ARC HYDRO TOOLS PROCESS PAPER #1

QUALITY CONTROL FOR VECTOR DATA USED IN ARC HYDRO TERRAIN PREPROCESSING



Environmental Systems Research Institute, Inc. (Esri) 380 New York Street Redlands, California 92373-8100 Phone: (909) 793-2853 E-mail: archydro@esri.com Copyright © 2013 Esri All rights reserved. Printed in the United States of America.

The information contained in this document is the exclusive property of Esri. This work is protected under United States copyright law and other international copyright treaties and conventions. No part of this work may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying and recording, or by any information storage or retrieval system, except as expressly permitted in writing by Esri. All requests should be sent to Attention: Contracts and Legal Services Manager, Esri, 380 New York Street, Redlands, CA 92373-8100 USA.

The information contained in this document is subject to change without notice.

Esri, the Esri globe logo, ArcGIS, ArcSDE, ArcCatalog, ArcMap, arcgis.com, esri.com, and @esri.com are trademarks, registered trademarks, or service marks of Esri in the United States, the European Community, or certain other jurisdictions. Other companies and products mentioned herein may be trademarks or registered trademarks of their respective trademark owners.



Quality Control for Vector Data Used in Arc Hydro Terrain Preprocessing

Arc Hydro Tools Process Document #1

Author	Dean Djokic
Date	August, 2013
Initial ArcGIS / Arc Hydro version	10.1 / 10.1.0.181
Comment	First release

Contents

1	Introduction	1
	1.1 Terrain preprocessing and need for clean data	2
	1.2 Types of QC checks	
	1.2.1 Basic stream geometry checks	3
	1.2.2 Basic polygon geometry checks	
	1.2.3 Stream geometry with respect to raster checks	
	1.2.4 Polygon geometry with respect to raster checks	5
	1.2.5 Stream/boundary/water body interaction checks	5
2	QC process operation	6
2	2.1 Introduction	
	2.1.1 Overview of ArcGIS Data Reviewer batch job execution	
	2.2 Prepare initial geodatabase with vector data to be QC'ed	
	2.3 Basic stream QC checks	
	2.3.1 Stream connectivity definition	
	2.3.2 Basic QC processes for streams	
	2.4 Basic QC processes for polygons	
	2.5 Advanced QC processes	
	2.5.1 Stream geometry with respect to raster QC checks	
	2.5.2 Polygon geometry with respect to raster QC checks	
	2.5.3 Stream/boundary/water body interaction QC checks	
	2.6 Loop through QC check and data fixes	
Арј	pendix 1. QC Check Definitions	
	Appendix 1.1 Stream Basic Checks (StreamBasicCheck.rbj)	
	Invalid Geometry Check Multipart Line Check	
	Polyline Closes on Self Check Duplicate Geometry Check	
	Short Line Check	-
	Pseudo Nodes Check	
	Outlets Check Flow Split Check	
	Appendix 1.2 Stream Area Checks (StreamAreaCheck.rbj)	
	Breaching Processing Area Check	
	Stream Boundary Proximity Check Vertex Snap Check	21

Quality Control for Vector Data Used in Arc Hydro Terrain Preprocessing

Appendix 1.3 Stream Advanced Checks (StreamAdvancedCheck.rbj)	22
Watershed Boundary Proximity Check	22
Appendix 1.4 Polygon Basic Checks	23
Invalid Geometry Check Multipart Geometry Check	23
Duplicate Geometry Check	24
Polygon Overlap Check	24
Polygon Touches Each Other Check	
Appendix 1.5 Polygon Area Checks	
Polygon Not Within Processing Area Check	
Appendix 1.6 Waterbody Advanced Checks (WaterbodyAdvancedCheck.rbj)	27
Non-Draining Waterbody Check	
Appendix 1.7 Watershed Advanced Checks (WatershedAdvancedCheck.rbj)	
Non-draining Watershed Check	
Appendix 2. A worked out example	
Initial data organization	29
Performing operations	
Initial project and database setup	
Running basic checks	
Running other checks	

List of Figures

Figure 1.	Create Initial QC GDB function user interface	9
Figure 2.	Stream Connectivity Parameters function user interface	10
Figure 3.	Stream QC function user interface.	12
Figure 4.	Streams Near Processing Area Boundary user interface	12
Figure 5.	Polygon To Line function user interface	14
Figure 6.	Non Draining Watersheds user interface.	15
Figure 7.	Streams Near Watershed Boundary user interface.	15
Figure 8.	Cache area of interest. Black are the HUC-8 boundaries	29

List of Tables

Table 1.	Basic stream geometry checks	4
Table 2.	Basic polygon geometry checks	4
Table 3.	Stream geometry with respect to raster checks	5
Table 4.	Polygon geometry with respect to raster checks	5
Table 5.	Stream/boundary/water body interaction checks	6
Table 6.	Advanced stream check vs. QCVALUE1	.3

Quality Control for Vector Data Used in Arc Hydro Terrain Preprocessing

1 Introduction

This document describes quality control (QC) processes for vector data used in Arc Hydro (AH) terrain preprocessing. They will also work when only vector preprocessing is performed (in which case only a subset of the QC processes will have to be performed).

The best way to fix data is not to have data problems in the first place. Using properly designed geodatabase and leveraging geometric networks and topology when the data are developed in house will go long ways in producing clean data from the get go. This document will explain why certain QC checks are being made and problems that can arise with Arc Hydro implementation if the data do not meet certain criteria. These explanations should help the reader in understanding the basic data requirements and their impact on Arc Hydro database development.

There are several important points to make:

- The QC processes presented here do NOT fix the data; they just identify known problems with the data that need to be reviewed and fixed before proceeding to Arc Hydro preprocessing. The users will have to perform data cleanup using standard ArcGIS data editing tools.
- Without clean input data, use of Arc Hydro tools will not generate correct and expected results and that effort should not be undertaken.
- Some of the QC processes are relative to the DEM resolution and parameters used in the preprocessing tools, so an iterative approach might have to be taken in the QC operations. The QC operations that are sensitive in this context are highlighted in the document.
- Not all the QC operations need to be performed. That will depend on the available data, terrain preprocessing workflows, and preprocessing tools to be used.
- The QC tools check for physical problems, not logical ones. That is still user's responsibility.
- Additional QC checks might be possible and needed and users are strongly encouraged to develop and capture them in more complex QC workflows that are specific to their data and data collection processes.

The Arc Hydro team is interested in learning about other possible QC checks, and if they are generic enough we will try to include them in the future releases of the tools. Please let us know what additional checks you find useful in your practice.

QC checks are implemented through a combination of Arc Hydro tools and processes and ArcGIS Data Reviewer extension (DR) checks. DR provides extensive and easy to configure data checking and result navigation capabilities. Use of DR also enables easy extension of the initial set of checks and was implemented with that particular approach in mind – the checks will grow over time and can be configured to meet specific user's data management and editing workflows. For those new to DR, there is a lot of online training and support documentation that can provide a quick start in its use.

- Using ArcGIS Data Reviewer to Assess Data Quality (ArcGIS Version 10.0): <u>http://training.esri.com/gateway/index.cfm?fa=catalog.webCourseDetail&courseid=2099</u>
- Introduction to GIS Data Reviewer (ArcGIS Version 9.3): http://training.esri.com/gateway/index.cfm?fa=catalog.webCourseDetail&courseid=1461
- ArcGIS Data Reviewer: An Introduction (video): <u>http://video.arcgis.com/watch/1698/arcgis-</u> <u>data-reviewer-an-introduction</u>

For those who want to understand and extend the QC processes embedded in the DR batch checks, they are presented in Appendix 1. The specific AH QC tools have been developed as python scripts and as such can be opened, reviewed, and modified if necessary. They are fairly simple scripts that should be easy to understand and extend. All AH QC tools and DR batch jobs are installed when AH is installed.

First chapter presents the overview of the QC checks. Second chapter presents the QC process in detail and presents individual steps in the process. The appendices present the configuration for each DR task/check, functionality of custom AH QC tools, and a worked out example.

1.1 Terrain preprocessing and need for clean data

Arc Hydro preprocessing establishes an Arc Hydro dataset that usually integrates raster and vector data into a single operational data system where Arc Hydro tools operate in fully optimized mode. Complexity of preprocessing depends on the available data and terrain morphology. In the simplest process, the only input into the preprocessing is the DEM and besides having a valid raster of the DEM, no other data are needed. For more complex processes where the user wants to impose known drainage structures into the DEM, additional vector data representing these structures are needed. The following vector data might be needed:

- Stream polylines representing known flow patterns.
- Sinks polygons representing known depressions collecting water.
- Waterbody polygons representing known areas collecting water that is drained out by a stream (as opposed to sinks that are not drained).
- Ridge polygons or lines representing know drainage divides.

