

# Lawrence Berkeley National Laboratory

## Recent Work

### Title

Interfacing 1990 U.S. Census TIGER Map Files with New S Graphics Software

### Permalink

<https://escholarship.org/uc/item/66w4d5jn>

### Authors

Rizzardi, M.  
Mohr, M.S.  
Merrill, D.W.  
et al.

### Publication Date

1992-07-01



# Lawrence Berkeley Laboratory

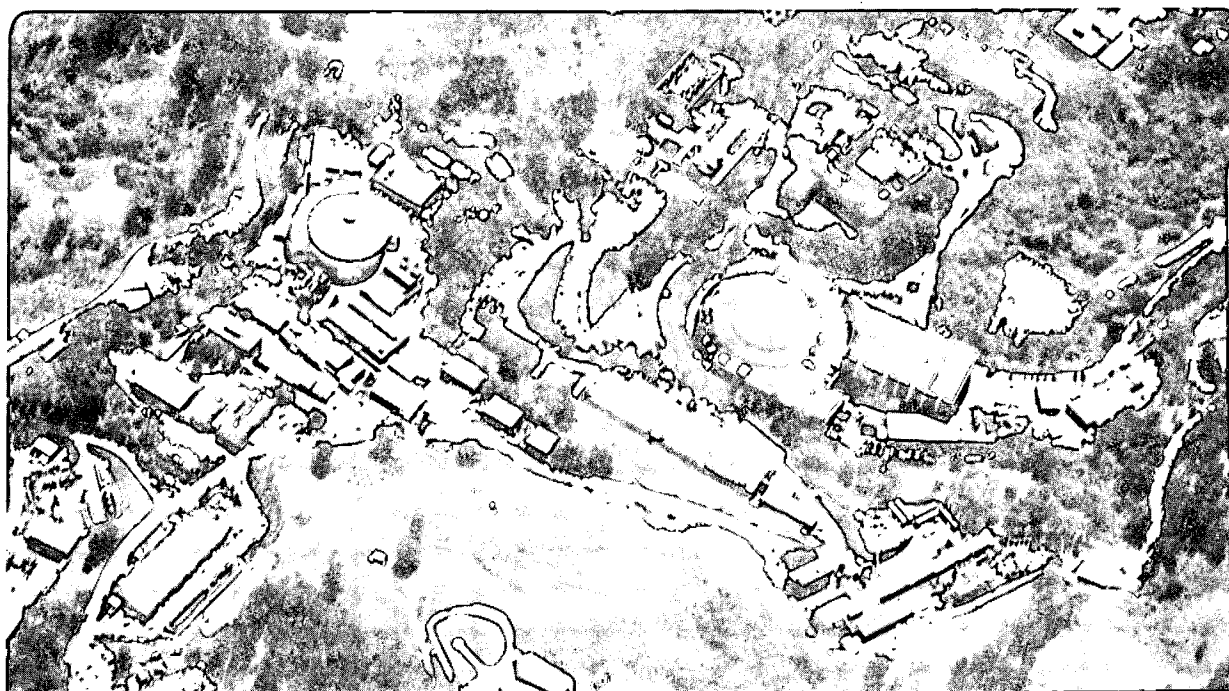
UNIVERSITY OF CALIFORNIA

## Information and Computing Sciences Division

### Interfacing 1990 U.S. Census TIGER Map Files with New S Graphics Software

M. Rizzardi, M.S. Mohr, D.W. Merrill, and S. Selvin

July 1992



REFERENCE COPY  
Does Not  
Circulate

Bldg. 50 Library

Copy 1

LBL-32641

## **DISCLAIMER**

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor the Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or the Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or the Regents of the University of California.

# Interfacing 1990 U.S. Census TIGER Map Files with New S Graphics Software

M. Rizzardi <sup>\*,§</sup>      M.S. Mohr <sup>†,§</sup>      D.W. Merrill <sup>†,§</sup>  
S. Selvin <sup>†,§</sup>

July 1992

---

<sup>\*</sup>Information and Computing Sciences Division  
Building 50B:3238  
Lawrence Berkeley Laboratory  
1 Cyclotron Road  
Berkeley CA 94720  
tel: 510-486-4635  
fax: 510-486-4004  
internet: marizzardi @ lbl.gov

<sup>†</sup> Information and Computing Sciences Division  
Lawrence Berkeley Laboratory

<sup>§</sup> School of Public Health  
University of California, Berkeley

This work was supported by the Director, Office of Epidemiology and Health  
Surveillance; Office of Health; Office of Environment, Safety and Health;  
U.S. Department of Energy under Contract No. DE-AC03-76SF00098.

## Interfacing 1990 U.S. Census TIGER Map Files with New S Graphics Software

*In 1990, the United States Bureau of the Census released detailed geographic map files known as TIGER/Line (Topologically Integrated Geographic Encoding and Referencing). The TIGER files are accessible to the public through purchase or Federal repository libraries such as the University of California, Berkeley library. Culminating a decade of development, the TIGER files contain 24 billion characters of data describing various geographic features including coastlines, hydrography, transportation networks, political boundaries, etc. covering the entire United States.<sup>1,2</sup> Many of these physical features are of potential interest in the study of geographic clusters of disease. Unfortunately, the large TIGER database only provides raw alphanumeric data; no utility software, graphical or otherwise, is included.*

*Recently AT&T's statistical graphics software New S, associated with the Unix operating system, has been extended to include map display functions.<sup>3,4</sup> The low cost of New S and its high-level approach toward statistical analysis of data and graphics are motives for wanting to interface TIGER maps with New S.*

*We discuss the approach, difficulties, and findings of our recent efforts to develop an interface for TIGER map file display using New S. Considerable attention is focused on attaining the New S map format requirement of complete polygons. Successful formation of complete polygons from the TIGER database depends, in large part, on the user's familiarity with the TIGER file structure and its feature classification system. New S census tract boundary maps of counties can be consistently created; however, choosing more specific boundaries, such as seashore edges, has been less successful because of TIGER's classification system for its diverse geographic data.*

# 1 Introduction

Advances in statistical software packages have made data analysis more readily available and less tedious for statisticians and other users. While it has become standard practice to provide numerical data in computer readable formats, applied analysis is often hampered by the lack of geographic maps in usable computer formats. The difficulty of obtaining maps that are easily used in a data analysis environment often discourages various statistical and epidemiological approaches to research. In particular, the lack of map databases, and of software to effectively utilize such databases, have placed substantial restraints on analysis.

