The Use of GIS for Integrated Watershed Analysis:
Integration of Environmental Models with GIS in the
Upper Roanoke River Watershed

Robert W. Dietz

Thesis submitted to the faculty of
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Master of Science

in

Environmental Science and Engineering

Dr. Randel L. Dymond
Dr. Darrell J. Bosch
Dr. Vinod K. Lohani

April 26, 2000
Blacksburg, VA, USA

Keywords: GIS, environmental modeling, watershed analysis, model integration
The Use of GIS for Integrated Watershed Analysis:  
Integration of Environmental Models with GIS in the  
Upper Roanoke River Watershed  

by  
Robert W. Dietz  

(ABSTRACT)  

Practitioners of watershed management are increasingly turning to computer models to understand and make decisions about the diverse problems that occur in the watershed. Such models can provide insight into how human interactions with the landscape affect water quality and quantity. Additional modeling tools trace how those effects ripple through ecosystems, economies, and other systems. In the past, models were stand-alone and process-specific, aimed at solving problems related to a narrow discipline. For example, hydrologic models analyzed the quantity of flow through waterways. Separate ecological models probed the cycling of nutrients in those waterways.

An emerging trend for watershed-based models is to link them to a geographic information system (GIS), which provides the basis for integrating data, algorithms, and methods from each discipline of interest. This integration capability makes GIS a very powerful tool for the watershed manager. The GIS in this study within the Upper Roanoke River Watershed integrates modeling efforts from the fields of hydrology, economics, and ecology.

The main goal of this study is to demonstrate the effectiveness of GIS as an integrating and computational aid for making sound decisions about a watershed. A secondary goal is to include GIS functionality in a prototype software application for evaluating the effects of land management decisions. The application, named DesktopL2W, can be a significant tool for choosing how and where development should occur within the boundaries of a watershed.

The three major results of the study are: (1) a library of spatial data that is valuable for watershed analysis; (2) a set of procedures for undertaking a GIS integration project; and (3) the DesktopL2W software product with its usefulness to planners and others who are interested in how development affects the watershed. In addition, discussion of technical issues, such as selection of data formats and spatial and temporal resolution, provides insight into the complexities associated with a GIS integration effort.
ACKNOWLEDGEMENTS

Within the project: I am thankful to be a part of this EPA/NSF-sponsored interdisciplinary project. It has given me the opportunity to work with many fine researchers and explore a topic of great interest to me. Special thanks go to my committee members, Randy Dymond, Darrell Bosch, and Vinod Lohani, for their outstanding guidance and suggestions. Fellow graduate students on the project, Jeff Chanat, Ioannis Kaltsas, Cameron Speir, Vann Stancil, and Newland Agbenowosi, also deserve my gratitude for helping me develop portions of the data and software.

Within the university: I am thankful to the faculty, students, and staff of Virginia Tech’s Environmental Engineering, Hydrosystems, and Geography programs. They provided all kinds of day-to-day help and inspiration for completing this project.

Family: Most of all, I want to acknowledge my family for their role. Thanks to Mom, Dad, Judith, Alice, Wen, and Chris for continually emphasizing the importance of education, believing in me, and reminding me about what is important. Most notably, thanks to Jen, who makes every day (even the ones spent writing a thesis into the wee hours of the night) a joy. My graduate degree and this thesis project are a direct result of her support and love. Thank you.
# TABLE OF CONTENTS

1.0 INTRODUCTION  
1.1 STATEMENT OF PROBLEM  
1.2 OBJECTIVES  

2.0 LITERATURE REVIEW  
2.1 SPREAD OF GIS TO ENVIRONMENTAL AND WATERSHED MODELING SYSTEMS  
2.1.1 URGENCY OF ENVIRONMENTAL DECISIONS  
2.1.2 SPATIAL NATURE OF WATERSHED ANALYSIS  
2.2 GENERAL LIMITATIONS OF GIS AS A TOOL FOR WATERSHED ANALYSIS  
2.3 AVAILABLE SPATIAL DATA FOR WATERSHED ANALYSIS  
2.3.1 HYDROGRAPHY  
2.3.2 ELEVATION  
2.3.3 LAND USE  
2.3.4 SOILS  
2.3.5 TRANSPORTATION  
2.4 INTEGRATING FUNCTIONS OF GIS IN ENVIRONMENTAL MODELING AND WATERSHED ANALYSIS  
2.4.1 SPATIAL INVENTORY/FEATURE CHARACTERIZATION  
2.4.2 PREPROCESSING AND DEVELOPMENT OF INPUT DATA  
2.4.3 SPATIAL ANALYSIS AND DATA MANIPULATIONS  
2.4.4 COMMUNICATION OF RESULTS AMONG MODELS  
2.4.5 PROVISION OF A COMMON USER INTERFACE  
2.5 TYPES OF MODELS – MODEL ARCHITECTURE  
2.5.1 COUPLED OR LINKED MODELS  
2.5.2 HOLISTIC GIS MODELS  
2.6 TECHNICAL CONSIDERATIONS IN BUILDING GIS-BASED ENVIRONMENTAL MODELS  
2.6.1 SELECTION OF SPATIAL DATA STRUCTURES  
2.6.2 SPATIAL AND TEMPORAL SCALE/RESOLUTION  
2.6.3 PROJECTION  
2.6.4 INTERDISCIPLINARY COLLABORATION  
2.6.5 OBJECT ORIENTED PROGRAMMING AND GIS  
2.7 MODELING EFFORTS/CASE STUDIES  

