# **Marshall University [Marshall Digital Scholar](http://mds.marshall.edu?utm_source=mds.marshall.edu%2Fetd%2F64&utm_medium=PDF&utm_campaign=PDFCoverPages)**

[Theses, Dissertations and Capstones](http://mds.marshall.edu/etd?utm_source=mds.marshall.edu%2Fetd%2F64&utm_medium=PDF&utm_campaign=PDFCoverPages)

1-1-2003

# Development of a GIS Geodatabase as a Tool for Analyzing Spatial Relationships in the Species Distributions of West Virginia Fishes

Nathan D. Bowe natebowe@gmail.com

Follow this and additional works at: [http://mds.marshall.edu/etd](http://mds.marshall.edu/etd?utm_source=mds.marshall.edu%2Fetd%2F64&utm_medium=PDF&utm_campaign=PDFCoverPages) Part of the [Aquaculture and Fisheries Commons](http://network.bepress.com/hgg/discipline/78?utm_source=mds.marshall.edu%2Fetd%2F64&utm_medium=PDF&utm_campaign=PDFCoverPages)

#### Recommended Citation

Bowe, Nathan D., "Development of a GIS Geodatabase as a Tool for Analyzing Spatial Relationships in the Species Distributions of West Virginia Fishes" (2003). *Theses, Dissertations and Capstones.* Paper 64.

This Thesis is brought to you for free and open access by Marshall Digital Scholar. It has been accepted for inclusion in Theses, Dissertations and Capstones by an authorized administrator of Marshall Digital Scholar. For more information, please contact [zhangj@marshall.edu.](mailto:zhangj@marshall.edu)

# **DEVELOPMENT OF A GIS GEODATABASE AS A TOOL FOR ANALYZING SPATIAL RELATIONSHIPS IN THE SPECIES DISTRIBUTIONS OF WEST VIRGINIA FISHES**

**Thesis submitted to The Graduate College of Marshall University**

**In partial fulfillment of the Requirements for the Degree of Master of Science Physical Science: Geobiophysical Modeling**

**by**

**Nathan D. Bowe**

**Dr. Michael L. Little, Committee Chairperson Dr. Ralph E. Oberly, Committee Member Dr. Ralph Taylor, Committee Member**

> **Marshall University Huntington, West Virginia**

> > **December 8, 2003**

## **ABSTRACT**

# **Development of a GIS Geodatabase as a Tool for Analyzing Spatial Relationships in the Species Distributions of West Virginia Fishes**

## **By Nathan D. Bowe**

One of the most complete references to date of fish species distribution in West Virginia is a 1995 book entitled "The Fishes of West Virginia" (Stauffer, et al.). In this project, geographic information systems (GIS) and relational database technology have been utilized to adapt that reference into a system where spatially arranged collection site features are related to distribution data through a series of common fields among tables in a geodatabase. The geodatabase is stored in an MDB-formatted database management system (DBMS) which is readable by Microsoft Access and useful in its own right as a means to query distribution data when there is no need for a mapping environment. A search form was built-in to assist in the most common search parameters, and results display in print-formatted reports. Much of the same search potential is achieved by loading the geodatabase into the GIS package ArcGIS®. In addition to traditional query power, the geodatabase houses the shapefile data layers used in mapping and spatial analysis. This system has been applied in a preliminary assessment to examine distribution data versus land use for four small streams under study by Marshall University.

# **TABLE OF CONTENTS**





# **LIST OF FIGURES**



## **CHAPTER I -Introduction**

#### **Background**

Ongoing studies assessing diversity of West Virginia fisheries are dependent on advancements in distribution data management. New technologies such as GIS and global positioning system (GPS) enable researchers to collect various data at higher precision and accuracy. Increasing accuracy of GIS spatial data sets such as land cover and elevation, coupled with better collection data, facilitates the combination of all data into a system where the highest number of statistical operations can be made. A system of this type allows researchers to relate spatial distributions of fish species that were difficult or impossible to process beforehand.

In this work, a geodatabase was developed as the initial step in the development of a comprehensive data system for West Virginia fishes. It was previously possible to store relational data in a standard relational database, but the associated GIS data was always file-based (McCain 2003). An "open GIS" movement was spawned shortly after the arrival of the first all-relational models. New models became capable of storing both spatial and attribute data in a relational database when standards organizations such as the Open GIS Consortium (OGC), the International Organization for Standardization, and the U.S. Federal Geographic Data Committee began promoting the idea of data sharing through spatial data standards ("Spatial Data Standards…" 2003). Environmental Systems Research Institute (ESRI) actively participated in these developments, and so the geodatabase structure was born. The geodatabase developed here houses tabular species distribution data along with spatial data layers of collection

sites, rivers, streams and lakes for the southern part of West Virginia. Discussed in this paper are the reasoning and methodology behind the system's creation, how to navigate the geodatabase in Microsoft Access and ESRI's ArcGIS for data retrieval, directions for adding new collection data, and a basic example of the system's use.

#### **GIS Capabilities**

The desktop mapping systems on the market today range from display-only systems such as electronic atlases to full-featured GIS applications. The dividing lines between one type of system and the next are not sharply defined. The systems do differ in a number of important ways: how they link geographic locations with information about those locations, the accuracy with which they specify geographic locations, the level of analysis they perform, and the way they present information as graphic drawings. Electronic atlases, for instance, allow you to display pictures of geographic areas on your computer screen. They provide limited information about the geographic areas, and limited ability to alter graphics. Without any tools for analyzing the information, these systems are most useful for providing graphics that can be used in presentations and reports.

Desktop GIS has the ability to analyze geographic location and the information linked to those locations. It is dynamic; maps can be created that are not limited to a single moment in time, information can be linked to a map, and data can be visualized and analyzed in new ways, revealing previously hidden relationships, patterns and trends. GIS supports formulation of a conceptual site model and evaluation of projected present and future use scenarios. Many valuable analysis and visualization algorithms are available as standard tools (Dunn & Klausmeier 2002).

#### **Study Area**

This project focuses on distribution data from the southern section of West Virginia, which has been defined for this purpose as the portion of the state south of (and including) the Little Kanawha, Elk and Greenbrier river drainages (Appendix A).

#### Big Sandy River/Tug Fork

Tug Fork arises in southern West Virginia and flows west to the Big Sandy River, which comprises a section of the West Virginia state boundary. The 196-km stretch is commonly considered as one stream. The Big Sandy River/Tug Fork drains approximately 11,092 km<sup>2</sup> of the Appalachian Plateau Physiographic Province (Stauffer et al. 1995).

#### Elk River

The Elk River, a major tributary of the Kanawha River, arises in Pocahontas County and flows 290 km to its confluence with the Kanawha River near Charleston, West Virginia. In the Elk River basin is a 600-ha impoundment, Sutton Lake. The Elk River drains approximately 3,968 km<sup>2</sup> of the unglaciated portion of the Appalachian Plateau Physiographic Province (Stauffer et al. 1995).

#### Gauley River

The Gauley River, a major tributary of the New River, originates on Gauley Mountain in northwest Pocahontas County and flows 171 km to its confluence with the New River at Gauley Bridge, West Virginia. Gauley River drains about 3,682 km<sup>2</sup> and its basin contains an 1100-ha impoundment, Summersville Lake (Stauffer et al. 1995).

#### Greenbrier River

The Greenbrier River, a major tributary of the New River, arises in Pocahontas County and drains approximately 4,289  $\text{km}^2$ . It flows south for 233 km, where it converges with the New River near Hinton, West Virginia (Stauffer et al. 1995).

#### Guyandotte River

The Guyandotte River arises in southwest Raleigh County and flows 268 km to its confluence with the Ohio River near Huntington, West Virginia. It drains about 4,348  $\text{km}^2$  of the Appalachian Plateau Physiographic Province. The Guyandotte River is characterized by narrow valley floors, and its basin holds R. D. Bailey Lake, a 255-ha impoundment (Stauffer et al. 1995).

#### James River

The James River drainage is represented by two small stream systems in West Virginia. Potts Creek originates at the confluence of the North and South Forks of Potts Creek and drains a total area of  $143.7 \text{ km}^2$ . Cove Creek arises in Monroe County near Wiley Church and drains a total of  $32.1 \text{ km}^2$ . All of the West Virginia tributaries of the James River are characterized by steep gradients, ranging from 17.3 m/km in Cove Creek to 41.6 m/km in the North Fork of Potts Creek (Stauffer et al. 1995).

#### Kanawha River

The Kanawha River begins at the confluence of the Gauley and New Rivers just upstream of Kanawha Falls. It flows 149 km to the Ohio River, draining 10,052 km<sup>2</sup> downstream of Kanawha Falls. The Kanawha River supports large commercial barge traffic and is controlled by three U.S. Army Corps of Engineers lock and dam facilities (Stauffer et al. 1995).

4

#### Little Kanawha River

The Little Kanawha River arises in Upshur County in north-central West Virginia and flows 270 km to its confluence with the Ohio River near Parkersburg, West Virginia. In the basin is a 392-ha impoundment, Burnsville Lake. The basin drains a total of 5,729  $\text{km}^2$  of the Appalachian Plateau Physiographic Province; it is characterized by steep hills and narrow valleys in the upstream areas and broad valleys near its confluence with the Ohio River (Stauffer et al. 1995).

#### New River

The New River arises in North Carolina, flows north through Virginia, and meanders approximately 140 km through West Virginia, where it joins with the Gauley River to form the Kanawha River. The New River drains  $17,918$  km<sup>2</sup>. In West Virginia, the river has been designated a Wild and Scenic River and the basin is completely in the Appalachian Plateau Physiographic Province. Of particular interest is the New River Gorge, which begins at Sandstone Falls below Hinton, West Virginia. The gorge contains a series of waterfalls and rapids, admired by rafting and kayaking enthusiasts.

