NOAA Technical Report NOS NGS 59

GEOCON v2.0 Technical Report

Dru Smith

Silver Spring, MD
2015

GEOCON v2.0

Technical Report

The purpose of this report is to document the creation of the program **GEOCON v2.0** for both scientific understanding of a general nature as well as to provide a roadmap for future NGS employees to be able to fully replicate its creation. As such, parts of the report which represent internal information (names of files and directories or location of internal documentation) are tagged with an endnote, and the endnotes[1] will be kept as a separate portion of this report for internal use only.

Purpose

GEOCON v2.0 incorporates, and expands upon, the functionality of GEOCON v1.1 and GEOCON11 v1.1. Driven primarily by customer demand in certain states, GEOCON v2.0 represents a philosophical shift from those two predecessor programs in many ways. First and foremost, the way by which GEOCON v2.0 incorporates realizations of NAD 83 between 1986 and 2011 is significantly different than GEOCON or GEOCON11. The change in philosophy can best be understood by explaining how the three programs dealt with the realizations between 1986 and 2011.

How realizations of NAD 83 between 1986 and 2011 were treated in different software…

GEOCON (v1.0 or v1.1)

> In this software, transformations were built for latitude, longitude and ellipsoid height in three regions (CONUS including DC, Alaska and Puerto Rico/Virgin Islands) connecting "the most recent published post-1986, pre-2007" NAD 83 realization on any point with NAD 83(NSRS2007) on that same point. That "most recent…" description was somewhat misleadingly called "NAD 83(HARN)" in the original documentation for GEOCON. The transformation might have (in fact did) include a vast mix of realizations in any given state. For example, in the state of Ohio, where the HARN was actually published as "NAD 83(1995)" and over 6000 NAD 83(1995) published points exist in that state, there are also over 1200 points whose "most recent post-1986, pre-2007" published realization were in NAD 83(1992), NAD 83(1993), NAD 83(1994), NAD 83(1996) and NAD 83(1997). This happened in every state. Transformations were performed on a region-wide basis, with no regard for what state or territory the data was in *within* that region.

GEOCON11 (v1.0 or v1.1)

> Functioning much like GEOCON, the GEOCON11 program transformed between NAD 83(NSRS2007) and NAD 83(2011) in the same three regions as GEOCON. As the two realizations supported were both from a single adjustment for the regions in question, there was no question of "multiple realizations" in this program. Transformations were

performed on a region-wide basis, with no regard for what state or territory the data was in *within* that region.

GEOCON v2.0

GEOCON v2.0 rigorously defines "supported realizations" of NAD 83 in each state and territory of the nation, has transformations between those supported realizations *only* and any data that did not fit within those realizations was discarded from influencing those transformations. In 19 states (all in CONUS) exactly 2 supported realizations were defined between 1986 and 2007. In the remaining 29 CONUS states as well as Alaska and DC, exactly 1 supported realization was defined between 1986 and 2007. For Puerto Rico there were 3 and for the US Virgin Islands there were 2. All of these states and territories also had an NAD 83(2011) realization. Within this software and this documentation and for the states and territories mentioned above, the last supported realization, prior to 2007, in a state or territory was always called the "FBN". If there was an earlier supported realization prior to the FBN it was referred to as the "HARN". (In the case of Puerto Rico, the earliest realization is called HARN0 and between that and the FBN was the realization called the HARN).

For the remaining states and territories of the USA, there was no NAD 83(NSRS2007) realization but there were NAD 83(MA11) or NAD 83(PA11) realizations (both from 2011), and so transformations were built for these areas which connected the MA11 or PA11 realizations to pre-2007 "supported realizations". For Hawaii, between 1986 and 2011 there was 1 supported realization. For American Samoa and CNMI there were 2 supported realizations between 1986 and 2011. For Guam, there were also 2 supported realizations (1993:HARN and 2002:FBN), but there was only 1 published point connecting these two, so no grids were created in GEOCON V2.0 to support the transformation between these two realizations.

Within this software and this documentation and for the states and territories mentioned above, the last supported realization, prior to 2011, in a state or territory was always called the "FBN". If there was an earlier supported realization it was referred to as the "HARN".

Prior to creating any transformations, all data in a state or territory that was not in a "supported realization" was discarded from consideration.

Every state or territory with a pre-FBN (that is a "HARN" or, in the case of Puerto Rico a "HARN0" and "HARN") supported realization had a distinct transformation (HARN to FBN) created just for that state or territory to get users to the FBN (for Puerto Rico, two transformations were built: HARN0 to HARN and then HARN to FBN.) Then, in much the same way as v1.0, all of the FBN's are tied in a regional transformation to NSRS2007 without regard for what state or territory they are in.

Other changes were made in GEOCON v2.0 from v1.0. They are:

1) *All* published data in supported realizations was used. For instance, some points had a published latitude and longitude in two supported realizations but not an ellipsoid height in both realizations. In GEOCON v2.0, such points were used to build the latitude and longitude transformation grids, but not used to create the ellipsoid height transformation grids. In GEOCON v1.0 such a point was simply discarded entirely.
2) When multiple transformations are performed in GEOCON v2.0, the "quality" of each transformation was added linearly into a final "quality". This is due to the systematic, and not random, nature of the quality estimates. Only the final transformed coordinates are reported in the standard output file.
3) When multiple transformations are performed, the option of an "extended output" is allowed, where each individual transformation is reported.

Overview

The functionality of GEOCON v2.0 was described in the previous section. However, it may be easiest to understand through the following graphics. In the two figures below arrows represent all available transformations in GEOCON v2.0. Blue arrows represent a transformation that is only valid in one state or territory. Red arrows represent transformations that span entire regions. Only 19 states in CONUS have a state-specific HARN/FBN transformation, but once in the FBN realization, all data in all states of CONUS may be transformed in one step to NSRS2007. Also note that the "regions" of Alaska, Hawaii and American Samoa all contain only one state or territory, but they are referred to as "regions" for the sake of consistency with other regional transformations in GEOCON v2.0.

The 2-letter state and territory codes in those figures should be generally familiar to readers, though some are special codes used in the NGS Integrated Database (specifically: CQ = CNMI, GQ = Guam, VQ = US Virgin Islands).

REGION: CONUS (48 states + DC)

AL		1992[FBN]	NSRS2007	2011
AR		1997[FBN]	NSRS2007	2011
AZ		1992[FBN]	NSRS2007	2011
CA	1992[HARN]	1998[FBN]	NSRS2007	2011
CO		1992[FBN]	NSRS2007	2011
CT	1992[HARN]	1996[FBN]	NSRS2007	2011
DC		1991[FBN]	NSRS2007	2011
DE		1991[FBN]	NSRS2007	2011
FL	1990[HARN]	1999[FBN]	NSRS2007	2011
GA		1994[FBN]	NSRS2007	2011
IA		1996[FBN]	NSRS2007	2011
ID	1992[HARN]	1999[FBN]	NSRS2007	2011
IL		1997[FBN]	NSRS2007	2011
IN		1997[FBN]	NSRS2007	2011
KS		1997[FBN]	NSRS2007	2011
KY		1993[FBN]	NSRS2007	2011
LA		1992[FBN]	NSRS2007	2011
MA	1992[HARN]	1996[FBN]	NSRS2007	2011
MD		1991[FBN]	NSRS2007	2011
ME	1992[HARN]	1996[FBN]	NSRS2007	2011
MI		1994[FBN]	NSRS2007	2011
MN		1996[FBN]	NSRS2007	2011
MO		1997[FBN]	NSRS2007	2011
MS		1993[FBN]	NSRS2007	2011
MT	1992[HARN]	1999[FBN]	NSRS2007	2011
NC	1995[HARN]	2001[FBN]	NSRS2007	2011
ND		1996[FBN]	NSRS2007	2011
NE		1995[FBN]	NSRS2007	2011
NH	1992[HARN]	1996[FBN]	NSRS2007	2011
NJ	1992[HARN]	1996[FBN]	NSRS2007	2011
NM		1992[FBN]	NSRS2007	2011
NV	1994[HARN]	1999[FBN]	NSRS2007	2011
NY	1992[HARN]	1996[FBN]	NSRS2007	2011
OH		1995[FBN]	NSRS2007	2011
OK		1993[FBN]	NSRS2007	2011
OR	1991[HARN]	1998[FBN]	NSRS2007	2011
PA		1992[FBN]	NSRS2007	2011
RI	1992[HARN]	1996[FBN]	NSRS2007	2011
SC	1995[HARN]	2001[FBN]	NSRS2007	2011
SD		1996[FBN]	NSRS2007	2011
TN	1990[HARN]	1995[FBN]	NSRS2007	2011
TX		1993[FBN]	NSRS2007	2011
UT		1994[FBN]	NSRS2007	2011
VA		1993[FBN]	NSRS2007	2011
VT	1992[HARN]	1996[FBN]	NSRS2007	2011
WA	1991[HARN]	1998[FBN]	NSRS2007	2011
WI	1991[HARN]	1997[FBN]	NSRS2007	2011
WV		1995[FBN]	NSRS2007	2011
WY		1993[FBN]	NSRS2007	2011

Figure 1: Transformations available in GEOCON v2.0 for the region CONUS. Blue arrows indicate transformations restricted to one state or territory. Red arrows indicate transformations which can be applied regionally.

REGIONS: Alaska, American Samoa, Guam and CNMI, Hawaii, Puerto Rico and USVI

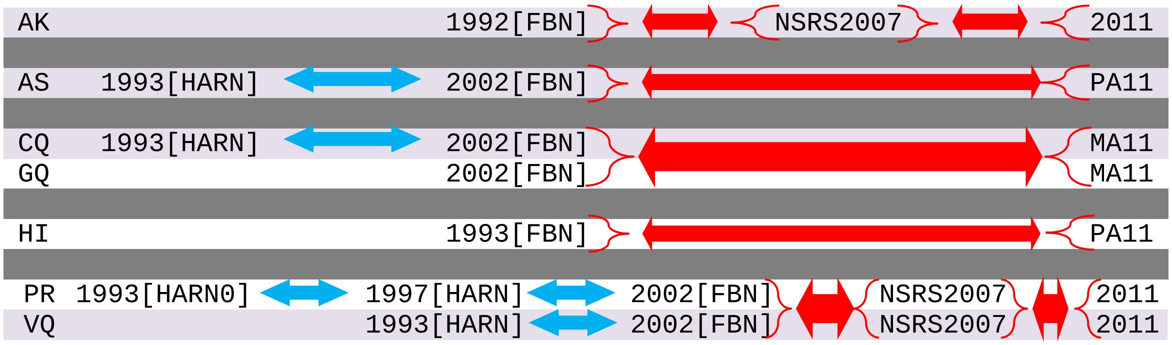

Figure 2: Transformations available in GEOCON v2.0 for the five regions of Alaska, American Samoa, Guam/CNMI, Hawaii and Puerto Rico/Virgin Islands. Blue arrows indicate transformations restricted to one state or territory. Red arrows indicate transformations which can be applied regionally.

<u>Deciding to create GEOCON v2.0</u>

<u>Brief history of v1.0</u>

Almost immediately after the February 2007 release of the NAD 83(NSRS2007) realization, the user community began requesting a transformation tool to connect pre-2007 realizations to NSRS2007. Because the known changes going from HARN or FBN into NSRS2007 were at the few cm level, NGS initially felt such a tool would not have great usefulness.[2]. However, in response to growing pressure from the NSRS user community, which peaked in 2011[3], NGS agreed to investigate the creation of such a tool. Dr. Dennis Milbert (former Chief Geodesist of NGS, retired) was hired under contract to investigate, and possibly create, such a tool. He began with a scientific feasibility study, which showed that the systematic nature of the transformation, while only at the few cm level, was definitely a signal that could be captured and distributed to the public in a gridded transformation tool. That result justified continuing the contract and creating the tool GEOCON v1.0.