3.0 STUDY AREA  
3.1 MAIN STUDY AREA  
3.2 STUDY AREA FOR THE PROTOTYPE SOFTWARE  

4.0 METHODS  
4.1 DATA SELECTION AND DEVELOPMENT  
4.1.1 DEVELOPMENT OF KEY BASE DATA LAYERS  

1.0 INTRODUCTION

As environmental problems associated with watershed development continue to be a cause of concern in the United States, researchers and policy makers rely more and more on the use of computer models to understand and cope with such problems. These models play a large role in managing and making decisions about water resources. They provide insight into how human and natural actions affect water quality and quantity in a watershed. Additional modeling tools and efforts trace how those effects ripple through ecosystems, economies, and other systems.

Concurrent to the increased use of models, a paradigm shift has transpired in how practitioners approach solutions to water resources problems in the U.S. In the past, the Environmental Protection Agency (EPA) attempted to manage water pollution and hydrologic phenomena by focusing on individual water bodies or dischargers. The agency now applies a Watershed Approach, which considers water quality and ecosystem problems at the watershed level. More specifically, EPA’s Watershed Approach consists of four main components (Davis, 1998):

1. Adoption of the watershed or stream basin as a base hydrologic unit;
2. Use of a comprehensive, integrated approach to address water quality and related natural resource management;
3. Involvement of stakeholders in making decisions along with a solid foundation of standards, regulation, and enforcement; and
4. Development of innovative solutions tailored to local needs, resources, and abilities.

These four components clearly show EPA’s commitment to managing water resources in a holistic way from the perspective of the watershed and the community.

The construction of models and their integration presents a challenge to modelers in following the Watershed Approach. Much like the water quality efforts that existed before the Watershed Approach, many computer models developed in the past to address environmental issues were narrow in scope and specific to a certain type of problem. The crux of the challenge is designing a model that can handle all the interdisciplinary tasks required by the Watershed Approach. Such a model must be able to integrate data, methods, and subroutines from a number of disciplines. This level of integration allows the model to function as a decision support system for planners, regulatory personnel, and policy makers involved in watershed development activities.

From the viewpoint of integration, the emerging trend for watershed-based models is to link them with or incorporate them into a geographic information system (GIS). The functionality of GIS enables the integrated model to handle the data management, computational aspects, and integration needs as emphasized in the Watershed Approach.

1.1 Statement of Problem

Building a GIS that is useful for conducting a watershed analysis is a complex task. Thorough analysis that accounts for the state of human modifications to the watershed minimally requires models from the fields of hydrology, ecology, and economics. The main purpose of this research is to develop a GIS that is capable of integrating models from these diverse fields of study. The
models focus explicitly on potential changes in land use, in the form of placement of new neighborhoods and developments, as perturbations on the landscape. The goal of the GIS, therefore, is to allow a modeler to visualize development changes to the landscape and produce resultant input values for the individual models. This research stems from a larger project at Virginia Tech sponsored by EPA and the National Science Foundation (NSF) to devise procedures for conducting multidisciplinary watershed analyses.

Efforts have been made in the past to link GIS to environmental models, but several gaps in the process remain. The present work seeks to address these gaps, which are:

- Lack of methods for selecting and obtaining appropriate spatial data;
- Absence of procedures for enhancing and customizing spatial data sets for use in models;
- Shortage of interface designs for GIS-based modeling, particularly with respect to modeling new development on the landscape; and
- Lack of information on how to handle differences in spatial and temporal resolution among models and the GIS.

1.2 Objectives

The objectives of this research are to:

1. Collect and enhance spatial data for use in a watershed modeling environment, with clear rationales for selection of specific data sets;
2. Collaborate with modelers in the fields of hydrology, economics, and ecology to develop custom spatial data sets for use in individual models;
3. Construct a working GIS software environment that provides an integrating interface for selected models/procedures used for evaluating hydrologic, economic, and ecological effects of land use changes;
4. Assemble disparate modeling approaches in a user interface to develop procedures for addressing watershed-scale issues; and
5. Identify the limitations of the GIS environment, especially with respect to the issue of spatial/temporal resolution in integration of multidisciplinary models.

The remainder of this document contains chapters that describe efforts to meet the objectives outlined above. Chapter 2 gives a review of the literature on GIS and its use in integrating environmental models. Chapter 3 describes the study area selected to demonstrate the techniques developed. Chapter 4 provides the methods used in the GIS integration process. Chapter 5 presents results, and Chapter 6 is a final discussion and scope for future work.
2.0 LITERATURE REVIEW

This literature review assembles information about how GIS has been used to help model and analyze watersheds. It focuses on the technical aspects of preparing a GIS that can integrate modeling efforts from assorted fields of science.