The Bluestone River, a major tributary of the New River, arises in Virginia and extends into Mercer and Summers counties in West Virginia. It drains a total of 969 km<sup>2</sup>, and contains in its basin an 830-ha impoundment, Bluestone Lake (Stauffer et al. 1995).

#### Ohio River

Most of West Virginia, comprising approximately 75% of the total area of the state, is drained by the Ohio River. The Ohio River drains approximately  $528,127 \text{ km}^2$ ,  $47,182 \text{ km}^2$  of which are in West Virginia. The main channel of the Ohio River begins

at the confluence of the Monongahela and Allegheny Rivers at Pittsburgh, Pennsylvania. It flows 981 km to its confluence with the Mississippi River in Illinois. A 443.2-km stretch of the main channel forms the western border of West Virginia. This portion of the river contains seven lock and dam facilities, operated the U.S. Army Corps of Engineers.

The Ohio River basin drains the Appalachian Plateau Physiographic Province, and is characterized by steep valleys and narrow flood plains. Raw materials found in the drainage include limestone, sandstone, sand, gravel, brine, petroleum, natural gas and coal. Two major impoundments, East Lynn Lake and Beech Fork Lake, are in the Twelvepole Creek watershed and occupy a total of 700 ha (Stauffer et al. 1995).

#### **Problem Statement**

The immediate need for this work stems from a research project at Marshall University wherein fish movement/migration is compared to natural stream structure and roadway culverts. The goal of the project is to determine the negative effects (if any) of particular culvert designs on the movement of fishes. New collections are being taken periodically, but researchers need to know what species have been collected on certain streams in the past. The best source for these data is "The Fishes of West Virginia" (Stauffer, et al. 1995), but it is very time-consuming to analyze distribution from paper maps in a book. Thus, this system, that adapts data from the book into digital format, has been developed. It also provides a structure that will be useful for appending and analyzing collection records in the future.

6

#### **CHAPTER II – Base Map Methodology**

#### **Collection of Data Layers**

 Most GIS data for this project were gathered from the West Virginia GIS Data Center (http://wvgis.wvu.edu). Sets from this source (Appendix D) include major water features (major rivers and lakes), watershed boundaries, West Virginia county boundaries, cities from the Geographic Names Information System (GNIS), mining permit polygons, underground mining limit polygons, valley fill polygons, and the National Land Cover Dataset (NLCD). Detailed river and stream data layers were collected from ESRI's "Data & Maps" 2002 and 2003 CDROM series (Appendix D). Additional layers that were used for quality control purposes (and not incorporated into the project) include DOQQs and digitized topographic maps stored by Marshall University.

#### **Data Layer Manipulation**

Some data sets required for this project were not prearranged in a suitable format; consequently, they were reformed to an appropriate level of detail. Often, a set must be manipulated in one fashion or another. A shapefile may hold more detail than is required and warrants merging or dissolving features together. Two data sets may collectively represent a feature better than either set does on its own; one set may gain better accuracy when edited with another, more accurate set (McCoy  $&$  Johnston 2002). In any case, decisions must be made on how to maneuver inside, around and among various sets of spatial data. This section summarizes the decisions made concerning the data used in this project.

7

#### Rivers and Streams

Although the 2003 ESRI data set is comprised of many more feature parts per feature (which makes analysis more difficult by presenting a more complex picture than required), the overall alignments are more accurate than those of the 2002 ESRI set. In cases that were problematic due to large missing segments (from 2002 set) such as Little Coal River and Pocatalico River, the improvement is substantial. However, the 2002 data set—derived from TIGER/Line files—provides more accurate and complete stream naming. Therefore, spatial alignments were corrected to the 2003 set as well as DOQQ imagery, and the names were checked against the 2002 set and digitized topographic maps as necessary.



**Figure 1: Meadow River segments from ESRI Data & Maps 2002.**



**Figure 2: Meadow River segments from ESRI Data & Maps 2003.**

For example: Meadow River (Figures 1 and 2), a tributary of Gauley River, had no named records in the 2003 set. The 2002 set had named records (so the general location was known) but the alignment was badly segmented; also, the alignment nearest the Gauley was totally missing. To solve this issue, both 2002 and 2003 sets were displayed together as layers in ArcMap. Using "Select by Attributes", all segments with "Meadow River" in field "Name" were selected in the 2002 set. Upon zooming to the selected region, it was obvious where the missing segments should align based on the improved spatial accuracy of the 2003 set. All segments representing Meadow River in the 2003 set (although unnamed) were selected and exported, and included in the Gauley River set. Segments of Meadow River were merged to one with the Editor Toolbar, and proper naming was added to the attribute table.

For those alignments that were already isolated based on the 2002 set (before the 2003 set became available), a variation of the aforementioned process for Meadow River was performed, where the spatial accuracy was checked against the 2003 set. For tributaries not completed as of receiving the 2003 set, alignments were isolated directly from that set and then naming was verified based on the 2002 set and topographic maps.

The variation of alignment updating is as follows:

Where streams isolated from the 2002 set were accurate except for some missing segments, a union operation was performed to add those missing segments from the 2003 set. Then, a merge operation joined all segments to one record and the naming was added. Extraneous fields in the attribute table, some of which were no longer accurate after the editing process, were deleted.

Where streams isolated from the 2002 data set were inaccurate in some places (i.e. segments branch off instead of maintaining only one alignment), all segments were imported from the 2003 data set using the paper maps (Stauffer et al. 1995) and the 2002 data set as templates only. Then, those segments were merged and names obtained from the 2002 set. Extraneous fields in the attribute table, some of which were no longer accurate, were deleted as a proactive approach to database quality control (McCain 2003).

#### Watersheds

 Sub-watersheds found in the original watershed shapefile (Appendix D) were merged together with the Editor Toolbar into the twelve main watersheds in the project (Appendix A). The "FNAME" field in the watershed attribute table, which assigns each sub-watershed to its parent watershed, was used to select all feature parts for an individual watershed. The Editor Toolbar's "Merge" command combined the parts into one polygon, and extraneous fields were deleted because of tabular inaccuracies due to the merge operation.

#### Conversion to Raster

The large, numbered collection site map (Stauffer et al. 1995) was scanned to the Tiff raster format. Because the map was black-and-white, the resultant was a raster with two pixel values, 0 and 1. By setting the pixel value representing white to a null color, the image could be overlaid as black text, points and lines on any GIS layer. However, the image was not in a geographic coordinate space until it was loaded and georeferenced (McCoy & Johnston 2002) to at least 20 control points in UTM Zone 17 North – mostly unique state boundary geography such as points and angles, but also stream intersections from finished tributary shapefiles. The Georeferencing Toolbar in

10

ArcMap was used for this process. A shapefile of the West Virginia outline (Appendix D) was used as the template spatial alignment:



**Figure 3: Raster collection site map georeferenced to digital layers.**

To georeference a raster dataset from image space to a real-world coordinate system, the location of recognizable features in both coordinate spaces must be known. These locations are used to create control points. The control points are used to build a polynomial transformation that will warp the image from one coordinate space to another.

Control points are locations that can be accurately identified on the raster dataset and in real-world coordinates. These identifiable locations may be road and stream intersections, building corners, bridges, the mouth of a stream, rock outcrops, and identifiable points on geometric landscape features such as the end of a jetty of land.

For each control point selected on the input raster dataset, the output location may be specified either by graphically selecting a point that is already in the desired output coordinate system (as in this project) or by typing in the known output coordinates. The relationship between the control points chosen in the raster dataset and the output coordinate space is then determined.

Using this relationship and a polynomial transformation, the raster dataset is converted from nonreal-world space to real-world space. A polynomial transformation computed using the specified control points is applied so the input locations approximate the specified output using a least-fit. This polynomial transformation yields two formulas, one for computing the output x-coordinate for an input  $(x, y)$  location and one for computing the y-coordinate for an input  $(x, y)$  location. The goal of the least-fit is to derive a general formula that can be applied to all points, usually at the expense of slight movement of the to-positions of the control points. The more control points of equal quality used, the more accurately the polynomial can convert the input data to output coordinates.

The cells of a raster dataset will always be square and of equal area with respect to the Cartesian coordinate system (map coordinate space) associated with the raster dataset. The shape and area a cell represents on the surface of the earth will never be

constant across a raster dataset. Since the area represented (on the face of the earth) by the cells will vary across the raster dataset, the output cell size and the number of rows and columns may change when projected.

Converting from one projection to another can also change the shape and area a cell represents on the surface of the earth. Each projection treats the relationship between a three-dimensional world and a two-dimensional one differently.

When displaying and performing analysis with raster datasets, they should be in the same coordinate space and in the same projection. If two raster datasets are in different coordinate systems, the values of the coordinates are on different scales. Errors will occur when comparing such datasets because they will represent different locations.

Upon inspecting the raster dataset sample points in comparison to the watershed and tributary data, a new method of quality control was noted. The scanned data points and alignments matched well enough with the isolated watershed tributaries that they were used in a reactive approach to further correct alignments previously chosen by eye from the source map prior to scan.

#### Collection Sites as a GIS Layer (Vectorization)

In many cases, an automatic vectorization operation can be performed to create vectors from a source raster backdrop. However, the raster image of the collection sites consisted of additional text (site numbers) and lines (drawn streams) that would have contributed to a sloppy vectorization. For this reason it had been decided to vectorize the point sites by hand.

