NAIP 2007 Arizona
Preliminary Seamline Shapefile
Inspection

For NAIP 2007, several pilot projects were initiated to develop NAIP into a better program. These pilots moved forward in Arizona. One of them, the delivery of a seamline shapefile, was completely new to NAIP. It became the task of APFO/GSB to develop a draft process to analyze and then inspect these seamline shapefiles. This document will outline that process.

Background

Until the Arizona pilot, APFO had always received a DOQQ polygon shapefile with each compressed county mosaic (CCM). The purpose of this shapefile was to determine acquisition dates of the DOQQs and to determine which DOQQs were used to create the CCM. The issue with the acquisition date was that it consisted of the majority dates of the imagery. If imagery for one DOQQ was collected on different dates, it did not matter; the majority acquisition date was entered into the polygon attributes. The seamline shapefile hopes to improve upon this issue. The seamlines are an adjoining set of polygons that identify the exact flying times of the associated image.

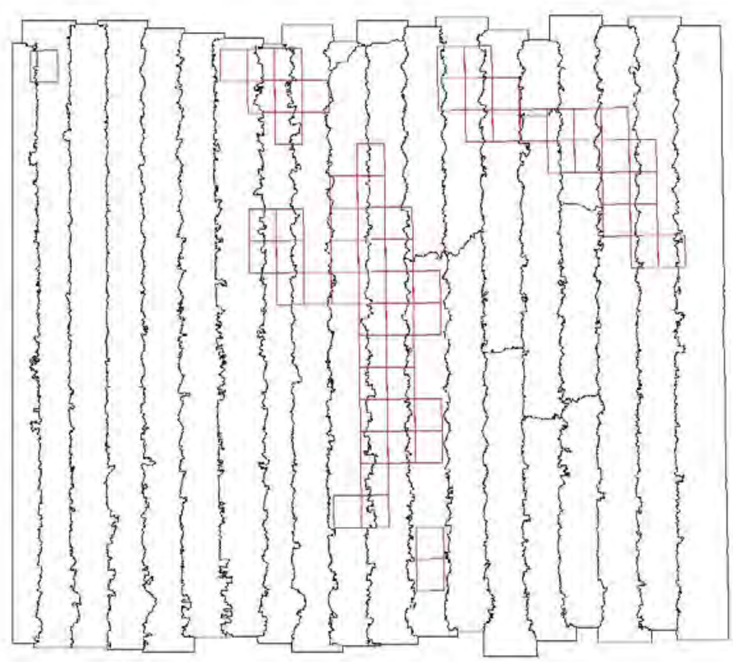


Figure 1: DOQQ polygons (RED) and seamline polygons (BLACK)

Each of the seamline polygons has their own unique attributes. The attributes are similar to those of the DOQQ polygons, but go into much more detail (see figures 2 and 3).

Attributes of naip_1-1_2n_s_az003_2006_1

FID	Shape	QQName	BCON	IDAT	DOQQ	QKEY
0	Polygon	REDINGTON SW	NC	2006-08-19	n3211037.sw	322230N1102615W
1	Polygon	DOUGLAS NE NW	NC	2006-08-27	n3110936.nw	312615N1093345W
2	Polygon	DOUGLAS NE SW	NC	2006-08-27	n3110936.sw	312230N1093345W
3	Polygon	SQUARE MOUNTAIN NW	NC	2006-08-27	n3210941.nw	321845N1095615W
4	Polygon	SQUARE MOUNTAIN NE	NC	2006-08-27	n3210941.ne	321845N1095230W
5	Polygon	SQUARE MOUNTAIN SE	NC	2006-08-27	n3210941.se	321500N1095230W
6	Polygon	WILLCOX NORTH NW	NC	2006-08-27	n3210942.nw	321845N1094845W
7	Polygon	OLGA SW	NC	2006-08-19	n3210946.sw	321500N1091845W
8	Polygon	OLGA SE	NC	2006-08-19	n3210946.se	321500N1091500W
9	Polygon	WILLCOX SOUTH SE	NC	2006-08-27	n3210950.se	320730N1094500W
10	Polygon	SIMMONS PEAK NW	NC	2006-08-27	n3210951.nw	321115N1094115W
11	Polygon	SIMMONS PEAK SW	NC	2006-08-27	n3210951.sw	320730N1094115W
12	Polygon	LITTLE WOOD CANYON NE	NC	2006-08-19	n3210954.ne	321115N1091500W
13	Polygon	VANAR NW NW	NC	2006-08-19	n3210955.nw	321115N1091115W
14	Polygon	VANAR NW NE	NC	2006-08-27	n3210955.ne	321115N1090730W
15	Polygon	VANAR NW SE	NC	2006-08-27	n3210955.se	320730N1090730W
16	Polygon	VANAR NW	NC	2006-08-27	n3210956.nw	321115N1090345W
17	Polygon	VANAR SW	NC	2006-08-27	n3210956.sw	320730N1090345W
18	Polygon	COCHISE NW	NC	2006-08-27	n3210957.nw	320345N1095615W
19	Polygon	COCHISE SW	NC	2006-08-27	n3210957.sw	320000N1095615W
20	Polygon	COCHISE NE	NC	2006-08-27	n3210957.ne	320345N1095230W
21	Polygon	COCHISE SE	NC	2006-08-27	n3210957.se	320000N1095230W
22	Polygon	SULPHUR SPRING NE	NC	2006-08-27	n3210958.ne	320345N1094500W

Record: 1 Show: All Selected Records (0 out of 66 Selected.) Options

Figure 2: DOQQ polygon attribute table.