These vector data will be inputs into the Arc Hydro tools for terrain preprocessing. For the tools to operate properly and produce expected results, the vector data have to have certain quality. Lack of that quality will result either in tools not working or producing unexpected results. Performing the quality control (QC) steps presented in this document will identify key input data problems that will make implementation of Arc Hydro preprocessing unpredictable. Data problems identified through the use of the presented QC check should be fixed before proceeding with Arc Hydro preprocessing.

1.2 Types of QC checks

There are two key types of checks:

- Checks on the quality of geometry of the data within a single feature class. These could be considered as "physical" checks. They include basic checks such as checking for null geometry as well as more complex checks such as checking for pseudo-nodes and end edges.
- 2. Checks on the relationship of data across multiple feature classes. These could be considered as "process" checks.

In both cases, understanding the impact of identifying a "problem" identified through the check is important, as often, the identified "problem" is not a problem at all but a desired and necessary feature. DR identifies the potential problems, but the user ultimately decides what action to take, whether it is to fix a problem or keep the feature as is.

The following sections list physical and process checks currently implemented in the Arc Hydro QC tools. It is expected that these will grow over time. The current checks focus on four feature classes that are most often used in Arc Hydro terrain preprocessing. Using DR capabilities, these checks can be quickly adapted for other input data if need be.

#	Check	AH issue/recommended fix	
	Physical check		
1.1	Zero geometry	Unnecessary features to process. Remove offending	
		features.	
1.2	Multipart	Unnecessary complexity, connectivity confusion. Split	
		into individual features and/or merge into a single	
		geometry.	
1.3	Coincidence	Connectivity, order confusion. Remove all but one	
		feature at each location.	
1.4	Self-intersection including circular line	Connectivity confusion. Clean up geometry.	
1.5	Short lines (not always a problem)	Connectivity confusion (lines can collapse when	
		burning into DEM). Drainage line and catchment size	
		issues. Merge into neighboring lines if possible.	
1.6	Pseudo node (not always a problem)	Drainage line and catchment size issues. Merge into	
		neighboring lines if possible.	
1.7	 Outlets (not always a problem) 	Can produce sinks.	
	 Overshoots (intersects) 	Flow direction, drainage line issues. Clean-up	
	 Non-end node connections 	geometry.	
1.8	Flow splits (not always a problem)	Flow direction issues. If really existing require specific	
		processing workflow. Remove if necessary.	
Process check			
1.9	Directionality (digitized direction	Flow direction issue. Fix line orientation in	
	should be in flow direction)	downstream direction.	
1.10	Braids (not always a problem)	Flow direction, short drainage line issues. Remove if	
		necessary.	

1.2.1 Basic stream geometry checks



1.11	Looping sequence (not always a	Flow direction, short drainage line issues. Remove if
	problem)	necessary.

 Table 1. Basic stream geometry checks

If ridges are being applied as linear features (as opposed to implementing them as watersheds – polygonal features discussed in the next section), checks 1-4 above should be performed on them.

1.2.2 Basic polygon geometry checks

These checks apply to polygon type feature classes used in the Arc Hydro preprocessing. This includes:

- Sinks
- Water bodies (non-sinks)
- Watersheds

#	Check	AH issue/recommended fix
	P	hysical check
2.1	Zero geometry	Unnecessary features to process. Remove offending features.
2.2	Multipart	Unnecessary complexity, ID and order confusion. Split into individual features and/or merge into a single geometry.
2.3	Coincidence	ID and order confusion. Remove all but one feature at each location.
2.4	Overlaps/gaps	Precedence confusion, incompleteness. Clean up geometry.
2.5	Touching (not always a problem)	Unnecessary complexity, incompleteness. Clean up geometry.

 Table 2. Basic polygon geometry checks

1.2.3 Stream geometry with respect to raster checks

These checks should be applied after all the issues identified in the basic checks have been resolved. It is important to note that results of these checks are a function of DEM cell size. Same vector geometry might have different issues (or none) depending on the DEM resolution.

#	Check	AH issue/recommended fix
	Physical check	
3.1	Vertices within a single raster cell that	Flow direction issues. Clean up geometry.
	form a loop	
3.2	Segments smaller than cell size (from	Lost drainage lines. Adjust the segment length if
	and to nodes within the same cell)	possible or just note the mismatch between the

		resolution of vector and raster data.
3.3	Parallel lines within a single cell	Lost drainage lines. Adjust distance between the lines
	sequence	if possible or just note the mismatch between the
		resolution of vector and raster data.
	Vertices from multiple lines occupying	Lost drainage lines and/or moved confluences. Adjust
	the same grid cell (confluences)	lines to enter a confluence from a single and different
		neighboring cell.
3.4	Identify lines that are extending over	Problem with "burning" streams. End streams that are
	the DEM's "edge" (usually a good	fully contained within the processing area will generate
	event)	internal drainage areas, which is valid only in fully
		deranged terrain morphology. Otherwise at least one
		end stream has to extend over the area (breach line).
3.5	Streams close to the edge of the DEM	Problem with "burning" streams. Streams within the
	extent	AGREE buffer distance from the full extent boundary
		might "breach" the boundary. Move the streams away
		from the boundary or use a narrower AGREE buffer.

Table 3. Stream geometry with respect to raster checks

1.2.4 Polygon geometry with respect to raster checks

#	Check	AH issue/recommended fix
Physical check		
4.1	All polygons have to be fully contained	Unexpected results. Edit/remove geometry to fit
	within the DEM extent	within the DEM extent.

Table 4. Polygon geometry with respect to raster checks

1.2.5 Stream/boundary/water body interaction checks

#	Check	AH issue/recommended fix
	Ph	ysical check
5.1	Water body (not sinks) polys that do not have streams going through them	These will not be used in "adjust flow direction in lakes" function. Extend the streams to "flow" through the water bodies.
5.2	Find if there are watershed boundary polygons (walls) that do not include either a sink poly or have a stream breaching its boundary (not always a problem)	Watersheds without a stream breach or a sink will generate an internal drainage area. Edit geometry if necessary, or just acknowledge internal drainage areas.
5.3	Streams close to watershed boundary (not always a problem)	Problem with "burning" streams. Streams within the AGREE buffer distance from the watershed boundary



(not outlet) might "breach" the boundary. Move the streams away from the boundary or use a narrower
AGREE buffer.

 Table 5.
 Stream/boundary/water body interaction checks

2 QC process operation

2.1 Introduction

Quality control for vector data used in the Arc Hydro preprocessing is a process that involves a sequence of steps. It is a combination of data preparation and tool operations and involves knowledge of how the data will be used in preprocessing tools. The process outline is:

- 1. Prepare the initial geodatabase with vector data to be QC'ed.
- 2. Run basic QC checks (checks on individual layers).
- 3. Fix identified basic data problems.
- 4. Run steps 2-3 until all the data have been fixed.
- 5. Run advanced QC checks (checks across multiple layers).
- 6. Fix identified advanced data problems.
- 7. Run steps 5-6 until all the data have been fixed.

In case that the user has several other layers that are used in data development and maintenance workflows but are not used in Arc Hydro terrain preprocessing, a larger data processing workflow needs to be developed to integrate the presented QC techniques into the overall data processing methodology.

2.1.1 Overview of ArcGIS Data Reviewer batch job execution

All DR batch job runs are performed in the following way (ArcGIS Data Reviewer extension has to be installed and enabled on the computer where the checks are being performed):

1. Make sure DR toolbar is visible.



2. Start an existing or new DR session (click on the "Reviewer Session Manager" tool and open a session). Whether to start a new or an existing session and whether to store the records about anomalies in the data in the same or different geodatabase from where the data are stored depends on user's overall data QC and editing strategy (e.g. if each layer is checked and fixed independently, generate a new session; if all are checked at once and then fixed, have all batch jobs reference the same session).



Reviewer Work C:\CurrWork	\ArcHydro\Tests\HealthCheck\CacheRiverHea Browse
User Name	
deand	
Session ID	Session 3 : Session 3 Vew
Name	Session 3
Reviewer Data	iset Version
	*

 Run the batch job by clicking on the "Batch Validate". Select one of several available AH DR batch job definition files (in the AH tools installation directory, "rbjfiles" folder – usually at C:\Program Files (x86)\ESRI\WaterUtils\ArcHydro\rbjfiles). Click on "Run" to run the batch validation.

n Batch Validate		
Features to Validate Selection Sets Current Extent Full Database Changed Features Only Batch Job Add from File Add from Product Library Remove Validate All Workspace Batch Job was successfully updated to current workspace.	Batch Jobs Default (Click 'Validate' to update) StreamChecks Volume Closes on Self Volume Closes on Self Volume Short Line Volume Closes	
Validation Status: Not validated	Run	Cancel



4. Upon completion, the validation process will provide a message about results being written to the reviewer table.

Table Writer	23
3336 Record(s) Written to Reviewer Table	
ОК	

5. Review the results of the validation process and fix any problems (click on "Reviewer Table" to see all the identified issues and navigate to them).

