-

A Tutorial for "TIGER in GRASS" with the Database Manager RIM

Jacinda Mountz Barbehenn

U.S. Army Corps of Engineers
Construction Engineering Research Laboratory
Environmental Division
Spatial Analysis Systems Team
P.O. Box 9005
Champaign, IL 61826-9005

January, 1993

Table of Contents

Table of Contents 2
1. Introduction 3
2. An Introduction to TIGER/Line 4
3. TIGER in GRASS 5
4. Moving RIM Databases and TIGER Vector Files 13
5. Conclusions and Recommendations 15
Sources, Resources, and References 16
Appendix A: Census Feature Class Codes (CFCC) 17
Appendix B: TIGER Database Screenlayout Format with Field Names and Descriptions 23
Appendix C: Summary of TIGER/Line Record Types 25

1. INTRODUCTION

1.1. Background

This document details how U.S. Bureau of the Census Topologically Integrated Geographic Encoding and Referencing File (TIGER)/Line data can be used within the Geographic Resources Analysis Support System (GRASS). GRASS is a public domain, image processing and geographic information system (GIS) originally developed by researchers in the Environmental Division of the U.S. Army Corps of Engineers Construction Engineering Research Laboratory (USACERL-EN) in Champaign, IL. The system is used to input, manipulate, analyze, and output geographic data by users in both military and non-military, public and private agencies, based in North America, Europe, and other parts of the world.

Although most GRASS development has been conducted at USACERL, system integration, development, testing, distribution, training, and support is performed by numerous publicly and privately operated sites, throughout the world. In particular, the GRASS programs for reading and importing TIGER/Line data (*v.db.rim, Gen.tractmap, tig.rim.sh*) were created by Dr. James Hinthorne and David Satnik of the GIS Laboratory at Central Washington University.

This release of GRASS implemented significant additions and modifications to system libraries and programming code. Mechanisms are needed to transfer current technology and information about GRASS to user and development sites. This document reflects modifications made to GRASS and guides the user through the use of GRASS programs to build GRASS data files from Census Bureau TIGER/Line data.

1.2. Objective

The objective of this work is to transfer knowledge of GRASS TIGER/Line import, manipulation, and conversion functions to the field. This document guides the user through the process of using GRASS programs and the RIM public domain database management system to which GRASS is interfaced to import, manipulate, and convert U.S. Census Bureau TIGER/Line data in GRASS.

1.3. Approach

Current GRASS census data import, manipulation, and conversion functions are examined in this report.

1.4. Scope

This document discusses GRASS programs which use the relational database manager RIM to import and manipulate TIGER/Line data in GRASS. In addition to creating map layers, this suite of programs builds and uses a RIM database of the associated TIGER/Line data. The building of the RIM, depending on the area the user is dealing with, can be very time consuming and, once built, take up a large amount of memory. For a few kinds of operations the suite of "TIGER in GRASS" programs described in this tutorial may be preferred, but for other operations, the suite of programs described in the tutorial "Importing TIGER/Line Maps and STF1 Data in GRASS" might be preferred.

Along with describing how to import TIGER/Line data in GRASS without using a database manager, this second tutorial ("Importing TIGER/Line Maps and STF1 Data in GRASS") describes the importation and use of TIGER/Line data to GRASS, both tutorials should be reviewed so that the user can decide the method that best fits her/his needs.

1.5. Mode of Technology Transfer

The information in this report will aid the technical transfer of USACERL's GRASS image processing and geographic information system. GRASS is being transferred to the field through the following mechanisms: training programs, hands-on experience, a user document center, newsletters, extensive documentation, institutional structures at the Army and Inter-agency levels, communication networks,

and other forums. User feedback on GRASS program capabilities, documentation, and other technology transfer mechanisms is important to the development of the system. Users are encouraged to communicate such feedback to GRASS development staff at USACERL, via existing electronic communication networks and via the GRASS Information Center at USACERL, P.O. Box 9005, Champaign, IL 61826-9005; phone 217-373-7220; fax 217-373-7222; or e-mail grassbug@zorro.cecer.army.mil.

2. An Introduction to TIGER/Line

The 1990 U.S. Census TIGER/Line file contains digital data that represent physical features and non-visible boundaries, such as political boundaries and census statistical areas, for the entire United States. The TIGER/Line file was originally intended as a tool to aid the U.S. Census Bureau in its acquisitions of demographic information for the 1990 decennial census.

The physical features and boundaries available in the TIGER/Line file include: 59 different classifications of roads; 20 railroad classes; 28 classes of hydrography; a wide range of landmarks such as hospitals, schools, airports, pipelines and power transmission lines; and nonvisible boundaries both political and statistical. For a complete list, see Appendices A and B.

2.1. TIGER/Line Versions

As of December 1991, there are four versions of the TIGER/Line file. The different versions feature slightly different data formats and are available on one or more media. Version 1 is the prototype version and is available on tape only. Version 2 is the pre-census version, available on tape and compact disk read-only memory (CD ROM). This version contains the first six record types (see section 1.2). Version 3 is the same as version 2 with the addition of a voting districts record. Version 3 is available only on tape. Version 4, the TIGER/Line Census Files, is the 1990 TIGER/Line file and contains all of the information from the previous version in their updated form, plus six additional record types. The GRASS 4.0 version of TIGER in GRASS expects version 2, the pre-census version. Version 5, the School District version, is expected out on CD-ROM in 1993

2.2. TIGER/Line Structure

Depending on the version, the TIGER/Line file consists, structurally, of either six or twelve record types. As noted in the previous section, GRASS 4.0 uses version 2 of the TIGER/Line file which contains the first six record types. All 12 are listed below for the user's information.