ArcCatalog is an application within ArcGIS Desktop that helps to organize and manage all GIS data; this includes the construction of new files in various GIS data formats, such as shapefiles, geodatabases, tables and layers. A new point shapefile was created in ArcCatalog and assigned the UTM Zone 17 North spatial reference. The shapefile was loaded in ArcMap and an editing session was started with it as the target. Before adding the new sample points to the shapefile, "snapping" was set to the stream layers. By doing so, it was ensured that new sites would gravitate toward streams when added—another method of quality control. The point site shapefile was also set to be the only selectable layer for the editing process.

The collection points were added with the pencil tool from the Editor Toolbar. Site numbers and descriptive fields (see "Sites Table", page 16) were added to the attribute table for the shapefile and filled as the procedure progressed. Numerous checks were made during the creation of the new data, and errors are expected to be unlikely due to the number of times each point was examined.



**Figure 4: Closeup of vector point collection sites.**

# **CHAPTER III – Geodatabase Methodology**

#### **Data Entry and Table Design**

The data are contained entirely in three tables: Fishes, Sites, and Collections, related by common fields. Microsoft Excel was used for data entry in the Fishes and Collections tables, and they were subsequently saved in the dBase IV file format (\*.dbf). Sites data were entered in the attribute table of the collection sites shapefile, which is automatically saved into a separate dBase file by ArcMap. The following section outlines the design of and data types contained in each table:

#### Fishes Table

The Fishes table contains indexed taxonomic and common names of fish species known to exist within the state of West Virginia as listed (Stauffer et al.). Fields include:

*FishID* – Indexed identification number assigned to each discrete fish species. It is formatted as F001, F002, F003, and so on. There is no repetition of records in this field and it is used as the primary key of the table.

*Family* – Taxonomic family name of fish, as a text string. This field may have duplicates.

*Genus* – Taxonomic genus name of fish, as a text string. This field may have duplicates.

*Species* – Taxonomic species name of fish, as a text string. This field does not have duplicates. Subspecies and hybrids are formatted as '[Sp]. [SUBSPECIES]' and '[SPECIES] x [SPECIES]', respectively.

*Common* – Common name of fish species, as a text string. This field does not have duplicates.

#### Sites Table

The Sites table originated as the attribute table for the collection sites shapefile. Descriptions of fields in this table are as follows:

*SiteID* – Indexed identification number assigned to each discrete collection site, as a number. It is formatted to match the site numbers taken from the foldout map (Stauffer et al.). There are cases where sites found on individual species maps are absent from the large foldout map. Those sites were assigned identification numbers starting at 2000, 2001, 2002, and so on. They were designated in order of when they were discovered. There are other cases in which a species is presumed to be in a region, or 'zone' of the state. These instances have been described on a case-by-case basis, and assigned identification numbers starting at 3000, 3001, and so on. This field has no duplicates, and is used as the primary key of the table.

*Location* – Description of the unique physical location of the collection site, as a text string. This field contains no duplicates. When the location is described using its position relative to a stream or city, it is always referring to a stream or city that exists in the data sets (Appendix D) used in this project. Identifiers such as 'above' or 'below' typically signify a location that is within one mile above or below a

feature. 'Near' implies that the closest municipality to the collection site is not streamside.

*County* – List of the county name in which collection site is located, as a text string. The field may contain duplicates, and in some cases holds more than one county name (e.g., sometimes the sample site lies on a county boundary.)

*Drainage* – Name of the major drainage (Appendix A) in which the collection site exists, as a text string. The field may contain duplicates.

*NearestRiv* (Nearest River) – This field holds the name of the larger stream one to three levels downstream of the collection site, as a text string. The motivation for the creation of this field was to boost query capability by narrowing the scope within a major watershed. Querying this field for "Mud River", for example, will return records from all immediate tributaries of the Mud River. Whenever the collection was made in the channel of one of the larger rivers (i.e. any stream named as a 'river'), this field will contain the name of that river. Because of this, users will have all sites within Mud River returned in addition to all its tributaries. The field will not contain the name of the next larger stream in these cases because the major rivers are large enough that the returned list of sites will often exceed those sites sought by users. This field may contain duplicates.

*StreamName* – Name of the stream where the collection took place, as a text string. This field may have duplicates.

#### Collections Table

The Collections table contains records relating collection sites per fish species. Primary references for this data were the individual species maps (Stauffer et al. 1995). It is the vital link that relates each fish species to its corresponding collection sites throughout southern West Virginia.

*FishID* – Individual fish species identification number, related to the field of the same name in the Fishes table.

*SiteID* – Individual collection site identification number, related to the field of the same name in the Sites table.

*Comments* – Description of the fish collection. Most sites are described as 'point sites', which mean that the SiteID number refers an actual point that exists in the species maps (Stauffer et al.). Other descriptive terms include zonal identifiers such as 'statewide' or 'New River drainage only'.

#### **Creating the Geodatabase**

In contrast to file-based formats (such as shapefiles, which store feature and attribute information in separate files) the geodatabase has the ability to store several types of information in an off-the-shelf database management system (DBMS) ("Working with the Geodatabase…" 2002). All crucial files for the project were eventually bundled into this geodatabase system, thereby condensing the project into one file available to different operations. For this project, the geodatabase is stored in MBD, the format native to Microsoft Access. It is technically referred to as a personal geodatabase by ArcGIS, a delineation that does not require the purchase of additional software programs to access or manage it ("Working with the Geodatabase…" 2002). However, it is sometimes desirable to gather specific tabular data without the need for a map interface. To accommodate this need, particular additions have been designed inside MS Access as described in the next

section of this document. A new personal geodatabase was produced in ArcCatalog by right-clicking in the "Contents" window, and the pertinent data were imported in the same fashion: collection site, river, lake and tributary shapefiles, along with the





Fishes and Collections dBase tables.

#### **Incorporation to a Traditional Relational Database**

The geodatabase (since it is an MDB file) can be directly opened by Access and modified within the application without affecting its interaction with ArcGIS ("Working with the Geodatabase..." 2002), as long as tables automatically created by ArcGIS are not manipulated or deleted. Most of the ArcGIS table names begin with "GDB\_".

## Description of Relationships

The record relationships (cardinality) between tables describe how individual record values relate to each other. Fishes and Collections have a one-to-many cardinality, which means that individual records from Fishes may be repeated for many related records in Collections. Collections and Sites are many-to-one, meaning that many records from Collections may reference a singular related record in Sites. Cardinality, while not as important inside the Access environment for this work, was vital in determining which table associations should be performed later in ArcGIS.



**Figure 6: Relationships among tables as seen in Access.**

#### Database Add-ons

## *Search Form*

Each text box in the search form is loaded with values from a select query – the "Nearest River" dropdown contains all unique values from the field of the same name from the Site table. The "Genus/Species" dropdown displays each unique FishID, Genus, Species and Common record. When the user selects a record from either of the dropdowns, it provides the criteria for a parameter query. Clicking on the "GO" button for either type of search activates the event procedure that displays a report based on the selected criteria. For example, selecting "Cherry River" from "Nearest River" will display a report containing information for every collection site on the Cherry River and its tributaries. Selecting "F112" from "Genus/Species" will return a report outlining information about every site where yellow bullheads were collected. The search form displays immediately when the geodatabase is opened in Access.



## *Reports*

MS Access reports are loaded with data borne from the results of parameter queries and activated (displayed) by the search form. Each is formatted for standard printing.



**Figure 8: Results from search form displayed in Access reports.**

## **CHAPTER IV – Geodatabase/GIS Interaction**

#### **ArcGIS Geodatabase Methodology**

Much of the same search parameters used in Access can be applied in ArcMap. However, there are some differences in procedure for the GIS end of this project that must be addressed beforehand. Most notably, relationships between tables must be forged differently in each software package. Data retrieval in ArcMap also requires an adequate GIS background, whereas users with no experience can simply choose criteria from the search form in Access.

#### Joins and Relates

ArcMap provides two methods to associate data stored in tables with geographic features: joins and relates. When two tables are joined, you append the attributes from one onto the other based on a field common to both tables. When you relate tables, a relationship is defined between the two tables – also based on a common field – but the attributes of one is not appended to the other. Instead, the related data can be accessed when necessary. Joins are designed for tables with one-to-one or one-to-many cardinality. For other cardinalities such as many-to-one or many-to-many, using a join will omit all records after the first match for each primary key value – a relate must be made instead.

In ArcGIS, similar relationships to those in Access must be prepared between 'Fishes' and 'Collections' as well as 'Sites' and 'Collections'. The procedure described here is slightly more complicated than in Access, and the more expensive licenses of ArcGIS, because the basic version of ArcGIS (ArcView 8.2) was used for the operation. More information about this choice can be found in chapter six (page 37).

#### Procedure

The relationships described herein must be reestablished each time the geodatabase is loaded into a new map. Repetition of this process can be avoided by saving the map document with the geodatabase loaded into the data frame or by saving a layer file for the sites shapefile included in the geodatabase (after the relationships have been made). Fishes and Collections have a one-to-many cardinality, while Collections and Sites have a many-to-one relationship. Thus, Fishes can be *joined* to Collections and the resulting temporary table "Collections\_Fishes" *related* to Sites. (Joins must always be performed before relates.) The most basic method, however, employs *relates* for both relationships. The latter choice will be discussed here, because a scenario of data retrieval exists that would benefit (species by site). To set up the relationships:

- (1) Open ArcMap and add the following files from the geodatabase Sites, Collections, and Fishes. Make sure the table of contents is set to "source" (otherwise, the Collections and Fishes tables will not be visible).
- (2) Right-click on Collections and choose "Joins and Relates >> Relate…" This brings up the relate dialog box. Fill in the dialog box to match Figure 9 and click OK (name of relate does not matter). Data from Fishes is related to Collections via the FishID field.
- (3) To continue this relationship to the site data, right-click on Sites and choose "Joins and Relates >> Relate…" This brings up the relate dialog box again.
- (4) Fill in the relate dialog box as shown in Figure 10 and click OK. Data from Collections (and consequently from Fishes) is related to Sites via the SiteID field. (Once again, the name of relate does not matter.) Data retrieval is enabled.