In 1990, however, the United States Bureau of the Census released detailed geographic base files known as TIGER/Line (Topologically Integrated Geographic Encoding and Referencing) which contain detail on the physical features and census tract boundaries of every county in the United States.<sup>1,2</sup> The TIGER database is attractive for two reasons. First, it is publicly available through the Bureau of the Census on tape or cd-rom for a minimal fee. Second, it contains 24 billion characters of data which describe geographic features of interest to the Census Bureau such as coastlines, hydrography, transportation networks, political boundaries, etc.<sup>5</sup> Unfortunately, the large TIGER database only provides raw alphanumeric data; no utility software, graphical or otherwise, is included.

On the other hand New S, a popular statistical software package by AT&T, has easily operated functions that permit advanced graphics in conjunction with data analysis.<sup>3</sup> New S has the ability to plot contours, lines, segments, and points. However, of special interest is the New S function *map* and its options.<sup>4</sup> Using the *map* function, which requires polygons as input, census tracts can be quickly selected, plotted, shaded, etc. New S graphics combined with the TIGER database has obvious potential.

This paper reports on our efforts to use the TIGER map files with New S, especially to construct census tract maps of counties. While census tract boundaries are inherently polygonal, they are not organized as such in the TIGER database. This conversion of the TIGER "line" format into New S "polygon/polyline" format is one facet of the work reported here. Also we discuss the selection and extraction of auxiliary geographic information from TIGER files for graphical display using New S.

## 2 Geopolitical Maps and New S

The data analysis computer system New S, associated with the UNIX operating system, includes a map-producing function which is easy to use once the geographic map is put into the required format. Other New S commands can then be used to analyze or enhance the resulting map. For example, the user can overlay additional lines, segments, points, or contours on top of an existing map display. Furthermore, which polygons and boundaries are displayed can be specified by the user. The map function's efficiency and flexibility in working with a database is conducive to productive data exploration.

New S uses a polygon representation for map storage and retrieval; i.e., a map is viewed as a collection of polygons. Each polygon is described as a set of connected chains or "*polylines*". A polyline consists of the connected line segments between two "*nodes*"; a node being a point where three or more line segments meet. In other words, a map is described by polygons which in turn are described by polylines.

Each map in New S is composed of a set of three files based on the general structure above. Briefly, the first file contains the name of each polygon and their respective polygon identification (ID) codes. The second file defines each polygon as a list of serially connected polyline ID codes. The third file gives the longitude-latitude coordinates of the points forming each polyline. A more detailed description of these three files follows using a small example map for illustration.

The first file, termed the "*name*" file, consists of three fields. The first field contains the regional name of the polygon, such as the census tract code. Some regions however may require several unconnected polygons to describe their borders. In this case, the region's name is repeated in a new record for each polygon of the region and a within-region polygon count kept in the second field. (Fields one and two are separated by a colon.) For illustration, refer to the example map shown in Figure 1 and notice that Tracts 001 and 003 each involve two polygons: an "island" forms the second polygon of tract 003 while the shared border of tract 002, which lies within tract 001, constitutes the second polygon of tract 001. The third field of the name file gives a running count of the polygons thus far listed and serves as a polygon/record ID code. (Fields two and three are separated by a tab.) Figure 2 lists the name file for the Figure 1 map.

In the second file, termed the "*polygon*" file, polygons are formed as lists of signed polyline ID codes. The sign indicates the direction of the polylines

so that when serially connected they trace out the polygon without “lifting pen” from paper (clockwise denoting positive area). The  $i^{th}$  record within the “name” file refers to the  $i^{th}$  record in the “polygon” file with the ASCII characters “EOR” denoting end-of-record in the polygon file. Thus, in Figures 1 and 2, polygon 1 is circumscribed by polyline 1 followed by polyline 3; the negative sign on 3 implying the coordinates of polyline 3 should be traced in reverse of the direction listed in the third file. Polygon 2 is circumscribed entirely by polyline 2. Polygon 3 is also entirely circumscribed by polyline 2 but in the reverse direction. Etc.

The third file, termed the “line” file, consists of the polyline records; the record numbers corresponding to the polyline ID codes used in the polygon file. Each polyline record begins by identifying the polygons which appear on each side of the (directed) polyline; the left side polygon ID code followed by the right side polygon ID code. (A polygon ID code of zero is used to denote external boundary.) The longitude-latitude coordinate pairs given in radians follow. “EOR” is again used as a record separator. For the example map of Figures 1 and 2, polyline 1 proceeds in a clockwise direction with polygon 1 on its right and external boundary to its left. Polyline 3, on the other hand, is shared by polygons 1 and 4 which, for the direction listed, has polygon 1 on its left and polygon 4 on its right. Thus, for a continuous clockwise tracing of polygon 1, polyline 1 must be followed by a reversely directed polyline 3 as shown in Figure 2’s polygon file.

### 3 TIGER Map Files

The New S library provides map databases for the states and counties of the continental United States. Though lacking in detail, these databases are ideal for quick sketches of state or county boundary maps. The additional geographic information available in TIGER files will often be desired.

TIGER maps, like New S maps, are composed of chains; in TIGER terminology “lines”. These lines are functionally equivalent to the polylines of New S. It is important to note that, because of the level of detail in TIGER maps, the polygons formed by these polylines may be as small as one square meter and usually do not correspond to political boundaries such as census tracts.

There are twelve separate files in the TIGER database for every county, of which the first two are of concern here. The first file contains the *basic data records* while the second contains the *shape coordinate points*. The



TIGER basic data record describes the characteristics of a line and reports the longitude-latitude coordinates of the line's two end-points to six places of implied decimal precision. The second file contains the longitude-latitude coordinates of the intermediate points of the line. Figure 3 illustrates the essential aspects of the two TIGER files corresponding to the example map of Figure 1.

The basic data record contains descriptive information that is helpful for selecting desired lines from the TIGER database. Every character position in the record designates a certain descriptive or organizational code. For the purposes at hand, four sections of the basic record are useful. Following their order of appearance in the record, they are: 1) the record identification number and feature name, 2) the census feature classification code (cfcc), 3) the state, county, and tract ID codes for the left and right side entities of the line, and 4) the longitude-latitude coordinates of the line's "from" and "to" points. Other codes, such as zip codes and the Alaska Native Regional Corporation codes might also prove useful.