Type 1	Basic Data Records
Type 2	Shape Coordinate Points
Type 3	Additional Decennial Census Geographic Area Codes
Type 4	Index to Alternate Feature Names
Type 5	Feature Name List
Type 6	Additional Address Range and ZIP Code Information
Type 7	Landmark Features
Type 8	Area Landmarks
Type A	Additional Polygon Geographic Area Codes
Type I	Area Boundaries
Type P	Polygon Location
Type R	Record Number Range

For further explanation of each of the record types, see Appendix C.

2.3. Naming Convention

The naming convention for the TIGER/Line file record types, each of which is a physically independent file, follows this format: "tgrSSCCC.fVN"; where "SS" is the two-digit FIPS state code, "CCC" is the three-digit FIPS county code, "V" is the TIGER/Line version number, and "N" is the record type (as listed above). Both "tgr" and "f" are constant. For Lawrence County, South Dakota, the record type files are:

tgr46081.f21
tgr46081.f22
tgr46081.f23
.
.

.
tgr46081.f2P
tgr46081.f2R

2.4. Amount of Data

The TIGER/Line file is conceptually one file, but physically many. The file has been sorted by state and divided into directories by county. The county directories are then divided into separate files for each of the record types associated with the particular counties.

The proportion of the TIGER/Line file dedicated to a particular state or county varies according to the size and population of the state or county. The average amount of data per state is 500 megabytes and ranges from 4 to 1,800 megabytes. The average size of the county files is 7.5 megabytes and ranges from 0.5 to 125 megabytes.

2.5. Source and Accuracy

A number of sources were used in developing the TIGER/Line file. The primary sources were United States Geological Survey (USGS) 1:100,000-scale maps and the 1980 GBF/DIME file (a digital data file of the urbanized areas of the United States created by the USGS and used by the Census Bureau for its 1980 decennial census). Other sources were maps from federal, state, and local agencies as well as private companies.

The positional accuracy of the information in the TIGER/Line file is no better than the National Map Accuracy standards for USGS 1:100,000-scale maps: 166 feet on the ground and 0.02 inches on a 1:100,000-scale map. Completeness of the TIGER/Line file for any particular area of the country is not guaranteed by the Census Bureau. Missing and inaccurate information is a possibility. Additionally, zip code information and address ranges are only available for urbanized areas and in many cases are still incomplete.

3. TIGER in GRASS

There are a series of programs within GRASS that facilitate the importation of TIGER data and RIM, a database management system. Programs within this series also allow the user to work with TIGER data in GRASS. In this version of TIGER in GRASS, maps can be built only in UTM locations. Forthcoming versions will accommodate map building in both UTM and latitude and longitude locations.

What follows are explanations and examples of these programs using the TIGER/Line file for Lawrence County, SD (the county which contains the area of the Spearfish sample data set used in GRASS). It is left to the user to provide the TIGER/Line input data to use when working through this tutorial. It will be slightly easier to use the Lawrence County TIGER/Line data when working through the tutorial, but TIGER/Line data for any county will work if the user makes the correct command-line input substitutions. In this tutorial it is also assumed that the user is familiar with concepts such as location and mapset and has some knowledge of GRASS. If this is not the case, the user should first read the GRASS Tool Access tutorial. If the examples are emulated in the following order, the result will be a complete set of TIGER vector files and a RIM database for Lawrence County, SD.

3.1. v.in.tig.rim

The GRASS program *v.in.tig.rim* converts the data for one county from the type 1 and type 2 TIGER/Line file into a RIM database and a master vector file, a map that contains every line segment in the TIGER/Line file that exists for a particular county. To create vector files of specific features such as roads or water, the program *v.db.rim* should be run after the completion of *v.in.tig.rim* (see section 3.7).

v.in.tig.rim requires four inputs: a database name, a type 1 input record file (basic data record), a type 2 input record file (shape coordinate points), and a UTM zone number. Inputs can be entered on the command line; alternatively, the user can simply enter the command name alone, and then *v.in.tig.rim* will prompt the user, interactively, for the inputs. Additionally, in order for *v.in.tig.rim* to work, the user must either be in the directory where the type 1 and type 2 records are located or specify the full

pathname of the input files. The command-line format and input associated with Lawrence County, SD is:

v.in.tig.rim dbname=sprfish in1file=tgr46081.f21 in2file=tgr46081.f22 zone=13

The database name is restricted to seven characters. For consistency, county names or their abbreviations make good database names, but any string of seven or fewer characters will work. *in1file* and *in2file* follow the file naming convention outlined in section 1.3. The UTM zone for the county that the user is processing can be found by running the program *m.tiger.region* (see section 2.3).

The master vector file (Figure 1) will be built and then stored under the user's current location and mapset with the name *dbname.Master* (in this example, *sprfish.Master*). The RIM database that *v.in.tig.rim* creates for the county is also built and stored in the user's current location and mapset in the rim/vect directory. It consists of three separate files with the naming convention: *dbname.rimdb1, dbname.rimdb2,* and *dbname.rimdb3* (in this example; *sprfish.rimdb1, sprfish.rimdb2,* and *sprfish.rimdb3*). Although they are three physically distinct files, the user never deals with them separately, but rather as a single database (named *dbname* or in this example *sprfish*) within the database manager RIM. In Figure 2, below, the data for one record (line segment) in the *sprfish* RIM database are listed. There is a similar about of data for each of the 8983 line segments in the sprfish RIM database. To interpret the abbreviations in Figure 2 refer to Appendix B.

Depending on the county, the size of both the RIM database and the master vector file can be quite large. As a general rule, the size of the two combined will be 50 percent larger than the combined size of the type 1 and 2 TIGER/Line records.

