9/30/2011

Task 3: Pilot Project- NHDPlus Tools Test in Idaho

COOPERATIVE AGREEMENT NO. G10AC00432

IDAHO DEPARTMENT OF WATER RESOURCES
322 EAST FRONT ST.
P.O. BOX 83720
BOISE, ID 83720

Linda Davis, Idaho NHD Principal Steward
Genna Ashley, Idaho NHD Tech. Point of Contact
Danielle Favreau, GIS Analyst

REPORT TO THE UNITED STATES GEOLOGICAL SURVEY

Jeffrey Simley
National Hydrography Dataset
U.S. Geological Survey , NGTOC
P.O. Box 25046, MS-510, DFC
Denver, Colorado 80225

Scott Van Hoff
Idaho Geospatial Liaison
U.S. Geological Survey
NSDI Partnership Office
230 Collins Road
Boise, Idaho 83702

Task 3: Pilot Project- NHDPlus Tools Test in Idaho

COOPERATIVE AGREEMENT NO. G10AC00432

TABLE OF CONTENTS

B. Draft NHDPlus User Toolbox User Guide

C. List of Edits That Need To Be Done To NHDflowlines And NHD Related Tables

Endnotes

Introduction:

The NHDPlus is a suite of geospatial datasets that contain hydrography with value added attributes such as flow volume, velocity estimates and catchments characteristics. This data is based on the National Hydrography Dataset (NHD), the National Elevation Dataset, the National Land Cover Dataset and the Watershed Boundary Dataset.[i] Currently, the NHDPlus is available at the 1:100,000 level. However, Horizon Systems Corporation (HSC), in association with the United State Geological Survey (USGS) and the Environmental Protection Agency (EPA) Office of Water, has developed Quality Assurance/Quality Control tools that can be applied to 1:24:000, high resolution NHD in preparation for creating high resolution NHDPlus data. These tools are called the NHDPlus NHD QAQC Tool (QA/QC Tool). The Idaho Department of Water Resources (IDWR) selected three 12-digit watersheds, one mountainous, one low relief agricultural, and one area that contain a mix of rural, urban and agricultural uses for testing of the QA/QC Tool.

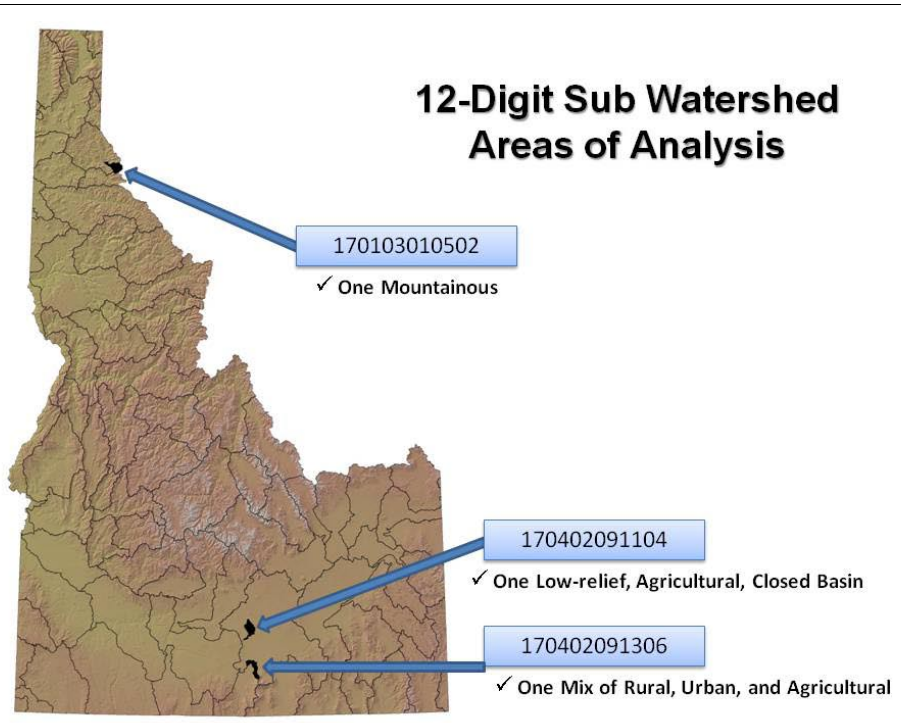


Figure 1: The three selected subwatersheds (6th fields) in Idaho

Areas of Analysis

I. Hydrologic Unit (HU) 170103010502 Butte Gulch-Prichard Creek:

The HU 170103010502 consists of Prichard Creek and all of its contributing tributaries. Prichard Creek flows into the North Fork of the Coeur d'Alene River, with the townsite of Prichard, Idaho at the confluence. The SubWatershed abuts the Idaho/Montana border and contains the townsites of Prichard and Murray, Idaho. Just 24 miles from the Shoshone County Seat of Wallace, Murray, Idaho is a small community of less than 100 and Prichard townsite has a population of zero (Figure 2).[ii] The majority of the HU is forested, with some areas of rangeland (less than 6%). The HU is primarily federally managed (78% USFS, 7% BLM) and has an extensive forest products and mining history. This HU has a maximum elevation of about 6800 ft. and a minimum elevation of about 2400 ft. HU 170103010502 is approximately 34060 acres (53 sq. miles) putting it in top 10% of the state in terms of size of 12 digit watersheds.

