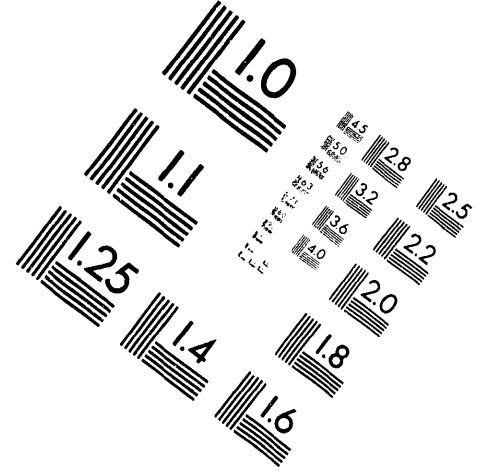
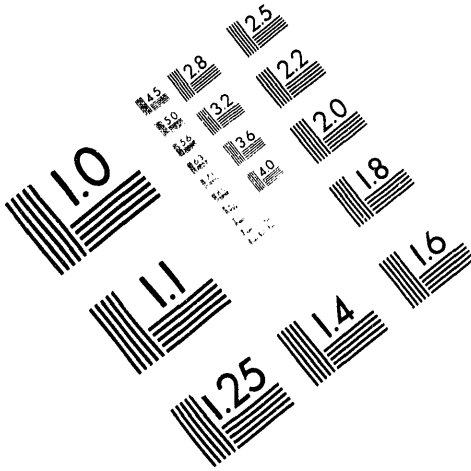




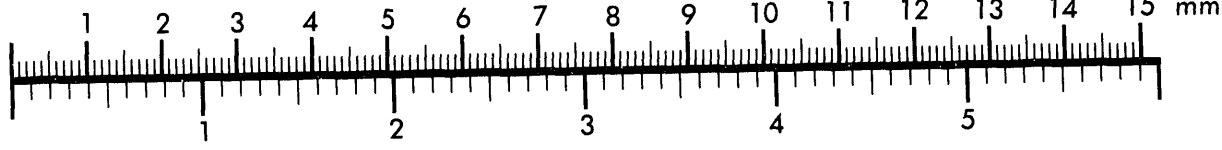
AIM

Association for Information and Image Management

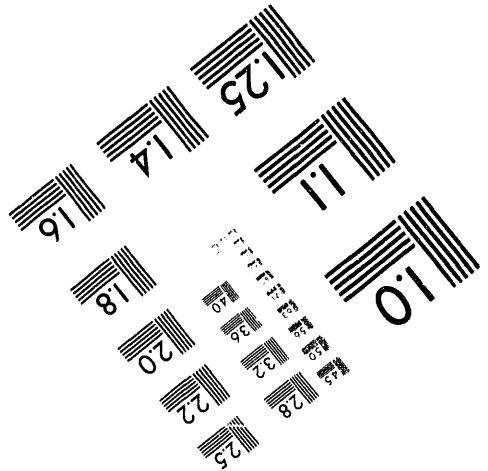
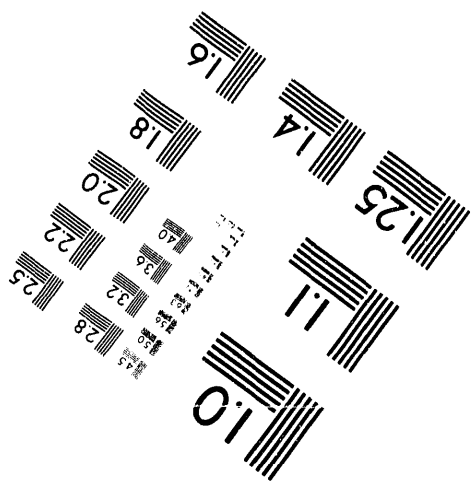
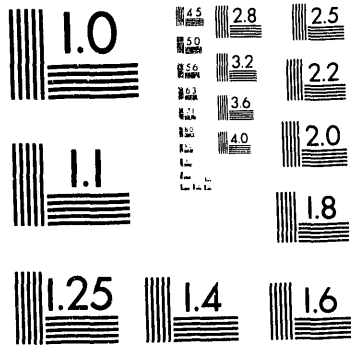
1100 Wayne Avenue, Suite 1100
Silver Spring, Maryland 20910
301/587-8202



Centimeter



Inches



MANUFACTURED TO AIM STANDARDS
BY APPLIED IMAGE, INC.

1 of 1

Interfacing U.S. Census Map Files with Statistical Graphics Software: Application and Use in Epidemiology

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

Information & Computing Sciences Division
Lawrence Berkeley Laboratory
University of California
Berkeley, CA 94720 USA

† School of Public Health
University of California, Berkeley
Berkeley, CA 94720 USA

March 1993

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.

MASTER *ds*
DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

Interfacing U.S. Census Map Files with Statistical Graphics Software: Application and Use in Epidemiology

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, accessible through purchase or Federal repository libraries, contain 24 billion characters of data describing various geographic features including coastlines, hydrography, transportation networks, political boundaries, etc. covering the entire United States.^{1,2} Many of these physical features are of potential interest in epidemiological case studies. Unfortunately, the TIGER database only provides raw alphanumeric data; no utility software, graphical or otherwise, is included.