2.2 Prepare initial geodatabase with vector data to be QC'ed

Arc Hydro data form an integrated analytical system. Both raster and vector data directly used by the Arc Hydro preprocessing tools have to be in the same spatial reference and vector data in the same geodatabase/feature dataset. The extent of the DEM dictates the area of interest and defines the extent of the vector data needed to support the development of the Arc Hydro preprocessed data system.

To automate the QC processes as much as possible, a fixed QC geodatabase structure has been defined and should be used to leverage the QC tools. Use of QC processes on non-structured data is possible but will require more manual work and thorough understanding of the implemented DR QC checks.

The following steps are performed to create the preformatted QC database. All QC jobs and data cleaning should be performed on that database.

- 1. Create an empty (file) geodatabase.
- 2. Create a feature dataset called "Layers" with spatial reference imported from the DEM that will be used in preprocessing. DEM should reside outside of the geodatabase and should cover at least the full extent of the area that will be preprocessed (or more).
- Import vector layers to be QC'ed into the feature dataset created in #2. Which layers need to be imported depends on the available data and terrain preprocessing workflows to be executed. Some or all can be added. They will be given the following fixed names. These names facilitate easy application of batch QC jobs.
 - a. "Stream" line feature class with streams to be used for "DEM Reconditioning" and "Create Drainage Line Structures" functions.
 - b. "Sink" polygon feature class defining sinks to be used in the "Level DEM" and "Create Sink Structures" functions.
 - "Waterbody" polygon feature class of "non-sinks" to be used in the "Adjust Flow Direction in Lakes".



- d. "Watershed" internal boundary polygon feature class to be used in "Build Walls" function.
- 4. Create "ProcessingArea" polygon feature class that defines the polygonal extent of DEM to be processed.
- 5. Create "ProcessingAreaOutline" polygon feature class that defines the extent of DEM to be processed as a linear feature set.

These steps are implemented in the "Create Initial QC GDB" AH QC tool (in "Vector QC" toolset within AH "Terrain Preprocessing" toolset).

Create Initial QC GDB	- 0	X	
Target Directory			^
		2	
Target GDB Name			
Reference DEM			
	-	2	
Stream FC (optional)		2	
J Sink FC (optional)	<u> </u>		
	•	2	
Waterbody FC (optional)		_	
Wetersted 50 (on FreeD)	<u> </u>	B	
Watershed FC (optional)	-	r i i i i i i i i i i i i i i i i i i i	
			-
OK Cancel Environments	Show He	lp >>	

Figure 1. Create Initial QC GDB function user interface.

Additional linear and polygonal features can be added to the QC geodatabase if necessary. At this point, no QC batch jobs are established for these additional layers, but can be developed by the user based on the jobs developed for the first four layers.

2.3 Basic stream QC checks

This is the process that is performed on stream feature class. It involves two steps in this particular order:

- 1. Run stream connectivity AH QC function
- 2. Run DR basic stream QC check batch job



These two steps should be performed each time after stream data have been edited.

2.3.1 Stream connectivity definition

These operations (Appendix 2.2) prepare the stream connectivity data for the basic stream QC tasks. They should be executed after any edits to the stream layer are made and before executing the DR QC checks. These steps are implemented in the "Stream Connectivity Parameters" AH QC function (in "Vector QC" toolset within AH "Terrain Preprocessing" toolset).

📑 Stream Connectivity Parameters		
Stream		
Stream		- 2
		T
	OK Cancel Environments	Show Help >>

Figure 2. Stream Connectivity Parameters function user interface.

This function will create HydroID, From_Node, To_Node, NextDownID, and FlowSplitCnt fields in the Stream feature class as well as creating the Stream_FS table and Stream_AllEndPnt feature class.

2.3.2 Basic QC processes for streams

This section covers checks presented in Table 1. Checks 1.1-1.8 are performed through DR batch job while checks 1.9-1.11 are processes performed using standard ArcGIS tools. QC function "Stream Connectivity Parameters" should have been run before DR QC checks.

The DR batch job configuration file for basic stream QC checks is called "StreamBasicCheck.rbj" (configuration is provided in Appendix 1).

Check #1.9 deals with directionality of the stream lines. This is important for flow direction enforcement. This is a redundant check as if all fixes based on check #1.7 have been made, all stream segments will be oriented correctly in the downstream direction and this check will not identify any further issues. The check is performed using the following general procedure using standard ArcGIS tools (no specific AH QC tools are provided for this purpose):

- 1. Perform basic checks 1.1-1.8 and fix anything that needs to be fixed.
- 2. Build geometric network out of streams.
- 3. Use Arc hydro tool to assign flow direction in the network based on digitized direction.
- 4. For each true outlet edge, place a network flag at the downstream node(s).
- 5. Set tracing results to generate selected set.

6. Trace upstream - this will identify all edges that flow into the outlet(s). If this is the complete set of edges, then all the edges are oriented "downstream" and flow into the outlets. If not all edges are selected, it means that there is a "miss-oriented" edge that needs to be inspected and potentially "flipped".

Check #1.10 deals with braids in the system. Braids are not necessarily a data problem, but often present more detail than is needed. Braids will be identified through check #1.8 (flow splits). When inspecting identified flow splits, make a decision whether the braids will be represented explicitly or not. If not, remove the "minor" braid stream elements and just keep the main stream. If braids are to be modeled, make sure you use appropriate terrain preprocessing workflow (you cannot use synthetic streams as that workflow does not allow for flow splits in auto generated drainage lines).

Check #1.11 deals with loops in the system. As with braids, this is not necessarily a data problem and have to be treated in the same manner, that is, remove one ("minor") side of the loop if loop is not desired, or keep them both if they are explicitly modeled.

As fixes to the data are made, the QC process will have to be repeated until all the fixes are completed. If new lines are added or deleted, or directionality of the lines is changed (flipped), before running the DR batch job, you have to rerun "Stream Connectivity Parameters" tool to recalculate the stream connectivity attributes.

2.4 Basic QC processes for polygons

This covers checks presented in Table 2. These checks are performed through DR batch jobs. No other steps need to be performed before running the basic polygon checks. The following batch jobs should be used for each appropriate type of polygon geometry. Each of these batch jobs performs the same set of checks (configuration is provided in Appendix 1).

- Sinks "SinkBasicCheck.rbj"
- Water bodies "WaterbodyBasicCheck.rbj"
- Watershed "WatershedBasicCheck.rbj"

It is best to fix any identified errors before proceeding to the advanced checks.

2.5 Advanced QC processes

The advanced QC processed deal with interaction of individual layers, for example interaction between the stream polylines and water body polygons. They should be performed after the basic checks and cleanup of data resulting from the basic checks have been completed.



2.5.1 Stream geometry with respect to raster QC checks

This covers checks presented in Table 3. These checks are performed through DR batch job ("StreamAreaCheck.rbj" - configuration is provided in Appendix 1). Before running this batch job, you first have to run:

1. "Stream QC" Arc Hydro QC tool.

🐔 Stream QC		
Input Raw DEM		^
dem	- 🖻	
Input Stream		
Stream	- 🖻	
Output Line		
C:\CurrWork\ArcHydro\Tests\HealthCheck\ttt.gdb\Layers\StreamAdjusted		
		Ŧ
OK Cancel Environments	Show Help >>	

Figure 3. Stream QC function user interface.

2. "Streams Near Processing Area Boundary" Arc Hydro QC tool (Appendix 2.3).

눡 Streams Near Processing Area Boundary		
Stream		×
Stream		▼ 🛃
ProcessingAreaOutline		
ProcessingAreaOutline		⊻ 🖻
Search Distance	1000	Feet -
	1000	
	OK Cancel Envi	ronments Show Help >>

Figure 4. Streams Near Processing Area Boundary user interface.

The "Stream QC" Arc Hydro QC tool will populate attribute "QCVALUE" in the input stream feature class defining the type of the QC issues for checks #3.1-3.3. This value is then used in the DR batch job. The result of the Stream QC function is also a "jaggedized" version of the input stream line feature class (default name of that feature class is "StreamAdjusted"). This represents stream lines where vertices are snapped on the center of the DEM cell that the vertex is in. Any potential geometry issues due to the cell

snapping will be documented in the QCVALUE field. The following values in the QCVALUE field correspond to the checks in table 3.

