

Standard Operating Procedure 15: Data Recording, Entry, Management, and Archiving, version 1.0

This SOP gives step-by-step instructions for managing the data collected in the Fish Monitoring program. This SOP describes the procedure for field checking data, data entry, data verification and correction, and database organization and documentation. Accurate collection and stewardship of data are critical to the success of the fish monitoring program. Ensuring data quality must be a priority during crew training, and throughout the data collection, entry, and verification and validation process.

Responsibilities for quality control

The fisheries biologist and field crew leader are responsible for ensuring that crew members:

- 1) can correctly identify fish species,
- 2) understand all field protocols,
- 3) understand data fields and values on paper data forms and in the database.

Correct taxonomic identification and documentation of fish species is an important aspect of monitoring. Frequent quality control assessments should be made by the fisheries biologist or lead technician, which will consist of field-checking crew member identifications and the quality of their entered data.

Crew members are responsible for learning species identification, field protocols, and all data fields and values on paper data forms and in the database. In addition, each crew member is responsible to do the following whenever they are unsure or unclear about the protocol:

- 1) Do not guess.
- 2) **Ask** the crew leader or other crew members while in the field.
- 3) Ask the biologist.
- 4) Take copious and clear notes including photographs about the uncertainty.
- 5) Follow through to securing guidance and correcting the uncertainty.
- 6) Err on the side of collecting too much data.

Data recording

Field crews record data on paper data forms (see SOP 18). Carry datasheets within an enclosed clipboard. **Print legibly** all information on datasheets using pencil. There may be some crew members who will not be able to record data due to poor handwriting skills. Waterproof/all-weather copies of blank datasheets should be carried into the field at all times. All corrections to the datasheet should be made by crossing through the error. Information should never be erased and incorrect or older information should never be overwritten.

Entering data in the field into hand-held data loggers and computers eliminates the need to transfer data from paper data forms into the project database, and thus may reduce the time needed to enter data later. However, field crews at SHEN have experimented with various data entry tools over the years and have always migrated back to recording data on paper.

Regardless of the data collection method, **all data values shall be called out and repeated** when one crew member is measuring and one crew member is recording. The measurer shall

clearly call out the value to be recorded. The recorder shall clearly repeat the value aloud as it is recorded, to ensure that the value was heard and recorded accurately. If the measurer does not hear a response from the recorder, the measurer shall repeat calling out the measurement until a response is obtained. There should be no extraneous chat from the rest of the crew while data is being recorded. This distracts from the data taking process and will result in erroneous data.

Field checking data

Field checking the data is the process of reviewing all paper data sheets before leaving the site to ensure all fields are complete and no obvious errors exist. A single person is needed to complete this step. Usually this task is completed by the field crew leader, but any field member who is knowledgeable of species in the park and the monitoring protocol, and has an eye for detail can also complete this task. Datasheets should be field checked by someone other than the data recorder.

On each datasheet, the person conducting the field check should ask the following:

- a. Is the header information complete – date, site number, stream name, crew?
- b. Is the site name used on the datasheet match the one used on the tree tag?
- c. Are multiple pages of the same datasheet numbered with site name and date filled in also?
- d. Are all species codes and numbers legible?
- e. Are all species codes valid codes on the park list?
- f. Do all species codes make sense for the given habitat?
- g. Are there any notes in the margins of the datasheets that need to be addressed, moved, or crossed out to eliminate future confusion?

Once it has been determined that the datasheet is complete, the initials of the crew member who field checked the datasheet should be added to the datasheet in the margin at the bottom right corner of each page.

Post-processing of field datasheets in office

After the site has been sampled, the datasheets go into the appropriate file cabinet drawer in the office. If questions or unknowns exist for the site, the datasheets should be flagged and filed appropriately. If possible, questions should be dealt with before leaving the office that day.

Electronic Files

The electronic files for this program can be found on the O: Drive in the working directory under “O:\WORKING\AQUATIC\”. There should be several directories with files for the program. Data management files are located in the “O:\WORKING\AQUATIC\Database\” directory (Figure 1). Electronic data files including the databases, scanned datasheets and photo monitoring images are stored in a directory structure on a file server with appropriate access levels granted to users. Prior to a data entry session an automated backup of the back-end data file is made and stored on the server to allow “roll-back” if database corruption is encountered. A scheduled back-up utility makes copies of changed files to a server in a different building at the SHEN HQ campus nightly. Annually these back-up directories are evaluated and archived or replaced by a “fresh” copy of the directory.