3.2. v.support

v.in.tig.rim creates only a dig and dig_att file for the master vector map. If the user wants to build a dig_plus file for the master vector map *v.support* should be run after *v.in.tig.rim* has completed. Although *v.support* is optional, it is a necessary set if the user intends to build vector maps within the interactive version of *v.db.rim*. In our example the command would be entered as follows:

v.support sprfish.Master

The result of this command would be a dig_plus file called *sprfish.Master* which is placed in the dig_plus directory of the user's current location and map set. v.support may take anywhere from one minute to several hours to run, depending on the user's machine platform and the number of type 1 TIGER/Line records. It should also be noted that v.support uses a tremendous amount of machine memory. This sometimes leads to the failure of v.support because the machine runs out of memory before the program is finished.

3.3. m.tiger.region

The program *m.tiger.region* determines the region coordinate extrema and the UTM zone(s) of given TIGER/Line input data. The program requires only the type 1 record file for the county of interest as input. Optionally, a zone and spheroid can be given as input. The program must be run from the directory in which the type 1 record is located or the full pathname must be specified. The output from this program will go only to the screen unless otherwise redirected (using the UNIX redirect function, ">"). For Lawrence County, SD the command would appear as follows:

m.tiger.region infile=tgr46081.f21

```
Number of calculated zones is: 1

INFO FOR ZONE 1:
  zone number:     13
  percentage of data points
  in this zone:        100.000000
  regional spread of points
  within this zone:
    north:    4939762.559994
    south:    4887763.314337
    east:     623820.776150
    west:      575000.425952
```

Figure 3. Output from *m.tiger.region*

Because the TIGER/Line coordinate data is given in latitude and longitude and then converted to UTMs by *v.in.tig.rim*, the user needs to know in which UTM zone(s) the coordinate data falls in order to set up a GRASS 4.0 location for the data. Referring to the output from *m.tiger.region* as shown above, if all of the TIGER/Line data are contained within one UTM zone, the information is given as above in "INFO FOR ZONE 1." If the data fall in more than one UTM zone, information is given for each zone: the UTM zone number, the percentage of data points in the zone, and the spread of the points within the zone. As a general rule, if the data for a county fall into more than one UTM zone, the zone

to choose for that county would be the zone in which the largest percentage of points fell.

3.4. tig.rim.sh

In addition to a county and a tract outline vector map, in Figure 4 *tig.rim.sh* also creates a blockgroup outline vector map for a given county and a series of user-specified tract-level vector maps. The individual tract maps come in three forms for every specified tract: a tract outline (T####), a block group outline map (####.###.bg) for each block group within the specified tract, and a map showing block boundaries (####.###.bk) within each block of the specified tract. Note that block groups 100, 200, 300, and so on correspond to block groups 1, 2, 3, etc., in some TIGER/Line documents and papers (from the U.S. Census Bureau and elsewhere).

tig.rim.sh takes as input the database name and the names of each of the tracts for which the user would like individual vector maps created. The tract numbers for a county can be found by running *tiger.info.sh* (see section 3.6). To produce vector maps for all of the census tracts for Lawrence County the input line would be:

tig.rim.sh sprfish 9961 9962 9963 9964 9965

This will produce the following 48 vector maps:

9961.100.bg	*9962.200.bg*	*9963.400.bg*	*9965.100.bg*
9961.100.bk	*9962.200.bk*	*9963.400.bk*	*9965.100.bk*
9961.200.bg	*9962.300.bg*	*9964.100.bg*	*9965.200.bg*
9961.200.bk	*9962.300.bk*	*9964.100.bk*	*9965.200.bk*
9961.300.bg	*9962.400.bg*	*9964.200.bg*	*T9961*
9961.300.bk	*9962.400.bk*	*9964.200.bk*	*T9962*
9961.400.bg	*9963.100.bg*	*9964.300.bg*	*T9963*
9961.400.bk	*9963.100.bk*	*9964.300.bk*	*T9964*
9961.500.bg	*9963.200.bg*	*9964.400.bg*	*T9965*
9961.500.bk	*9963.200.bk*	*9964.400.bk*	*sprfish.bg*
9962.100.bg	*9963.300.bg*	*9964.500.bg*	*sprfish.county*
9962.100.bk	*9963.300.bk*	*9964.500.bk*	*sprfish.tract*

If no tract numbers are specified on the command line, only the county, tract, and blockgroup outlines for the county will be created.

An example of the blockgroup outline is given in Figure 5. In addition, Figure 5 also depicts examples of the blockgroup outline map *9961.100.bg* and the block map *9961.100.bk*.

3.5. tiger.info.sh

tiger.info.sh provides the tract numbers and census classification codes (CFCC) associated with a particular county. The type 1 TIGER/Line data file is used as input and the program must be run from the directory in which the type 1 record file is located or the full pathname must be specified. The output from *tiger.info.sh* will go to the screen unless otherwise redirected (using the UNIX redirect function, ">"). The command line for our example would be:

tiger.info.sh tgr46081.f21

Identified the following tract numbers:

9961
9962
9963
9964
9965

Found the following Classification Feature Codes:

A00 Road, Classification Unknown or Not Elsewhere Classified
A11 Primary road, interstate highway and limited access road,
A11 undivided
A21 Secondary road, U.S. highway not classified A11, and state
A31 A11 or A21, undivided
A13 Primary road, interstate highway and limited access road,
A13 undivided, underpassing
A23 Secondary road, U.S. highway not classified A13, and state
A32 A13 or A23, undivided, underpassing
A21 Secondary road, U.S. highway not classified A11, and state
A21 roads, undivided
A31 A11 or A21, undivided
A31 Connecting road, county roads, and roads not classified as
A31 A11 or A21, undivided
A33 Connecting road, county roads, and roads not classified as
A40 Neighborhood roads, city streets and unimproved roads: This
A40 category includes city streets in built-up areas, unpaved roads
A40 that are passable with an automobile in non built-up areas, and
A40 all other remaining improved roads
A41 Neighborhood roads, city streets and unimproved roads,
A41 undivided
A51 Class 5 road, undivided
A63 Cloverleaf or interchange
A71 Walkway
B00 Railroad, Classification Unknown or Not Elsewhere Classified
B01 Railroad track
B11 Railroad main track
B13 Railroad main track, underpassing
B21 Railroad spur track
C20 Power Transmission Line
D51 Airport or airfield
F10 Nonvisible Political Boundary
F11 Offset corporate boundary
F23 Closure extension
F40 Property Line
F72 1990 Block boundary
H01 Shoreline of perennial water feature
H11 Perennial stream
H12 Intermittent stream or wash
H21 Perennial canal, ditch, or aqueduct
H22 Intermittent canal, ditch, or aqueduct

Figure 6. Output from *tiger.info.sh*

3.6. v.db.rim

v.db.rim is a database management/query interface, which allows the user to access and use both the RIM DBMS and GRASS vector files in a linked environment. *v.db.rim* allows users to create, manage and query information about labeled lines and areas from a number of different vector maps in a batch mode or interactively. Operations done on the database can be done on the command line or through an interactive menu. For more detail on *v.db.rim* see pages 359 to 375 of the *GRASS 4.0 User's Reference Manual*, Volume 2.

As alluded to earlier in this tutorial, both RIM and *v.db.rim* are used while processing and working with the TIGER/Line file in GRASS. Although invisible to the user, *v.in.tig.rim*, and *tig.rim.sh* use *v.db.rim* and access RIM databases. When running *v.in.tig.rim* for a particular county, a RIM database (consisting of TIGER data for that county) is created. Once the RIM database is created, the user can access it through the GRASS/RIM interface *v.db.rim* and create any number of different TIGER map layers (vector files).

This section will provide an example of how to create a TIGER/GRASS vector file using the command-line version of the interface, *v.db.rim*. For information on the interactive version of *v.db.rim*, refer to pages 368 to 375 of the *GRASS 4.0 User's Reference Manual*. The example provided below will produce a vector map of the power transmission lines found in Lawrence County, SD.

3.6.1. The Command-Line Version of v.db.rim

To enter the command-line version of *v.db.rim*, enter *v.db.rim* and the database name of the database that is to be accessed. If just *v.db.rim* is entered the interactive version of the interface will be opened. For our example the user would enter the following:

v.db.rim sprfish

The user will then be in the command-line version and be prompted with the following prompt:

v.db>

Once inside *v.db.rim*, all commands (with the exception of the where clause within the query command) start with a dot (".").

To build a vector map of power transmission lines for Lawrence County, the user will first need to query the relevant line segments from the RIM database. To do this, the user will need to enter the query mode. This is done by entering ".query" or the command's abbreviated form ".q".

v.db> **.query**

The user will then be within the query mode and be prompted by:

query>

It is at this point that the actual query is performed. For more detail about the query mode see page 369 of the *GRASS 4.0 User's Reference Manual*. To query line segments from the RIM database that represent power transmission lines enter:

query> **where CFCC=C20**
query> **.end**

The ".end" (or its abbreviated form ".e") exits from the query mode and returns to the *v.db.rim* command prompt. A successful query will yield the following lines:

184 records now selected from database <sprfish>
v.db>

The user then enters a command that would process the selected records, creating a GRASS vector map. In addition to the command itself, ".vect_map" (or ".v"), the user must also specify the name of the vector map to be built. In this example *sprfish.power* will be the name of the new vector map, but any user-specified name will be good. For the other options associated with this command refer to page 366 of the *GRASS 4.0 User's Reference Manual*. The simpliest form of the command is:

v.db> **.vect_map sprfish.power**

After completion of this task, the interface will come back with the following message:

Vector map <sprfish.power> has been built.
Run v.support on this map to build the topology information.

v.db>

At this point the user is done with *v.db.rim*. To exit the interface enter ".exit" or ".ex" on the command line, as follows:

v.db> **.exit**

Upon returning to GRASS, the user can run *v.support* on the vector file to create a topology (dig_plus) file by entering:

v.support sprfish.power

The resulting *sprfish.power* vector map can be displayed using *d.vect* and will look the same as the internal lines of Figure 7. The county outline, vector map sprfish.county, has been added to Figure 7 for reference purposes.

Figure 7. *sprfish.power*

4. Moving RIM Databases and TIGER Vector Files

4.1. Moving and Renaming the Master TIGER Vector Files

The mapset (directory), in which the master vector file for a county is located is coded into the RIM database for that county when it is created by *v.in.tig.rim*. If the user changes the name of the mapset or moves the mapset to a different location, the user must also change the name of the mapset in the RIM database for that county. This must be done because *v.db.rim* looks-up and uses the master vector file mapset name in the RIM database when it builds new vector files.

To change the mapset (directory) name in the RIM database, the user, once the mapset has been moved, simply needs to modify one line in the RIM database for that county. The following instructions will step the user through this change. Again the Lawrence County, SD database (sprfish) will be used in the example.