Part way through the creation of GEOCON v1.0, the decision was made to also develop GEOCON11 v1.0 under the same contractual work agreement with Dr. Milbert.[4]

<u>Feedback on v1.0</u>

GEOCON v1.0 and GEOCON11 v1.0 were released on NGS's beta website[5] on January 18, 2013. A "feedback link" was provided and almost immediately, feedback came in. While the tools received a generally positive response, the primary request for improvement was that the tool allow for other input and output formats than just FGCS Bluebook[6]. This request was the driver for the creation of GEOCON v1.1 and GEOCON11 v1.1, which was created by Dr. Dru Smith, Chief Geodesist of NGS between January and March of 2014 which used identical transformation files as their v1.0 counterparts, but allowed users "free format" input and output, consolidated the information about "quality" of the transformation into the output file (reducing 4 output files to 3) and consolidated manuals (going from 3 reports to 2). Other than that, v1.1 operated identically to v1.0.

The second most frequent piece of feedback came from those states[7] whose users were had exclusively adopted either the HARN or FBN realizations in their state to the exclusion of using the other realization. Those users wished to have GEOCON distinguish between those realizations, transform between them and then transform into NSRS2007. This eventually led to the creation of GEOCON v2.0 which was created by Dr. Dru Smith, Chief Geodesist of NGS between April 2013 and August 2014.

Supported Realizations

Of all the issues which needed to be addressed for GEOCON v2.0 to become a reality, the decision to define "supported realizations" for each state and territory was one of the key ones. But it was not a decision easily or lightly made. Changing GEOCON to recognize and support multiple realizations was obviously seen as a daunting task just from a software standpoint. User needs on a state by state (or territory by territory) basis had to be made. Decisions to keep or discard data not in those realizations was weighed against the usefulness of that data and expectations of the user community.

It was primarily through surveying the advisors and coordinators of NGS that decisions were made which reflected the needs and desires of users in every state and territory in the country, and led to a planned way forward.

Defining that way forward marked the point when a decision to build GEOCON v2.0 was made, and the work of actually constructing the tool began. The final decision on what constituted a "supported realization" in each state or territory, and whether transformations would be applied regionally or restricted to one state or territory, can be seen earlier in Figures 1 and 2.

Creating GEOCON v2.0[8]

The decision to create GEOCON v2.0 can be traced to about April of 2013, when the choice to define "supported realizations" in each state or territory was made. By June 6, 2013, the supported realization list was complete.[9]

To begin, Figure 3 will help distinguish earlier versions of GEOCON and GEOCON11 from GEOCON v2.0.

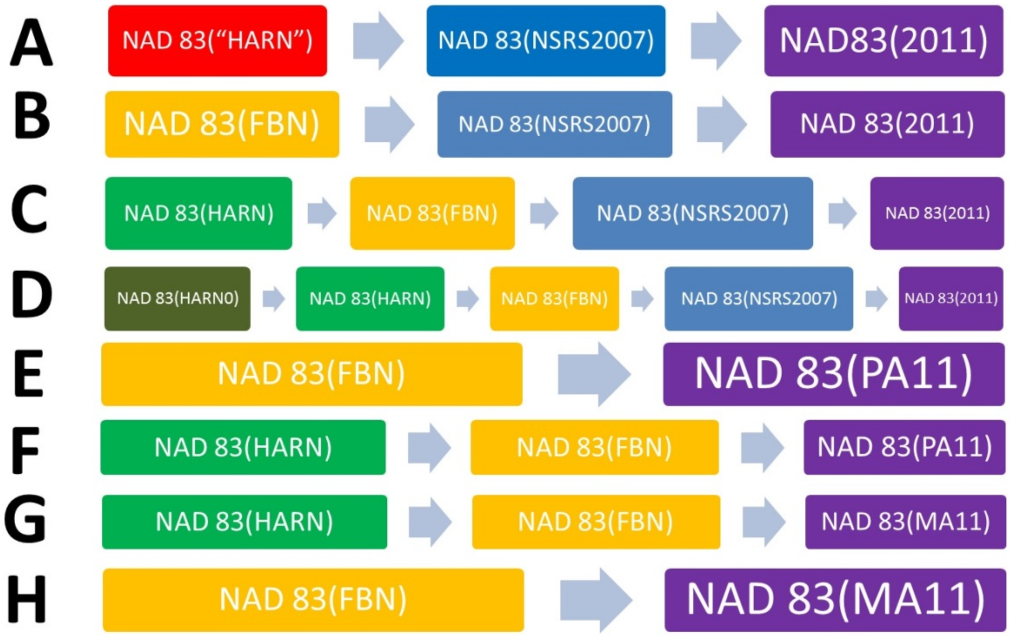

Figure 3: Transformation paths available through GEOCON & GEOCON11. Path "A" represents v1.1, while paths B-H represent GEOCON v2.0.

In Figure 3, the eight paths (A-H) represent the following:

A) This is the method of **GEOCON v1.0/v1.1** and **GEOCON11 v1.0/v1.1.** "HARN" remains in quotes because it is by no means distinctly just the HARN of the state/territory in which the point resides. It is valid only in CONUS, Alaska and PR/VI.

<u>Paths B through H represent all of the various functionality of GEOCON v2.0:</u>

B) This is the method of **GEOCON v2.0** for those states where a HARN was originally published, but no discernable difference between the HARN and the FBN exists, so that there is effectively just one realization of NAD 83 between 1986 and 2007. For the sake of consistency (specifically allowing the FBN-to-NSRS2007 transformation be done on a regional, not statewide, basis) this one realization is called the "FBN" in this state. The states which follow this path, and the year of their HARNs are found in Table 1.

C) This is the method of **GEOCON v2.0** for those states or territories where the HARN was published, but where a new, second supported realization was published during the FBN survey. Then no new supported realization of NAD 83 was published in that state until NAD 83(NSRS2007). The states and territories which follow this path, and the year of their HARNs and FBNs are found in Table 2.

D) This is the method of **GEOCON v2.0** specifically for Puerto Rico, where the HARN was published in 1993 and stood for four years but then a significant systematic error (on the order of 30 cm horizontal) was found and corrected in 1997. Because the first HARN (to be called "HARN0" herein) stood for years and the entire territory needed correction of such a significant size, the users in this territory have specifically requested that GEOCON support transformations between the old (erroneous) HARN and the new (corrected) HARN. In addition to the corrected HARN, a new, third supported realization was published during the FBN survey in 2002. Then no new supported realization of NAD 83 was published in Puerto Rico until NAD 83(NSRS2007).

E) This is the method of **GEOCON v2.0** specifically for Hawaii. Hawaii did not participate in the HARN surveys, but a 1993 GPS survey defined a realization of NAD 83 for the state. Hawaii was not part of the NSRS2007 project, so the next supported realization was PA11.

F) This is the method of **GEOCON v2.0** specifically for American Samoa, with two realizations after 1986 (in 1993 and 2002), but without participating in the NSRS2007 project. As such, the FBN is connected directly to the PA11 realization.

G) This is the method of **GEOCON v2.0** for the Commonwealth of the Northern Mariana Islands, with two realizations after 1986 (in 1993 and 2002), but without participating in the NSRS2007 project. As such, the FBN is connected directly to the MA11 realization.

H) This is the method of **GEOCON v2.0** for Guam, with one realization after 1986 (in 2002), but without participating in the NSRS2007 project. As such, the FBN is connected directly to the MA11 realization. (In truth, there was one other realization, NAD 83(1993), in Guam where a singular point had a latitude and longitude published in NAD 83(1993) and also in NAD 83(2002). No support for the transformation between these two realizations was built into GEOCON v2.0 due to the fact that a single point is not a good data set for building an island-wide transformation. Nonetheless, the coordinate difference information is provided in the GEOCON v2.0 User's Guide.)

Table 1: States and territories with only one realization of NAD 83 [called NAD 83(FBN)] between NAD 83(1986) and NAD 83(NSRS2007), and the year of that realization.

STATE	FBN
AK	1992
AL	1992
AR	1997
AZ	1992
CO	1992
DC	1991
DE	1991
GA	1994
IA	1996
IL	1997
IN	1997
KS	1997
KY	1993
LA	1992
MD	1991
MI	1994
MN	1996
MO	1997
MS	1993
ND	1996
NE	1995
NM	1992
OH	1995
OK	1993
PA	1992
SD	1996
TX	1993
UT	1994
VA	1993
WV	1995
WY	1993

Table 2: States and territories with two realizations of NAD 83 [called NAD 83(HARN) and NAD 83(FBN)] between NAD 83(1986) and NAD 83(NSRS2007), and the year of the HARN and FBN.

STATE	HARN	FBN
CA	1992	1998
CT	1992	1996
FL	1990	1999
ID	1992	1999
MA	1992	1996
ME	1992	1996
MT	1992	1999
NC	1995	2001
NH	1992	1996
NJ	1992	1996
NV	1994	1999
NY	1992	1996
OR	1991	1998
RI	1992	1996
SC	1995	2001
TN	1990	1995
VQ	1993	2002
VT	1992	1996
WA	1991	1998
WI	1991	1997

Examples

It may be easiest to fully understand how the new GEOCON v2.0 philosophy of "supported realizations" works by way of a few examples. All of the points discussed below have both an NAD 83(1986) and an NAD 83(NSRS2007) set of coordinates, but between these two realizations of NAD 83 they have a variety of histories. Coordinate histories that are indented and in red are ignored in GEOCON v2.0

Example 1: NT0239 – Good point

This is an Idaho point, first participating the original NAD 83(1986) network, then in the original Montana/Idaho HARN (1992) and then participating in the Idaho FBN (1999). Finally, it was included in the national adjustment in 2007. Its history of published coordinates in NAD 83 from 1986 to 2007 looks like this:

NAD 83(1986) -- Original nationwide adjustment of NAD 83
NAD 83(1992) – Montana/Idaho HARN
NAD 83(1999) – Idaho FBN
NAD 83(NSRS2007) -- Nationwide adjustment

In this case, this point *will* be used in creating the GEOCON v2.0 transformation software (both to connect HARN to FBN and to connect FBN to NSRS2007), as it does have both an NAD 83(HARN) coordinate and an NAD 83 (FBN) coordinate for the state in which it resides (Idaho = 1992 and 1999 respectively).

Example 2: JZ2060 – Good point

This is an Indiana point first participating the original NAD 83(1986) network then feathered into Kentucky, then into Ohio/WV before it became part of its own HARN. Finally, it was included in the national adjustment in 2007. Its history of published coordinates in NAD 83 from 1986 to 2007 looks like this:

NAD 83(1986) -- Original nationwide adjustment of NAD 83
    NAD 83(1993) -- Kentucky HARN
    NAD 83(1995) -- Ohio/WV HARN
NAD 83(1997) -- Indiana HARN
NAD 83(NSRS2007) -- Nationwide adjustment

In this case, this point *will* be used to build the GEOCON v2.0 transformation software as it does have a NAD 83(HARN) coordinate for the state in which it resides (Indiana = 1997). *The NAD 83(1993) and NAD 83(1995) coordinates will be* ***ignored****.*[10]

Example 3: JZ3259 – Excluded point

This is an Indiana point first participating the original NAD 83(1986) network then "feathered" into Kentucky's HARN, then into the Ohio/WV HARN but which was not adjusted as part of the Indiana HARN. Finally, it was included in the national adjustment in 2007. Its history of published coordinates in NAD 83 from 1986 to 2007 looks like this:

NAD 83(1986) -- Original nationwide adjustment of NAD 83
    NAD 83(1993) -- Kentucky HARN
    NAD 83(1995) -- Ohio/WV HARN
NAD 83(NSRS2007) -- Nationwide adjustment

In this case, <u>this point *will not* be used to build the GEOCON v2.0 transformation software,</u> as it has *no* NAD 83(HARN) coordinate for the state in which it resides (Indiana = 1997), which therefore prevents the HARN-NSRS2007 connection at this point. One might argue that this point could inform a 1986-NSRS2007 transformation, but it is generally NGS policy that transformations are done step-wise in chronological order between datums or datum realizations, and not "leap over" a datum or realization.[11]

<u>Example 4: BC1730 – Excluded point</u>

This is a Florida point, first participating the original NAD 83(1986) network then then participating in the Florida HARN (1990), but then feathered into the Georgia HARN (1994), but then *not* incorporated into the Florida FBN(1999). Finally, it was included in the national adjustment in 2007. Its history of published coordinates in NAD 83 from 1986 to 2007 looks like this:

NAD 83(1986) -- Original nationwide adjustment of NAD 83
NAD 83(1990) -- Florida HARN
    NAD 83(1994) -- Georgia HARN
NAD 83(NSRS2007) -- Nationwide adjustment

In this case, <u>this point *will not* be used to build the GEOCON v2.0 transformation software,</u> as it has *no* NAD 83(FBN) coordinate for the state in which it resides (Florida = 1999), which therefore prevents both the HARN-FBN connection and the FBN-NSRS2007 connection at this point.