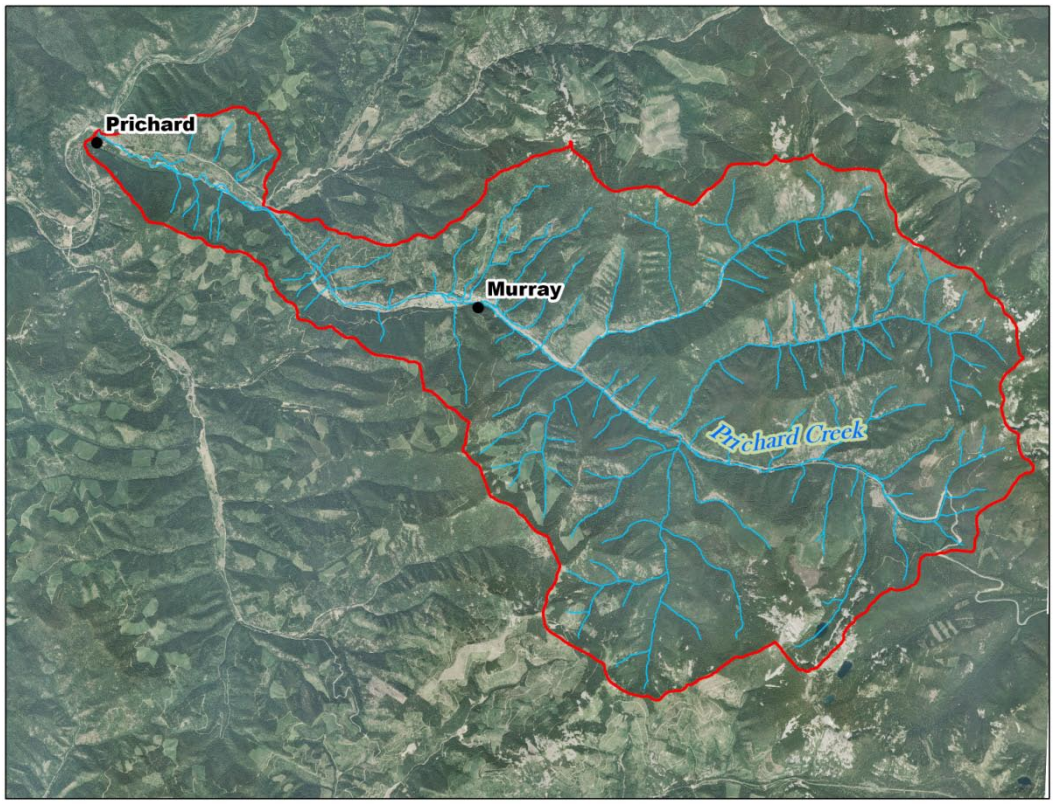


Figure 2: Prichard Creek 6th field Subwatershed

II. Hydrologic Unit (HU) 170402091104 Town of Kimama:

The HU 170402091104 is a closed basin that contains no named streams or creeks (figure 3). The southern tail of the HU is crossed by Idaho State Highway 24 and the Union Pacific Railroad. The HU contains the town site of Kimama, although the 2010 Census population is zero.[iii] Land Use in the HU is approximately 25% agriculture and 75% rangeland. Agriculture is concentrated in the southern portion of the HU and the northwestern portion of the HU is dominated by lava fields. Nearly 66% of the HU is managed by the BLM. This HU has a maximum elevation of about 4650 ft. and a minimum elevation of 4225 ft. HU 17040209104 is approximately 33949 acres (53 sq. miles) putting it in the top 10% of the state in terms of size.

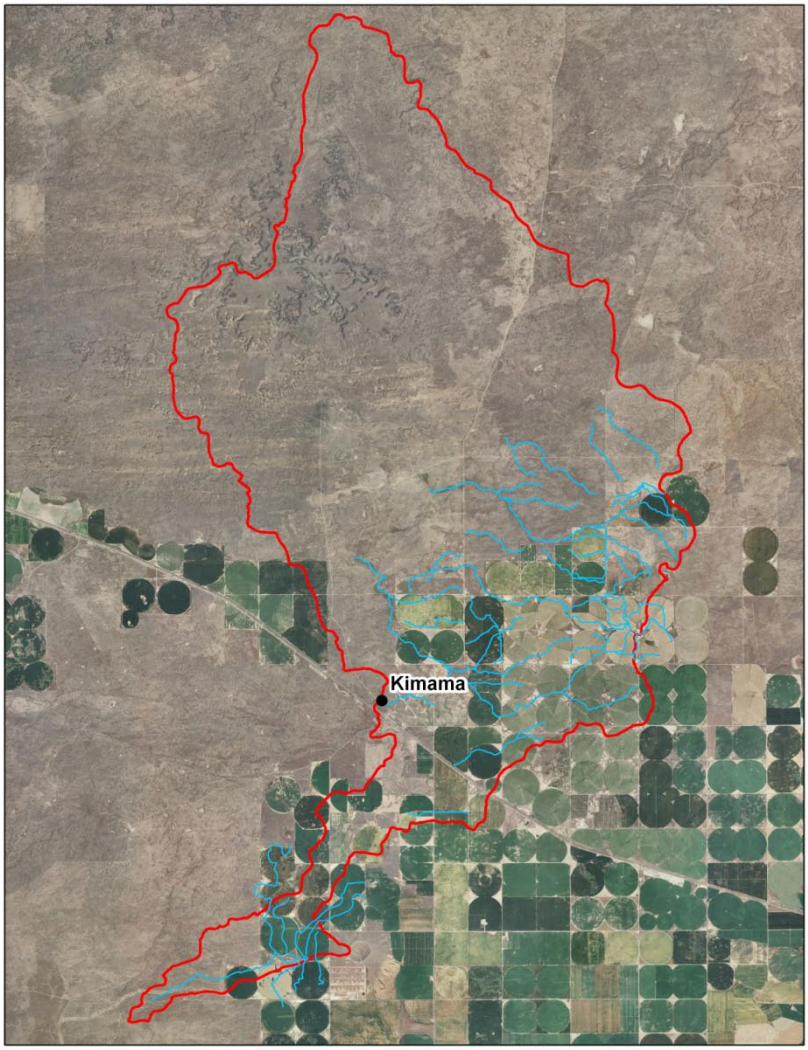


Figure 3: Town of Kimama 6th field Subwatershed

III. Hydrologic Unit (HU) 170402091306 Snake River - Burley:

A principal feature of HU 170402091306 is the Snake River which bisects the northern portion of the HU (figure 4). The City of Burley (population approximately 10,000) dominates the south side of the river and the City of Heyburn (pop. approximately 3000) is on the north side.[iv] The Southern portion of the HU is primarily agriculture (Land Use: 71% agriculture, 15% rangeland, and 10 % urban). The natural hydrology of HU 170402091306 has been significantly changed due to agricultural and urban growth and consists mainly of canals and laterals of the Burley and Minidoka Irrigation Districts. The majority of the HU is privately owned with some BLM and USFS in the southernmost tip of the HU. The minimum elevation for this HU is about 4100 feet. Although this HU has a maximum elevation of about 7400 feet, only the southern 20% of the HU is above 4300 feet. This area includes Antelope Hill and the foothills of the Albion Mountains. HU 170402091306 is approximately 29008 acres (45 sq. miles) putting it in the top 20% of the state in terms of size.