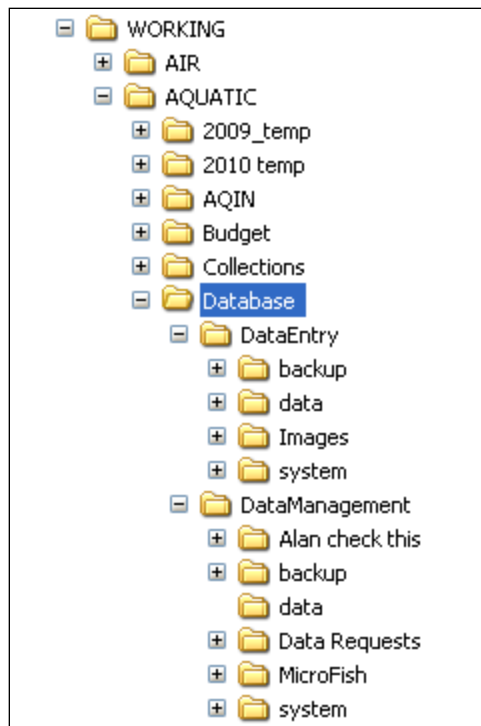


Figure 1. O Drive directories.

Database System

The Shenandoah National Park fish data system consists of three main components, the data entry application, the temporary storage databases, and the master database. Field data is entered into the first temporary database using the data entry application. Then to help assure accurate data the same data is entered a second time into a second temporary database. The two temporary databases are compared via script and any discrepancies can be reviewed and fixed. At the end of the field season, the data in the temporary databases are reviewed, spot checked, and then added to the master database by the program staff (see below). An ESRI shapefile is maintained to house the point locations of the sample sites. The shapefile can be joined to data exports in ArcGIS to display attributes of the sites spatially.

The current SHEN fish monitoring database was developed in Microsoft Access 95 but has been converted and updated to the current version of MS Access (Figure 2). This database predates the NPS I&M Natural Resources Database Template (NRDT) but many of the design principles are present.