<u>In summary:</u> It is hoped by these examples that the concept of a rigid and unbending rule of adherence to "supported" realizations is hoped to be achieved in GEOCON v2.0.

<u>Analysis of Database</u>

To build GEOCON v2.0 using only supported realizations, it was necessary to fully understand the coordinate history of the NSRS. Prior to attempting to analyze the entire database, multiple pulls of all relevant points only in the state of Delaware in the NGS IDB (meaning points in Delaware with a geodetic latitude and/or geodetic longitude and/or ellipsoid height in any datum) was performed[12] in an attempt to get "the right" data as needed for building GEOCON v2.0. As each pull was performed, it was analyzed[13], and a growing understanding of the loading history of the NGS IDB became apparent. Two issues were frequently encountered:

1) Occasionally a point would have a latitude, longitude or ellipsoid height loaded in an *older* datum realization which would supersede the existing published control on that point which was already in a *newer* datum realization.

a. Example: Point AA9308: On Dec 1, 1995 it had NAD 83(*1993*) coordinates loaded into the IDB. Then 2 ½ years later, on April 7, 1998, an entirely new set of coordinates were loaded for this point, but in the older datum realization of NAD 83(*1991*)!
   i. Question to be solved: **What should be done when coordinates in two datum realizations exist on a point out of chronological order?**

2) Occasionally a point would have its existing latitude, longitude or ellipsoid height superseded but the datum realization would remain the same.
   a. Example: Point AA9308 (again): On April 7, 1998 an ellipsoid height (of -28.046) was loaded for this point in NAD 83(1991). This value stood for 4 ½ years, until on Aug 21, 2002 a new ellipsoid height (of -28.040) was loaded while maintaining the datum realization of NAD 83(1991). This new value then stood for another 4 ½ years until being superseded by NSRS2007.
      i. Question to be solved: **When multiple different values for a coordinate exist on a point in the same datum realization, which one should be used?**

Discussions inside NGS with those most familiar with NGS's history of loading data[14] confirmed that the above observations were not mistakes but conscientious choices made over the years. Each choice seems to have been an individual decision on a case by case basis.[15] In order to build GEOCON v2.0, answers to the above questions (and others) needed answers. The answers to these first two questions were as chosen as:

1) For a point with a load history where a newer datum realization was loaded chronologically earlier than an older datum realization, ignore the loading of the newer datum realization. In the example above, the load of NAD 83(1993) coordinates, occurring before the NAD 83(1991) load, is struck out of consideration.

2) For a point with multiple loads of a coordinate in a single datum realization, keep only the latest load. In the example above, the loading of ellipsoid height -28.046, in NAD 83(1991), was struck out of consideration leaving intact the load of the ellipsoid height of -28.040, also in NAD 83(1991).

One final observation was made in this analysis – the pull of the IDB which was needed had to account for every time any of the three coordinates (geodetic latitude, geodetic longitude or ellipsoid height) were loaded into the database. In this way, a history of what the datasheet would have looked like for a point after every load was constructible.

On December 20, 2013 the final pull of the IDB was made which would contain all the data needed to build GEOCON v2.0[16]. Three files were pulled, due to the nature of how data are stored in the NGS IDB: one for published control, one for FAA control and one for CORS data[17]. The CORS data did not contain the sort of information used to build the transformations and so was dropped. The other two files contained records (one record per line) of every time a latitude and/or longitude and/or ellipsoid height were loaded for a point. After some analysis[18], it was determined that, while the FAA data contained some actual changes to coordinate data sets on points in the IDB, that these FAA-based values were never published[19]. This means that FAA changes to coordinates were never accessible to external users (never appearing on datasheets) and therefore any transformation program should ignore them. A decision was therefore made[20] to build transformation software only on the pull of data from the published control data set.

Even working with just the one database pull, a number of oddities were discovered, such as ellipsoid heights which had been loaded with no datum realization ("POS_TAG" = "N/A") or loading ellipsoid heights months or even years before loading latitudes and longitudes or loading them as separate loads but on the same day. A complete addressing of these oddities is discussed later. Additionally, hand-held and scaled coordinates are a huge part of the database (primarily on leveled benchmarks), and they needed to be dropped from consideration in making a transformation tool. Overall, the final database pull containing both oddities and undesired data required evaluation and the creation of a set of rules to keep or toss data based on each situation.

Ultimately, the chain of files which led to the raw working file (`pull4_us`) is as follows:

Table 3: Chain of files from initial database pull to final usable data file.

File Name	Created / How	Notes
`dru_us.csv`[21]	Sept 21, 2013[22] from an IDB pull for every "load" record of lat, lon or h, moved from DOS to UNIX after receipt.	7,202,010 Lines 3,601,003 Records of a "load" In alphabetical PID order 518,544,720 bytes
`dru_us`	Sept 23, 2013 using the "dos2unix" command on the above file	7,202,010 Lines 3,601,003 Records of a "load" In alphabetical PID order 511,342,710 bytes
`dru_us_b`	Dec 6, 2013 using program "conv_us_1b.f" on the above file to turn 2-row load records into 1-row load records	3,601,005 Lines 3,601,003 Records of a "load" In alphabetical PID order 529,347,735 bytes
`pull4_us`	Dec 20, 2013 using program "conv_us_2.f" on the above file to change it into a more useful format, strip off the headers drop all scaled or other unadjusted positions, and fix a few minor typographical glitches. Note the massive drop of records (due to so many points being "scaled" in the IDB).	1,265,217 Lines 1,265,217 Records of a "load" In alphabetical PID order 139,173,870 bytes

The top ten lines from each of these four files is shown below.

```
,PID   ,STATE,COUNTY,LATITUDE     ,LONGITUDE     ,POS_SOURCE,DATUM,ELLIP_HT  ,ELL_SOURCE,POS_TAG,EPOCH1  ,ADJ_ID       ,ADJ_DATE
         LOAD_DATE,
,------,-----,------,-------------,-------------,----------,-----,----------,----------,-------,--------,-------------,--------
         ---------,
,AA0001,FL   ,087   ,N243349      ,W0814752      ,S         ,27   ,N/A       ,N/A       ,N/A    ,N/A     ,P            ,NULL
         19940924 ,
,AA0001,FL   ,087   ,N243351      ,W0814751      ,S         ,83   ,N/A       ,N/A       ,1986   ,N/A     ,Q            ,NULL
         20030107 ,
,AA0002,FL   ,087   ,N243338      ,W0814814      ,S         ,27   ,N/A       ,N/A       ,N/A    ,N/A     ,P            ,NULL
```

```
         19940924 ,

,PID   ,STATE,COUNTY,LATITUDE     ,LONGITUDE     ,POS_SOURCE,DATUM,ELLIP_HT  ,ELL_SOURCE,POS_TAG,EPOCH1  ,ADJ_ID       ,ADJ_DATE
         LOAD_DATE,
,------,-----,------,------------,-------------,----------,-----,----------,----------,-------,--------,------------,--------
         ---------,
,AA0001,FL   ,087   ,N243349     ,W0814752     ,S         ,27   ,N/A       ,N/A       ,N/A    ,N/A     ,P            ,NULL
         19940924 ,
,AA0001,FL   ,087   ,N243351     ,W0814751     ,S         ,83   ,N/A       ,N/A       ,1986   ,N/A     ,Q            ,NULL
         20030107 ,
,AA0002,FL   ,087   ,N243338     ,W0814814     ,S         ,27   ,N/A       ,N/A       ,N/A    ,N/A     ,P            ,NULL
         19940924 ,
```

```
,PID   ,STATE,COUNTY,LATITUDE     ,LONGITUDE     ,POS_SOURCE,DATUM,ELLIP_HT  ,ELL_SOURCE,POS_TAG,EPOCH1  ,ADJ_ID       ,ADJ_DATE LOAD_DATE,
,------,-----,------,------------,-------------,----------,-----,----------,----------,-------,--------,------------,-------- ---------,
,AA0001,FL   ,087   ,N243349     ,W0814752     ,S         ,27   ,N/A       ,N/A       ,N/A    ,N/A     ,P            ,NULL     19940924 ,
,AA0001,FL   ,087   ,N243351     ,W0814751     ,S         ,83   ,N/A       ,N/A       ,1986   ,N/A     ,Q            ,NULL     20030107 ,
,AA0002,FL   ,087   ,N243338     ,W0814814     ,S         ,27   ,N/A       ,N/A       ,N/A    ,N/A     ,P            ,NULL     19940924 ,
,AA0002,FL   ,087   ,N243340     ,W0814813     ,S         ,83   ,N/A       ,N/A       ,1986   ,N/A     ,Q            ,NULL     20030107 ,
,AA0003,FL   ,087   ,N243338.16  ,W0814819.24  ,O         ,83   ,N/A       ,N/A       ,1986   ,N/A     ,GPS1909/120  ,NULL     20110919 ,
,AA0004,FL   ,087   ,N243336.0   ,W0814822.3   ,O         ,83   ,N/A       ,N/A       ,1986   ,N/A     ,GPS1909/120  ,NULL     20110919 ,
,AA0005,FL   ,087   ,N243330.2   ,W0814824.5   ,O         ,83   ,N/A       ,N/A       ,1986   ,N/A     ,GPS1909/120  ,NULL     20110919 ,
,AA0006,FL   ,087   ,N243323     ,W0814824     ,S         ,27   ,N/A       ,N/A       ,N/A    ,N/A     ,P            ,NULL     19940924 ,
```

```
AA0015 FL 087 N243244.15887 W0814813.61148   X 27        N/A N/A  N/A      N/A 14831         1979     19870101
AA0015 FL 087 N243245.68706 W0814812.94316   X 83        N/A N/A 1986      N/A 17289         19860719 19870101
AA0015 FL 087 N243245.68047 W0814812.95293   A 83        N/A N/A 1990      N/A 17478         19910501 19910529
AA0028 FL 087 N243405.63501 W0814607.17515   X 27        N/A N/A  N/A      N/A 14831         1979     19870101
AA0028 FL 087 N243407.16083 W0814606.50343   X 83        N/A N/A 1986      N/A 17289         19860719 19870101
AA0028 FL 087 N243407.15445 W0814606.51366   X 83    -20.649    X 1990      N/A GPS0170       19900913 19910430
AA0028 FL 087 N243407.15627 W0814606.51362   X 83    -20.849    X 1999      N/A GPS1378/2     20010531 20010605
AA0028 FL 087 N243407.15592 W0814606.51353   X 83    -20.792    X 2007 2002     GPS2300       20070210 20070210
AA0028 FL 087 N243407.15572 W0814606.51185   A 83    -20.806    A 2011 20100101 GPS2800       20120627 20120627
AA0033 FL 087 N243323.27881 W0814600.39080   X 83    -21.165    X 1990      N/A GPS0750       19941121 19950209
```

Once a final data file (**`pull4_us`**) was created, additional analysis and working files were created. Specifically, a rules-based program was built[23] which analyzed the load records for every PID in file **`pull4_us`**, created a working file of coordinate sets in supported realizations that could be used in any transformation software (including rebuilds of NADCON, for example) and logged what changes and deletions were necessary to arrive at this subset of load records.

