Bureau of the Census, Geography Division
The Global Positioning Systems Test Project
A Report of the Newberry County, South Carolina Field Test

Introduction

The Newberry County, South Carolina field test (Newberry Field Test) is the second field test conducted by the Geospatial Research and Standards Staff (GRaSS) in support of the Global Positioning Systems (GPS) Test Project. The results of the initial field test conducted in Hampshire County, West Virginia (Hampshire Field Test) are reported in the GRaSS document titled "*Bureau of the Census Global Positioning Systems (GPS) Field Test Results.*" This Newberry Field Test report summarizes field procedures modified as a result of the Hampshire Field Test and used in the Newberry Field Test, identifies shortcomings and recommends improvements in field methodology, and reports the status of the analysis and evaluation of the collected GPS positional data.

Objectives

The objectives of <u>GRaSS-conducted field tests</u> are:

- Collect GPS positional data for use in the GPS Test Project.
- Develop procedures, and ultimately specifications, for collecting GPS positional data on a national basis.
- Evaluate the GeoLink® PowerMap™ software of the CRADA partner, GeoResearch, Inc., for use in Census GPS positional data collection activities and work with GeoResearch, Inc. to enhance GeoLink PowerMap to improve its ability to collect GPS positional data for Census specific applications.
- Gain firsthand experience in using GPS technology to collect positional data.

The further objectives <u>specific to the Newberry Field Test</u> are:

- Compare Geographic Data Technology, Inc.'s (GDT) Dynamap/2000® dataset with TIGER/Line '97 and on-site observations (ground truth) to determine the currentness and correctness of both datasets and to use the resulting information to develop strategies for future updating of the respective datasets.
- Develop the costs and determine the benefits of using GPS positional data and compare these to the alternative of using Digital Orthophoto Quadrangle data (DOQs) for collecting anchor points (anchor points are pivotal to the proposed GRaSS methodology for spatially enhancing TIGER.)

The overreaching goals of the <u>GPS Test Project</u> are:

- Spatially enhance and update the TIGER database with positional data captured using either GPS technology or DOQs by
  - adjusting the existing coordinates of TIGER,
  - improving the coordinates in TIGER (through one cell curve point substitution), and
  - updating (adding) features to TIGER.
- Investigate the viability of collecting the locations (latitude/longitude coordinates) of housing units and other structures for insertion in the TIGER database.

The Newberry Field Test

A Second Field Test

Although originally planned as part of a series of tests, the importance for a second field test arose from problems encountered during the first field test in Hampshire County, West Virginia. The Hampshire Field Test served as an important proving ground for Census collection of GPS positional data and improved the staff's understanding of GPS technology and its potential; lessons learned in Hampshire greatly improved collection procedures. However, two factors limited the usability of the collected data from the Hampshire Field Test for the GPS Test Project:

- GeoLink PowerMap allows for the capture of coordinate data (longitude/latitude), in decimal degrees, to double precision. Due to inexperience, equipment defaults were incorrectly set and, in actuality, data was captured only to single precision, which generally does not provide greater positional accuracy than that already in TIGER[1].
- The learning curve for GPS data collection coupled with the extreme terrain in Hampshire (steep canyons, narrow winding roads, isolated and difficult to reach housing units) slowed data capture far below the anticipated level. Therefore, an insufficient portion of the road network and housing units were captured for a viable test.

Because of problems with collecting enough data and with the collected data, a second field test became mandatory.

Selecting Newberry

The Census Bureau selected the city and area around Columbia, South Carolina as a Dress Rehearsal (DR) Site. The Columbia, South Carolina DR Site contains the city of Columbia in its entirety, the Town of Irmo, and 11 contiguous counties. Newberry County is one of the contiguous counties.

The GPS Test Project was described during DR preparation meetings and both State and local officials indicated an interest in participating. Newberry County's E911 service is actively involved in geographic database issues, as are the Office of Research and Statistics of the State of South Carolina and faculty at the University of South Carolina (located near Newberry County in Columbia, SC.) Further, Newberry County had USGS DOQ data coverage dated 1990 or more recently. Newberry County was found to meet GRaSS criteria for becoming a field test site.

Relevant Statistics

Important factors in selecting Newberry were:

- A rural county, sufficiently different in geographic character from the initial Hampshire Field Test site, was desired.

---

[1] Single precision, 32-bit numbers can only resolve location down to about +/- 32 feet. That is inadequate for positioning road centerlines relative to structure centroids or for matching with certainty to coordinates of cadastral datasets. Double precision, 64-bit numbers are capable of resolving locations to centimeters.

- A size (area) considered sufficient to complete the field test within a two-week timeframe. Number of staff available for the field test was a limiting factor, making the road network distances and number of housing units in a county an important consideration.
- Proximity to Census Bureau headquarters (within a day's drive) so that the equipment for the field test could be carried, not shipped. Additionally, air access and lodging were considerations.
- State and local agency cooperation was considered critical to the success of the field test.
- USGS DOQ coverage of 1990 or more recent was necessary to allow for comparison of data collection techniques.

Newberry County, South Carolina (FIPS.State = 45, FIPS.County = 071) is a rural county containing five Block Numbering Areas (BNAs) approximately equal in terms of geographic area, population, and housing units. The total 1990 population of the county is 33,172 with the majority of the population located in areas considered rural by the Census Bureau. The number of housing units (a 13,777 count in 1990) and total road network distances measured in miles (1,545 miles in TIGER/Line 97) fell within the workload criteria. (Refer to Appendix A for complete Newberry County statistics.)

Field Test Overview

The Newberry Field Test was conducted from June 2nd through June 11th, 1998. Seven Geography Division staff (six from GRaSS) participated.

Monday, June 1st and Friday, June 12th were travel days. Sunday, June 14th was a rest day.

Six Geography Division staff teamed with South Carolina Geodetic Survey staff to form six "Rover" teams collecting GPS positional data.

The seventh Geography Division staff member was the Operations Center Manager. Operations Center Manager duties included manning an Operations Center for the Rover teams, briefing visitors (State and local agency officials, Census officials, local politicians, and university faculty) and processing daily the Rover collected data.

An additional Geography Division staff member provided support at the Operations Center during the first three days of fieldwork.

A representative from GeoResearch, Inc. was on-site for consultations during the first three days of data collection.