Note: After this process is completed, it would be a good idea to save a map document or layer file as explained in the introduction to this section.



**Figure 9: Relating Fishes to Collections in ArcMap.**



**Figure 10: Relating Collections to Sites in ArcMap.**

### **Retrieval of Records**

This section is a step-by-step user guide for isolating specific site or collection

records from the geodatabase in the ArcMap 8.x environment:

#### Species by Site

*Scenario: The site location is known, and the user wishes to list all species collected.*

*Option 1, the single site point-and-click "Identify" method:*

- 1) With the geodatabase loaded in ArcMap, choose the Identify tool  $\bullet$ . This changes the cursor to an Identify pointer. Click on the site point. This opens the Identify Results window, which displays attributes of the site.
- 2) Navigate through the left side of the results window to reveal collection records.



**Figure 11: Identify results window.**

*Option 2, selecting one or more sites through the Sites attribute table:*

- 1) Open the Sites attribute table. Select the site (or sites) in question, either by hand or "Options >> Select by Attributes". Selected records are blue by default.
- 2) In the Sites attribute table, navigate to "Options >> Related Tables >> Collections". This opens Collections and automatically selects all records related to the chosen sites.

3) In the Collections attribute table, navigate to "Options >> Related Tables >> Fishes". This opens Fishes and selects all species collected at the chosen sites.



4) Click on "Show: Selected" to show only the species collected (Figure 12).

#### **Figure 12: Selected attributes of Fishes.**

5) To export the results as a dBase table or text file, navigate to "Options >> Export", choose file name, and save. (Make sure to choose to export "selected records" unless the entire species list is desired.)

#### All Sites of Specific Species

Whereas a potential user can choose a site(s) and find all species collected there, it is also possible to choose a species (or multiple) and find all its collection sites. The process is similar to "Species by Site" option 2, with the only difference being that the user begins by selecting records from Fishes instead of from Sites. After selecting the desired (single or multiple) species in Fishes, navigate through "Options >> Related Tables >> Collections". Then, in Collections, proceed to "Options >> Related Tables >> Sites". The sites that satisfy the user's query will automatically highlight both on the map and in the Sites attribute table.

#### Combinational Queries

As in a RDMS such as Access, data can be segregated by more than one field in more than one related table. The same process takes multiple steps in ArcMap using the geodatabase system developed here.

#### *Scenario: How many collection sites in the Guyandotte drainage held rainbow darters?*

- 1) Select the record for rainbow darter in the attribute table of Fishes either by choosing it from the list or "Options >> Select by Attributes"
- 2) Navigate through "Options >> Related Tables" (as in "All Sites of Specific Species") until the rainbow darter collection sites are highlighted.
- 3) In the Sites attribute table, navigate to "Options >> Select by Attributes". Ensure that Method = "Select from current selection" and formulate the expression, [Drainage] = 'Guyandotte' (Figure 13). Rainbow darter collection sites within the Guyandotte river system are highlighted.



Note: This is one example of a combinational query. Permutations of this procedure will allow the user to filter data in other ways, providing that queries originate in the attribute tables of either Sites or Fishes. The collections table is a "bridge" between the two, and only useful for querying if either the SiteID or FishID is known.

**Figure 13: Select by Attributes dialog window.**

#### Tabular Export

Any tabular records can be exported as a text file or dBase table. Also, selected records (blue) can be exported independently of the dataset as a whole. With the desired table open in ArcMap, navigate to "Options >> Export…" Choose either the selected records or all records and the preferred path/filename for the export file (Figure 14). Click

OK to finish the procedure.



**Figure 14: Export tabular data dialog window.**

#### **Appending the Collection Data**

One of the advantages of the geodatabase structure is that geometry is stored as a tabular attribute (Joselyn 2002) in the MDB – a feature that allows for the manipulation of the database in either Access or ArcMap. Still, a potential user must account for the type of data that will be appended. New collections at old sites can be added in Access or ArcMap, with no mapping required. New collection sites, however, must include the spatial characteristics of the points which the user cannot edit in Access. Therefore, new collection site vectors must be added in ArcMap first. Then, tabular data can be appended in either Access or ArcMap.

*Scenario: After revisiting sites already in the geodatabase, new collection records are ready to be added.*

Before adding new records, use the Fishes table to find the FishID of those collected and the Sites table to find the SiteID of the corresponding collection sites.

*Option 1, in Access:*

- 1) After geodatabase is loaded and the search form pops up, navigate to "Window >> Unhide…" on the main menu bar. Choose to unhide the database window.
- 2) Double-click on Collections in the "Tables" section to open it.
- 3) Scroll to the bottom and enter each FishID/SiteID pair, as well as any comments.
- 4) Save and exit.

*Option 2, in ArcMap:*

- 1) Load only the Collections table from the geodatabase.
- 2) If the Editor Toolbar is not visible, navigate to "View >> Toolbars >> Editor" to open it (Figure 15).



- 3) Click "Editor >> Start Editing". The "Target" should be blank as in Figure 15.
- 4) Open Collections and scroll to the bottom. There will be a blank row where new records can be added. Enter the FishID/SiteID pairs.
- 5) When finished, select "Editor >> Stop Editing" and click "Yes" when asked to save edits.

**Figure 15: Editor Toolbar.**
*Scenario: New collections are taken at new sites.*

*In ArcMap:*

- 1) Load the Sites and all water shapefiles and open the Editor Toolbar.
- 2) Choose "Editor >> Start Editing" and make sure the target is set to Sites.
- 3) Navigate to "Editor >> Snapping…" and set snapping to the edge of major river and tributaries (Figure 16).



**Figure 16: Snapping Environment window.**

- 4) Choose the "Create New Feature" tool  $\blacksquare$  and add new sites. By changing "View >> Data Frame Properties >> General >> Units >> Display" to decimal degrees or degrees minutes seconds, the real-time coordinates of the pointer will be displayed at the bottom of the screen as it is scrolled across the data frame.
- 5) For each site, add descriptive data (especially SiteID) into Site's attribute table.
- 6) When finished adding sites, go to "Editor >> Stop Editing" and choose "Yes" when asked to save edits.
- 7) To associate new collections of species to the new sites, update Collections using either of the two methods described in the previous scenario (page 30).

# **CHAPTER V – Distribution and Land Use**

#### **Overview**

Four small streams were chosen by researchers at Marshall University for a study of the effects of highway development – culverts, in particular – on the migration of fishes. Each stream was selected because of the predominant land use (from land covers "agricultural", "developed", or "mining") of its area:

- 1) Agricultural
	- Whites Creek of Big Sandy River (3 collection sites)
- 2) Developed (urban/suburban)
	- Hurricane Creek of Kanawha River (1 site)
- 3) Coal Mining
	- Huff Creek of Guyandotte River (1 site)
	- Pinnacle Creek of Guyandotte River (2 sites)

To test whether each stream is truly representative of its intended land use, the geodatabase developed in this project was applied to (1) query the list of fish species collected in each stream through Access, (2) highlight all sites throughout southern West Virginia where each species was collected, and (3) compare the results with spatial data layers, chiefly the NLCD (Appendix D). See Appendix C for images used in the analysis described in this chapter. It should be noted that especially common species (distributions described as "statewide") were not included in the investigation. The NLCD is a satellite sensor-derived 30-meter ground resolution raster image (Appendix D). Pixel values were assigned to a class system (Appendix D) using unsupervised classification by the U.S.

Geological Survey (USGS), thereby enlisting computers to detect pixel value ranges representing homogeneous features. The system was designed to function primarily by the interpretation of remote sensor data, without collecting data *in situ* (Jensen 1996). SQL commands through Raster Calculator (McCoy & Johnston 2002) were used to create three tiff rasters, isolating the pixel value range representing each of the three land covers in question. These rasters were converted to 1000-meter cell size vector features with Spatial Analyst, which generalized the land cover locations (McCoy & Johnston 2002). A new polygon shapefile was prepared in ArcCatalog, and polygons were created that encompass each predominant land cover (most prevalent of the three land covers).



**Figure 17: Generalized land cover of West Virginia.**

Low spatial accuracy of the original site map and georeferencing inaccuracies due to the digital conversion process enables this broad generalization. Future collections appended to the geodatabase should be analyzed using datasets of higher resolution.

#### Agricultural

Three collection sites on Whites Creek yielded the following species: *Etheostoma caeruleum* (rainbow darter), *Labidesthes sicculus* (brook silverside), *Lampetra aepyptera* (least brook lamprey), and *Lythrurus umbratilus* (redfin shiner).

As Figure 17 illustrates, there are two large zones of agricultural land cover on either side of the mining zone. All species with exception of *E. caeruleum* were found to exist is streams with predominantly agricultural land cover. *L. umbratilus* collections were taken 95% of 38 times in the agricultural zones on the west side of the state. *L. sicculus* and *L. aepyptera* were collected in agricultural zones 84% and 86%, respectively. All but two sites are on the western side of the mining zone. The one exception to this pattern, *E. caeruleum*, seems to be correlated (60% of 147 sites) with areas of mine activity in the state. This is not reason for surprise, however, for Whites Creek flows on the outskirts of mining areas. In fact, there is possible mining activity near the mouth of Whites Creek, according to the NLCD and mining permit records. Also, 35% of *E. caeruleum* collections were in agricultural areas.