2.1 Spread of GIS to Environmental and Watershed Modeling Systems

A geographic information system, commonly called a GIS, is a general purpose computer technology for manipulating digital geographic data. A useful GIS has the ability to manage specific tasks such as (Goodchild, 1993):

- Preprocessing of data into a form that can be analyzed (e.g., projection or reprojection of coordinates and reclassification of data);
- Analysis of data and modeling of physical and social processes; and
- Postprocessing of results, including production of map displays and data reports.

These capabilities of GIS reveal its potential as a tool for modeling processes that occur within the watershed. The urgency that accompanies environmental problems and the spatial nature of watershed analysis also promote the use of GIS.

2.1.1 Urgency of Environmental Decisions

Managers who are responsible for tackling environmental problems often face a high degree of urgency in understanding and solving the problems. The problems themselves are usually complex, not only fraught with fragmentary and hotly contested data, but also filtered by various institutions. On top of these convoluting characteristics, managers are often pressed to make quick decisions about these problems for an impatient public. This desire to arrive at decisions that are both timely and beneficial drives the need to co-adapt GIS and environmental models (Parks, 1993).

2.1.2 Spatial Nature of Watershed Analysis

Most environmental/natural resource problems are spatial phenomena. Such problems are inextricably linked to the landscapes, watersheds, and other geographic units in which they occur. Spatial distributions may be even more important in watershed-related issues, such as nonpoint source water pollution. By definition, nonpoint source pollution originates from diffuse sources within a geographic area. Rainfall or snowmelt that moves over or through the ground in the area carries pollutants into receiving waters. The problem of nonpoint source pollution, therefore, is closely tied to the characteristics of the landscape in which it transpires. Examples of such characteristics include slopes, soil types, amount of precipitation, land use, and land cover.

Similar to nonpoint source modeling, general hydrologic modeling is spatial in nature. Hydrologic modeling seeks to understand three main issues, pollution control, water utilization, and flood control. In addressing these issues, hydrologic models use four spatial components: surface watersheds, pipes or stream channels, subsurface aquifers, and lakes or estuaries. Each
component is a three-dimensional landscape characteristic which can generally be represented satisfactorily in a two-dimensional GIS (Maidment, 1993).

GIS excels at explicit examination of spatial characteristics (Parks, 1993). Landscape characteristics can be generalized and analyzed by a GIS. When linked with environmental models, GIS provides an efficient way to handle the complex spatial and temporal heterogeneities of the earth’s surface and subsurface (Loague, Corwin, and Ellsworth, 1998).

2.2 General Limitations of GIS as a tool for Watershed Analysis

Judging by the growing acceptance of GIS in environmental modeling (Loague, Corwin, and Ellsworth, 1998), its advantages outweigh the disadvantages, but it is still important to understand the characteristics of GIS that limit its value. Some of the major limitations are highlighted below:

- Acquisition of the data and information necessary to use GIS and simulation models can be costly in both a monetary and time sense (Lovejoy, 1997). Some of the reasons for this potentially high cost are: (1) the data needed for simulations are site specific, and data may not be readily available for a given site; (2) spatial data come from many sources and are developed at different qualities, scales, and projections, so compiling a complete and compatible set of data is difficult; and (3) development of useable spatial data often requires substantial processing such as tiling/mosaic building, reprojection, and error removal.

- Soil, climate, and chemical data that are sparsely collected often have to be estimated for use in a GIS-based model (Loague, Corwin, and Ellsworth, 1998).

- GIS can be an implement of simplistic thinking if an analyst is unable to include numerical models, time-variant relationships, three-dimensional objects, and other advanced properties (Parks, 1993). Modeling approaches that are built into GIS often appear to be overly simple and restricted (Fedra, 1993).

- Output maps of GIS-based models have a given level of uncertainty in them. This uncertainty stems from two types of errors – model errors and data errors. Model errors can arise from oversimplification of processes that are represented in the model, while data errors can arise from faulty input or lack of information. Both types of errors weaken GIS-based models (Loague, Corwin, and Ellsworth, 1998). Indeed, the greatest challenge to modeling watershed processes cost-effectively is to obtain sufficient data to characterize spatial distributions with a knowledge of their uncertainty (Corwin and Wagenet, 1996).

- There are few opportunities to determine the reliability of results in a GIS-based model. Environmental models developed from simple empirical relationships are frequently more successful at making predictions than complex, physically based models. These physically based models have rarely been fully evaluated for predictive ability. In linking such models to GIS, therefore, the sole criterion of quality is often the cartographic display of results (Burrough, van Rijn, and Rikken, 1996).
2.3 Available Spatial Data for Watershed Analysis

As noted above, the costs of acquiring spatial data can be a major drawback to using GIS for environmental modeling. As the use of GIS has become more widespread for a number of enterprises, however, access to free and inexpensive spatial data sources has increased tremendously. Examples of useful types of spatial data that are readily available for the United States are hydrography, elevation, land use, soils, and transportation. Many other types and sources of data are also available from various public and private sources. Where possible, URLs for information or direct download of data sources are provided.