Recently, the S statistical software package has been extended to include a map display function.^{3,4} The map function augments S's high-level approach toward statistical analysis and graphical display of data. Coupling this statistical software with the map database developed for U. S. Census data collection will facilitate epidemiological research.

We discuss the technical background necessary to utilize the TIGER database for mapping with S. Two types of S maps, segment-based and polygon-based, are discussed along with methods to construct them from TIGER data. Polygon-based maps are useful for displaying regional statistical data; e.g., disease rates or incidence at the census tract level. Segment-based maps are easier to assemble and appropriate if the data are not regionalized.

Census tract data of AIDS incidence in San Francisco (CA) and lung cancer case locations relative to petrochemical refinery sites in Contra Costa County (CA) are used to illustrate the methods and potential uses of interfacing the TIGER database with S.

1 Introduction

Advances in statistical software packages have made data analysis more readily available and less laborious for statisticians and other users. While it has become standard practice to provide numerical data in computer readable formats, that has not been the case for map data. In particular, the lack of digitized map databases, and of software to utilize such databases effectively, have placed substantial restraints on analytic techniques and display in spatial epidemiologic work.⁵ Linking geographic databases with statistical software packages encourages the development of more innovative methods for the analysis and exploration of epidemiologic spatial data.

In 1990, the United States Bureau of the Census released the 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 and territory of the United States.¹ The TIGER database is attractive because of its extensive detail and its direct correspondence with the 1990 U. S. census. It is publicly available through the Bureau of the Census on tape or cd-rom or may be accessed on-line from certain Federal repository libraries. TIGER 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.¹ Unfortunately, the TIGER database only provides the maps in alphanumeric form; no utility software, graphical or otherwise, is included.

On the other hand *S*, a popular statistical software package, has easy to use functions which facilitate advanced graphics in a data analysis setting.³ *S* has the ability to plot contours, lines, segments, and points, as well as the usual compliment of statistical graphs. Of special interest is the *S* function *map*⁴ which allows polygons to be quickly selected, plotted, shaded, etc. Complex statistical modelling of data can also be done in the *S* environment. Therefore, linking the TIGER map files with *S* would facilitate spatial epidemiologic research.

This paper reports on the interfacing of TIGER map files with *S*. Particular attention is devoted to the nuts and bolts of piecing together TIGER census tracts into a polygon format suitable for use with the *S* function *map*. While census tract maps and the TIGER database are both defined as a collection of closed polygons, TIGER polygons are defined at a much finer level; occasionally as small as a few square meters. Therefore, conversion of the TIGER "line" format into *S* "polygon/polyline" format is one facet of

the work reported here. However, if shading or other polygon-based displays are unnecessary for an analysis, the task of getting TIGER geographic data into a form usable by *S* is greatly simplified and we cover this as well. In either case, an understanding of the TIGER feature classification system will be required. We discuss this classification system and how it can be used to obtain subsets of the TIGER database for the map features of interest.

Sections 2 through 5 of this paper provide the technical support for the above topics. Two example applications are then provided in section 6. In the first example, a census tract polygon map of San Francisco (CA) was prepared and the census tracts shaded according to tract-level AIDS incidence data in the city. In the second example, a segment-based map of Contra Costa County (CA) was assembled and the locations of petrochemical refineries and lung cancer cases shown. Refinery related exposure was modelled and contoured over the map.

2 Geopolitical Maps and *S*

A state-level and a county-level map of the continental U. S. are included as part of the *S* software package. Although lacking the detail and accuracy of TIGER, these databases are ideal for quick sketches of state or county boundaries. Because they are in *S* polygon format, the *S* function *map* can be used to obtain subsets of these databases, shade prescribed regions, extract surrounding boundaries, and so forth. Other *S* display functions such as *segments*, *points*, *contour*, and *text*, can be used to overlay on these maps user-specific geographic information.

Map databases such as TIGER can be imported to *S* in one of two ways depending on their intended use. If regions are to be shaded according to statistical data or if polygon-based statistical moment calculations will be required, then the map will first have to be put into the *S* polygon format. If not, the map can simply be input as a collection of segments, and functions other than *map* used to display the features. Readers interested only in the latter approach can skip to the next section where the structure of the TIGER map files is discussed. For the rest of this section, we discuss in detail the *S* polygon format with our ultimate objective being the construction of TIGER census tract maps.

In the *S* polygon format, 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 usually 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 polygon-based map in S is defined by a set of three files based on the general structure above. Briefly, the first file contains the name of each polygon and its respective polygon identification (ID) code. 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 is 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 (shaded) 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.) The name file for the Figure 1 map is included in Figure 2.

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 traced in a clockwise fashion by polyline 1 followed by a reversely directed polyline 3. Polygon 2 is traced entirely by polyline 2. Polygon 3 is also entirely traced by polyline 2 but in the reverse direction.

The third file, termed the "*line*" file, consists of the polyline records; i.e., 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; i.e., 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.

The "polygon" and "name" files must be converted into indexed binary files before they can be accessed by the *S* function *map*. *S* provides directly for this ASCII to binary conversion. Details of the necessary steps are given in the Appendix.

3 TIGER Map Files

TIGER maps, like *S* polygon maps, are composed of chains, or "*lines*" in TIGER terminology. These lines are functionally equivalent to the polylines of *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 a few square meters 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 attributes 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 Figure 1 map.