Attributes of ortho_1-1_1n_s_az003_2007_1

FID	Shape	IDAT	SDATE	EDATE	CAM_TYPE	CAM_MAN	CAM_MOD	SENSNUM	SHAPE_AREA
0	Polygon	6/9/2007	06/09/2007 15:37	06/09/2007 16:10	Digital	Leica Geosystems	ADS52	30022	1373900000
1	Polygon	6/9/2007	06/09/2007 16:14	06/09/2007 16:38	Digital	Leica Geosystems	ADS52	30022	1132200000
2	Polygon	6/9/2007	06/09/2007 16:43	06/09/2007 17:15	Digital	Leica Geosystems	ADS52	30022	1136900000
3	Polygon	6/9/2007	06/09/2007 17:19	06/09/2007 17:43	Digital	Leica Geosystems	ADS52	30022	766640000
4	Polygon	6/9/2007	06/09/2007 17:47	06/09/2007 18:19	Digital	Leica Geosystems	ADS52	30022	521530000
5	Polygon	6/9/2007	06/09/2007 18:22	06/09/2007 18:46	Digital	Leica Geosystems	ADS52	30022	569680000
6	Polygon	6/10/2007	06/10/2007 15:13	06/10/2007 15:46	Digital	Leica Geosystems	ADS52	30022	420920000
7	Polygon	6/10/2007	06/10/2007 15:50	06/10/2007 16:13	Digital	Leica Geosystems	ADS52	30022	416480000
8	Polygon	6/10/2007	06/10/2007 16:18	06/10/2007 16:37	Digital	Leica Geosystems	ADS52	30022	50555000
9	Polygon	6/10/2007	06/10/2007 16:41	06/10/2007 16:58	Digital	Leica Geosystems	ADS52	30022	27035000
10	Polygon	6/25/2007	06/25/2007 16:18	06/25/2007 16:46	Digital	Leica Geosystems	ADS52	30022	860710000
11	Polygon	6/25/2007	06/25/2007 16:50	06/25/2007 17:19	Digital	Leica Geosystems	ADS52	30022	1241700000
12	Polygon	6/25/2007	06/25/2007 17:16	06/25/2007 17:45	Digital	Leica Geosystems	ADS52	30102	719610000
13	Polygon	6/25/2007	06/25/2007 17:22	06/25/2007 17:50	Digital	Leica Geosystems	ADS52	30022	1071600000
14	Polygon	6/25/2007	06/25/2007 17:54	06/25/2007 18:22	Digital	Leica Geosystems	ADS52	30022	1102900000
15	Polygon	6/25/2007	06/25/2007 18:26	06/25/2007 18:54	Digital	Leica Geosystems	ADS52	30022	1069300000
16	Polygon	6/25/2007	06/25/2007 18:57	06/25/2007 19:24	Digital	Leica Geosystems	ADS52	30022	1082600000
17	Polygon	6/25/2007	06/25/2007 19:28	06/25/2007 19:56	Digital	Leica Geosystems	ADS52	30022	986610000
18	Polygon	6/30/2007	06/30/2007 15:14	06/30/2007 15:33	Digital	Leica Geosystems	ADS52	30022	333340000
19	Polygon	6/30/2007	06/30/2007 15:37	06/30/2007 15:46	Digital	Leica Geosystems	ADS52	30022	544800000
20	Polygon	6/30/2007	06/30/2007 16:04	06/30/2007 16:23	Digital	Leica Geosystems	ADS52	30022	592820000
21	Polygon	6/30/2007	06/30/2007 16:34	06/30/2007 16:43	Digital	Leica Geosystems	ADS52	30022	684610000
22	Polygon	6/30/2007	06/30/2007 16:47	06/30/2007 17:00	Digital	Leica Geosystems	ADS52	30022	699940000

Record: 1 Show: All Selected Records (0 out of 26 Selected.) Options

Figure 3: Seamline polygon attribute table.

Notice in the above tables the differences between the fields. The **IDAT** field shows the majority flying date in the DOQQ table. In the seamline table, the **IDAT** field exists, but the **SDATE** (seam polygon collection start time) and **EDATE** (seam polygon collection end time) fields supplement the **IDAT**. The date of acquisition is known, as well as the start and end times of a particular image acquisition. This allows end users to pinpoint exactly when a particular part of an image was collected.

Inspection Methodology

The creation of an inspection process for the 33 shapefiles encompassing Arizona required some thought. Previously, the inspection involved loading the shapefile to see if it covered the imagery on the CCM. Because the seamline files are so much different than the DOQQ files, a new process had to be created. Several questions came about regarding the seamline inspection:

1. Is there complete county (CCM) coverage?
2. Is there complete state coverage?
3. Is the polygon topology correct?
4. Will the format of the attributes pass through a parser?

Answering each of these questions became part of the inspection process methodology. Eventually, the process developed into four major steps: checking that the CCM has complete seamline shapefile coverage, checking that all attribute fields are correctly populated and formatted, checking the validity of the topology for the shapefiles, and checking to see if all polygons are at least 10 acres in size. For the purpose of this inspection, the four steps were done in ArcGIS 9.1.

Inspection Process

Part I

The first step is to check for complete seamline coverage on the CCM. In Arizona, several counties require multiple volumes in order to provide complete county coverage. In this case, all shapefile volumes were loaded along with all CCM volumes at the same time to determine coverage.