Check #	QCVALUE
3.1	1
3.2	2
3.3	4

Table 6.Advanced stream check vs.QCVALUE

QCVALUE can contain additional values besides 1, 2, and 4. These values are the sums of individual values indicating that more than one issue is present for that feature. For example, QCVALUE = 7 for a feature indicates issues 1, 2, and 3 (= 1 + 2 + 4) are present.

The DR batch task "Vertex Snap" will identify all features that have issues 3.1-3.3 in a single batch job task. By looking at the QCVALUE field, you can do further refinement of the issues. If desired, a single QC task for each of the issues can be established if that will facilitate data editing once the issues have been identified.

When evaluating results of check #3.4, there are two cases. One case is when the intersection of the stream line with the boundary is with a line that is "flowing" into the processing area. That can cause problems as this often indicates that after the streams are "burned" into the DEM, these lines will introduce additional area into the system (this of course might not be an issue if there is no DEM outside of the processing area polygon). The second case is when the intersection is with the line that is "flowing" out of the processing area. This is a desired situation for dendritic and combined terrains, but not for completely deranged terrains.

Check #3.5 is a general case of check #3.4. The "Streams Near Processing Area Boundary" Arc Hydro QC tool will populate attribute "NearBnd" in the input stream feature class defining whether the stream is within specified distance from the processing area boundary (=1 if it is and "Null" if it is not). The specified buffer distance should reflect the buffer distance that will be used in the "DEM Reconditioning" AH terrain preprocessing function (it should be the same or slightly bigger). When the streams are close to the boundary (whether external or internal), they can breach that boundary during the DEM reconditioning (buffer) process. If the streams are outlet streams, that is usually a good situation, while if they are not, then these streams need to be fixed.

2.5.2 Polygon geometry with respect to raster QC checks

This covers checks presented in Table 4. These checks are performed through DR batch job. The following batch jobs should be used for each appropriate type of polygon geometry. Each of these batch jobs performs the same check (configuration is provided in Appendix 1).

- Sinks "SinkAreaCheck.rbj"
- Water bodies "WaterbodyAreaCheck.rbj"
- Watershed "WatershedAreaCheck.rbj"

No specific functions need to be run before performing these checks (so after edits to participating features, they can be immediately reapplied).

If the "Watershed" feature class is being used, after all the QC and cleaning is completed on it, the line representation of this feature class needs to be created for later QC tasks. Standard gp function "Polygon To Line" should be used. The resulting line feature class must be named "WatershedOutline" and has to reside in the same feature dataset where the input "Watershed" feature class is.

🔨 Polygon To Line		2	
Input Features			^
Watershed	-	2	
Output Feature Class			
C:\CurrWork\ArcHydro\Tests\HealthCheck\CacheRiverHealthCheckT1.gdb\Layers\WatershedOu	utline	2	
dentify and store polygon neighboring information (optional)			
			-
OK Cancel Environments	Show He	elp >>	

Figure 5. Polygon To Line function user interface.

2.5.3 Stream/boundary/water body interaction QC checks

This covers checks presented in Table 5. These checks are performed through several DR batch jobs (configuration is provided in Appendix 1).

- "WaterbodyAdvancedCheck.rbj" performs check #5.1.
- "WatershedAdvancedCheck.rbj" performs check #5.2.
- "StreamAdvancedCheck.rbj" performs check #5.3.

Check #5.1 does not require any specific functions to be run before performing it. It will identify water bodies that do not have a stream running through them. Such water bodies will be ignored when performing "adjust flow direction in lakes" function. User has three options:

- 1. Ignore these issues and keep those polygons in the feature class (they will just be ignored)
- 2. Remove the "offending" polygons it results in a "cleaner" dataset (but the result on the preprocessed DEM will be the same as for option #1)
- 3. Extend the streams to "drain" (intersect) the offending polygons. This will then include these polygons in the flow direction adjustment.

Check #5.2 requires that "Non Draining Watersheds" Arc Hydro QC tool is run first. This tool will populate attribute "NonDraining" in the input watershed feature class defining whether the watershed is not drained (intersected by) neither by a stream nor a sink (=1 if it is not drained and "Null" if it is). Such watersheds might indicate a problem with any of participating datasets (watershed, stream, sink) and such cases need to be reviewed and problem data fixed. By not having a "draining" feature within a watershed, streage results might happen depending on the type of terrain preprocessing being used as these

watersheds represent internal draining areas, and if they are truly such, a deranged processing needs to be used – in which case having a sink within that watershed is desired to explicitly direct the flow into it.

I Non Draining Watersheds			
Input Watershed			^
Watershed		- 🖻	
Input Stream (optional)			
Stream		- 🖻	
Input Sink (optional)			
Sink		- 2	-
	OK Cancel Environments SI	how Help >>	

Figure 6. Non Draining Watersheds user interface.

Check #5.3 requires that "Streams Near Watershed Boundary" Arc Hydro QC tool is run first. This tool will populate attribute "NearWsh" in the input stream feature class defining whether the stream is within specified distance from the watershed boundary (=1 if it is and "Null" if it is not). The specified buffer distance should reflect the buffer distance that will be used in the "DEM Reconditioning" AH terrain preprocessing function. When the streams are close to the boundary, they can breach that boundary during the DEM reconditioning (buffer) process. If the streams are watershed outlet streams, that is usually a good situation, while if they are not, then these streams need to be fixed. This function requires a "WatershedOutline" polyline feature class to exist (section 2.5.2).

Stream	A
Stream	_
WatershedOutline	
WatershedOutline	I 🖆
Search Distance	1000 Feet -
ОК	Cancel Environments Show Help >>

Figure 7. Streams Near Watershed Boundary user interface.

2.6 Loop through QC check and data fixes

After a set of checks is completed and data are fixed, the checks should be run again until no fixes are needed. Remember that some of the checks are dependent on the DEM that the vector data are being applied to. The process should always be such that the vector data are first cleaned on their own and only then they are checked in the context of the DEM and other data that they will be used in conjunction with. This ensures that the initial layer cleaning is not influenced by the later processing as that same layer might be used for other purposes or processing scales.

After the basic checks and data cleanup resulting from them have been completed and before performing advanced checks, it is advised to make a backup of all vector layers. Since advanced checks are related to interaction of different data layers, they might affect the individual layers in different ways that are not relevant to the dataset by itself. Having the backup will allow the user to go back to the original unaffected dataset if some of other vector layers in advanced checks are changed, or the data are applied to a different resolution DEM.

Appendix 1. QC Check Definitions

This appendix presents definitions of each DR check.

Checks	Batch job file	Batch job tasks
1.1 - 1.8	StreamBasicCheck.rbj	Stream Checks Invalid Geometry Multipart Line Polyline Closes on Self Duplicate Geometry Short Line Pseudo Node Outlet Flow Split
3.1 – 3.5	StreamAreaCheck.rbj	 Stream Area Checks Breaching Processing Area Boundary Proximity Vertex Sanp
5.3	StreamAdvancedCheck.rbj	□ Stream Advanced Checks Watershed Boundary Proximity
2.1 – 2.5	SinkBasicCheck.rbj	 Sink Checks Invalid Geometry Multipart Duplicate Geometry Overlap Touches Each Other
4.1	SinkAreaCheck.rbj	 Sink Area Checks Not Within Processing Area
2.1 – 2.5	WaterbodyBasicCheck.rbj	Waterbody Checks Invalid Geometry Multipart Upplicate Geometry Verlap Touches Each Other
4.1	WaterbodyAreaCheck.rbj	□ Waterbody Area Checks Not Within Processing Area
5.1	WaterbodyAdvancedCheck.rbj	Waterbody Advanced Checks Non-Draining Waterbody
2.1 – 2.5	WatershedBasicCheck.rbj	Watershed Checks Invalid Geometry Multipart Duplicate Geometry Overlap Touches Each Other
4.1	WatershedAreaCheck.rbj	Watershed Area Checks Not Within Processing Area
5.2	WatershedAdvancedCheck.rbj	 Watershed Advanced Checks Non-Draining Watershed



Appendix 1.1 Stream Basic Checks (StreamBasicCheck.rbj)