The user must be in the directory in which the RIM database is stored. This will be in the user's current location and mapset under the directory rim/vect. Once in that directory, the user should enter RIM and open the database.

grass 4.0> **rim**
Rim: **open sprfish**

Once in RIM, one item in the *referencemaps* table needs to be modified. To view the table enter:

Rim: **select * from referencemaps**

The item to be modified is under the *map_mapset* column. This item is the name of the mapset (directory) in which the master vector file is stored. When the master vector file is moved or the entire mapset is moved or renamed, the *map_mapset* item will need to be changed as well. To do this, enter the following:

Rim: **change map_mapset to newname in referencemap where map_mapset eq lawrence**

In the above command the mapset name is changed in the RIM database from lawrence (the old mapset or directory) to newname. After making the above change view the table again to verify the change by using the *select* command as shown above, then exit RIM by entering:

Rim: **exit**

4.2. Moving RIM Databases to Different Machine Platforms

Moving RIM databases to different machine platforms is a more involved process than simply moving databases to different directories on the same machine or type of machine. Because RIM databases have different binary forms on different machine platforms, if a database is moved from one machine platform to another, the second machine may not be able to read the binary form of the database correctly. To move RIM databases from one machine platform to another and to be able to use them on the second platform, the databases will need to be translated (downloaded) into ASCII format, moved to the different machine, and then translated back (uploaded) into RIM on the second platform.

To accomplish the moving of a RIM database from one machine platform to another, the command-line version of the GRASS interface *v.db.rim* should be used. Within *v.db.rim* the user queries an entire RIM database and writes it to a file (in this example; *big.tempfile*). In addition the user needs to write the screen layout information to a file (in this example; *layout.tempfile*). The example below will use sprfish as the database name and lawrence as the mapset (directory) name.

v.db.rim sprfish
v.db> **.query**
v.db> **.end**
v.db> **.out big.tempfile**
v.db> **.p -a**
v.db> **.out**
v.db> **.out layout.tempfile**
v.db> **.show**
v.db> **.out**

v.db> **.exit**

These steps will create two ASCII files; one very large file with all the rows from the data table of the RIM database and one screen layout file from the *screenlayout* table, both of which will be in an ASCII form that can be read on any machine platform.

Before moving the files to the new machine the *layout.tempfile* will need to modified. In the editor, *vi*, edit the *layout.tempfile* by adding a **.make** as the first line of the file and a **.end** as the last line of the file.

After the *layout.tempfile* has been edited and both files have been moved to the second machine, use the following commands to recreate the RIM database on the second machine:

v.db.rim sprfish
v.db> **input layout.tempfile**
v.db> **input big.tempfile**
v.db> **.exit**

Finally, the user needs to load the *referencemaps* table into the RIM database. This is done manually in RIM so the user should be in the current location and mapset in the rim/vect directory and then enter the following series of commands:

rim
Rim: **open sprfish**
Rim: **load referencemaps**
Rim_load: **1 sprfish.Master lawrence**
Rim_load: **end**
Rim: **exit**

On the line *Rim_load: 1 sprfish.Master lawrence*, the *1* is a constant, *sprfish.Master* is the *map_name* or the name of the master vector file, and *lawrence* is the *map_mapset* or the mapset in which the master vector file is stored.

5. CONCLUSIONS AND RECOMMENDATIONS

This document guides the user through the process of using GRASS programs and the RIM public domain database management system to which GRASS is interfaced to import, manipulate, and convert U.S. Census Bureau TIGER/Line data in GRASS. GRASS programs that extract five five additional record types included in the 1990 TIGER/Line Census files (Version 4): i.e., record types 7, 8, A, I, and P are explained in the tutorial, "Importing/Line Maps and STF1 Census Data to GRASS." Programs discussed in that tutorial use TIGER/Line data in either latitude and longitude or convert the coordinates to UTMs. Other programs discussed in that tutorial deal with importing and using 1990 STF1 Census Dta in GRASS.

This manual is part of an ongoing process to document Geographic Resources Analysis Support System program functions. It is one of many technology transfer mechanisms being implemented to explain current, and future, GRASS program capabilities to resouce managers and system users.

Documentation, along with training programs, hands-on use, a user document center, newsletters, institutional structures at the Army and inter-agency levels, communication networks, and other forums, can be used to aid this on-going technology transfer effort.

User feedback on GRASS program capabilities, documentation, and other technology transfer mechanisms is also needed. Users are encouraged to communicate such feedback to GRASS development staff at USACERL, via existing electronic communication networks and via the GRASS Information Center at USACERL, P.O. Box 9005, Champaign, IL 61826-9005, Phone: 217-373-7220, FAX: 217-373-7222, Email: grassbug@zorro.cecer.army.mil.

SOURCES AND REFERENCES

The following document was used in the preparation of this tutorial and its appendices and should be referred to by the user when working with the TIGER/Line file.

U.S. Department of Commerce, Bureau of the Census, *TIGER/Line Census Files, 1990 Technical Documentation* (Washington, D.C.).

A copy of this document can be obtained by writing or calling:

Data User Services Division
Bureau of the Census
Washington, D.C. 20233
Phone: (301) 763-1580

Other, general GRASS-related references are listed below.

Ruiz, Marilyn S., and Jean M. Messersmith, *Cartographic Issues in the Development of a Digital GRASS Database*, USACERL Special Report N-90/16 (U.S. Army Construction Engineering Research Laboratory [USACERL], September 1990).

Westervelt, James, and William D. Goran, *Introduction to GRASS 4*, Draft ADP Report (USACERL, July 1991).

Westervelt, James, Mary Martin, and Deborah Brinegar, *GRASS 4.0 Programs*, in *GRASS User's Reference Manual*, ADP Report N-87/22 rev (USACERL, September 1988, last revised December 1991).