The *record identification number* (record ID) of a line is a ten digit code that is unique to a line within the county and is used to reference its shape coordinate records in the second file. No flag in the basic data record indicates how many shape points, if any, augment the line. The only indication that a shape coordinate record exists for the line is if the ID of the basic data record is shared by a record in the shape coordinate file. (In the example map files of Figure 3, no shape points exist for record ID 07.) Each shape coordinate record consists of ten shape points given as longitude-latitude pairs. A line is traced beginning at the basic data record "from" point, through the sequential shape points, and ends at the basic data record "to" point. "From" and "to" line end-points imply direction, thus distinguishing between the left and right sides of the line. For example, line record ID 06 in Figure 3 traces the island of the example map. When a line requires more than ten shape points, the next record in the shape coordinate points file will have the identical record ID number and the additional shape points added to the sequence. If the number of shape points is not a multiple of ten, the remaining longitudes and latitudes of the record are assigned values of zero.

The record ID's in the basic data file are not necessarily ordered sequentially. Fortunately for computational speed and algorithm design, the record ID's in the shape coordinate points file have the same ordering as those in the basic data file. For the example map, record ID 05 precedes record ID 04 in both TIGER files (Figure 3).

Included in the basic data record are the *geopolitical codes* for the left and right side of each line. The state, county, and census tract codes are of fundamental interest. Interest often focuses on mapping the census tracts of a particular county and, in this case, one would select only those lines for which the census tract code on the left and right sides differ. Similarly, for a county-level map, one would select only those lines for which the county codes on the left and right sides differ.

## 4 Census Feature Classification System

The *census feature classification code* (cfcc) describes the physical “*feature*” of a line. Roads, railroads, power transmission lines, hydrographic features, cartographic features, and many others are described by the cfcc. Great detail exists for certain features; there are more than fifty different codes for various types of roads ranging from four types of jeep trails to nine types of county roads. Right and left side cfcc descriptions are not given because the cfcc describes only the singular feature type of the line. The cfcc is a source of strength because of its detail, and confusion because of the occasional line falling under the optional description of more than one cfcc.

The tiered coding system of the cfcc, however, makes it simple to isolate features of a certain class and level. The cfcc consists of a character followed by two integers. The character specifies the class of the feature; for example, roads begin with “A” and railroads begin with “B”. The first integer is for subclassification, so jeep trails would start with “A5” and interstate highways with “A1”. The second integer allows the user to discriminate within a subclassification. Computer code is easily written to select lines that are of a specific cfcc classification. An important use of the cfcc, discussed later, is for selecting auxiliary feature lines to plot on top of an existing New S map.

Caution must be exercised when using the census feature classification codes. As large as the TIGER database is, feature misclassification certainly exists. In addition, because each line has but a single cfcc, the Census Bureau uses a hierarchical system in designating the cfcc of a line.<sup>1,2</sup> For example, a perennial shoreline label, “H01”, would have priority designation over a property line designation, “F40”. No formal documentation has yet been published on the process of feature classification.

Also, a specific feature may not be traceable as a continuous line via its cfcc. An interstate highway, for example, may cross an entire county and

*appear* as a single line. Nevertheless, each segment may be described by one of the nine class codes specific to interstate highways. In these instances it is advantageous to use a more general level of classification when specifying the cfcc selection criteria.

Care is also required when dealing with bodies of water. For example, census tracts often extend three miles out over the ocean instead of ending at the shorelines. Although these are the official tract boundaries, they are most often not of practical use. Epidemiologists for example, would want the beach water-lines to be used as the tract boundaries in a population-based investigation. In this case, the census feature classification codes can be used to ignore tract boundaries that exist over water and keep lines that are classified as shoreline; the water-side of the line now denoted as external boundary. Census "*block*" codes may also prove useful for water identification as their trailing two digits are designated "99" for a water area.

## 5 Formatting TIGER Maps for New S

Two primary steps are involved in transforming TIGER file data into a New S map database: filtering and polygon construction. Although lines are linked into polygons in the TIGER database, the sheer number and minute scale of these polygons render them impractical for most uses. For example, a TIGER-level polygon could be a rectangle with boundaries formed by a fence line, a zip code boundary, a census tract edge, and a ditch. Unwanted TIGER lines, therefore, must be identified as such and filtered out. Subsequently, for use with New S, the remaining lines need to be assembled into complete polygons. Figures 4 and 5 show an unfiltered and filtered demonstration map, each with complete polygons. Here, the tract boundaries that extended over the ocean and bay were replaced by the shoreline, the lake retained, and other features, such as rivers, railroads, etc. were removed.

Writing conditional computer commands that filter extraneous features under ideal circumstances is not difficult; however, filtering the data so that complete polygons are formed is usually not so straightforward. If a polygon cannot be completed, it may be difficult to track down the source of the error due to the complexity of the TIGER files. It has been our experience, however, that errors are most often the result of the cfcc filtering process; errors in the TIGER files are infrequent. The inconsistency of the cfcc labels makes filtering a delicate job and, for advanced filtering, experience with the

cfcc labels is a requirement and familiarity with the region helpful.

Once unnecessary TIGER detail has been filtered, the remaining lines must be linked into complete polygons. Working first with the filtered basic data records, lines that have a common census tract code are grouped together to create a tract-level map. Traveling clockwise, lines within a tract are linked end-to-end until the polygon closes. As the example map illustrates (Figure 1), some census tracts may consist of more than one polygon. In these cases, the lines remaining after the initial polygon has been formed are linked to form other polygons until the set of lines is exhausted. Data from the shape coordinate points file are ignored during this phase because they are not necessary to determine the assembly order of the polygon lines.

After the polygons have been identified, the defined map can be put into the database format required by New S. The tract codes used to group the lines for polygon formation can be used as names in the New S “*name*” file. Whether a polygon was formed using the left or right side of a line determines the sign of the polyline ID code in the New S “*polygon*” file listing. Recall that each TIGER line consists of at least a “to” and a “from” coordinate. If shape points exist for the line, they are simply inserted between the two end-points in the New S “*line*” file. New S assumes that the “*line*” file coordinates are in radians but plots the map in degrees.

Once the three New S ASCII map files have been made, they must be converted into indexed binary files before they can be accessed by the New S map function. New S provides directly for this ASCII to binary conversion. Details of this process are given in the Appendix.

The map function’s convenience and numerous options makes its use desirable for data analysis. New S map databases can, however, be enhanced by auxiliary TIGER geographic information which cannot be incorporated into the polygon/polyline format. Using the cfcc, more detailed geographic and political features pertinent to the current map database can be obtained from the TIGER files. Once extracted, these coordinates may be plotted on top of the overall map using the New S functions *points*, *segments*, or *lines*. This approach can greatly augment an existing map and provide flexibility in what features are plotted.