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 class code (cfcc), 3) the state, county, tract, and block 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 attributes, including zip codes and address ranges, might also be of epidemiologic interest in obtaining subsets of the TIGER database.

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 the presence of its ID code in the Shape Coordinate Points File. (For example, in the Figure 3 map files, 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, continuing through the sequential shape points, and ending 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 Figure 1 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 Record 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 Record File. For the example map, record ID 06 precedes record ID 05 in both TIGER files (Figure 3).

4 Census Feature Classification System

The *census feature class code* (cfcc) describes the physical feature of a TIGER 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 that could be described by more than one cfcc.

The hierarchical coding system of the cfcc, however, can be used 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 subclass; for example jeep trails would start with "A5" and interstate highways with "A1". The second integer allows the user to discriminate within a subclass. Computer code is easily written to select lines that are of a specific feature class.

Caution must be exercised when using the *cfcc* to extract desired portions of the TIGER database. Only a single *cfcc* accompanies each TIGER line and thus the Census Bureau uses a "priority ranking system" in declaring the *cfcc*. For example, a perennial shoreline label, "H01", has priority over a property line designation, "F40". Unfortunately, no formal documentation has yet been published by the Census Bureau on the priority structure they use for feature classification.⁶

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 line may be described by one of nine class codes specific to interstate highways (e.g. divided, undivided, tunnel, underpass, etc). In these instances it is advantageous to use the more general class level when specifying the *cfcc* selection criteria.

In any particular context, simply plotting out all lines having one of the *cfcc*'s of interest often provides the clues necessary for a more systematic selection of desired TIGER lines. For further details regarding the census feature classification system and codes we refer the reader to the TIGER technical documentation⁷.

5 Filtering and Reformatting TIGER Files for *S*

TIGER file data can be displayed by *S* as a collection of polygons or as a set of many simple segments. In either case, unwanted TIGER lines must be identified and filtered out. Use of the *S map* function further requires that the selected lines be assembled into complete polygons.

Filtering is accomplished by scanning the Basic Data Record File and keeping only those records that match a set of user-specific criteria involving *cfcc*'s, feature ID's, and left-right state, county, tract, block ID's. If a basic data record meets the selection criteria, the Shape Coordinate Points File can then be referred to for the intermediate points shaping the line.

For example, a tract-level map may be constructed by selecting only those lines for which the tract code on the left and right sides of the line differs. However, if a census tract is next to water, its boundary line often extends out for some distance over the water instead of ending at the shore-

line. These extensions, while delineating official tract boundaries, are rarely of practical interest in epidemiological investigations. Fortunately, a block code ending in "99" signifies a water feature. Therefore, such extensions are identifiable as lines having left *and* right block codes ending with "99". Perennial shoreline can be identified by lines with a cfc of "H01".

For a segment-based map in *S*, simply reformat the selected lines as a set of from-to coordinates and use the *S* function *segments* to display them. Without the requirement of closed polygons, the ordering and direction of the segments is unimportant. Other features and data can then be superimposed with the *S* functions *points*, *contour*, *text*, etc.

Polygon-based maps require the selected lines to be linked into complete polygons. For tract-level maps, lines with a common census tract code are pooled and clockwise linked end-to-end to form a polygon(s). After all polygons have been identified, the map needs to be put into the database format required by *S*. The census codes used to group the lines for polygon formation can be used as names in the *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 *S* "polygon" file. If shape points exist for the line, they are simply inserted between the two end-points in the *S* "line" file. External boundary lines need be given an ID code of 0. *S* assumes that the "line" file coordinates are in radians but plots the map in degrees.

6 Example Applications

The California counties of San Francisco and Contra Costa were chosen to illustrate the methods and potential uses of interfacing TIGER with *S*. In the San Francisco example, we chronicle the filtering process used and form an *S* polygon map. With Contra Costa, we present a segment-based map and overlay additional feature information and statistical quantities.

For San Francisco, our goal was to produce a tract-level *S* polygon map so that tract-level data on AIDS incidence could be displayed. We began by using the *segments* function to plot all line segments defined in the 1990 San Francisco County TIGER database (Figure 4). The "bubbles" seen in the west of Figure 4 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 4, they are obscured by the deluge of other TIGER features. We then filtered through the TIGER database to retain only those

lines that were census tract boundaries – lines having different tract codes on their left and right side. 2,830 of 16,885 Basic Data Record File records and 150 of 1,335 Shape Coordinate Points File records define the census-tract boundaries. The census tract boundary lines were then assembled into polygons, formatted as an *S* map database, and displayed using the *map* function (Figure 5). Notice that the state territorial water boundaries give a distorted view of where people actually live and work in San Francisco.

To generate a map representative of the city proper, we restarted the process with a stricter set of filtering conditions. All TIGER lines inside the eastern-most census tract (017902) were not retained, while its census tract borders were kept only if they were shoreline (cfcc="H01"); that is, if the tract border was shared with tracts of the city proper. Next, boundary extensions over water were removed by ignoring all lines that had block ID's ending in "99" on both the right and left sides. Finally, lines that successfully met the above criteria were retained only if they were shoreline or census tract boundary. Shorelines that were not tract borders were denoted as exterior boundary on the side with the "99" block code.