The rules set by this program were:

1) No stand-alone loads were kept
    a. Rationale: Since two different realizations would need to have been loaded for the data to inform a transformation program.
2) When an ellipsoid height was loaded by itself, and the POSTAG field (which stores the "datum realization" information) was set to N/A, if there were earlier loads of latitude and longitude on that point with a valid POSTAG value, then bring forward those latitude and longitude values and that POSTAG value and associate them with the ellipsoid height load. This yielded a latitude, longitude and ellipsoid height in one datum realization for a point.
    a. Rationale: All loads should be in a datum realization. It was NGS policy to skip the POSTAG value when loading *just* an ellipsoid height, even though the implication was that the ellipsoid height was being "adjoined" to the already existing latitude and longitude loaded earlier.
3) When an ellipsoid height was loaded alone, with a POSTAG value set to N/A and there were no loads of latitude or longitude prior to this that ellipsoid height was deleted.
    a. Rationale: With no way to know what datum realization the ellipsoid height was in, the point is effectively useless.
4) When a load occurs with a POSTAG chronologically out of order, delete the early load of a newer POSTAG.
    a. Rationale: See earlier.

5) When a load occurs with a POSTAG equal to an existing loaded POSTAG, keep only the most recent load record.
    a. Rationale: See earlier.
6) When a load occurs for an unsupported datum realization, drop that record.
    a. Rationale: See earlier.
7) When a load occurs for a nonsensical datum (such as NAD 27 in Hawaii), drop that record.
    a. Rationale: Certain datums simply do not exist in certain areas, and anything in the IDB which says differently is an error needing correction
8) When a load occurs with an unsupported state code drop that record.
    a. Rationale: The IDB contains data in regions of the world besides the USA, or in places with defunct state codes (such as the "Trust Territories of the Pacific"). Such data were not used in this transformation tool.
9) When a load occurs but is unpaired with either the previous supported realization or next supported realization drop that load record.
    a. Rationale: A transformation tool, by its nature, is built on records of paired realizations. The policy of NGS is that transformation tools generally be built between chronologically adjacent datums or datum realizations.

The output files were **`pull4.log`**, **`pull4.out`** and **`pull4.io`**. Each of these files is described below:

`pull4.log`
Is a log file of every change made[24] to **`pull4_us`** in order to create **`pull4.out`**.

`pull4.io`
Is a report file, on a PID-by-PID basis of every load record that was found in **`pull4_us`** and every output record that was placed into **`pull4.out`**.

`pull4.out`
Is the working file containing all data from the IDB which can, and should, inform every transformation program between any supported datum realizations. This file is a critical part of building GEOCON v2.0 and should be used if or when NADCON is rebuilt and merged with GEOCON or at least to validate the current state of NADCON prior to such a merge.

Examples of records found in these three files are found on the next pages.

Explaining **`pull4.log`**

In yellow below are examples of the lines found in file **`pull4.log`**. An explanation of each line follows it.

An examination of **`pull4_us`** reveals that PID "AA1702" has just one load entry. As such, being unpaired, it is deleted from consideration.

An examination of **`pull4_us`** reveals that PID AA2323 has the following four load records:

```
AA2323 AL 011 N320459.72243 W0854246.16808   X 83   110.923   X 1992      N/A GPS0746       19941130 19950313
AA2323 AL 011           N/A             N/A N/A 83   110.868   X  N/A      N/A GPS1250/4C    20020729 20020809
AA2323 AL 011 N320459.72278 W0854246.16747   X 83   110.883   X 2007 2002     GPS2300       20070210 20070210
AA2323 AL 011 N320459.72280 W0854246.16693   A 83   110.872   A 2011 20100101 GPS2800       20120627 20120627
```

The load of ellipsoid height "110.868" with a POSTAG of N/A occurred on 8/9/2002, but a good load of a latitude and longitude with POSTAG=1992 had occurred 7 ½ years earlier on 3/13/1995. As such, the latitude, longitude and POSTAG are brought forward, so that the new set of load records looks like this (changes in green):

```
AA2323 AL 011 N320459.72243 W0854246.16808   X 83   110.923   X 1992      N/A GPS0746       19941130 19950313
AA2323 AL 011                                  83   110.868   X           N/A GPS1250/4C    20020729 20020809
AA2323 AL 011 N320459.72278 W0854246.16747   X 83   110.883   X 2007 2002     GPS2300       20070210 20070210
AA2323 AL 011 N320459.72280 W0854246.16693   A 83   110.872   A 2011 20100101 GPS2800       20120627 20120627
```

But now there are two "1992" POSTAG records for this point, so the next entry in **`pull4.log`** can, and did, occur.

As mentioned above, this entry goes hand in hand with the previous. By Rule #5, we keep only the latest entry when there are two POSTAG values.

```
Killing unfixable POSTAG=N/A point: AW5607 TX 201           N/A             N/A N/A 83    12.294   X  N/A       N/A 17564         19930121 19930217
```

An examination of **pull4_us** reveals that PID AW5607 has the following three load records:

```
AW5607 TX 201           N/A             N/A N/A 83    12.294   X  N/A       N/A 17564         19930121 19930217
AW5607 TX 201 N294645.89146 W0952558.74186   X 83    13.364   X 2007 2002      GPS2300       20070210 20070210
AW5607 TX 201 N294645.89203 W0952558.74040   A 83    13.313   A 2011 20100101 GPS2800       20120627 20120627
```

The load of an ellipsoid height with POSTAG=N/A occurred on 2/17/1993. Without any prior loads with any POSTAG values, it is impossible to know what realization this load is to be associated with, and therefore this load record is deleted.

```
Killing NAD83 entry with chron. too early POSTAG: AA5438 PA 027 N405116.75794 W0775015.57083   X 83   330.630   X 1993      N/A GPS0743/B     19951130 19951201
```

In **pull4_us**, PID AA5438 has these seven load records:

```
AA5438 PA 027 N405116.75813 W0775015.57076   X 83   330.465   X 1992      N/A GPS0748       19950523 19950626
AA5438 PA 027 N405116.75794 W0775015.57083   X 83   330.630   X 1993      N/A GPS0743/B     19951130 19951201
AA5438 PA 027 N405116.75854 W0775015.57093   X 83   330.575   X 1992      N/A GPS1060/A     19980324 19980325
AA5438 PA 027           N/A             N/A N/A 83   330.480   X  N/A      N/A GPS1060/16B   20010917 20010925
AA5438 PA 027           N/A             N/A N/A 83   330.456   X  N/A      N/A GPS1462/2     20011004 20011005
AA5438 PA 027 N405116.75868 W0775015.57097   X 83   330.479   X 2007 2002      GPS2300       20070210 20070210
AA5438 PA 027 N405116.75838 W0775015.57021   A 83   330.459   A 2011 20100101 GPS2800       20120627 20120627
```

The 2nd load record, with a POSTAG=1993 occurrs before the 3rd load record, with a POSTAG=1992. As such, by Rule #4, the 2nd load record is deleted.

```
Unsupported realization.  Killing: AA1850 LA 019 N300745.37418 W0934148.52875 N/A 83   -24.689   X 1993      N/A GPS1264/5     20020220 20020221
```

In **pull4_us** there are five load records for PID AA1850:

```
AA1850 LA 019 N300745.37418 W0934148.52875   X 83   -24.600   X 1992      N/A GPS0766       19950113 19950203
AA1850 LA 019 N300745.37418 W0934148.52875   X 83       N/A N/A 1993      N/A 17596         19960216 19960228
AA1850 LA 019           N/A             N/A N/A 83   -24.689   X  N/A      N/A GPS1264/5     20020220 20020221
AA1850 LA 019 N300745.37596 W0934148.52941   X 83   -24.720   X 2007 2002      GPS2300       20070210 20070210
AA1850 LA 019 N300745.37609 W0934148.52785   A 83   -24.769   A 2011 20100101 GPS2800       20120627 20120627
```

The 2nd record, with POSTAG=1993 (and by extension, using the same rules mentioned, load record #3 is also in the 1993 realization and is the one being deleted) is in an unsupported realization (for the state of LA, only NAD 83(1992) is supported between 1986 and 2007.) As such, this load record is deleted.

```
Unpaired realization.  Killing: AA1859 CA 085 N370542.04343 W1213416.11887   X 83    85.118   X 1992 19910507 GPS0796       19940930 19950420  6 4
```

The load records for PID AA1859 in **pull4_us** are:

```
AA1859 CA 085 N370542.04343 W1213416.11887   X 83    85.118   X 1992 19910507 GPS0796       19940930 19950420
AA1859 CA 085 N370542.05121 W1213416.12933   X 83    85.004   X 2007 20070101 GPS2300       20070210 20070210
AA1859 CA 085 N370542.05523 W1213416.13149   A 83    85.017   A 2011 20100101 GPS2800       20120627 20120627
```

But since the supported realizations for NAD 83 in California are 1992, 1998, 2007 then 2011, without a 1998 set of coordinates on this point, the 1992 set of coordinates is unpaired and cannot inform the 1992/1998 transformation portion of GEOCON v2.0. The POSTAG=1992 record is therefore dropped.

```
Nonsensical Datum for this Point.  Killing: TV0039 VQ 001 N174528.90182 W0643402.35625   X 27       N/A N/A  N/A      N/A 16377         1979     19870101
```

There is no evidence that NAD 27 was ever expanded to the US Virgin Islands. This NAD 27 set of coordinates on a point in the USVI is deleted.

```
Bad state.  Killing entry: AA4434 TQ 000 N070358.21280 W1884316.71491   X 83    30.942   X 1993 19930814 GPS0667/D     19941130 19950607
```

The state code “TQ” is defunct, and reflects the “Trust Territories of the Pacific” region. As such, this load record is deleted.

Explaining pull4.io

Every PID in **pull4_us** has a set of lines in **pull4.io** explaining what the PID’s load records looked like in **pull4_us** and what they look like in **pull4.out**. Here is an example for AA2323:

```
Prior to editing, this PID has these entries:
AA2323 AL 011 N320459.72243 W0854246.16808   X 83   110.923   X 1992      N/A GPS0746       19941130 19950313
AA2323 AL 011           N/A            N/A N/A 83   110.868   X  N/A      N/A GPS1250/4C    20020729 20020809
AA2323 AL 011 N320459.72278 W0854246.16747   X 83   110.883   X 2007 2002     GPS2300       20070210 20070210
AA2323 AL 011 N320459.72280 W0854246.16693   A 83   110.872   A 2011 20100101 GPS2800       20120627 20120627
After    editing, this PID has these entries:
AA2323 AL 011 N320459.72243 W0854246.16808 N/A 83   110.868   X 1992      N/A GPS1250/4C    20020729 20020809
AA2323 AL 011 N320459.72278 W0854246.16747   X 83   110.883   X 2007 2002     GPS2300       20070210 20070210
AA2323 AL 011 N320459.72280 W0854246.16693   A 83   110.872   A 2011 20100101 GPS2800       20120627 20120627
```

See that the four load records are reduced now to three records, each reflecting the coordinates to be used in each supported realization in creating any transformation tools, including GEOCON v2.0. In this example, a point in Alabama ends up with an adjusted latitude, longitude and ellipsoid height in NAD 83(1992), NAD 83(NSRS20007) and NAD 83(2011), so it can inform the 1992/NSRS2007 transformation and the NSRS2007/2011 transformation in GEOCON v2.0.