## 6 Example Applications

The California counties of San Francisco and Tulare are used below to illustrate the TIGER file filtering process and the capabilities of New S graphics

functions. In the case of San Francisco, the goal was to produce a census tract map of the city proper including shorelines. Using the New S function *segments*, every line segment defined in the 1990 San Francisco County TIGER database was plotted to make Figure 6(a). The “bubbles” seen in the west of Figure 6(a) are features associated with the Farallon Islands. The city proper, including its three mile offshore state territorial water boundaries, lies in the map’s eastern half. Although the census tract boundaries are present in Figure 6(a), they are obscured by the deluge of other TIGER features. The TIGER file line segments were easily filtered by a FORTRAN program to retain line segments that were tract boundaries – different tract codes on the left and right side. The census tract boundary segments were then assembled into polygons, formatted as a New S map database by a FORTRAN program, and then plotted by the *map* function (Figure 6(b)). The census tracts of the city proper were brought into focus using *map*’s polygon subsetting option *regions* (Figure 6(c)). Notice that the state territorial water boundaries, and the absence of freshwater lakes and nearby islands, give a distorted view of where people actually live and work in the City of San Francisco.

Generating the more representative city map required further conditional statements to be used in the filtering program (Figure 6(d)). We retained TIGER census tract boundary lines and perennial shorelines (cfcc=“H01”), except those with “water blocks” on both sides (left and right block codes end with “99”), and excluded features west of -122.7 degrees longitude (Farallon Islands). When shoreline took the place of census tract boundaries, the census tract code on the water side was designated as outer boundary. Only 2,818 out of 16,885 Basic Data File records and 212 out of 1,335 Shape Points File records remained as a result of this process. The *map* plot of this revised database appears in Figure 6(d). Geographic features for which the City of San Francisco is famous are now visible: Seal Rocks and the lakes of Golden Gate Park lie to the west, the Presidio and Alcatraz Island lie to the north, while the piers and Treasure Island lie to the north-east.

For Tulare County, our goal was to depict specified TIGER features, in this case railroads, within selected geographic regions. The 1990 Tulare County TIGER database was filtered for census tract boundary segments, polygons assembled, and then formatted as a New S map database. The filtering process resulted in 3,190 Basic Data records and 902 Shape Coordinate Points records – significantly less than the 45,636 and 24,036 records in the unfiltered files. The census tracts were then plotted using *map* (Figure 7(a)). A second call to *map*, using the *regions* and *fill* options, superimposed

a shading on tract 15 in Figure 7(a) for future reference. Figure 7(b) shows the Tulare County boundaries, the railroad features within the county (thin internal lines), and the darkened region of tract 15. For Figure 7(b), the *map* function, with option *internal* set to False, was used to display the county boundary with the internal tract boundaries absent. A second call to *map* was used to shade tract 15. The TIGER files were then filtered for railroad features (cfcc's beginning with "B") and these line segments were superimposed onto Figure 7(b) using the function *segments*.

For Figure 7(c), the area around tract 15 in Figure 7(b) was magnified by setting the longitude and latitude range options of the *map* function. For Figure 7(d), a detailed portrait of tract 15 was brought into view by plotting only tract 15 (using the *regions* option of *map*), shading it, and then superimposing all of the county railroad segments that were displayed in Figure 7(b). Those railroad feature coordinates outside the plot axes limits, as determined by the longitude and latitude range of tract 15, were ignored by New S. The level of railroad feature detail displayed in Figure 7(d) might be useful, for example, in a study of the effects of herbicides sprayed on railroad beds or of pollutants spread by trains.

Again, it is important to emphasize that filtering the TIGER files in order to produce a polygon-based census tract map is straightforward; however, producing a polygon-based map which treats water masses in a natural way requires a customized filtering process. Tulare County was simple to filter, but a region composed of a large river delta system, for example, would be much more complicated to successfully filter and assemble into complete polygons. On the other hand, if the various features and options of the *map* function are not needed, New S can still be used to directly display filtered line segments. This was the approach taken in Figure 6(a), where every line segment in the TIGER database was plotted directly using the New S function *segments*, and in Figures 7(b), 7(c), and 7(d) where filtered railroad features were superimposed onto existing *map* plots using the *segments* function.

## 7 Summary

The 1990 TIGER map files contain a large amount of cartographic and geopolitical data useful for mapping. The New S map function provides a means to make the complicated TIGER files manageable to plot and manipulate. The large quantity of often unwanted lines in the TIGER map database

and the requirement of complete polygons in a New S map database are the main hurdles in coupling the two systems. Census feature classification codes and census tract codes can be utilized to reduce the complexity of the TIGER data and yet maintain complete polygons. New S commands, other than the map function, can overlay additional, non-polygonal, geographic features of interest onto an existing map.

## 8 Acknowledgments

We extend our thanks to Richard Birdsong, U.S. Bureau of the Census, for his patience and willingness to share with us his expert knowledge of the TIGER database.

This work was supported by the Director, Office of Epidemiology and Health Surveillance; Office of Health; Office of Environment, Safety and Health; U.S. Department of Energy under Contract No. DE-AC03-76SF00098.

## References

1. Marx, R.W. (Ed.) *The Census Bureau's TIGER System*, Cartography and Geographic Information Systems 17, Vol. 1, 1990.
2. U.S. Bureau of the Census. *TIGER/Line Census Files, 1990 Technical Documentation*, Bureau of the Census, Washington, 1991.
3. Becker, R.A., Chambers, J.M., and Wilks, A.R. *The New S Language: A programming environment for data analysis and graphics*, Wadsworth & Brooks/Cole, Pacific Grove CA, 1988.
4. Becker, R.A. and Wilks, A.R. 'Maps in S', AT&T Bell Laboratories, Murray Hill NJ, 1991.
5. Merrill, D.W. 'Public Census Data on CD-ROM at Lawrence Berkeley Laboratory', Report LBL-32165 (Rev.1), July 1992.