For counties with complicated geopolitical borders, several attempts at filtering will likely be required. In this example, lines representing Marin County shoreline, on the north side of the Golden Gate, satisfied the above conditions but were not desired. Thus, we imposed an additional condition to ignore TIGER lines that were tract borders of the north-west census tract (060100) and, in this case, had the feature name "San Francisco Bay". The final step was to retain a specific TIGER Basic Data Record File record (ID 192281768) which represented where the Bay Bridge entered San Francisco from the east. Rather than having a shoreline cfcc, it was designated as Interstate Highway. The resulting map of San Francisco, along with its internal lakes and the Farallon Islands then formed a set of complete polygons which were put into *S* polygon format. To remove the various islands and internal lakes from the *S* database, we simply deleted their corresponding records in the *S* "names" file. Needless to say, forming polygon maps for inland counties tends to be easier than for shoreline counties.

The prepared San Francisco census tract map was then displayed using the *S* function *map* (Figure 6) and census tracts were shaded according to the incidence of AIDS during the years 1987-89 among gay and bisexual men age 18-29.⁸ Higher counts are apparent in the central region of the city, most noticeably near the Castro District. The visual impacts of different grouping and shading levels on the plot can quickly be assessed once the *S* database is in place. With the 1990 census data, demographic cofactors

might also be utilized in a graphical analysis and presentation.

The association between air emissions at petrochemical refineries and cancer incidence in Contra Costa County has been the subject of epidemiological investigation⁹. A segment-based outline map of Contra Costa County is presented in Figure 7 showing the locations of 5 petrochemical refineries and 23 cases of lung cancer. Cases were defined as white female residents in the age category of 30-49 who had adenocarcinoma lung cancer between the years 1976 and 1980. The mean distance from each case to the nearest refinery was 8 kilometers with a standard error of 7 kilometers. The case data was obtained from the San Francisco Surveillance, Epidemiology and End Results (SEER) cancer registry which provided the number of cases in each census tract and the tract's geographic centroid.

A simple model of exposure for these data would assume that exposure is inversely proportional to nearest refinery distance. To illustrate this model, nearest refinery distance was calculated for a 100 x 100 grid of regularly spaced locations, the reciprocals taken, and the *contour* function used to display lines of constant exposure (Figure 7). The difference in exposure between neighboring contours is constant, and, the inner- and outer-most contours correspond to refinery distances of 2 and 30 kilometers, respectively.

A more realistic model of exposure for these data would be based on direct measurements of refinery emissions, prevailing wind patterns, and other environmental data. While these types of measurements would not likely be collected at regularly spaced intervals, they could easily be interpolated and then contoured. A complete analysis, of course, would account for population density and other demographic factors.

Our intention with these examples has not been to arrive at definitive epidemiological conclusions but rather to illustrate the kinds of spatial data display that are possible in the *S* environment. Be it regional incidence or locations of cases relative to contours of exposure, spatial data displays often express interpretable patterns and stimulate ideas for further investigation.

7 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, Vol. 17, No. 1, 1990.
2. Merrill, D.W. 'Public Census Data on CD-ROM at Lawrence Berkeley Laboratory', Report LBL-32165 (Rev.3), January 1993.
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. Besag, J., and Newell, J. 'The Detection of Clusters in Rare Diseases', *Journal of the Royal Statistical Society, A.*, **154**, 143-155 (1991).
6. Birdsong, Richard. personal communication, Bureau of the Census, Geography Division, (March 2, 1993).
7. U.S. Bureau of the Census. *TIGER/Line Census Files, 1990 Technical Documentation*, Bureau of the Census, Washington, 1991.
8. San Francisco AIDS Office, San Francisco Department of Public Health.
9. Selvin, S., Shaw, G., Schulman, J., and Merrill, D.W. 'Spatial Distribution of Disease: Three Cases Studies', *Journal of the National Cancer Institute*, **79**, 417-423 (1987).

Appendix

An *S* polygon map is defined by 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 San Francisco example, the names of these files would be, say, “sf.N”, “sf.gon”, and “sf.line”. The *S* function *map* 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 sf.line sf.L
% $SHOME/library/maps/Gmake b sf.gon sf.G sf.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.