Invalid Geometry Check

Invalid Geometry Invalid Geometry Feature Class Feature Class/Subtype Stream - CacheRiverHealthCheckT1.gdb Always Run on Full Database Where Clause SQL Returns features with invalid geometry. This includes rull geometries, empty envelopes, and non-simple geometries. Notes Severity	Check Title	Check Descript	tion	
adure Class (Subtype Stream - CacheRiverHealthCheckT1.gdb) Always Run on Full Database here Clause SQL Retures with invalid geometry, eviewer Remarks Notes averity 1 - High	Invalid Geometry	OBJECTID	SHAPE	FCSUBTYPE
Stream - CacheRiverHealthCheckT1gdb Always Run on Full Database 3 here Clause SQL Returns features with invalid geometry, this includes null geometry, eviewer Remarks Notes 1 - High		1 1	POLYLINE	
Always Run on Full Database 3 NULL A0040- A040- BRIDGE LINE Arere Clause SQL Returns features with invalid geometry. This includes null geometries, features with empty envelopes, and non-simple geometries. Notes 1 - High			POLYLINE	
SQL Returns features with invalid geometry. This includes null geometries, entry geometries, features with empty envelopes, and non-simple geometries. Notes 1 - High			NULL	
		SQL This includes nu geometries, feat	ull geomet tures with	ries, empty empty

ultipart Line Check Propert	es	Check Description		
Check Title		Check Description		
Multipart Line		Number of Parts > 1		
Feature Class		Number of Parts > 1		
Feature Class/Subtype				
Stream - CacheRiverHe	althCheckT1.gdb			
Always Run on Full Data	base			
Where Clause		-		
	SQL	Returns polyline features with more than one part.		
Reviewer Remarks				
Notes				
	1 - High			
Severity 1	5 - Low			
	OK Cancel			

Multipart Line Check

Polyline Closes on Self Check

Polyline or Path Closes on Self Check Properties	And in case of the second second	Duplicate Geometry Check Properties	And the state of the
Check Title Polyline Closes on Self Feature Class Feature Class/Subtype Stream - CacheRiverHealthCheckT1.gdb Always Run on Full Database Where Clause SQL Error Type Closed Closed Closed Reviewer Remarks Notes 1 - High Severity 1	Check Description	Check Title Duplicate Geometry Feature Class 1 Feature Class/Subtype Stream - CacheRiverHealthCheckT1.gdb	Check Description
ОК	Cancel	ОК	Cancel

Duplicate Geometry Check



Short Line Check

A 20m length is used as an example. Other lengths can be used as well (but will require rule change).

Evaluate Polyline Length Check Properties	Evaluate Polyline Length Check Properties	
Check Title Check Description	Polyline Check Description	
Short Line	O Part/Path	
Feature Class Feature Class/Subtype Stream - CacheRiverHealthCheckT1 gdb • Where Clause SQL	○ Segment Length Meters ○ Operation <(ess than) Less Than: 20	
Type Polyline Part/Path Segment	geometries for polyline features where the length is within specified parameters. E Reviewer Remarks Notes Severity 1 ▼ 1 - High Severity 1 ▼ 5 - Low	
OK Cancel	OK Cancel	

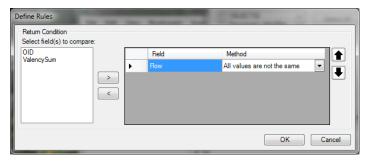
Pseudo Nodes Check

This will identify all two lines connected at a single end point. This will include valid case of a terminal confluence.

Valency Check Properties	A CONTRACTOR OF THE OWNER.	Valency Check Properties
Check Title	Check Description	Fields to Include for Feature Class 2
Pseudo Node	Point Valency: 1	OBJECTID Permanent_Identifier FDate Clear All Clear All
Feature Class 1 Feature Class/Subtype Stream_AllEndPt - CacheRiverHealthChec -		Validation Method Valency at point location
✓ Always Run on Full Database	Line Valency: 3	Valency at nonexistent point location Line Valency: 3
Where Clause SQL	Searches for points or nodes of linear features that intersect with a specified number of linear features.	Valency © Operation = (equal to) Searches for points or nodes of linear features that intersect with a specified number of linear features.
Fields to Include for Feature Class 1	-	Equal to: 2
OBJECTID A Permanent_Identifier FDate Clear All		Customize Table Query
Feature Class 2 Feature Class/Subtype		Advanced Settings
Stream - CacheRiverHealthCheckT1.gdb		Use XY tolerance
Always Run on Full Database		Use comparison rules Define Rules
Where Clause		Reviewer Remarks
SQL		Notes
ОК	Cancel	OK Cancel

١





Outlets Check

Flow Split Check

Execute SQL Check Properties	Execute SQL Check Properties
Check Title Outlet Feature Class Feature Class/Subtype Stream - CacheRiverHealthCheckT1.gdb Vhere Clause	Check Title Fow Split Feature Class Feature Class/Subtype Stream - CacheRiverHealthCheckT1.gdb Vhere Clause Returns features or rows resulting from an
NextDownID = -1 SQL Reviewer Remarks SQL Notes 1 - High Severity 3 • 5 - Low	Reviewer Remarks SQL SQL Query where clause Notes 1 - High Severity 3 • 5 - Low
OK Cancel	OK Cancel



Appendix 1.2 Stream Area Checks (StreamAreaCheck.rbj)

Breaching Processing Area Check

Geometry on Geometry Check Properties		Geometry on Geometry Check Properties	A
Check Title Breaching Processing Area Feature Class/Subtype Stream - CacheRiverHealthCheckT1.gdb Always Run on Full Database Where Clause SQL Feature Class/Subtype ProcessingArea - CacheRiverHealthCheckT Always Run on Full Database Where Clause SQL Spatial Relation Check Type Wathin Custom Spatial Relation	Check Description	Not - find features not in this relationship Merge features from feature class 2 Tolerance Meters Attributes None Compare All Attributes Attributes to Ignore (Feature Level Metadata) Select Attributes Compare Attributes Select Attributes Reviewer Remarks Notes 1 - High Severity 1 - 5 - Low	 Check Description Check Description Check Description Check Description Check Description Returns relationship specific geometries for features in Feature Class 1 that have a user-defined relationship to features in Feature Class 2. Relationships include: crosses; intersects, touches, and others. Any 9-Intersection relationship can be used by providing a strain containing 1; "" f," or "" An optional tolerance can be applied to the relationship for features in both feature classes.
OK Car	ncel	ОК	Cancel

Stream Boundary Proximity Check

Execute SQL Check Properties Execute SQL Check Properties Check Title Check Description Check Title Check Description Boundary Proximity Vertex Sanp Feature Class Feature Class Feature Class/Subtype Feature Class/Subtype Stream - ttt.gdb Stream - CacheRiverHealthCheckT1.gdb -. -. Always Run on Full Database Always Run on Full Database TYPE=4 TYPE=4 Where Clause Where Clause SQL Returns features or rows resulting from an SQL Query whereclause Returns features or rows resulting from an SQL Query whereclause NearBnd = 1 QCVALUE >= 1 SQL Reviewer Remarks Reviewer Remarks Notes Notes 1 - High 1 - High Severity 5 -Severity 5 -5-Low 5 - Low OK Cancel OK Cancel

Vertex Snap Check



Appendix 1.3 Stream Advanced Checks (StreamAdvancedCheck.rbj)

Watershed Boundary Proximity Check

Execute SQL Check Properties				
Check Title	Check Description			
Watershed Boundary Proximity				
Feature Class Feature Class/Subtype				
Stream - ttt.gdb	• •			
☑ Always Run on Full Database	TYPE=4			
Where Clause NearWsh = 1	Returns features or rows resulting from an SQL Query whereclause			
Reviewer Remarks Notes				
Severity 3				
ОК	Cancel			



Appendix 1.4 Polygon Basic Checks

These checks apply to these three polygon geometries:

- Sink (SinkBasicCheck.rbj)
- Water body (WaterbodyBasicCheck.rbj)
- Watershed (WatershedBasicCheck.rbj)

The examples are presented for the water body, but they are exactly the same as for the other two feature classes (the feature class name will be pointing to the appropriate layer).

