NOAA Technical Memorandum NOS NGS 88

GEOCON 1.0 Operating Instructions and User Guide

Dennis Milbert
Silver Spring, MD
July 2012

Program GEOCON
Version 1.0
July 6, 2012

Operating Instructions

Author:
Dr. Dennis Milbert
Chief Geodesist, Retired
Rockville, Md

1. Introduction

GEOCON performs three-dimensional coordinate transformations between NAD 83 (HARN) coordinates and NAD 83(NSRS2007) coordinates. GEOCON also issues information about the quality of the transformation at each point, and notifications regarding poor quality results.

GEOCON is written as a simple demonstration program that recognizes the NGS Blue Book *80* and *86* position records. The coordinate transformation and the associated quality values are obtained through biquadratic interpolation within a series of grids. Third party applications may obtain identical results if they use the same grids and algorithms.

GEOCON employs high resolution grids (1' by 1') to obtain unprecedented fidelity in modeling coordinate differences. Frequently, one may see that the reported quality is extremely high (e.g. 1 cm or better), and could be considered comparable to a geodetic readjustment of survey measurements. Nonetheless, the National Geodetic Survey considers actual readjustment of survey measurements, and not coordinate transformations, as "best practice".

2. Example GEOCON Session

```
C>geocon
 program geocon - v1.0 - 2012jul06
----------------------------------------
input  80/86 file  (input) : bb8086
output 80/86 file (output) : zout
quality file name (output) : zqual
clip file name    (output) : zclip
notification file (output) : znote
 Region number 1 -- CONUS
 Region number 2 -- Alaska
 Region number 3 -- Puerto Rico/Virgin Islands
 ---------------------------------------------
Enter region number 1 through 3   : 1
Convert from HARN into NSRS2007? (Y/N): n
 Program will convert from NSRS2007 into HARN
 ----------------------------------------
 transformation grids successfully loaded
         29  points interpolated
          2  notifications issued

Press ENTER key to stop program:
```

3. Operation

Assume that geocon.exe and its supporting grid files have been installed in a folder (or subdirectory) named "MyDir" on a PC with a Windows operating system. Open the "MyDir" folder. Copy or move a file of *80* and *86* position records into the "MyDir" folder. Assume, for the sake of this example, that the *80* and *86* records are named "bb8086". Simply double left click on the "geocon" file in "MyDir" (full file name is "geocon.exe"). This will create a command prompt console window, and will automatically start GEOCON in the console window.

GEOCON will ask for a series of file names. The first file name is the input file of the *80* and *86* records to be transformed. The remaining file names are output files that will be created by GEOCON. Note, do not use an output file name of your grid or info files. These files are write-protected, and can not be overwritten. On output, GEOCON will overwrite a file if you reuse a file name and it is not write-protected. File names use standard Windows OS conventions, but are limited to 88 characters. File name extensions (e.g. "zout.txt") are supported if the user provides them.

After prompting for the file names, GEOCON will prompt for the region to be used for the coordinate transformation. The regions are:

1	CONUS	24-50N, 66-125W
2	Alaska	46-77N, 128-194W
3	Puerto Rico/Virgin Islands	17-20N, 62-68W

Enter the region number for the points in the "bb8086" file. They are in the conterminous U.S., so enter a "1". Points in the selected region will be transformed, and points outside the region will be written to the clip file. If the wrong region is selected, then all of the input points will be clipped, and no points will be transformed or written into the output file.

The last prompt is for the direction of the coordinate transformation. The prompt is "Convert from HARN into NSRS2007? (Y/N):". If you wish to transform in that sense, then respond with a "y" or "Y". If, on the other hand, you wish to transform from NSRS2007 into HARN, then respond with an "n" or "N". GEOCON will then confirm the direction of the transformation in the subsequent line.

GEOCON will then issue a message indicating that it has successfully loaded the coordinate transformation grids. GEOCON will transform the points in the input file and write them into the output file. For each transformed point, an associated transformation quality record will be written into the quality file. If a point is clipped, due to falling outside the designated region boundaries, the untransformed point will be written into the clip file. Supplemental information regarding potential sources of poor quality results will be listed in the notification file if appropriate.

Finally, GEOCON will indicate its completion by issuing the prompt:

```
Press ENTER key to stop program:
```

When you press the enter key, GEOCON will stop, and the command prompt console window will close. The output files created by GEOCON are formatted, ASCII text files, and may be inspected by any suitable application, such as Notepad.

Note it is also possible to manually launch a command prompt console window within Windows. If this is done, then use the "cd" command to maneuver to the "MyDir" subdirectory. One may then start GEOCON by typing "geocon" at the command prompt. If a console window is manually created, then it will not close when GEOCON stops.

A footnote about Hawaii:

There is no official NSRS2007 coordinate for Hawaii. However, a control point, KOKEE, (PID = TT3487) was inadvertently published in the NGS database as NSRS2007. This NSRS2007 position was subsequently removed from the NGS database in 2012. Hence, there are no GEOCON coordinate transformations for Hawaii.

4. The Input *80* and *86* File

This file must be provided by the user. The file consists solely of *80* and *86* records as described in Chapter 2, Horizontal Observation (HZTL OBS) Data, of the Input Formats and Specifications of the National Geodetic Survey Data Base (the NGS Blue Book), Volume I - Horizontal Control. Further detail may be found below in the sections "Format of Blue Book *80* Control Point Record" and "Format of Blue Book *86* Orthometric Height, Geoid Height, Ellipsoid Height"

GEOCON requires the input file to consist solely of *80* and *86* records, entered pair wise. That is, *80* will always be the odd numbered records, and *86* records will always be the even numbered records. And, any given *86* record will be associated with the horizontal position of the *80* record immediately preceding it.

It is required that all the records in the input file be in the same coordinate set. And, those input coordinates must be either NAD 83(NSRS2007) or NAD 83(HARN). It is absolutely critical that the response to the prompt

```
Convert from HARN into NSRS2007? (Y/N):
```

is correct. If your input coordinates are HARN, you must respond "y" (or "Y"). And, if your input coordinates are NSRS2007, you must respond "n" (or "N").

It is not required that all of the input records be in the same geographic region. However, it may be useful to organize data in that way. Any record pairs that do not fall in the geographic regions defined above will be clipped, and written to the clip file.

5. The Output *80* and *86* File

This file is generated by GEOCON, and consists of the transformed *80* and *86* records. Provided that all the input data were in the selected geographic region, and that there were no clipped points, then there will be a one-to-one correspondence between the input records and the output records. GEOCON does not modify the orthometric height or the geoid height in the *86* record.

Each specific transformation is obtained through biquadratic interpolation from the associated coordinate transformation grid (latitude, longitude, and height). Those grids were obtained by fitting splines in tension to sets of coordinate differences in the various regions. In cases of clusters of points that fell within 1' x 1' cells, a modified median filter was used to select the point passed to the spline model. More detail on the fitting process may be found in the GEOCON Users Guide.

6. The Transformation Quality File

This file is generated by GEOCON, and its records express the quality of the coordinate transformation. While these records (*94*) appear to be Blue Book records, they are not defined as Blue Book records. The format of these records may be found below in the section, "Format of *94* Transformation Quality Record".

There will be a one-to-one correspondence between each *94* record and each *80*/*86* pair in the output 80/86 file. And, assuming that SSN's (Station Serial Numbers) were provided in the input *80* records, those SSN's can be used to distinguish each *94* record.

The quality values in latitude, longitude, and height are obtained through biquadratic interpolation from the associated transformation error grids. Those grids were obtained by fitting splines in tension to sets of cross-validation errors in the various regions.

Cross-validation is a statistical technique to assess how accurately a model may predict. Imagine a data set of 69540 latitude differences. One could withhold one difference, and model a unique grid with the remaining 69539 differences. One can then compare the withheld difference with the predicted difference interpolated from the unique grid. While it is prohibitive to do this for every single point in a large data set, one may use this process and randomly select disjoint sets of points to reduce the computational effort. The GEOCON errors were obtained from 69-fold cross-validation

on sets of 69540 coordinate differences. More detail on the cross-validation may be found in the GEOCON Users Guide.

In cases of clusters of cross-validation errors that fell within 1' x 1' cells, a worst case algorithm was used to select the point passed to the spline model. First, a modified median filter was used to select the most likely error. Then, the error in the cluster with the greatest difference in magnitude from the median error was passed to the error spline model. More detail on the error fitting process may be found in the GEOCON Users Guide.

Note that for the coordinate transformations, a median condition is used to develop the grids. But, for the transformation quality, a worst-case condition is used to develop the grids. This provides the most likely transformation and a conservative estimate of its error.

Also note that the reported transformation quality values are signed quantities. These are reported in the sense of: actual coordinate difference – gridded coordinate difference. (As described above, the gridded coordinate difference is obtained when the point in question is withheld from the grid computation.) It was decided to retain the signs to show cases of systematic under or over prediction which may help in data problem diagnosis.

The transformation quality values should be treated as systematic error. One should add the absolute value of the quality number to the network accuracy of the pre-transformed point. A coordinate transformation never improves the quality of the underlying data.

7. The Clip File

The clip file contains copies of input *80*/*86* record pairs that fall outside the selected region. If all points are within the selected region, the file will be empty. The number of records in the clip file plus the number of records in the output transformed point file should equal the number of records in the input file.

Note that if the input file has points from a mixture of regions, then one may use the generated clip file as input to a subsequent run of GEOCON with a different selected region. In this way, one may easily transform all the input points.

8. The Notification File

This file contains supplemental information that explains the source of large transformation errors. This information is not issued for all instances of lower quality transformations. Rather, it is issued when a cluster of points is nearby, and one member of that cluster is a worst-case contributor to the transformation quality grid. Thus, most

members of a cluster of points may be very well modeled by the coordinate transformation. But, if there is an outlier in the cluster, that error will be expressed in the error grid. The notification file gives specific information on the reason why.

For example, consider the transformation of:

```
003480*80*5084NOT METOMEN GPS           40115697143N104433324639W 31273AWIBA
      *86*5084   312725A  N88       -35761W  1503844A32A
```

This generates a transformation quality record:

```
   *94*5084          0.00137     4.22   0.01766     41.66     -9.54
```

Where one can see sizable longitude and vertical error values (41.66 cm and -9.54 cm).

The notification file issues the messages:

```
003480*80*5084NOT METOMEN GPS           40115697110N104433324608W 31273AWIBA
Note - poor quality due to LL1465 - unmodeled hztl. error:    72.8 cm.
      *86*5084   312725A  N88       -35761W  1503865A32A
Note - poor quality due to LL1477 - unmodeled vert. error:   -12.6 cm.
```

Thus, one can see the longitude transformation quality of 41.66 cm was issued due to the 72.8 cm coordinate modeling problem at LL1465. And the height transformation quality of -9.54 cm was due to LL1477.

Notification file messages are informational in nature. There is no specific action a user need take due to them. However, if the transformation error exceeds the desired error budget, then the user is strongly encouraged to investigate exactly how the user's input data were connected to the geodetic control network. More detail on this issue may be found in the GEOCON Users Guide.

Notification messages are generated through input *80*/*86* records, the appropriate transformation error grid, and the appropriate information file. Information files are provided by NGS, and their format is described below in the section, "Format of Information File Records". The information files are plain ASCII files that list troublesome points that are members of clusters. These points were dropped when creating the coordinate transformation grids. And, they were retained when creating the transformation quality grids.

Notification messages are issued when an input point has horizontal or vertical error of 5 cm or more and is within about 5 km of a troublesome point.

9. Format of Blue Book *80* Control Point Record

These records are standard ASCII text with fixed column formatting. They are a legacy format more fully described in Chapter 2, Horizontal Observation (HZTL OBS) Data, of the Input Formats and Specifications of the National Geodetic Survey Data Base

(the NGS Blue Book), Volume I - Horizontal Control. This information is available online at http://www.ngs.noaa.gov/FGCS/BlueBook/

GEOCON only considers columns 7 through 14 and 45 through 69 of this format. It is possible to not fill the remaining fields at all, or to fill them with alternative information. However, it will be difficult or impossible to interpret transformation quality and any supplemental vertical notification messages without a Station Serial Number (SSN).

GEOCON requires the input file to consist solely of *80* and *86* records, entered pairwise. That is, *80* will always be the odd numbered records, and *86* records will always be the even numbered records. And, any given *86* record will be associated with the horizontal position of the *80* record immediately preceding it.

```
CC 01-06    SEQUENCE NUMBER. OPTIONAL. RIGHT JUSTIFIED. INCREMENT BY 10
            FROM THE PREVIOUS RECORD.
CC 07-10    DATA CODE. MUST BE *80*.
CC 11-14    SSN. SEE PAGES 1-1, JOB CODE AND SURVEY POINT NUMBERING AND
            2-12, ASSIGNMENT OF STATION SERIAL NUMBERS.
CC 15-44    STATION NAME. MUST NOT EXCEED 30 CHARACTERS. THE NAME OF A
            HORIZONTAL CONTROL POINT WITH PERIPHERAL REFERENCE MARKS
            AND/OR AZIMUTH MARKS MUST NOT EXCEED 24 CHARACTERS TO ALLOW
            FOR ADDING RM 1, RM 2, AND/OR AZ MK TO THE NAME WITHOUT
            EXCEEDING THE 30-CHARACTER LENGTH LIMIT.
CC 45-55    LATITUDE. DEGREES, MINUTES, SECONDS (DDMMSSsssss).
CC 56       DIRECTION OF LATITUDE. RECORD CODE "N" FOR NORTH OR CODE
            "S" FOR SOUTH.
CC 57-68    LONGITUDE. DEGREES, MINUTES, SECONDS, (DDDMMSSsssss).
CC 69       DIRECTION OF LONGITUDE. RECORD CODE "E" FOR EAST OR CODE
            "W" FOR WEST.

THE *86* RECORD IS TO BE USED FOR THE ELEVATION (ORTHOMETRIC HEIGHT)
AND ELEVATION CODE, WHICH WERE FORMERLY DISPLAYED IN THE FOLLOWING TWO
FIELDS.

CC 70-75    BLANK.
CC 76       BLANK.
CC 77-78    STATE OR COUNTRY CODE. IF THE CONTROL STATE IS LOCATED IN
            THE UNITED STATES/CANADA, ENTER THE CODE FROM ANNEX A FOR
            THE STATE/PROVINCE OR TERRITORY WHICH CONTAINS THE STATION.
            IF NOT, ENTER THE CODE FROM ANNEX A FOR THE COUNTRY WHICH
            CONTAINS THE STATION. SEE ANNEX A.
CC 79-80    STATION ORDER AND TYPE. REFER TO PAGES 2-35 THROUGH 2-38,
            STATION ORDER AND TYPE AND SEE ANNEX E.
```

Example:

```
003650*80*5120FRIENDSHIP S GPS          43500699703N088295718725W 24000KWI1A
     *86*5120   240003K  N88      -35880W   204115A32A
003660*80*5121VAN DYNE GPS              43521622332N088301532627W 24358KWI1A
     *86*5121   243581K  N88      -35960W   207618A32A
```

8

10. Format of Blue Book *86* Orthometric Height, Geoid Height, Ellipsoid Height

These records are standard ASCII text with fixed column formatting. They are a legacy format more fully described in Chapter 2, Horizontal Observation (HZTL OBS) Data, of the Input Formats and Specifications of the National Geodetic Survey Data Base (the NGS Blue Book), Volume I - Horizontal Control. This information is available online at http://www.ngs.noaa.gov/FGCS/BlueBook/

GEOCON only considers columns 7 through 10 and 46 through 52 of this format. It is possible to not fill the remaining fields at all, or to fill them with alternative information. However, it will be difficult or impossible to interpret transformation quality and any supplemental vertical notification messages without a Station Serial Number (SSN).

GEOCON requires the input file to consist solely of *80* and *86* records, entered pairwise. That is, *80* will always be the odd numbered records, and *86* records will always be the even numbered records. And, any given *86* record will be associated with the horizontal position of the *80* record immediately preceding it.

```
CC 01-06    SEQUENCE NUMBER. OPTIONAL. RIGHT JUSTIFIED. INCREMENT BY 10
            FROM THE PREVIOUS RECORD.
CC 07-10    DATA CODE. MUST BE *86*.
CC 11-14    SSN OF CONTROL POINT.
CC 15-16    BLANK
CC 17-23    ORTHOMETRIC HEIGHT. IN METERS (MMMMmmm).
CC 24       ORTHOMETRIC HEIGHT CODE. SEE FOLLOWING TABLES.
CC 25-26    ORTHOMETRIC HEIGHT ORDER AND CLASS. USE PUBLISHED VERTICAL
            ORDER AND CLASS, OTHERWISE LEAVE BLANK.
CC 27       ORTHOMETRIC HEIGHT NGSIDB INDICATOR. SEE FOLLOWING TABLES.
CC 28-29    ORTHOMETRIC HEIGHT DATUM. SEE FOLLOWING TABLES.
CC 30-35    ORGANIZATION WHICH ESTABLISHED AND/OR MAINTAINS THE
            ORTHOMETRIC HEIGHT OF THE CONTROL POINT. ENTER THE
            ABBREVIATION LISTED IN ANNEX C OR ON THE DATASET
            IDENTIFICATION RECORD.
CC 36-42    GEOID HEIGHT. ABOVE (POSITIVE) OR BELOW (NEGATIVE) THE
            REFERENCE ELLIPSOID. IN METERS (MMMMmmm).
CC 43       GEOID HEIGHT CODE. SEE FOLLOWING TABLES.
CC 44-45    BLANK.
CC 46-52    ELLIPSOID HEIGHT. IN METERS (MMMMmmm).
CC 53       ELLIPSOID HEIGHT CODE. SEE FOLLOWING TABLES.
CC 54-55    ELLIPSOID HEIGHT ORDER AND CLASS. SEE ANNEX G.
CC 56       ELLIPSOID HEIGHT DATUM. SEE TABLE, P. 2-85.
CC 57-80    COMMENTS.
```

Example:

```
003650*80*5120FRIENDSHIP S GPS          43500699703N088295718725W 24000KWI1A
     *86*5120   240003K  N88      -35880W   204115A32A
003660*80*5121VAN DYNE GPS              43521622332N088301532627W 24358KWI1A
     *86*5121   243581K  N88      -35960W   207618A32A
```

11. Format of *94* Transformation Quality Record

These output records are standard ASCII text with fixed column formatting. They are a new format. While they follow the general structure of the HZTL OBS Blue Book, they are not Blue Book records. The Station Serial Number (SSN) of this record is obtained from the ASCII contents of columns 11 through 14 of the input *80* record.