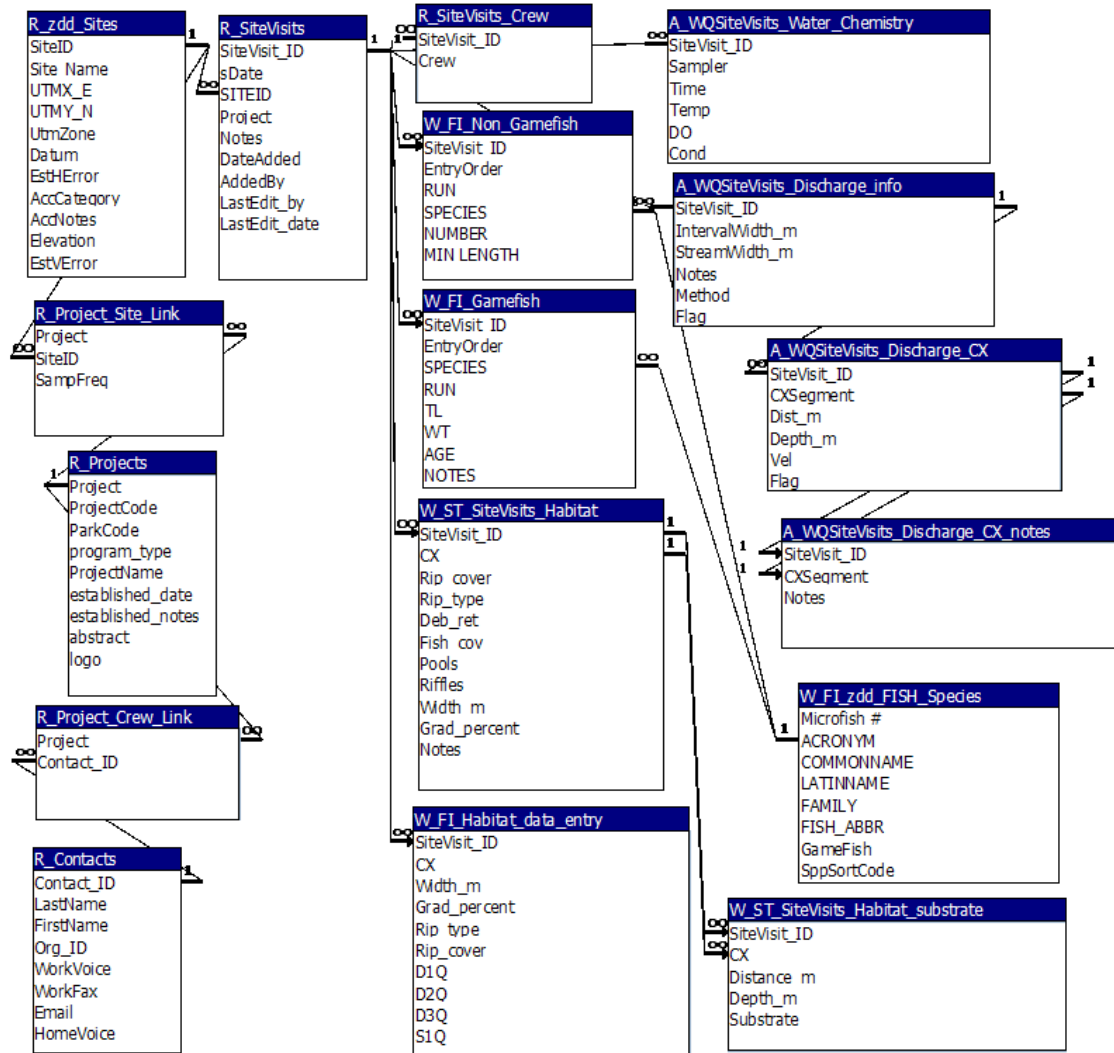


Figure 2. Back-End table design and relationships for the fisheries database.

Data Entry and Verification

The Shenandoah National Park fish data entry program consists of three main components, the data entry application, the temporary storage databases, and the master database. Data is entered into the first temporary database using the data entry application. Then to help provide accurate data the data is entered a second time into a second temporary database. The two temporary databases are compared and any discrepancies can be reviewed and fixed. At the end of the field season the data in the temporary databases are reviewed by the program staff and then added to the master database by the data manager.

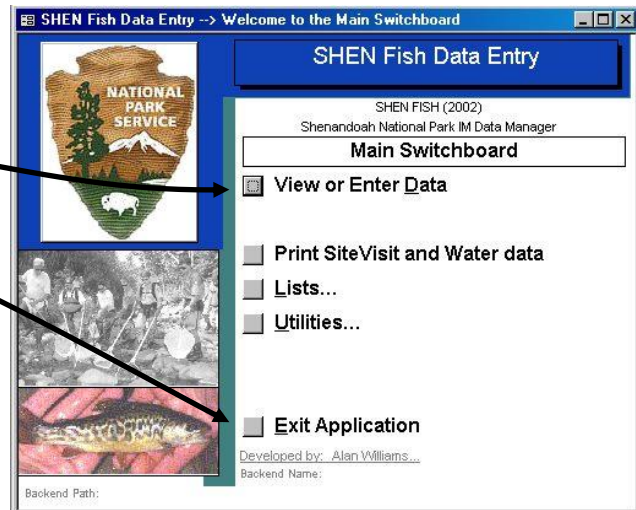
Data entry –

1) Overview:

To start data entry press
<View or Enter Data>

To quit press
<Exit Application>

The other buttons are used by the data
or program managers.



2) The main data entry screen has
basic date and location information at
the top, followed by a series of tabs
with different types of data on each tab.

To enter data for a new date and site
press the <Add New> button.

To search for data that has already been
entered use the pull down list.

You can also use the navigation buttons
at the bottom of the form to move
between records.

Data is automatically saved as you go,
row by row, so there is no need to save
manually.

3) **Water Chemistry:** fill in the information for each of the fields from the data sheet.

4) **Crew:** Select the members of the crew that were present from the right side (A) then press the “move” button (B). If you need to add someone to the list use the “Edit” button (C).