Explaining `pull4.out`

The final output coordinates in supported realizations to be used in all transformation software is found in **`pull4.out`**. Let us pick one line at random to demonstrate the format of the file:

Table 4: Column-by-column format of each record in `pull4.out`

Columns	Value	Possible Values	IDB Table	Description
1-6	EC1408	Alphanumerics	PID	PID of the point
7	(space)	(space)		- Purposefully blank -
8-9	NC	Any state/territory code	STATE	State or Territory code in which the point falls
10	(space)	(space)		- Purposefully blank -
11-13	179	001 through 999	COUNTY	Numeric code for which county of the state the point falls
14	(space)	(space)		- Purposefully blank -
15	N	N or S	LATITUDE	Hemisphere identifier for latitude
16-17	34	00 through 90	LATITUDE	Integer degrees of latitude of point
18-19	54	00 through 59	LATITUDE	Integer minutes of latitude of point
20-27	21.46734	00.00000 through 59.99999	LATITUDE	Decimal seconds of latitude of point
28	(space)	(space)		- Purposefully blank -
29	W	W or E	LONGITUDE	Hemisphere identifier for longitude
30-32	080	000 through 180	LONGITUDE	Integer degrees of longitude of point
33-34	19	00 through 59	LONGITUDE	Integer minutes of longitude of point
35-42	39.77514	00.00000 through 59.99999 or N/A	LONGITUDE	Decimal seconds of longitude of point
43	(space)	(space)		- Purposefully blank -
44-46	A	A or X or N/A	POS_SOURCE	Code describing whether the point's lat/lon came from an adjustment
47	(space)	(space)		- Purposefully blank -
48-49	83	27, 83, US, GU, PR, AS, HI	DATUM	Datum code
50	(space)	(space)		- Purposefully blank -
51-59	119.206	Varied numbers or N/A	ELLIP_HT	Ellipsoid height in meters
60	(space)	(space)		- Purposefully blank -

61-63	A	A or X or N/A	ELL_SOURCE	Code describing whether the point's ellip. ht. came from an adjustment
64	(space)	(space)		- Purposefully blank -
65-68	2011	A year or N/A	POS_TAG	The datum realization identifier[25]
69	(space)	(space)		- Purposefully blank -
70-77	20100101	A date as YYYYMMDD or just YYYY or N/A	EPOCH1	The datum realization's epoch.
78	(space)	(space)		- Purposefully blank -
79-91	GPS2800	A number or GPS#### or DOPPLER or N/A	ADJ_ID	The adjustment identifier for how the coordinate was adjusted.
92	(space)	(space)		- Purposefully blank -
93-100	20120627	Any date as YYYYMMDD or just YYYY or NULL	ADJ_DATE	The date when the adjustment was performed.
101	(space)	(space)		- Purposefully blank -
102-109	20120627	Any date as YYYYMMDD or just YYYY or NULL	LOAD_DATE	The date the data was loaded into the IDB (does not usually equal the adjustment date).
110	(space)	(space)		- Purposefully blank -
111-112	31	01 through 56	-not in IDB-	State Number. This is a number invented solely for use of this data file in creating transformation software.[26]
113	(space)	(space)		- Purposefully blank -
114-114	7	1 through 7	-not in IDB-	Realization Number. This is a number invented solely for use of this data file in creating transformation software.

With that information at hand, let's look at the entire suite of entries in **`pull4.out`** for point AA2323:

```
AA2323 AL 011 N320459.72243 W0854246.16808 N/A 83   110.868   X 1992      N/A GPS1250/4C    20020729 20020809  2 4
AA2323 AL 011 N320459.72278 W0854246.16747   X 83   110.883   X 2007 2002     GPS2300       20070210 20070210  2 5
AA2323 AL 011 N320459.72280 W0854246.16693   A 83   110.872   A 2011 20100101 GPS2800       20120627 20120627  2 6
```

In this case, the point has data to inform the transformations between supported realizations 1992/NSRS2007 (realizations 4 and 5 in this state) and NSRS2007/2011 (realizations 5 and 6 in this state) in the state of Alabama. However, not all points have a full set of latitudes, longitudes and ellipsoid heights to support all transformations. Consider point AA1012:

```
AA1012 FL 087 N243936.55501 W0812521.94293   X 83       N/A N/A 1990      N/A GPS0174       19910204 19910514 12 4
AA1012 FL 087 N243936.55736 W0812521.94188   X 83   -16.354   X 1999      N/A GPS1378/5     20011213 20011213 12 5
AA1012 FL 087 N243936.55674 W0812521.94261   X 83   -16.373   X 2007 2002     GPS2300       20070210 20070210 12 6
AA1012 FL 087 N243936.55675 W0812521.94133   A 83   -16.384   A 2011 20100101 GPS2800       20120627 20120627 12 7
```

The four supported realizations 1990, 1999, NSRS2007 and 2011 are all represented at this point, but no ellipsoid height is present in the 1990 realization. As such only the latitude and longitude will inform the 1990/1999 transformation. For the 1999/NSRS2007 and NSRS2007/2011 transformations, however, this point will inform transformations in all three coordinates.

Creating final transformation-by-transformation input files

***As a reminder:* "`pull4.out`" represents all data in the NGS IDB which should inform all transformations between all supported realizations in all states and territories**

The file **`pull4.out`** is a large and somewhat unwieldy file with which to work. Therefore, as a way to break down the information into usable and digestible sizes, individual files were created from **`pull4.out`** on a transformation-by-transformation case (that is, one file for every one state/territory or region, and two chronologically adjoined supported realizations). A new program[27] was written to parse **`pull4.out`** into these sub-files. A log of this parsing (showing how many pairs of latitudes, longitudes and ellipsoid heights occurred for each transformation) was kept in a separate file. The final output files had the general naming convention of:

NADCON5.*AREA*.*OLDDATUM*.*NEWDATUM*.in

Where

AREA is frequently a 2 digit state or territory code, but sometimes was regional or multi-regional
OLDDATUM was the chronologically older supported datum (or datum realization) name
NEWDTUM was the chronologically newer supported datum (or datum realization) name

The prefix "**NADCON5**" was chosen purposefully, under the presumption that GEOCON v2.0 would eventually be merged with NADCON (which currently is in version 4.2), making some future "NADCON v5.0" product.

For example, the file **NADCON5.CA.NAD83_1998.NAD83_2007.in** contains all of the data in the state of California capable of building transformation models between NAD 83(1998) and NAD 83(NSRS2007) in that state. The first few lines of this file look like the following:

```
                            NAD 83(1998:CA)             |                NAD 83(NSRS2007)
AA1871 CA 085 N371952.98270 W1220459.33068   100.202 | N371952.98670 W1220459.33410   100.138
AA1872 CA 085 N372114.03381 W1215425.18785   -14.080 | N372114.03685 W1215425.19086   -14.136
AA1873 CA 085 N372320.17305 W1215911.89487   -24.660 | N372320.17608 W1215911.89823   -24.720
AA2147 CA 001 N373234.73786 W1220054.41178   -19.286 | N373234.74098 W1220054.41511   -19.349
```

This shows that each such file will have one header line, expressing the name of the Old Datum on the left and the New Datum on the right. Note that, in this case, the realization is called "NAD 83(1998:CA)" to explicitly differentiate it from other NAD 83(1998) realizations in other states. Each line represents one point, showing the PID, state or territory code, county code and then the latitude, longitude and ellipsoid height, first in the old datum then in the new datum. Any missing values (usually an ellipsoid height) will be given the value "N/A". As long as there are at least one good pair (2 latitudes or 2 longitudes or 2 ellipsoid heights), the point will be in this file.

Some of the files are regional, and this is only the case when a regionally-valid realization (like NAD 83(NSRS2007) is paired with another regionally-valid realization (like NAD 83(2011)). For instance, this file **NADCON5.CONUS_PR_VQ.NAD83_2007.NAD83_2011.in** contains all the points in all of CONUS, Puerto Rico and the US Virgin Islands which can be used to transform between NAD 83(NSRS2007) and NAD 83(2011.)[28] The first and last few lines of this file show that the data are stored in alphabetical order by state/region, and then within that state/region by alphabetical order of PID.

```
                           NAD 83(NSRS2007)             |                  NAD 83(2011)
AA1903 AL 009 N335033.39073 W0865526.29797   210.953 | N335033.39088 W0865526.29723   210.938
AA1904 AL 009 N335025.27146 W0865509.08231   215.666 | N335025.27161 W0865509.08157   215.651
AA1905 AL 009 N334918.55191 W0864310.13104   145.024 | N334918.55200 W0864310.13029   145.008
AA1906 AL 009 N334917.67125 W0864252.06421   145.628 | N334917.67134 W0864252.06345   145.611
.
```

```
.
.
PY1281 WY 029 N444005.13918 W1101535.30987  2509.161 | N444005.13940 W1101535.30861  2509.124
PY1282 WY 039 N442447.59438 W1105721.12230  2609.224 | N442447.59461 W1105721.12102  2609.189
PY1284 WY 039 N440745.59760 W1104911.91675  2191.822 | N440745.59803 W1104911.91604  2191.783
PY1285 WY 039 N440754.05307 W1104901.33058  2192.280 | N440754.05350 W1104901.32987  2192.240
```

There were a total of 145 such files created[29]. Some informed only very old datums (for possible use in the future), but otherwise each was used as the primary working file when building one transformation for one state, territory or region in GEOCON v2.0.

Inventory of existing code and data used to create v1.0

Once working files were created, a careful inventory of the source code used in building GEOCON v1.0 (and GEOCON11 v1.0) was created. This was because the overall flow of *regional* grid production (ala "v1.0") was seen as essentially the same as the creation of the *state-by-state* transformation grids[30] which would be used to transform between the HARN and the FBN realizations in those states or territories that had indicated such a transformation was desired. This series of programs, beginning with a database pull and ending with grids and pictures was carefully documented[31].

[Note: Work on GEOCON v2.0 was temporarily paused at this point so that GEOCON v1.1 and GEOCON11 v1.1 could be created[32]. GEOCON v1.1 was turned over to the NGS Systems Development Division in March of 2014. It was finally released to the Beta NGS web page on September 30, 2014.]

Prior to creating any transformation grids for GEOCON v2.0, it was critical to make sure that the new data was either identical to, or else explainably different (yet better) than the data used in the creation of the grids used in GEOCON v1.0. A quick test was made between the v1.0 input data[33] and the v2.0 input data.[34] A comprehensive analysis of every difference seen in every point in the District of Columbia was performed first.[35] Every difference was explainable, and the v2.0 data clearly was complete and proper ("correct territory", "only in the supported datum realizations"). The amount of manual effort to perform this check made its expansion to the entire dataset too large to consider. As such, it was decided to proceed under the presumption that the v2.0 data was good.

Creating the HARN-to-FBN transformations, aka "BRIDGE"

The first major re-build of GEOCON was to create the state-by-state (or territory-by-territory) transformations which "bridged" the HARN and FBN realizations. This included a 2-step transformation for Puerto Rico, bridging HARN0 to HARN and HARN to FBN. This effort was known during its construction as "building the BRIDGE".

All files created to this point contained coordinate information. A work file was needed which contained a variety of pre-computed values, such as coordinate differences (in various units) and azimuths of a combined latitude/longitude shift. To begin with, a list of the "**`NADCON***.in`**" files which would be used was created. A simple list file (called "**`BRIDGE.filelist`**") was manually created which listed all of the input files to become part of the "work file".[36] Its contents were:

```
NADCON5.AS.NAD83_1993.NAD83_2002.in
NADCON5.CA.NAD83_1992.NAD83_1998.in
NADCON5.CQ.NAD83_1993.NAD83_2002.in
NADCON5.CT.NAD83_1992.NAD83_1996.in
NADCON5.FL.NAD83_1990.NAD83_1999.in
NADCON5.ID.NAD83_1992.NAD83_1999.in
NADCON5.MA.NAD83_1992.NAD83_1996.in
NADCON5.ME.NAD83_1992.NAD83_1996.in
NADCON5.MT.NAD83_1992.NAD83_1999.in
NADCON5.NC.NAD83_1995.NAD83_2001.in
NADCON5.NH.NAD83_1992.NAD83_1996.in
NADCON5.NJ.NAD83_1992.NAD83_1996.in
NADCON5.NV.NAD83_1994.NAD83_1999.in
NADCON5.NY.NAD83_1992.NAD83_1996.in
NADCON5.OR.NAD83_1991.NAD83_1998.in
NADCON5.PR.NAD83_1993.NAD83_1997.in
NADCON5.PR.NAD83_1997.NAD83_2002.in
NADCON5.RI.NAD83_1992.NAD83_1996.in
NADCON5.SC.NAD83_1995.NAD83_2001.in
NADCON5.TN.NAD83_1990.NAD83_1995.in
NADCON5.VQ.NAD83_1993.NAD83_2002.in
NADCON5.VT.NAD83_1992.NAD83_1996.in
NADCON5.WA.NAD83_1991.NAD83_1998.in
NADCON5.WI.NAD83_1991.NAD83_1997.in
```

For each of these 24 files, a separate set of transformation grids would need to be created for GEOCON v2.0.[37] A program took this file list, and created a "work" file[38], containing all of the pre-computed information necessary to grid the transformations.