It is suggested by these results that Whites Creek is an acceptable choice to represent a stream draining primarily agricultural land. However, this study also proposes that Whites Creek be used in characterizing the western agricultural land of West Virginia and that a stream in perhaps the Greenbrier watershed would be more informative in characterizing the eastern side of the state.

#### Developed

One collection site on Hurricane Creek yielded a single species, *Percopsis omiscomaycus* (trout-perch). Of 76 sites where this fish was collected, 20% are in one of the outlined zones of development (urban/suburban). The other 80% lies on western agricultural land. Approximately the lower half of Hurricane Creek meanders through sporadic farmland. Based on the limited number of non-"statewide" fish records on this stream, it is difficult to accept or reject Hurricane Creek as representatively "developed". It can be inferred from the NLCD, however, that it flows through a more transitional land cover than the other streams in the study. A stream closer to the Charleston area may show to have a more uniformly urban/suburban watershed.

#### Mined

One collection site on Huff Creek yielded a single species, *Etheostoma caeruleum* (rainbow darter). Two sites on Pinnacle Creek yielded three species:*Clinostomus funduloides* (rosyside dace), *Etheostoma caeruleum* (rainbow darter), and *Etheostoma variatum* (variegate darter).

Both streams drain to the Guyandotte River and have similar dense spatial patterns of mining activity based on the NLCD and shapefiles of mining permits. They are approximately forty miles apart, Pinnacle Creek above Huff Creek. As stated in the "Agricultural" section, *E. caeruleum* collections were made in mining areas for 60% of

147 sites. Of 111 *E. variatum* collections, 46% were from the mining zone; most taken outside the zone were found in the Little Kanawha system (32% of total). *C. funduloides* collections were in the generalized mining area 23% of 26 times and in eastern agricultural use for 77%. Due to the combination of land cover and relative fish distribution, this study suggests that Huff and Pinnacle creeks are suitable for representing mining-impacted streams.

# **Chapter VI – Conclusions**

This project has been designed as a straightforward solution for querying and analyzing fish distribution records. A potential user with modest GIS background should be able to obtain any requested information from the system with little difficulty. In this respect, the project is a success. As with any software-related endeavor, however, one can always depend on the need for updating and upgrading the system. Software versions change over time and ultimately become obsolete; us ers tend to prefer more and more simplified graphical user interfaces (GUIs). Prospective development of the geodatabase is suggested to include new collection records, as well as data sets related to West Virginia stream health. Efforts must be made to incorporate temporal (time/date of collection) and locational (map coordinate) characteristics of new records into the geodatabase so that full GIS analysis capability can be realized.

#### **Software Limitations**

New versions of ArcGIS (8.x) can be accessed using three software products, all sharing common architecture and each providing a higher level of functionality. So, while data manipulated with any of the three can be viewed with any of the three, certain editing procedures are only allowed in the higher-end products. ArcView provides comprehensive mapping and analysis tools along with simple editing and geoprocessing tools. It is the most basic (and least expensive) of the three products. This project was designed with simplicity and the widest availability to users in mind, and therefore employed ArcView for the process. The decision to do so left ample room to expand the project using one of the two remaining products.

ArcEditor includes the full functionality of ArcView, with the addition of advanced editing capabilities for coverages and geodatabases. It encompasses the potential for the greatest leap in the system's capability. ArcInfo extends the functionality of both to include advanced geoprocessing and applications from what was previously distributed as the ArcInfo Workstation.

An example of system limitation encountered by employing ArcView rather than ArcEditor or ArcInfo is the "relationship class". Relationship classes store associations between objects in a geodatabase that can include spatial objects (features in feature classes), nonspatial objects (rows in a table), or spatial and nonspatial objects. They can be viewed in ArcView, but created and edited only in either of the more expensive products. As a result, this system makes use of the join/relate structure.

## **References**

- Dunn, J. L., & Klausmeier, M. E. (2002). *An Integrated Data Management System for Environmental and Geospatial Data*. Proceedings of the 22<sup>nd</sup> Annual ESRI International User Conference. 8-12 July 2002. Retrieved from http://gis.esri.com/library/userconf/proc02/pap0970/p0970.htm
- Jenson, J. R. (1996). *Introductory Digital Image Processing: A Remote Sensing*  Perspective (2<sup>nd</sup> ed.). Upper Saddle River, NJ: Prentice Hall.
- Joselyn, Mark. (2002). *Building a Geodatabase for Watershed Management*. Proceedings of the 22<sup>nd</sup> Annual ESRI International User Conference. 8-12 July 2002. Retrieved from http://gis.esri.com/library/userconf/proc02/pap0929/p0929.htm
- Kennedy, H. (Ed.). (2001). *Dictionary of GIS Terminology*. Redlands, CA: ESRI Press.
- McCain, M. (2003). *Geodatabase Quality Control, Now More Important than Ever*. Proceedings of the 23<sup>rd</sup> Annual ESRI International User Conference. 7-11 July 2003. Retrieved from http://gis.esri.com/library/userconf/proc03/p0711.pdf
- McCoy, J., & Johnston, K. (2002). *Using ArcGIS Spatial Analyst*. Redlands, CA: ESRI Press.
- *Spatial Data Standards and GIS Interoperability*. (2003, January). ESRI White Paper. Retrieved from http://downloads.esri.com/support/whitepapers/ao\_/spatial-datastandards.pdf
- Stauffer, J. R. Jr., Boltz, J. M., & White, L. R. (1995). *The Fishes of West Virginia*. Philadelphia: Academy of Natural Sciences of Philadelphia.
- *Working with the Geodatabase: Powerful Multiuser Editing and Sophisticated Data Integrity*. (2002, February). ESRI White Paper. Retrieved from http://www.esri.com/library/whitepapers/pdfs/arcgis\_geodb\_multiuser.pdf

# **Software Inventory**



**Geodatabase:** WVFishGeo.mdb (3.81 MB)

# **Appendix A – Southern West Virginia Watersheds**

Southern West Virginia, for the purpose of this document, is defined as the portion of the state south of and including the Little Kanawha, Elk and Greenbrier River drainages:



Big Sandy River and Tug Fork – includes tributaries:

- ? Beech Creek
- $\mathscr{L}$  Big Creek
- $\mathscr{L}$  Dry Fork
- ? Elkhorn Creek
- $\mathscr{A}$  Jacobs Fork
- $\mathscr{A}$  Laurel Fork
- ? Mill Creek
- $\mathscr{A}$  Panther Creek
- $\mathscr{L}$  Pigeon Creek
- ? War Creek
- ? Whites Creek



Coal River – includes Big Coal and Little Coal Rivers and tributaries:

- $\mathscr{\mathscr{E}}$  Camp Creek
- $\ll$  Clear Fork
- ? Hewett Creek
- $\mathscr{L}$  Hopkins Fork
- ? Joes Creek
- ? Laurel Creek
- $\mathscr{L}$  Marsh Fork
- $\mathscr{D}$  Pond Fork
- $\approx$  Sandlick Creek
- $\mathscr{L}$  Spruce Fork
- ? Spruce Laurel Fork
- $\mathscr{L}$  Turtle Creek
- ? West Fork Pond Fork
- $\ll$  White Oak Creek

Elk River – includes Birch and Holly Rivers, Sutton Lake, and tributaries:

- $\mathscr{L}$  Back Fork Elk River
- $\mathscr{L}$  Big Otter Creek
- $\approx$  Big Sandy Creek
- $\mathscr{\mathscr{E}}$  Big Spring Fork
- $\mathscr{L}$  Blue Creek
- ? Buffalo Creek
- $\mathscr{\mathscr{E}}$  Falling Rock Creek
- ? Laurel Creek
- $\ll$  Left Hand Run
- ? Little Sandy Creek
- $\mathscr{L}$  Strange Creek

Gauley River – includes Cherry, Cranberry, Meadow and Williams Rivers, Summersville Lake, and tributaries:

- ? Bells Creek
- $\approx$  Big Beaver Creek
- $\approx$  Big Clear Creek
- $\mathscr{L}$  Big Creek
- $\mathscr{L}$  Big Ditch Run
- $\mathscr{L}$  Boggs Creek
- ? Brushy Meadow Creek
- ? Collison Creek
- ? Deer Creek
- ? Dogwood Creek





42

- ? Grassy Creek
- $\mathscr{L}$  Hominy Creek
- ? Little Clear Creek
- ? Little Sewell Creek
- ? Muddlety Creek
- ? Panther Creek
- $\mathscr{\mathscr{E}}$  Patterson Creek
- ? Peters Creek
- ? Sewell Creek
- $\mathscr{L}$  South Fork Big Clear Creek
- $\mathscr{\mathscr{E}}$  Twentymile Creek

Greenbrier River – including Little River and tributaries:

- $\mathscr{\mathscr{A}}$  Abes Run
- $\mathscr{L}$  Anthony Creek
- ? Cochran Creek
- ? Deer Creek
- ? Douthat Creek
- $\mathscr{L}$  Dry Creek
- ? Galford Run
- $\mathscr{L}$  Harts Run
- $\mathscr{L}$  Howard Creek
- ? Kitchen Creek
- $\mathscr{L}$  Knapp Creek
- ? Laurel Creek
- $\mathscr{L}$  Little Creek
- ? Mill Creek
- $\mathscr{L}$  Muddy Creek
- $\mathscr{L}$  North Fork
- ? North Fork Anthony Creek
- ? Robbins Run
- ? Second Creek
- $\mathscr{\mathscr{E}}$  Shellington Creek
- $\mathscr{L}$  Shock Run
- $\mathscr{L}$  Sinking Creek
- $\mathscr{L}$  Spring Creek
- $\mathscr{L}$  Stony Creek
- $\mathscr{\mathscr{E}}$  Sulphur Lick Run
- ? Thorny Creek
- ? Wolf Creek



Guyandotte River – including Mud River, R.D. Bailey Lake, and tributaries:

- $\mathscr{L}$  Barkers Creek
- $\mathscr{L}$  Big Creek
- $\mathscr{L}$  Big Harts Creek
- $\mathscr{L}$  Big Ugly Creek
- ? Buffalo Creek
- $\mathscr{\mathscr{E}}$  Clear Fork
- $\mathscr{L}$  Devils Fork
- $\mathscr{B}$  Elk Creek
- $\mathscr{A}$  Garrett Fork
- ? Gooney Otter Creek
- $\mathbb Z$  Huff Creek
- $\mathscr{L}$  Indian Creek
- $\mathscr{L}$  Island Creek
- $\mathscr{L}$  Laurel Fork
- $\mathscr{\mathscr{E}}$  Little Huff Creek
- $\mathscr{L}$  Pinnacle Creek
- $\mathscr{L}$  Slab Fork
- ? Tommy Creek
- $\mathscr{L}$  Trace Fork
- $\mathscr{\mathscr{E}}$  Winding Gulf

James River – does not flow within West Virginia. However, some tributaries of the James exist in Monroe County:

- ? Cove Creek
- $\mathscr{\mathscr{E}}$  Dunlap Creek
- ? Potts Creek

Kanawha River – includes Pocatalico River and tributaries:

- ? Cabin Creek
- $\mathscr{\mathscr{E}}$  Campbells Creek
- $\mathscr{A}$  Cherry Fork
- ? Davis Creek
- $\mathscr{L}$  Eightmile Creek
- $\mathscr{L}$  Five and Twenty Mile Creek
- $\mathscr{L}$  Hughes Creek
- $\mathscr{\mathscr{E}}$  Hurricane Creek
- ? Kellys Creek
- $\mathscr{L}$  Lens Creek
- ? Loop Creek
- $\mathscr{A}$  Mudlick Fork





- ? Paint Creek
- $\mathscr{L}$  Pond Branch
- $\mathscr{L}$  Poplar Fork
- $\mathscr{\mathscr{E}}$  Sixteenmile Creek
- $\mathscr{L}$  Tenmile Creek
- $\ll$  Thirteenmile Creek
- $\ll$  Threemile Creek
- $\ll$  Twomile Creek
- $\ll$  Witcher Creek

Little Kanawha River – including Hughes River, Burnsville Lake, and tributaries:

- $\mathscr{L}$  Beech Fork
- ? Cedar Creek
- $\mathscr{\mathscr{E}}$  Copen Run
- $\mathscr{L}$  Cove Creek
- $\ll$  Fink Creek
- $\mathscr{L}$  Henry Fork
- $\mathscr{A}$  Indian Fork
- $\mathscr{\mathscr{E}}$  Leading Creek
- ? Left Fork Reedy Creek
- $\mathscr{L}$  Left Fork Spring Creek
- $\mathscr{\mathscr{E}}$  Left Fork Steer Creek
- $\&$  Left Fork West Fork Little Kanawha River
- $\mathscr{\mathscr{E}}$  Little Bear Fork
- $\mathscr{L}$  Middle Fork Reedy Creek
- $\mathscr{L}$  Millstone Creek
- $\approx$  Oil Creek
- $\mathscr{L}$  Reedy Creek
- $\mathscr{L}$  Right Fork Little Kanawha River
- $\mathscr{L}$  Right Fork Steer Creek
- ? Saltlick Creek
- $\mathscr{L}$  Sand Fork
- $\mathscr{\mathscr{E}}$  Sinking Creek
- $\mathscr{\mathscr{E}}$  Spring Creek
- ? Steer Creek
- $\ll$  Stillwell Creek
- ? Tanner Creek
- $\mathscr{L}$  Tygart Creek
- ? Walker Creek
- $\mathscr{A}$  West Fork Little Kanawha River
- $\mathscr{L}$  Worthington Creek



New River – including East, Bluestone and Little Bluestone Rivers, Bluestone Lake, and tributaries:

- $\approx$  Adair Run
- $\approx$  Arbuckle Creek
- ? Beaver Creek
- $\mathscr{L}$  Brush Creek
- $\mathscr{\mathscr{E}}$  Camp Creek
- $\ll$  Dropping Lick Creek
- $\mathscr{L}$  Cry Creek
- $\mathscr{\mathscr{E}}$  Dunloup Creek
- ? Glade Creek
- $\mathscr{L}$  Hans Creek
- $\mathscr{A}$  Indian Creek
- $\ll$  Laurel Branch
- ? Laurel Creek
- $\mathscr{\mathscr{E}}$  Lick Creek
- ? Little Beaver Creek
- ? Manns Creek
- $\ll$  Meadow Creek
- $\ll$  Mill Creek
- $\mathscr{\mathscr{E}}$  Mountain Creek
- $\mathscr{L}$  Piney Creek
- $\mathscr{L}$  Pipestem Creek
- $\mathscr{L}$  Rich Creek
- $\approx$  Rock Camp Creek
- ? Rockbottom Creek
- $\ll$  Toms Run
- $\mathscr{L}$  Turkey Creek
- ? Wolf Creek

Ohio River South – including tributaries:

- $\mathscr{\mathscr{E}}$  Crab Creek
- $\mathscr{L}$  Eighteenmile Creek
- $\mathbb Z$  Elk Fork
- $\mathscr{\mathscr{E}}$  Guyan Creek
- ? Lee Creek
- $\mathscr{L}$  Left Fork Sandy Creek
- $\ll$  Little Mill Creek
- ? Little Sandy Creek
- ? Middle Fork Lee Creek
- $\mathscr{\mathscr{E}}$  Mill Creek
- $\mathscr{L}$  Mill Run
- $\ll$  Nesselroad Run





- ? North Fork Lee Creek
- $\mathscr{L}$  Parchment Creek
- ? Pond Creek
- $\mathcal{L}$  Right Fork Sandy Creek
- ? Sandy Creek
- $\mathscr{\mathscr{E}}$  Sixteenmile Creek
- ? South Fork Lee Creek
- $\mathscr{L}$  Tug Fork

Twelvepole Creek – including Beech Fork and East Lynn Lakes, and tributaries:

- $\mathscr{A}$  Beech Fork
- $\mathscr{L}$  Brush Creek
- $\mathscr{\mathscr{A}}$  Buffalo Creek
- ? Cove Creek
- ? Fisher Bowen Branch
- ? Garrett Creek
- $\mathscr Z$  Millers Fork



# **Appendix B – Glossary of Terms**

(Kennedy 2001)

alignment – For this project, alignment refers to the linear position of a river or stream channel.

ArcGIS® - The geographic information system (GIS) software package developed by Environmental Systems Research Institute (ESRI). Included in ArcGIS are applications ArcMap, ArcCatalog, ArcToolbox and others.

attribute  $- A$  piece of information describing a map feature. Attributes of a river might include its name, length, and average depth.

attribute table – A table containing descriptive attributes for a set of geographic features, usually arranged so that each row represents a feature and each column represents one attribute. Each cell in a column stores the value of that column's attribute for that row's feature.

background image – A satellite image or aerial photograph (such as a DOQQ) over which vector data is displayed. Although the image can be used to align coordinates, it is not linked to attribute information and is not part of the spatial analysis in a GIS.

basemap – A map depicting geographic features such as landforms, drainage, roads, landmarks, and political boundaries, used for locational reference.

cardinality– In a relationship between objects in a database, the number of objects of one type that are associated with objects of another type. A relationship can have a cardinality of one-to-one, one-to-many (many-to-one), or many-to-many.

control point – A point that represents the identical location on two different spatial layers; used in georeferencing one layer to another.