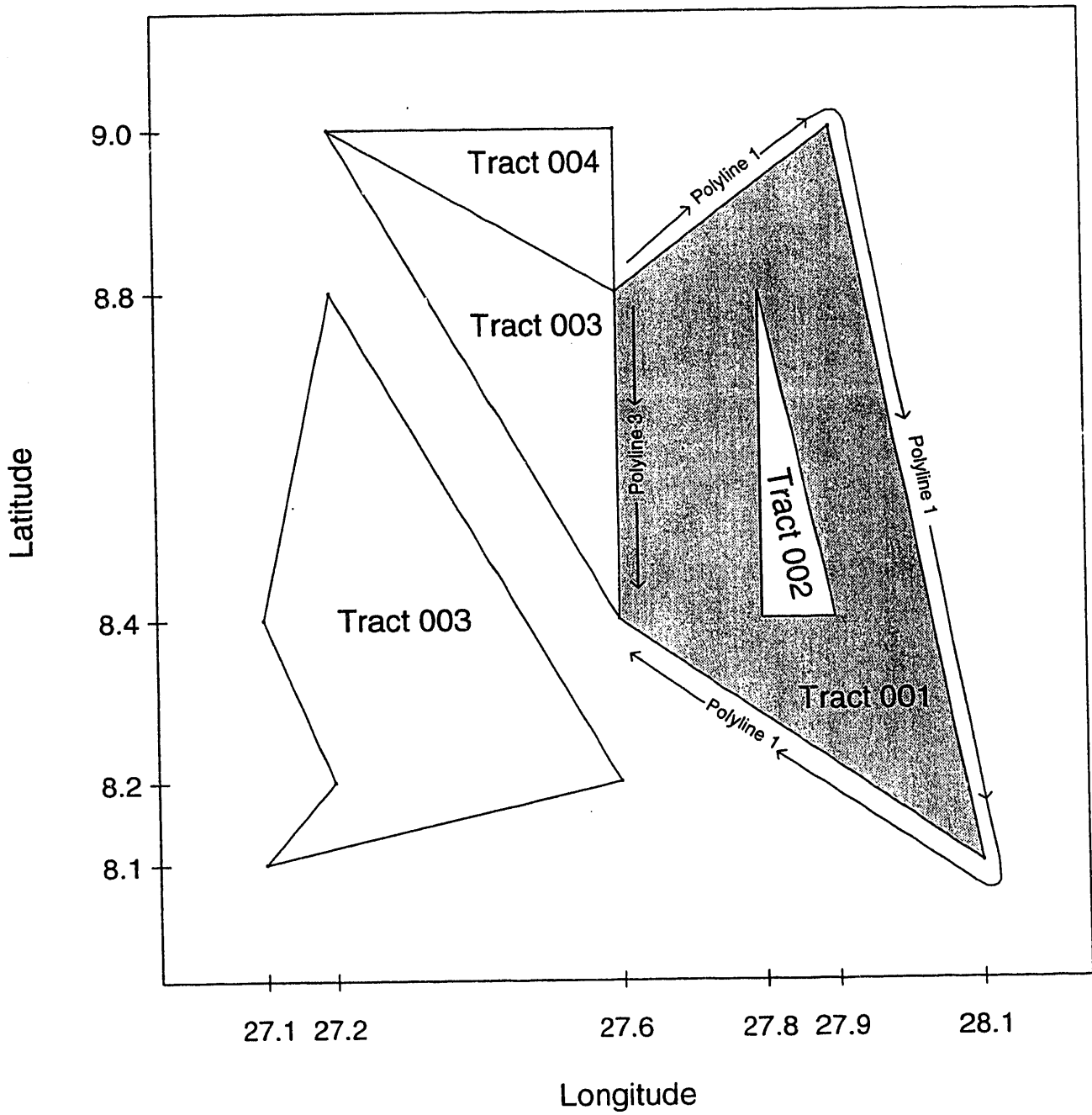


Figure 1. Map used to illustrate *S* polygon and TIGER file structures (see Figures 2 and 3).

Name file:

001:01	1	# polygon 1
001:02	2	# polygon 2
002:01	3	# polygon 3
003:01	4	# polygon 4
003:02	5	# polygon 5
004:01	6	# polygon 6

Polygon file:

1 -3	# polygon 1
EOR	
2	# polygon 2
EOR	
-2	# polygon 3
EOR	
4 6 3	# polygon 4
EOR	
5	# polygon 5
EOR	
-6 7	# polygon 6
EOR	

Line file:

0 1					# polyline 1
27.6 8.8	27.9 9.0	28.1 8.1	27.6 8.4		
EOR					# polyline 2
3 1					
27.8 8.8	27.8 8.4	27.9 8.4	27.8 8.8		
EOR					# polyline 3
1 4					
27.6 8.8	27.6 8.4				
EOR					# polyline 4
0 4					
27.6 8.4	27.2 9.0				
EOR					# polyline 5
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					# polyline 6
6 4					
27.2 9.0	27.6 8.8				
EOR					# polyline 7
0 6					
27.2 9.0	27.6 9.0	27.6 8.8			
EOR					

Figure 2. S database for the map shown in Figure 1. Here, coordinates are expressed in degrees (a working S database uses radians).

Basic Data TIGER file (simplified) of Example map.

```
=====
```

record	feature		left/right	"from"	"to"
id	name	cfcc	tract id	long lat	long lat
02	TRACT BORDER	F10	000 004	+ 2720+ 900	+ 2760+ 880
03	TRACT BORDER	F10	000 001	+ 2760+ 880	+ 2760+ 840
04	KING BEACH	H01	000 003	+ 2760+ 840	+ 2720+ 900
06	KING ISLE	H01	000 003	+ 2710+ 810	+ 2710+ 810
05	VENICE CANAL	H20	001 002	+ 2780+ 880	+ 2780+ 880
07	BAKER ROAD	A30	003 001	+ 2760+ 840	+ 2760+ 880
08	BAKER ROAD	A30	004 003	+ 2720+ 900	+ 2760+ 880

Shape Coordinate Points file (simplified) of Example map.

```
=====
```

record	id	long	lat	long	lat	long	lat
02		+ 2760+	900	+00000+	0000	+00000+	0000
03		+ 2790+	900	+ 2810+	810	+00000+	0000
06		+ 2720+	820	+ 2710+	840	+ 2720+	880
06		+ 2760+	820	+00000+	0000	+00000+	0000
05		+ 2790+	840	+ 2780+	840	+00000+	0000

Figure 3. TIGER file structure for the map shown in Figure 1. Here, coordinates have an implied decimal precision of 2 (actual TIGER files have precision 6).