This record expresses the quality of the coordinate transformation by use of cross-validation errors. In cases of clusters, the worst-case error is reported. These values should be used to increase the base network accuracy of the pre-transformed coordinates.

```
cc 01-06  Sequence number.  Optional. See Blue Book
cc 07-10  Data Code.  Must be *94*
cc 11-14  SSN.  See Blue Book
cc 15-20  BLANK
cc 21-30  Latitude Error, arc-sec              9999.99999  f10.5
cc 31-40  Latitude Error, cm                   9999999.99  f10.2
cc 41-50  Longitude Error, arc-sec             9999.99999  f10.5
cc 51-60  Longitude Error, cm                  9999999.99  f10.2
cc 61-70  Ellipsoid Height Error, cm           9999999.99  f10.2
```

Example:

```
   *94*5120       0.00000      0.00   0.00001     0.02     -0.07
   *94*5121      -0.00005     -0.16  -0.00011    -0.25     -0.06
```

12. Format of Binary Grid Files

These input files are unformatted files, written as FORTRAN unformatted sequential files. As such, these FORTRAN records automatically contain prefix and suffix information that indicate the lengths of each record. The binary files were created on a PC, and have the "little-endian" representation of x86 architectures. The binary grid files are provided by NGS. They are a legacy form that has been used internally.

The first record is a header record that defines and georeferences the remaining grid data.

Header record elements:

glamn   geodetic latitude, positive N., deg., double precision floating point
glomn   geodetic longitude, positive E., deg., double precision floating point
dgla   delta latitude, positive N., deg., double precision floating point
dglo   delta longitude, positive E., deg., double precision floating point
nla   number of rows, integer
nlo   number of columns, integer
ikind   type code of grid data, integer
           ikind = 0 indicates the grid data are unformatted integer
           otherwise, the grid data are unformatted single precision floating point

Note that for the GEOCON files, `ikind` = 1 has been used throughout.

Grid records:

Each record corresponds to a row in the geographic raster grid. Thus, one will always have exactly "nla" unformatted sequential records. Each record will contain exactly "nlo" elements. And, since `ikind` = 1 for the GEOCON grids, each element will be a single precision (32 bit) floating point value.

Records are written sequentially from South to North. And, within each record, the elements are written sequentially from West to East (row major, column minor order). Thus, all the elements of the first row correspond to the geodetic latitude of `glamn`. Elements of the second row correspond to latitude `glamn + dgla`, and so on. Similarly, the first element of any given row corresponds to a geodetic longitude of `glomn`. The second element corresponds to a longitude of `glomn + dglo`, and so on. One may see that the maximum latitude of the grid is `glamn + (nla − 1) * dgla`. And, the maximum longitude of the grid is `glamn + (nla − 1) * dgla`.

Each element corresponds to the exact latitude and longitude implied by the row and column order in the set of sequential records. There are no cells, nor is there any registration to a cell corner. The elements are associated to an exact point. Of course, grid elements may be point values, or they may be obtained from some area averaging process.

For all coordinate shift grids, the shift is stored as NAD 83(NSRS2007) − NAD 83(HARN). For all transformation error grids, the error is stored as actual coordinate difference − gridded coordinate difference. Latitudes are considered positive North, longitudes are positive East, and heights are positive up. For latitude and longitude, the shifts and errors are in units of 0.00001 arc seconds. For heights, the shifts and errors are in units of 1 cm.

13. Format of Information File Records

These input records are standard ASCII text with fixed column formatting. They are provided by NGS. They are a new format. The records may be in any order. These records report points in clusters, where the coordinate shifts of those points do not match their neighbors. Information records are used to generate the GEOCON notification messages.

```
cc 01-15  Longitude, positive E., deg,      99999.999999999 f15.9
cc 16-30  Latitude, positive N., deg        99999.999999999 f15.9
cc 31-40  Value, see note for units         9999999.99      f10.2
cc 42-47  PID. (Permanent Identifier)       ASCII characters
```

Note: Value represents the difference between the actual coordinate shift and the predicted coordinate shift obtained from biquadratic interpolation from the gridded coordinate shifts. Units of "`Value`" are 0.00001 arc seconds for horizontal information files, and 1 cm for vertical information files.

Example (fragment from infoh.txt):

```
256.197346503   40.330059214   -331.58 LK0770
255.274098225   40.182492064   2357.81 LL1465
```

14. HARN Coordinates

There is no single HARN coordinate realization. The coordinate computations of the NSRS (National Spatial Reference System) underwent an explosion in the early GPS years. The name, HARN, itself, underwent evolution from High Precision Network (HPN) to High Precision Geodetic Network (HPGN) to High Accuracy Geodetic Network (HARN). Despite the name changes, they refer to a GPS backbone campaign observed to the highest standards. As the effort approached completion, GPS reduction software and models improved to the level that significantly better ellipsoid heights were obtained. This lead to an additional cycle of nationwide GPS HARN surveys, referred to as the FBN/CBN (Federal Base Network/Cooperative Base Network).

From the geodetic network standpoint, it would have been ideal to complete all the observations and re-observations, and then issue a new set of national coordinates. However the public demand for higher accuracy geodetic control could not be met if NGS waited for all the surveys to be completed. So, in 1990 distinct coordinate sets for least squares adjustments of one or more states began to be issued. NGS continued this process through the 15 year HPN, HPGN, HARN, HARN FBN/CBN evolution.

The coordinates in this period are expressed as a datum tag in the form: NAD 83(nnnn), where "nnnn" is a notional date of the GPS survey observations. It is important to note that the datum tag NAD 83(1994) in Utah refers to a different realization than NAD 83(1994) in Michigan.

In addition, a majority of the states were re-observed as part of the HARN FBN/CBN. To reduce the impact of the inevitable coordinate changes, a flexible, 5 centimeter horizontal coordinate tolerance was adopted. If newer horizontal coordinates did not change by more than 5 cm, then the old horizontal coordinates were retained. In these cases the older datum tag was retained, even though the ellipsoidal heights were *always* updated with the latest GPS surveys and reductions. In other cases, all horizontal and vertical coordinates were updated, and a new datum tag was issued. Note that one can not look at an older datum tag, such as NAD 83(1994) in Ohio, and immediately tell if the associated coordinates are traceable to an earlier HARN adjustment, or to the later HARN FBN/CBN phase.

GEOCON was developed with the latest coordinates available in the NGS database in both the NSRS2007 and the HARN coordinate sets.

To assist users in understanding HARN coordinate sets, and to help track the variety of coordinates, the principal datum tags are listed below. In some cases, more than one datum tag is listed. These are cases where a significant number of points were found in a secondary coordinate set. In these cases, the first coordinate entry represents the majority of the points, and the second coordinate entry represents the lesser number of points.

```
LIST OF PRINCIPAL DATUM TAGS FOR THE HARN'S
-------------------------------------------------------
AK  Alaska          NAD 83(1992)
AL  Alabama         NAD 83(1992)
AR  Arkansas        NAD 83(1997)
AZ  Arizona         NAD 83(1992)
CA  California      NAD 83(1998)   NAD 83(1992)
CO  Colorado        NAD 83(1992)
CT  Connecticut     NAD 83(1996)
DC  Dist.Columbia   NAD 83(1993)   NAD 83(1991)
DE  Delaware        NAD 83(1991)
FL  Florida         NAD 83(1999)
GA  Georgia         NAD 83(1994)
HI  Hawaii          NAD 83(1992)
IA  Iowa            NAD 83(1996)
ID  Idaho           NAD 83(1999)
IL  Illinois        NAD 83(1997)
IN  Indiana         NAD 83(1997)
KS  Kansas          NAD 83(1997)
KY  Kentucky        NAD 83(1993)
LA  Louisiana       NAD 83(1992)
MA  Massachusetts   NAD 83(1996)
MD  Maryland        NAD 83(1991)
ME  Maine           NAD 83(1996)
MI  Michigan        NAD 83(1994)
MN  Minnesota       NAD 83(1996)
MO  Missouri        NAD 83(1997)
MS  Mississippi     NAD 83(1993)
MT  Montana         NAD 83(1999)
NC  North Carolina  NAD 83(2001)
ND  North Dakota    NAD 83(1996)
NE  Nebraska        NAD 83(1995)
NH  New Hampshire   NAD 83(1996)
NJ  New Jersey      NAD 83(1996)
NM  New Mexico      NAD 83(1992)
NV  Nevada          NAD 83(1999)
NY  New York        NAD 83(1996)
OH  Ohio            NAD 83(1995)
OK  Oklahoma        NAD 83(1993)
OR  Oregon          NAD 83(1998)
PA  Pennsylvania    NAD 83(1992)   NAD 83(1996)
PR  Puerto Rico     NAD 83(2002)
RI  Rhode Island    NAD 83(1996)
SC  South Carolina  NAD 83(2001)
SD  South Dakota    NAD 83(1996)
```

```
TN  Tennessee       NAD 83(1995)
TX  Texas           NAD 83(1993)
UT  Utah            NAD 83(1994)
VA  Virginia        NAD 83(1993)
VI  Virgin Is.      NAD 83(2002)   NAD 83(1997)
VT  Vermont         NAD 83(1996)
WA  Washington      NAD 83(1998)
WI  Wisconsin       NAD 83(1997)
WV  West Virginia   NAD 83(1995)
WY  Wyoming         NAD 83(1993)   NAD 83(1992)
```

15. Inventory of GEOCON Support Files

dela.b	binary grid, transformation error, latitude, CONUS region
delaa.b	binary grid, transformation error, latitude, Alaska region
delap.b	binary grid, transformation error, latitude, PR/VI region
delo.b	binary grid, transformation error, longitude, CONUS region
deloa.b	binary grid, transformation error, longitude, Alaska region
delop.b	binary grid, transformation error, longitude, PR/VI region
dev.b	binary grid, transformation error, vertical, CONUS region
deva.b	binary grid, transformation error, vertical, Alaska region
devp.b	binary grid, transformation error, vertical, PR/VI region
dsla.b	binary grid, coordinate transformation, latitude, CONUS region
dslaa.b	binary grid, coordinate transformation, latitude, Alaska region
dslap.b	binary grid, coordinate transformation, latitude, PR/VI region
dslo.b	binary grid, coordinate transformation, longitude, CONUS region
dsloa.b	binary grid, coordinate transformation, longitude, Alaska region
dslop.b	binary grid, coordinate transformation, longitude, PR/VI region
dsv.b	binary grid, coordinate transformation, vertical, CONUS region
dsva.b	binary grid, coordinate transformation, vertical, Alaska region
dsvp.b	binary grid, coordinate transformation, vertical, PR/VI region
infoh.txt	information file, horizontal, CONUS region
infoha.txt	information file, horizontal, Alaska region
infohp.txt	information file, horizontal, PR/VI region
infov.txt	information file, vertical, CONUS region
infova.txt	information file, vertical, Alaska region
infovp.txt	information file, vertical, PR/VI region

Program GEOCON
Version 1.0
July 6, 2012

User Guide

Author:
Dr. Dennis Milbert
Chief Geodesist, Retired
Rockville, Md

Contact:
Dr. Dru Smith
Chief Geodesist
National Geodetic Survey, NOAA

1. Introduction

GEOCON performs three-dimensional coordinate transformations between the NAD 83(HARN) coordinates and the NAD 83(NSRS2007) coordinates. GEOCON also issues information about the quality of the transformation at each point, and notifications regarding poor quality results.

This document ("The User Guide") is a companion to the GEOCON Operating Instructions. Please refer to the latter document for details on core elements, such as formats. It may be helpful to read the Operating Instructions before reading this guide.

2. Coordinate Transformation

The coordinate sets, both HARN and NSRS2007, are defined quantities (as of a particular date). Good or bad, the coordinate differences between the HARN and the NSRS2007 are constants. There can be no expectation of smoothness in a coordinate transformation. Plots of high resolution coordinate transformations may not look pleasing. Blunders in the older coordinates that are corrected in the newer coordinates will appear as abnormal shifts. But, they are not to be excluded from consideration. They are actual differences between two disseminated coordinate sets. Further, abnormal coordinate differences exist because some errors were corrected when making the newer coordinate set.

To support our user communities, NGS decided to create a coordinate transformation between the NAD 83(HARN) and the NAD 83(NSRS2007). This, in turn, implies seeking a mathematical mapping between the two coordinate sets, *irrespective of their values.*

3. Coordinate Differences: Between the Points

One could argue that a coordinate transformation only exists at the points that define the coordinate differences. While being a legitimate argument, such a perspective would provide no guidance on how to treat intervening points that were directly or indirectly tied to the defining coordinates. For guidance we must consider field practice. And, this practice will not just include surveyors, but all practitioners who create georeferenced data sets.

It is natural to define the coordinate transformation between defining points as being an intermediate value. As such, one does not want to see extraneous oscillations in a transformation function, even when there are large, local excursions at the defining points. In essence, one wishes to "connect the dots".

To satisfy the needs of honoring the data and generating intermediate values, we select a gridding method of splines in tension (Smith and Wessel, 1990). This models the physical behavior of a thin, flexible plate that passes through the defining points. However, such a model, by itself, is subject to overshoots and undershoots when data differences occur near gaps in irregularly spaced data. By mathematically applying

tension at the edges of a grid, it is possible to suppress the oscillations, and generate representative intermediate values. For GEOCON, after some tests, a 1' x 1' grid with a tension parameter of 0.4 was selected.

Basically, the fitted grid is modeling the coordinate differences a practitioner would obtain when performing two different least squares adjustments of the same survey data when controlled by two different control point coordinate sets. Or, the grid models the differences for photogrammetric data, or synthetic aperture radar, or LIDAR, or any other coordinate measurement system that ties to the control point coordinate set.

If, however, a practitioner only performs a single point tie, then all of the geospatial data should be transformed by the unique coordinate difference of the source control point.

Consider an extreme example; the ellipsoidal height for the point M 123 (PID = TT2413) as obtained from the NGS database in November 2011 (Appendix A.1.):

NAD 83(HARN) ellipsoidal height: 486.945 meters   (06/20/05)
NAD 83(NSRS2007) ellipsoidal height: 641.786 meters   (07/17/09)

The change of nearly 155 meters is due to the re-measurement and readjustment of M 123.

Now suppose a practitioner has some regional data in a HARN coordinate set. The data are tied to the abnormal point M 123 and other nearby control points with much smaller coordinate differences. The ellipsoidal height coordinate differences will get larger and larger for the data points nearer and nearer M 123. On the other hand, if the practitioner connected the work solely to M 123, then *every* ellipsoidal height must be increased by exactly 154.841 meters irrespective of the distance from M 123.

This is one reason why the National Geodetic Survey considers actual recomputation of geospatial data, and not coordinate transformations, as "best practice". The coordinate transformation is, at its heart, only a model of actual geospatial measurement and processing.

4. Coordinate Differences: Inside the Cell

The analysis of the NAD 83(NSRS2007) National Readjustment (Milbert 2008) showed a number of surprising results. One was the highly local character of the network. Over 50% of all GPS vectors were 31 km or less in length. Figure 5.3 of Milbert (2008, pg. 12) is reproduced below as Figure 4.1. Note the significant number of GPS vectors of just 1 or 2 km in length.
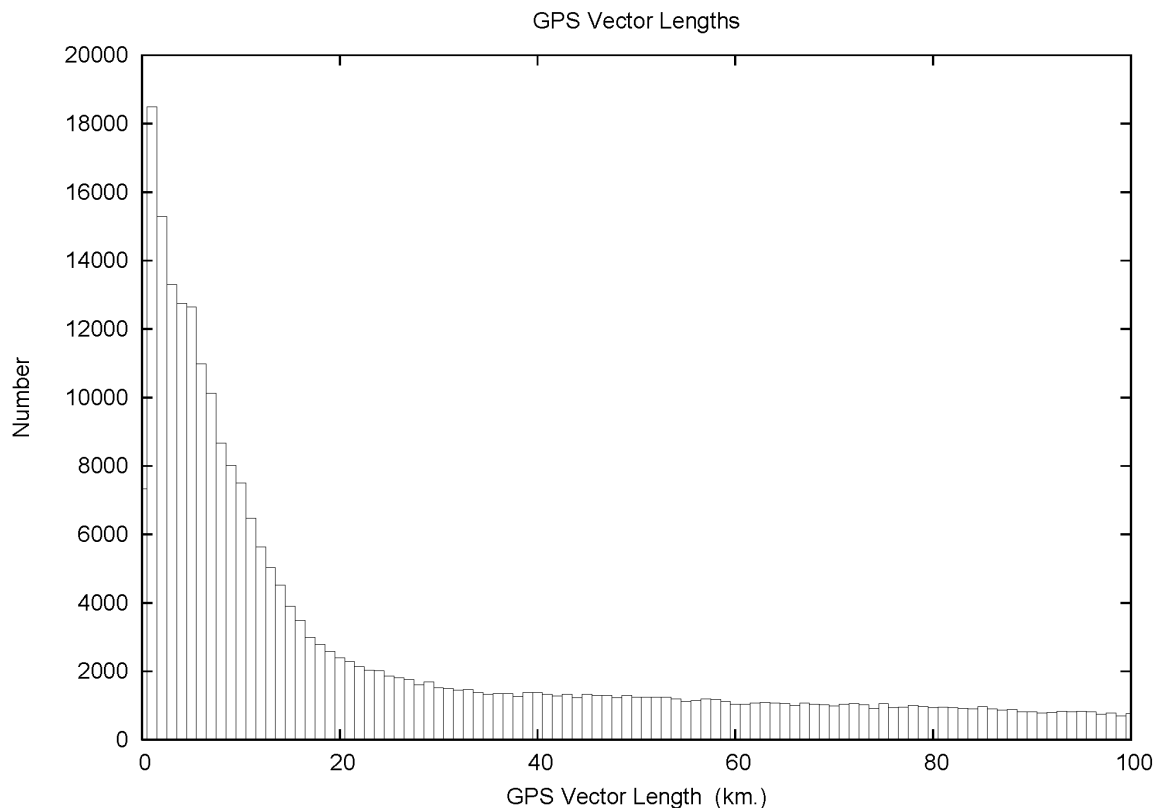
Figure 4.1. Distribution of GPS vector lengths, 0 to 100 km. 1 km bin size.