As a first step, coverage plots were created, to ensure that data fell in the state or territory which the IDB claimed[39]. As a next step, to control edge effects, a "mask" file was needed to buffer around each state. This first required that a reasonably good boundary file for each state or territory be created. Some pre-existing state boundary polygons were found online, and some needed to be hand-digitized. All were converted to a usable format[40]. A 1x1 degree buffer was put around each boundary, defining a region in which the transformations would work, and outside of which the transformations would be forced to fall off. The next step was to create, from the "work" file a set of plots showing the vector changes in latitude, longitude, ellipsoid height and a combination of latitude/longitude (aligned along the proper azimuth). This helped identify outliers and other troublesome points. The residual values (delta latitude, delta longitude, delta ellipsoid height, delta horizontal) were then run through a block-median filter at a 1 x 1 arcminute basis. Points that were kept in this filter are called "**kept points**". Points dropped by the filter are called "**dropped points**"[41]. The first three grids actually became the final transformation grids for this portion of GEOCON v2.0, while the fourth was simply created for the purposes of plotting and visualization. All four grids were plotted using the GMT tool into colorized representations.

Cross-Validation

The use of cross-validation, as first presented in GEOCON v1.0, was continued in GEOCON v2.0. In brief, it can be summed as follows: Randomly shuffle all of the points in a region (both **kept points** and **dropped points)** and then, systematically, pull a small number of points ("**nlucky**") out of that group, create a new experimental transformation grid using this thinned data set, and then compare said grid to the extracted points (through interpolation off of said grid). Record the difference in value. Do this repeatedly (for "**nloops**") until every point has been checked against a grid created without itself. This process is the essence of cross-validation.

For each state or territory, the number of kept points, dropped points and the number of loops and the number of lucky points pulled out in each loop is listed in Table 5. Note that on the last loop, if the number of points left to be cross-validated is less than "nlucky" then only the remaining points are pulled and cross validated. The number of loops (nloops) was purposefully forced to be 50 or less, to make the computational burden of performing so many cross-validations manageable.

Once all cross-validation differences were collected, one final step was taken – a "worst case" filter (called an "anti-median" in the GEOCON v1.0 documentation) was run on all of these values. These worst case values were then gridded to create the so-called "quality" grids. The intent, as mentioned in GEOCON v1.0, was to provide highly conservative estimates of error to the end user who may or may not have tied into egregiously bad outliers. That is, the **transformation grid** shows the median transformation values, but the **quality grid** shows the worst possible transformation error.

To clarify by using an example: In the state of Wisconsin, the number of points with a published coordinate in both NAD 83(1991) and NAD 83(1997) is 816 (latitude), 816 (longitude) and 747 (ellipsoid height). When these data were run through the block-median filter,the number of **kept points** was 752 (latitude), 752 (longitude) and 700 (ellipsoid height), so the number of **dropped points** was 64 (latitude), 64 (longitude) and 47 (ellipsoid height). Note that the median filter works on each type of coordinate separately, so there is nothing forcing the 816 kept points in latitude to be identical to the 816 kept points in longitude. The **kept points** were then gridded into the 3 actual **transformation grids** for Wisconsin between NAD 83(1991) and NAD 83(1997).

Once the transformation grids were created, the cross-validation went back to the 816 latitude points, 816 longitude points and the 747 ellipsoid height points and randomly pulled out small samples (17 latitude or longitude points at a time; 15 ellipsoid height points at a time), re-ran the median filter, gridded the data into a temporary grid and then compared the temporary grid to the withheld points. This was done 48 times in latitude and longitude, and 50 times in ellipsoid height. Once completed, the cross-validation values were collected, run through the ***anti-median*** filter (which keeps and drops the same *number* of points as a median filter, only this time keeping the *worst* point, not the *median* point) and gridded to make Wisconsin's three **quality grids**.

Table 5:Gridding point counts and cross-validation loop information for states and territories that have a HARN/FBN transformation.

State/Territory Old Realization New Realization	Latitude Counts	Longitude Counts	Ellipsoid Height Counts	Number of Cross-Validation Loops ("nloops")	Number of points pulled out in each loop ("nlucky")
AS 1993 2002	Total: 193 Kept: 73 Dropped: 120	Total: 193 Kept: 73 Dropped: 120	Total: 38 Kept: 15 Dropped: 23	Lat: 49 Lon: 49 Eht: 38	Lat: 4 Lon: 4 Eht: 1
CA 1992 1998	Total: 544 Kept: 539 Dropped: 5	Total: 544 Kept: 539 Dropped: 5	Total: 479 Kept: 476 Dropped: 3	Lat: 50 Lon: 50 Eht: 48	Lat: 11 Lon: 11 Eht: 10
CQ 1993 2002	Total: 6 Kept: 6 Dropped: 0	Total: 6 Kept: 6 Dropped: 0	Total: 6 Kept: 6 Dropped: 0	Lat: 6 Lon: 6 Eht: 6	Lat: 1 Lon: 1 Eht: 1
CT 1992 1996	Total: 3603 Kept: 1643 Dropped: 1960	Total: 3603 Kept: 1643 Dropped: 1960	Total: 9 Kept: 8 Dropped: 1	Lat: 50 Lon: 50 Eht: 9	Lat: 73 Lon: 73 Eht: 1
FL 1990 1999	Total: 4480 Kept: 3489 Dropped: 991	Total: 4480 Kept: 3489 Dropped: 991	Total: 4019 Kept: 3167 Dropped: 852	Lat: 50 Lon: 50 Eht: 50	Lat: 90 Lon: 90 Eht: 81
ID 1992 1999	Total: 181 Kept: 172 Dropped: 9	Total: 181 Kept: 172 Dropped: 9	Total: 115 Kept: 111 Dropped: 4	Lat: 46 Lon: 46 Eht: 39	Lat: 4 Lon: 4 Eht: 3
MA 1992 1996	Total: 6982 Kept: 2604 Dropped: 4378	Total: 6982 Kept: 2604 Dropped: 4378	Total: 75 Kept: 64 Dropped: 11	Lat: 50 Lon: 50 Eht: 38	Lat: 140 Lon: 140 Eht: 2
ME 1992 1996	Total: 6900 Kept: 2838 Dropped: 4062	Total: 6900 Kept: 2838 Dropped: 4062	Total: 146 Kept: 109 Dropped: 37	Lat: 50 Lon: 50 Eht: 49	Lat: 138 Lon: 138 Eht: 3
MT 1992 1999	Total: 201 Kept: 196 Dropped:5	Total: 201 Kept: 196 Dropped: 5	Total: 125 Kept: 123 Dropped: 2	Lat: 41 Lon: 41 Eht: 42	Lat: 5 Lon: 5 Eht: 3
NC 1995 2001	Total: 2926 Kept: 1646 Dropped: 1280	Total: 2926 Kept: 1646 Dropped: 1280	Total: 306 Kept: 247 Dropped: 59	Lat: 50 Lon: 50 Eht: 44	Lat: 59 Lon: 59 Eht: 7
NH 1992 1996	Total: 727 Kept: 414 Dropped: 313	Total: 727 Kept: 414 Dropped: 313	Total: 41 Kept: 36 Dropped: 5	Lat: 49 Lon: 49 Eht: 41	Lat: 15 Lon: 15 Eht: 1
NJ 1992 1996	Total: 930 Kept: 495 Dropped: 435	Total: 930 Kept: 495 Dropped: 435	Total: 381 Kept: 323 Dropped: 58	Lat: 49 Lon: 49 Eht: 48	Lat: 19 Lon: 19 Eht: 8
NV 1994 1999	Total: 106 Kept: 94 Dropped:12	Total: 106 Kept: 94 Dropped: 12	Total: 97 Kept: 85 Dropped: 12	Lat: 36 Lon: 36 Eht: 49	Lat: 3 Lon: 3 Eht: 2

NY 1992 1996	Total: 4989 Kept: 2124 Dropped: 2865	Total: 4989 Kept: 2124 Dropped: 2865	Total: 101 Kept: 100 Dropped: 1	Lat: 50 Lon: 50 Eht: 34	Lat: 100 Lon: 100 Eht: 3
OR 1991 1998	Total: 366 Kept: 332 Dropped: 34	Total: 366 Kept: 332 Dropped: 34	Total: 337 Kept: 308 Dropped: 29	Lat: 46 Lon: 46 Eht: 49	Lat: 8 Lon: 8 Eht: 7
PR 1993 1997	Total: 1199 Kept: 601 Dropped: 598	Total: 1199 Kept: 601 Dropped: 598	Total: 50 Kept: 42 Dropped: 8	Lat: 50 Lon: 50 Eht: 50	Lat: 24 Lon: 24 Eht: 1
PR 1997 2002	Total: 51 Kept: 42 Dropped: 9	Total: 51 Kept: 42 Dropped: 9	Total: 51 Kept: 42 Dropped: 9	Lat: 26 Lon: 26 Eht: 26	Lat: 2 Lon: 2 Eht: 2
RI 1992 1996	Total: 1535 Kept: 560 Dropped: 975	Total: 1535 Kept: 560 Dropped: 975	Total: 94 Kept: 85 Dropped: 9	Lat: 50 Lon: 50 Eht: 47	Lat: 31 Lon: 31 Eht: 2
SC 1995 2001	Total: 3478 Kept: 2820 Dropped: 658	Total: 3478 Kept: 2820 Dropped: 658	Total: 3191 Kept: 2584 Dropped: 607	Lat: 50 Lon: 50 Eht: 50	Lat: 70 Lon: 70 Eht: 64
TN 1990 1995	Total: 2848 Kept: 2073 Dropped: 775	Total: 2848 Kept: 2073 Dropped: 775	Total: 134 Kept: 123 Dropped: 11	Lat: 50 Lon: 50 Eht: 45	Lat: 57 Lon: 57 Eht: 3
VQ 1993 2002	Total: 9 Kept: 8 Dropped: 1	Total: 9 Kept: 8 Dropped: 1	Total: 9 Kept: 8 Dropped: 1	Lat: 9 Lon: 9 Eht: 9	Lat: 1 Lon: 1 Eht: 1
VT 1992 1996	Total: 993 Kept: 555 Dropped: 438	Total: 993 Kept: 555 Dropped: 438	Total: 174 Kept: 154 Dropped: 20	Lat: 50 Lon: 50 Eht: 44	Lat: 20 Lon: 20 Eht: 4
WA 1991 1998	Total: 992 Kept: 915 Dropped: 77	Total: 992 Kept: 915 Dropped: 77	Total: 975 Kept: 907 Dropped:68	Lat: 50 Lon: 50 Eht: 49	Lat: 20 Lon: 20 Eht:20
WI 1991 1997	Total: 816 Kept: 752 Dropped: 64	Total: 816 Kept: 752 Dropped: 64	Total: 747 Kept: 700 Dropped:47	Lat: 48 Lon: 48 Eht: 50	Lat: 17 Lon: 17 Eht:15

Creation of all post-BRIDGE portions of GEOCON v2.0

The post-BRIDGE portions of GEOCON v2.0 were developed in three separate waves. Each wave ended up with a wave name used to describe it as the process progressed (much like how the HARN/FBN work was done under the wave name "BRIDGE"). These names were used mostly in file naming during the development itself. But once development was complete, a massive re-naming of files took place to provide a systematic file naming convention across all of the released GEOCON v2.0 files. This information is mostly provided to explain any reference to these code names in the development notebooks.