## Appendix

### (a) ASCII to binary conversion

A New S map database consists of the three ASCII files described in the text: “name”, “polygon”, and “line”. The names of these files are conventionally suffixed as “.N”, “.gon”, and “.line”. For the Tulare County example, the names of these files would be “tulare.N”, “tulare.gon”, and “tulare.line”. The New S map function requires, however, that the “polygon” and “line” ASCII files be converted to binary form and suffixed, respectively, as “.G” and “.L”. The following two shell commands produce these two binary files:

```
% $SHOME/library/maps/Lmake 6 s b tulare.line tulare.L
% $SHOME/library/maps/Gmake b tulare.gon tularé.G tulare.L
```

In the first command line above, “6” specifies the level of map precision, “s” denotes spherical map units (radians), and “b” denotes binary output.

### (b) New S code change

As of June 1, 1992, the C source file *map.h* restricted the number of polylines per polygon to 256. The following simple change in *map.h* removes this constraint and will be required for most maps produced from the TIGER database. Change

```
typedef unsigned char Line; /*polylines in a region*/
to
typedef unsigned short Line; /*polylines in a region*/
```

Figure 1. Example map

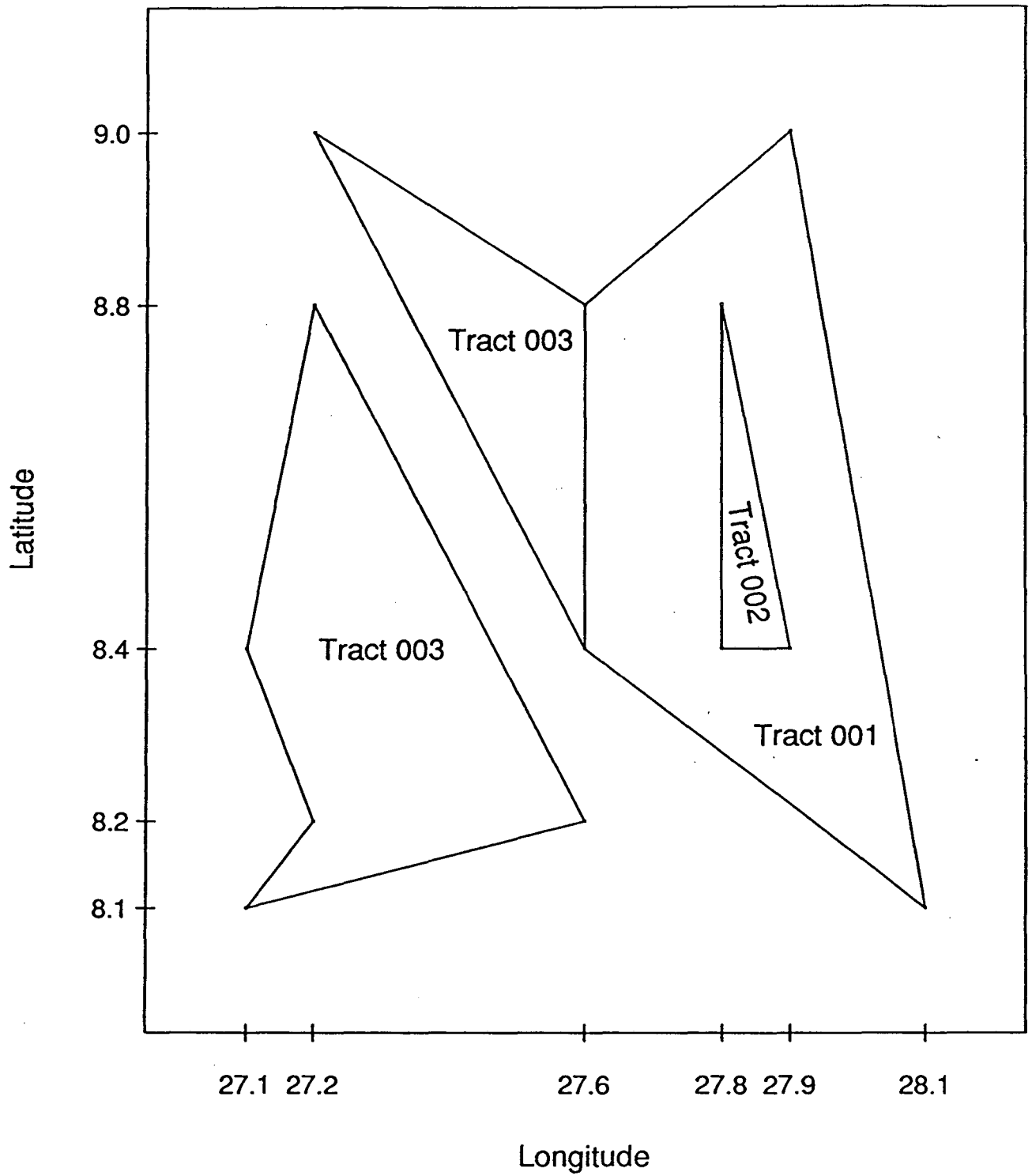


Figure 2. New S database for Example map (see Figure 1).

Name file:

-----

001:01	1
001:02	2
002:01	3
003:01	4
003:02	5

Polygon file:

-----

```
1 -3
EOR
2
EOR
-2
EOR
4 3
EOR
5
EOR
```

Line file:

-----

(Degrees for the polyline coordinates have been used to maintain the example's simplicity. A working New S database would use radians.)

```
0 1
27.6 8.8 27.9 9.0 28.1 8.1 27.6 8.4
EOR
3 1
27.8 8.8 27.8 8.4 27.9 8.4 27.8 8.8
EOR
1 4
27.6 8.8 27.6 8.4
EOR
0 4
27.6 8.4 27.2 9.0 27.6 8.8
EOR
0 5
27.1 8.1 27.2 8.2 27.1 8.4 27.2 8.8
27.6 8.2 27.1 8.1
EOR
```

Figure 3. TIGER database for Example map (see Figure 1).

(An implied decimal precision of two, rather than six,  
is used to maintain the example's simplicity.)