SHEN Fish Data Entry

Project: FISH Jump to (order by): date site Select a Site Visit Add New DONE

DATE (mmddyy): 06/27/01 SITEID: 1F138 Edit Stream: Piney River SiteVisit_ID: 2001-06-27-1F138

Water Chemistry and Crew Discharge Habitat GameFish Non-Gamefish Site Info and Notes

Sampler	Time	Temp	DO	Cond	pH	TDS	Level	Color	Notes
HYD1	9:00	14.63	9.73	26.00	6.73	0.0166			

Crew Working at that Site Visit

Crew
Carswell_Ben
Cyphert_Mike
Daley_Michelle
Degel_Kristin
Demarest_David
McBride_Ryan
Reider_Trevor
Rowles_Matt
Sager_Scott

Select >> then click below to Move to Current list

Potential Crew List

Name	Organization
Jim	SHEN
Mike	SHEN
David	SHEN
McBride	SHEN
Eckles	USOS
Hall	Volunteer
Ryan	SHEN
Reider	SHEN
Rowles	SHEN
Sager	SHEN

Edit Potential Crew list

Use these buttons to move between SiteVisits

Record: 19 of 55

Data Entry User: abw

5) Discharge:

Enter the stream width and the interval that was selected.

Then press the <Add Stream...> button and the CX and Dist data should be filled in by the computer.

After that data is confirmed enter in the depth and velocity for each distance.

SHEN Fish Data Entry

Project: FISH Jump to (order by): date site Select a Site Visit Add New DONE

DATE (mmddyy): 07/09/01 SITEID: 2FVA5 Edit Stream: Rapidan River SiteVisit_ID: 2001-07-09-2FVA5

Water Chemistry and Crew Discharge Habitat GameFish Non-Gamefish Site Info and Notes

Interval Width_m	Stream Width_m	Notes	Method	Flag	Add Stream
0.3	3.6		Flowmeter		CXs
1-00	0.00	0.00			
1-01	0.30	0.06			
1-02	0.60	0.16			
1-03	0.90	0.57			
1-04	1.20	0.55			
1-05	1.50	0.50			
1-06	1.80	0.39			
1-07	2.10	0.39			
1-08	2.40	0.33			
1-09	2.70	0.33			
1-10	3.00	0.22			
1-11	3.30	0.19			
1-12	3.60	0.00			

Use these buttons to move between SiteVisits

Record: 26 of 55

Data Entry User: abw

6) Habitat:

Enter the total length of the transect.

Then press the <Add Stream...> button and the CX data should be filled in by the computer.

After that data is confirmed enter the rest of data for each of the CX row of data.

SHEN Fish Data Entry

Project: FISH Jump to (order by): date site Select a Site Visit Add New DONE

DATE (mmddyy): 01/01/02 SITEID: 1F003 Edit Stream: Piney River (Current site 1996-present) SiteVisit_ID: 2002-01-01-1F003

Water Chemistry and Crew Discharge Habitat GameFish Non-Gamefish Site Info and Notes

Stream Distance: Add Stream Distances Use Auto DropList

CX	Width	Grad %	Rip	Cover	D1Q	D2Q	D3Q	S1Q	S2Q	S3Q	P	R	Notes
0	4.0		T	1	1.00	2.00	1.20	B	BR	C			
10	6.0	5	T	1	2.00	0.20	0.36	B	C	C	5	5	
20	4.0	50	T	3	1.20	1.20	0.20	B	BR	O	6	4	(Fake Data)
30			T										
40			T										
50			T										
60			T										
70			T										
80			T										
90			T										
100			T										
101			T										
109			T										

Use these buttons to move between SiteVisits

Record: 54 of 55

Data Entry User: abw

7) Game Fish:

Enter the total length of the transect.

Then press the <Add number...> button and the EntryOrder data should be filled in by the computer. This should be a number where the last two digits match the line number of the data sheet and the first 1 or 2 digit(s) is the page number.

The last Run and Species entered will be carried down so be sure to confirm they are correct.

EntryOrder	RUN	SPECIES	TL	WT	NOTES
101	1	BRT			
102	1	BRT			
103	1	BKT			
104	1	BKT			
105	1	BKT			
106	1	BKT			
107	1	BKT			
108	1	BKT			
109	1	BKT			
110	1	BKT			
111	1	BKT			
112	1	BKT			
113	1	BKT			
114	1	BKT			
115	1	BKT			

6) Non-Game Fish:

Enter the data for each run from the data sheet. The last Run entered will be carried down so be sure to confirm it is correct for each line.

If there are generic notes about that day at that site including data collection, equipment problems, water level, fishing signage etc, enter it in the space provided on the last tab.