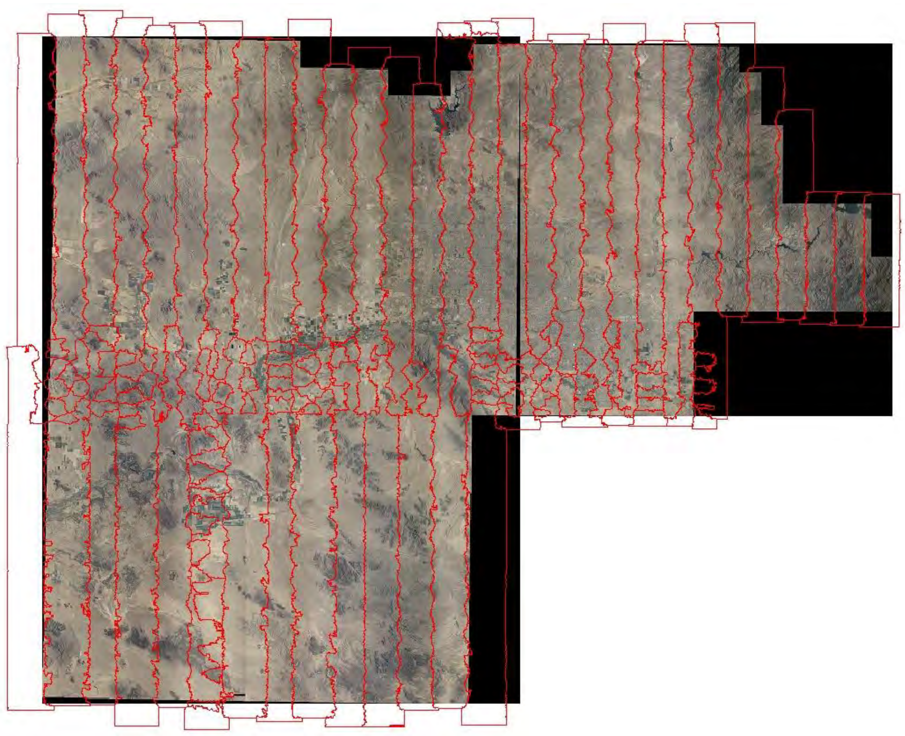


Figure 4: Both volumes of Maricopa County (04013) with seam coverage.

This step begins with loading the CCMs and the seamline shapefiles into an ArcMap session. Once all appropriate data layers are loaded, the coverage is then checked.

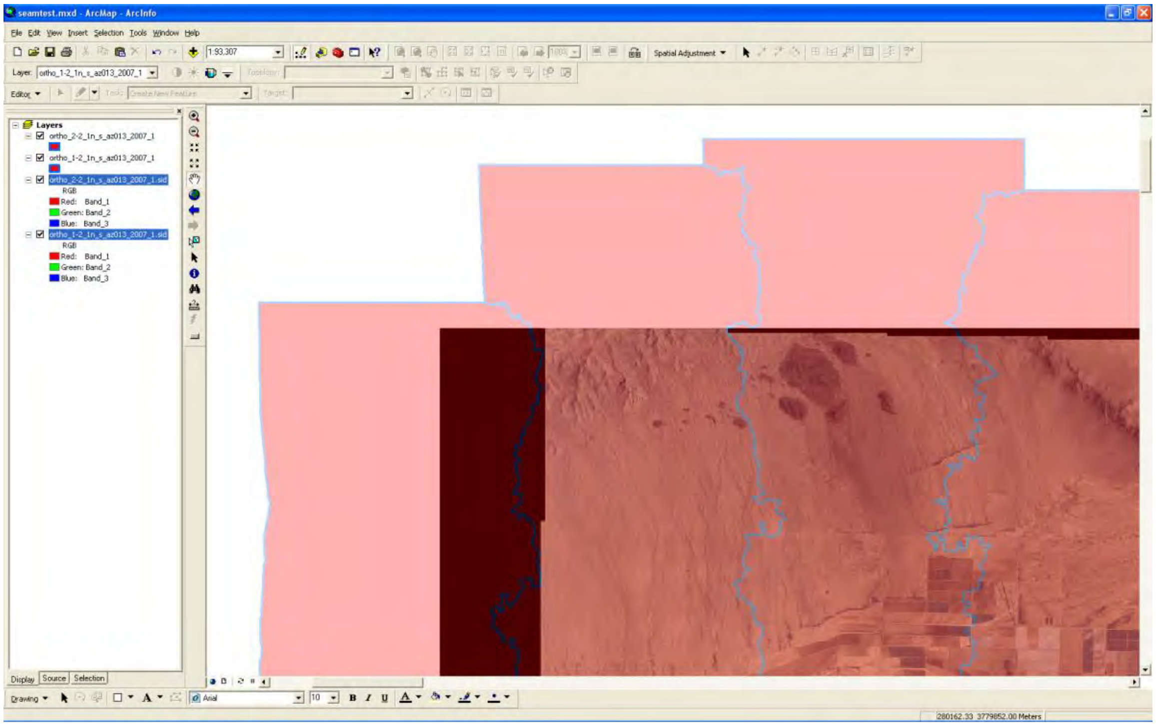


Figure 5: Checking CCM for complete seamline coverage

The check is done by panning around the edge of the CCM coverage looking for any image areas that are not covered by seam polygons. In the above example, a transparency was applied to the polygons to aid in determining coverage.

Part II

This part of the inspection process is the most time consuming. It required the development of a new inspection method. In this step, the seam polygons were checked based on the validity of their respective topologies. To accomplish this, topologies had to be created based upon seamline shapefile specifications in the NAIP contract. Then, the topologies were validated based upon two rules: no gaps between polygons and no overlaps between polygons. All of the inspection in part two was done in ArcGIS 9.1. After Part II inspection was completed, the results were then compiled into a spreadsheet.

The following illustrates the steps used in part two of the inspection. First, a personal geodatabase is created in ArcCatalog (see figure 6).