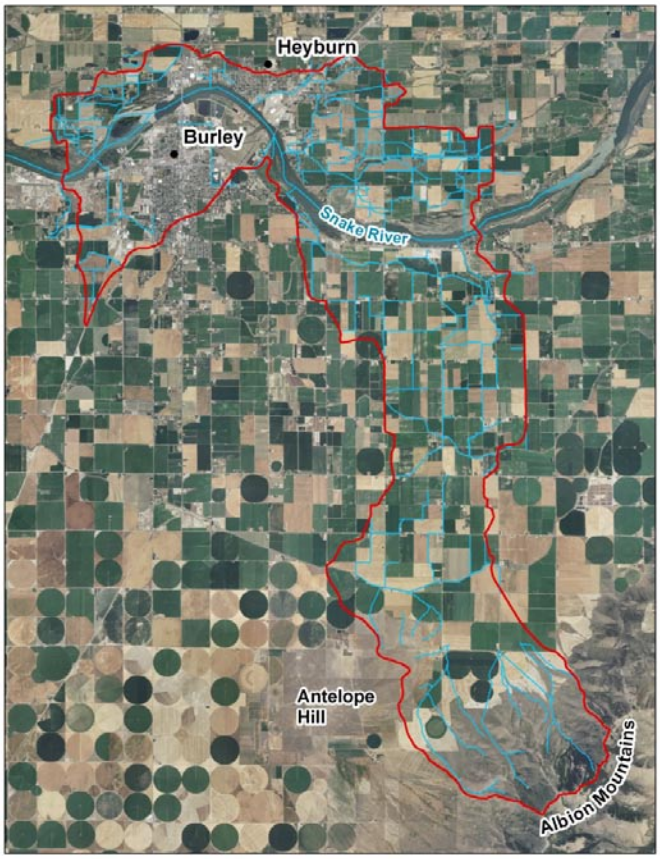


Figure 4: Snake River-Burley 6th field Subwatershed

General Description of QA\QC Tools

The NHDPlus NHD QAQC Tool Version 1.0.3 (QA\QC Tool) was provided via FTP download to IDWR from Horizon Systems Corporation (HSC). The QA\QC Tool performs a series of quality control and assurance checks on a NHD geodatabase focused primarily on the network described by the NHD flow table, but also included feature geometry checks and feature name checks.[v] The tools were provided in an ArcGIS Toolbox format compatible with ArcGIS 9.3.1.

Methods & Results:

Criteria for SubWatershed Selection

IDWR considered several conditions when selecting the three 12-digit watersheds to be used in the analysis of the QA\QC Tool. The goal was to have a sample set of watersheds that represent different conditions found in Idaho. A general summary of preferences are listed below:

A) Primary preferences:

- One mountainous Subwatershed
- One low relief, agricultural Subwatershed
- One Subwatershed that contain a mix of rural, urban and agricultural uses
- At least one of the watersheds would have an inflow

B) Other preferences:

- One closed basin
- One headwaters basin
- One Subwatershed with recently updated NHDflowline data. (Local data submissions, not including Maintenance -Lite I or II)
- Subwatershed completely within Idaho
- Subbasin not checked out
- Watersheds with denser hydrography
- Watersheds within the same subbasin
- Watersheds with supplemental datasets

The primary preferences were conditions originally described in the project description. Other preferences were determined in order to address Idaho's unique and diverse landscape or to address technical requirement in order to run the QA/QC Tool.

Some watersheds within Idaho, especially in southern Idaho, have limited hydrography due to lava fields, alluvial fans and agricultural practices. A more densified hydrography is required to be able to run the QA/QC Tools and get a snapshot of possible problems that may need to be addressed. Watersheds in the same subbasin were preferred in order to limit the effects of different demographic factors, such as agricultural practices, may have on the results of the QA/QC Tool. Also, watersheds that had supplemental datasets that would allow for the determination of conditions such as the amount of agriculture, relief, updated linework, and land use in order to identify if a watershed met our primary conditions.

There are 86 Subbasins (HU-8) and approximately 2500 SubWatersheds (HU-12) completely or partially within the State of Idaho. Of those Subbasins, only four had a significant amount of updated NHDFlowline work. Two of those four Subbasins were significantly covered by lava fields. Since these subbasins within lava fields tend to have disjointed sparse hydrography, they were eliminated. The two remaining Subbasins, 17040206 & 17040209, are low relief, high agricultural areas when compared to other Subbasins in the state.

Next, the closed basins within Subbasins 17040206 & 17040209 were reviewed. Most of the closed basins did not have a significant flowline density. If two closed basins had similar flowline density, SubWatersheds with more traditional/connected flowlines were preferred. SubWatershed 170402091104 was selected as a pilot area because it is a closed basin of sufficient flowline density, has low relief, and land use within it is predominantly agriculture.

SubWatershed 170402091306 was selected next. This SubWatershed was chosen because it contained the City of Burley with outlying rural areas, the primary land use is agriculture, and it has an inflow.

Although it would have been preferred to select watersheds all within the same Subbasin, HU 17040209 did not contain a watershed that could be considered high relief when compared to other watersheds in the state. More extreme relief in Idaho is found in Sub-Regions 1706 and 1701. Sub-Region 1706 eliminated because it was checked out by USGS for Maint-Lite II. In Sub-Region 1701, IDWR has access to Land Use data comparable to that available for Subbasin 17040209 for Basin 170103. Subbasins 17010305, 17010306, and 17010308 were eliminated because significant portions of the subbasins were outside the Idaho State boundary. Subbasin 17010303 was eliminated as it was checked out by IDWR for another project.

SubWatershed 170103010502 was selected as the pilot subwatershed to represent high relief, headwaters areas. This SubWatershed is the third largest Headwaters HU in subbasins 17010301, 17010302, and 17010304.

Data Preparation and Use of the QA/QC Tool

All data preparation and use of the QA\QC Tool was done within the ArcGIS 9.3.1 environment.