RUN	SPECIES	NUM	MIN LEN	MAX LEN	WEIGHT	MORT	NOTES	Order
1	BKD	94	38	89	309.5	17		1
1	LGD	29	61	140	317	0		2
1	AME	2	642	663	1152	0		3
1	TED	4	44	58	7	0		4
2	BKD	71	34	82	198.7	17		5
2	LGD	27	70	120	259.8	0		6
2	AME	2	250	640	333.7	0		7
2	TED	5	47	52	7.5	0		8
3	BKD	51	46	87	129.9	16		9
3	LGD	6	55	137	89.8	0		10
3	AME	2	319		49.9	0		11
3	TED	4	39	61	6.7	0		12
3								13

7) Second phase data entry:

Follow the same basic procedure for entering data in both data entry phases.

The main differences on the 2nd phase are:

- The form will be yellow indicating a 2nd phase.
- You will not be entering the Date or site data again, rather you will be choosing from the list of already entered dates and sites.
- Once you have completed entering all of the data for a site press the <Check> button to compare it to the first phase.

Sampler	Time	Temp	DO	Cond	pH	TDS	Level	Col	Notes
hyd1	12:30	1.00	0.00	0.00	0.00				
*		0.00	0.00	0.00	0.00				

8) **Checking** the 1st and 2nd data entry phase:

After pressing the <Check> button (see above) wait a moment and this form will appear.

At the top should show the date and site it is checking.

If there are any listings in the 2) or 3) areas then each of those tables will need to be fixed. To fix a table select the table name and press the button below or just double click the table name.

9) Un-Matched Records:

Each table will have a custom form. Check the records in question against the data sheets. Either keep or delete each one.

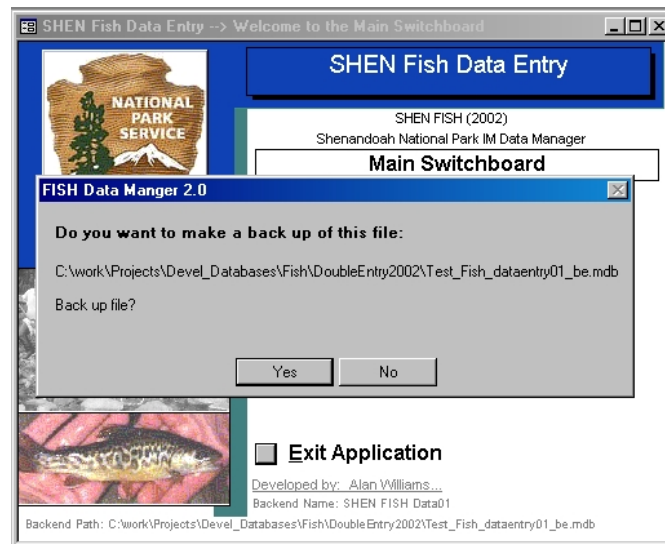
10) Resolve Errors:

Each table will have a custom form. Yellow fields will indicate the problem data. Check the records in question against the data sheets. Keep either the record from the 1st or 2nd phase.

When you have fixed all of the problems the form should close returning you to the “**Checking**” form. Fix all of the tables then press <Done>.

11) Exiting and Backing up:

When you are done entering data for that day you should return to the Switchboard by closing the data entry form using the <Done> button. Then press the <Exit Application> button. You will be asked if you want to make a back-up of a file, press <Yes>. This saves a copy of the data from the network to your local computer.



For help contact the data manager:
Phone: 540-999-3500 x3431

Scanning Field datasheets

All of the field datasheets will be scanned to provide a permanent back up of the raw data. This should be done after the site data has been entered twice (and reviewed) in order to ensure all site and data entry notes are captured for future reference. Use the copier/printer/scanner in the hallway of the NCR Lower building for scanning data sheets. Press the “scan/email” button on the far left. Using the touch-pad, press “T Drive > NCR”. Quality Adjust > Scanner settings: Gray Scale, Resolution= 600, Contrast=Middle, Type=Text/Photo. Load the stack of pages to be scanned face up with the top of the documents to the left side of the scanner (or going into the document feeder head first) making sure corners are not folded and sheets are flat. Press the green button and wait for the prompt that the scan is complete then proceed to the next set of documents to scan.

After documents have been scanned, log into a computer and go to “T Drive” (\\Inpsenhq2\ftp\Scanjobs\NCR). Search for new scanned documents with the scanner generated names like, “0800230291F0-170503112037-0035.pdf”. Open these, confirm the site and date information for the file, close the file then rename. Rename each of the files using the format, “SiteID_YYYY-MM-DD_Date Sampled.pdf” (example, 1F114_2018-07-02.pdf). Double check that all sites that were scanned were saved and re-named. Then, move the pdf files into the O:\Working\AQUATIC\New_Data\Scanned_Datasheets\folder. After all this is completed, file the site folders after the “Scanning Complete – To be Filed” section of the fish data file drawer.