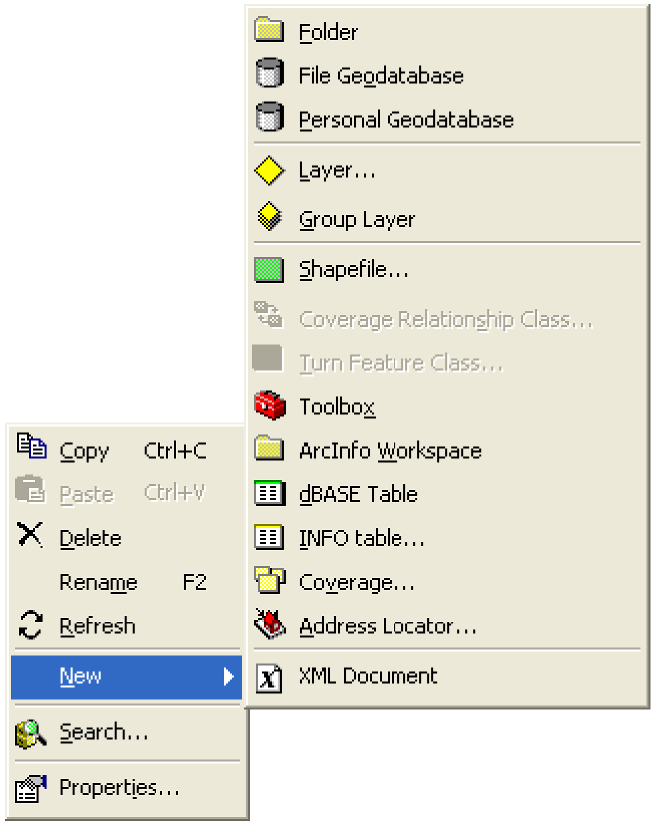


Figure 6: Creating a personal geodatabase

Next, a new feature dataset is created in the personal geodatabase.

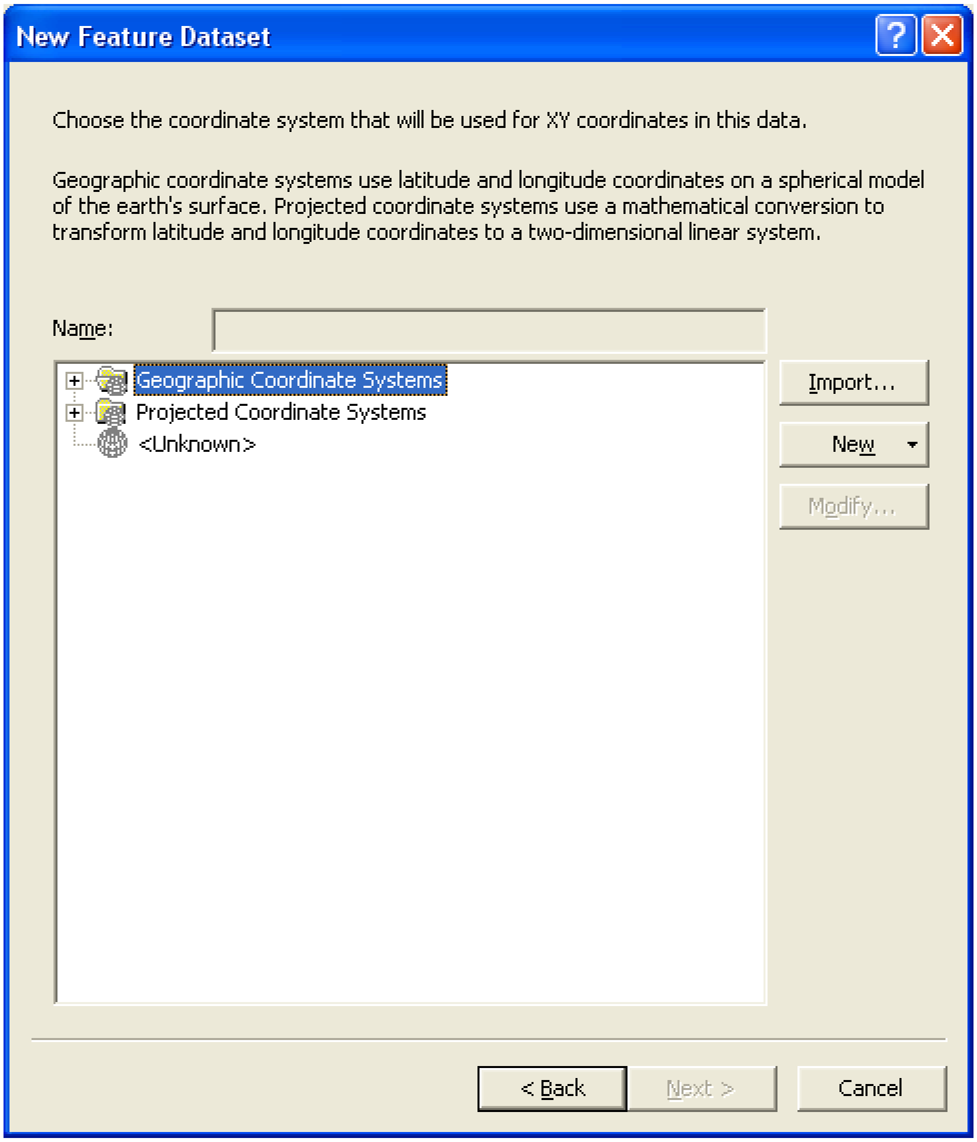


Figure 7: Creating a new feature dataset

When naming the output feature dataset, the county name or FIPS code should be used. Eventually, the naming convention for the feature datasets will be standardized. Next, the seamline shapefile is imported into the new feature dataset as a single feature class (see figure 8). This new feature class will be used in the topology validation. Once again, the naming of the output feature class should be standardized.