2.3.1 Hydrography

Hydrography databases provide locations and attributes of surface water, such as rivers, streams, and lakes. The U.S. Geological Survey (USGS) produces digital hydrography data as part of its series of digital line graphs (DLGs). DLGs contain points, lines, and polygons that were originally digitized from USGS topographic maps (Kemp, 1993), and are available in several different scales. The public can download files from the U.S. GeoData web site (http://edc.usgs.gov/doc/edchome/ndcdb/ndcdb.html). Additional hydrologic data are contained in the National Hydrography Dataset (http://nhd.usgs.gov/).

The U.S. Bureau of the Census produces a series of spatial data called TIGER/Line files (Topologically Integrated Geographic Encoding and Referencing System) (Robinson and Ragan, 1993). These files include a hydrography component at 1:100,000 scale that is based on the USGS DLGs. Technical and practical information about the TIGER/Line files, including how to obtain files, is available from the U.S. Census web site (http://www.census.gov/geo/www/tiger/).

2.3.2 Elevation

The USGS is the primary provider of elevation data for the United States. The agency produces two distinct sets of digital elevation data – digital elevation models (DEMs) and contour lines in the form of DLGs. DEMs supply data as a grid of point elevations (Kemp, 1993), while DLGs represent terrain topography by displaying lines of equal elevation. DEMs are available in two scales, 1:24,000 and 1:250,000, to coincide with USGS topographic quad sheets. DEMs may be directly downloaded from the U.S. GeoData web site: (http://edc.usgs.gov/doc/edchome/ndcdb/ndcdb.html).

2.3.3 Land Use

In its role as the central agency in charge of digital spatial data, the USGS provides Land Use and Land Cover (LULC) data files. These files classify land surfaces by vegetation, water, natural surface, and cultural features. LULC files are available at scales of 1:250,000 and 1:100,000, although not all quadrangles have been digitally mapped. The U.S. GeoData web site (http://edc.usgs.gov/doc/edchome/ndcdb/ndcdb.html) provides downloads.

For many watershed modeling purposes, land use is extremely important, and often requires collection of data beyond what is available. Such data collection efforts may include acquisition and classification of satellite imagery, such as LANDSAT Thematic Mapper and SPOT HRV. Modelers may also use digital orthophotos to develop land use data. Digital orthophotos are
scanned aerial photographs that have been corrected to remove visual errors caused by elevation changes and camera tilt (Kemp, 1993).

2.3.4 Soils

Databases developed by the U.S. Department of Agriculture’s Natural Resource Conservation Service (NRCS) are the cheapest source of soils data. The databases vary in scale. National level, state and regional level, and county level databases exist, but digital data are not available for many counties. The three sets of NRCS databases are the National Soil Geographic Database (NATSGO), the State Soil Geographic Database (STATSGO), and the Soil Survey Geographic Database (SSURGO) (Lytle, 1993). STATSGO data were intended to replace existing, inconsistent hard-copy soil maps with consistent digital information (Lytle, Bliss, and Waltman, 1996). These databases provide a great deal of attribute data for soils, such as bulk density and organic matter, to accompany spatial information (Loague, Corwin, and Ellsworth, 1998). STATSGO, SSURGO, and other soil files may be downloaded from the National Soil Survey Center (http://www.statlab.iastate.edu/soils/nssdf/), which is part of the U.S. Department of Agriculture’s Natural Resource Conservation Service.

Soil databases sometimes do not provide the required inputs for advanced models. In such cases, several methods can be used to obtain supplemental data. If spatial data are lacking, modelers can digitize NRCS county soil map sheets (Robinson and Ragan, 1993). If soil attributes are lacking, researchers can collect data by field measurements or by remote sensing techniques. Remote sensing techniques include electromagnetic induction, ground-penetrating radar, and seismic reflectance (Corwin and Wagenet, 1996). Hard-to-measure soil characteristics can often be estimated from parameters that are more easily measured. For example, water retention can be estimated from soil particle size distribution, bulk density, and organic carbon content by equations called pedotransfer functions (Loague, Corwin, and Ellsworth, 1998).

2.3.5 Transportation

Spatial transportation databases are available from the same sources as hydrography databases. USGS DLGs and Census TIGER/Line files provide roads, rail, utility lines, and miscellaneous transportation. State agencies, such as the Virginia Department of Transportation (VDOT), may also provide transportation files. VDOT has a detailed set of spatial data on roads and other transportation infrastructure.

2.4 Integrating Functions of GIS in Environmental Modeling and Watershed Analysis

The interdisciplinary nature and complexity of environmental and water resource problems require a modeling approach that can integrate information from each of the involved scientific fields (Leavesley et al., 1996). The functions of a GIS, or the tasks that it is capable of accomplishing, make it a highly effective tool in performing such integration. The major functions that GIS provides for integration are:

- Spatial inventory/feature characterization;
- Preprocessing and development of input data;
- Spatial analysis and data manipulations;
• Communication of results among models; and
• Provision of a common user interface.

These functions allow modeling steps such as data development, parameterization, and visualization to be integrated in simulation runs.