Westervelt, James, Michael Shapiro, William D. Goran, et al., *GRASS User's Reference Manual*, ADP Report N-87/22 rev (USACERL, September 1988, last revised December 1991).

Appendix A: Census Feature Class Codes (CFCC)

This is a series of codes that provides more detailed information on the classification of the line segment, such as class of road, class of stream, and so forth. This field provides more information than the 1980 GBF/DIME-File "NS" code (nonstreet feature code) field. For line segments that originated with the USGS DLG-3 files, the CFCC is based on the USGS classification code in the DLG-3 file. For line segments that originated with the 1980 GBF/DIME-Files, the CFCC is based on the NS code and other feature identification content of the GBF/DIME-File. A list of these codes follows.

6. CFCC Classification A = Road Features

A00 Road, Classification Unknown or Not Elsewhere Classified
A01 Road, undivided
A02 Road, undivided, in tunnel
A03 Road, undivided, underpassing
A04 Road, undivided, with rail line in center
A05 Road, divided
A06 Road, divided, in tunnel
A07 Road, divided, underpassing
A08 Road, divided, with rail line in center

A10 Primary road, interstate highway, and limited-access road; this category includes interstate highways, primary U.S. highways, primary state highways, most multi-lane roads, and most other limited-access roads
A11 Primary road, interstate highway, and limited-access road, undivided
A12 Primary road, interstate highway, and limited-access road, undivided, in tunnel
A13 Primary road, interstate highway, and limited-access road, undivided, underpassing
A14 Primary road, interstate highway, and limited-access road, undivided, with rail line in center
A15 Primary road, interstate highway, and limited-access road, divided
A16 Primary road, interstate highway, and limited-access road, in tunnel
A17 Primary road, interstate highway, and limited-access road, divided, underpassing
A18 Primary road, interstate highway, and limited-access road, divided, with rail line in center

A20 Secondary road, U.S. highway not classified A10, and state roads; this category includes the U.S. highways not classified as A10 and state roads. Most of the roads in this category tend to be state-level roads
A21 Secondary road, U.S. highway not classified A11, and state roads, undivided
A22 Secondary road, U.S. highway not classified A12, and state roads, undivided, in tunnel

A23 Secondary road, U.S. highway not classified A13, and state roads, undivided, underpassing
A24 Secondary road, U.S. highway not classified A14, and state roads, undivided, with rail line in center
A25 Secondary road, U.S. highway not classified A15, and state roads, divided
A26 Secondary road, U.S. highway not classified A16, and state roads, divided, underpassing
A27 Secondary road, U.S. highway not classified A17, and state roads, divided, underpassing
A28 Secondary road, U.S. highway not classified A18, and state roads, divided, with rail line in center

A30 Connecting road, county roads, and roads not classified as A10 or A20; this category includes county roads, roads not classified A10 or A20 that connect towns or major features, and principal non-A10/A20 roads through built-up areas. Most of the roads in this category are county roads
A31 Connecting road, county roads, and roads not classified as A11 or A21, undivided
A32 Connecting road, county roads, and roads not classified as A12 or A22, undivided, in tunnel
A33 Connecting road, county roads, and roads not classified as A13 or A23, undivided, underpassing
A34 Connecting road, county roads, and roads not classified as A14 or A24, undivided, with rail line in center
A35 Connecting road, county roads, and roads not classified as A15 or A25, divided
A36 Connecting road, county roads, and roads not classified as A16 or A26, divided, in tunnel
A37 Connecting road, county roads, and roads not classified as A17 or A27, divided, underpassing
A38 Connecting road, county roads, and roads not classified as A18 or A28, divided, with rail line in center

A40 Neighborhood roads, city streets, and unimproved roads; this category includes city streets in built-up areas, unpaved roads that are passable with an automobile in non built-up areas, and all other remaining improved roads
A41 Neighborhood roads, city streets, and unimproved roads, undivided
A42 Neighborhood roads, city streets, and unimproved roads, undivided, in tunnel
A43 Neighborhood roads, city streets, and unimproved roads, undivided, underpassing
A44 Neighborhood roads, city streets, and unimproved roads, undivided, with rail line in center
A45 Neighborhood roads, city streets, and unimproved roads, divided
A46 Neighborhood roads, city streets, and unimproved roads, divided, in tunnel
A47 Neighborhood roads, city streets, and unimproved roads, divided, underpassing
A48 Neighborhood roads, city streets, and unimproved roads,

divided, with rail line in center

A50 Class 5 Road - (Jeep Trail)
A51 Class 5 road, undivided
A52 Class 5 road, undivided, in tunnel
A53 Class 5 road, undivided, underpassing

A60 Special Road Feature
A61 Cul-de-sac
A62 Traffic circle
A63 Cloverleaf or interchange
A64 Service drive
A65 Ferry crossing

A70 Other Thoroughfare
A71 Walkway
A72 Stairway
A73 Alley

7. CFCC Classification B = Rail Features

B00 Railroad, Classification Unknown, or Not Elsewhere Classified
B01 Railroad track
B02 Railroad track, in tunnel
B03 Railroad track, underpassing

B10 Railroad Main Track
B11 Railroad main track
B12 Railroad main track, in tunnel
B13 Railroad main track, underpassing

B20 Railroad Spur Track
B21 Railroad spur track
B22 Railroad spur track, in tunnel
B23 Railroad spur track, underpassing

B30 Railroad Yard
B31 Railroad yard
B32 Railroad yard, in tunnel
B33 Railroad yard, underpassing

B40 Railroad Ferry Crossing

B50 Other Rail Feature
B51 Carline
B52 Cog railroad, incline railway, or logging tram