Even though a grid size of 1' x 1' was selected to model the coordinate differences, it can be expected that there will be clusters of defining points at some of the grid nodes. This is because of the close spacing of some of the control points. This naturally raises a question. How can one "honor the data" for multiple coordinate differences near a grid node?

It was decided to select the most representative control point by means of a median filter. The median value of a set is the value that delimits the higher half and lower half of the set. It can be obtained by sorting the set and choosing the middle value. In the case of an even number of points, the mean of the two central points is reported. The advantage of the median is that it is a robust estimate when more than two data are present. An outlier does not disturb the median.

To further increase the robustness, a modified median procedure was used prior to gridding. In the case of a cluster of exactly 2 points, a search was performed in the 1' to 2' ring surrounding the central 1'x1' cell. If the search found a point, it was used as a tiebreaker to select the winning median point in the central cell. If the 1' to 2' ring was insufficient, then the 2' to 3' ring was searched for a tiebreaker. If, after two ring searches no tiebreaker was found, then the central cell median was selected at random. It was found that that the ring search procedure was able to reduce 8177 pairs to 1534 pairs. And, by using a random selection, the possibility of the median being influenced by an abnormal point in the pair is additionally halved.

The philosophy in choosing a median procedure is rooted in the likely practice of geospatial professionals. In the presence of a cluster of control points, connections

should be made to a sufficient number to confirm the connection to the control network. Depending upon the accuracy of the positioning measurements, significant discrepancies may be identified. And, ties to suspect control points would be discarded. This is standard practice in surveying.

5. Transformation Quality

The preceding two sections illustrate the challenge in describing the quality of coordinate transformation.

It is possible that georeferenced data may be tied to control points with both standard coordinate differences and abnormal coordinate differences.

It is possible that georeferenced data may be tied solely to control points with standard coordinate differences, but be near points with abnormal coordinate differences.

Conversely, it is possible that georeferenced data may be tied solely to control points with abnormal coordinate differences, but be near points with standard coordinate differences.

Further, there can be highly variable scale. "Near" may refer to spacing of a few hundred meters or a few hundred kilometers. "Standard" and "abnormal" may refer to coordinate differences of less than a millimeter to over 100 meters. User accuracy requirements may vary from millimeter to multi-meter network accuracy, and may address only horizontal or only vertical components.

To represent transformation quality, a statistical resampling procedure known as cross-validation was selected. Cross-validation is useful in assessing predictive models (Efron and Tibshirani, 1998). It is appropriate to consider coordinate transformation as an exercise in prediction. The NGS database of coordinates with multiple coordinate realizations represents prior knowledge. Other practitioners, whose data are never known to NGS, will establish coordinates traceable to the NGS database. We seek to predict the transformed coordinates from those unknown coordinates.

In its simplest form, cross-validation consists of cutting a data set in half. Call the first half the *training set*, and build the prediction model from the first half. Then compare the second half of the data, called the *validation* or *testing set*, to the model predictions. Similarly, one may exchange the two data halves, and repeat the process.

In the extreme case, one can imagine taking a set of data, withholding a single data point, producing a unique model, and then computing the difference between the single withheld data point and the model. And, one can imagine doing this sequentially for every single point in the data set. This method is known as the jackknife (Efron 1979).

Needless to say, the jackknife can entail a large computational burden for sizable data sets. However, there is a middle ground. One can compute K-fold cross-validation (Efron and Tibshirani, *ibid*). The data set is partitioned into K subsets. In sequence, each

subset is designated as a testing set, and is temporarily withheld from the data set. The model is computed from each reduced data set, and differences are computed between the temporarily withheld data and the model prediction. The process is sequenced K times until a prediction error is established for each data point.

For GEOCON 69-fold cross-validation was performed. The master data set was 69540 point pairs. So each testing subset was a little over 1000 points. This means 69 training grids were computed for each component, for the regions of CONUS, Alaska, and Puerto Rico/Virgin Islands. For example, the 207 training grids for CONUS took about 8 hours to compute. The result was a set of cross-validation errors for the coordinate differences in latitude, longitude, and height.

The appeal of cross-validation error is that it quantifies the abnormality of a point. Consider the case of M 123. The ellipsoid height coordinate difference (2007 - HARN) is +154.797 meters. The cross-validation error is +154.8192 meters, showing that the point is definitely abnormal when compared to its neighbors.

By also gridding the cross-validation error as a quality measure, it is possible to quantify the uncertainty associated with varying field procedures and processing. If one is in the midst of normal control points, then the cross-validation errors will tend to be small. As one approaches an abnormal point, the quality measure will get larger, indicating increasing uncertainty about the proper coordinate transformation value to use.

For the case of a cluster of points, it is important to reflect the possibility that one of the points may be abnormal. Recall, when gridding the coordinate differences, a modified median procedure was used to select the points. Such a procedure will drop abnormal points in clusters. For the quality grids, the median is not used. Rather the worst case cross-validation error in a cluster is gridded. The worst case is selected by choosing the error that is furthest in magnitude from the median error (an "anti-median").

In this way, the coordinate transformation grids of GEOCON provide the most likely transformations to apply to geospatial data. And, the transformation quality grids give a conservative idea of the quality of the transformation.

6. An Example in Northeast Colorado

It is useful to consider some raw coordinate data and its conversion into grids as illustration. Consider a region in northeast Colorado (Figure 6.1):
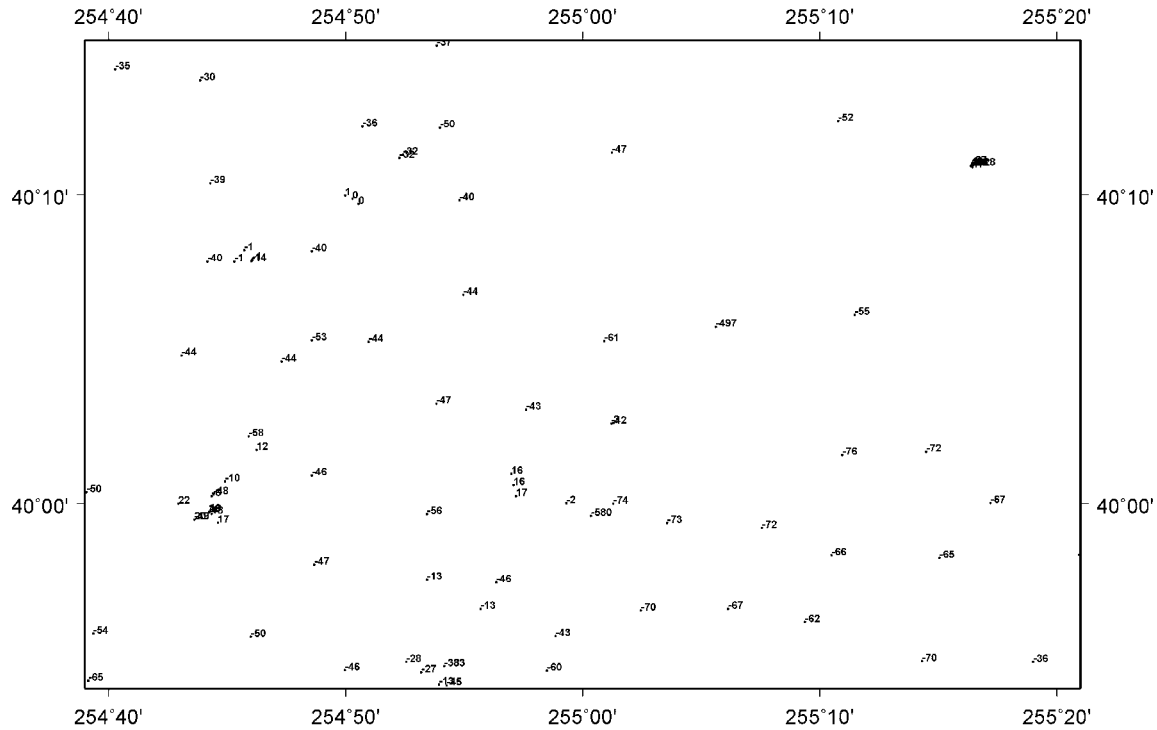
Figure 6.1. Longitude Coordinate Differences, Northeast Colorado.

To aid in discussion, the permanent identifiers (PID) of the points are displayed in Figure 6.2. The values portrayed in Figure 6.1 are the coordinate differences in longitude taken in the sense of NAD 83(NSRS2007) – NAD 83(HARN). In computing the shifts, the sense of the longitudes are positive East. The units are 0.00001 arc seconds. Thus, the numerical values in Figure 6.1 (as well as in subsequent figures) are integers. The small dots present near each number are the locations of the points. The small dots are *not* decimal points.

NSRS2007 - HARN PID's

Figure 6.2. Permanent Identifiers (PID), Northeast Colorado.

Considerable variation in the coordinate longitude shifts is seen in Figure 6.1. However a nominal shift of -0.00050 arc sec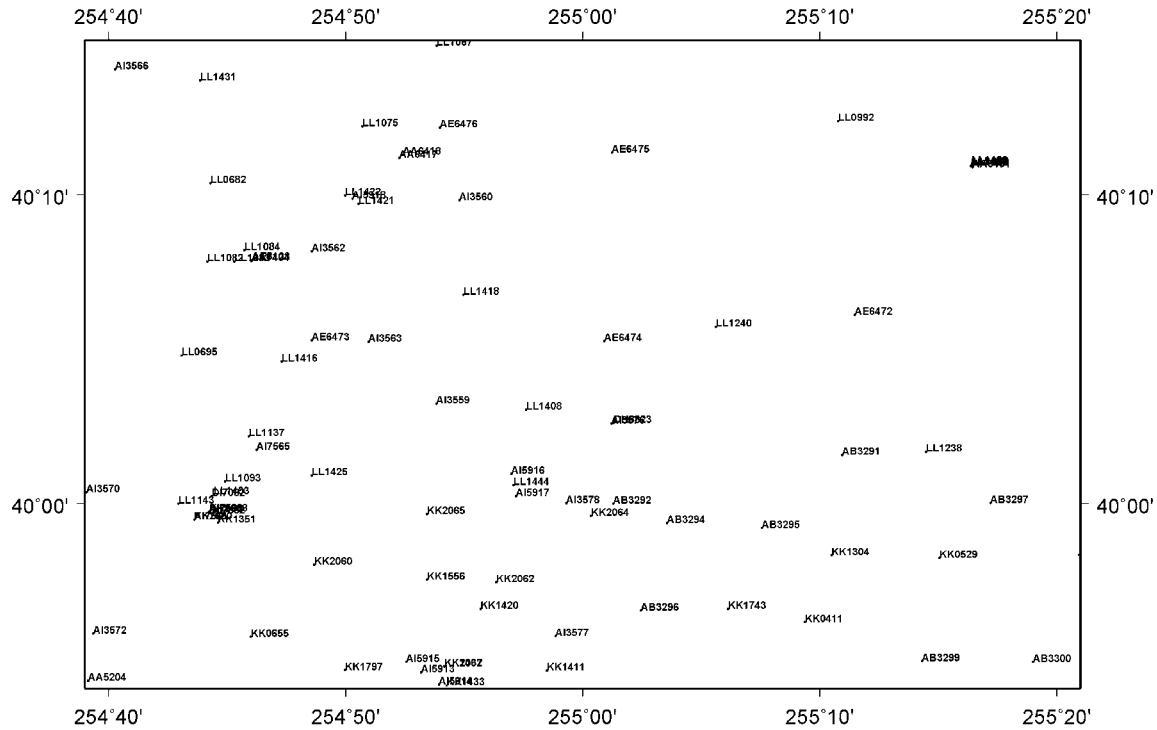onds is evident. Close inspection of the shifts will show some abnormal values. In some cases, the points are clustered so tightly, it is impossible to resolve the numbers.

There are four abnormal longitude shifts of particular interest. These are plotted in Figure 6.3. And, their PID's are plotted in Figure 6.4. Recall from the earlier discussion, these are actual differences in the NGS database. And, most likely, subsequent surveys, maps, and other geospatial data have been controlled by both sets of coordinates. While they are abnormal in a regional sense, they must be considered in a coordinate transformation.

Table 6.1 – Abnormal Longitude Coordinate Shifts

PID	Long. Shift (0.00001 arc sec)	Long. Shift (cm)
KK2067	-333	-7.8
KK2064	-580	-13.7
LL1240	-497	-11.8
LL1465	+2328	+55.1

8

Select NSRS2007 - HARN Long. Shifts (10^-5 arcsec)

254°40' 254°50' 255°00' 255°10' 255°20'

40°10'

.2328

.-497

40°00'

.-580

.-333

254°40' 254°50' 255°00' 255°10' 255°20'
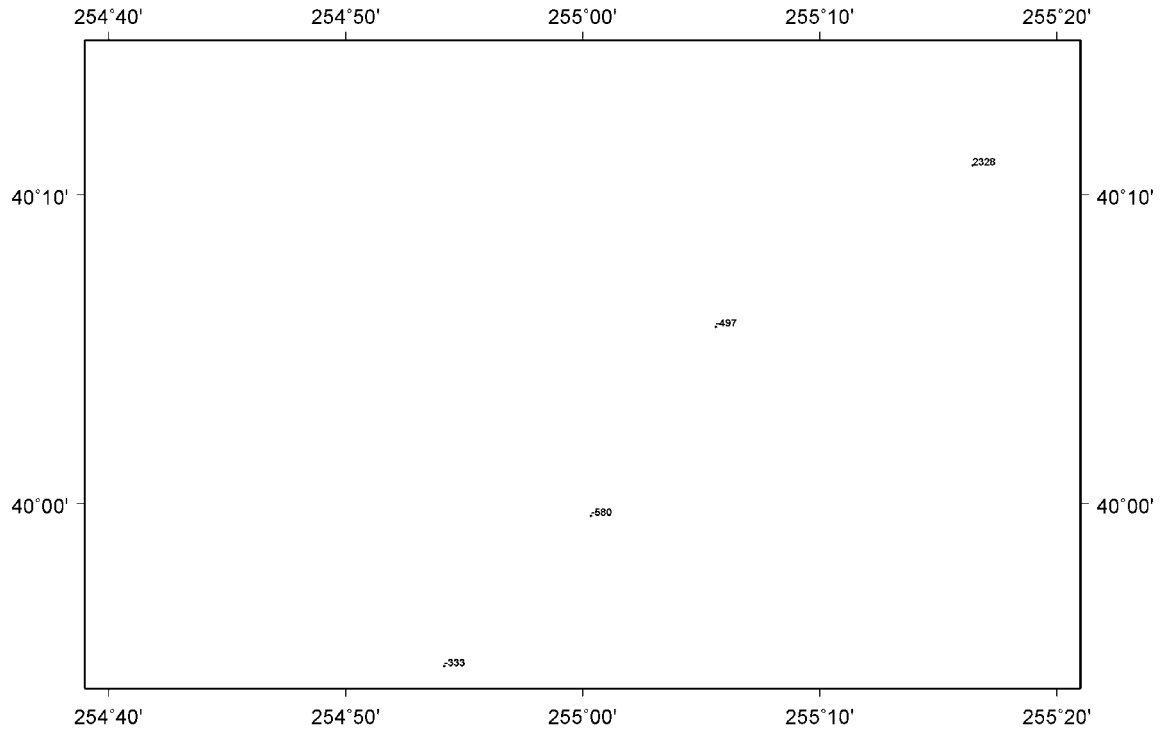
Figure 6.3. Abnormal Longitude Coordinate Differences, Northeast Colorado.