### 2.4.1 Spatial Inventory/Feature Characterization

The most obvious function that a GIS provides in a watershed model is inventory and characterization of spatial features. A GIS has built-in applications for description and management of environmental data. In fact, some highly practical GIS-based models rely mainly on this function to evaluate water resource problems. Australia’s LANDCARE GIS synthesizes and characterizes data on land conditions from many sources in an attempt to promote sustainable land use practices and prevent pollution (East and Wood, 1998).

### 2.4.2 Preprocessing and Development of Input Data

Another principal function of spatial data in GIS-based environmental models is to provide input data about geographic characteristics. In supplying this function, the GIS simplifies the heterogeneity of the land unit into a single value. These values, in turn, feed into mathematical equations (Kemp, 1993). The goal is to have the GIS obtain necessary parameters for hydrologic and other modeling processes through analysis of terrain, land cover, and other features (Maidment, 1993). Watershed models often need information on soils, land use, weather, topography, ecology, and other issues. GIS excels at managing, assimilating, and analyzing such massive quantities of data (Lovejoy, 1997).

This data assembly step is a crucial part of the modeling process, which is why it can be referred to as preprocessing. Data must be prepared such that they are compatible with the various models with which they interact. Preprocessing can include not only preparation of spatial data, but also development of tabular or attribute databases. The end result of preprocessing is the availability of a variety of databases (spatial and otherwise) suitable for the prescribed needs of environmental models (Leavesley et al., 1996).

### 2.4.3 Spatial Analysis and Data Manipulations

Spatial analysis functions of a GIS allow the user to understand interrelationships among ground features based on their locations and attributes (Aspinall, 1994). There are four base-level operations in spatial analysis:

1. **Reclassification** of map categories, which consists of repackaging existing information on the map (e.g., reclassifying ten types of land cover into two – disturbed and undisturbed).
2. **Overlay** of two or more maps, which results in delineation of new boundaries (e.g., a soil map is combined with a slope map to produce new areas showing where steep slopes and erodible soils might lead to serious sediment runoff).
3. Measurement of **distance and connectivity** (e.g., measuring the area of a watershed, determining which streams flow into the river, or finding which farms are adjacent to bodies of water).
4. Characterization of *cartographic neighborhoods*, which creates new maps where the value assigned to each location is a function of surrounding values (e.g., the overall slope of a land area can be computed by averaging slopes for surrounding areas) (Berry, 1993).

Many specialized forms of spatial analysis that are related to these four operations also provide functionality for environmental modeling. Examples include buffering, determination of containment, calculation of slope and aspect, watershed delineation, and determination of downstream flows in a river network.

2.4.4 Communication of Results among Models

The burden of developing a spatial environmental decision support system is on development of interfaces that allow data transfer among applications (Djokic, 1996). Such transfer processes enable individual environmental models to communicate with each other. The complexity of environmental processes, rife with feedback loops and multifaceted relationships, demands that a robust model contains mechanisms for communication. GIS software, which typically possesses interfaces to common database formats, is a proficient setting for developing such a model (Djokic, 1996).

2.4.5 Provision of a Common User Interface

In working with a set of environmental models or a decision support system, a user needs an interface to interact with the system. Such interaction occurs mainly in two processes: defining input scenarios and viewing/analyzing results of simulation runs. GIS, with its highly visual character, provides the user a flexible environment for relating to all the components of the system.

2.4.5.1 Input Scenarios

Models can be very adept at analyzing processes that occur in nature and society, but such models often require vast amounts of input data. A GIS interface provides efficient and cost-effective methods for supplying such data (Krummel et al., 1996).

2.4.5.2 Display of Results

The display functions of GIS are a major reason for the technology’s proliferation. GIS can produce powerful visual results that, if well rendered, are easy to analyze. In fact, maps based on GIS display functions can be so simple to generate and precise looking, that they may hide underlying errors in the spatial data or overstate the dependability of model results. Common tools in GIS display functions, such as zooming and panning (Nyerges, 1993) allow modelers great flexibility in producing desired output. Other more advanced functions, like the ability to generate model results in 3-D perspective, can aid interpretation of the results (Kemp, 1993).

2.5 Types of Models – Model Architecture

In using GIS for integrated environmental modeling, two categories of model type or architecture are apparent. The categories are distinguished by the amount of integration between the GIS
software and the environmental modeling software. The less integrated model type is the coupled or linked model, and the more integrated type is the holistic model.

2.5.1 Coupled or Linked Models

Coupled models represent the conceptually simplest way to integrate GIS and environmental models. They involve passing relevant information (in the form of data files) between a GIS and a stand-alone environmental model. In the coupling process, the GIS minimally provides:

- A consistent spatial element or map unit for the modeling process;
- Input parameters for the environmental model; and
- Map displays of output from the model (Poiani and Bedford, 1995).

Four general steps are required to couple a GIS with an environmental model (Corwin and Wagenet, 1996):

1. Description of transformations required to make data compatible between the two applications;
2. Specification of software to import and export data between the two;
3. Development of the software to run without intervention; and
4. Formation of bidirectional data transfer capabilities.