Nine days were devoted to collecting GPS positional data. A synopsis of total staff days:

- Geography Division Staff – 66 staff days
- GeoResearch, Inc. Staff – 3 staff days
- South Carolina Geodetic Survey Staff – 56 staff days

Significant Changes

Several aspects of field-testing were modified following a review of the Hampshire Field Test. The modifications significantly adding to the success of data collection in the Newberry Field Test are categorized and described below (Many procedures did not change. Refer to the "*Bureau of the Census Global Positioning Systems (GPS) Field Test Results*" and/or field manuals for a more complete field methodology.)

Pre-Planning

Pre-planning is an important aspect of conducting an efficient and complete field test. For the Newberry Field Test, the entire GRaSS staff:

- Attended and participated in weekly planning meetings, beginning two months prior to the field test.
- Jointly identified and implemented procedural change decisions.
- Developed detailed activity flow charts and work plans.
- Were given oversight of and responsibility for specific tasks.

A major benefit of involving the entire staff in pre-planning was an increased knowledge base of all planned activities. At the conclusion of the Newberry Field Test, Census Bureau participants reported their understanding of the plan aided in data capture as they were able to make informed decisions when faced with on-site problems.

Procedures and Manuals

Three types of manuals were developed and maintained:

- The Rover Manual – A concise "how to" manual carried by the Rovers in their vehicles. The Rover Manual contained safety tips, a general schedule, explained the initial on-site briefing and equipment check-out and inventory procedure, summarized the daily morning briefing and use of progress maps and progress reports, provided general rules and methods of data collections, described data check-in, and explained the final site close-out briefing.
- A Project Manual – A documentary of the Newberry Field Test. The Project Manual contained background information, rationale for decisions made, ongoing records of all contact, meeting notes, and information about equipment.
- A GPS Base Station Set-Up Guide – A concise "how to" guide on setting up and maintaining the GPS Base Station and data collection procedures.

Dress Rehearsal (DR)

Before the Newberry Field Test, a DR was held at the Census Bureau's Geography Division Headquarters. Each staff member was responsible for developing one aspect of the training, which included both lectures and collecting GPS positional data.

Preparing for and having a DR was critical to identifying problems and resolving the identified problems prior to departing for the field. GeoResearch, Inc. delivered a new version of GeoLink PowerMap to GRaSS for use in Newberry County. Having a DR allowed GRaSS to fully test and evaluate the new version. GRaSS determined its reliability was questionable when running on Census Bureau laptop computers.

As GRaSS was unable to consistently collect GPS positional data using the new version, the decision was made not to risk the field test. Discovering the problem prior to the actual field test allowed GRaSS to revert to the earlier version of the software (used in the Hampshire Field Test), as well as to develop more extensive procedures for coping with software problems.

Staff reported the DR gave them confidence. The amount of data collected during the first day of the Newberry Field Test indicates a significantly lowered on site, data collection learning curve. The staff's obvious competence impressed the local drivers with whom they were teamed and added to the spirit of cooperation with local officials.

GeoResearch, Inc's Continued Involvement

GeoResearch, Inc extended its involvement in the Newberry Field Test at the request of GRaSS. As stated above, GeoResearch, Inc. delivered an updated version of its software for use in Newberry. The updated version incorporated some changes suggested by GRaSS as a result of the Hampshire Field Test. GeoResearch, Inc. worked closely with GRaSS in attempting to correct the problems identified during the DR[2]. GeoResearch, Inc. is investigating the possibility that a modification in the way in which information is displayed may overload data storage and interrupt GPS signal reception; another possibility being checked is an incompatibility between a function of the updated software and the brand of laptop (Samsung) used by GRaSS.

The presence of GeoResearch, Inc. staff on site during the first few days of the field test was invaluable. GeoResearch, Inc. assisted in resolving minor procedural problems and responded to the many questions that arose when the field test began.

The Importance of Local Participation

The Newberry County experience differed from the Hampshire County experience in the amount of advanced coordination for cooperation and support from State and local government officials. In the Hampshire County Field Test, the Census Bureau made only limited efforts to incorporate local participation. Newberry County was initially selected for a field test because of its Census Bureau DR status. State and local officials had expressed interest in the goals of the GPS Test Project and in data sharing. Excellent cooperation between the Census Bureau and South Carolina State and local government agencies and the on-site support received was a major factor in the success of the Newberry Field Test. Support was received from:

- Newberry County Administrator
  The initial Newberry County contact assisted in identifying State and local organizations to assist in the field test.
- Newberry County Law Enforcement Center/South Carolina
  Provided a secure, accessible location for the Census Bureau Base Station equipment.

[2] The reason for the difficulties remains unknown and is still being investigated by GeoResearch, Inc. GRaSS and GeoResearch, Inc. conducted additional tests of the new software version in Newberry. The "old" and "new" versions of the software were simultaneously used by two Rovers in a single vehicle to collect GPS positional data. Collection problems with the "new" software were verified.

- South Carolina State Budget and Control Board, Office of Research and Statistics Provided “up-to-date” State maintained TIGER/Line 95 files for Newberry County.
- South Carolina State Budget and Control Board, South Carolina Geodetic Survey Provided drivers with local knowledge and the 4-wheel drive vehicles necessary to successfully navigate back roads (often muddy) in rural areas. (This is important since the State’s Geodetic Survey vehicles are a familiar sight, so that the Census Bureau did not encounter the distrust of “Feds from Washington” encountered during the Hampshire Field Test.)
  Additionally, daily provided Base Station information in RYNEX format.

Staff from the above-mentioned agencies visited the field test during GPS data collection, offering their insights, support, camaraderie, and knowledge of local dining spots!

Operations Center Operation

The field activities were run out of an Operations Center. It is at the Operations Center that daily morning briefings are given to the Rover teams, visitors are briefed, and data processing is accomplished. It also serves as a central site for the Manager of the field test and normal or emergency communications with the staff. As a result of the experience of running an Operations Center for the Hampshire Field Test, significant changes were made for the Newberry Field Test.

In the Newberry Field Test a desktop PC was used as the processing computer; during the Hampshire Field Test processing was done on a laptop. The desktop’s software included GeoResearch, Inc.’s software for performing differential correction and for translating the resulting data into ESRI shape files. The software also included a copy of ArcView™ GIS and general office software such as a word processor.