8. CFCC Classification C = Pipelines, Power Transmission Lines, and Miscellaneous Transportation Features

C00 Special Transportation Feature, Classification Unknown, or Not Elsewhere Classified

C10 Pipeline

C20 Power Transmission Line

C30 Other Special Transportation Feature
C31 Aerial tramway, monorail, or ski lift

9. CFCC Classification D = Landmark Features

D00 Feature Not Elsewhere Classified

D50 Transportation Terminal
D51 Airport or airfield

10. CFCC Classification E = Other Physical Features

E00 Physical Feature, Classification Unknown, or Not Elsewhere Classified

E10 Fence Line

E20 Topographic Feature
E21 Ridge line

11. CFCC Classification F = Nonvisible Boundaries

F00 Nonvisible Boundary, Classification Unknown, or Not Elsewhere Classified

F10 Nonvisible Political Boundary
F11 Offset corporate boundary
F12 Corporate corridor
F13 Nonvisible interpolated boundary

F20 Feature Extension (extensions not otherwise classified)
F21 Automated extension
F22 Manually-added block extension
F23 Closure extension
F24 Nonvisible separation line
F25 Nonvisible corporate corridor centerline

F30 Point-to-Point Line

F40 Property Line

F50 ZIP Code Boundary

F60 Map Edge

F70 Statistical Area Boundary
F71 1980 statistical area boundary
F72 1990 block boundary
F73 Urbanized area land use boundary
F74 1990 statistical area boundary

12. CFCC Classification H = Hydrographic Features

H00 Water Feature, Classification Unknown or Not Elsewhere Classified

H01 Shoreline of perennial water feature
H02 Shoreline of intermittent water feature

H10 Stream
H11 Perennial stream
H12 Intermittent stream or wash
H13 Braided stream

H20 Canal, Ditch, or Aqueduct
H21 Perennial canal, ditch, or aqueduct
H22 Intermittent canal, ditch, or aqueduct

H70 Nonvisible Water Definition Boundary
H71 USGS closure line
H72 Census-computed center line
H73 Census international water boundary
H74 Census water boundary
H75 3-mile limit water boundary

13. CFCC Classification X = Features Not yet Classified

X00 Feature Not Yet Classified

Appendix B: TIGER Database Screenlayout Format with Field Names and Descriptions

(The field names used below refer to the data contained in the Type 1 record of the Precensus TIGER/Line file).

where:
~ = space marker
s = site identification number
x = easting coordinate
y = northing coordinate
f = double precision floating point decimal
i = integer
t = text string
v = vector type
m = reference map

Precensus TIGER/Line Record Type 1 v.db.rim Screenlayout

NAME	DESCRIPTION
RECTYP	Record Type (value "1")
VERSION	Version Number
RECNUM	Record Number (permanent)
SIDECDE	Single Side Segment Code (value "1" signifies data exists only for one side of the segment)
SOURCE	Source Code
DIRPRE	Feature Direction, prefix
FEANME	Feature Name
FEATYP	Feature Type
DIRSUF	Feature Direction, suffix
CFCC	Census Feature Class Code
FRADDL	From Address Left
TOADDL	To Address Left
FRADDR	From Address Right
TOADDR	To Address Right
FRIADDFL	From Imputed Address Flag Left
TOIADDFL	To Imputed Address Flag Left
FRIADDFR	From Imputed Address Flag Right
TOIADDFR	To Imputed Address Flag Right
ZIPL	Zip Code Left (only when an address range is present)
ZIPR	Zip Code Right (only when an address range is present)
FAIRL	FIPS American Indian Reservation (AIR), Alaska Native Village Statistical Area (ANVSA), Tribal Jurisdiction Statistical Area (TJSA), Tribal Designated Statistical Area (TDSA), left
FAIRR	FIPS American Indian Reservation (AIR), Alaska Native Village Statistical Area (ANVSA), Tribal Jurisdiction Statistical Area (TJSA), Tribal Designated Statistical Area (TDSA), right
ANRCL	Alaska Native Regional Corporation, code left
ANRCR	Alaska Native Regional Corporation, code right
STL	FIPS State Code, left
STR	FIPS State Code, right
COL	FIPS County Code, left
COR	FIPS County Code, right
FMCDL	FIPS MCD/CCD Code, left
FMCDR	FIPS MCD/CCD Code, right
FSMCDL	FIPS SUB-MCD Code, left
FSMCDR	FIPS SUB-MCD Code, right
FPLCDEL	FIPS Place Code, left
FPLCDER	FIPS Place Code, right
CTBNAL	Census Tract/BNA Code, left
CTBNAR	Census Tract/BNA Code, right
BLKL	Tabulation Block Number, left
BLKR	Tabulation Block Number, right
FROMLONG	Longitude From (implied 6 decimal places)
FROMLAT	Latitude From (implied 6 decimal places)
TOLONG	Longitude To (implied 6 decimal places)
TOLAT	Latitude To (implied 6 decimal places)

Precensus TIGER/Line Record Type 1 Field Names and Descriptions

Appendix C. Summary of TIGER/Line Record Types

14. Record Type 1: Basic Data Record

Record Type 1 provides a single record for each unique line segment in the 1990 Census TIGER/Line file. The endpoints of the line segments are expressed in latitude/longitude coordinate values in degrees and decimal fractions of a degree to six decimal places. This record also contain address ranges and ZIP codes (in selected areas) and geographic area codes for each side of the line segment. By convention, if one is standing at the "from" coordinate position facing the "to" coordinate position, data listed in the fields carrying a right qualifier would indeed be found to the right of the line segment. Data users can collect the necessary segments to construct polygons and features that intersect from the information contained in this basic record.