Select NSRS2007 - HARN PID's

254°40' 254°50' 255°00' 255°10' 255°20'

40°10'

.LL1465

.LL1240

40°00'

.KK2064

.KK2067
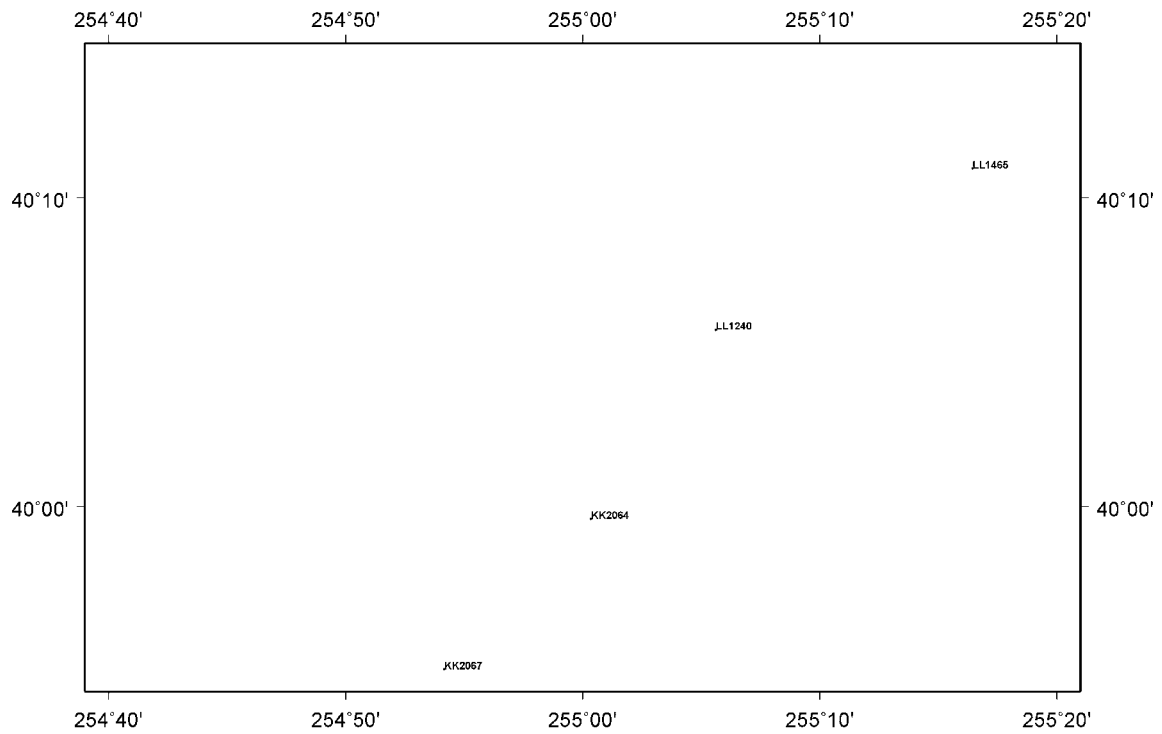
254°40' 254°50' 255°00' 255°10' 255°20'

Figure 6.4. Permanent Identifiers (PID) of Abnormal Longitude Differences.

9

When the data in Figure 6.1 are gridded, we obtain the results portrayed in Figure 6.5. Once again, the units are 0.00001 arc seconds, the values displayed are integers, and the small dots represent grid nodes, and *not* decimal points.

NSRS2007 - HARN Long. Shifts (10^-5 arcsec)

```
       254°40'            254°50'            255°00'            255°10'            255°20'

       36 35 34 33 32 30 31 31 32 33 35 37 39 40 41 41 43 45 46 47 48 49 52 56 60 65 67 68 67 64 60 54 49 45 41 38 36 34 33 32 32 33
       37 36 36 35 33 32 31 30 30 30 31 34 38 40 43 46 47 47 47 47 48 50 53 59 67 74 79 80 79 75 68 60 52 47 43 39 36 33 32 32 33 33
       38 38 37 37 35 33 31 28 26 25 26 28 34 38 44 49 49 48 48 48 48 49 53 64 76 87 95 98 96 90 81 68 56 50 45 41 36 32 31 32 33 34
       39 39 40 39 38 35 30 25 22 19 16 15 20 28 36 42 45 47 48 50 50 51 53 72 92 108 119 123 121 113 100 86 71 61 52 44 37 30 31 32 33 34
40°10' 40 41 42 42 41 38 29 21 18 15 10 2 5 18 29 36 41 46 50 54 59 66 77 97 119 139 153 159 155 143 125 105 87 72 60 50 43 37 35 35 35 36   40°10'
       42 43 44 44 43 36 21 14 15 19 21 18 18 24 30 36 42 47 53 59 68 81 99 124 152 179 200 208 201 180 153 125 100 80 66 56 48 43 40 38 37 37
       44 45 46 47 46 39 10 4 16 29 39 35 32 33 35 39 43 48 54 62 73 90 114 146 186 227 261 276 260 224 181 139 105 82 67 57 51 46 43 40 39 38
       45 46 46 46 44 37 24 20 26 35 42 41 39 39 40 41 44 48 53 60 71 88 116 158 212 273 333 366 331 270 204 144 98 75 63 57 52 48 45 42 40 38
       46 46 46 46 44 39 33 32 37 42 46 45 43 42 42 43 45 47 50 54 61 74 99 149 216 296 386 461 387 297 213 139 78 64 59 56 53 50 46 43 40 38
       47 46 45 45 44 39 36 36 42 47 51 48 44 44 44 44 44 46 47 48 51 64 118 185 255 320 357 325 265 201 142 96 75 65 60 56 52 49 45 40 36
       45 42 39 37 34 30 27 27 33 40 45 47 46 46 46 45 43 41 42 43 45 56 96 148 199 242 262 250 216 175 134 102 82 71 65 60 56 52 47 42 36
       43 38 33 27 22 17 12 11 20 31 40 45 48 48 48 45 39 34 35 42 44 46 44 78 117 153 180 193 188 170 145 118 95 81 74 69 64 60 55 51 46 41
       42 35 26 17 10 5 1 7 9 27 41 47 50 50 47 40 29 18 14 25 50 84 81 87 102 121 137 145 143 133 117 99 81 76 73 72 68 64 60 55 51 47
       44 33 20 6 3 1 9 7 17 33 48 52 53 53 50 37 20 1 15 2 51 165 119 95 92 99 107 112 112 106 97 87 78 74 72 71 69 66 64 60 56 52
40°00' 49 34 16 4 22 19 6 6 19 33 44 50 52 54 55 39 21 2 15 30 11 334 128 85 78 81 86 89 89 86 82 78 74 71 70 69 69 68 67 63 59 55   40°00'
       45 34 19 2 11 16 9 7 22 35 43 46 45 41 37 30 23 15 9 16 64 176 127 93 78 72 74 75 74 71 71 70 69 68 68 68 67 66 65 62 59 57
       45 37 26 16 7 5 9 19 30 40 47 44 38 29 20 24 28 30 31 41 67 104 99 85 76 71 70 70 68 67 66 66 66 67 66 65 64 62 60 58 57
       49 43 36 29 25 24 28 33 39 44 46 43 37 31 28 32 38 45 45 48 57 72 75 73 70 69 68 68 66 65 64 64 65 66 67 67 65 62 59 57 54 52
       54 50 45 41 39 40 43 47 48 48 47 44 39 34 33 37 41 46 47 47 46 55 62 64 65 65 66 65 64 63 62 63 65 67 68 68 65 60 56 51 48 45
       60 57 53 51 50 50 52 54 54 52 49 45 39 33 31 36 39 45 51 55 52 54 57 59 61 63 64 64 63 63 62 63 64 66 68 70 62 56 50 43 38 32
       65 62 59 57 57 58 60 61 59 55 50 45 38 30 21 15 28 40 49 53 53 52 53 55 58 61 62 63 62 62 62 62 62 60 61 58 52 46 40 30 18 11

       254°40'            254°50'            255°00'            255°10'            255°20'
```
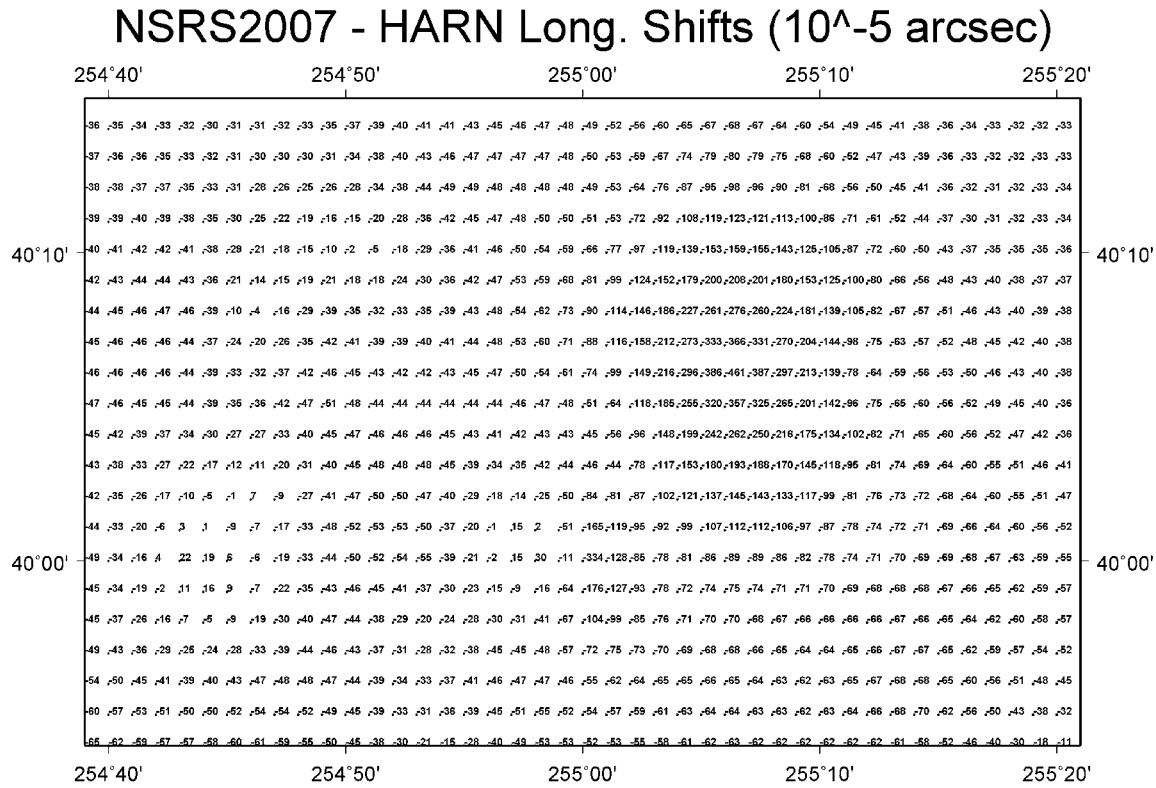
Figure 6.5. Gridded Longitude Coordinate Differences, Northeast Colorado.

We note that 1' x 1' grid captures the general magnitude of the coordinate shifts. It models the smaller shifts in the vicinity of 40° 00' N, 254° 45' E. Two of the four abnormal shifts (KK2064 and LL1240) are also evident as depressions in the grid. In particular, it is seen that the depression formed by LL1240 extends for some distance in all directions. This is because there are no neighbors nearby.

The use of a grid model will perform well if, for example, a local survey connects to LL1240 in addition to its neighbors: AE6472, AB3291, AB3295, AB3294, etc. One can see that a survey that solely connects to AE6472 (-0.00055 arc sec), but does not connect to LL1240 (-0.00497 arc sec), has a nominal longitude shift of -0.00055 arc seconds. But as those local survey points approach LL1240, they will get progressively worse shifts. This situation must be reported to the coordinate transformation user. This is done by means of the transformation quality grid.

In addition, while two of the abnormal shifts were modeled in the coordinate transformation grid, two other shifts were not (KK2067 and LL1465). These points were in clusters within the 1' x 1' grid resolution. They were dropped by the modified median filter prior to the actual gridding. The points are plotted in Figure 6.6.
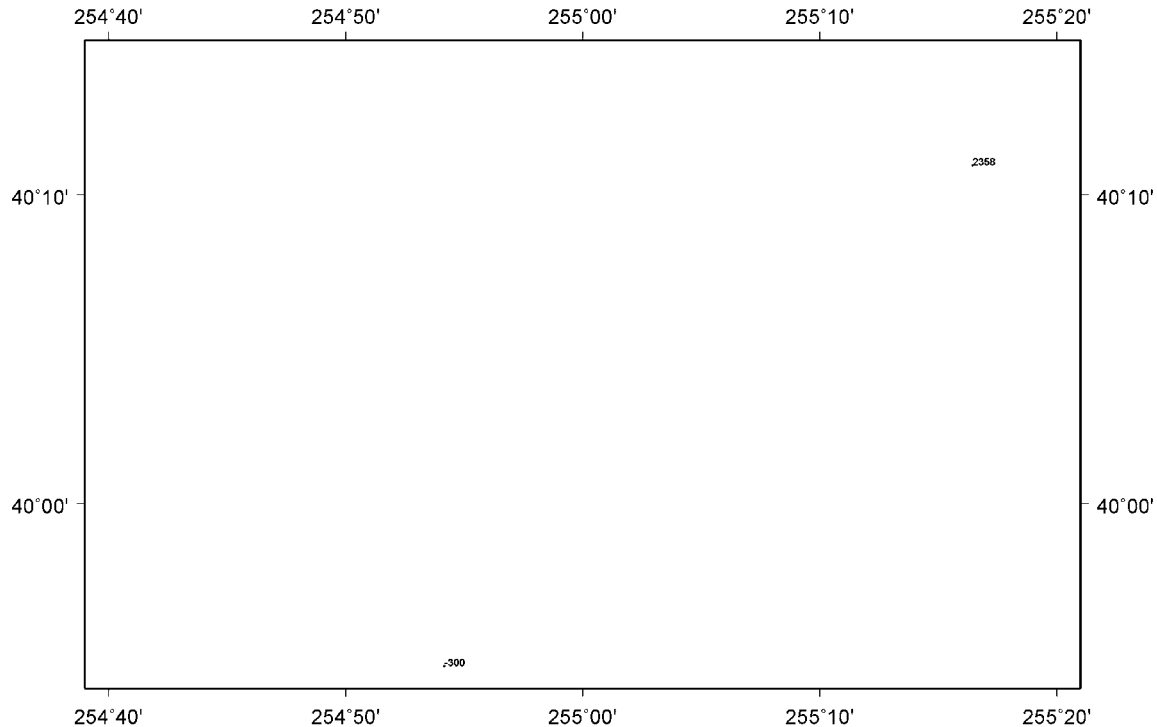
Figure 6.6. Abnormal Longitude Differences Dropped by Median Filter.

Note that the values in Figure 6.6 are not the actual longitude coordinate differences. Rather, they are the difference between the abnormal longitude differences of Figures 6.1 and 6.3, and the gridded longitude differences of Figure 6.5. Also note that the gridded values in Figure 6.5 are undisturbed by these abnormalities.

Once again, these abnormal quantities must be communicated to the coordinate transformation user. It is assumed that local surveying, mapping, and other geospatial work will have made multiple connections to nearby points, and have detected the issues. So the filtered transformation grid represents the most likely situation. As discussed earlier, the situation is reported through the transformation quality grid.

Figure 6.7 plots the cross-validation errors in the sense of measured difference – gridded difference, where the grid was obtained after dropping that specific measured difference (and around 1000 other randomly selected points throughout the entire data set). As before, units are 0.00001 arc seconds, and longitudes are positive East. To assist in inspection, the cross-validation errors of the four abnormal points are shown in Figure 6.8.

11

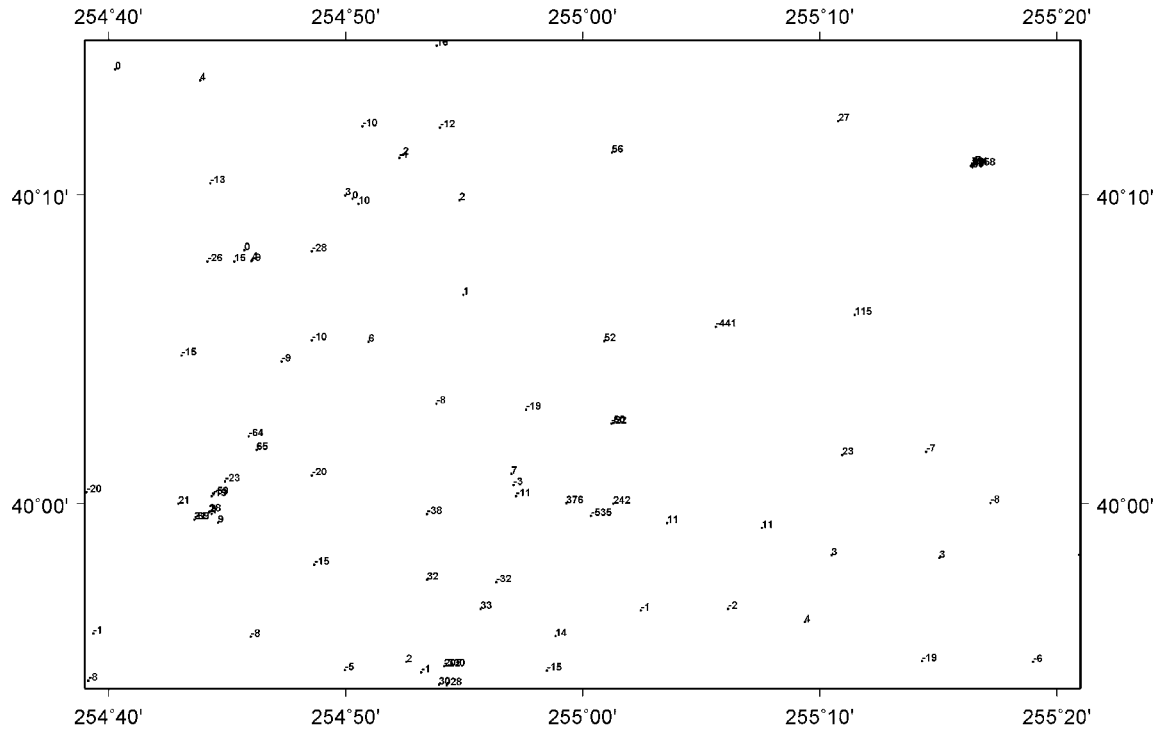## Cross Valid. Long. Shifts (10^-5 arcsec)

Figure 6.7. Cross-Validation Error of Longitude Coordinate Differences.