A small, page-size, color ink-jet plotter was part of the equipment at the Operations Center. It was used to make maps from ArcView and to provide output from the word processor software. This proved to be invaluable by allowing each Rover team to have plots of the processed, readable data and to provide briefing items for visitors.

Although not integrated as part of the test procedures, a digital camera was available that allowed recording of examples and actions of the field test.

Base Station Operation and Data Processing

The GPS Base Station Set-up Guide, previously mentioned, provided a guide to setting up and operating the base station, including the equipment, cables, tripod, receiver, and amplifier. The Guide explains how to determine the position of the antenna as a control point for use in differential correction. It describes the data files and their purpose. This Guide was an elaboration of the guidelines that were used during the Hampshire Field Test. The Guide is probably adequate for future use.

Each day the base station system was initialized and a new file begun. This usually took place before the Rover teams began their data collection, and the base station collection continued until after the Rover teams came in from the field. Each evening the base station data was copied to a disk and brought to the Operations Center.

<u>Use of State-supplied Base Station Information</u>
Ordinarily the base station data would be used to apply differential correction to the Rover collected data. However, in the Newberry Field Test the opportunity arose to use base station data of much higher precision from the South Carolina Geodetic Survey—approximately +/- 10 centimeters. Therefore, data was collected from the Census Bureau base station, but South Carolina Geodetic Survey base station data was used for the differential correction work.

It turned out that this double collection was useful. Since the South Carolina Geodetic Survey data was not available over the weekend, the Census Bureau base station data was used to do differential correction so the Rover teams could see where they had worked. The following week GRaSS went back and applied the South Carolina Geodetic Survey data so all files would have the same accuracy.

Areas Requiring Further Consideration
As a post-field exercise, participating staff was invited to evaluate the field test and offer suggestions for improvement. The majority of comments received can be grouped into the following concerns:

- *GeoResearch, Inc. Software*
  Specifics on software functionality, errors encountered, and suggested modifications and enhancements to assist the Census Bureau in GPS data collection.
- *Maps*
  The usability of and suggested improvements for both the Geography Division's TIGER Mapping Branch (TMB) generated paper products and the TIGER/Line background maps displayed by GeoResearch, Inc. software.
- *Vehicles and Drivers*
  Accolades to the local drivers and their State 4-wheel drive vehicles (both courtesy of the South Carolina Geodetic Survey.)
- *General Equipment*
  Documenting minor problems encountered using miscellaneous general equipment.
- *Procedures and Manuals*
  Targeting procedures that require continued clarification.
- *Base Station*
  Hints on improving the placement and usability of the base station.
- *Dress Rehearsal*
  Ideas on improving the DR.
- *Team Member Knowledge*
  General knowledge was good, but more would be better!

Appendix B lists detailed comments (excluding those concerning GeoResearch, Inc. software) received from staff after the Newberry Field Test.

Additionally, the following recommendations by the Operations Center Manager apply to operating an Operations Center:

1. Provide a separate room or building for the Operations Center. For the Newberry Field Test a large motel room was used with couches and tables to one side and sleeping and bath to the other. This did not prove very satisfactory in that work in the Operations Center could not continue while the room occupant wanted to sleep.
2. Establish the base station antenna as close as possible to the Operations Center. While the antenna site must be secure because the base station data are read and stored by a laptop (or desktop), it should be as convenient as possible to go to for data copying and to assure that the system is still collecting. The best arrangement would be for the base station antenna and associated computer to be in the same location. However, the antenna cannot share a computer with the processing activity because any failure of the processing software may interrupt data collection from the antenna.
3. Change the processing procedures so that the Rovers are provided with feedback on their work that is no later than one day. This would help the Rovers to identify areas that must be redone and to change their collection plans to do so.
4. All manuals should emphasize that the source of the differential correction data should be consistent. If not, the resulting differentially corrected GPS points may be significantly different from each other at places such as road intersection points.

Status of Objectives

The status of the listed objectives is:

<u>GRaSS-conducted field test objectives</u>

- *Collect GPS positional data for use in the GPS Test Project*

  GRaSS's goal for Newberry County was collection of GPS data for 100% of intersection anchor points, road centerlines, and housing units. Data collection was accomplished in two phases. The Rovers first collected all intersection anchor points and road centerlines in their assigned areas. During the first "pass" Rovers additionally annotated on paper maps the streets having housing units. When intersection anchor point and road centerline collection was complete, the Rovers began a second "pass" of their assigned areas and collected GPS data for the housing units.

  At this point, having completed only a preliminary review of the data, it appears the intersection anchor points and road centerline collection objective was achieved. It is interesting to note that many of the roads in TIGER/Line '97 no longer exist (these roads were dirt logging roads, did not lead to housing units, and have not been maintained as passable roads.) Some roads in TIGER/Line '97 were not collected as they were gated; project specifications were that local guidance (usually from the South Carolina Geodetic Survey drivers on the Rover teams) be used to determine whether Rovers should enter a gated community; in most instances local guidance was to not enter and collect gated areas. Road centerlines were collected to a gate and the road centerline segment commented accordingly.

There may be a problem with the first day's collection of road centerlines, similar to the problem encountered during the Hampshire Field Test. Parameters for the project appear to have been incorrectly set, consequently road centerlines collected during the first day may be single, rather than double, precision. GRaSS is investigating the possibility of using this data for completeness, but has not made a final decision.

A solid start at collecting housing units was made. Housing units were collected in approximately ½ the geographic area of Newberry County. The percentage of collected housing units is unknown. A problem using the GeoLink PowerMap software was encountered during housing collection. Housing unit locations appear positionally incorrect, being aligned in perfectly straight lines rather than along roads. GRaSS is working with GeoResearch, Inc. to determine the cause of and ways to possibly correct the data. The problem appears to have occurred randomly.

- *Develop procedures, and ultimately specifications, for collecting GPS positional data on a national basis.*
  A concerted effort was made to develop procedures and enhance specifications prior to departing for the Newberry Field Test. Excellent suggestions for additional improvements were received from the Rovers in their post-field operations comments. GRaSS will update manuals based on the comments and add a new manual concerning pre-deployment planning.

