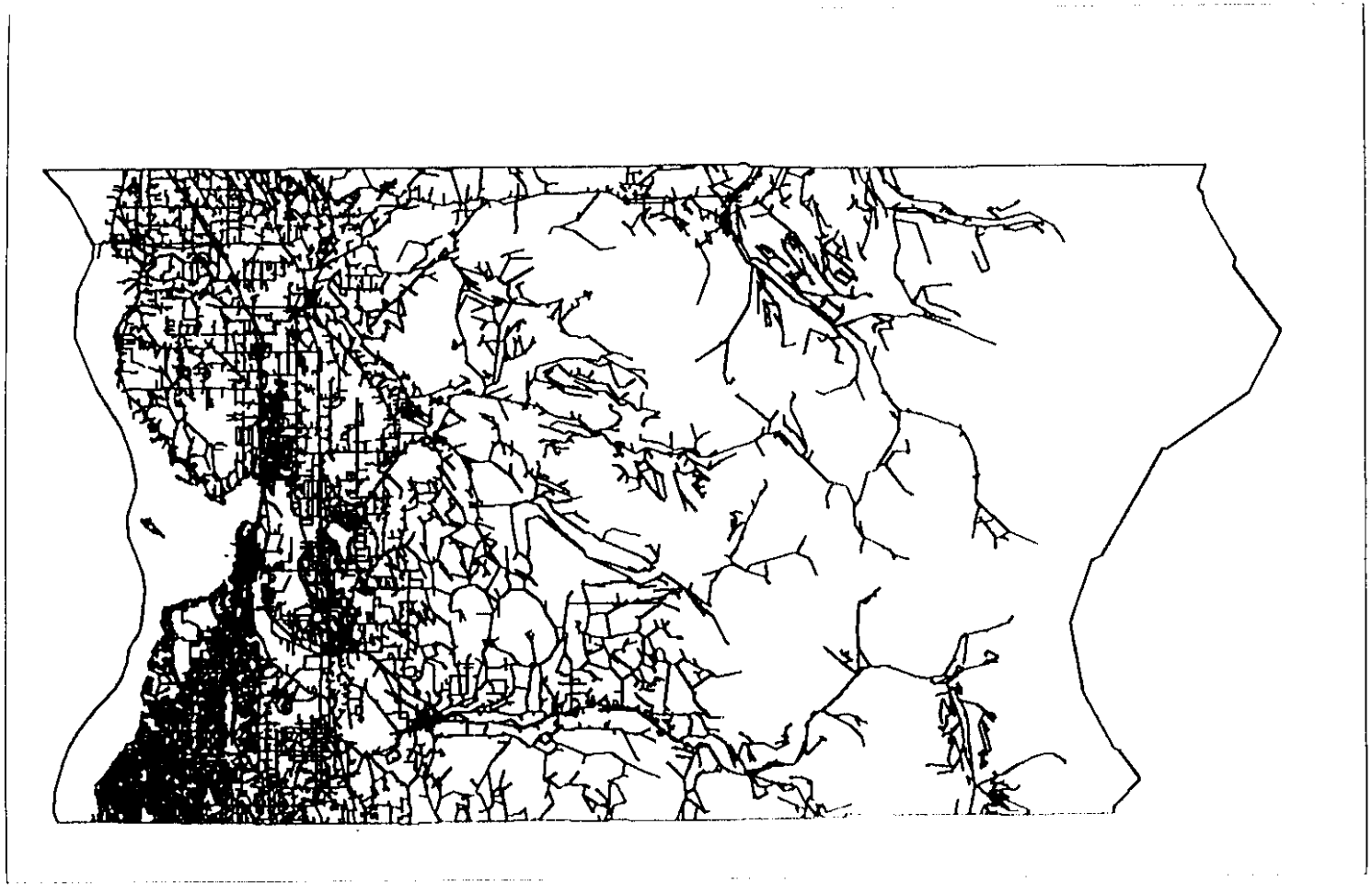

Figure 11. Snohomish County - All TIGER/Line Roads

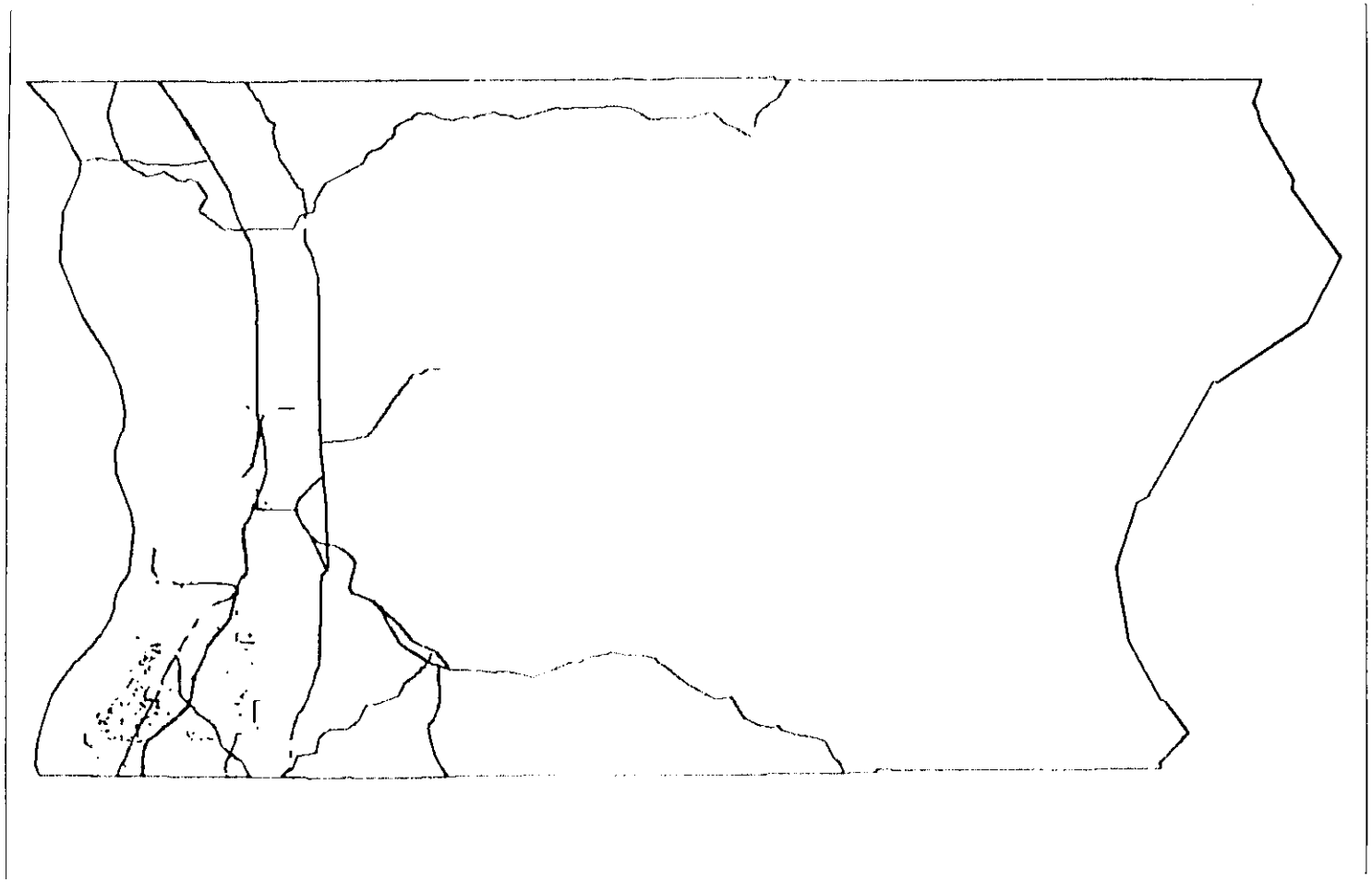

Figure 12. Snohomish County - Extracted Segments

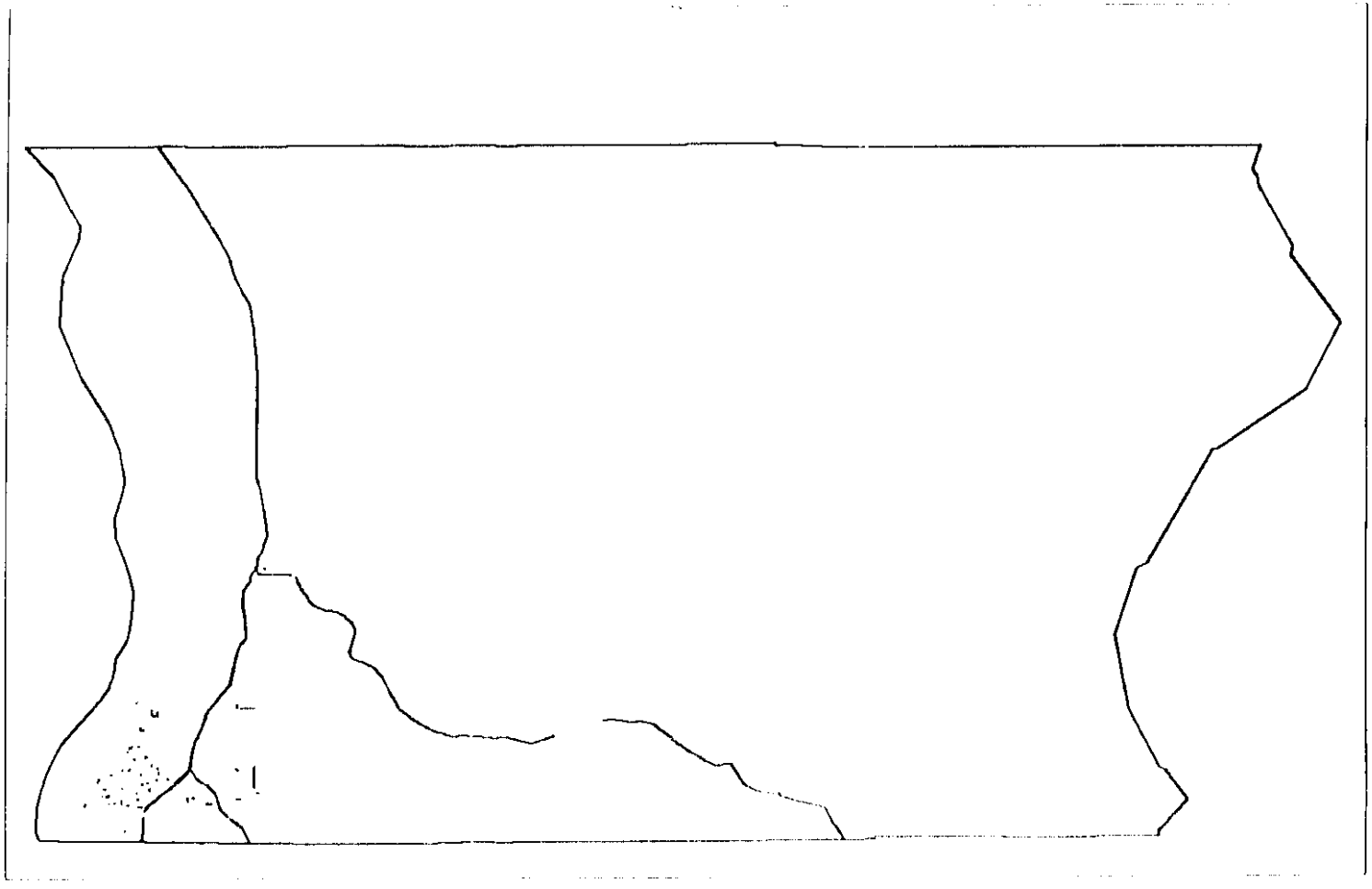