Digital Orthophoto Quarter Quadrangle (DOQQ) – An aerial photograph in one-quarter increments of the standard USGS quadrangle format. An orthophotograph has the same scale throughout and can be used as a map.

editor – In this project, refers to the Editor toolbar in ArcMap used for creating and manipulating vector shapefile features.

export – To move data from one computer system to another, and often, in the process, from one file format to another.

feature – A shape in a spatial data layer, such as a point, line, or polygon, that represents a geographic object.

feature class – In a shapefile, a collection of spatial data with the same shape type (e.g., point, line, or polygon).

geodatabase – An ArcInfo 8 data storage format. A geodatabase represents geographic features and attributes as objects and is hosted inside a relational database management system.

geographic coordinate system– A reference system using latitude and longitude to define the locations of points on the surface of a sphere or spheroid.

geographic information system (GIS) – A configuration of computer hardware and software that captures, stores, analyzes and displays geographic information

georeference – To assign coordinates from a known reference system, such as latitude/longitude, UTM, or State Plane to the page coordinates of an image or planar map.

import – To load data from one computer system or application into another, often involving some form of data conversion.

land use – The classification of land according to how it is used; for example, agricultural, industrial, residential, urban, rural or commercial. Natural features of the land such as forest, pastureland, brushland and bodies of water are also often classified in this manner.

layer – 1. A set of vector data organized by subject matter, such as roads, rivers, or political boundaries. Vector layers act as digital transparencies that can be laid atop one another for viewing or spatial analysis. 2. A set of raster data representing a particular geographic area, such as an aerial photograph or a remotely sensed image. In both (1) and (2), layers covering the same geographical space are registered to one another by means of a common coordinate system.

merge – In GIS, a method by which unnecessary boundaries between features are removed.

parameter query – A query that when run displays its own dialog box prompting you for information, such as criteria for retrieving records or a value you want to insert in a field; handy when used as the basis for forms, reports, and data access pages.

personal geodatabase – A geodatabase that stores data in a single-user relational database management system (RDBMS). A personal geodatabase can be read simultaneously by several users, but only one user at a time can write data into it.

primary key – The attribute column that uniquely identifies each row in a table, such as the unique number assigned to each parcel within a county.

projected coordinate system– Latitude and longitude coordinates projected to x,y coordinates in a planar coordinate system.

query (attribute query) – A statement or logical expression used to select features or records from a database. In ArcMap, queries are typically written in Structured Query Language (SQL).

raster image – A matrix of pixels whose values represent the level of energy reflected or emitted by the surface being photographed, scanned, or otherwise sensed.

raster-to-vector conversion (vectorization) – The conversion of cell data into points, lines, and polygons.

relational database – Data stored in table that are associated by shared attributes. Any data element can be found in the database through the name of the table, the attribute (column)

name, and the value of the primary key. In contrast to hierarchical and network database structures, the data can be arranged in different combinations.

segment – A line that connects vertices.

select query– The most common type of query. It retrieves data from one or more tables and displays the results in a datasheet where you can update the records (with some restrictions).

shapefile – A vector file format for storing the location, shape, and attributes of geographic features. It is stored in a set of related files and contains one feature class.

snapping – Moving a feature, or a portion of it, to coincide with the coordinates of another feature.

spatial analysis – Studying the locations and shapes of geographic features and the relationships between them.

spatial data – Information about the locations and shapes of geographic features, and the relationships between them; usually stored as coordinates and topology.

Structured Query Language  $(SOL) - A$  syntax for defining and manipulating data in a relational database. Developed by IBM in the 1970s, it has become an industry standard for query languages in most relational database management systems.

topographic map – A map showing relief, often as contour lines, along with other natural and human-made features; map sheets published by the U.S. Geological Survey in the 7.5 minute or 15-minute quadrangle series.

topological overlay – Superimposing two or more geographic data sets in order to produce a new geographic layer with a new set of attributes; method of spatial analysis.

transformation (rectification) – Converting the coordinates of a map or an image from one system to another , typically by shifting, rotating, scaling, skewing or projecting them.

union - Combines features from different layers into one feature while maintaining the original features and attributes.

universal transverse Mercator (UTM) – A commonly used projected coordinate system that divides the globe into sixty zones, starting at -180 degrees longitude. Each zone extends north-south from 84 degrees north to 80 degrees south, spans 6 degrees of longitude, and has its own central meridian. Most of the state of West Virginia, including the entire portion of the state represented in this project, is in UTM Zone 17 North.

vector – A data structure used to represent linear geographic features. Features are made of ordered lists of x,y coordinates and represented by points, lines, or polygons; points connect to become lines, and lines connect to become polygons. Attributes are associated with each feature (as opposed to a raster data structure, which associates attributes with grid cells).



# **Appendix C – Glossary of Spatial Analysis**



**Image 7** – Huff Creek with "barren" polygons, mine permit boundaries, valley fill permit boundaries



**Image 9** – Distribution of *L. sicculus* (brook silverside)





**Image 8** – Pinnacle Creek with "barren" polygons, mine permit boundaries, valley fill permits



**Image 10** – Distribution of *L. aepyptera* (least brook lamprey)





**Image 13** – Distribution of *P. omiscomaycus* (trout-perch)



**Image 15** – Distribution of *E. variatum* (variegate darter)



# **Appendix D – Metadata**

# **Data Sources Made Available by the West Virginia GIS Data Center: (http://wvgis.wvu.edu)**

#### **Watersheds**

## Description

"Watershed" hydrologic units, a subdivision within a sub-basin, represent the 5th level (10-digits) in the hydrologic unit hierarchy. Watersheds range in size from 40,000 to 250,000 acres.

Scale

1:100,000

Location

**Statewide** 

Data Source Lineage

During the late 1970's the Natural Resources Conservation Service (NRCS), formerly the Soil Conservation Service, initiated a national program to further subdivide HUC's into smaller watersheds for water resources planning, resulting in the 5th level, or 10 digit HUCs. In the mid-1990's the NRCS delineated and digitized watersheds (Level 5 hydrologic units) on 1:100,000-scale USGS topographic maps. In 2001 new national standards for delineating watershed hydrologic units adopted a 10-digit format instead of 11 digits. This coverage will be updated when more detailed watershed/subwatershed mapping is completed as part of the Watershed Boundary Dataset, a national geospatial database containing the hydrologic unit boundaries for the 1st through 6th level units. This dataset was downloaded from the WV Department of Environmental Protection and converted from NAD 27 to NAD 83. Published January 2002.

Coordinate System UTM Zone 17 (NAD 83) in meters

File Format

Compressed ESRI shapefiles

# **Counties**

Description West Virginia County Boundaries Scale 1:24,000 Location Statewide Data Source Lineage Digitized from USGS 1:24,000-scale digital raster graphics (scanned topographic maps) by the West Virginia Department of Environmental Protection. Published January 2002. Coordinate System

UTM Zone 17 (NAD 83) File Format Compressed ESRI shapefile

#### **Cities**

**Description** Populated Places recorded in the Geographic Names Information System Scale 1:24,000 Location Statewide Data Source Lineage The Geographic Names Information System (GNIS), developed by the U.S. Geological Survey (USGS) in cooperation with the U.S. Board on Geographic Names (BGN), contains information about physical and cultural geographic features. Publication Date: 05-01-1981. Coordinate System UTM Zone 17 (NAD 83) File Format Compressed ESRI point shapefiles (double precision)

## **Major Water**

Description Major water bodies intended for use in making statewide cartographic maps. Scale 1:100,000 Location Statewide Data Source Lineage Line and polygon coverages derived from the USGS National Hydrological Dataset (NHD) and Digital Line Graphs (DLG). Compiled and edited by the WV GIS Technical Center. Published 2002. Coordinate System UTM Zone 17 (NAD 83) File Format Compressed ESRI line and polygon shapefiles

# **Mining Permit Boundaries**

Description Permit oundaries digitized from West Virginia Department of Environmental Protection's (WVDEP) Division of Mining and Reclamation (DMR) permit maps. Scale 1:24,000 Location Statewide Data Source Lineage Permit Boundaries from 1:24,000-scale or larger DMR permit maps were digitized by the WVU Natural Resource Analysis Center (NRAC) and then validated by DMR personnel for completeness. Coordinate System UTM Zone 17 (NAD 27) File Format Compressed ESRI shapefiles

#### **Underground Mining Limits**

Description

Underground Mining Limits digitized from West Virginia Department of Environmental Protection's Division of Mining and Reclamation (DMR) permit maps.

Scale 1:24,000 Location Statewide Data Source Lineage Underground Mining Limits from 1:24,000-scale or larger DMR permit maps were digitized by the WVU Natural Resource Analysis Center (NRAC) and then validated by DMR personnel for completeness. Coordinate System UTM Zone 17 (NAD 27) File Format Compressed ESRI shapefiles

#### **Valley Fills**

Description Valley fills digitized from West Virginia Department of Environmental Protection's Division of Mining and Reclamation (DMR) permit maps. Scale 1:24,000 Location Statewide Data Source Lineage Valley fills from 1:24,000-scale or larger DMR permit maps were digitized by the WVU Natural Resource Analysis Center (NRAC) and then validated by DMR personnel for

completeness.

Coordinate System UTM Zone 17 (NAD 27) File Format Compressed ESRI shapefiles

# **National Land Cover Dataset (NLCD) 1992**

Description

The National Land Cover Dataset (NLCD), produced as a cooperative effort between the U.S. Geological Survey (USGS) and the U.S. Environmental Protection Agency (US EPA), provides a consistent, land cover data layer for the conterminous U.S. using early 1990s Landsat thematic mapper (TM) data purchased by the Multi-resolution Land Characterization (MRLC) Consortium. This data can be used for landscape scale analysis in various disciplines such as wildlife ecology, forestry, or land use planning. Scale

1:50,000 Location Statewide Data Source Lineage The National Land Cover Dataset was compiled from Landsat satellite TM imagery

(circa 1992) with a spatial resolution of 30 meters and supplemented by various ancillary data. Map projection of original NLCD data set converted from Albers Conical Equal Area to a UTM (Zone 17) and Geographic coordinate system by WVGISTC. Information: WV NLCD Version 05-27-99 (Updated: January, 1996; February, 1997; July, 1998, May, 1999; March, 2000). Coordinate System UTM Zone 17 (NAD 83) File Format

Compressed ESRI ARC/INFO grid

# **National Land Cover Data Key:**

NOTE - All Classes May NOT Be Represented in a specific state data set.