The first wave was the creation of the FBN-to-NSRS2007 transformation. Like the HARN-to-FBN creation process having the name "BRIDGE", this wave of development was called "**geocon2**". It was the original intent that GEOCON v2.0 functionality would end at this point.

However, as development progressed, it became obvious that the more efficient way to provide transformations to the user community was to envelop the entire GEOCON11 functionality into GEOCON. With that decision, came the next logical decision to support not just NAD 83(2011) but also NAD 83(PA11) and NAD 83(MA11).

Building grids to support PA11 and MA11 was the second wave in developing GEOCON v2.0 and was called "**geocon2a**" while it was ongoing.

Finally, for consistency, it was necessary to not just envelop the old GEOCON11 grids, but actually re-create the NSRS2007-to-2011 grids, using the methods and criteria of all other work in GEOCON v2.0. Thus a third and final wave of development, called "**geocon2b**" was undertaken, where the NSRS2007-to-2011 transformation of GEOCON11 was re-created from scratch.

Wave 1: Creating the FBN-to-NSRS2007 transformations[42] (a.k.a. "geocon2")

As pointed out repeatedly, GEOCON v1.0 transformed from "HARN" to NSRS2007 on a regional basis (over CONUS, Alaska or PR/VI). A similar approach was taken for GEOCON v2.0, with one distinct difference – rather than connect to the ambiguously named "HARN" points, it connected to the rigorously defined supported FBN realization for each state (see earlier), and only to that data. This necessitated a strict filtering process in the input data, which significantly changed the coverage, but not so much the final transformation grid, from GEOCON v1.0.

The FBN-to-NSRS2007 transformation grids were built in the three regions where NSRS2007 exists: CONUS (including DC), Alaska and PR/VI. For each region, the data in the proper "`NADCON5…in`" files were selected. For CONUS, there were 49 files, for Alaska just 1 file and for PR/VI there were 2 files.[43] A similar approach to that already mentioned was followed: Block-median filter the data, grid the thinned data into transformation grids, then return to all the data and randomly shuffle it, then run a cross-validation, run an anti-median filter on the cross-validation results, and grid thinned data into the quality grids, all while creating plots of relevant information along the way.

Similar to the "bridge" work, a work file was needed which contained a variety of pre-computed values, such as coordinate differences (in various units) and azimuths of a combined latitude/longitude shift. To begin with, a list of the "`NADCON***.in`" files which would be used was created. A simple list file (called "`GEOCON2.filelist`") was manually created which listed all of the input files to become part of the "work file"[44]. A program took the file list, and created a "work" file[45], containing all of the pre-computed information necessary to grid the transformations.

Prior to moving on, an analysis between the work file for GEOCON v1.0 ("HARN" to NSRS2007) and the work file for the FBN to NSRS2007 component of GEOCON v2.0 (aka "geocon2") was conducted.

GEOCON v1.0 vs GEOCON v2.0 work files

It is to be expected that the work files between these two builds of GEOCON will differ, as the "old datum" information in v1.0 will be a mix of HARN, FBN, both or neither whereas the "old datum" information in v2.0 will be explicitly for the supported FBN realization in each state. Note that the work file for v1.0 contained a number of points not in the 48 CONUS states, nor in Alaska, DC, Puerto Rico nor the Virgin Islands. These were not included in the comparison[46].

# points in GEOCON v1.0 work file[47]	:	69,320
# points in GEOCON v2.0 work file	:	63,068
# points in common	:	62,188

The reasons for the different counts are (a) v1.0 allowed any published "old datum" information to stand whereas v2.0 only took FBN information, reducing the count of v2.0 below v1.0 (b) new data may have been loaded into the NGS IDB between the release of v1.0 and v2.0, which could account for there being 880 points in v2.0 that are not in v1.0 or (c) v2.0 allowed points to be used that only had a good lat/lon pair, wheras v1.0

required that lat/lon/eht all had to exist, again possibly accounting for some of the 880 points in v2.0 that are not in v1.0.

Of the 62,188 common points, the coordinate differences (NSRS2007 minus "old datum") were compared to one another[48]. The difference between "HARN" and FBN should account for the entirety of these differences. The statistics of the differences are seen in Table 6.

Table 6: Comparison of GEOCON v1.0 and v2.0

	[ϕ(NSRS2007)-ϕ("HARN")] - [ϕ(NSRS2007)-ϕ(FBN)]	[λ(NSRS2007)-λ("HARN")]- [λ(NSRS2007)-λ(FBN)]	[h(NSRS2007)-h("HARN")]- [h(NSRS2007)-h(FBN)]
N1*	62,188	62,188	62,118
N2**	62,188	62,188	61,778
Tolerance	0.001 arcseconds	0.001 arcseconds	0.1 meters
N3***	92	206	111
The statistics below ignore the outliers listed above.			
N4****	62,096	61,982	61,667
AVE	0.00000 arcsec	0.00000 arcsec	0.00020 meters
STD	0.00004 arcsec	0.00004 arcsec	0.00693 meters
RMS	0.00004 arcsec	0.00004 arcsec	0.00693 meters

* N1 = Number of points in common between the work files for GEOCON v1.0 and GEOCON v2.0
**N2 = Of N1, the number of points that do not have a rejection criteria in v1.0 nor in v2.0
***N3 = Of N2, the number of points outside of the tolerance (in absolute value)
****N4 = N2-N3 (number of common, non-rejected points that fall inside the tolerance). The statistics in the table below N4 refer only to the N4 points themselves.

Note that aside from a very small number of outliers (< 4% of all non-rejected points in all cases) the agreement in transformative data between version 1.0 ("HARN") and 2.0 (FBN) is excellent. With solidly zero averages, the 0.00004 arcseconds translates into about 1.2 millimeters in latitude and lonigitude, while ellipsoid heights have about a 6.9 millimeter agreement (around a mean of 0.2 mm). Outliers were briefly studied, but no significant issues were found with the data. The source of the outliers was always due to a difference between the use of "HARN" versus the use of FBN as the "old datum" in the transformation.

<u>Creation of grids</u>

After the work files had been compared, creation of the transformation grids commenced. The basic workflow follows that of the individual BRIDGE grids, but is now done for the three regions (CONUS, Alaska, PR/VI) in which there are FBN/NSRS2007 pairs.

As a first step, coverage plots were created.[49] The next step was to create, from the "work" file a set of plots showing the vector changes in latitude, longitude, ellipsoid height and a combination of latitude/longitude (aligned along the proper azimuth).[50] As was done in GEOCON v1.0, the CONUS vector plots were generated in a mosaic of 9 overlapping sub-regions, just to provide some clarity. This helped identify outliers and other troublesome points. An analysis of some obvious differences between v1.0 and v2.0 was performed, and confidence was gained that all of the differences were explainable.[51]

The residual values (delta latitude, delta longitude, delta ellipsoid height, delta horizontal) were then run through a block-median filter at a 1 x 1 arcminute basis. As in the BRIDGE build there were "**<u>kept points</u>**". and "**<u>dropped points</u>**". For more details, see the BRIDGE section of this report, earlier.

As a next step, to control edge effects, a "mask" file was needed to buffer around each region. Alaska, being a one-state "region" had its own mask already. However, the 49 masks of CONUS (48 states + DC) were combined into one mask for CONUS. Similarly, the two masks of Puerto Rico and the US Virgin Islands were combined into one mask for PRVI.[52]

The differential coordinates were thinned and gridded, one grid each for latitude, longitude, ellipsoid height and horizontal magnitude. The grids were compared to those from v1.0 and a brief analysis of their major differences performed[53]. Significant differences were discussed with Dr. Dennis Milbert (the author of GEOCON v1.0) and were always explainable. Confidence in the grids of v2.0 was achieved so that the build could continue.

Cross-Validation

The cross validation methodology followed the same scheme used in v1.0 and in the BRIDGE portion of v2.0 and will not be repeated here. The statistics of how many points were good, how many thinned points were kept, and the size of cross validation data sets is found below. The three regions each have a 2 character code (CONUS = CS, Alaska = AK and PR/VI = PV).

Table 7:Gridding point counts and cross-validation loop information for regions with an FBN/NSRS2007 transformation.

REGION **Old Realization** **New Realization**	**Latitude Counts**	**Longitude Counts**	**Ellipsoid Height Counts**	**Number of Cross-Validation Loops ("nloops")**	**Number of points pulled out in each loop ("nlucky")**
CS FBN* NSRS2007	**Total: 62,244** **Kept: 49,752** **Dropped: 12,492**	**Total: 62,244** **Kept: 49,752** **Dropped: 12,492**	**Total: 60,998** **Kept: 48,797** **Dropped: 12,201**	**Lat: 63** **Lon: 63** **Eht: 63**	**Lat: 988** **Lon: 988** **Eht: 969**
AK 1992 (FBN) NSRS2007	**Total: 722** **Kept: 611** **Dropped: 111**	**Total: 722** **Kept: 611** **Dropped: 111**	**Total: 720** **Kept: 611** **Dropped: 109**	**Lat: 63** **Lon: 63** **Eht: 63**	**Lat: 12** **Lon: 12** **Eht: 12**
PV 2002 NSRS2007	**Total: 101** **Kept: 70** **Dropped: 31**	**Total: 101** **Kept: 70** **Dropped: 31**	**Total: 101** **Kept: 70** **Dropped: 31**	**Lat: 101** **Lon: 101** **Eht: 101**	**Lat: 1** **Lon: 1** **Eht: 1**

*FBN here means that the 49 separate FBN realizations in the 48 states and DC of CONUS were all used to form one regional transformation within CONUS connected to the NSRS2007 realization of NAD 83.

After cross validation was finished, the data were also gridded into "quality" grids (aka "worst case error" grids) and plots made[54].

Coding "geocon.v2.0.f"

As mentioned earlier, the FBN-to-NSRS2007 transformation was originally planned to be the final set of supported transformations for GEOCON v2.0. As such, at this point, work on the actual code for running the transformations (in FORTRAN90, called "`geocon.v2.0.f`") was begun[55]. However shortly after beginning coding, it became apparent that the scope of changes between `geocon.v1.0.f` and `geocon.v2.0.f` was going to be so extensive that it made sense to just envelop all of GEOCON11 into GEOCON, as well as any other transformations supporting NAD 83(2011), since an entire re-write was occurring. As such, work was paused on "`geocon.v2.0.f`" and work begun on making all supported realizations of NAD 83 through 2011 part of GEOCON v2.0.

Wave 2: Creating the FBN-to-MA11 and FBN-to-PA11 transformations[56] (a.k.a. "geocon2a")

GEOCON11 only covered transformations between NAD 83(NSRS2007) and NAD 83(2011), and only in the regions of CONUS, Alaska and PR/VI. Because it was decided to support all of the 2011 realizations of NAD 83 (including 2011, MA11 and PA11) it was not enough to simply move the functionality of GEOCON11 into GEOCON v2.0. Entirely new transformations needed to be created for three regions: American Samoa (AS), Hawaii (HI) and Guam/CNMI (GC).

The creation of these grids followed similar development waves. Transformations were built for the three regions of American Samoa (aka "AS", between the FBN of 2002 and NAD 83(PA11)), Hawaii (aka "HI", between the FBN of 1993 and NAD 83(PA11)) and for Guam/CNMI (aka "GC", between their two FBN's of 2002 and 2002 respectively, and NAD 83(MA11)).