Basic Data TIGER file (simplified) of Example map.

record ID	feature name	cfcc	left/right tract ID	"from" long lat	"to" long lat
03	TRACT BORDER	F10	000 001	+ 2760+ 880	+ 2760+ 840
05	VENICE CANAL	H20	001 002	+ 2780+ 880	+ 2780+ 880
04	KING BEACH	H01	000 003	+ 2760+ 840	+ 2760+ 880
06	KING ISLE	H01	000 003	+ 2710+ 810	+ 2710+ 810
07	BAKER ROAD	A30	003 001	+ 2760+ 840	+ 2760+ 880

Shape Coordinate Points file (simplified) of Example map.

record ID	long lat	long lat	long lat
03	+ 2790+ 900	+ 2810+ 810	+00000+0000
05	+ 2790+ 840	+ 2780+ 840	+00000+0000
04	+ 2720+ 900	+00000+0000	+00000+0000
06	+ 2720+ 820	+ 2710+ 840	+ 2720+ 880
06	+ 2760+ 820	+00000+0000	+00000+0000

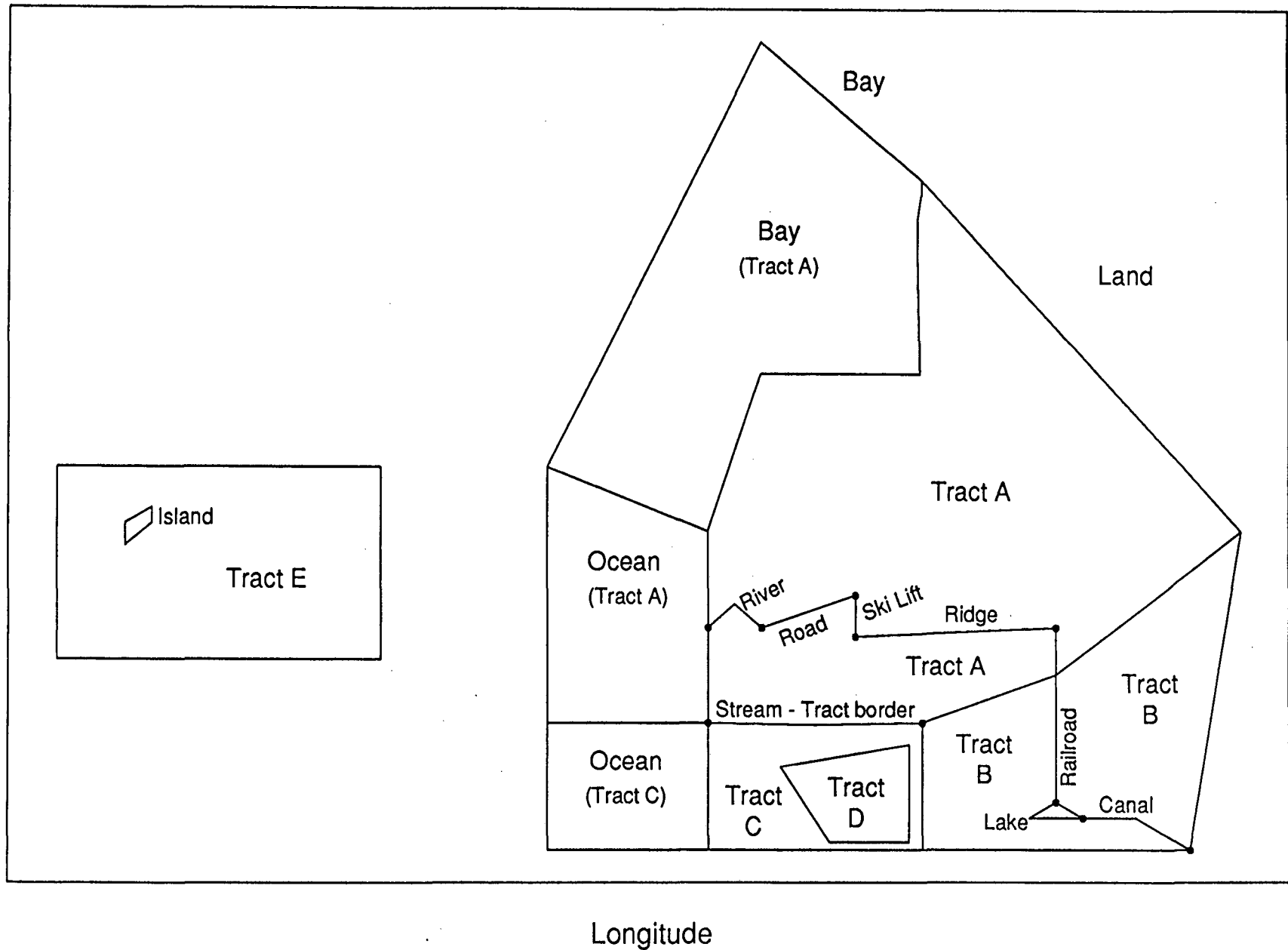


Figure 4. Unfiltered demonstration map.