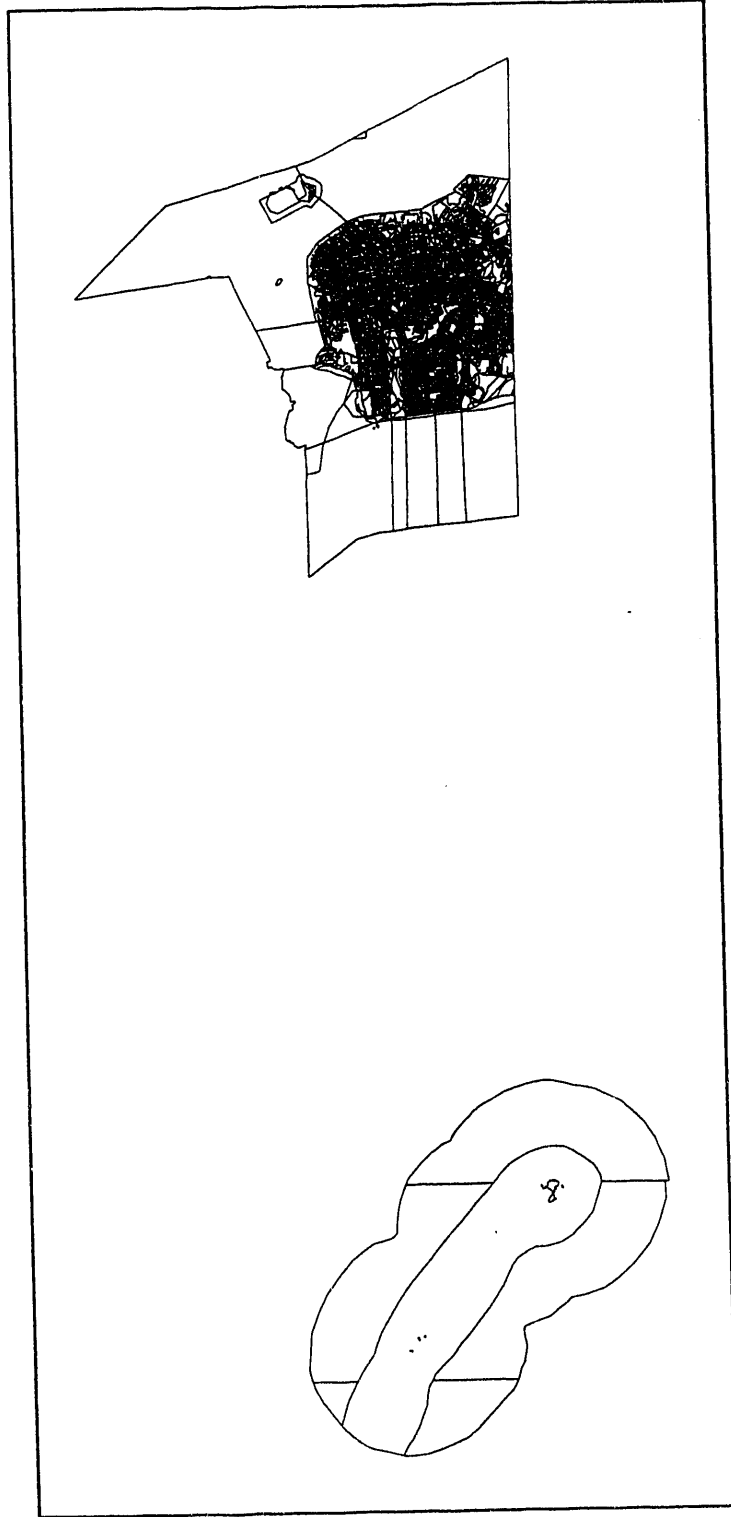


Figure 4. Unfiltered 1990 TIGER map of San Francisco County (CA).