In a straightforward version of a coupled model, therefore, the environmental model may read input data from the GIS and produce output files that can be displayed by the GIS (Figure 1. Fedra, 1993). Coupling models with this type of integration may require specialized software or additional computer code to let the GIS format its data in a way that the modeling software can understand or vice versa (Goodchild, 1993).

![Figure 1. Coupled Model with Integration Occurring through Shared Files.](image-url)
More deeply integrated applications can be developed if the GIS and modeling software use the same database structures (Goodchild, 1993). These applications share a common user interface and pass information between the GIS and the environmental model(s) more freely (Figure 2, Fedra, 1993). If the goal of integrating GIS and environmental models is to produce a decision support system, this type of model architecture may be particularly useful. Most decision support systems do not require the full functionality of a complete GIS. Limited functionality, such as map displays, can be integrated efficiently in this architecture while preserving flexibility and performance (Fedra, 1996).

![Diagram of coupled model with deeper integration and a common user interface.](image)

**Figure 2. Coupled Model with Deeper Integration and a Common User Interface.**

2.5.1.1 Advantages

The coupled architecture is advantageous when the process being modeled is complex, the process needs to be modeled by experts, and/or processing on dedicated hardware or use of special programming languages is required (Burrough, van Rijn, and Rikken, 1996). Perhaps a more notable advantage of a coupling architecture is that it allows the designer to use time-tested, accepted modeling software.

2.5.1.2 Disadvantages

There are several disadvantages of the coupled approach. The first is that linkage of environmental models that may not have been designed for such linkages can be unwieldy and error-prone. This trait frequently necessitates extensive debugging. In addition, in assuring that components of a coupled model work together, files need to be converted to compatible data formats. Coupled systems also do not allow user interaction and query during a simulation, and often do not effectively encapsulate temporal relationships (Bennett, Armstrong, and Weirich, 1996). Another potential disadvantage is that modelers must overcome differences in spatial and temporal resolution that occur among models.
2.5.1.3 Example

An example of a coupled model is the U.S. Environmental Protection Agency’s BASINS (Better Assessment Science Integrating Point and Nonpoint Sources) model, which combines GIS with national watershed data and environmental assessment and modeling tools (USEPA, 1999). Specifically, BASINS couples HSPF (a hydrologic model), TOXIROUTE (a model of chemical dilution and decay in streams), and QUAL2E (a water quality and eutrophication model) with ArcView GIS. BASINS is designed for use by regional, state, and local agencies in performing watershed-based and water quality studies. The objectives of the software design are (1) to facilitate examination of environmental information; (2) to provide an integrated watershed and modeling framework; and (3) to support analysis of point and nonpoint source pollution management alternatives, especially in developing total maximum daily loads (TMDLs) for waterways (USEPA, 1999). BASINS uses the coupled approach because it models complex processes and uses special, preexisting water quality models. More information on BASINS is available at EPA’s web site (http://www.epa.gov/ostwater/BASINS/).

2.5.2 Holistic GIS Models

Holistic GIS/environmental models are those which are run solely by GIS software. Analysis, calibration, and other model functions are processed by the GIS’s command language (Goodchild, 1993). Such a holistic GIS requires a specialized interface to be practical for an environmental modeler. For a holistic nonpoint source pollution model, the minimum suggested abilities of the interface include (Adapted from Goodchild, 1993):

- **Use of a single variable to represent a continuous geographic variation** -- In spatial distributions, the term *continuous variation* refers to surfaces with values that continuously change with distance, such as elevation or temperature (as opposed to *discrete variation*, which refers to surfaces that have constant values within a given spatial boundary, such as a county or watershed). For example, the user should be able to assign a sole variable to represent elevation.

- **Calculation functions** – The GIS should allow variables that represent continuous geographic variation to be combined and altered in equations specified by the user.

- **Computations with common environmental equations** – The GIS’s programming language should have built-in capabilities to apply equations often used in environmental modeling.

Holistic models require sufficiently modular tools to allow various software components to share memory and functions within a single application with a common user interface. Because of the level of integration in these models, they can be expensive to develop. Often the expense can be worthwhile, however, since holistic models can be very powerful as planning and decision support tools (Fedra, 1993).

2.5.2.1 Advantages

The holistic architecture allows an environmental process to be modeled using generic tools on a single integrated database (Burrough, van Rijn, and Rikken, 1996).
2.5.2.2 Disadvantages

The most apparent disadvantage of the holistic approach is that it may be difficult to write robust environmental models based on the available GIS functionality and command language. This disadvantage is often what prevents development of holistic models. Another limitation of holistic models has to do with the representation of time (McDonnell, 1995). 3-D environmental models that use time as a major parameter do not have a parallel data scheme in 2-D GIS systems (Nyerges, 1993). While currently marketed GIS systems have excellent flexibility in representing spatial relationships, they do not have full data representation flexibility for space and time (Corwin and Wagenet, 1996) – see the section below on spatial and temporal resolution.