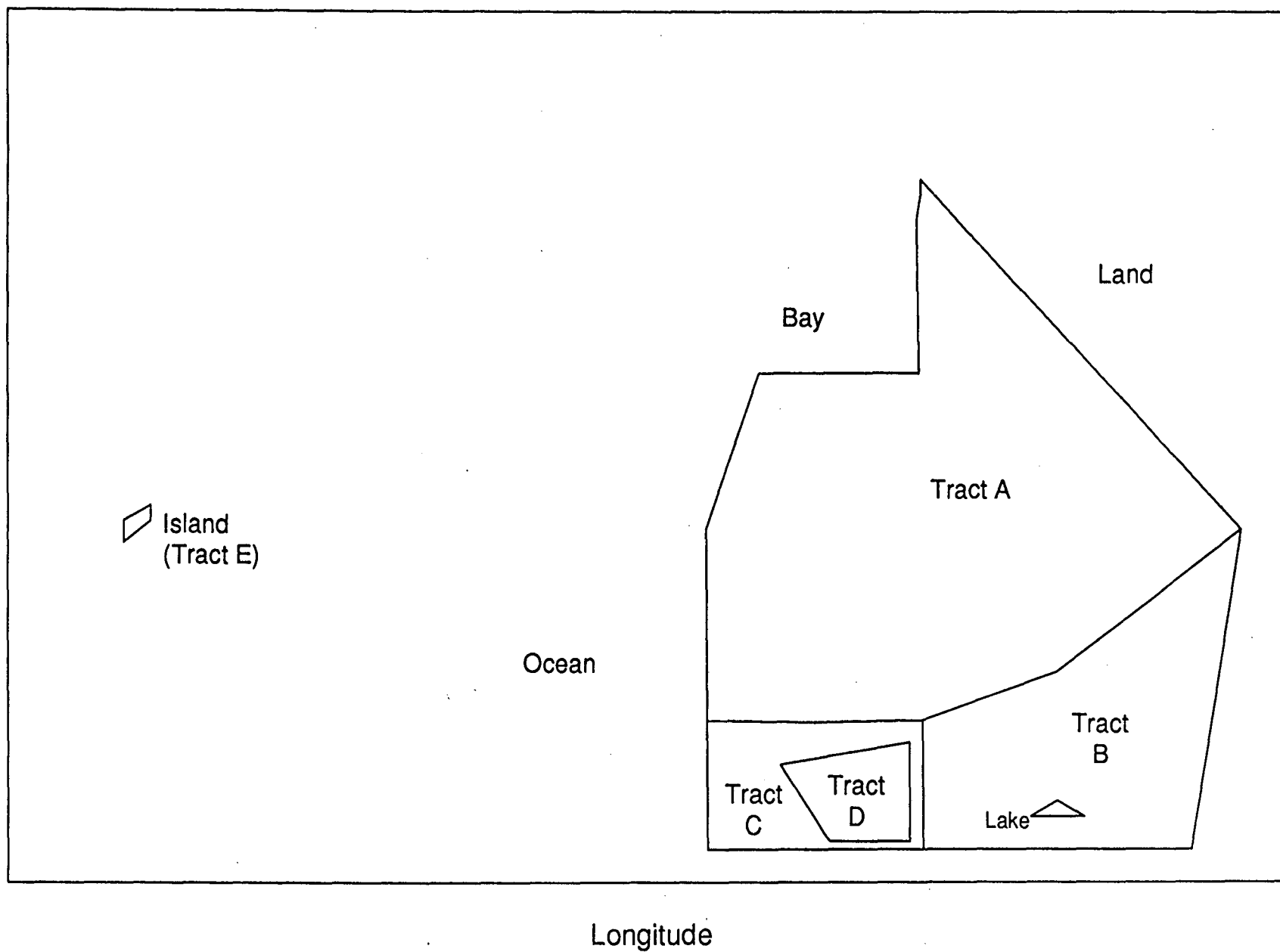
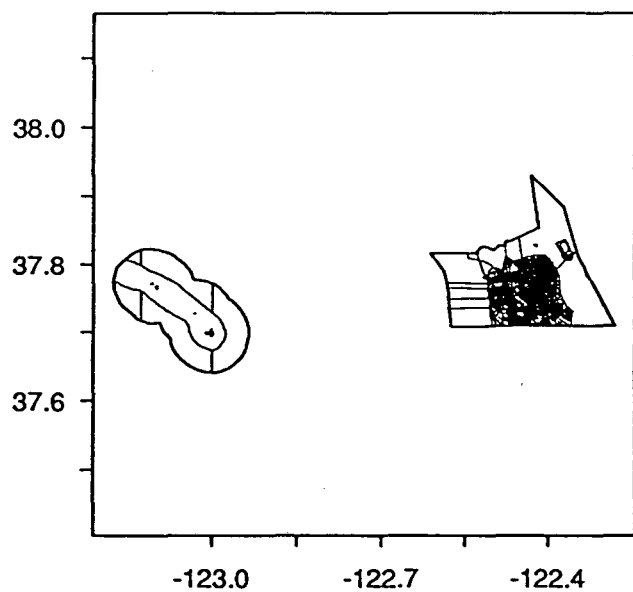
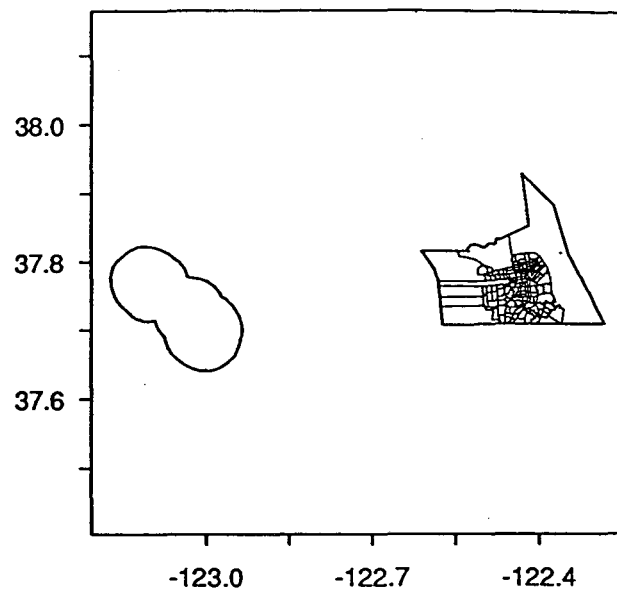


Figure 5. Filtered demonstration map.

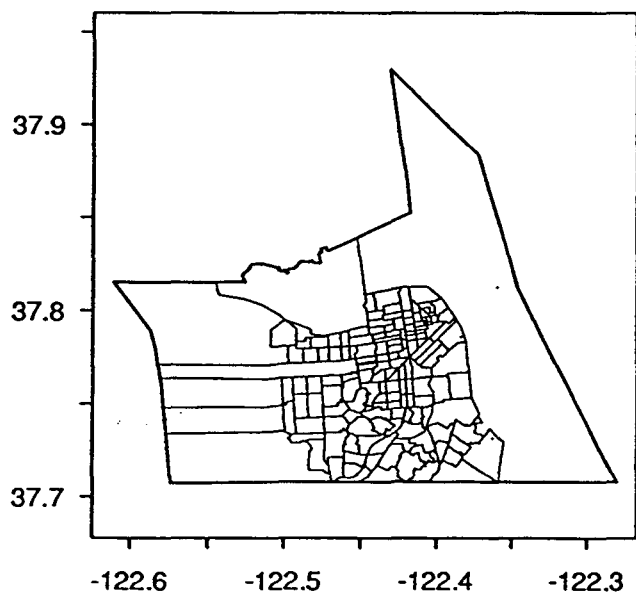
Figure 6. San Francisco County, California