Figure 13. Snohomish County - Segments Selected by Condition

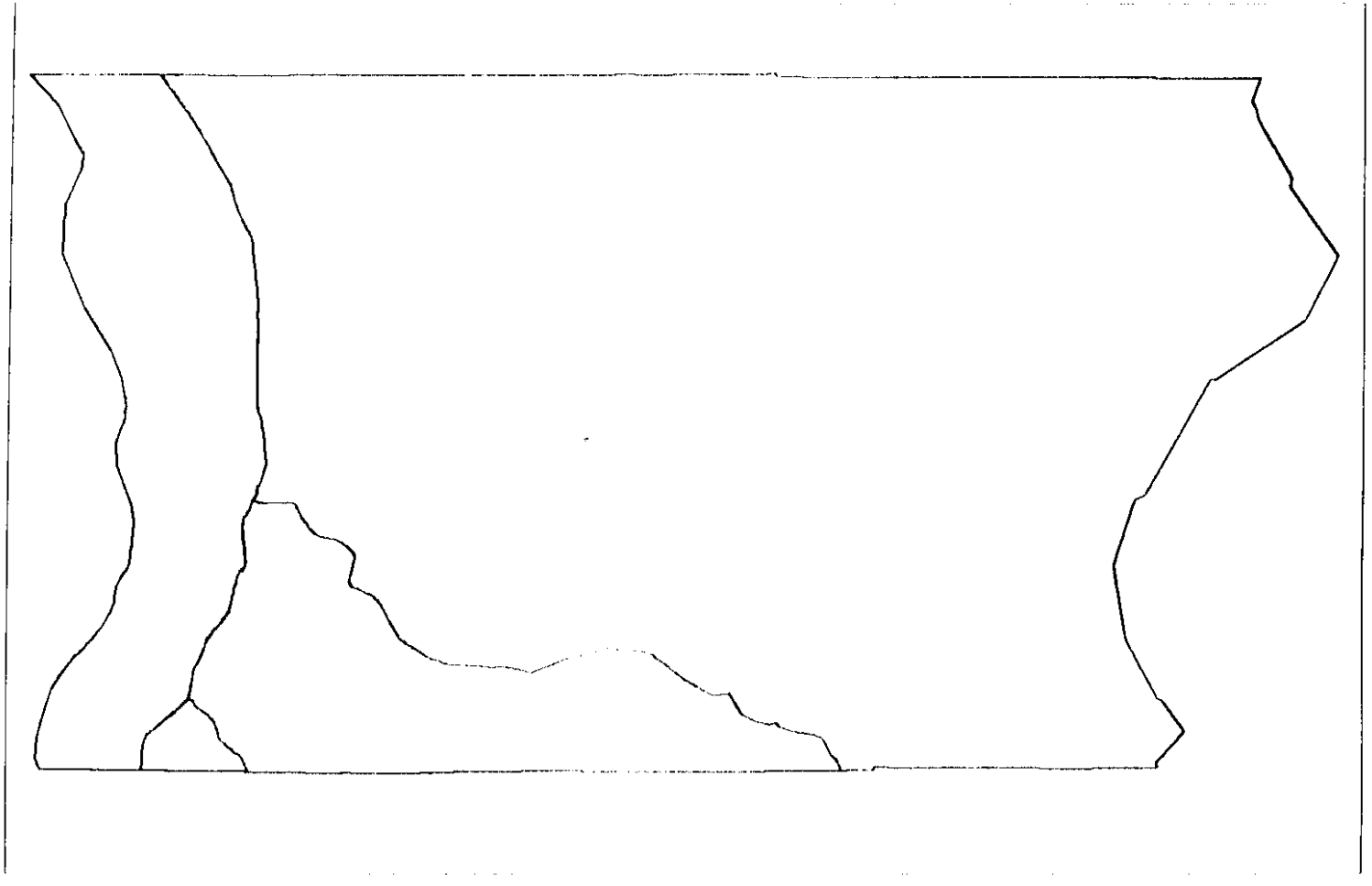

Figure 14. Snohomish County - Selected Segments

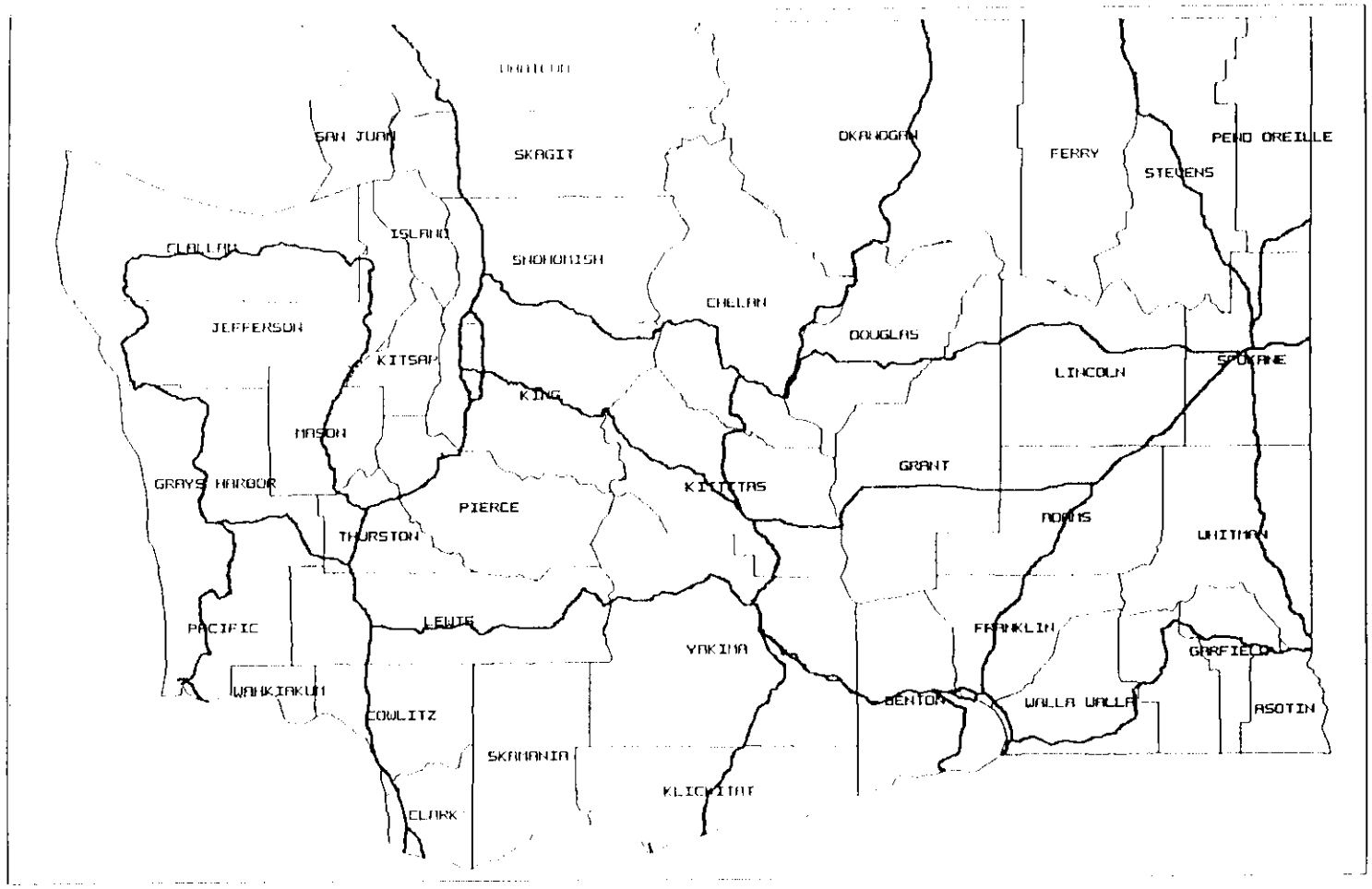

Figure 15. Selected Segments