15. Record Type 2: Shape Coordinate Points

Record Type 2 provides an additional series of latitude and longitude coordinate values that describe the shape of each line segment that is not straight for the associated Record Type 1. All coordinate values are expressed in degrees and decimal fractions of a degree of latitude and longitude. The decimals are carried to six places to permit the representation of lines that are very close to one another. If the segment in Record Type 1 is a straight line, there will not be a Record Type 2.

16. Record Type 3: Additional Decennial Census Geographic Area Codes

Record Type 3 includes the 1990 voting district codes provided to the Census Bureau for the 1990 Census Redistricting Data Program. Record Type 3 also includes some 1980 census geographic area codes and 1990 census geographic area codes not included on Record Type 1. The 1980 census block numbers are available only for areas covered by the 1980 GBF/DIME-Files. During the conversion of the 1980 GBF/DIME-Files to the TIGER data base format, some 1980 census block numbers may have been deleted or changed. Users are advised to check all 1980 census geographic area codes, especially any 1980 block numbers, before using them in a planned application. The Census Bureau has not verified any of the 1980 census geographic codes in these files. There will be discrepancies between the geographic area boundaries and codes in these files and the 1980 census maps, which are the basis for the 1980 census tabulations.

The TIGER/Line files may contain 1980 block numbers for portions of the country where the Census Bureau did not tabulate 1980 census data by block or block group. These situations occur because these portions of the Tiger file originated from the 1980 GBF/DIME-Files that extended beyond the 1980 block-numbered area. Data users concerned about the validity of 1980 block numbers are advised to discard all 1980 block numbers that do not correspond to block numbers in the 1980 MARF or 1980 STF files.

17. Record Type 4: Index to Alternate Feature Names

Record Type 4 provides an index to alternate feature names associated with the basic line segment (Record Type 1). A Record Type 4 will not exist for a Record Type 1 that has only one name. A line segment can have more than one alternate name. For example, Record Type 1 may have the name Main Street, while Record Type 4 points to the segment’s alternate names, such as 5th Street and State Highway 101, contained in the list of feature names (Type 5 records). If no Type 1 record has an alternate name, there will be no Record Type 4 file. There is no assurance that a feature that consists of multiple-line segments will consistently have the same name as the primary (on Record Type 1) and alternate names on the Record Type 4 index.

18. Record Type 5: Feature Name List

Record Type 5 contains a list of all unique feature names for line segments in the 1990 Census TIGER/Line file. Each name has a Feature ID code. Record Type 4 uses these Feature ID codes as as "shorthand" for the alternate names. Feature names for landmarks are in Record Type 7.

19. Record Type 6: Additional Address Range and ZIP Code Data

Record Type 6 provides additional address range information for a street segment when the information cannot be presented as a single address range (e.g., the house/building numbers are not uniformly arranged to form an address range). Additional ZIP Codes, if any, also appear in Record Type 6 for corresponding address ranges. Record Type 6 appears only for those counties that have address ranges and ZIP Code information in the TIGER data base. There is no assurance that the address ranges provided on Record Type 6 will be "shorter" than those appearing on Record Type 1. Data users must use Record Type 6 to obtain the complete picture of the potential address ranges along a segment.

20. Record Type 7: Landmark Features

Record Type 7 contains the area and point landmarks in the Census Bureau's TIGER data base. During the extraction of this data, we assigned a temporary Landmark Identification Number that we use to link the landmark attributes to the polygons that comprise the landmark. If there is no landmark in a county file, there will be no Record Type 7 or Record Type 8 for that county file.

21. Record Type 8: Area Landmarks

We use Record Type 8 to link the polygon identification codes with the area landmark identification codes. Thus, there will be from one to n Type 8 records for each area landmark, depending on the number of internal polygons in the Census Bureau's TIGER data base. If a county file does not have Record Type 7, it also will not have Record Type 8.

22. Record Type A: Additional Polygon Geographic Area Codes

Record type A exists for every polygon in the TIGER data base. We also are providing the basic 1990 census geographic area codes (State, county, county subdivision, place, American Indian/Alaska Native area, census tract/BNA, block) on this record type to assist those data users who are interested only in polygon information. We have reserved several fields for possible future use; we have not established a schedule for future versions of the TIGER/Line file.

23. Record Type I: Area Boundaries

Record Type I links the segments in Record Type 1 to the polygons that are provided in Record Type P. We provide a Record Type I for every Record Type 1. When Record Type I is linked to a single-sided (county boundary) Record Type 1, only the left or right polygon identifier is provided.

24. Record Type P: Polygon Location

Record Type P exists for every polygon in the TIGER data base. Because the file includes all lines in the TIGER data base in the TIGER/Line file as Record Type 1 records, there will be a one-to-one relationship between the polygons constructed from Record Types 1 and 2 and Record Type P. In constructing polygons from Record Types 1 and 2, users are cautioned to be sure their software can handle the necessary coordinate precision and does not "snap" lines that

are close together.

Record Type P that represents a polygon contains an internal point for that polygon. Some of the polygons (approximately 400 nationwide) are so small that the internal point given may be identical to a point on one of the lines bounding the polygon or identical to one of the nodes. Depending upon the software or hardware employed by the data user, the internal point may be identified to the data user as being outside the correct polygon or a neighboring polygon may be identified as possessing two internal points.

25. Record Type R: Record Number Range

Record Type R contains the minimum, maximum, and current high values for the permanent Record Number within each census file partition. Because each TIGER/Line file can reside in multiple partitions, there will be one record type R for each CENID in the TIGER/Line file for any county or statistically-equivalent entity.