- *Evaluate the GeoLink PowerMap software of the CRADA partner, GeoResearch, Inc. for use in Census GPS positional data collection activities and work with GeoResearch, inc. to enhance GeoLink PowerMap to improve its ability to collect GPS positional data for Census specific applications.*
  GRaSS met with GeoResearch, Inc. at Geography Division Headquarters for a debrief following the Newberry Field Test. GRaSS informally provided GeoResearch, Inc. with a copy of staff comments detailing specific problems encountered during GPS data collection. GRaSS is preparing a formal report to GeoResearch, Inc., and is prepared to work with GeoResearch, Inc. in determining the cause of problems encountered.

  GeoResearch, Inc. indicated during the debrief that their staff is working on developing an extension to the existing CRADA, focusing on their possible role in contributing to spatially improving TIGER. GeoResearch, Inc. will be submitting an addendum to the CRADA for Geography Division review in the near future. GRaSS is prepared to work with GeoResearch, Inc. to further the stated goals of the CRADA.

- *Gain firsthand experience in using GPS technology to collect positional data.*
  This objective was met. All junior members of GRaSS participated in the field test as Rovers collecting data using GPS technology.

Newberry Field Test-specific Objectives (Additional)

- *Compare Geographic Data Technology, Inc.'s (GDT) Dynamap/2000 dataset with TIGER/Line '97 and on-site observations (ground truth) to determine the currentness and correctness of both datasets and to use the resulting information to develop strategies for future updating of the respective datasets.*

  GDT offered its Dynamap/2000 dataset as a background dataset for use during the Newberry Field Test. Prior to the field test, GRaSS determined Dynamap/2000 lacked the necessary links to the TIGER database (TIGER/Line IDs) for use in its long-term objective of improving the spatial quality of TIGER. GRaSS did, however, overlay TIGER/Line '97 and Dynamap/2000 and felt it would be of value, and further the CRADA between the Census Bureau and GDT, to investigate the types of differences between the datasets and to compare both datasets against one another and "ground truth".

  A GRaSS staff member was assigned the task of coordinating this activity. Prior to the Newberry Field Test, differences between the datasets were analyzed and categorized and areas for on-site investigation identified (Appendix C contains the initial findings of the catalogued differences and identified areas). A Rover team devoted one day in the field to investigating the areas. GRaSS is preparing a detailed report of its findings along with recommendations for strategies for future updating of each dataset.

- *Develop the costs and determine the benefits of using GPS positional data and compare these to the alternate of using Digital Orthophoto Quadrangle data (DOQs) for collecting anchor points (anchor points are pivotal to the proposed GRaSS methodology for spatially enhancing TIGER).*

  GRaSS obtained USGS DOQs for Newberry County. Procedures have been developed for using ArcView GIS to overlay the DOQs, TIGER/Line, and the file of TIGER/Line intersection anchor points; determine a spatially "correct" position of each intersection anchor point using the DOQ; and rate the "quality" of each determination. GRaSS has begun the intersection anchor point determination and is maintaining cost and staff time records. At the completion of the DOQ work, GRaSS will have two "corrected" anchor point files to compare to the file of TIGER/Line '97 anchor points: the first obtained from GPS data collection, the second obtained from DOQ analysis. Additional processing of the two anchor point files will be performed in parallel to achieve the first objective of the GPS Test Project (adjust the existing coordinates of TIGER). An analysis will be made of the quality of the DOQ anchor point file against that of anchor points gathered using GPS data collection techniques. GRaSS will additionally consider the costs and benefits of each method and make recommendations as to the viability of each method to met Census Bureau needs.

  Staff performing the DOQ anchor point location function have been instructed to additionally consider the use of DOQs as a possible method of obtaining information to achieve the second and third objectives of the GPS Test Project (improve the coordinates of TIGER and update features in TIGER).

GPS Test Project Objectives

GRaSS has developed a methodology to achieve the multiple goals of the GPS Test Project and is beginning the short-term tasks required to meet these goals. The short term tasks, as well as a brief summary of the long term tasks, is outlined in the following section, *Using the Newberry County, SC GPS Positional Data*.

Using the Newberry County, SC GPS Positional Data

Short Term Tasks

The following short term tasks are considered vital to accomplishing the first of the three GPS Test Project Goals and are currently assigned to GRaSS Staff:

- Newberry County ESRI shape and associated file creation.
    1. Process the individual traverse information into a single data file.
    2. Produce CDs containing complete sets of data. Four CDs are required, containing (1) alpha files (GPS collected data that has not been differentially corrected), (2) beta files (GPS collected data that has been differentially corrected data), (3) Base Arc files (differentially corrected, GPS collected data translated into ESRI shape and associated files), and (4) "Clean" ESRI shape and associated files (refer to the Database Edit task.)
- Database Edit (or clean-up)
    1. Identify and catalogue types of errors or potential processing problems in the Newberry County dataset (examples: zingers, inadequately filtered or weeded data, incorrectly recorded housing unit locations.)
    2. Develop and test rules for filtering point sequences (traverses). Consider acceleration between points, distance between points (segment length), and averaging within. Create a batch process to filter the data, develop a manual filtering procedure, and test the efficiency of each.
    3. Develop procedures to manually edit the errors (these are things a filter would not catch).
    4. Edit the data and develop a "clean", workable dataset.
- Usable anchor points
    1. Determine rules for selecting the GPS collected (or DOQ) anchor points to use in correcting the existing TIGER/Line coordinates. This task requires an analysis of the distances between existing TIGER/Line intersection points and the collected anchor points. Additionally, determine rules for handling the excluded anchor points.
    2. Create a file of "usable" GPS collected (or DOQ) anchor points from all anchor points collected. A Visual Basic program is being developed for this.
- Triangulation Work
    1. Create a file of "sorted" usable GPS collected (or DOQ) anchor points, listing the records in order of smallest to longest.
    2. Build logical basis for adjustment operation: This requires determining "edge effect" rules, triangulation rules, and algorithms.
    3. Triangulate the TIGER/Line file based on the adjustment of existing TIGER/Line intersection anchor points to usable GPS collected (or DOQ) anchor points. The theory to be tested is that this should improve the spatial accuracy of the TIGER/Line coordinates without effecting topology.

- Adjustment Check
    1. Develop rules for checking for topological errors introduced by the coordinate adjustment.

Longer Term tasks

Longer-term tasks will lead to the second and third GPS objectives (improve the coordinates of TIGER and update features in TIGER). Because these objectives depend upon the accomplishment of the first objective, GRaSS is focusing on accomplishment of the first objective before beginning detailed development work on the remaining tasks.

Timeframe

GRaSS hopes to accomplish the first objective of the GPS Test Project in a six to eight month time frame.