Data preparation began by downloading a copy of the NHD for the appropriate SubBasins from the NHD website (http://nhd.usgs.gov/data.html). Because the 8-digit SubBasins are the minimum unit currently downloadable from the NHD website, the SubBasin data was modified to the SubWatershed level via the following steps:

- Deleted the existing network (via ArcCatalog)
  - Remove any old network files and/or connection tables
- Modified NHDFlowline Feature Class
  - ArcGIS Process: Select By Location. Removed all line segments that did not intersect the 12-digit SubWatershed of interest.
- Rebuild the network and flow tables using the NHDGeoEdit Tools. The steps used to rebuild the network were necessary because this process required that the dataset be in NHD data-model format. The steps required to complete this task are described in Appendix A.

Once the NHDFlowline data has been modified and the network rebuilt, The QA/QC Tool was run from the ArcToolbox Menu (NHDPlus User Toolbox) on each of the SubWatersheds. The following inputs were used:

- Click on 05. NHDQAQC Version 1.0.3.2
  - Dialog Box appears (figure 1):
    - Build Name – Pre-populated
    - NHD Data Format: Personal Geodatabase
    - Input File: Navigate to the .mdb file of a SubWatershed
    - Temporary Work Area: D:/Temp
- All Checks, with the exception of Check #16 (Flowing Connections and Elevations Disagree) were populated. Check #16 was permanently grayed; this option was not selectable.

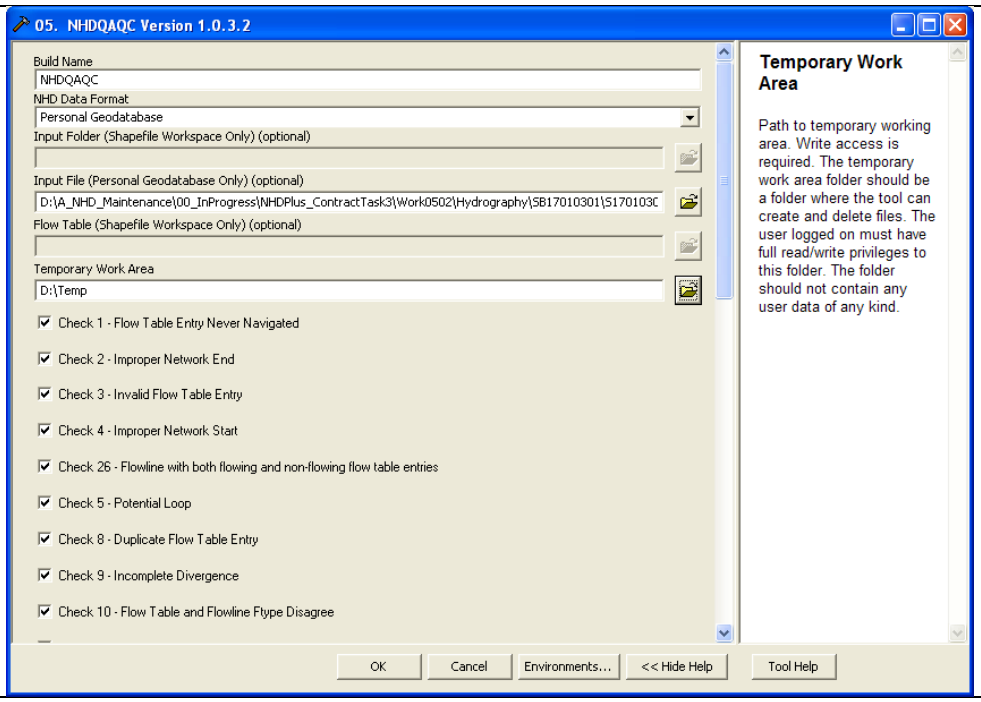


Figure5 – NHD Plus QA/QC Tools Dialog Box

Results

The results of each running of the QA/QC Tool results in a table added to the geodatabase of the input file. Errors are indicated in the table by an Error Code. The errors reported in each HU are described below:

I. Hydrologic Unit (HU) 170103010502
   a. Error Code 12 (Isolated Network), 5 Errors
II. Hydrologic Unit (HU) 170402091104
   a. Error Code 12 (Isolated Network), 20 Errors
III. Hydrologic Unit (HU) 170402091306
   a. Error Codes: 12 (Isolated Network), 65 Errors; Error Code 14 (Flowlines Relate in Flow Table But Do Not Touch), 1 Error; 19 (Non-Linear Reach), 1 Error; & 25 (Nonlinear Named Path), 16 Errors

Table 1: Summary of All Errors for all three Subwatersheds			
Error Code	**Subwatershed**		
	170103010502	170402091104	170402091306
Error Code 12 *Isolated Network*	5	20	65
Error Code 14 *Flowlines Relate in Flow Table but do not Touch*	0	0	1
Error Code 19 *Non-Linear Reach*	0	0	1
Error Code 25 *Non Linear Named Path*	0	0	16
Subwatershed Totals	**5**	**20**	**83**
Grand Total = 108			

Of the four Error Codes reported for the three Subwatersheds, three were considered severe and are required to be fixed before moving forward to NHDPlus creation; Error Codes 14, 19 and 25. Error Code 14 (Flowlines Relate in Flow Table But Do Not Touch) checks for an agreement between the flowline geometry and/or attributes and the flow table, Error Code 19 (Non-Linear Reach) checks to make sure that there are no gaps between segments with the same reachcode, and Error Code 25 (Non-linear named Path) checks to make sure that there are no gaps between segments with the same GNIS numbers.

For Subwatershed 170103010502, there were 5 errors; all Error Code # 12 “Isolated Network”. In four of these cases, each NHDFlowline segment was correct, and was, in fact, an isolated network. This includes the mouth of Prichard Creek, where the subbasin was clipped to the 6th field sub watershed boundary line for this study. For the 5th isolated network error, it is uncertain whether the flowline goes under a road and flows into an NHDWaterbody, or flows into Prichard Creek. This information has been sent to the local data steward to ask them to review the aerial photo and NHD linework to clarify the flow direction, and verify that the NHDWaterbody in this area is accurate. As expected, this subbasin produced the least amount of errors, as it is a forested, classic subwatershed in a high-relief area.