Invalid Geometry Check

Multipart Geometry Check

Invalid Generates Check Properties			Multipart Polygon Check Properties
Invalid Geometry Check Properties	Check Description	n	Check Title
Invalid Geometry	OBJECTID SH	APE FCSUBTYPE	Multipart Feature Class
Feature Class Feature Class/Subtype Waterbody - CacheRiverHealthCheckT1.gd		LYLINE AQO40- BRIDGE LINE LYLINE AQO40- BRIDGE LINE	Feature Class/Subtype Waterbody - CacheRiverHealthCheckT1.g
Always Run on Full Database Where Clause SQL	This includes null geometries, featur	BRIDGE LINE with invalid geometry. geometries, empty	Where Clause Where Clause SQL Search Goal
Reviewer Remarks Notes Severity 1 - High 5 - Low		un angro goononio.	Only Multiple Parts Only Holes Parts and Holes Both
OK Cancel			OK Cancel



Duplicate Geometry Check

Duplicate Geometry Check Properties	
Check Title	Check Description
Duplicate Geometry	
Feature Class 1 Feature Class/Subtype Waterbody - CacheRiverHealthCheckT1.gd • Where Clause SQL Feature Class 2 Feature Class/Subtype Waterbody - CacheRiverHealthCheckT1.gd • Where Clause Where Clause	Returns geometries from features in Feature Class 1 that are co-located with geometries in Feature Class 2. Optionally, leatures must share identical non-spatial attributes with the option to ignore feature level metadata.
SQL Attributes	-
Attributes to lanore	Cancel

Polygon Overlap Check

Polygon Overlap/Gap is Sliver Check Properties	Polygon Overlap/Gap is Sliver Check Properties	
Polygon Overlap/Gap is Sliver Check Properties Check Title Overlap Feature Class 1 Feature Class/Subtype Waterbody - CacheRiverHealthCheckT1.gd Feature Class 2 Feature Class 2 Feature Class 2 Feature Class 2 Feature Class 8 Where Clause Sal Search Goal: Find Overlaps OK Cancel	Where Clause SQL Search Goal; Find Gaps and Overlaps Find Overlaps Find Overlaps Find Gaps Returns overlap/gap geometries between polygon features from two feature classes that have a Thinness Ratio	



Polygon Touches Each Other Check

Geometry on Geometry Check Properties	Geometry on Geometry Check Properties
Check Title Touches Each Other Feature Class 1 Feature Class Subtype Waterbody - CacheRiverHealthCheckT1gd Image: Solution of Full Database Where Clause Solution Feature Class 2 Feature Class 2 Feature Class/Subtype Waterbody - CacheRiverHealthCheckT1gd Waterbody - CacheRiverHealthCheckT1gd Waterbody - CacheRiverHealthCheckT1gd Waterbody - CacheRiverHealthCheckT1gd Where Clause Solution Spatial Relation Check Type Custom Spatial Relation OK	Not - find features not in this Image: Provide the stature of



Appendix 1.5 Polygon Area Checks

These checks apply to these three polygon geometries:

- Sink (SinkAreaCheck.rbj)
- Water body (WaterbodyAreaCheck.rbj)
- Watershed (WatershedAreaCheck.rbj)

The examples are presented for the water body, but they are exactly the same as for the other two feature classes (the feature class name will be pointing to the appropriate layer).

Geometry on Geometry Check Properties	Geometry on Geometry Check Properties
Check Tile Check Tile Npt within Processing Area Feature Class 1 Feature Class/Subtype Waterbody - CacheRiverHealthCheckT1.gd Always Run on Full Database Where Clause SqL Feature Class 2 Feature Class 2. ProcessingArea - CacheRiverHealthCheckT Always Run on Full Database Where Clause SqL ProcessingArea - CacheRiverHealthCheckT Always Run on Full Database Where Clause Spatial Relation Check Type Within Custom Spatial Relation OK Cancel	Not - find features not in this Image: relationship Merge features from feature class 2 Tolerance Meters Attributes Image: Compare All Attributes Compare All Attributes Select Attributes Select Attributes Select Attributes Select Attributes Select Attributes Select Attributes Select Attributes Select Attributes Select Attributes Select Attributes Select Attributes Select Attributes Select Attributes Select Attributes Select Attributes Select Attributes Select Attributes Motes Select Attributes Select Attributes Select Attributes Select Attributes Motes Select Attributes Select Attributes

Polygon Not Within Processing Area Check



Appendix 1.6 Waterbody Advanced Checks (WaterbodyAdvancedCheck.rbj)

Non-Draining Waterbody Check

Geometry on Geometry Check Properties	Geometry on Geometry Check Properties
Check Title Non-Draining Waterbody Feature Class 1 Feature Class/Subtype Waterbody - CacheRiverHealthCheckT1.gd Always Run on Full Database Where Clause Solut Feature Class 2 Feature Class 2 Feature Class 2/Subtype Stream - CacheRiverHealthCheckT1.gd Always Run on Full Database Where Clause Stream - CacheRiverHealthCheckT1.gd Non-Draining Waterbody Stream - CacheRiverHealthCheckT1.gd Nore Clause Solut Spatial Relation Check Type Intersects Custom Spatial Relation OK	Not - find features not in this relationship Image: Second Seco



Appendix 1.7 Watershed Advanced Checks (WatershedAdvancedCheck.rbj)

Non-draining Watershed Check

Execute SQL Che	eck Properties	
Check Title		Check Description
Non-Draining	g Watershed	
Feature Class Feature Class		
	Watershed - ttt.gdb	•
	un on Full Database	TYPE=4
Where Clause		Returns features or rows resulting from an
NonDrained >	>= 1 S(GQL SQL Query where clause
Reviewer Rem	narks	
Notes		
Saurity (5 🔻 1 - High	
Severity	5 - Low	
	ОК	Cancel



Appendix 2. A worked out example

Initial data organization

The set of presented QC processes has been applied to the Cache watershed in Illinois. Cache is located at the southern tip of Illinois at the confluence of Mississippi and Ohio rivers. It is partially contained in subregion 07 (Mississippi) and 05 (Ohio). It covers all of 07140108 and portion north of Ohio River in 05140206 HUC-8. Red is the general area of the Cache watershed used in this example. Imagery shows Mississippi River (west boundary) and Ohio River (south boundary).



Figure 8. Cache area of interest. Black are the HUC-8 boundaries.

The vector data were put together from the NHD Hires (24K) dataset from USGS (<u>ftp://nhdftp.usgs.gov/DataSets/Staged/SubRegions/FileGDB/HighResolution/</u>). They were clipped based on the red Cache boundary (extent is slightly larger than the watershed boundary derived from the DEM) from the two subregion file geodatabases (05 and 07). NHDFlowline and NHDWaterbody feature classes were used as a source for streams, sinks, and water bodies for QC checks.

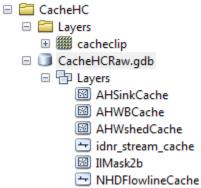
All NHDFlowlines were considered as streams (5465 features). NHDWaterbodys were split into sinks and water bodies based whether they were overlying the terminal (outlet) stream or not respectively. 42 sink and 8143 water body features were thus identified.



HUC12 that reside within Cache watershed were selected from the WBD dataset to represent the watershed boundaries.

The DEM is the 10m NED that was projected into Ill projection using bilinear interpolation and clipped to the red boundary.

The starting data were stored in the following data structure (top directory is called CacheHC):



Folder "Layers" contains the DEMcalled "cacheclip" and the "raw" starting geodatabase (CacheHCRaw.gdb) with the following data structure:

- AHSinkCache sinks
- AHWBCache water bodies
- AHWshedCache watersheds
- idnr_stream_cache original IDNR streams used in the previous stream characterization project
- IIMask2b Cache extent used for data clipping ("red" polygon)
- NHDFlowlineCache streams

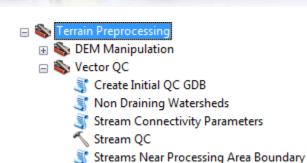
Naming of spatial layer for the "raw" data is arbitrary. It is important that the raster dataset is in projected system that will be used in the QC process. Vector data can be in any projection as they will be projected into the DEM projection during the initial processing step.

Performing operations

AH tools used in these operations are in the "Arc Hydro Tools -> Terrain Preprocessing -> Vector QC" toolset.

30



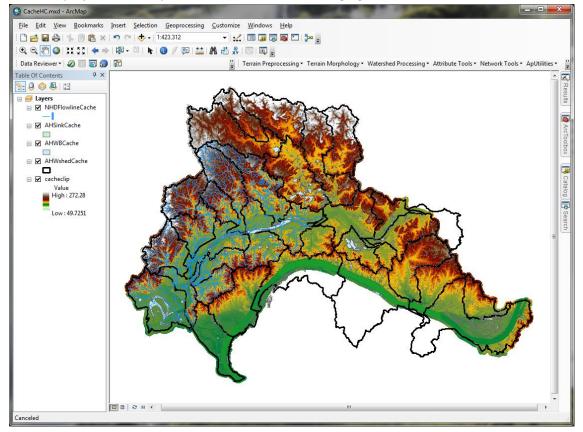


- Streams Near Processing Area Bounda
- 💐 Streams Near Watershed Boundary

Initial project and database setup

Steps 1-4 are not required. You can jump into step 5 directly from an existing ArcMap project that has the required layers in it.