The Need for Future Field Tests

At this time, GRaSS does not anticipate a need for additional field testing. Once the data from Newberry County has been analyzed and the methodology for improving the spatial accuracy of TIGER proven GRaSS will reevaluate a need for more field tests.

A decision must still be made concerning the completion of Hampshire County data collection. It was once felt a limited amount of time would be required to complete GPS data collection and meet the Census Bureau’s promise of providing the collected data to State and local officials. However, the quality of the collected data indicates extensive field work is required. If additional proving in of either procedures or methodology appears necessary, repeating the Hampshire Field Test should be considered.

Appendix A

Newberry County Statistics

Page 1 of appendix A

Page 2 of appendix A

Page 3 of appendix A

Page 4 of appendix A

Page 5 of appendix A

Page 6 of appendix A

Page 7 of appendix A

Page 8 of appendix A

Appendix B

Newberry County Post-Field Test Detailed Comments

Summary of GRaSS Recommendations

Maps

- The (TMB supplied paper) maps are cumbersome to use. It would be helpful to bind them into sets and tab the sheets by sheet number.
- Availability of both small and large scale (TMB supplied paper) maps is helpful.
- Having both small and large-scale (TMB supplied paper) maps was invaluable though coordinating a large BNA paper map and a series of individual map sheets was cumbersome.
- The scale of the maps used for annotation and reference was good, though the size of the (TMB supplied paper) index map made it somewhat difficult to work with. Folding the map several times to a more manageable size resolves this but does cause some ripping. In retrospect, a solution would be to cut the map to a size that could be attached to an 8 x 10" clipboard.
- Use of on-the-fly shape files as background maps (displayed by GeoLink PowerMap software) is beneficial to data collection. The shape files are more useful in showing progress than the annotated paper maps.

Vehicles and Drivers

- 4-wheel drive trucks (with drivers knowledgeable about this utility) significantly improve data collection and should be used.
- Use of 4-wheel drive vehicles is very beneficial.
- 4 wheel-drive proved useful many times.
- Vehicles with State-emblems provide a high "comfort level" to local residents and should be used.
- Use of a State vehicle is helpful, particularly during interactions with local residents.
- A pickup with a bench seat and a 'club cab' may be an ideal type of vehicle for the project. The bench seat is the important component as it allows for the storage of all maps on the front seat between the driver and operator.
- The speed of the vehicle. Sometimes it is difficult to enter all the data required for a particular feature before arriving at the next feature.
- As there is a "learning curve" for drivers (regarding the procedures, what the driver should look for, driving speed considerations), it is most efficient to have as little driver turnover as possible in the field.
- Having the same driver for an extended period of time is helpful, eliminating the need to constantly train drivers.
- Having a single driver is useful, as a minimum amount of training is required.
- It is helpful to allow a driver to switch roles and collect data for a few hours. Perhaps this should be done on the first morning to help the driver understand what kind of driving is required for successful data collection. Additionally, the operator experiences what it is like to be the driver.
- Local knowledge supplied by the South Carolina Geodetic Survey was very helpful.

General Equipment

- Before departing for the work site on the first morning of work, each Rover should test the power supply to ensure its compatibility with the vehicle.
- The manual range finders consistently lose calibration and are more trouble than they are worth.
- The mechanical range finders lost calibration frequently. A more accurate and reliable method is needed.
- Resolve the inverter (pocket socket) problems.

Procedures and Manuals

- Rovers experienced problems in accurately identifying intersections (anchor points) on a first drive through. Develop a procedure whereby a rover can flag the last intersection collected as invalid to alleviate the problem of incorrect collections.
- Anchor points are not instantaneously dropped when the operator strikes the crossroad hotkey. Develop procedures to improve this. One possible solution is to collect the intersection slightly before entering it.
- There is confusion as to where comments should be attached to a road segment. Clarify the procedure in the rover manual.
- Clarify confusion on traversal of side roads and private roads and the rationale behind the procedures (improve the procedures). Clarification is required on collection of (1) small segments on side roads and private drives and (2) long private drives.
- Develop collection rules addressing what to do when a road cannot be located.
- Improve collection methodology for inaccessible roads.
- Improve collection methodology for urban data, particularly for structures. An area of concern leading to capture inconsistencies is housing units on top of businesses.
- Collection of main roads during a single traverse worked well.
- Distribute information on features requiring re-collection to rovers on a more consistent and continual basis to avoid long distance driving within Block Numbering Areas (BNAs were used as work assignment areas).
- On the whole, road centerlines and housing units should be collected concurrently. This would eliminate duplication of effort. (This should not be done in limited situations such as on high speed roads or in city centers.)

Base Station

- Expand the documentation for operating the base station and retrieving data in the rover manual.
- Locate the base station at the Operations Center. The base station site and access to it in Newberry County were both inconvenient, resulting in time-consuming data retrieval. Ensure ease of entry to the base station in the future.
- Use the closest secure facility to locate base station-related equipment to expedite post-collection processing.
- Streamline the process of retrieving RYNEX (or any) data required for post processing purposes. Look into using data from a Census-operated base station and using external data only if differences in the results are significant.

Dress Rehearsal

- A dress rehearsal held immediately prior to fieldwork allowing hands-on experience is integral to the success of the fieldwork.
- It would be optimal to use the same version of the software during both dress rehearsal and in the field.
- Test and analyze the PowerMap system more thoroughly before entering the field (possibly in a county adjacent to Geo Division headquarters.)
- Software problems during dress rehearsal resulted in team members not being completely comfortable at the beginning of fieldwork. Further instruction in data capture rather than recovery from loss of GPS signal would have been helpful.
- Dress Rehearsal should be designed to help each Rover deal with his/her problems in the field rather than the following day or several days later (i.e. editing mistakes, redos.)

Team Member Knowledge

- Stratification of knowledge within the project team should be alleviated. It would be beneficial for all team members to have a working knowledge of all aspects of the project (i.e. hotkey programming, data translation, base station management.) This does not mean everyone must be involved in every aspect of the project, simply that they be able to assist in accomplishing tasks as required.
- In theory, there is justification and a need for the training of all field members on the post-collection data processing. However, data management is one of the keys to the success of this project and multiple data managers are not optimal. Some duties are sharable (example – the collection of RYNEX data) but data management should remain with a few team members for the sake of consistency. In this area the project was probably understaffed.
- Train all team members in all aspects of the process (for example, post processing) to heighten understanding and prepare them for unexpected situations.
- Knowledge of all base stations used and their unique data characteristics by all team members would be helpful. Additionally, it would be useful for all team members to know how to set up/close down the Census base station.