End of Season - Database Tasks and Data Quality Control (QC)

Following data entry, 100% verification, scanning of field datasheets, and renaming of site photos for the current season, a QC data review will be conducted to identify additional potential mistakes. This QC step should include, running a set of predefined QC summaries to catch common errors, omissions, and problems. The summaries will include:

- List of sites entered in the database checked against the crew leader's notes and the schedule for the year.
- List of renamed digital photos checked against photo logs or visited site list
- List of scanned datasheets checked against folders of entered data and database
- Identify out-of-range values
- Search for missing values
- Evaluate outliers

After identified issues are addressed or corrected then the data records can be Labeled “Certified” and the records can be locked by the program lead. Future edits to the data will need to “un-lock” the record and describe why the edit is occurring in the notes, then relock the record after edits are made. Then data in the temporary / seasonal databases are added to the master database using pre-programmed macros (Figure 2). The master database is used for all data summaries and analysis. Scanned datasheets and renamed site photos should then be moved from the “O:\Working\AQUATIC_New_Data” folder into the master directories for each site such as “O:\Working\AQUATIC\Database\Scanned_Datasheets\AQIN\Site_Folders” and “O:\Working\AQUATIC\Aquatic_Images\Site_Folders”.

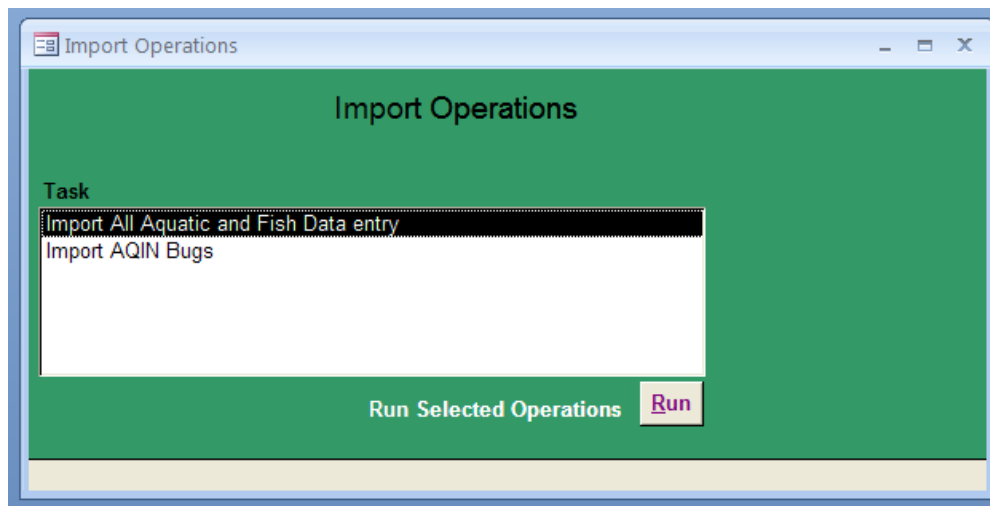


Figure 2. Import data operations form.

Metadata Procedures

Any dataset compiled by SHEN must be accompanied by metadata. This includes both spatial and non-spatial datasets. For metadata associated with geospatial data, we abide by Executive Order 12906, which mandates that every federal agency document all new geospatial data it collects or produces using the Federal Geographic Data Committee (FGDC) Content Standard for Digital Geospatial Metadata (CSDGM; www.fgdc.gov/metadata/contstan.html). All GIS data layers will be documented with applicable FGDC and NPS metadata standards. SHEN will also generate FGDC-style metadata for non-spatial datasets that meet this standard, without the geospatial-specific elements.

SOP Revision History

All protocols, regardless of how sound they are originally, require modification as new and different information becomes available. SOP version numbers increase incrementally by tenths (e.g. version 1.1, version 1.2...etc.) for minor changes. Major revisions should be designated with the next whole number (e.g., version 2.0, 3.0, 4.0 ...). Small changes or additions to existing methods will be reviewed in-house by Shenandoah or Mid-Atlantic Network staff. However, if a complete change in methods is sought, an outside review is required.

Revision History Log:

Prev. Version #	Revision Date	Author	Changes Made	Reason for Change	New Version #