- 1. Open a new ArcMap project and add the DEM (cacheclip).
- Save the project. It is recommended to save it in the directory one level above the directory containing the DEM. In this case it would be in the "CacheHC" directory. The name of the project can be anything. In this case it is "CacheHC.mxd".
- 3. Add vector data to be used in the QC process. Modify symbology as needed.
- 4. Save the project. The map should look like the following figure.





5. Run "Create Initial QC GDB" AH QC tool (make sure that Arc Hydro toolbox is available). Populate the form.

💕 Create Initial QC GDB				x
Target Directory				*
C:\CurrWork\IlDNR\CacheHC				3
Target GDB Name				
CacheHCQCRun1				
Reference DEM				_
cacheclip			-	3
Stream FC (optional)				
NHDFlowlineCache			-	3
Sink FC (optional)			_	
AHSinkCache			- 2	3
Waterbody FC (optional)			_	
AHWBCache			-	3
Watershed FC (optional)			_	
AHWshedCache			-	3
			_	
	OK Cancel	Environments	Show Help	>>

6. Run the tool to create the QC database.

```
Executing: CreateInitialQCGDB C:\CurrWork\IlDNR\CacheHC
CacheHCQCRun1 cacheclip NHDFlowlineCache AHSinkCache AHWBCache
AHWshedCache
Start Time: Tue Aug 06 15:57:27 2013
Running script CreateInitialQCGDB...
  Creating QC geodatabase...
      Creating QC geodatabase completed in 1.72 seconds.
  Defining processing area polygon and outline polyline...
      Defining processing area polygon and outline polyline
completed in 18.38 seconds.
  Importing optional vector layers...
          Importing optional stream layer ...
          Importing optional sink layer ...
          Importing optional waterbody layer ...
          Importing optional watershed layer ...
      Importing optional vector layers completed in 15.58 seconds.
Completed script CreateInitialQCGDB...
Succeeded at Tue Aug 06 15:58:03 2013 (Elapsed Time: 36.00 seconds)
```

7. This step creates a new geodatabase with processing area and its outline and loads the optional vector data into it if provided. The geodatabase structure is as follows (layer names are fixed). Note that any of the optional vector data will be copied (and projected if necessary) into the new geodatabase with all the fields as they are. If you want to simplify the data structure of the



input layers, you can remove all unnecessary fields. QC tools do not require any fields to be present as they will create them on as needed basis.

CacheHCQCRun1.gdb

🗆 🖶 Layers

- 🖾 ProcessingArea
- 🛨 ProcessingAreaOutline
- 🖾 Sink
- 🛨 Stream
- Waterbody
- 🖾 Watershed
- 8. Save the project.
- 9. Close ArcMap.

Running basic checks

- 1. Start a new ArcMap project. Make sure that DR toolbar is visible and extension is active. Make sure that Arc Hydro toolbox is available.
- 2. Add data from the geodatabase created in step 6 above.
- 3. Save the project with the same name as the geodatabase created in step 6 (in this case "CacheHCQCRun1"). This will ensure that the HydroID dispenser tables are created and managed in the proper place (QC geodatabase).
- 4. Run "Stream Connectivity Parameters" AH QC tool.

Stream Connectivity Parameters	
Stream Stream	
,	
	OK Cancel Environments Show Help >>



```
Executing: StreamConnectivityParameters Stream

Start Time: Tue Aug 06 16:52:15 2013

Running script StreamConnectivityParameters...

Populating stream attributes...

Populating HydroID ...

Populating FROM_NODE and TO_NODE ...

Populating NextDownID and generating Stream_FS table ...

Populating FlowSplitCnt ...

Populating stream attributes completed in 29.50 seconds.

Defining stream line end points...

Defining stream line end points completed in 4.59 seconds.

Adding results to the map ...

Cleaning up...

Completed script StreamConnectivityParameters...

Succeeded at Tue Aug 06 16:52:49 2013 (Elapsed Time: 34.00 seconds)
```

5. Start a new DR session. Specify the reviewer workspace to be the same as the geodatabase with the data to be QC'd (this can be pointing to a different gdb – it will depend on the user preferences and data editing workflows). Populate other fields in the form as desired and start the session.

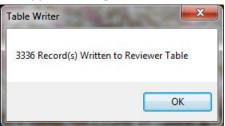
Reviewer Session Manager	×
Reviewer Workspace	
CacheHCQCRun1.gdb	Browse
User Name	
Cache	
Session	
ID Session 1 : Session 1	▼ New
Name Run1	
Reviewer Dataset Version	
· · · · · · · · · · · · · · · · · · ·	
Advanced Start Session	Apply Close

- 6. To run the basic stream checks, run "Batch Validate" from the DR menu.
 - a. Once the form is presented, select "Add from File ..." and navigate to the AH tools installation directory, "rbjfiles" folder (usually at C:\Program Files (x86)\ESRI\WaterUtils\ArcHydro\rbjfiles) where the AH HC rbj files are and select "StreamBasicCheck.rbj". You should see a message in the lower left corner of the form that "Batch Job was successfully updated to current workspace". This indicates that the



checks could recognize the current workspace and layer in it to which to apply the checks to. You can click on "Workspace ..." to review the target layers.

- b. Click on "Full Database" radio button to ensure that the checks will be performed on the full extent of the data.
- c. Click on "Run" to execute the checks.
- d. After about one minute (performance will vary depending on the hardware) a message will appear stating that a number of records have been written to the reviewer table.



- 7. To review the records in the reviewer table, click on "Reviewer Table" on the DR menu. That will open the reviewer table and you will be able to navigate through identified records and use all DR capabilities. Following figures demonstrate few examples of reviewer table interface.
 - a. The "CHECKTITLE" shows which check "triggered" a feature. Clicking in the left-most box in each row will navigate and zoom to that feature on the map.

General	Drag a column I	header here to	group by that col	umn.			
-/-	RECORDID	OBJECTID	CHECKTITLE	ORIGINTABLE	ORIGINCHECK	REVIEWSTATUS	SU
	3037	5062	Outlet	Stream	Execute SQL	(NextDownID = -1)	
	3038	5141	Outlet	Stream	Execute SQL	(NextDownID = -1)	
	3039	5142	Outlet	Stream	Execute SQL	(NextDownID = -1)	
	3040	5193	Outlet	Stream	Execute SQL	(NextDownID = -1)	
	3041	5350	Outlet	Stream	Execute SQL	(NextDownID = -1)	
	3042	5450	Outlet	Stream	Execute SQL	(NextDownID = -1)	
a fair	3043	3	Flow Split	Stream	Execute SQL	(FlowSplitCnt >= 1)	
	3044	18	Flow Split	Stream	Execute SQL	(FlowSplitCnt >= 1)	
-	3045	62	Flow Split	Stream	Execute SQL	(FlowSplitCnt >= 1)	
	10040	חד	FI 0.10	C+	Final col	/FILC-100-15 11	1

b. By moving the "CHECKTITLE" column header into the white box with "Drag a column header here to group by that column" text in it, you can group the records by the "problem" type and get an overview of problem types and their counts.

Reviewer Ta	ble
General	CHECKTITLE A
://	CHECKTITLE : Flow Split (125 items)
	CHECKTITLE : Outlet (86 items)
	CHECKTITLE : Pseudo Node (2956 items)
	CHECKTITLE : Short Line (169 items)



- 8. Run other basic checks following steps in #6. Depending on the desired editing workflows, you can start a new DR session (step #5) or keep current one active. In this example, all the checks will be done in a single session.
 - a. Basic sink checks SinkBasicCheck.rbj
 - b. Basic water body checks WaterbodyBasicCheck.rbj. This check produced the following errors:

Reviewer Processing Errors	
Batch Job Processing Errors	
Warnings Touches Each Other: Unable to build error geometry for feature 3212 Touches Each Other: Unable to build error geometry for feature 4984	
Close	

- c. Basic watershed checks WatershedBasicCheck.rbj
- 9. Upon completion of basic checks, the following list of "issues" can be produced.