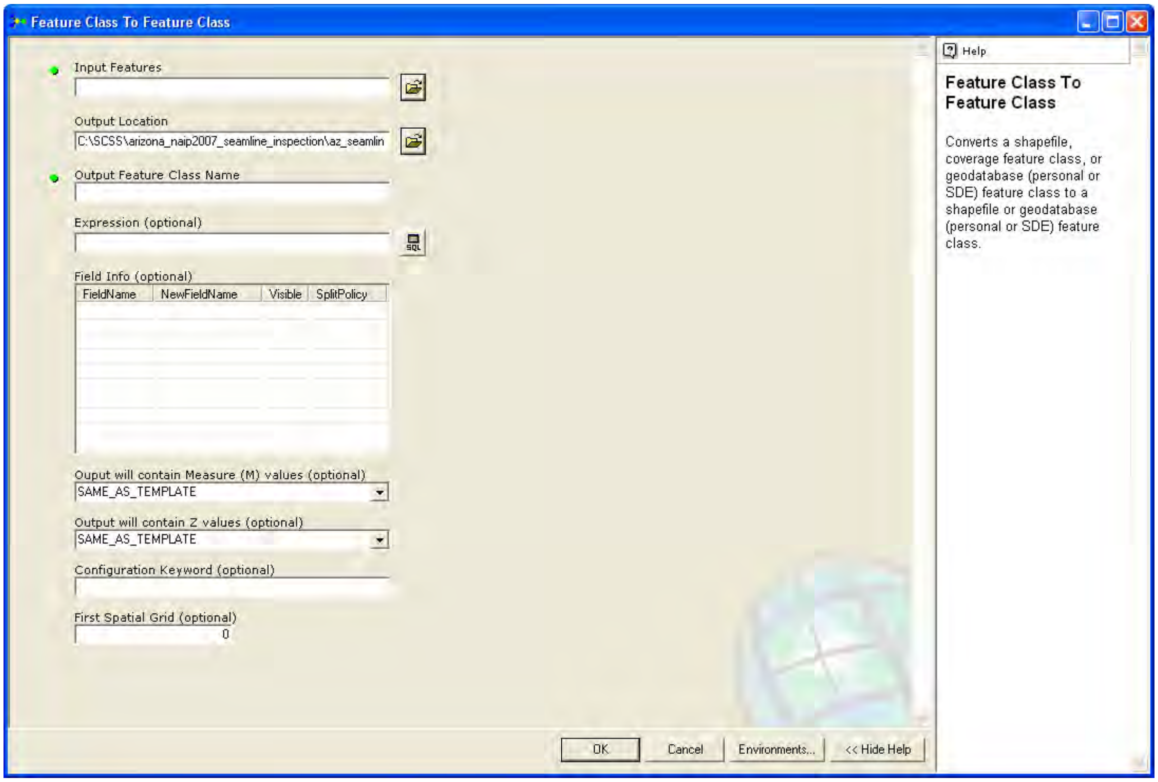


Figure 8: Importing a feature class

Once a feature class is imported, then a topology needs to be created. The topology is created in ArcCatalog within the feature class using the "New Topology" wizard (figure 9). Eventually, the wizard will prompt for "rules". These allow the inspector to set topological rules for the inspection process. In this inspection, the two rules are "must not overlap" and "must not have gaps" (see figure 10). Once a topology is created, it is then loaded into an ArcMAP session along with the feature class layer. With the "Editor" active and the "Topology" toolbar accessible, the topology can be validated (see figure 11). By validating the topology, the inspector can see where the errors fall based upon the rules created earlier. These errors are displayed according to the Symbology in the table of contents. Once the errors are validated, the errors can be viewed in more detail using the error inspector tool (see figure 12).