2.5.2.3 Example

The Watershed Delineator, developed by the Texas Natural Resource Conservation Commission and the Environmental Systems Research Institute, is an example of an environmental model with a holistic architecture (TNRCC and ESRI, 1997). The Watershed Delineator is a set of procedures, tools, and utilities for ArcView GIS software that allow a user to delineate watershed boundaries on a digital elevation model. The Delineator is a series of algorithms developed in ArcView’s own command language. Inclusion of the delineation algorithms within the GIS precludes sharing of files between the GIS and a separate delineation program. This holistic approach makes the process of modeling watershed boundaries more efficient, but it requires alterations to the GIS user interface.

2.6 Technical Considerations in Building GIS-Based Environmental Models

In building successful GIS-based environmental models, modelers must address a number of technical considerations that affect the overall model development process. Modelers often must make decisions about such technical considerations based on the particular characteristics of their projects.

2.6.1 Selection of Spatial Data Structures

GIS themes generally use one of two common spatial data structures – raster or vector. Raster data structures typically lay a rectangular grid of homogenous cells over a map area. Any point location in the map area falls into a particular cell, and each cell in the raster holds a piece of attribute data. For this reason, the raster data structure is often used to represent continuous phenomena like elevation, soils, and land use. In contrast, the vector data structure uses points, lines, and polygons to represent spatial features. These spatial features do not necessarily cover the entire map area, so vector structures are often used to portray discrete phenomena (objects that have clearly defined boundaries, such as trees, streams, buildings, and counties) (Kemp, 1993). An example of the difference between the two structures can be demonstrated by how each portrays a lake. The raster structure defines the lake by its interior, identifying cells that represent open water. The vector structure defines the lake by its border, which separates the exterior from the interior (Berry, 1993).

Deciding which of the two data structures to use for a particular data theme is a significant issue for the modeler. In some cases, spatial data sets are provided to the modeler in one form or the
other, but often the modeler must choose a data structure based on how the environmental model discretizes space. For example, finite difference hydrologic models break up areas into rectangular grids (Maidment, 1993) that are well suited for interaction with raster data in a GIS.

Relatively new structures, called hierarchical spatial data structures, seek to extend the capabilities of raster and vector structures. Hierarchical structures (e.g., quadtrees) apply the principle of recursive decomposition of space, a characteristic that is similar to the conceptual framework of many environmental models (Csillag, 1996).

2.6.2 Spatial and Temporal Scale/Resolution

Scientists who are attempting to combine remotely sensed and ground observation data into a GIS for simulation and analysis are confronted with data of varying spatial and temporal scales and resolutions (Bromberg et al., 1996). This variation can generate serious difficulties in efforts to integrate GIS data and environmental models.

The familiar concept of a map scale (ratio of length on a map to length on the ground) assumes a major role when GIS and watershed models work together. Many environmental models are dependent on scale for setting parameters and representing processes. Some models work on a global scale, while others operate at a much finer level. For example, climate models often function globally, while nutrient cycling models can operate at the scale of a single plant form. (Nemani et al., 1993). Scale, furthermore, determines the predominant processes that are influential and need to be modeled (Loague, Corwin, and Ellsworth, 1998). In a GIS, scale determines the level of detail and amount of generalization in spatial data. These functions of scale in a GIS-based model make choosing an overall scale for a particular project very difficult. The modeler must account for the processes being modeled and be aware of the availability of data at desired scales (Nemani et al., 1993).

GIS databases interpret the physical world as a set of basic entities, such as points, lines, polygons, or pixels that carry attributes. This interpretation is limited because it excludes proper treatment of many continuous and stochastic phenomena, such as spatial and temporal variation of land forms, soil, biological systems, and water. For some of the processes in environmental modeling, variations of spatial patterns, both across distances and over time, are essential. In GIS, continuous variation is represented by discrete units. The size of units for approximating continuous variation determines the level of spatial and temporal detail that can be resolved. (Burrough, van Rijn, and Rikken, 1996). At the extreme ends of the spectrum, patterns mapped at a spatial resolution that is too fine appear as noise, while patterns mapped at a spatial resolution that is too coarse appear as constants (Aspinall and Pearson, 1996).

In addition to representing the level of detail in a feature’s position (spatial resolution), a GIS has to deal with the fact that physical features of the natural environment tend to change over time. For example, a coastline that has been pummeled by a large storm can take an altered shape. The level at which a GIS is able to represent changes over time is referred to as temporal resolution. Unfortunately for environmental modelers, most commercially available GIS software packages have very limited temporal resolutions and cannot adequately model landscape features that undergo rapid changes. A GIS often contends with temporal changes by storing a feature that exhibits different states on multiple layers (Raper and Livingstone, 1996). For instance, land use
scenes of a particular area are to be used in a GIS-based ecological model. The area has experienced increasing deforestation over time. The GIS stores “snapshots” of the area from different times as individual data layers to depict the changing nature of the scene. This method of handling change can be insufficient for interaction with environmental models that require data at a fine level of temporal detail.

Issues of spatial and temporal resolution can be complex, as these two classes of resolution are often related in describing a physical phenomenon. In fact, some scientists define spatiotemporal hierarchies that describe the interplay of spatial and temporal resolutions. For example, researchers in the field of landscape ecology have portrayed a series of forest processes (in a space-time diagram) as occurring at varying spatiotemporal resolutions (Figure 3) (Johnson, 1996). Modelers can face stiff challenges in trying to account for these spatiotemporal phenomena.