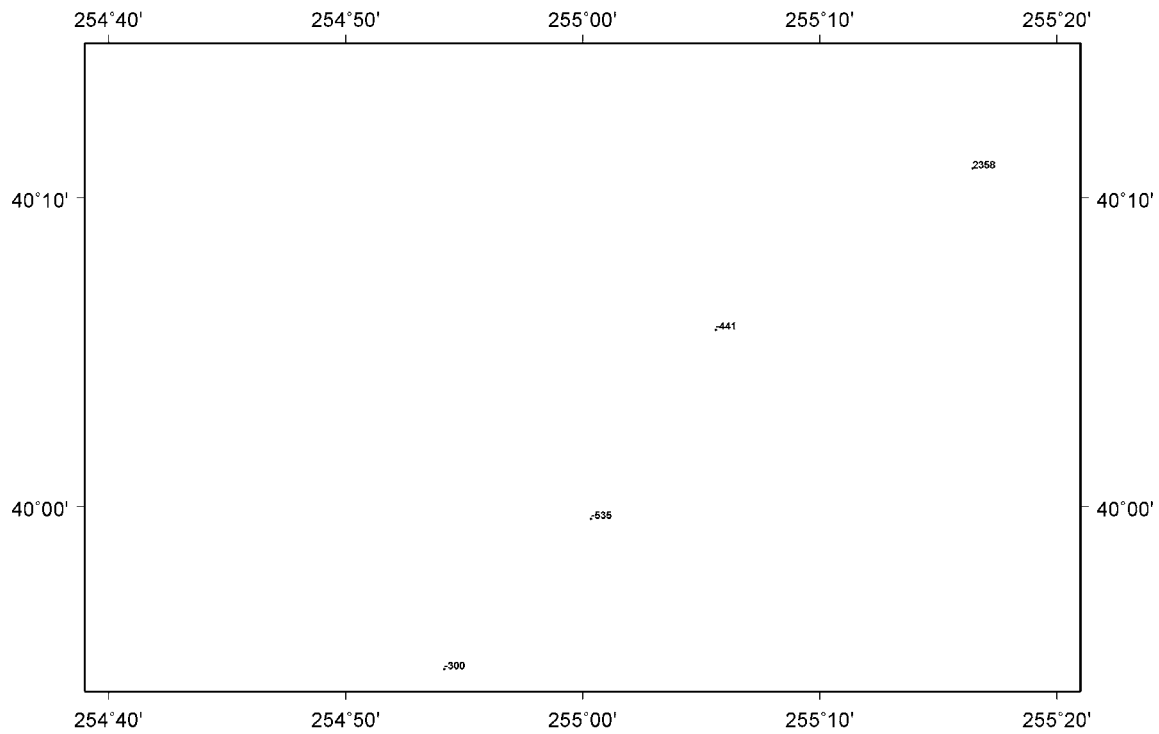## Select Cross Valid. Long. Shifts (10^-5 arcsec)

Figure 6.8. Abnormal Cross-Validation Error of Longitude Differences.

The first thing to be noticed in Figure 6.7 is that the nominal values of the cross-validation errors are typically much smaller than the nominal -0.00050 arc second longitude shift. This shows the grid is doing a very good job of predicting the coordinate shift at a point when that point was withheld from the gridding computation.

In addition, the cross-validation errors at the abnormal points do, indeed, look abnormal. They show the systematic characters of the quality value, including the sign of the value. For example, consider `LL1465`, with an actual longitude shift of +0.02328 arc seconds. The gridded transformation in Figure 6.5 is about -0.00030 arc seconds. So the systematic error for `LL1465` is assessed at +0.02358 arc seconds.

Recall that the transformation quality grid is obtained by gridding the cross-validation error, and that abnormalities were dropped by a median filter. However, in the case of clusters of the cross-validation error, the worst-case error is passed to the gridding algorithm. This makes a distinct difference in the locations around `KK2067` and `LL1465`. This is displayed in Figure 6.9.
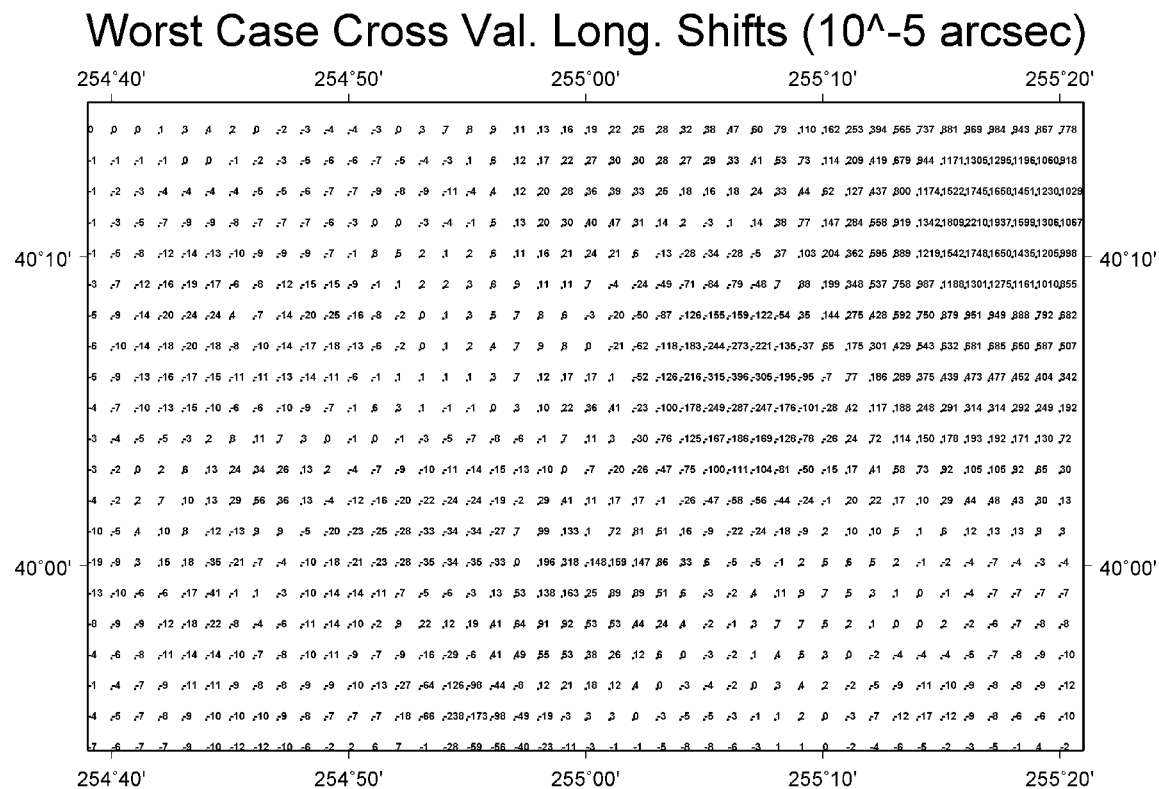
Worst Case Cross Val. Long. Shifts (10^-5 arcsec)

Figure 6.9. Gridded Worst Case Cross-Validation Error of Longitude Differences.

The points `KK2064` and `LL1240` have isolated abnormal shifts. So, their gridded errors spread out over some distance. The southern point, `KK2067`, is the worst case in a cluster. But the gridded error quickly decreases as one approaches the normal neighbors. The northeast point, `LL1465` is in an isolated cluster. Its worst case cross-validation error passes to the gridding algorithm. And, because the cluster was isolated, the error estimate spreads over a larger distance in the quality grid.

Therefore, we see that the user is warned about the presence of abnormal coordinate shifts in the vicinity of a geospatial project. Large values in the quality grid are created by coordinate shifts that don't agree with neighboring values. And, those large values are reported whether they are isolated, or present in a cluster of normal coordinate shifts.

These quality values indicate potential systematic error in the reported transformation. The error is *not* due to uncertainty in a nearby coordinate shift. The coordinate shifts used to create the GEOCON transformations are, by definition, error-free. Rather, the large quality values indicate uncertainty in how a local project connected into the existing control, and what procedures were followed by the practitioner regarding misfits. The quality values should be considered systematic, and absolute values of the quality numbers should be used to increase the base network accuracy of the pre-transformed coordinates.

As a final part of this example, consider the notification messages. As discussed in the GEOCON Operating Instructions, notification messages may be generated when converting near abnormal coordinate shifts present in a cluster. The pair of abnormal points (KK2067 and LL1465) seen in Figure 6.6 are stored in an information file. Notification messages are issued when an input point has horizontal or vertical error of 5 cm or more and is within about 5 km of an abnormal point in an information file. The notification messages are purely informational, and help indicate the source of a large quality value when a nearby cluster would create some ambiguity in diagnosis.

7. General Quality of the Coordinate Transformations

The quality grids, obtained from 69-fold cross-validation, are highly variable. Even so, it is worthwhile to get a general view of how well the transformation grid is able to predict at withheld points.

The two-tailed percentiles of the distributions of the cross-validation for the conterminous U.S. are collected in Table 7.1. Approximately 68490 points were validated.

Table 7.1 – Percentiles of Cross-Validation Error, CONUS

Percentile	Latitude (0.00001 arc sec)	Longitude (0.00001 arc sec)	Height (cm)
50%	3.1	3.8	0.2
68%	7.2	9.0	0.5
90%	31.2	40.9	1.9
95%	62.5	82.0	3.5
99%	324.0	405.6	8.9
99.9%	895.7	1003.6	23.6

It is seen that the 95% limits are remarkably good. We have 95% bounds of +/- 1.9 cm in latitude, +/- 2.0 cm in longitude, and +/- 3.5 cm in height. Note that the distribution is not Gaussian. It is very peaked (leptokurtic), with long tails. The 68% bounds are almost 10 times smaller than the 95% bounds. In general, the quality of the

coordinate transformation is remarkably good. In fact, at the 90% level it is comparable to the network accuracy of NAD 83(NSRS2007). Of course, as discussed earlier, the cross-validation error should be treated as systematic. Therefore, the absolute value of the quality numbers should be used to increase the base network accuracy of the pre-transformed coordinates.

The two-tailed percentiles of the distributions of the cross-validation for Alaska are collected in Table 7.2. Approximately 770 points were validated.

Table 7.2 – Percentiles of Cross-Validation Error, Alaska

Percentile	Latitude (0.00001 arc sec)	Longitude (0.00001 arc sec)	Height (cm)
50%	28.5	61.3	1.4
68%	67.8	162.7	3.9
90%	246.7	534.4	13.7
95%	323.7	735.3	23.3
99%	1091.3	1338.0	95.6
99.9%	1739.5	2255.1	8602.7

It is seen that the 95% limits are much poorer than for the conterminous U. S. We now have 95% bounds of +/- 10.0 cm in latitude, +/- 10.8 cm in longitude, and +/- 23.3 cm in height. Even so, these are sufficient quality to transform many types of geospatial data. Note that because of the small sample size (770), when the limits are established at the 99.9% boundary, height outliers are seen to appear.

The two-tailed percentiles of the distributions of the cross-validation for Puerto Rico/Virgin Islands are collected in Table 7.3. Approximately 145 points were validated.

Table 7.3 – Percentiles of Cross-Validation Error, Puerto Rico/Virgin Islands

Percentile	Latitude (0.00001 arc sec)	Longitude (0.00001 arc sec)	Height (cm)
50%	5.3	7.9	0.4
68%	8.7	15.6	0.7
90%	37.1	82.3	2.3
95%	64.3	169.0	3.1
99%	112.0	400.6	5.2
99.9%	120.3	434.7	9.9

Here the 95% limits fall between those of the conterminous U. S. and Alaska. The 95% bounds are +/- 2.0 cm in latitude, +/- 5.2 cm in longitude, and 3.1 cm in height. Note that the relatively small values at the 99% and 99.5% limits can not be given much interpretation, since the sample size is so small (145).

Note that the limits reported in Tables 7.1, 7.2, and 7.3 are general statistics. The cross-validation error is highly variable. For this reason, one should refer to the *94* transformation quality records issued by GEOCON for specific projects.

15

8. Hawaii

There is no official NSRS2007 coordinate for Hawaii. However, a control point, KOKEE, (PID = TT3487) was inadvertently published in the NGS database as NSRS2007. This NSRS2007 position was subsequently removed from the NGS database in 2012. Hence, there are no GEOCON coordinate transformations for Hawaii.

9. Images of the Transformation and Quality Grids

At this point it is useful to provide some images of the coordinate transformations from the NAD 83(HARN) to the NAD 83(NSRS2007), and the associated transformation quality grids. In this guide, only images for the conterminous U. S. are provided. However, many more images are available in the Electronic Support Material (ESM) of the GEOCON Technical Report. Figure 9.1 displays horizontal coordinate differences.
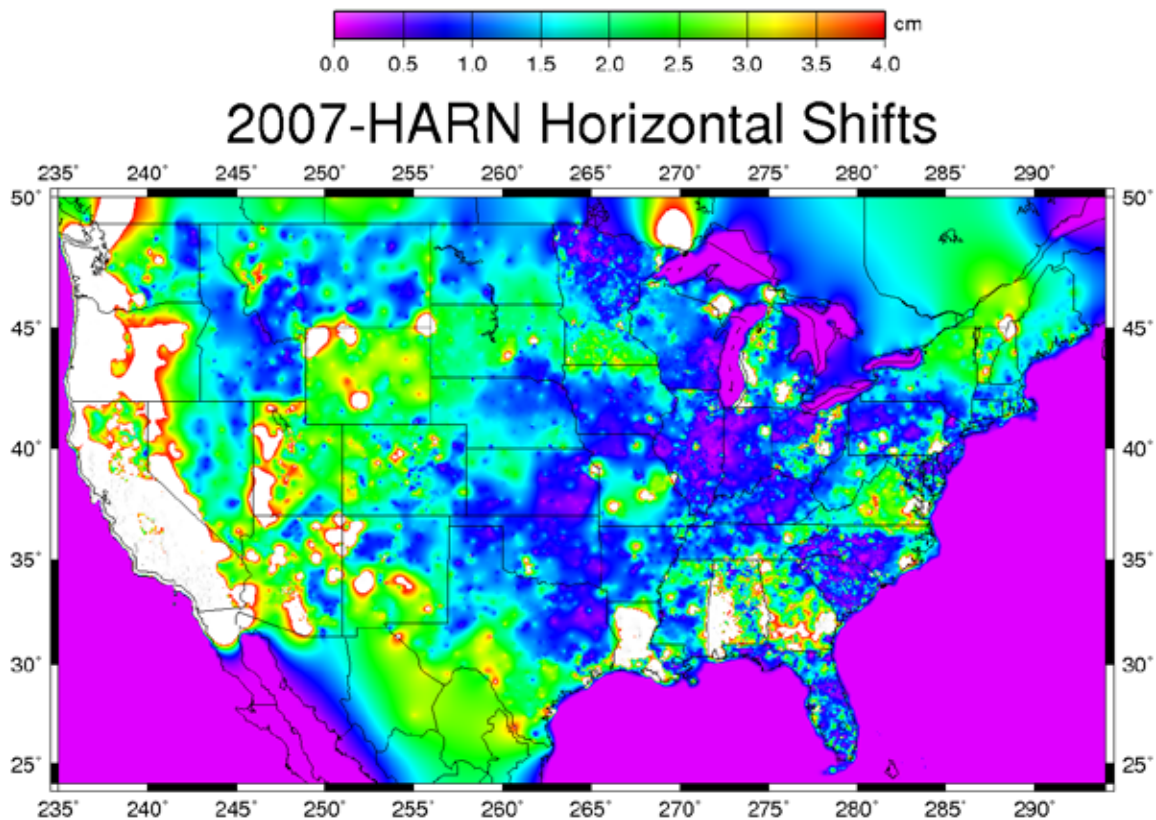
Figure 9.1. Horizontal Coordinate Differences.

The white areas in Figure 9.1 are where the horizontal shift exceeds the color scale of 4 cm. The most obvious shifts are along the West Coast of the U. S. This reflects the crustal motion of the Pacific plate between the years of 2002 and 2007 (the epochs of the HARN and the NSRS2007). Values in Canada and Mexico are to be disregarded. The coordinate shifts in the oceans and Great Lakes were set to near-zero to control the gridding.

16

Also notable in Figure 9.1 are large shifts in Louisiana and portions of Alabama and Georgia. It is also seen that the color transitions sometimes conform to state boundaries. This is because the HARN adjustments were performed on groupings of one or more states (Milbert and Milbert, 1994).

And, it must be pointed out that the horizontal shifts in most of the country exceed the typical horizontal network accuracy of 1 cm (Milbert 2008). This demonstrates that the NSRS2007 National Readjustment was necessary to obtain those excellent network accuracies.