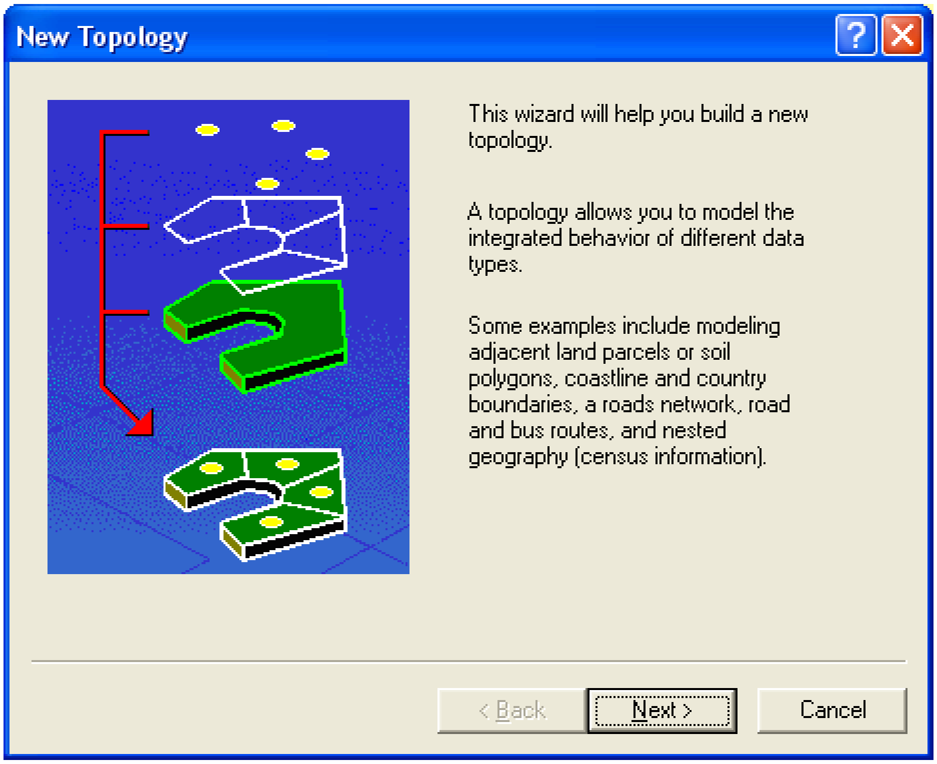


Figure 9: Creating a new topology using the wizard

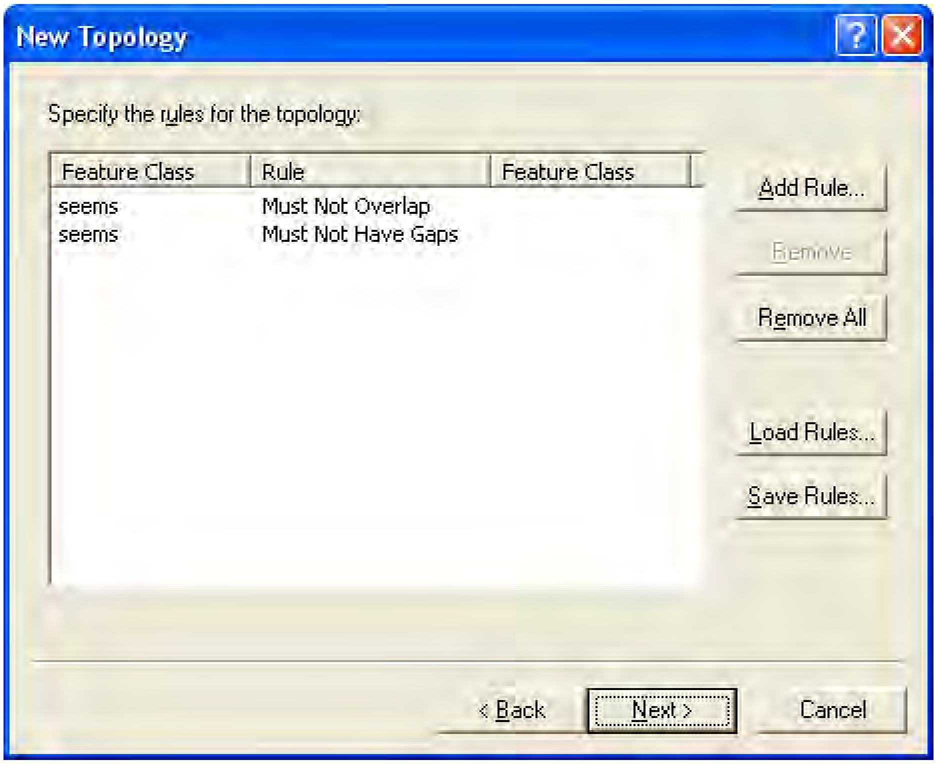


Figure 10: Topology rules

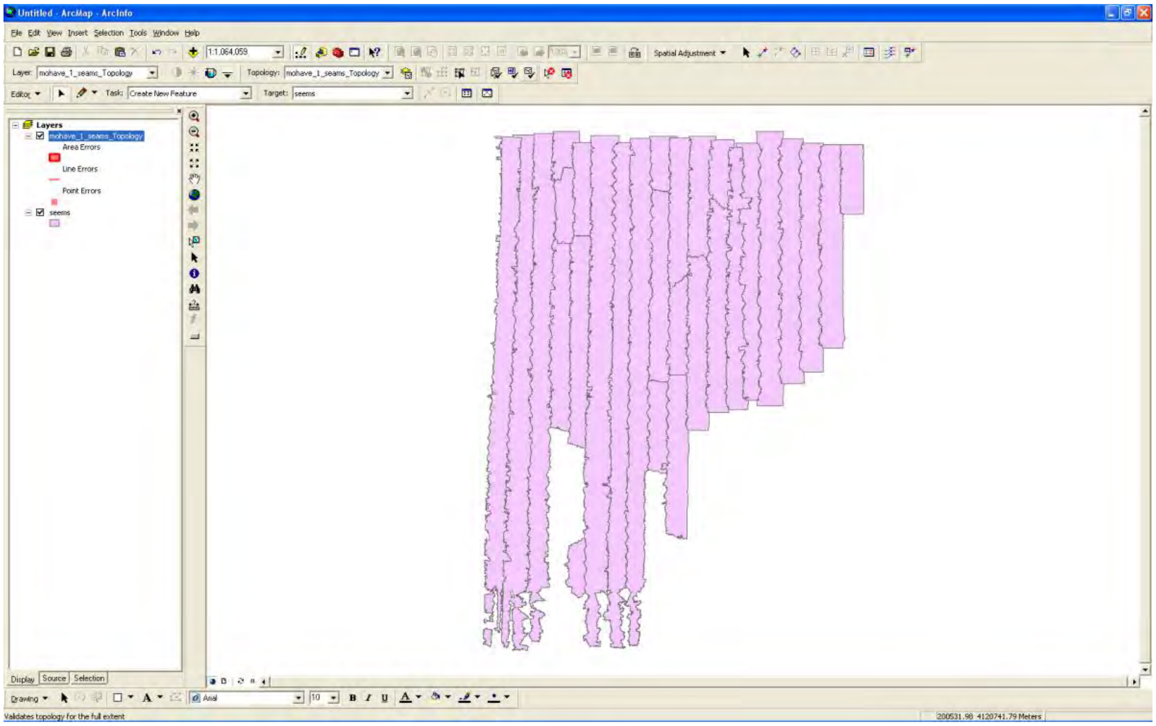
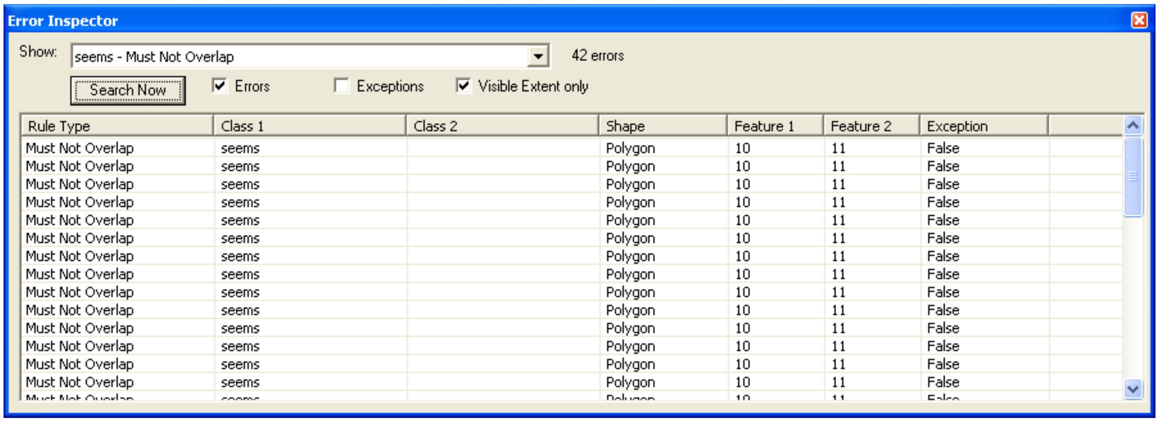