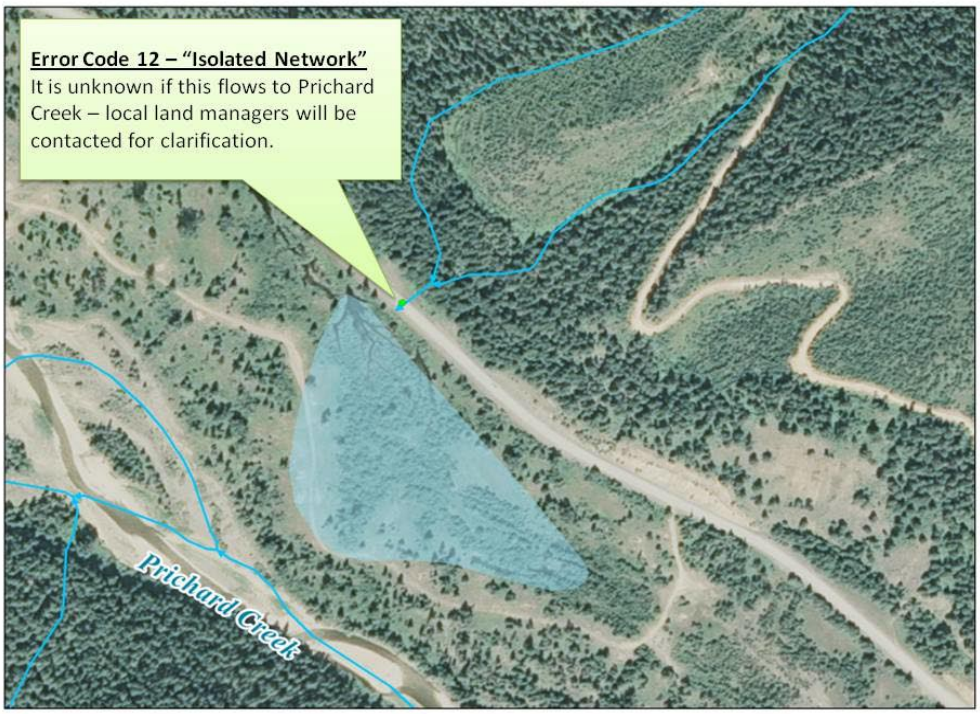


Figure 6 – Example of Error Code 12 - Questionable

Subwatershed 170402091104 was chosen as a representative of a closed basin 12-digit unit. All errors in this subwatershed were warning errors; Error Code #12 "Isolated Network", with no other error codes present. Of the 20 errors presented, 11 of these segments were, in fact, isolated networks, and part of "closed basins" and did not connect to another canal or stream in the network. There were 9 instances where this error was identified but it is unclear as to how they are isolated networks. Examples of this unknown condition are depicted in figure 7.

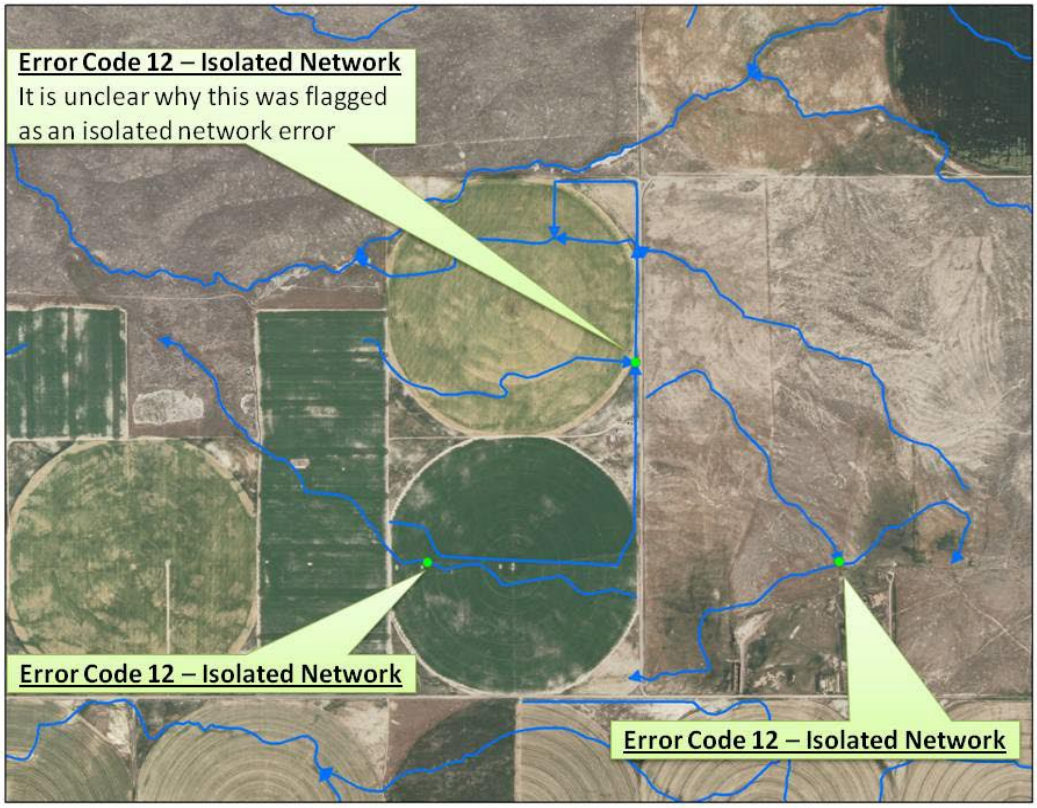


Figure 7 - Error Code 12 Examples

A segment may be identified as participating in an isolated network if it crosses another water feature but does not interact with that water feature, such as a raised aqueduct or siphon. Under these conditions, water is not exchanged into another adjacent canal at the intersection (Figure 8). Another area where a linear feature could be incorrectly identified as an isolated network is where the subwatershed boundary creates an artificial break in the network (see figure 9). In this case, the remaining portion of the network is missing, only because of the subwatershed boundary clip, not because it is actually missing from the network.

Figure 8 - Example of flume/siphon location at the intersection - over/under segments.