Miscellaneous

- Resolve understaffing. A continually manned Operations Center is necessary.
- Streamline project preparation for future projects now that the process has been extensively beta-tested.
- Personnel having local knowledge are beneficial to the project.
- It is vital to test the loaded PowerMap “project” so that serious errors can be corrected prior to fieldwork.
- Logistical preparations for future GPS tests should expend time and energy necessary to include local, State, or other participation for in-field collection.
- Dividing the BNAs into work assignment areas proved beneficial.

Appendix C

Objectives, Catalogued Differences and Ground Truth Areas
TIGER/Line '97 and GDT's Dynamap/2000

Purpose

The purpose of the activity is to compare two datasets against one another and on-site observations (ground truth) to determine the currentness and correctness of the datasets. The information will be used to develop strategies for future updating of each dataset.

Overview

Two datasets of Newberry County, South Carolina are being compared. The datasets are the TIGER/Line '97 Newberry County file (TIGER97) and the GDT file of Newberry County (GDT). The project has three parts:

- Compare the datasets to one another using a Geographic Information System (GIS). Identify types of differences and areas having major discrepancies.
- Field-check several areas identified as having major discrepancies while collecting data using Global Positioning Systems (GPS) in Newberry County, South Carolina.
- Report on the comparisons between the datasets and ground truth.

Basics of Datasets

The TIGER/Line '90 file was the source file for both TIGER97 and GDT.

TIGER97 has been updated using enumerator updates (Census enumerators reporting changes encountered during on-site enumeration) and local input (local officials submitting updates to Census staff). Both types of updates were accepted without question.

GDT, Inc. staff reports continuous updates to its dataset have been received from local officials. As updates to GDT have been on the content, rather than positional accuracy, of features, GDT, Inc. staff believes GDT has essentially the same positional accuracy as TIGER/Line '90.

Part 1 – Dataset Comparison Using GIS

- Examine TIGER97 and GDT linear vehicular features.
- Observe the differences between the two datasets.
- Produce a qualitative summary of the kinds of differences between the datasets.
- Quantitatively describe the qualitative differences, and produce a quantitative summary of line segment, polygon, and data intersection differences.

Methodology

ArcView GIS was used. The two datasets were viewed as an overlay (one map over the other) to detect differences between the datasets. An initial determination was made that discrepancies existed and were significant. Several distinct types of discrepancies were noted.

Five regions of 25 square miles were chose for more in-depth study. The discrepancies within the five areas were classified into nine unique types describing differences between the datasets. Unique line segments and polygons within the five areas were totaled.

Discrepancies

Several types of discrepancies were repeatable and are classifiable. The nine types of discrepancies are:

Extensions – A road segment exists in both datasets; however, the road segment ends in one dataset but continues (or "extends") in the second. The extension does not create a new polygon. See Figure 1.

Scalar differences and shifted lines – Similarly shaped road segments exist in both datasets; however, the road segments begin and end in significantly different spaces or appear to take up a significantly different amount of space. Scalar differences and sifted lines do not create a new polygon in either dataset. See Figure 2.

Returning loops – A road segment exists in one dataset only. It originates from and returns to a road segment found in both datasets. The return loop creates a new polygon. See Figure 3.

Generic branches – A road segment exists in one dataset only. It originates from a road segment found in both datasets. The generic branch does not create a new polygon. See Figure 4.

Crossing branches – A road segment exists in one dataset only. It "crosses" a road segment found in both datasets. Technically, the road segment is two road segments beginning at a common intersection point from the road found in both datasets. The crossing branches may create a new polygon. See Figure 5.

Branch of branch – A road segment exists in one dataset only. It originates from a road segment that exists only in the same dataset. Branch of branch may create a new polygon. See Figure 6.

Connector – A road segment exists in one dataset only. It is a "connection" between two road segments that exist in both datasets. Connector creates a new polygon. See Figure 7.

Interchange area – An interchange exists in one dataset only and none of the road segments comprising the interchange area have been included in any of the above-listed discrepancies. The Interchange area is in and of itself. Interchange area creates one or multiple polygons. See Figure 8.

Figure 10

V-Intersection-An intersection exists and both roads are included in both dataset, however there is an additional short line segment included in one dataset producing a 'V' at the point of intersection. See Figure 9.

Figures and qualitative data concerning the above discrepancies are found at the end of this analysis.

Differences in Line segments, polygons, and dataset intersections.
Notable differences in the numbers of line segments and polygons were found between TIGER97 and GDT.

Line segments
What appears to be a single road may be comprised of many line segments. One line segment is terminated at the intersection with any other line segment; at this point another line segment begins.

The following example contains six line segments:

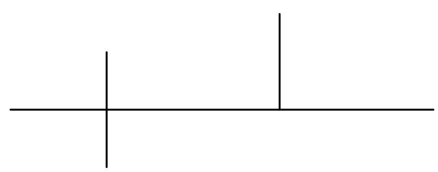


Figure 10

Exclusive line segments are line segments which exist in one, not both, datasets.

There are 124 exclusive line segments in TIGER97 (51% of total).
There are 121 exclusive line segments in GDT (49% of total).

Polygons
New polygons are created when an exclusive line segment "closes off" or creates a new area.

An example of exclusive line segments that do not create a new polygon.

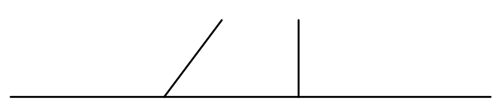


Figure 11

An example of exclusive line segments that do create a new polygon (note, it only took one more exclusive line segment to create a new polygon!)

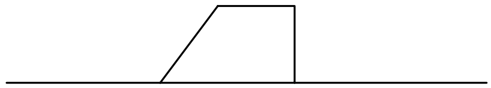


Figure 12

There are 15 polygons created by exclusive line segments in TIGER97 (29% of total).
There are 37 polygons created by exclusive line segments in GDT (71% of total).

Dataset intersections

Dataset intersections are created when an exclusive line segment intersections with a line segment existing in both datasets or exclusive to the other dataset (upon a merge of the two datasets.)