Next, Figure 9.2 shows the ellipsoidal height coordinate differences. White areas are where the vertical shift is greater than the color scale, and black areas are where the shift is smaller than the color scale. Of note are vertical shifts in portions of California, southern Minnesota, and Alabama.
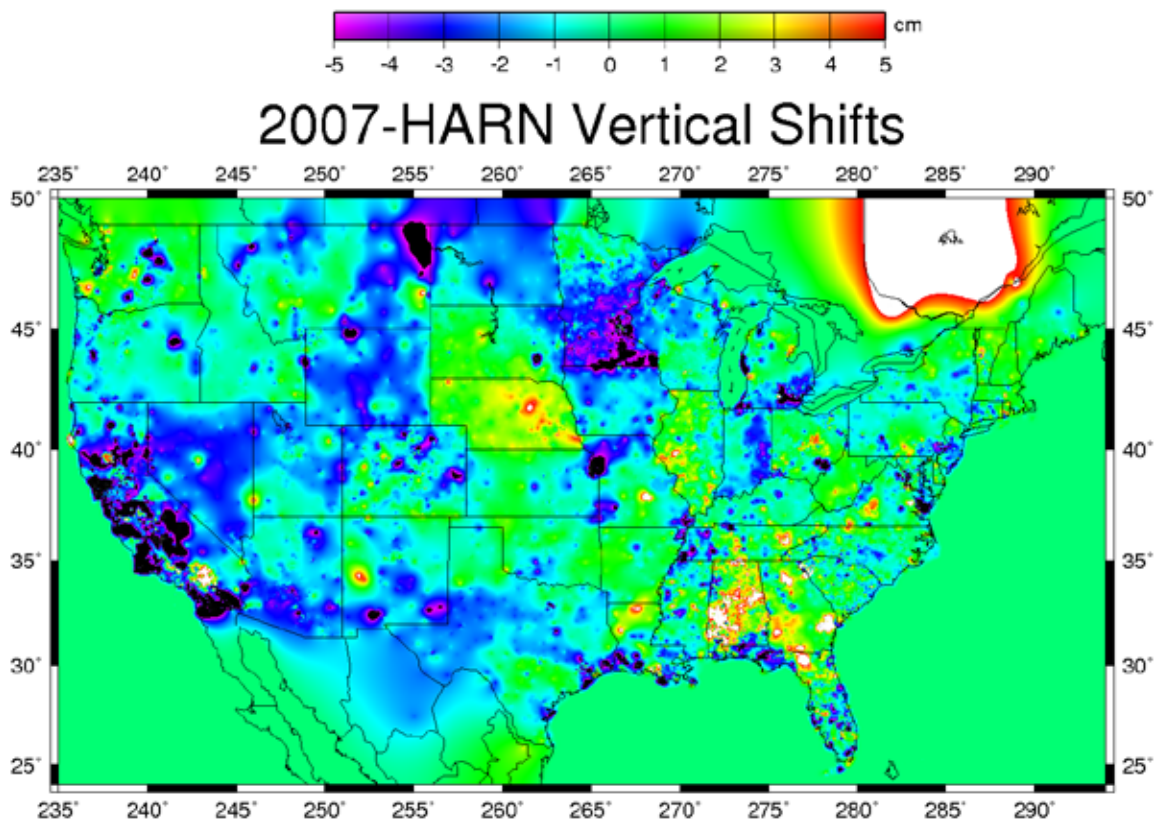

Figure 9.2. Ellipsoidal Height Coordinate Differences.

Figures 9.3, 9.4, and 9.5 plot the worst case cross-validation errors in latitude, longitude, and ellipsoidal height. These portray the transformation quality grids.
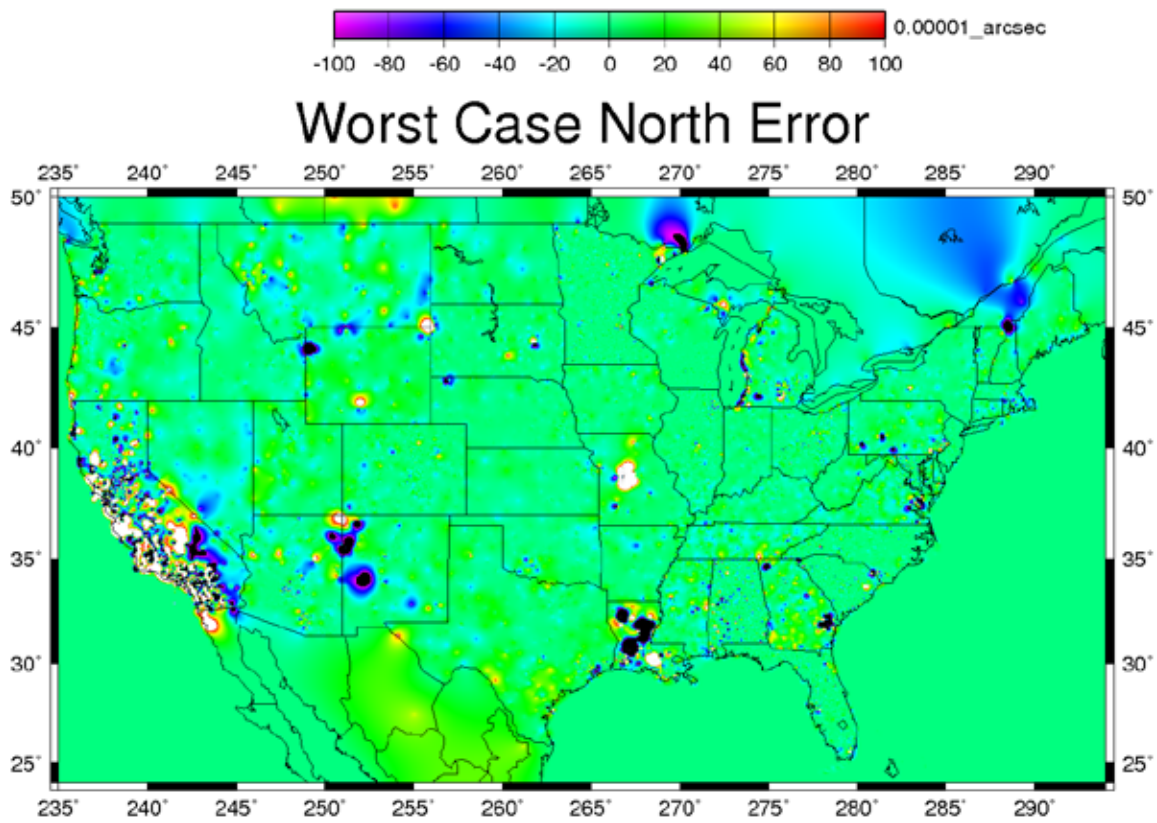
17

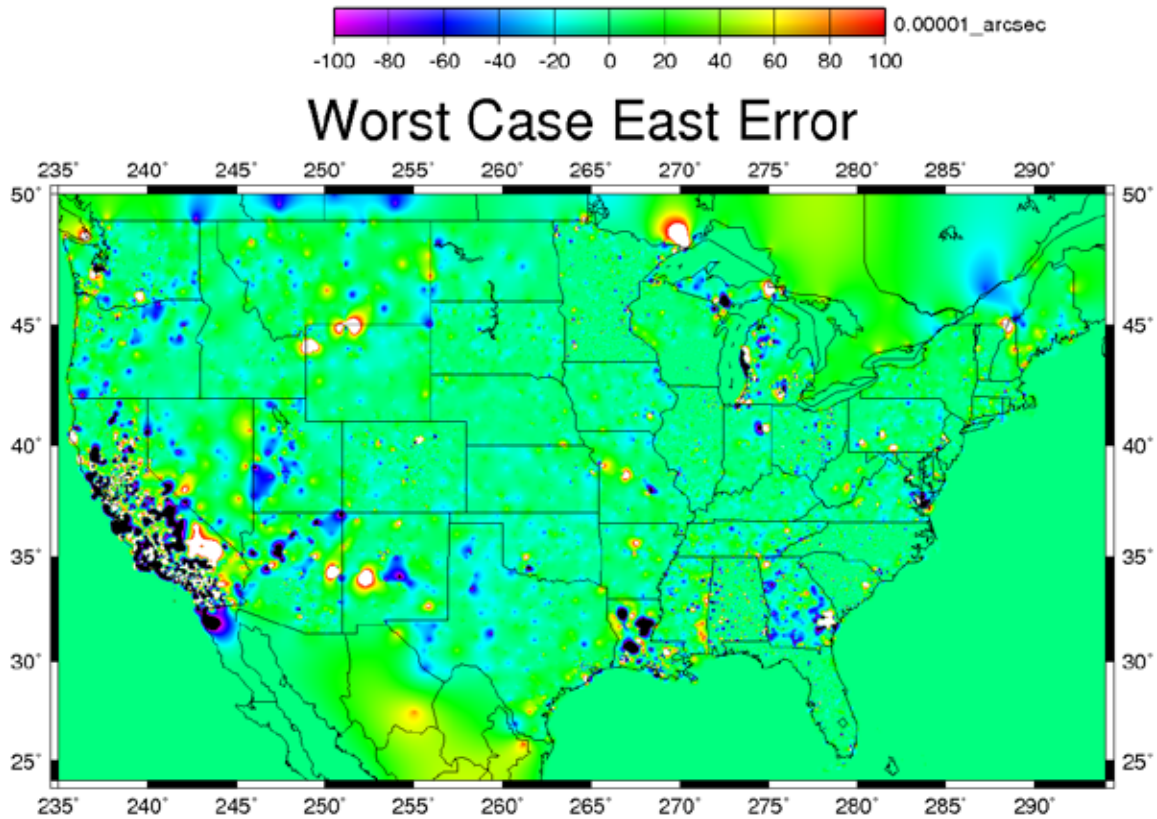Figure 9.3. Worst Case Cross Validation Error, Latitude.

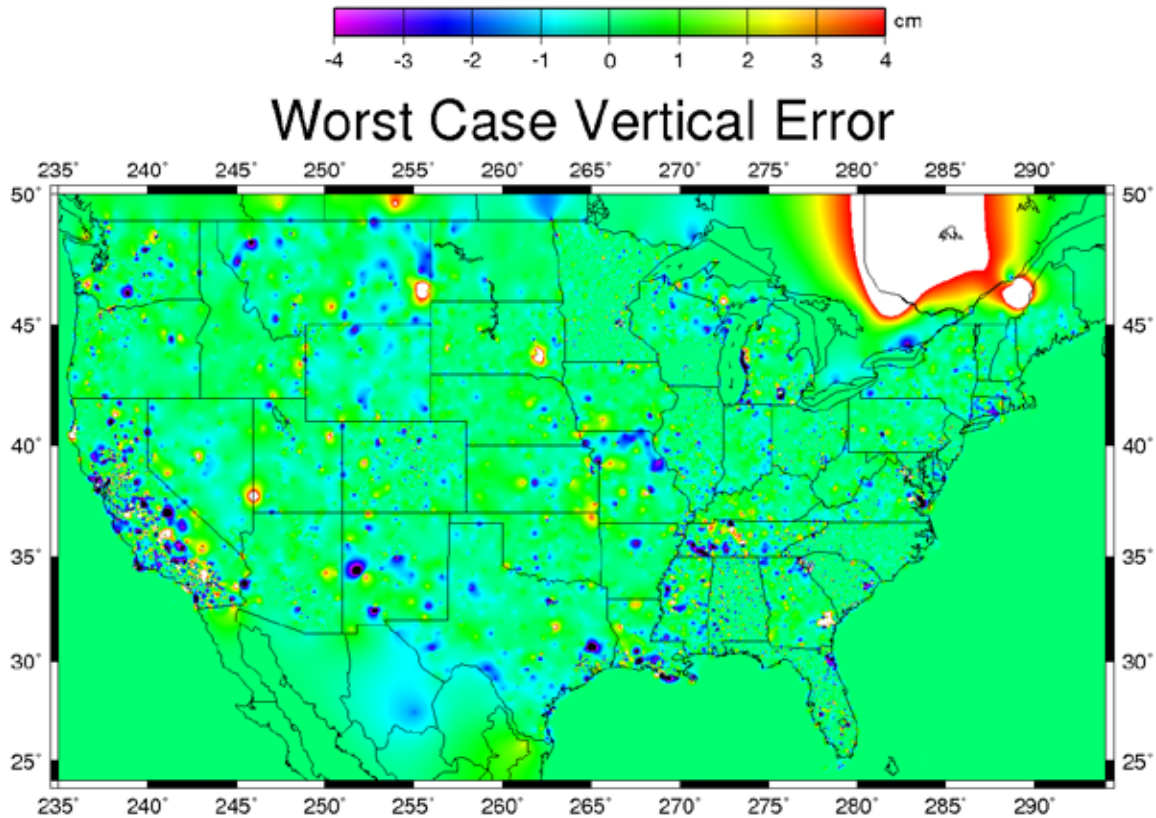Figure 9.4. Worst Case Cross-Validation Error, Longitude.

Figure 9.5. Worst Case Cross-Validation Error, Ellipsoidal Height.

The error grids should be inspected in conjunction with the percentiles of Table 7.1. It is seen that the coordinate shifts can be predicted quite well. The abnormal cases are quite sporadic. California and Louisiana have the most troublesome horizontal coordinate shifts. However, the vertical shifts in Louisiana are seen to be modeled somewhat better in GEOCON.

10. Case Study 1

For the first case study, consider a GPS relative carrier phase survey in Arizona: Project GPS2828. It was observed in February 2006, and was processed in December 2010. The survey extends about 1 degree in latitude and 1.7 degrees in longitude. This survey includes 45 points, including 2 CORS. Eight passive marks were existing control points.

This project provides a means of testing the predictive capability of the GEOCON coordinate transformation. While the raw data existed in 2006, it was not included in the NSRS2007 National Readjustment. The GPS vectors were not generated until the end of 2010. The test will be to see how well HARN coordinates can be transformed from starting NSRS2007 coordinates.

First, a constrained adjustment by means of program ADJUST (Milbert and Kass, 1987) was computed for the NSRS2007 coordinates. Eight passive control points were

held fixed to the NGS database NSRS2007 values. The resulting set of 45 coordinates in NSRS2007 is denoted `bbook2007.txt`.

Then, the *80*/*86* NSRS2007 coordinate records were extracted from `bbook2007.txt`, and transformed from NSRS2007 into HARN by program GEOCON. These transformed coordinates are designated `8086-harn.txt`. No points were clipped, and no notifications were issued. The quality values were all sub-centimeter, and were typically only a few millimeters.

Then, a "truth" coordinate set is generated by another constrained adjustment computed for the HARN coordinate. The same set of 8 control points were held fixed to the database HARN values. This control set in the HARN is named `bbookharn.txt`.

Finally, a simple program, `harnstat`, was written to compare the HARN coordinates between the transformed set and the control set obtained by the HARN constrained adjustment. The output of `harnstat` is presented in Appendix A.2. The sense of the signs is: transformed – adjusted. Longitudes are taken as positive East. All units are millimeters. The results, including the Root Mean Square (RMS) about zero and about the mean, are summarized in Table 10.1.

Table 10.1 – Summary Statistics for the Transformation of GPS2828

Percentile	Latitude (mm)	Longitude (mm)	Height (mm)
Average	-0.3	-0.6	1.1
RMS, zero	1.5	4.2	2.6
RMS, mean	1.5	4.1	2.4

It is seen that GEOCON did a very good job of transforming the NSRS2007 coordinates into the HARN. The RMS about the mean is roughly comparable to the variety of quality values issued by GEOCON.

Of course, in this first example, the GPS survey happens to tie into points that behave very much like their neighbors. None of the 8 fixed points can be considered as having abnormal coordinate shifts between HARN and NSRS2007.

11. Case Study 2: Abnormal Points

The second case study will consist of 3 parts. However, all 3 parts will use an identical set of GPS vectors. Consider a synthetic GPS project in Northeast Colorado shown in Figure 11.1.
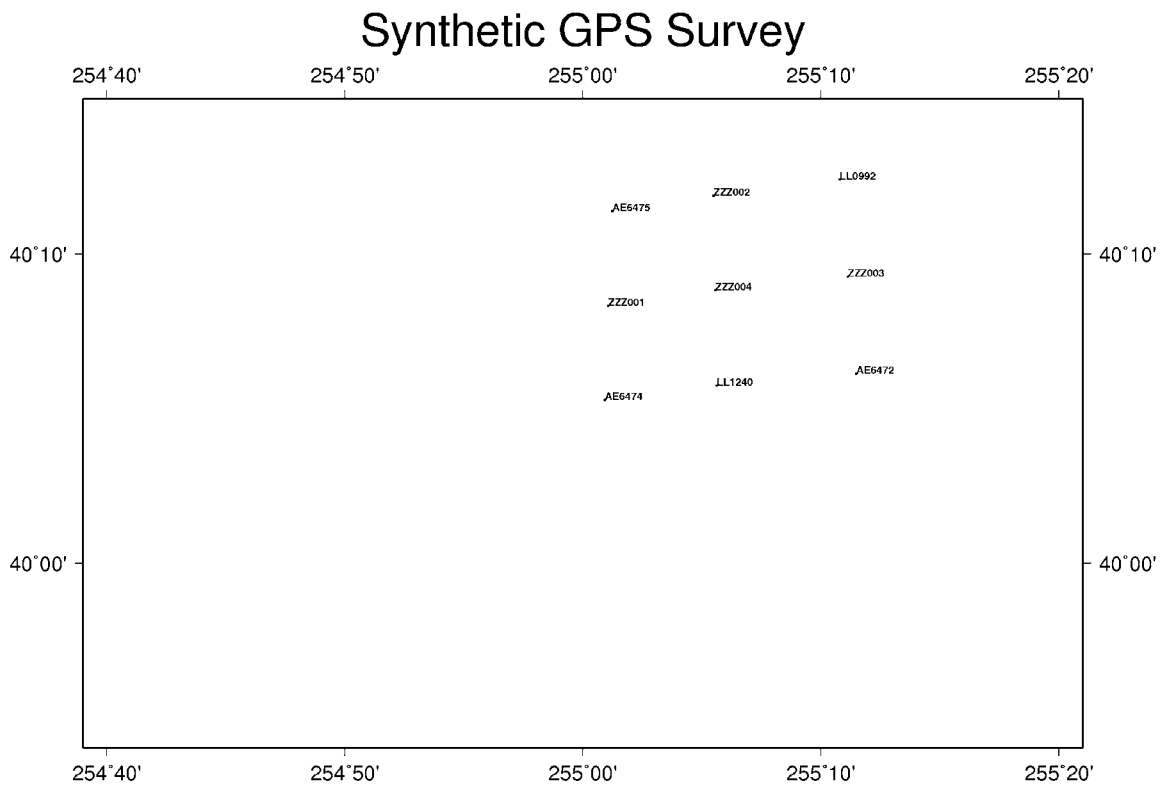
Synthetic GPS Survey

Figure 11.1. A Simulated GPS Survey.