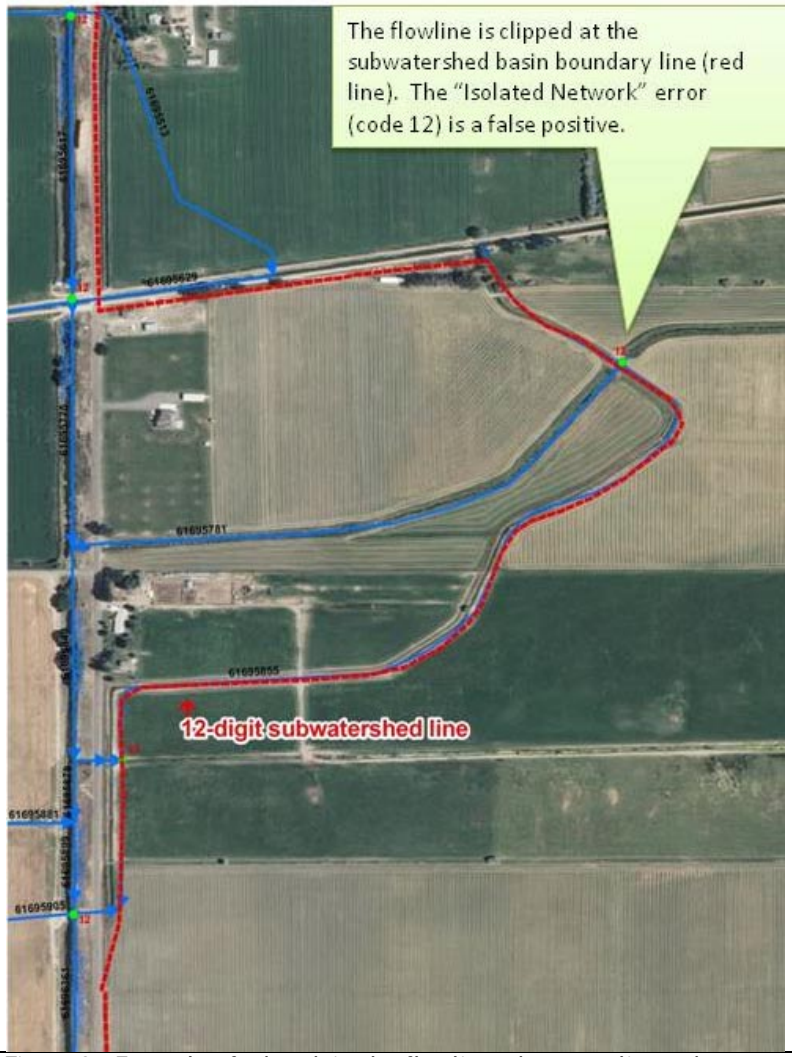


Figure 9 - Example of a break in the flowlines due to a clip at the subwatershed boundary line.

Subwatershed 170402091306 was chosen because extensive work has been done in the Burley Irrigation District to update the canal and lateral network within this subbasin and it is a mix of rural, urban and agricultural use in Idaho. The intent was to run the QA/QC Tool on a subwatershed where IDWR had worked closely with the local stewards and were fairly confident that correct flow direction and connectivity had been identified. This subwatershed produced significantly more errors than the high relief, classic subwatershed unit 170103010502 or the closed basin subwatershed 170402091104. This was not unexpected because this subwatershed contains a complex network of man-made canals and laterals that do not necessarily mimic the natural flow of a classic watershed consisting of streams and tributaries.

Although this subwatershed produced the most errors, the majority of those errors were warning errors (Error Code 12), and have been spot-checked and determined to be correct as-is within the NHD. Additionally, there were three severe error codes; Error Codes 14, 19, and 25.

The first severe error was Error Code 14 – "Flowlines Relate in Flow Table But Do Not Touch". NHD flow tables contain information on the "from" and "to" relationships of the flowline features and keep track of the relationship of flow within the NHD flowline network. When the data is extracted from the National Database, a flow network is built. The flow network can be removed from the database and rebuilt, using the NHD Utilities. If the "fromcomid" and "tocomid" attributes do not contain a common endpoint, the result is error code 14. Upon inspection of this particular error (figure 10), this data has been sent to the local data steward to determine if these flowlines should be connected, i.e., if water is exchanged from one lateral to another (Figure 10) or if water is not exchanged into the adjacent lateral.

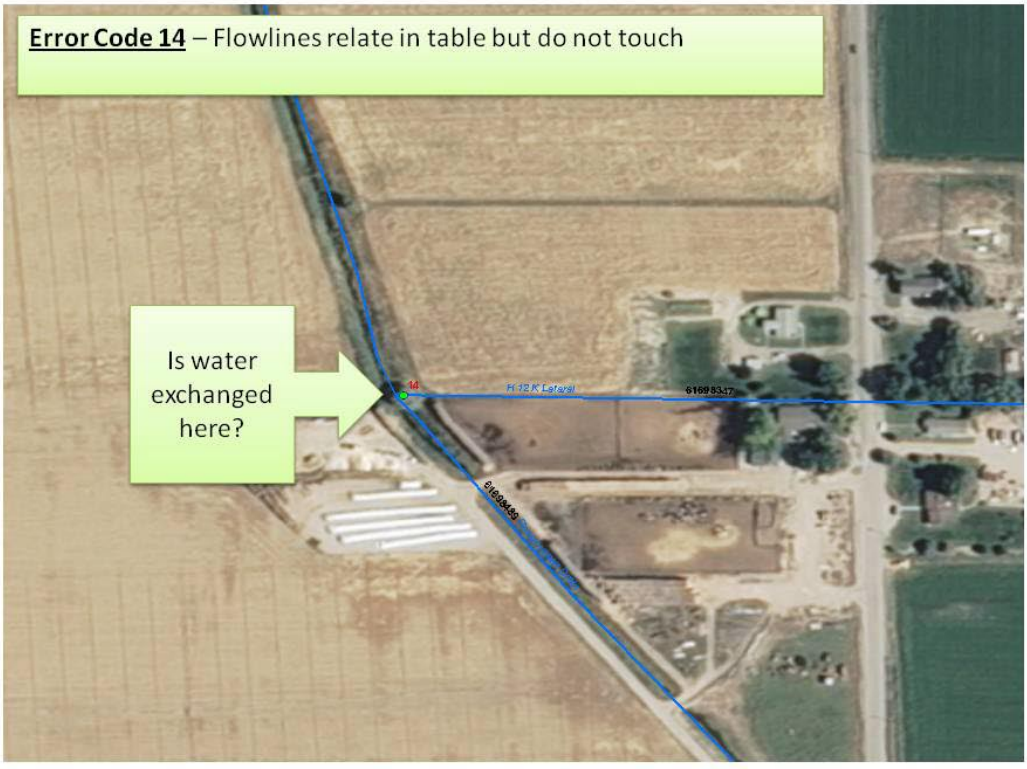


Figure 10 - Error Code 14 - Water exchange needs to be determined here

The second severe error code that was present in this data was Error Code 19 “Nonlinear Reach”. This condition occurs when the merged geometry for all NHDFlowlines with a particular reachcode has other than 2 endpoints (and not 0 endpoints). Upon inspection, this error is a result of one lateral with the same reachcode flowing opposite directions where it intersects with another lateral (Figure 6). This data has been sent to the local data steward for clarification, and further edits may be needed. If edits are necessary, they will be done using a checked out version of the NHD high resolution dataset, and edited using the GeoEdit tools.