The dataset intersection tabulations do not include data intersections that are a result of *Scalar differences and shifted lines* discrepancies. Dataset intersections appear visually to be "stray lines" and look like lines digitized beyond what is intended.

An example of a dataset intersection. Both datasets contain a line segment (double line). The single line is an exclusive line segment, which intersects with the common line segment.

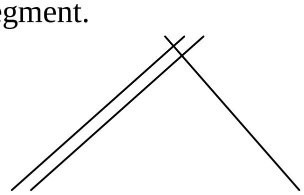


Figure 13

There are 7 dataset intersections in the sample areas.

Part 2 – Field-Checking

Data discrepancies will be evaluated on-site by identifying ground truth.

- Census staff will be in Newberry County, South Carolina to GPS road and housing units.
- While in Newberry County, Census Staff will, in addition to collecting data using GPS, visually inspect and report the ground truth in the five regions identified in Part 1. This information will be used, along with GPS data, to evaluate the currentness and correctness of the two datasets.

Methodology

The Census Staff will be provided with maps of the five regions. Census Staff will "correct" the map to reflect ground truth.

Discrepancy Figures

Extensions

Figure 1

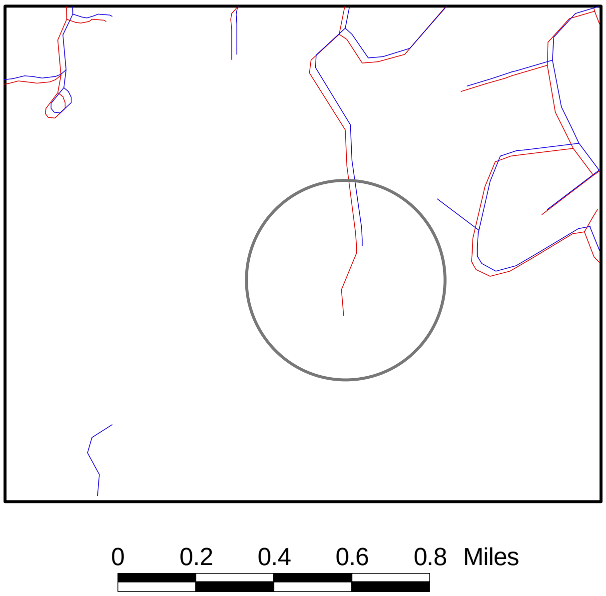
Extension-type discrepancies found exclusively in:

TIGER/Line '97 40%

GDT 60%

Extension-Type discrepancies as a percent of total discrepancies:

5%



Scalar-Differences and Shifted Lines

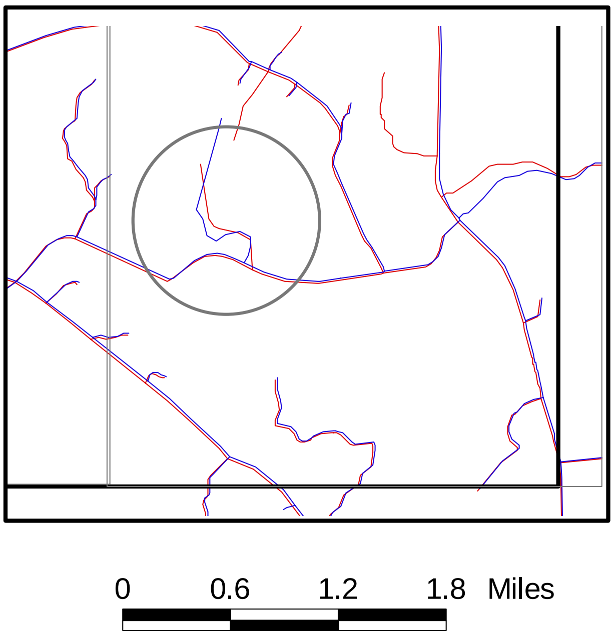
Figure 2

Scalar-differences and shifted line discrepancies found in the two data sets:

15

Scalar-differences and shifted line discrepancies as a percent of total discrepancies:

8%



Returning Loops

Figure 3

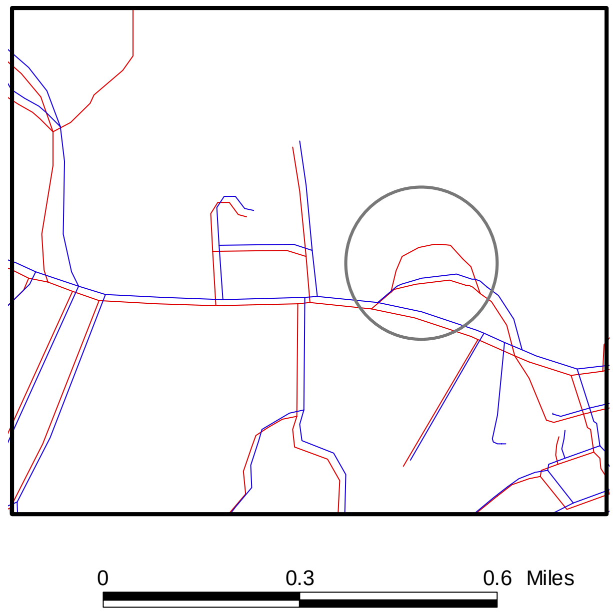
Returning loop-type discrepancies found exclusively in:

TIGER/Line ’97 60%

GDT 40%

Returning loop-type discrepancies as a percent of total discrepancies:

5%



Generic Branches

Figure 4

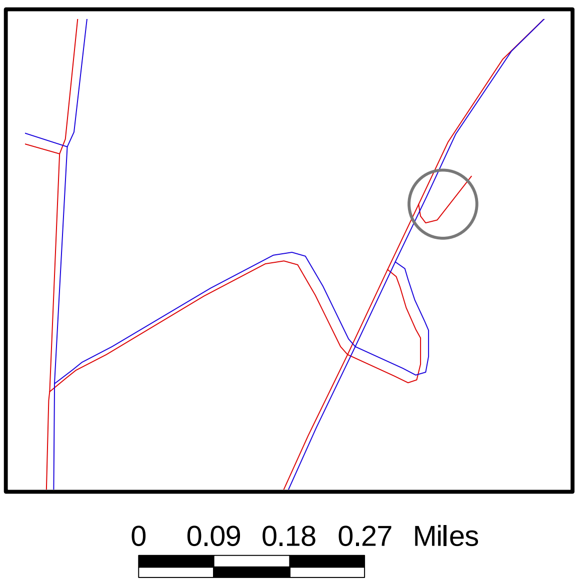
Generic branch-type discrepancies found exclusively in:

TIGER/Line ’97 50%

GDT 50%

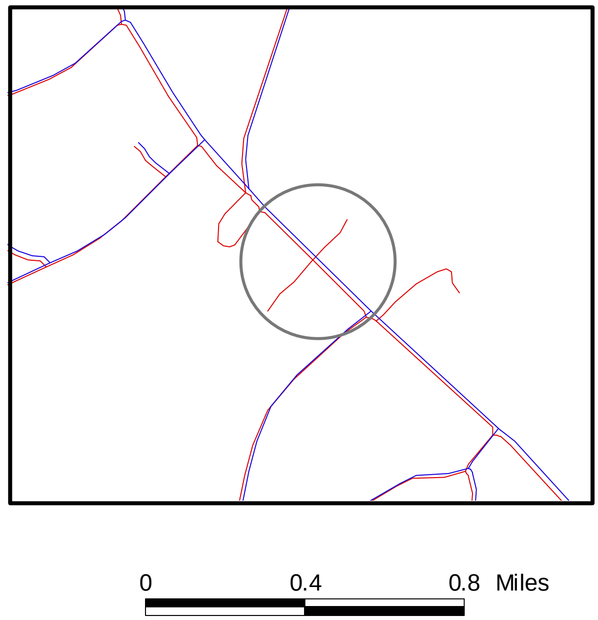
Generic branch-Type discrepancies as a percent of total discrepancies:

49%



Crossing Branch

Figure 5



Crossing branch-type discrepancies found exclusively in:

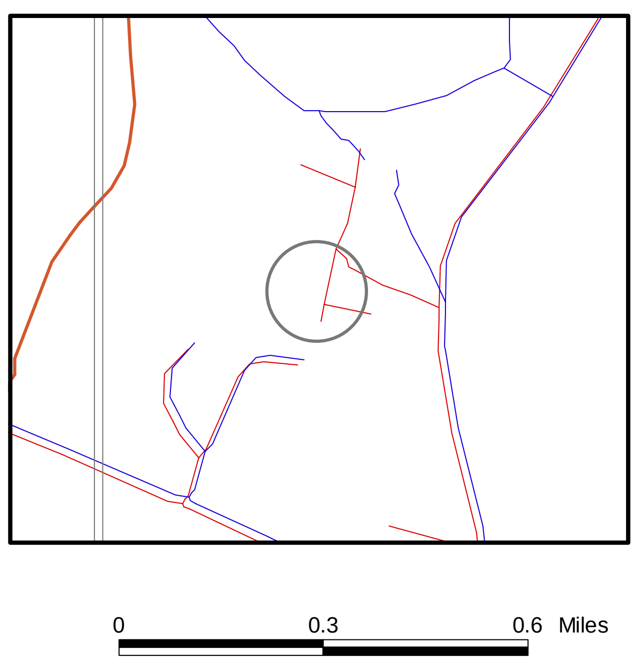
TIGER/Line ’97	40%
GDT	60%

Crossing Branch-type discrepancies as a percent of total discrepancies:

2%

Branch of Branch

Figure 6



Branch of branch-type discrepancies found exclusively in:

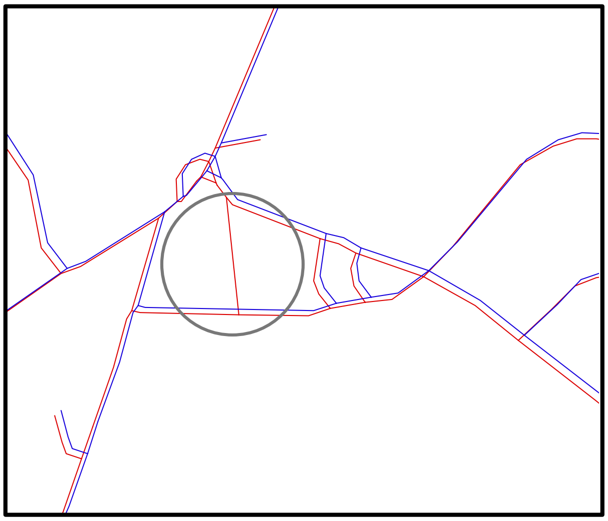
TIGER/Line ’97	63%
GDT	37%

Crossing Branch-type discrepancies as a percent of total discrepancies:

22%

Connector

Figure 7



0 0.2 0.4 0.6 Miles

Connector-type discrepancies found exclusively in:

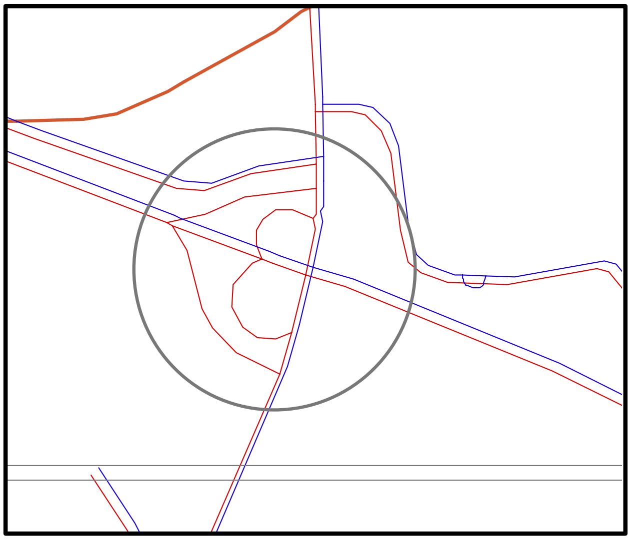
TIGER/Line ’97	29%
GDT	71%

Connector-type discrepancies as a percent of total discrepancies:

4%

Interchange Area

Figure 8



0 0.2 0.4 Miles

Interchange area discrepancies found exclusively in:

TIGER/Line ’97	0%
GDT	100%

Interchange area discrepancies as a percent of total discrepancies:

2%

V-Intersection

Figure 9

V-type connector discrepancies found exclusively in:

TIGER/Line ’97	0%
GDT	100%

V-type connector discrepancies as a percent of total discrepancies:

2%