This example represents densification with 4 new points labeled with the prefix "ZZZ". The remaining 5 points are the same control points found in Section 6. The objective is to get a set of HARN coordinates for the new points from NSRS2007 coordinate values using GEOCON. Recall that the coordinate shift difference at LL1240 was abnormal (about 12 cm in longitude). It is important to note that the new points' survey connections to the established control are excellent.

The synthetic GPS data consists of 12 vectors, each in a separate session, with no correlations between components. Each vector connects northward or eastward to an adjacent point. This creates figures of quadrilaterals with 2 vectors in each of the 3 rows and 3 columns. The vector components were created from preliminary coordinates based on NSRS2007. Gaussian random noise of 0.5 cm standard deviation was computed by a Box-Mueller method (Forsythe *et al.*, 1977, pg. 247), and added to each synthetic component.

The first part of this case study begins with a constrained adjustment of AE6474, LL1240, AE6472, AE6743, and LL0992 in NSRS2007. The resulting set of 9 coordinates in NSRS2007 is denoted bbook2007.txt.

As in the first case study, the *80*/*86* NSRS2007 coordinate records were extracted from bbook2007.txt, and transformed from NSRS2007 into HARN by program GEOCON. These transformed coordinates are designated 8086-harn.txt. No points were clipped, and no notifications were issued. But, the quality values ranged up to 8 cm. This is due to the influence of the longitude shift at LL1240.

The control coordinate set is generated by another constrained adjustment computed for the HARN. The same set of 5 control points were held fixed to the database HARN values. This control set in the HARN is named `bbookharn.txt`.

As with the first case study, the transformed coordinates are compared against the HARN-adjusted coordinates. Results in millimeters are displayed in Table 11.1

Table 11.1 – Part 1, 5 Fixed Control Points (units of mm)

```
        program harnstat -- 2012jan18
        input  80/86 file  (input) : 8086-harn.txt
        input  bbook file  (input) : bbookharn.txt
                9  points loaded
                       test comparison      transformation quality
        AE6474  0001      0.0    1.9    0.0        4.6     9.5    -9.9
        LL1240  0002      0.6  -16.7    2.0        3.8   -89.7    12.8
        AE6472  0003      0.3    3.1   -1.0        3.1    32.2   -10.3
        ZZZ001  0004      0.3    2.8    0.0        2.9    -4.4    -3.5
        ZZZ004  0005      0.0    5.3   -3.0        3.7   -22.8     1.1
        ZZZ003  0006      1.9   -2.8   -2.0       11.2    89.9   -24.7
        AE6475  0007      0.0    1.9    0.0        0.4    11.1     1.3
        ZZZ002  0008      0.3    0.0   -5.0        2.4     3.4    -0.6
        LL0992  0009      0.3    0.9   -1.0        4.6    23.0   -10.7
                9  points processed
        ave:     0.4   -0.4   -1.1
        rms0     0.7    6.1    2.2
        rms:     0.5    6.1    1.9
```

GEOCON has done a surprisingly good job in transforming the coordinates from NSRS2007 into the HARN despite the abnormal shift at `LL1240` (serial number `0002`). In fact, the -16.7 mm longitude difference between the two coordinate sets at `LL1240` shows GEOCON has some difficulty in exactly modeling the full amount of the shift at that point. The reason for these excellent results is that the synthetic survey tied into all the surrounding control. This is exactly the scenario that is modeled by the spine fitting method used to develop the GEOCON grids.

In the second part, we keep the same GPS survey. But suppose that the survey was unable to connect into `LL1240`. Perhaps the point couldn't be recovered. Perhaps it was destroyed when the synthetic survey was performed. In any case, at that location a new point is established, named `ZZZ005` (serial number `0002`).

The same procedures are followed. An adjustment is performed with 4 fixed NSRS2007 control points. The 9 coordinate records are transformed from NSRS2007 to HARN with GEOCON. To compare, the same 4 fixed point adjustment is done in the HARN. The transformed coordinates are compared to the adjusted coordinates and given in millimeters in Table 11.2.

Table 11.2 – Part 2, 4 Fixed Control Points (units of mm)

```
        program harnstat -- 2012jan18
        input  80/86 file  (input) : 8086-harn.txt
        input  bbook file  (input) : bbookharn.txt
                9  points loaded
                       test comparison      transformation quality
        AE6474  0001      0.0    1.9    0.0        4.6     9.5    -9.9
```

```
ZZZ005  0002      -3.4  119.2  -12.0        3.8  -89.7   12.8
AE6472  0003       0.3    3.1   -1.0        3.1   32.2  -10.3
ZZZ001  0004      -0.3   17.9   -2.0        2.9   -4.4   -3.5
ZZZ004  0005      -1.2   50.7   -8.0        3.7  -22.8    1.1
ZZZ003  0006       1.5   12.4   -3.0       11.2   89.9  -24.7
AE6475  0007       0.0    1.9    0.0        0.4   11.1    1.3
ZZZ002  0008       0.0   15.1   -7.0        2.4    3.4   -0.6
LL0992  0009       0.3    0.9   -1.0        4.6   23.0  -10.7
           9  points processed
ave:      -0.3   24.8   -3.8
rms0       1.3   44.1    5.5
rms:       1.3   36.5    4.0
```

GEOCON applied a very large longitude shift where ZZZ005 (serial number 0002) is located. And it also applied about half that amount at ZZZ004 (serial number 0005). These happen to be erroneous shifts. This is because in this second part, the synthetic GPS survey was connected solely to points that had common coordinate shifts. And, GEOCON is modeling the longitude abnormality at LL1240.

In the third part of the case study, we keep the same GPS survey. But, suppose that the survey was connected to only one point, LL1240. In this case we have new points ZZZ005, ZZZ006, ZZZ007, and ZZZ008 replacing the old control at the same locations. This is a total of 8 new points, and one old point.

The same adjustment and transformation procedures are followed as before, however with just one fixed point. The transformed coordinates are compared to the adjusted coordinates and given in millimeters in Table 11.3.

Table 11.3 – Part 3, 1 Fixed Control Point (units of mm)

```
  program harnstat -- 2012jan18
input  80/86 file  (input) : 8086-harn.txt
input  bbook file  (input) : bbookharn.txt
          9  points loaded
                 test comparison      transformation quality
ZZZ005  0001     4.6 -132.8    7.0      4.6    9.5   -9.9
LL1240  0002     0.6  -16.7    2.0      3.8  -89.7   12.8
ZZZ006  0003     2.8 -133.1   26.0      3.1   32.2  -10.3
ZZZ001  0004     4.6 -119.2    0.0      2.9   -4.4   -3.5
ZZZ004  0005     3.1  -86.2    0.0      3.7  -22.8    1.1
ZZZ003  0006     5.6 -124.5    8.0     11.2   89.9  -24.7
ZZZ007  0007     5.3 -137.4   -8.0      0.4   11.1    1.3
ZZZ002  0008     4.9 -122.9   -6.0      2.4    3.4   -0.6
ZZZ008  0009     5.6 -136.5    2.0      4.6   23.0  -10.7
          9  points processed
ave:      4.1 -112.2    3.4
rms0      4.4  118.0   10.0
rms:      1.6   36.8    9.4
```

This third part of this case study created a HARN coordinate set that was completely anchored to LL1240. It was only at LL1240 that GEOCON was able to issue a coordinate shift consistent with that particular adjustment.

The longitude transformation quality values in this second case study have an estimated error of -9 cm at LL1240, and -2.3 cm at ZZX004. These errors are much too

23

large for part one of this case study. However, we had ideal connections to both normal and abnormal network control points in part one. These same errors are a little small, but about right for part two. In part two, the survey was only connected to control points with normal longitude shifts. And, these same quality values fail to show the reverse distribution of error seen in part three. Of course, part three is a pathological situation; where the survey is connected solely to an isolated, abnormal control point, and all the existing control in the area was ignored.

As described in Section 6, the cross-validation errors estimate how well one can predict a value at a withheld point. Hence, they can quantify abnormality of a coordinate difference. This quality model is most appropriate when one only connects to normal points. The quality values overestimate when one connects into both normal and abnormal points. But, cross-validation was deliberately selected to provide conservative estimates of error. This is similar to the decision to choose the worst case cross-validation error in the case of a point cluster: provide maximal warning to the user. In the end, the quality values do not reflect survey accuracy. Rather, they indicate our lack of knowledge about how the input coordinates are connected to the national network.

These case studies provide a deeper explanation of why the National Geodetic Survey considers actual recomputation of geospatial data, and not coordinate transformations, as "best practice". The coordinate transformation is, at its heart, only a model of actual geospatial measurement and processing.

In closing this section, it should be noted that the results would have been the same if the densification consisted of a grid of 400 new, interconnected points instead of 4. And, photogrammetric mosaics, or other interconnected geospatial positioning data, would behave similarly to synthetic GPS vectors.

12. Case Study 3: Clusters

The third case study illustrates behavior when control points are clustered. This study also consists of 3 parts. The general approach is similar to the second case study. The study involves a synthetic GPS project somewhat further to the northeast than seen in the second case study. This example also presents a densification survey with 3 new points labeled with "zzz". The arrangement of points is depicted in Figure 12.1.
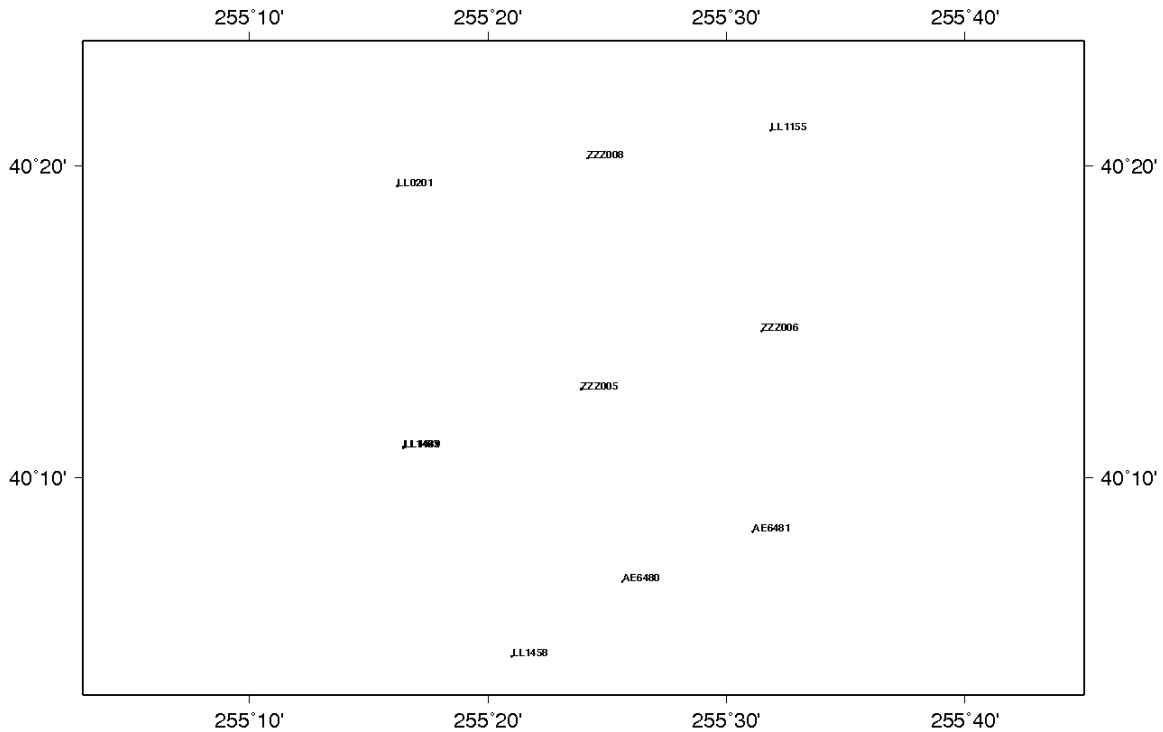
Synthetic GPS Survey

Figure 12.1. Another Simulated GPS Survey.

The West-central location in the array is a cluster with 7 points. For the purposes of this case study, the cluster is thinned down to 3 points: `LL1439`, `LL1465`, and `LL1477`. `LL1439` is a point that generally agrees with its neighbors in the network. `LL1465` and `LL1477`, however, have abnormal coordinate shifts.

The synthetic GPS data consists of 12 vectors with the same general arrangement as in the second case study. However, due to the cluster at the West-Central location, the choice of control point changes for each part of this three-part study. Therefore the synthetic GPS data are regenerated anew for each part.

In the first part of the study, the synthetic survey ties to `LL1439`, a normal point. A set of synthetic data was generated using the NSRS2007 coordinates and adding Gaussian random noise. A constrained adjustment was performed with `LL1458`, `AE6480`, `AE6481`, `LL1155`, `LL0201`, and `LL1439` fixed in NSRS2007. The resulting coordinates were transformed with GEOCON into the HARN. As in the second case study, a separate adjustment of the synthetic GPS data was computed in the HARN. The transformed coordinates are compared to the adjusted coordinates and given in millimeters in Table 12.1.

Table 12.1 – Part 1, All Ties to Normal Points (units of mm)

```
program harnstat -- 2012jan18
input  80/86 file  (input) : 8086-harn.txt
input  bbook file  (input) : bbookharn.txt
          9   points loaded
                      test comparison      transformation quality
```

```
LL1458   0001      0.3    0.3   0.0        5.3     7.2   -10.7
AE6480   0002      0.0   -0.3   0.0        1.8     5.6    -2.8
AE6481   0003      0.0    0.0   0.0        1.4     0.9    -6.2
LL1439   0004      3.4    0.6   2.0       51.2   510.3  -116.8
ZZZ005   0005      1.9   -0.9   0.0       12.5   115.1   -27.3
ZZZ006   0006      0.9   -0.3   2.0        3.7    26.1    -8.6
LL0201   0007      0.3    0.0   0.0        4.5     4.7     2.6
ZZZ008   0008      7.4   -4.6 -10.0        1.5    30.1    -3.9
LL1155   0009      0.3   -0.3   0.0        3.8    -2.8     0.5
             9  points processed
ave:       1.6   -0.6   -0.7
rms0       2.8    1.6    3.5
rms:       2.3    1.5    3.4
```

It is seen that the transformed set of HARN coordinates have excellent agreement with the coordinates generated by the readjustment of the synthetic data in the HARN coordinate. However, very large transformation quality values are issued by GEOCON. In addition, GEOCON issues a pair of notification messages:

```
   *80*0004LL1439                        40105728646N104433127706W150124
Note - poor quality due to LL1465 - unmodeled hztl. error:    72.8 cm.
   *86*0004  1501245              000   1501267
Note - poor quality due to LL1477 - unmodeled vert. error:   -12.6 cm.
```

Both messages refer to station ID 0004, which is LL1439. That point is completely benign and had a good transformation. However, it is known that that NGS database included points LL1465 and LL1477, and they both had abnormal coordinate shifts. The GEOCON transformation grids do not model the behavior of those 2 points. They were in a cluster of 7 points, and were dropped in the gridding process. But, the transformation quality grid is based on the worst case behavior of the cross-validation error. The large values in the transformation quality are a warning survey connections to the troublesome points will lead to coordinate conversion problems. The notifications explain that LL1439 is near those troublesome points.

It happens that the survey in the first part only connected to control with normal coordinate shifts. Hence, the transformation warnings seem much too large. Consider a similar situation, where the synthetic survey does tie to an abnormal point, LL1465.

In the second part, point, LL1465 replaces LL1439. A new set of synthetic data was generated using the NSRS2007 coordinates and adding Gaussian random noise. A constrained adjustment was performed with LL1458, AE6480, AE6481, LL1155, LL0201, and LL1465 fixed in NSRS2007. The resulting coordinates were transformed with GEOCON into the HARN. As in the first part, a separate adjustment of the synthetic GPS data was computed in the HARN. The transformed coordinates are compared to the adjusted coordinates and given in millimeters in Table 12.2.

Table 12.2 – Part 2, Tie to Abnormal Point, LL1465 (units of mm)

```
program harnstat -- 2012jan18
input  80/86 file  (input) : 8086-harn.txt
input  bbook file  (input) : bbookharn.txt
          9  points loaded
                    test comparison      transformation quality
LL1458   0001      0.3    0.3   0.0        5.3     7.2   -10.7
```

```
AE6480  0002      0.0   -0.3    0.0        1.8    5.6   -2.8
AE6481  0003      0.0    0.0    0.0        1.4    0.9   -6.2
LL1465  0004     55.6  728.3  -28.0       51.4  512.2 -117.2
ZZZ005  0005     17.0  217.8   -8.0       12.5  115.1  -27.3
ZZZ006  0006      5.9   72.6   -1.0        3.7   26.1   -8.6
LL0201  0007      0.3    0.0    0.0        4.5    4.7    2.6
ZZZ008  0008     12.7   68.3  -13.0        1.5   30.1   -3.9
LL1155  0009      0.3   -0.3    0.0        3.8   -2.8    0.5
           9  points processed
ave:     10.2  120.7   -5.6
rms0     19.9  255.6   10.6
rms:     17.1  225.2    9.1
```

Inspection of Table 12.2 shows that the transformation quality errors issued in parts one and two of this case study were not too large. In fact, the quality values underestimate the actual error between the transformed and the adjusted coordinate sets. Recall, the same GEOCON transformation grids were used in parts one and two. In the first part, the transformation performed excellently. In the second part, the transformation did not. The difference is due to the way the local data tied into the network. NGS can not know how such projects connect into the national network. But, NGS can warn the user that potential problems might exist.