Figure 11. Example of a Flowline with opposing flow, same reachcode

The third severe error that was encountered for this subwatershed was Error Code 25 “Nonlinear Named Path”. Similar to error code 19, this happens when the merged geometry for all NHDFlowlines have other than 2 endpoints (and not 0 endpoints), but

the routine checks for the GNIS-id field instead of the reachcode field (Figure 12). For this particular error, it was discovered that a segment was missing from the J-19 lateral. This data has been sent to the local data steward for clarification, and further edits may be needed. If edits are necessary, they will be done using a checked out version of the NHD high resolution dataset, and edited using the GeoEdit tools.

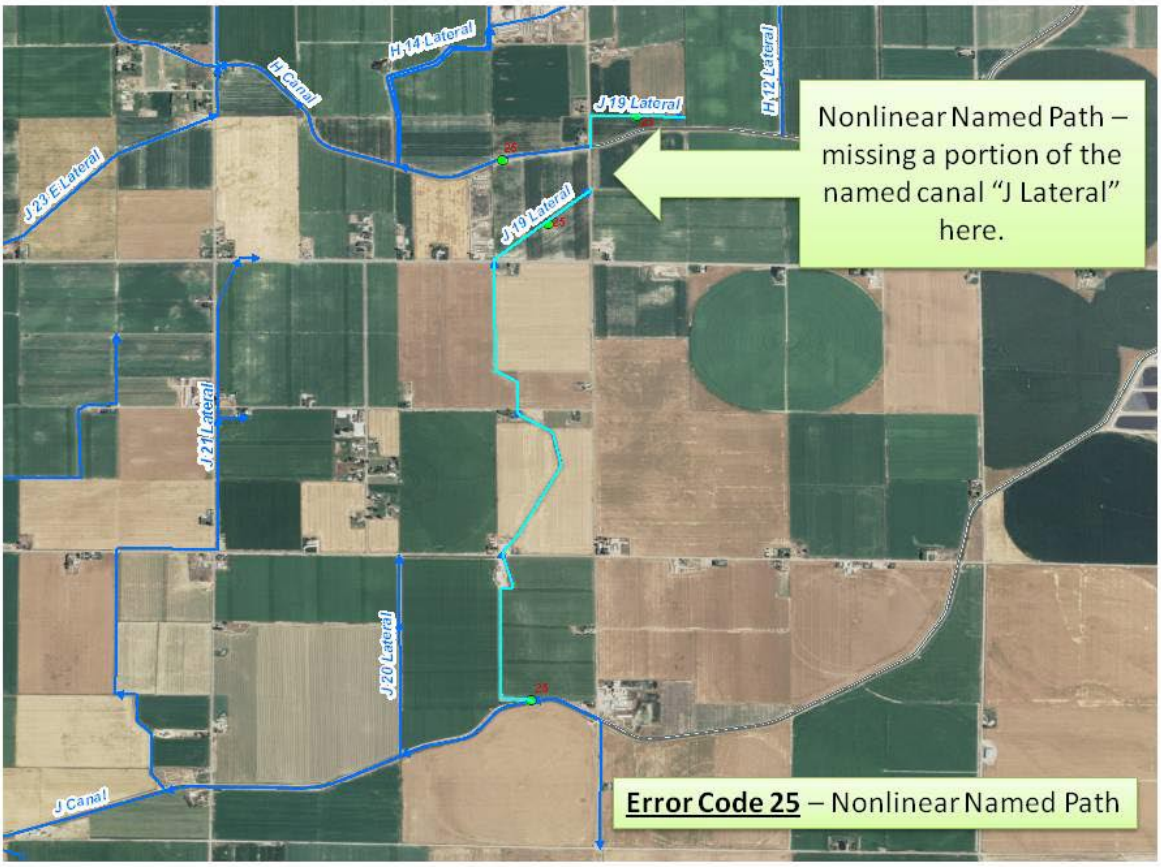


Figure 12. Missing a segment of the J Lateral - Nonlinear Named Path error

A complete list of edits that need to be done to NHDflowlines and NHD Related Tables can be found in Appendix C.

Conclusions:

Issues and Challenges

One major challenge was rebuilding the network and flow tables. The order of the steps taken to rebuild the network affected the number of errors returned by the QA/QC Tool. The steps outlined in Appendix A resulted in the fewest errors returned by the QA/QC Tool.

Additionally, it was determined that a better understanding of the conditions causing the errors to occur is needed before adjustments should be made to the NHDflowlines. It was discovered late in the process that the user could click on each of the errors in the NHDQAQC dialog box to get a short explanation of each of the error conditions in the tool tip box, but a more in-depth description was needed to effectively rule out areas that were questionable. In the cases of the non-severe errors, a better understanding of how to indicate a validated NHDflowline geometry is required. Many of these NHDFlowlines are canal segments, and set to uninitialized. It was unclear if setting the non-networked flowlines to "uninitialized" would help in this process, enabling the user to drop them from the network.

Clarification is also needed to determine which stage of the NHD update process these edits should be incorporated into prior to submission to Horizon Systems Corporation (HSC) for NHDPlus creation. For example, do the edits need to be added via an official "checked out" version of the NHD subbasin, or to a downloaded copy of the subbasin? If a checked-out version is required, do those edits need to be done (using the USGS GeoEdit Tools) and checked back into the USGS before it can be submitted to HSC?

Conclusion

The QA/QC Tool ran successfully on all three subwatersheds selected. The majority of the errors were warning errors. Upon investigation of most warning errors, adjustments to the NHD are not necessary. For those errors that do require edits to the NHD, some clarification may be required from local data steward. It is unclear as to what level any required edits need to go through in order to clear the QA/QC Tool and then be sent on to HSC for NHDPlus processing.