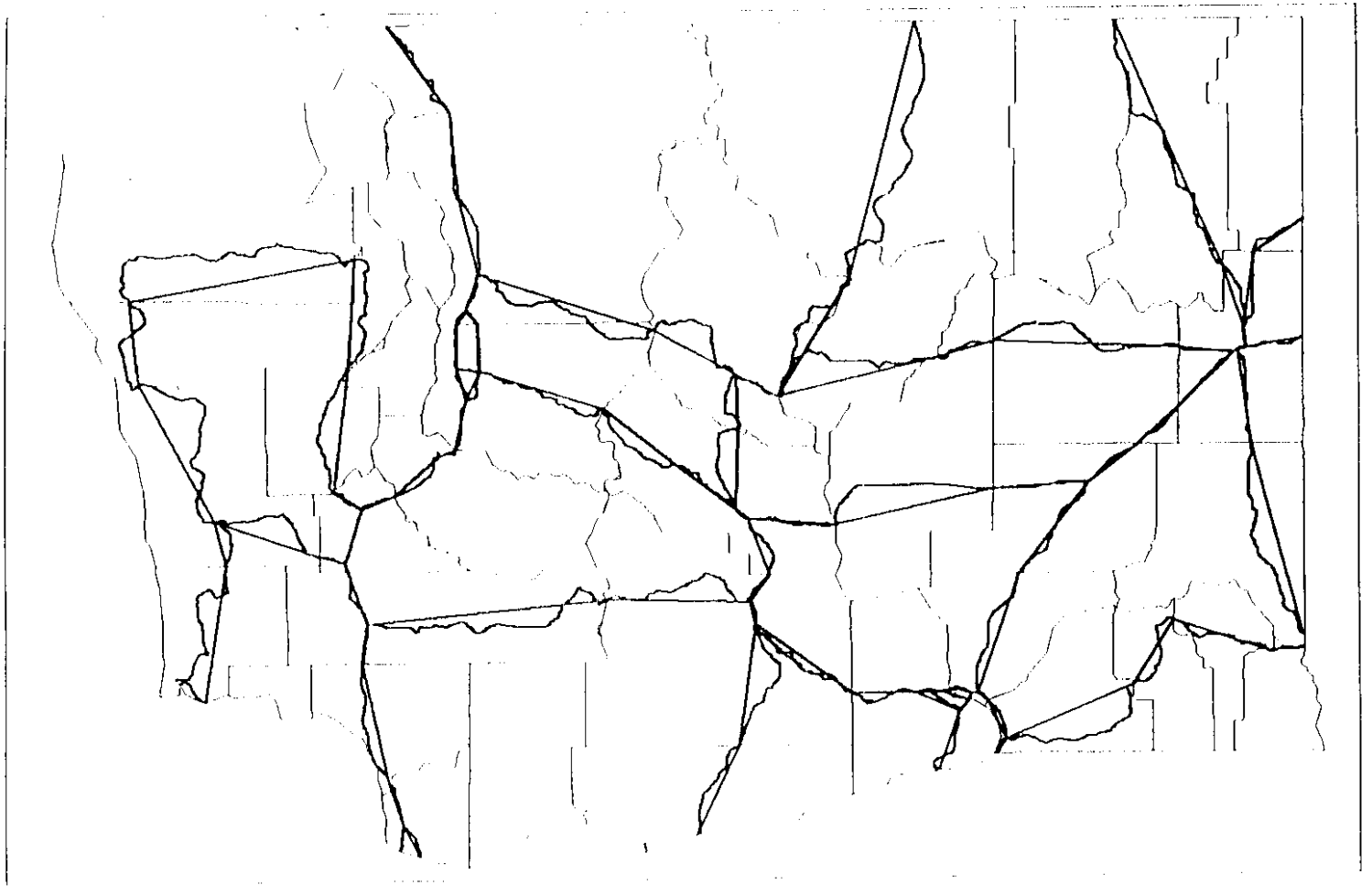

Figure 16. Selected Segments and Abstract Links

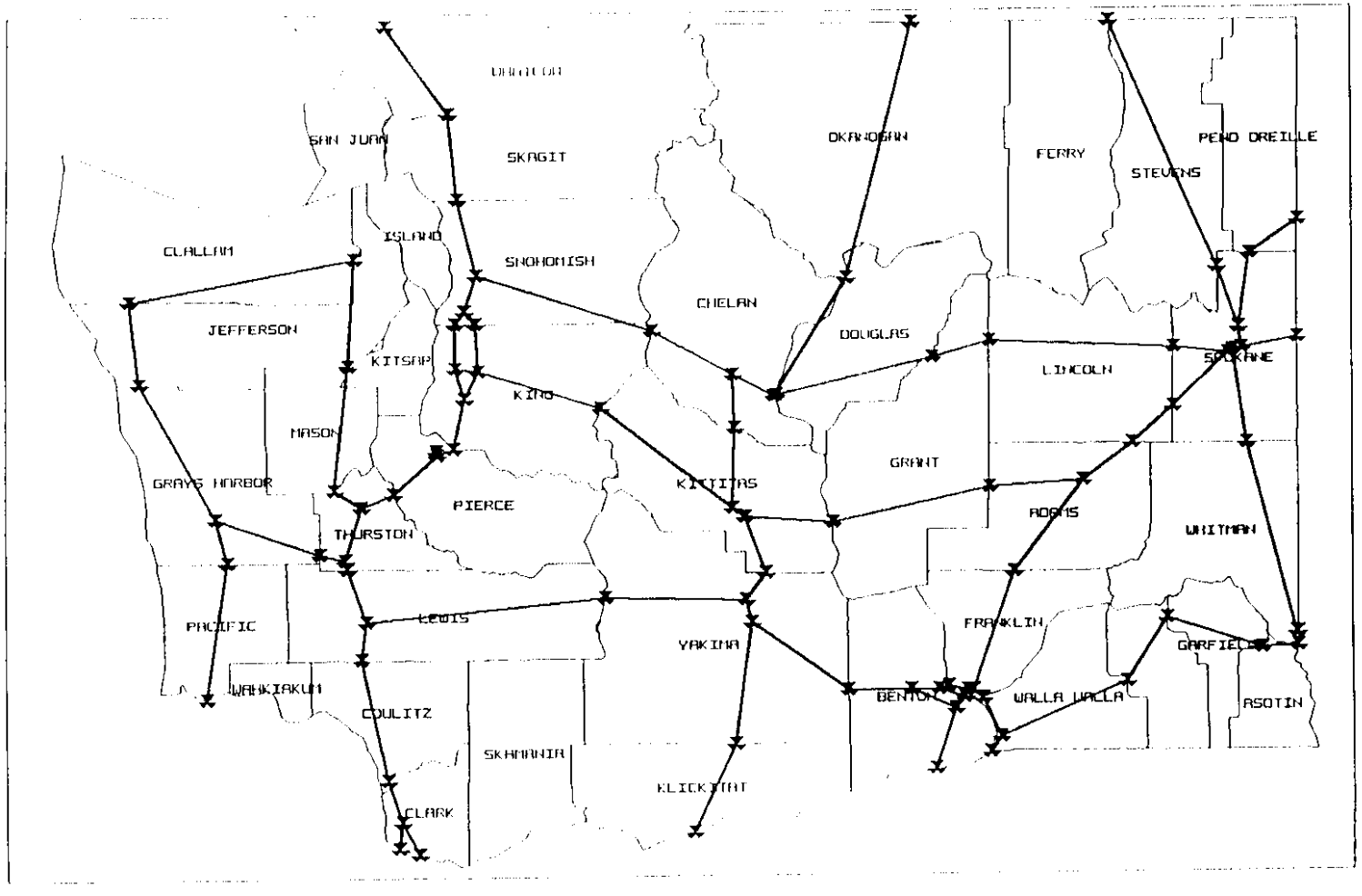

Figure 17. State Abstract Network

census block) in which a link or node of interest was located. Since the census tract (and block) information would be in the TIGER/Line files, the software could be fairly simple. Once a network link or node has been assigned to a census tract, with further processing the census socio-economic data could be linked to the model network.

Since the traditional transportation modeling process depends on a zonal division of a study area, a link between socio-economic data from the 1990 census to a model network would assist modelers in readily modifying the size and nature of a zoning scheme to better match both the spatial variation of the travel behavior factors in the study area and the focus of the study. Ideally, the use of census data would result in more accurate zoning structures and trip generation. In addition, modelers would be able to better understand and analyze link volumes and other model output using census data.