In the third part, point, LL1477 replaces LL1439 and/or LL1465. New synthetic data was generated as before. A constrained adjustment was performed with LL1458, AE6480, AE6481, LL1155, LL0201, and LL1477 fixed in the NSRS2007 coordinate. The remaining procedures follow parts one and two. The transformed coordinates are compared to the adjusted and given in millimeters in Table 12.3.

Table 12.3 – Part 3, Tie to Abnormal Point, LL1477 (units of mm)

```
program harnstat -- 2012jan18
input  80/86 file  (input) : 8086-harn.txt
input  bbook file  (input) : bbookharn.txt
           9  points loaded
                test comparison      transformation quality
LL1458  0001      0.3    0.3    0.0        5.3    7.2  -10.7
AE6480  0002      0.0   -0.3    0.0        1.8    5.6   -2.8
AE6481  0003      0.0    0.0    0.0        1.4    0.9   -6.2
LL1477  0004     -0.9   -4.3 -126.0       51.4  512.6 -117.3
ZZZ005  0005      0.3   -2.2  -38.0       12.5  115.1  -27.3
ZZZ006  0006      0.3   -0.6  -11.0        3.7   26.1   -8.6
LL0201  0007      0.3    0.0    0.0        4.5    4.7    2.6
ZZZ008  0008      7.1   -4.9  -22.0        1.5   30.1   -3.9
LL1155  0009      0.3   -0.3    0.0        3.8   -2.8    0.5
           9  points processed
ave:      1.6   -0.6   -0.7
rms0      2.8    1.6    3.5
rms:      2.3    1.5    3.4
```

The 7 point cluster that is a foundation for this study contains 2 different anomalous points: LL1465 and LL1477. The first is largely anomalous in latitude and the second is anomalous in height coordinate shift. Table 12.3 illustrates what occurs when the survey ties to the point with the anomalous height shift. The transformation error is distributed very much the same as seen in part two, only in the height component. The transformation quality errors are reflecting both anomalous points.

13. Summary User Guidance

If at all possible, consider reprocessing your geospatial data using control points in the coordinate set you wish to realize. In the case of geodetic survey data, this would entail performing a least squares adjustment with new coordinates as constraints. Notwithstanding the high quality of the GEOCON transformation, the National Geodetic Survey considers actual readjustment of survey measurements, and not coordinate transformations, as "best practice".

It is understood that for many cases it will not be economical, or may be impossible to work with the original geospatial measurements. Further, the network accuracy of the geospatial data set may be a much larger number than the expected quality value of the coordinate transformation. GEOCON provides an attractive solution.

The quality values reported by GEOCON in the *94* records are systematic, and their absolute values should be added to the base network accuracy of the pre-transformed coordinates. If your transformed accuracies meet your needs, then congratulations, you are done.

However, there may be some spots where the transformation does not have the desired quality. These are caused by proximity to abnormal coordinate differences. The abnormal point or points may be isolated or in clusters. Either way, the abnormality will be expressed in the quality grids. Short of reprocessing the raw data, one is faced with researching how the raw data was connected, directly or indirectly, to the network.

In the case of isolated anomalous points, the transformation grid is modeling the general behavior of a data set connected to every point in the region including the anomalous point. If this is the case for the original raw data, the coordinate transformation will be better than the reported quality. On the other hand, if the raw data connected solely to the normal points, or solely to the isolated anomalous point, then the transformation should be applied manually.

In the case of an anomalous point in a cluster, the transformation grid is modeling the connection to the normal points, including the median normal point in the cluster. If this is the case for the original raw data, the coordinate transformation will be better than the reported quality. However, if the raw data connected to the anomalous point and other normal points, then one must estimate how the anomalous coordinate shift will be distributed between the anomalous point and its neighbors. And, if the raw data connected solely to the anomalous point in the cluster, then the transformation based on the anomalous point should be applied manually.

ACKNOWLEDGEMENTS

REFERENCES

Efron, B., and R. J. Tibshirani, 1998: *An Introduction to the Bootstrap*. Chapman & Hall/CRC, New York, 436 pp.

Efron, B., 1979: Computers and the theory of statistics: thinking the unthinkable. *SIAM Rev.*, 21(4), 460-480. Web page (PDF document) at: http://statistics.stanford.edu/~ckirby/techreports/BIO/BIO%2039.pdf

Forsythe, G. E., M. A. Malcolm, and C. B. Moler, 1977: *Computer Methods for Mathematical Computations*. Prentice-Hall, Englewood Cliffs, NJ, 259 pp.

Milbert, D., 2008: An analysis of the NAD 83(NSRS2007) national readjustment. *Sponsored report*, National Geodetic Survey, NOAA, Silver Spring, MD, January 2008, pp.182. Web page (PDF document) at: http://www.ngs.noaa.gov/PUBS_LIB/NSRS2007/NSRS2007Analysis.pdf

Milbert, D. G., W. G. Kass, 1987: ADJUST: the horizontal observation adjustment program. *NOAA Technical Memorandum* NOS NGS-47, National Geodetic Survey, NOAA, Silver Spring, MD, September 1987, 53pp. Web page (PDF document) at: http://www.ngs.noaa.gov/PUBS_LIB/Adjust_TheHorizontalObservationAdjustmentProgram_TM_NOS_NGS_47.pdf

Milbert, K. O. and D. G. Milbert, 1994: State readjustments at the National Geodetic Survey. *Surv. Land Info. Sys.*, 54(4), 219-230.

Smith, W. H. F., and P. Wessel, 1990: Gridding with continuous curvature splines in tension. *Geophysics*, 55(3), 293-305.

Wessel, P. and W. H. F. Smith, 1995: New version of the Generic Mapping Tools released. *EOS Trans. Amer. Geophys. U.*, 76(33), 329 pp.

About the Author

Dr. Dennis Milbert received his Bachelors degree in Physics from the University of Colorado, and his M.S. and Ph.D. degrees in Geodetic Science from The Ohio State University. He worked for over 29 years at the National Geodetic Survey (NGS) of the National Oceanic and Atmospheric Administration, where he was promoted to the position of Chief Geodesist. In his federal career he developed accuracy standards, adjustment software, gravity and geoid models, GPS kinematic surveys, and vertical datum transformations. Dr. Milbert served on numerous federal technical and policy working groups and he was an alternate representative to the Senior Steering Group of the Interagency GPS Executive Board. He served for eight years on the joint Editorial Board for Manuscripta Geodetica/Bulletin Geodesique and the Journal of Geodesy. Dr. Milbert is a recipient of the Kaarina and Weikko A. Heiskanen Award, the NOAA Administrator's Award, the Department of Commerce Bronze Medal, and two Department of Commerce Silver Medals. He is a member of the American Geophysical Union, the International Association of Geodesy, and the Institute of Navigation. Dr. Milbert retired from government service in 2004 and pursues research in various geodesy topics.

APPENDIX

A.1. Datasheet for M 123

```
1        National Geodetic Survey,   Retrieval Date = NOVEMBER 25, 2011
 TT2413 *************************************************************************
 TT2413  DESIGNATION -  M 123
 TT2413  PID        -  TT2413
 TT2413  STATE/COUNTY-  AK/DENALI BOROUGH
 TT2413  USGS QUAD  -  HEALY C-4
 TT2413
 TT2413                          *CURRENT SURVEY CONTROL
 TT2413  _____
 TT2413* NAD 83(2007)-  63 43 21.43573(N)    148 57 43.76796(W)    ADJUSTED
 TT2413* NAVD 88     -        628.189  (meters)   2060.98  (feet)  ADJUSTED
 TT2413  _____
 TT2413  EPOCH DATE  -        2007.00
 TT2413  X          -  -2,426,212.599 (meters)                     COMP
 TT2413  Y          -  -1,459,997.451 (meters)                     COMP
 TT2413  Z          -   5,696,668.363 (meters)                     COMP
 TT2413  LAPLACE CORR-        -8.38  (seconds)                      DEFLEC09
 TT2413  ELLIP HEIGHT-        641.786 (meters)         (07/17/09)   ADJUSTED
 TT2413  GEOID HEIGHT-         13.66  (meters)                      GEOID09
 TT2413  DYNAMIC HT  -        629.112 (meters)   2064.01  (feet)    COMP
 TT2413  MODELED GRAV-    982,033.7  (mgal)                         NAVD 88
 TT2413
 TT2413  HORZ ORDER  -  FIRST
 TT2413  VERT ORDER  -  FIRST     CLASS II
 TT2413  ELLP ORDER  -  FOURTH    CLASS I
 TT2413
 TT2413.The horizontal coordinates were established by GPS observations
 TT2413.and adjusted by the National Geodetic Survey in July 2009.
 TT2413
 TT2413.The datum tag of NAD 83(2007) is equivalent to NAD 83(NSRS2007).
 TT2413.The horizontal coordinates are valid at the epoch date displayed above.
 TT2413.The epoch date for horizontal control is a decimal equivalence
 TT2413.of Year/Month/Day.
 TT2413
 TT2413.The orthometric height was determined by differential leveling and
 TT2413.adjusted in June 1991.
 TT2413.WARNING-GPS observations at this control monument resulted in a GPS
 TT2413.derived orthometric height which differed from the leveled height by
 TT2413.more than one decimeter (0.1 meter).
 TT2413
 TT2413.The X, Y, and Z were computed from the position and the ellipsoidal ht.
 TT2413
 TT2413.The Laplace correction was computed from DEFLEC09 derived deflections.
 TT2413
 TT2413.The ellipsoidal height was determined by GPS observations
 TT2413.and is referenced to NAD 83.
 TT2413
 TT2413.The geoid height was determined by GEOID09.
 TT2413
 TT2413.The dynamic height is computed by dividing the NAVD 88
 TT2413.geopotential number by the normal gravity value computed on the
 TT2413.Geodetic Reference System of 1980 (GRS 80) ellipsoid at 45
 TT2413.degrees latitude (g = 980.6199 gals.).
 TT2413
 TT2413.The modeled gravity was interpolated from observed gravity values.
 TT2413
 TT2413;                   North          East     Units Scale Factor Converg.
 TT2413;SPC AK 4    - 1,083,325.355   551,279.408   MT  0.99993219  +0 55 50.2
 TT2413;UTM  06     - 7,067,593.809   403,087.880   MT  0.99971501  -1 45 34.3
 TT2413
 TT2413!           - Elev Factor  x  Scale Factor =   Combined Factor
 TT2413!SPC AK 4   -  0.99989959   x   0.99993219  =   0.99983179
 TT2413!UTM  06    -  0.99989959   x   0.99971501  =   0.99961463
 TT2413
 TT2413                    SUPERSEDED SURVEY CONTROL
 TT2413
 TT2413  NAD 83(2007)-  63 43 21.43619(N)    148 57 43.77233(W) AD(2007.00) 0
 TT2413  ELLIP H (02/10/07)  486.989  (m)                       GP(2007.00)
 TT2413  NAD 83(1992)-  63 43 21.43370(N)    148 57 43.78028(W) AD(      ) 1
```

```
TT2413  ELLIP H (06/20/05)  486.945  (m)                        GP(      ) 4 1
TT2413  NAVD 88 (07/17/09)  628.19   (m)      2061.0  (f) LEVELING    3
TT2413  NAVD 88 (06/20/05)  474.1    (m)      1555.   (f) GPS OBS
TT2413  NGVD 29 (??/??/92)  626.505  (m)      2055.46 (f) ADJ UNCH   1 2
TT2413
TT2413.Superseded values are not recommended for survey control.
TT2413.NGS no longer adjusts projects to the NAD 27 or NGVD 29 datums.
TT2413
TT2413_U.S. NATIONAL GRID SPATIAL ADDRESS: 6VVR0308767593(NAD 83)
TT2413_MARKER: DD = SURVEY DISK
TT2413_SETTING: 36 = SET IN A MASSIVE STRUCTURE
TT2413_SP_SET: SET IN BRIDGE
TT2413_STAMPING: M 123
TT2413_MAGNETIC: N = NO MAGNETIC MATERIAL
TT2413_STABILITY: B = PROBABLY HOLD POSITION/ELEVATION WELL
TT2413_SATELLITE: THE SITE LOCATION WAS REPORTED AS SUITABLE FOR
TT2413+SATELLITE: SATELLITE OBSERVATIONS - May 31, 1989
TT2413
TT2413  HISTORY    - Date     Condition        Report By
TT2413  HISTORY    - 1965     MONUMENTED       CGS
TT2413  HISTORY    - 19890531 GOOD             JOA
TT2413
TT2413                         STATION DESCRIPTION
TT2413
TT2413'DESCRIBED BY COAST AND GEODETIC SURVEY 1965
TT2413'1.8 MI W FROM MCKINLEY PARK.
TT2413'1.8 MILES WEST ALONG THE MCKINLEY PARK HIGHWAY FROM THE RAILROAD
TT2413'STATION AT MCKINLEY PARK, 0.2 MILE NORTHEAST OF THE ENTRANCE TO THE MT
TT2413'MCKINLEY NATIONAL PARK HEADQUARTERS, 12 FEET SOUTHEAST OF THE CENTER
TT2413'LINE OF THE HIGHWAY, SET IN THE TOP OF THE EAST CURB OF A CONCRETE
TT2413'BRIDGE OVER ROCK CREEK, 15 FEET SOUTHWEST OF THE NORTHEAST END OF A
TT2413'CONCRETE GUARDRAIL, AND ABOUT 0.6 FOOT HIGHER THAN THE ROAD.
TT2413
TT2413                         STATION RECOVERY (1989)
TT2413
TT2413'RECOVERY NOTE BY JOHN OSWALD AND ASSOCIATES, LLC 1989
TT2413'THE STATION IS LOCATED 0.2 KM (0.10 MI) NORTH EAST OF THE MCKINLEY
TT2413'PARK HEADQUARTERS, 6.3 KM (3.90 MI) WEST OF THE PARKS HIGHWAY, IN A
TT2413'BRIDGE OVER ROCK CREEK. OWNERSHIP, NATIONAL PARK SERVICE. TO REACH
TT2413'THE STATION FROM THE INTERSECTION OF THE PARKS HIGHWAY AND THE
TT2413'MCKINLEY PARK ROAD, DRIVE WEST ON THE ACCESS ROAD 6.3 KM (3.90 MI) TO
TT2413'THE ROCK CREEK BRIDGE. STATION IS ON THE LEFT. THE STATION IS A 7.6
TT2413'CM BRASS DISK SET IN THE SOUTH EAST BRIDGE ABUTMENT OF THE ROCK CREEK
TT2413'BRIDGE, LOCATED 3.6 M (11.8 FT) SOUTHEAST OF THE CENTERLINE OF THE
TT2413'ROAD, 4.5 M (14.8 FT) SOUTHWEST OF THE END OF THE ABUTMENT, 0.3 M (1.0
TT2413'FT) NORTHEAST OF A WITNESS DECAL ON THE RAIL OF THE BRIDGE. NOTE, THE
TT2413'MAGNETIC PROPERTIES OF THE STATION ARE UNKNOWN.
```

A.2. Output for the GPS2828 Case Study

```
C>harnstat
 program harnstat -- 2012jan18
input  80/86 file  (input) : 8086-harn.txt
input  bbook file  (input) : bbookharn.txt
         45  points loaded
1001     0.6     1.5     0.0
1002     0.0    -3.7     1.0
1003     1.2    -4.6    -2.0
1004     0.3    -0.6     0.0
1005     0.9     0.6     1.0
1006     0.3    -3.7    -1.0
0001     0.0    -0.6    -2.0
1007    -0.9    -4.6     0.0
1008     0.0    -0.3     0.0
1009    -0.9    -3.4     4.0
1010     0.3     0.9     0.0
0002    -7.7    17.9     7.0
1011     0.0     2.8     1.0
1012     0.0     1.2     0.0
1013     0.3     0.0     0.0
1014     0.0    -2.8     3.0
1015     0.6    -0.6     1.0
1016    -2.8    -5.6     2.0
1017    -2.2    -5.3     6.0
1018     0.3     0.0     1.0
1019    -0.9    -4.6     3.0
1020     0.0     2.2     1.0
1021    -2.2    -4.9     4.0
1022     0.0    -1.2     2.0
1023     0.0     0.0     0.0
1024    -1.2    -3.7     5.0
1025    -0.6     3.7     1.0
1026    -1.2    -4.0     7.0
1027     0.0    -0.6     0.0
1028    -0.9     7.1     2.0
1030     0.9    -4.3    -3.0
1029    -0.3    -2.2     4.0
1031     0.3     1.5     0.0
1032     1.2    -0.3    -1.0
1033    -0.3     4.9     2.0
1034     0.3     0.9     1.0
1035    -2.2    -4.9     3.0
1036     0.0    -4.0    -1.0
1037     2.5    -6.5    -5.0
1038     0.0    -0.3     0.0
1039     0.6     0.0     0.0
1040     0.0     0.0    -1.0
1041     0.3    -0.6     1.0
1042     0.0     1.5     2.0
1043    -0.3     3.1     1.0
         45  points processed
ave:    -0.3    -0.6     1.1
rms0     1.5     4.2     2.6
rms:     1.5     4.1     2.4
```