Figure 11: Validating a topology



Error Inspector

Show: seems - Must Not Overlap 42 errors

Search Now ☑ Errors ☐ Exceptions ☑ Visible Extent only

Rule Type	Class 1	Class 2	Shape	Feature 1	Feature 2	Exception
Must Not Overlap	seems		Polygon	10	11	False
Must Not Overlap	seems		Polygon	10	11	False
Must Not Overlap	seems		Polygon	10	11	False
Must Not Overlap	seems		Polygon	10	11	False
Must Not Overlap	seems		Polygon	10	11	False
Must Not Overlap	seems		Polygon	10	11	False
Must Not Overlap	seems		Polygon	10	11	False
Must Not Overlap	seems		Polygon	10	11	False
Must Not Overlap	seems		Polygon	10	11	False
Must Not Overlap	seems		Polygon	10	11	False
Must Not Overlap	seems		Polygon	10	11	False
Must Not Overlap	seems		Polygon	10	11	False
Must Not Overlap	seems		Polygon	10	11	False
Must Not Overlap	seems		Polygon	10	11	False

Figure 12: Viewing the error inspector

Once the errors have been calculated, they are then logged into a custom spreadsheet (see figure 13). This spreadsheet is a preliminary draft, and could be updated or modified in the future.

Microsoft Excel - Arizona_Inspection_Report.xls

Arizona Seam File Inspection Report

County	FIPS	UTM Zone	Volumes	County Coverage	Volume Coverage	Error Occurs	Pass Rule 1 (No Overlap)	Pass Rule 2 (No Gaps)	Attribute Data	Hyperlink	Remarks
Apache	04001	12	3	Yes	No	Vol1	Yes	Yes	Yes		
Cochise	04003	12	1	Yes	Yes	NA	Yes	Yes	Yes		one part of multiple polygon smaller than 10 acres
Coconino	04005	12	7	Yes							
- vol 1	04005	12			Yes	No	Yes	Yes	Yes		
- vol 2	04005	12			No	Yes	No - 767	Yes	Yes		
- vol 3	04005	12			No	Yes	No - 383	Yes	Yes		
- vol 4	04005	12			Yes	Yes	No - 568	Yes	Yes		
- vol 5	04005	12			Yes	Yes	No - 522	Yes	Yes		
- vol 6	04005	12			Yes	No	Yes	Yes	Yes		
- vol 7	04005	12			Yes	Yes	No - 318	Yes	Yes		
Gila	04007	12	2	Yes							
- vol1	04007	12			Yes	No	Yes	Yes	Yes		
- vol 2	04007	12			Yes	No	Yes	Yes	Yes		
Graham	04009	12	1	Yes	Yes	No	Yes	Yes	Yes		
Greenlee	04011	12	1	Yes	Yes	No	Yes	Yes	Yes		
La Paz	04012	11,12	1	No	N/A	Yes	No-95 errors	Yes	Yes		Small part of county not covered
Maricopa	04013	12	2	Yes							
- vol 1	04013	12			Yes	No	Yes	Yes	Yes		
- vol 2	04013	12			Yes	No	Yes	Yes	Yes		to halves for topology check for outside polygon
Mohave	04015	11,12	4	Yes							Image acquisition times appear to be wrong
- vol 1	04015	11,12			Yes	Yes	No- 42 errors	Yes	Yes		
- vol 2	04015	11,12			Yes	Yes	No- 70 errors	Yes	Yes		
- vol 3	04015	11,12			Yes	Yes	No- 76 errors	Yes	Yes		two halves for topology check for outside polygon
- vol 4	04015	11,12			Yes	Yes	No- 79 errors	Yes	Yes		two halves for topology check for outside polygon
Navajo	04017	12	3	Yes							
- vol 1	04017	12			Yes	No	Yes	Yes	Yes		
- vol 2	04017	12			Yes	No	Yes	Yes	Yes		
- vol 3	04017	12			Yes	No	Yes	Yes	Yes		
Pima	04019	12	2								
- vol 1	04019	12									
- vol 2	04019	12									
Pinal	04021	12	1								
Santa Cruz	04023	12	1								
Yavapai	04025	12	2								
- vol 1	04025	12									
- vol 2	04025	12									
Yuma	04027	11,12	2	Yes							
- vol 1	04027	11,12			Yes	No	Yes	Yes	Yes		
- vol 2	04027	11,12			Yes	No	Yes	Yes	Yes		

Figure 13: Error log spreadsheet

Part III

The third part of the inspection involves checking the attribute table. There are 10 fields in the table when viewing it in ArcGIS. Ideally, all of the fields in the table are populated with the correct attributes (see figure 14).