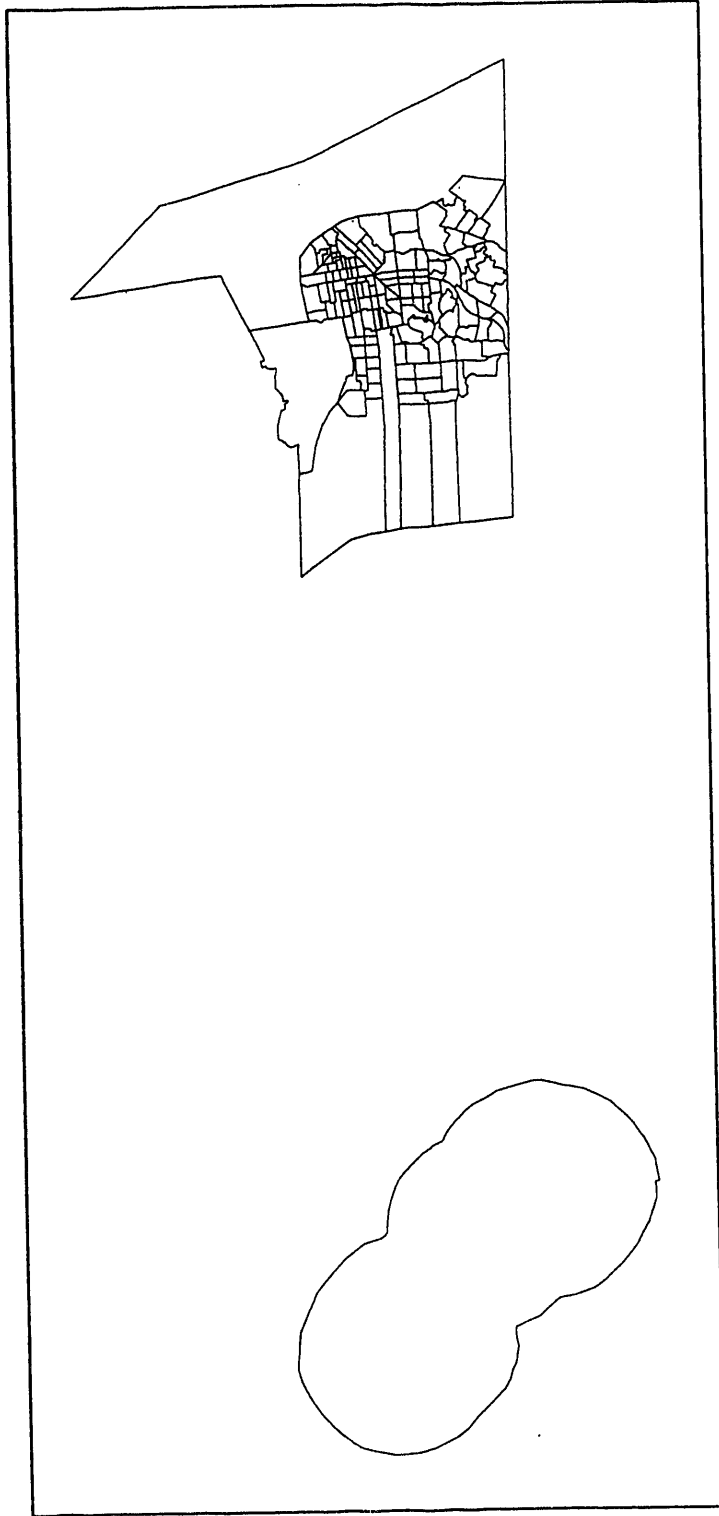


Figure 5. 1990 census tract boundaries of San Francisco County (CA).