Similar to the "bridge" work, a work file was needed which contained a variety of pre-computed values, such as coordinate differences (in various units) and azimuths of a combined latitude/longitude shift. To begin with, a list of the "`NADCON***.in`" files which would be used was created. A simple list file (called "`GEOCON2A.filelist`") was manually created which listed all of the input files to become part of the "work file"[57]. For this functionality (FBN-to-PA11 or FBN-to-MA11) three separate sets of transformation grids would need to be created for GEOCON v2.0. A program took this file list, and created a "work" file[58], containing all of the pre-computed information necessary to grid the transformations.

Coverage plots and vector plots were created[59]. Data were then thinned, plotted and cross-validated as before. Statistics on thinning and cross-validation loop sizes are found in Table 8.

Table 8: Gridding point counts and cross-validation loop information for regions with an FBN/MA11 or FBN/PA11 transformation.

REGION Old Realization New Realization	Latitude Counts	Longitude Counts	Ellipsoid Height Counts	Number of Cross-Validation Loops ("nloops")	Number of points pulled out in each loop ("nlucky")
AS 2002 (FBN) PA11	**Total: 55 Kept: 23 Dropped: 32**	**Total: 55 Kept: 23 Dropped: 32**	**Total: 55 Kept: 23 Dropped: 32**	**Lat: 55 Lon: 55 Eht: 55**	**Lat: 1 Lon: 1 Eht: 1**
HI 1993 (FBN) PA11	**Total: 248 Kept: 137 Dropped: 111**	**Total: 248 Kept: 137 Dropped: 111**	**Total: 244 Kept: 134 Dropped: 110**	**Lat: 248 Lon: 248 Eht: 244[60]**	**Lat: 1 Lon: 1 Eht: 1**
GC FBN* MA11	**Total: 10 Kept: 7 Dropped: 3**	**Total: 10 Kept: 7 Dropped: 3**	**Total: 10 Kept: 7 Dropped: 3**	**Lat: 10 Lon: 10 Eht: 10**	**Lat: 1 Lon: 1 Eht: 1**

*FBN here means that the 2 separate FBN realizations in the 2 territories (Guam and CNMI) of region GC were both used to form one regional transformation within GC connected to the MA11 realization of NAD 83.

Wave 3: Creating the NSRS2007-to-2011 transformations[61] (a.k.a. "geocon2b")

For the purposes of consistency with the general methodology of other grids in GEOCON v2.0, it was seen as necessary to re-create the NSRS2007-to-2011 transformation grids, rather than simply relying on the pre-existing grids of GEOCON 11 v1.0/v1.1. Very little disagreement was expected between the new grids and the old grids, due to there being no difference in the actual realizations supported (NSRS2007 and 2011) for v1.0/1.1 and v2.0.

Similar to previous development waves of GEOCON v2.0, a work file was needed which contained a variety of pre-computed values, such as coordinate differences (in various units) and azimuths of a combined latitude/longitude shift. To begin with, a list of the "**`NADCON***.in`**" files which would be used was created. A simple list file (called "**`GEOCON2B.filelist`**") was manually created which listed all of the input files to become part of the "work file"[62].

A comparison[63] program generated the statistical differences between GEOCON11 v1.0 and GEOCON v2.0 grids, seen in Table 9.

Table 9: Comparison of GEOCON11 v1.0 and GEOCON v2.0

	[ϕ(2011)-ϕ(NSRS2007)] v2.0-[ϕ(2011)-ϕ(NSRS2007)]v1.0	[λ(2011)-λ(NSRS2007)]v2.0-[λ(2011)-λ(NSRS2007)]v1.0	[h(2011)-h(NSRS2007)]v2.0-[h(2011)-h(NSRS2007)]v1.0
N1*	74,059	74,059	74,059
N2**	74,059	74,059	74,003
Tolerance	0.001 arcseconds	0.001 arcseconds	0.1 meters
N3***	11	22	8
The statistics below ignore the outliers listed above.			
N4****	74,048	74,037	73,995
AVE	0.00000 arcsec	0.00000 arcsec	0.00020 meters
STD	0.00001 arcsec	0.00002 arcsec	0.00206 meters
RMS	0.00001 arcsec	0.00002 arcsec	0.00206 meters

* N1 = Number of points in common between the work files for GEOCON11 v1.0 and GEOCON v2.0 for the NSRS2007/2011 transformation.
**N2 = Of N1, the number of points that do not have a rejection criteria in v1.0 nor in v2.0
***N3 = Of N2, the number of points outside of the tolerance (in absolute value)
****N4 = N2-N3 (number of common, non-rejected points that fall inside the tolerance). The statistics in the table below N4 refer only to the N4 points themselves.

Note that aside from a very small number of outliers (< 0.03% points in all cases) the agreement in transformative data between GEOCON11 v1.0 and GEOCON v2.0 is excellent, but it is not perfect. With solidly zero averages, the 0.00001 arcseconds translates into about 0.3 millimeters in latitude and lonigitude, while ellipsoid heights have about a 2.1 millimeter agreement (around a mean of 0.2 mm). Because the data in both GEOCON11 v1.0 and GEOCON v2.0 should be (theoretically) identical for the NSRS2007-to-2011 transformation, it is worth investigating why there are any discrepencies at all, even though they are generally small. A few points common to GEOCON11 v1.0 and GEOCON v2.0 were chosen where their transformation data exceeded the outlier criteria. In every point checked the data showed that either an error was made in GEOCON11 v1.0 which had been corrected in GEOCON v2.0, or that data used in GEOCON11 v1.0 had since been superseded with new data since the release of GEOCON11 v1.0 and that GEOCON v2.0 was using the new data. With spot checks validating GEOCON v2.0 in every case, it was presumed that the GEOCON v2.0 build was correct.[64]

For this final functionality (NSRS2007-to-2011) three separate sets of transformation grids would need to be created for GEOCON v2.0, one each in CONUS, Alaska and PR/VI. A program took the file list, and created a "work" file[65], containing all of the pre-computed information necessary to grid the transformations.

Coverage plots and vector plots were created. Data were then thinned, plotted and cross-validated as before. Statistics on thinning and cross-validation loop sizes are found in Table 10.

Table 10: Gridding point counts and cross-validation loop information for regions with a NSRS2007/2011 transformation.

REGION Old Realization New Realization	Latitude Counts	Longitude Counts	Ellipsoid Height Counts	Number of Cross-Validation Loops ("nloops")	Number of points pulled out in each loop ("nlucky")
CS NSRS2007 2011	**Total: 74,132 Kept: 58,753 Dropped: 15,379**	**Total: 74,132 Kept: 58,753 Dropped: 15,379**	**Total: 74,075 Kept: 58,708 Dropped: 15,367**	**Lat: 75 Lon: 75 Eht: 75**	**Lat: 989 Lon: 989 Eht: 988**
AK NSRS2007 2011	**Total: 830 Kept: 672 Dropped: 158**	**Total: 830 Kept: 672 Dropped: 158**	**Total: 830 Kept: 672 Dropped: 158**	**Lat: 70 Lon: 70 Eht: 70**	**Lat: 12 Lon: 12 Eht: 12**
PV NSRS2007 2011	**Total: 127 Kept: 87 Dropped: 40**	**Total: 127 Kept: 87 Dropped: 40**	**Total: 127 Kept: 87 Dropped: 40**	**Lat: 127 Lon: 127 Eht: 127**	**Lat: 1 Lon: 1 Eht: 1**

After cross validation was finished, the data were also gridded into "quality" grids (aka "worst case error" grids) and plots made[66].

Re-naming Everything

In hindsight, a common, simple, comprehensive naming scheme should have been selected prior to work. To correct this oversight, prior to the release of GEOCON v2.0, massive batch files[67] were used to rename every major output file (including graphics) into a new naming convention. The naming convention chosen is described below.

Every file will have the following naming style, broken down into seven parts, each separated by a period:

AA.RR.OOOOOOOOOO.NNNNNNNNNN.EEE.YYYYMMDD.end

Part 1 "AA" (2 characters):

A 2 digit code telling the user what kind of data is in the file:

- **`ds`** : file is a binary transformation grid *or* a jpg of the transformation grid
- **`de`** : file is a binary transformation quality grid *or* a jpg of the transformation quality grid
- **`ol`** : file is a text outlier list
- **`cv`** : file is a jpg of the coverage of points
- **`vc`** : file is a jpg of the transformation vectors

Part 2 "RR" (2 characters):

A 2-character state, territory or region code describing the area in which this data file is valid. In the case of the 50 states, the standard 2-character state codes are used, and are not repeated here. Other codes are:

Territories:

- **`DC`** : District of Columbia
- **`PR`** : Puerto Rico
- **`VQ`** : U.S. Virgin Islands

AS : American Samoa
GQ : Guam
CQ : Commonwealth of the Northern Mariana Islands

Regions (multi-state or multi-territory):
CS : CONUS (lower 48 states plus the District of Columbia)
GC : Guam and Commonwealth of the Northern Mariana Islands
PV : Puerto Rico and the U.S. Virgin Islands

Part 3 "OOOOOOOOOO" (10 characters):

This is where the "old datum" or "old datum realization" name resides. All transformation data (particularly the grids) follow the sign convention of "new minus old". In the case of GEOCON v2.0, all data used are realizations of NAD 83, but in the future, should other transformation software be merged with GEOCON, these ten characters can hold any datum or datum realization.

For GEOCON v2.0, three naming schemes will be seen in this location:

- **NAD83_YYYY** which means "NAD 83(YYYY)" where "YYYY" falls between 1986 and 2007, non-inclusively.
- **NAD83_2007** which means "NAD 83(NSRS2007)"
- **NAD83_FBN_** which means "a collection of data across states or territories, each on its own FBN" and "FBN" is used to mean the last supported realization in a state prior to 2007.

That third selection occurs because it was decided to allow the regional transformation of the state-specific FBN realizations to connect to the *regional* NSRS2007 and 2011/PA11/MA11 realizations.

Part 4 "NNNNNNNNNN" (10 characters)

This is where the "new datum" or "new datum realization" name resides. All transformation data (particularly the grids) follow the sign convention of "new minus old". In the case of GEOCON v2.0, all data used are realizations of NAD 83, but in the future, should other transformation software be merged with GEOCON, these ten characters can be used to hold any datum or datum realization.

For GEOCON v2.0, four naming schemes will be seen in this location:

- **NAD83_YYYY** which means "NAD 83(YYYY)" where "YYYY" falls between 1987 and 2011, inclusively, but excluding 2007.
- **NAD83_2007** which means "NAD 83(NSRS2007)"
- **NAD83_PA11** which means NAD 83(PA11)
- **NAD83_MA11** which means NAD 83(MA11)

Part 5 "EEE" (3 characters)

These three characters will describe what type of data the file contains:

lat : geodetic latitude
lon : geodetic longitude
eht : ellipsoid height
hor : horizontal (combination of latitude and longitude)

Part 6 "YYYYMMDD" (8 characters)

An eight character date, indicating when the file was created.

YYYYMMDD : Is the month (MM), day of month (DD) and year (YYYY) the file was created. (Not to be confused with the YYYY used in the old datum or new datum part of the file.)

Part 7 "end" (1 or 3 characters)

A 1 or 3 character ending, telling what kind of file this is.

b : A binary grid, in a format standard to NGS grids
txt : An ASCII text file
jpg : A jpg image file

Example

Using the above information, **ds.WI.NAD83_1991.NAD83_1997.eht.20140825.b** can be seen to be the binary transformation grid between NAD 83(1991:WI) and NAD 83(1997:WI), for ellipsoid heights, as created on August 25th, 2014. Data in said grid is of the form h(NAD 83(1997:WI)) minus h(NAD 83(1991:WI)).

Release

Coding of the primary executable (**geocon.v2.0.f**) was finished on October 27, 2014. Release on the NGS development website occurred on ?????????????. A complete set of JPG files showing coverage, vectors, and transformation magnitudes was released in association with GEOCON V2.0 as well at ?????????????????.