Reviewer Ta	ble
General	
	CHECKTITLE : Duplicate Geometry (1 items)
1	ORIGINTABLE : Waterbody (2 items)
	CHECKTITLE : Flow Split (1 items)
	ORIGINTABLE : Stream (125 items)
	CHECKTITLE : Outlet (1 items)
	ORIGINTABLE : Stream (86 items)
A.	CHECKTITLE : Overlap (1 items)
	ORIGINTABLE : Waterbody (16 items)
	CHECKTITLE : Pseudo Node (1 items)
	ORIGINTABLE : Stream_AllEndPt (2956 items)
	CHECKTITLE : Short Line (1 items)
	ORIGINTABLE : Stream (169 items)
	CHECKTITLE : Touches Each Other (3 items)
	ORIGINTABLE : Sink (2 items)
	ORIGINTABLE : Waterbody (80 items)
	ORIGINTABLE : Watershed (33 items)

10. The identified problems can now be navigated and should be fixed before proceeding to the other checks.

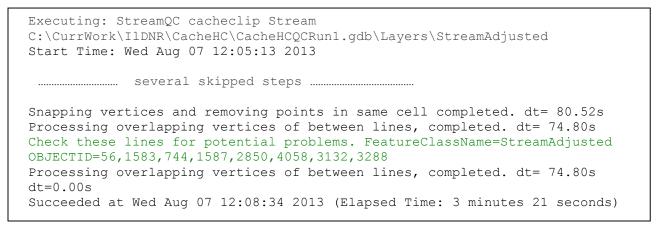


Running other checks

Ideally, all real problems identified in basic checks should be fixed before proceeding to the advanced checks. In this example, that was not done and the checks are performed on "raw" data. If necessary, open ArcMap with the QC project and start the reviewer session (same or new one).

- 1. Stream geometry with respect to raster
 - a. Add DEM to the map (if not already in it).
 - b. Run "Stream QC" AH QC tool. This is a complex function that can take time to finish. It provides a lot of feedback on what is being performed while the function is running.

🖌 Stream QC	
Input DEM	^
cacheclip	- 🖻
Input Stream	
Stream	- 🖻
Output Line	
C:\CurrWork\IDNR\CacheHC\CacheHCQCRun1.gdb\Layers\StreamAdjusted	
	Ψ.
OK Cancel Environments	Show Help >>



c. Run "Streams Near Processing Area Boundary" AH QC tool. 160 ft search distance is used as a 5 cell buffer is being planned for use in the DEM Reconditioning function (so the search distance is slightly larger than 5 times the cell size).



Input Stream Layer					_	_ ^
Stream					-	3
Input Processing Area Outline Layer					_	
ProcessingAreaOutline						3
Input Search Distance						
			160	Feet		•
	ОК	Cancel	Envir	onments	Show Help	>>

Executing: StreamsNearProcessingAreaBoundary Stream ProcessingAreaOutline
"160 Feet"
Start Time: Wed Aug 07 13:15:50 2013
Running script StreamsNearProcessingAreaBoundary
Populating NearBnd stream attribute
Populating NearBnd stream attribute completed in 4.07 seconds.
Cleaning up
Completed script StreamsNearProcessingAreaBoundary
Succeeded at Wed Aug 07 13:15:54 2013 (Elapsed Time: 4.00 seconds)

- d. To run the stream area checks, run "Batch Validate" from the DR menu with "StreamAreaCheck.rbj" file. (11 more records will be written to the Reviewer Table).
- 2. Polygon geometry with respect to raster.
 - a. To run the sink area checks, run "Batch Validate" from the DR menu with "SinkAreaCheck.rbj" file. (No additional records will be written to the Reviewer Table).
 - b. To run the water body area checks, run "Batch Validate" from the DR menu with "WaterbodyAreaCheck.rbj" file. (11 more records will be written to the Reviewer Table). Note that this check takes a long time to process if the number of water bodies is large (as in this case). It is possible to perform similar check using out of the box spatial selection and skipping use of DR for the task but that would then eliminate the benefits of using DR for management of all the checks. It would also be possible to modify the currently implemented DR check and implement a different technique while still using DR for the overall check management.
 - c. To run the watershed area checks, run "Batch Validate" from the DR menu with "WatershedAreaCheck.rbj" file. (6 more records will be written to the Reviewer Table).
- 3. Advanced checks
 - To run the water body advanced checks, run "Batch Validate" from the DR menu with "WaterbodyAdvancedCheck.rbj" file. (7564 more records will be written to the Reviewer Table).
 - b. To run the watershed advanced checks:
 - i. Run "Non Draining Watersheds" AH QC tool.



💐 Non Draining Wat	ersheds
Input Watershed Watershed Input Stream (option Stream Input Sink (optional) Sink	
	OK Cancel Environments Show Help >>
Executing: NonDrainingWater Start Time: Wed Aug 07 14:2	sheds Watershed Stream Sink 7:03 2013

Start Time: Wed Aug 07 14:27:03 2013
Running script NonDrainingWatersheds
Populating NonDrained watershed attribute
Populating NonDrained watershed attribute completed in 1.50 seconds.
Cleaning up
Completed script NonDrainingWatersheds
Succeeded at Wed Aug 07 14:27:05 2013 (Elapsed Time: 2.00 seconds)

- ii. Run "Batch Validate" from the DR menu with "WatershedAdvancedCheck.rbj" file. (1 more records will be written to the Reviewer Table).
- c. To run the stream advanced checks:
 - Run "Polygon To Line" standard geoprocessing function (Data Management Tools -> Features toolset) to create watershed outlines (as linear features).
 Make sure that the resulting feature class is saved in the proper (QC) geodatabase.

Polygon To Line	x
Input Features	^
Watershed	3
Output Feature Class	
C: \CurrWork\IDNR \CacheHC \CacheHC \QCRun 1.gdb \Layers \WatershedOutline	3
Identify and store polygon neighboring information (optional)	
	-
OK Cancel Environments Show Hel	p >>

Executing: PolygonToLine Watershed C:\CurrWork\IlDNR\CacheHC\CacheHCQCRun1.gdb\Layers\WatershedOutline IGNORE_NEIGHBORS Start Time: Wed Aug 07 14:57:35 2013 Succeeded at Wed Aug 07 14:57:35 2013 (Elapsed Time: 0.00 seconds)



ii. Run "Streams Near Watershed Boundary" AH QC tool.

I Streams Near Watershed Boundary
Input Stream Layer Stream Input Watershed Outline Layer Watershed Outline Input Search Distance 160 Feet OK Cancel Environments Show Help >>
Executing: StreamsNearWatershedBoundary Stream WatershedOutline "160 Feet" Start Time: Wed Aug 07 15:03:09 2013
Running script StreamsNearWatershedBoundary
Populating NearWsh stream attribute
Populating NearWsh stream attribute completed in 30.64 seconds.
Cleaning up

iii. Run "Batch Validate" from the DR menu with "StreamAdvancedCheck.rbj" file.(212 more records will be written to the Reviewer Table).

Succeeded at Wed Aug 07 15:03:40 2013 (Elapsed Time: 31.00 seconds)

Upon completion of all checks, the following list of "issues" can be produced.

Completed script StreamsNearWatershedBoundary...



Reviewer Tal	ble
General	
	CHECKTITLE : Boundary Proximity (1 items)
1	ORIGINTABLE : Stream (1 items)
	CHECKTITLE : Duplicate Geometry (1 items)
	ORIGINTABLE : Waterbody (2 items)
	CHECKTITLE : Flow Split (1 items)
	ORIGINTABLE : Stream (125 items)
A.	CHECKTITLE : Non-Draining Waterbody (1 items)
	ORIGINTABLE : Waterbody (7564 items)
	CHECKTITLE : Non-Draining Watershed (1 items)
126	ORIGINTABLE : Watershed (1 items)
===	CHECKTITLE : Not Within Processing Area (2 items)
	ORIGINTABLE : Waterbody (11 items)
	ORIGINTABLE : Watershed (6 items)
	CHECKTITLE : Outlet (1 items)
	ORIGINTABLE : Stream (86 items)
	CHECKTITLE : Overlap (1 items)
	ORIGINTABLE : Waterbody (16 items)
	CHECKTITLE : Pseudo Node (1 items)
	ORIGINTABLE : Stream_AllEndPt (2956 items)
	CHECKTITLE : Short Line (1 items)
	 ORIGINTABLE : Stream (169 items)
	CHECKTITLE : Touches Each Other (3 items)
	ORIGINTABLE : Sink (2 items)
	ORIGINTABLE : Waterbody (80 items)
	ORIGINTABLE : Watershed (33 items)
	CHECKTITLE : Vertex Sanp (1 items)
	ORIGINTABLE : Stream (10 items)
	CHECKTITLE : Watershed Boundary Proximity (1 items)
	ORIGINTABLE : Stream (212 items)

DR now allows prioritization and navigation through the identified issues. Standard ArcGIS editing capabilities should be used to resolve any real issues. The most important issues to be resolved are the ones related to the flow patterns (stream connectivity, directionality, and flow splits, and existence and location of sink polygons). This is an iterative process. As features are edited, relevant checks have to be rerun.