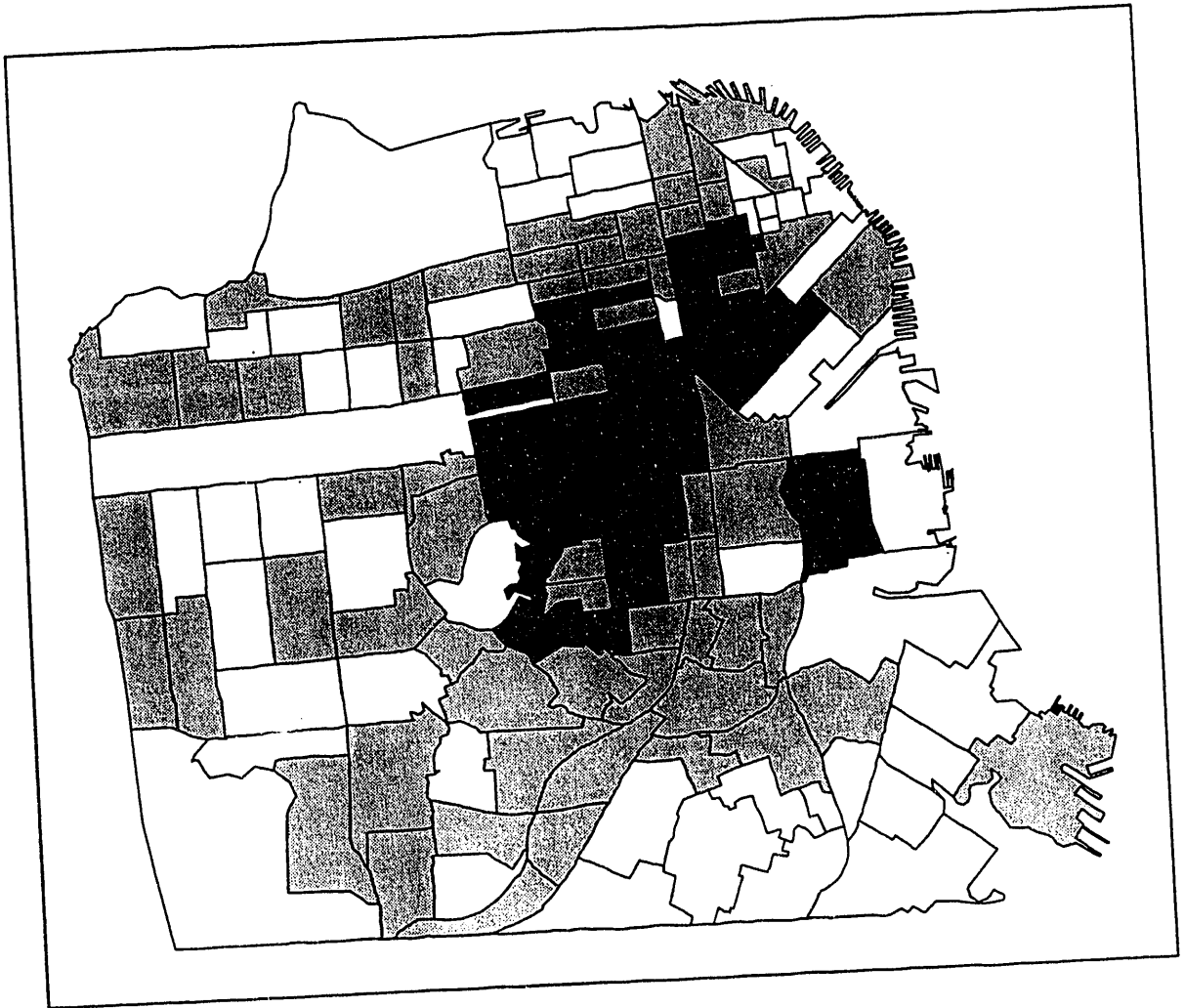


Figure 6. San Francisco (CA) AIDS Incidence, 1987-89, Gay/Bisexual Men, Age 18-29, by census tract. Shading levels: white = 0 cases; light-gray = 1-4 cases; dark-gray = 5-9 cases; black > 9 cases.

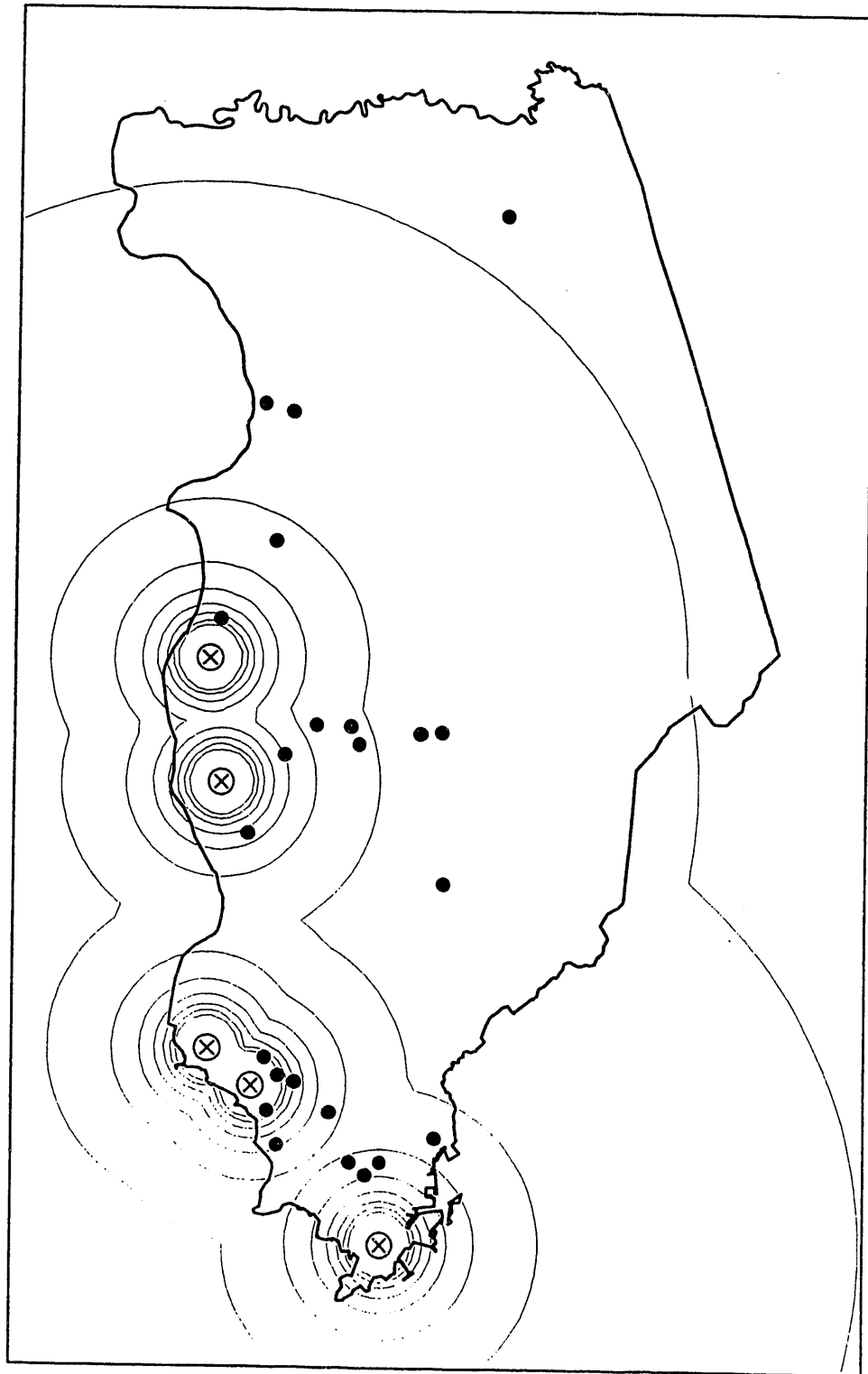


Figure 7. Contra Costa County (CA) with locations of petrochemical refineries (⊗) and cases (●) of adenocarcinoma lung cancer in white females age 30-49 (1976-1980). Contours represent lines of equal exposure assuming exposure is inversely proportional to nearest plant distance.

**DATE
FILMED**

7 / 27 / 93

END