The class number represents the digital value of the class in the data set. NLCD Land Cover Classification System Key - Rev. July 20, 1999

# **Water**

 11 Open Water 12 Perennial Ice/Snow

# **Developed**

 21 Low Intensity Residential 22 High Intensity Residential 23 Commercial/Industrial/Transportation

#### **Barren**

 31 Bare Rock/Sand/Clay 32 Quarries/Strip Mines/Gravel Pits 33 Transitional

#### **Forested Upland**

41 Deciduous Forest

42 Evergreen Forest

43 Mixed Forest

#### **Shrubland**

51 Shrubland

#### **Non-natural Woody**

61 Orchards/Vineyards/Other

#### **Herbaceous Upland**

71 Grasslands/Herbaceous

## **Herbaceous Planted/Cultivated**

 81 Pasture/Hay 82 Row Crops 83 Small Grains 84 Fallow 85 Urban/Recreational Grasses

#### **Wetlands**

 91 Woody Wetlands 92 Emergent Herbaceous Wetlands

#### **NLCD Land Cover Classification System Land Cover Class Definitions:**

**Water**- All areas of open water or permanent ice/snow cover.

11. Open Water - All areas of open water; typically 25 percent or greater cover of water (per pixel).

12. Perennial Ice/Snow - All areas characterized by year-long cover of ice and/or snow.

**Developed** - Areas characterized by a high percentage (30 percent or greater) of constructed materials (e.g. aspha lt, concrete, buildings, etc).

21. Low Intensity Residential - Includes areas with a mixture of constructed materials and vegetation. Constructed materials account for 30-80 percent of the cover. Vegetation may account for 20 to 70 percent of the cover. These areas most commonly include single-family housing units. Populationdensities will be lower than in high intensity residential areas.

22. High Intensity Residential - Includes highly developed areas where people reside in high numbers. Examples include apartment complexes and row houses. Vegetation accounts for less than 20 percent of the cover. Constructed materials account for 80 to 100 percent of the cover.

23. Commercial/Industrial/Transportation - Includes infrastructure (e.g. roads, railroads, etc.) and all highly developed areas not classified as High Intensity Residential.

**Barren**- Areas characterized by bare rock, gravel, sand, silt, clay, or other earthen material, with little or no "green" vegetation present regardless of its inherent ability to support life. Vegetation, if present, is more widely spaced and scrubby than that in the "green" vegetated categories; lichen cover may be extensive.

31. Bare Rock/Sand/Clay - Perennially barren areas of bedrock, desert pavement, scarps, talus, slides, volcanic material, glacial debris, beaches, and other accumulations of earthen material.

32. Quarries/Strip Mines/Gravel Pits - Areas of extractive mining activities with significant surface expression.

33. Transitional - Areas of sparse vegetative cover (less than 25 percent of cover) that are dynamically changing from one land cover to another, oftenbecause of land use activities. Examples include forest clear-cuts, a transition phase between forest and agricultural land, the temporary clearing of vegetation, and changes due to natural causes (e.g. fire, flood, etc.).

**Forested Upland** - Areas characterized by tree cover (natural or semi-natural woody vegetation, generally greater than 6 meters tall); tree canopy accounts for 25-100 percent of the cover.

41. Deciduous Forest - Areas dominated by trees where 75 percent or more of the tree species shed foliage simultaneously in response to seasonal change.

42. Evergreen Forest - Areas dominated by trees where 75 percent or more of the tree species maintain their leaves all year. Canopy is never without green foliage.

43. Mixed Forest - Areas dominated by trees where neither deciduous nor evergreen species represent more than 75 percent of the cover present.

**Shrubland** - Areas characterized by natural or semi-natural woody vegetation with aerial stems, generally less than 6 meters tall, with individuals or clumps not touching to interlocking. Both evergreen and deciduous species of true shrubs, young trees, and

trees or shrubs that are small or stunted because of environmental conditions are included.

51. Shrubland - Areas dominated by shrubs; shrub canopy accounts for 25-100 percent of the cover. Shrub cover is generally greater than 25 percent when tree cover is less than 25 percent. Shrub cover may be less than 25 percent in cases when the cover of other life forms (e.g. herbaceous or tree) is less than 25 percent and shrubs cover exceeds the cover of the other life forms.

**Non-natural Woody** - Areas dominated by non-natural woody vegetation; non-natural woody vegetative canopy accounts for 25-100 percent of the cover. The non-natural woody classification is subject to the availability of sufficient ancillary data to differentiate non-natural woody vegetation from natural woody vegetation.

61. Orchards/Vineyards/Other - Orchards, vineyards, and other areas planted or maintained for the production of fruits, nuts, berries, or ornamentals.

**Herbaceous Upland** - Upland areas characterized by natural or semi-natural herbaceous vegetation; herbaceous vegetation accounts for 75-100 percent of the cover.

71. Grasslands/Herbaceous - Areas dominated by upland grasses and forbs. In rare cases, herbaceous cover is less than 25 percent, but exceeds the combined cover of the woody species present. These areas are not subject to intensive management, but they are often utilized for grazing.

**Planted/Cultivated**- Areas characterized by herbaceous vegetation that has been planted or is intensively managed for the production of food, feed, or fiber; or is maintained in developed settings for specific purposes. Herbaceous vegetation accounts for 75-100 percent of the cover.

81. Pasture/Hay - Areas of grasses, legumes, or grass-legume mixtures planted for livestock grazing or the production of seed or hay crops.

82. Row Crops - Areas used for the production of crops, such as corn, soybeans, vegetables, tobacco, and cotton.

83. Small Grains - Areas used for the production of graminoid crops such as wheat, barley, oats, and rice.

84. Fallow - Areas used for the production of crops that are temporarily barren or with sparse vegetative cover as a result of being tilled in a management practice that incorporates prescribed alternation between cropping and tillage.

85. Urban/Recreational Grasses - Vegetation (primarily grasses) planted in developed settings for recreation, erosion control, or aesthetic purposes. Examples include parks, lawns, golf courses, airport grasses, and industrial site grasses.

**Wetlands** - Areas where the soil or substrate is periodically saturated with or covered with water.

91. Woody Wetlands - Areas where forest or shrubland vegetation accounts for 25-100 percent of the cover and the soil or substrate is periodically saturated with or covered with water.

92. Emergent Herbaceous Wetlands - Areas where perennial herbaceous vegetation accounts for 75-100 percent of the cover and the soil or substrate is periodically saturated with or covered with water.

Data Sources Obtained from Environmental Systems Research Institute, Inc. Data & Maps Publications:

ESRI Data & Maps 2002

# **U.S. Rivers**

Description

U.S. Rivers provides a data set of the U.S. rivers for each state of the United States and the District of Columbia.

Scale

1:100,000

Location United States

Data Source Lineage

This data set was extracted from the standard GDT line water layer and includes only FCC codes H10 (this code no longer exists, but at the time was described as stream) and H11 that represent perennial rivers, streams, and streams of unknown minor category (perennial, intermittent, or braided). These line segments were chained and thinned to serve as a reference cartographic layer. Many of the cartographic problems inherent in the original 1:100,000-scale source data remain; however, for small areas these rivers make a nice cartographic addition to many basemaps.

Coordinate System GCS North American 1983 File Format

ESRI shapefile

#### **West Virginia Rivers**

#### Description

West Virginia Rivers represents streams and rivers within West Virginia. It provides a reference cartographic layer that is chained and thinned to reduce size and improve draw performance.

Scale

Largest scale when displaying the data: 1:100,000.

Location

**Statewide** 

Data Source Lineage

Geographic Data Technology, Inc. (GDT)

The data included in the boundary files was extracted from the 1990 Census Version TIGER/Line files (CD–ROM edition). The following steps were performed by ESRI: Extracted data sets from GDT Dynamap/2000 v7.3. Ran ArcView Avenue request ReturnConnected, and generalized. Converted all states to North American Datum of 1983 (NAD83). Recalculated any attributes based on length. Created ArcGIS™ layer file (.lyr). Created ArcView legend file (.avl). Created projection file (.prj). Ran ArcView Avenue request ExportClean, and created spatial indices. The geospatial part of this data set was extracted from the 1990 Census Version TIGER/Line™ files (CD–ROM edition). The positional accuracy of the TIGER/Line coordinates at best meets the established National Map Accuracy standards (+/- 167 feet approximately) for the 1:100,000-scale source maps from the USGS. The level of positional accuracy in the TIGER/Line files is not suitable for high-precision measurement applications such as engineering problems, property transfers, or other uses that might require highly accurate measurements of the earth's surface. No duplicate features are present. The shapefile is exported using Avenue request ExportClean. This request verifies and enforces the correctness of shapes. After processing, the data set is checked for drawing display and number of records and file sizes compared with source materials. Some of the cartographic problems with this data inherent in the original 1:100,000 USGS Topographic Quadrangle source data remain in that the density of features can change abruptly.

Coordinate System GCS North American 1983 File Format

ESRI shapefile

ESRI Data & Maps 2003

#### **U.S. Rivers and Streams**

#### **Description**

U.S. Rivers and Streams represent detailed rivers and streams in the United States. The data set provides a database of linear water features that interconnects and identifies the stream segments or reaches that comprise the surface water drainage system of the

United States. The detailed and comprehensive rivers and streams are from the National Hydrography Dataset by the U.S. Geological Survey in cooperation with the U.S. Environmental Protection Agency.

Scale

1:24,000

Location

United States

Data Source Lineage

The following steps were performed by ESRI: Extracted NHD Route DRAIN from National Hydrography Dataset (NHD). Added STRM\_LEVEL and NAME attributes from NHD Route RCH. Removed unneeded attributes. Split the data set by U.S. states (for easier processing). Unsplit features based on combining NAME, FTYPE, FCODE, and STRM\_LEVEL attributes. Merged the data sets into one. Recalculated the lengths for the METERS attribute. Formatted the attributes.

Coordinate System GCS North American 1983

File Format ESRI shapefile

# **Appendix E – Data Tables**






















