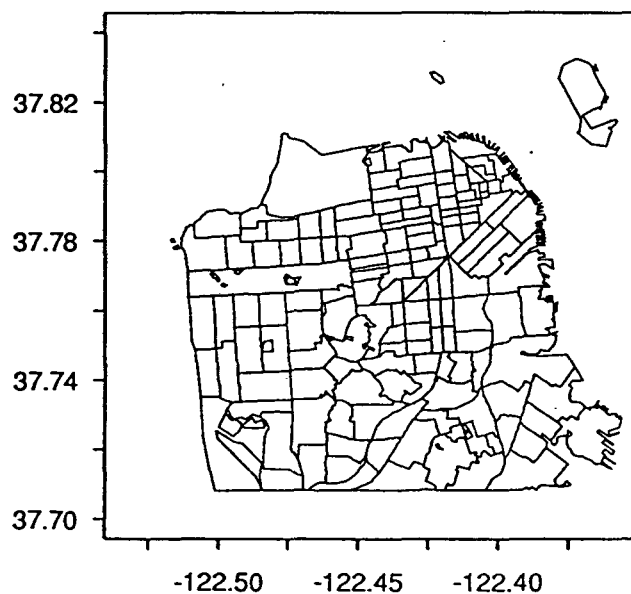
(a) All TIGER features



(b) Census tract boundaries only

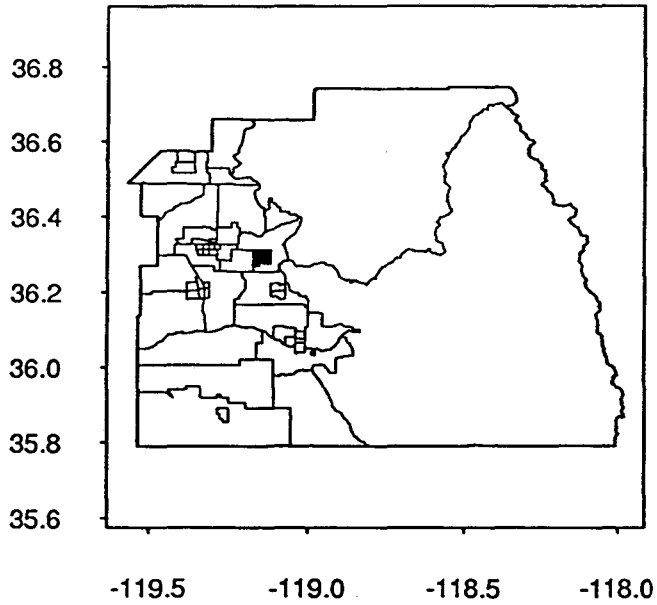


(c) San Francisco City region

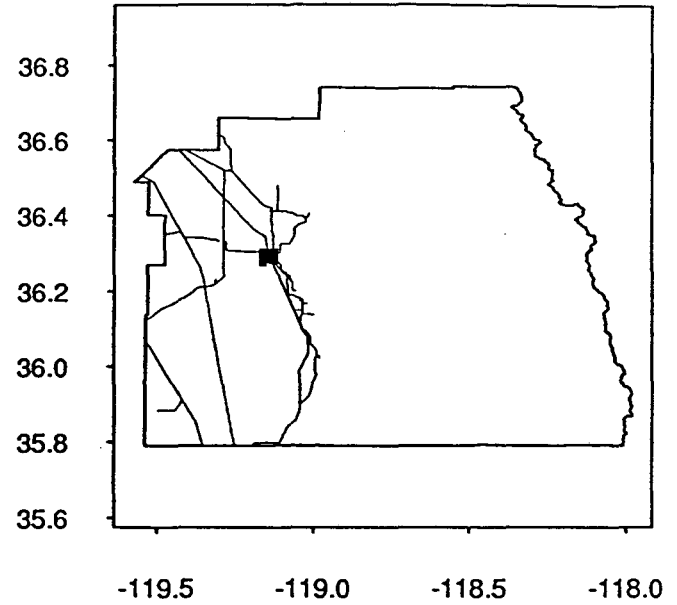


(d) With shoreline boundaries

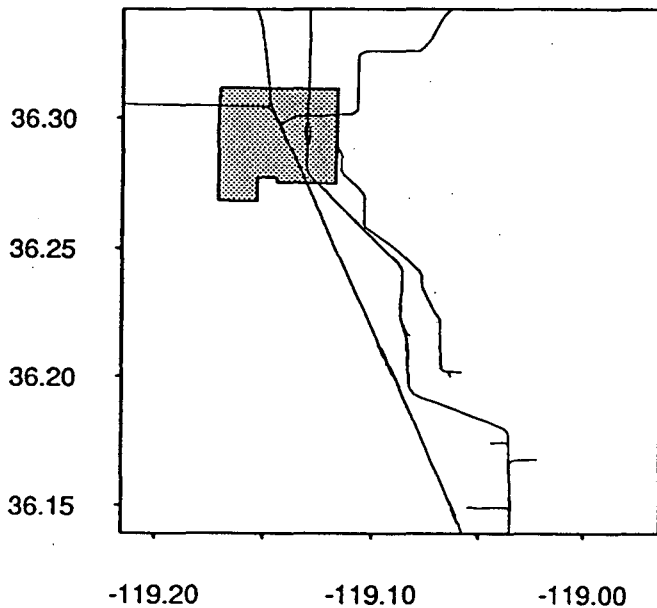
Figure 7. Tulare County, California



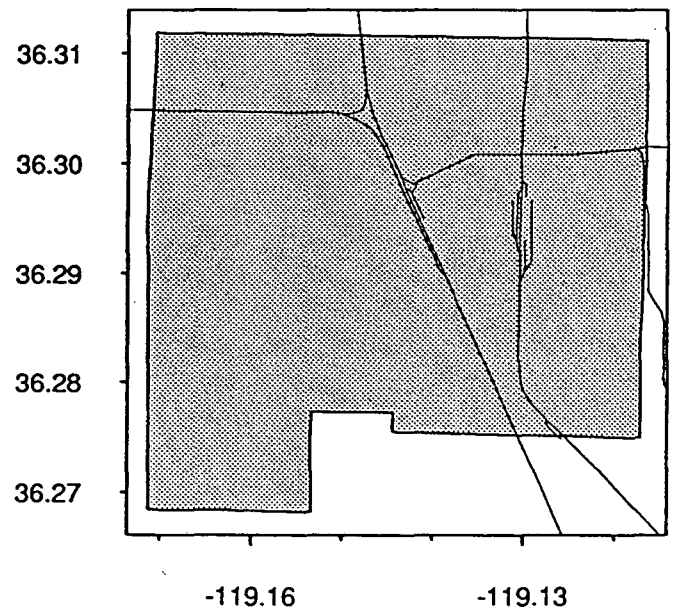
(a) Census tract polygons  
(tract #15 shaded)



(b) Railroad features  
(tract #15 shaded)



(c) Railroad features near  
tract #15



(d) Close-up of tract #15  
and railroad features



LAWRENCE BERKELEY LABORATORY  
UNIVERSITY OF CALIFORNIA  
TECHNICAL INFORMATION DEPARTMENT  
BERKELEY, CALIFORNIA 94720