It was determined that the QA/QC Tool are very useful as a stand-alone application that can be used in addition to the NHD Flowcheck utility to locate and investigate areas where the NHDFlowlines may not be properly networked.

Test of QA/QC Tool in ArcGIS 10.0

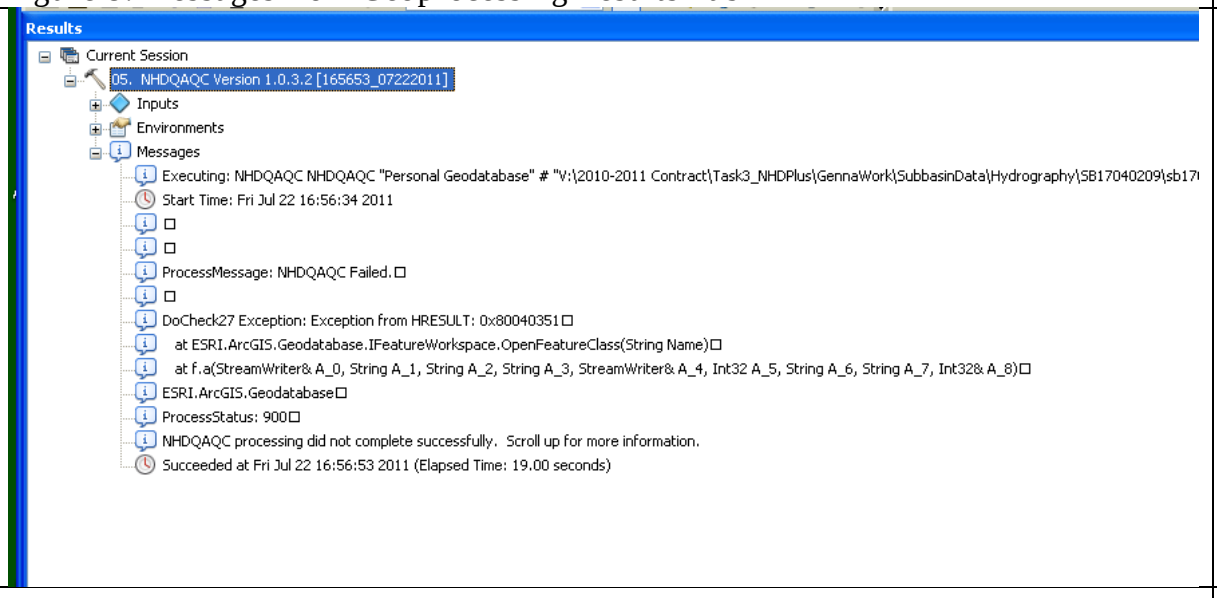
The NHDPlus NHD QAQC Tool Version 1.0.3 (QA\QC Tool) was also tested in ArcGIS 10.0. The tools did appear to run but several errors occurred.

The Toolbox containing the QA\QC Tool loaded into ArcGIS 10.0 without errors. When the tools were clicked on, the input dialog box opened normally. The following inputs were provided:

- Build Name – Pre-populated
- NHD Data Format: Personal Geodatabase
- Input File: Navigate to the .mdb file of a SubWatershed
- Temporary Work Area: D:/Temp
- All Checks, with the exception of Check #16 (Flowing Connections and Elevations Disagree) were populated. Check #16 was permanently grayed-out; this option was not selectable.

The QA\QC Tool was run on all three SubWatersheds. The tool did not error as it ran but in reviewing the Messages, several errors were noted (Figure 8).

Figure 8: Messages from Geoprocessing Results Tab



Additionally, the following errors were found:

1) Unlike in 9.3.1, the QA/QC Tool Results were dropped into a text file in a folder in workspace, not added to the input geodatabase.

2) Upon comparing results, 99% were identical with the results using ArcGIS 9.3.1. The exception was an Error Code 14 result for Hydrologic Unit 170402091306 which was not identified when using ArcGIS 10.0. Identical results for Error Codes 12, 19, and 25 were identified when using either ArcGIS 9.3.1 or ArcGIS 10.0.

Upon looking at the QA/QC Tools installation folder, the 9.3.1 .dll files appear to be packaged with the tool. There were significant changes to many .dll files from 9.3.1 to 10.0. Horizon Systems Corporation has no scheduled plan for upgrading the QA\QC Tools to ArcGIS 10.0[vi]

If 9.3 tools are migrated to Arc10, we have found the following to be useful

- Tools should be written in .net of Java
- ESRI has a page that details all of the steps of moving from 9.3 to 10:
    - http://help.arcgis.com/en/sdk/10.0/arcobjects_net/conceptualhelp/index.html#/Migrating_your_code_from_9_3_to_10/000100000ms1000000/
- If you have desktop ArcObjects software for standalone programs that consume ESRI licenses, you need to explicitly target whether the license is for Desktop, Engine, or Server, and bind to a particular version of ArcGIS.

Acknowledgements:

Thank you to Cindy McKay (HSC) for providing the QA/QC Tool.

Appendix:

A. Steps Followed To Rebuild The Geometric Network
   http://www.idwr.idaho.gov/GeographicInfo/NHD/Projects/PDF/appendix_A_Task3.pdf
B. Draft NHDPlus User Toolbox User Guide
   http://www.idwr.idaho.gov/GeographicInfo/NHD/Projects/PDF/NHDPlus_User_ToolBox_Users_Guide_NHDQAQC_Tool_20100726.pdf
C. List of Edits That Need To Be Done To NHDflowlines And NHD Related Tables
   http://www.idwr.idaho.gov/GeographicInfo/NHD/Projects/PDF/Edits_List_Task3.pdf

End Notes:

[i] NHD Users Guide: http://nhd.usgs.gov/userguide.html
[ii] http://factfinder2.census.gov/faces/nav/jsf/pages/index.xhtml
[iii] http://factfinder2.census.gov/faces/nav/jsf/pages/index.xhtml
[iv] http://factfinder2.census.gov/faces/nav/jsf/pages/index.xhtml

[v] NHDPlus User Toolbox Users Guide
http://www.idwr.idaho.gov/GeographicInfo/NHD/Projects/PDF/NHDPlus_User_ToolBox_Users_Guide_NHDQAQC_Tool_20100726.pdf

[vi] Email Communication between Linda Davis, IDWR and Cindy McKay, HSC on Aug. 23, 2011