Non-census Data

The integration of non-census data structures with a TIGER derived model network would require a common data element. One element that is included for a portion of the TIGER/Line file is address information. Using existing address matching techniques, modelers could integrate other survey sources with mailing address information to the TIGER data. An example of this procedure would be a mail-out housing survey in which each respondent's address was known. By using an address matching program, the geographic location of housing data could be linked to the address information in the TIGER/Line file.

Unfortunately, addressing information is only available in the TIGER data for the geographic areas that match the 1980 DIME file coverage. As a result, rural areas and new urban areas do not have address information in TIGER/Line files.

An alternative data integration method might utilize the latitude and longitude coordinates contained in the TIGER/Line files. A number of states have mapping data stored in a coordinate format, which, with projection conversion

techniques, could be converted to latitude/longitude coordinates. Once in a latitude/longitude format the data could be integrated into a TIGER derived network.

However, before data sets could feasibly be integrated into a TIGER derived network, a number of issues and technical question would need resolving. Among these issues would be the following:

- Census tract boundaries are often roadways. If a link in a transportation network overlays a census boundary, how does one assign tract characteristics to that link?
- The TIGER/Line files have address information, but only in areas that correspond with the 1980 DIME file. If one wishes to use addressing to integrate a data structure, how should rural areas and new urban areas that do not have address information be incorporated?
- The use of latitude and longitude data as an element to link TIGER data to other data sets raises concerns about the conversion between coordinate systems. Particularly on a state level, coordinate conversion can be difficult because of the mapping projections necessary at this scale.

APPLICATION

If a transportation modeler was planning to use the TIGER/Line files as a basis for a state or corridor level transportation planning effort, the following issues would have to be addressed:

- The extraction of a state level network involved considerable "hand" coding and correcting by members of the research team. Assuming that software capable of handling the scope of a state level project was available, the extraction and development of a network capable of handling a entire state highway system for a state the size of Washington (350 MB) would take an estimated 2 to 3 person-months. Since the size of state level TIGER files varies from 24 MB (Delaware) to 1,542 MB (Texas), the time necessary to extract a state level roadway network would also vary considerably between states. For most states the TIGER files are large enough that they tend to overwhelm the storage capacity of a microcomputer hard disk. Frequently modelers would have to manipulate the data in pieces and to extract only the elements of the data that were of interest.
- To the best knowledge of the research team, software capable of manipulating the TIGER network and census data on a state level or for a large corridor does not exist. At this time, people interested in integrating data sets, would have to develop their own programs. Since the TIGER/Line files are a census product, it is possible that vendors will eventual develop and sell TIGER oriented software and data sets in a form that would be usable for a state level transportation effort.
- Because of the level of detail present in the TIGER/Line files, a TIGER derived network might have more applications for a smaller

study area (state) or corridor study that needed to utilize the high level of roadway detail available in the TIGER/Line files.

- Before census and, especially, non-census data sets data were integrated with a TIGER derived model network, a number of issues would need to be addressed, computer programs written, and procedures developed.

REFERENCES

Data User's Service Division, Bureau Of the Census; TIGER/Line File: Boone County, Missouri -- Technical Documentation; 1988.

Fleck, John; A Demonstration of the TIGER File in Boone County Missouri, Presented at the Second Conference of Transportation Planning Methods, Orlando Florida, April 24-28, 1989.

Nyerges, Timothy L and Dueker, Kenneth J.; Geographic Information Systems in Transportation; Paper Supported by Federal Highway Administration, Office of Planning, July 6, 1988.

Simkowitz, Howard J.; Using the Tiger/Line Files for Transportation Planning, Presented at the National Conference on Effective Planning Techniques for Small and Medium-Sized Areas, Phoenix Arizona, December 6-8, 1988.

Simkowitz, Howard J.; Geographic Information Systems: An Important Technology for Transportation Planning and Operations, Draft August 5, 1988.