Attributes of ortho_2-2_1n_s_az013_2007_1

FID	Shape	IDATE	SDATE	EDATE	CAM_TYPE	CAM_MAN	CAM_MOD	SENSNUM	SHAPE_AREA
0	Polygon	6/4/2007	06/04/2007 21:21	06/04/2007 21:35	Digital	Leica Geosystems	ADS52	30011	648922180
1	Polygon	6/4/2007	06/04/2007 21:40	06/04/2007 21:57	Digital	Leica Geosystems	ADS52	30011	638751833
2	Polygon	6/4/2007	06/04/2007 21:56	06/04/2007 22:10	Digital	Leica Geosystems	ADS52	30011	734713320
3	Polygon	6/7/2007	06/07/2007 20:01	06/07/2007 20:20	Digital	Leica Geosystems	ADS52	30011	63256456
4	Polygon	6/7/2007	06/07/2007 20:24	06/07/2007 20:41	Digital	Leica Geosystems	ADS52	30011	23177472
5	Polygon	6/7/2007	06/07/2007 22:31	06/07/2007 22:48	Digital	Leica Geosystems	ADS52	30011	30713940
6	Polygon	6/7/2007	06/07/2007 22:52	06/07/2007 23:25	Digital	Leica Geosystems	ADS52	30011	625391775
7	Polygon	6/7/2007	06/07/2007 23:29	06/07/2007 23:57	Digital	Leica Geosystems	ADS52	30011	683275532
8	Polygon	6/16/2007	06/16/2007 15:06	06/16/2007 15:27	Digital	Leica Geosystems	ADS52	30102	772493199
9	Polygon	6/16/2007	06/16/2007 15:30	06/16/2007 15:51	Digital	Leica Geosystems	ADS52	30102	714935956
10	Polygon	6/16/2007	06/16/2007 15:55	06/16/2007 16:16	Digital	Leica Geosystems	ADS52	30102	778206469
11	Polygon	6/16/2007	06/16/2007 16:19	06/16/2007 16:39	Digital	Leica Geosystems	ADS52	30102	763111513
12	Polygon	6/16/2007	06/16/2007 16:42	06/16/2007 17:01	Digital	Leica Geosystems	ADS52	30102	761573300
13	Polygon	6/16/2007	06/16/2007 17:04	06/16/2007 17:23	Digital	Leica Geosystems	ADS52	30102	877084980
14	Polygon	6/17/2007	06/17/2007 16:43	06/17/2007 17:13	Digital	Leica Geosystems	ADS52	30011	711848000
15	Polygon	6/17/2007	06/17/2007 17:17	06/17/2007 17:45	Digital	Leica Geosystems	ADS52	30011	666414738
16	Polygon	6/17/2007	06/17/2007 17:49	06/17/2007 18:19	Digital	Leica Geosystems	ADS52	30011	695178082
17	Polygon	6/17/2007	06/17/2007 18:23	06/17/2007 18:50	Digital	Leica Geosystems	ADS52	30011	741975570
18	Polygon	6/17/2007	06/17/2007 18:55	06/17/2007 19:25	Digital	Leica Geosystems	ADS52	30011	627144604
19	Polygon	6/17/2007	06/17/2007 19:28	06/17/2007 19:56	Digital	Leica Geosystems	ADS52	30011	704290214
20	Polygon	6/24/2007	06/24/2007 15:41	06/24/2007 16:05	Digital	Leica Geosystems	ADS52	30102	752457950
21	Polygon	6/24/2007	06/24/2007 16:08	06/24/2007 16:33	Digital	Leica Geosystems	ADS52	30102	814463857
22	Polygon	6/24/2007	06/24/2007 16:36	06/24/2007 17:00	Digital	Leica Geosystems	ADS52	30102	732270621
23	Polygon	6/24/2007	06/24/2007 17:03	06/24/2007 17:27	Digital	Leica Geosystems	ADS52	30102	1013086286

Record: 1 Show: All Selected Records (0 out of 37 Selected) Options

Figure 14: Seamline shapefile attribute table

Currently, a quick check of the attribute table is done only to see if all fields are populated. The formats of the different fields are not checked. This should be implemented in future inspections.

Part IV

The final part of the inspection is determining the size of the seamline polygons. The NAIP contract states that "***the smallest area covered by a single polygon shall not be smaller than 40,470 square meters (approximately 10 acres***)". The **SHAPE_AREA** field lists the areas of each polygon in square meters. Using the "Sort Ascending" option on the field, it was determined if any polygons do not meet this requirement.

Future Inspections and Recommendations

This preliminary inspection was intended to be an outline of how the inspection can be done. Each of the steps can be done in a different manner. Once an offline manual inspection process has been firmly established, most of the parts can be automated.

Inspection Recommendations

- Modify as needed, and then perform this offline inspection process on the 2008 seamline shapefile pilot states. This inspection will be done first by the SCSS, and then the QA section. It should be the task of SCSS to document the offline (non-automated) inspection before QA begins so that good reference material is on hand.
- Automate the process. Even though this inspection is somewhat labor intensive, most of the steps can be automated. Parts III and IV could be done upon data ingestion. This will require the assistance of TSB and QA.
- All inspection data collected should be entered into a formatted spreadsheet, database, or whatever means for storing data.
- Once these recommendations are in place, the QA section should add the seamline inspection to their NAIP imagery inspection process.

Contract Language Recommendations

The following is taken from the NAIP 2007 contract, task order 1, revised task order:

"(b) Seamline Shapefile. The index shall be topologically correct and contain a polygon for each exposure used in the creation of the CCM. The polygons shall completely cover the entire area represented by the extents of the visible imagery and not have gaps (slivers) in the shapefile within the area represented by the extents of the visible imagery. ***If the size of the county requires multiple volumes, then the polygons shall completely cover the entire area***

represented by that volume to the extent of the visible imagery and not have gaps (slivers) in the shapefile within the area represented by the extents of the visible imagery. There shall be no overlapping polygons in the shapefile. ***There shall be no multiple part polygons in the shapefile.*** Unless approved in writing by the Contractor Officer, the smallest area covered by a single polygon shall not be smaller than 40,470 square meters (approximately 10 acres). The shapefile shall be attributed with the following data columns…"

The parts in italics with the bold font are the contract language recommendations regarding the seamline shapefile for NAIP 2008.