![Space-Time Diagram for Forest Processes](image)

**Figure 3. Space-Time Diagram for Forest Processes.**

2.6.3 Projection

In using GIS for environmental modeling, many spatial data themes may be collected, analyzed, and displayed together. For such processes to occur successfully, each data theme must be
projected in the same way. Projection is the process that allows a three-dimensional section of the earth to be displayed on a flat map or computer screen. A theme’s projection provides a coordinate system that identifies map locations which can be referenced to locations on the earth. Examples of common U.S. coordinate systems that are based on different projections are Universal Transverse Mercator (UTM) coordinates (Kemp, 1993) and the State Plane Coordinate System. If one theme is not projected to the same coordinate system as another theme, the GIS functions will not work properly on that theme.

2.6.4 Interdisciplinary Collaboration

Attempts to solve complex problems, such as those that occur in the watershed, require interdisciplinary and interagency collaboration for successful analysis. Modeling systems need to process individual perspectives in an integrative domain (Bromberg et al. 1996). Construction of such a model, therefore, requires cooperation among GIS analysts and scientists interested in the watershed, such as hydrologists, biologists, and economists.

2.6.5 Object Oriented Programming and GIS

Object orientation is a powerful tool for environmental modeling, especially for water resources. To represent a water resources system digitally, the structure of the system and its behavior must be captured in the digital domain. The structure of the water resources system refers to the absolute and relative locations of entities (e.g., rivers), along with thematic attributes of those entities (e.g., name). The behavior refers to the dynamic processes associated with the entities of the system, such as flow characteristics (Bennett, Armstrong, and Weirich, 1996). The object-oriented approach, which treats these entities as programmed objects, is a natural fit for water systems.

Beyond programming benefits, object oriented approaches to data representation and modeling have two major advantages over other approaches in GIS-based environmental modeling. These are (1) separation of attribute information from process information and (2) use of inheritance hierarchies. The first advantage can be explained by an example. In an object oriented model of a river basin, reservoirs are represented as objects. Databases store attributes about reservoirs, such as volume or elevation. Separate data structures store process information in the form of functions, such as evaporation or mass balance. Both types of data structures, attribute and process, can be related or tied to reservoir objects. With access to both attribute and process information, each reservoir object can monitor its own state. The second advantage, inheritance hierarchies, helps make modeling more efficient by allowing objects to inherit properties and functionality. For instance, general process information, such as evaporation algorithms, can be preserved at the reservoir class level. Each reservoir object in this class, then, can inherit that process information (Reitsma, 1996).

In an object oriented structure for integrating GIS and environmental models, there are two main classes of objects, spatial and thematic. Examples of spatial objects are the globe, countries, cities, river basins, and forests (Fedra, 1996). These objects can be described as having political or natural boundaries. Examples of thematic objects are indicators, raw time series data, and descriptors. The modeling system contains methods for linking these two classes of objects, providing a high degree of flexibility (Fedra, 1996).
2.7 Modeling Efforts/Case Studies

Numerous studies have been conducted in recent years that use GIS in modeling water resources. Many of these studies involve innovative coupled models, and others use holistic GIS models. Table 1 provides summary-level information about selected coupled-model studies, and Table 2 lists information about holistic models (adapted and expanded from tables provided by Poiani and Bedford, 1995). The information in these two tables demonstrates recent advances in GIS-based models and illustrates their widespread acceptance by the research community.

Table 1. Selected Studies Involving Coupled GIS/Environmental Models.

<table>
<thead>
<tr>
<th>Reference</th>
<th>GIS</th>
<th>Environmental Model(s)</th>
<th>Area</th>
<th>Problem</th>
<th>Focus of Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bromberg et al., 1996</td>
<td>Genamap, S-Plus, and Geo-EAS</td>
<td>CENTURY ecosystem model</td>
<td>Flexible – can be applied to any ecosystem size</td>
<td>Analysis of carbon and nutrient cycling</td>
<td>Development of a user friendly ecosystem simulation environment.</td>
</tr>
<tr>
<td>Dikshit and Loucks, 1996-7</td>
<td>ERDAS</td>
<td>CNPS NPS pollution model</td>
<td>Fall Creek watershed, NY; 330 km²</td>
<td>NPS – sediment, nutrients</td>
<td>Creation of a tool to predict the influence of land use changes on N and P loadings.</td>
</tr>
<tr>
<td>Djokic, 1996</td>
<td>Arc/Info</td>
<td>HEC-1 rainfall and runoff model</td>
<td>General purpose – watershed scale</td>
<td>Runoff tracking in a watershed</td>
<td>Creation of a flexible tool for modeling rainfall and runoff in a watershed.</td>
</tr>
<tr>
<td>Engel et al., 1993</td>
<td>GRASS</td>
<td>ANSWERS, AGNPS, and SWAT NPS models</td>
<td>Indian Pine Nat. Resources Field Station, IN; 830 acres</td>
<td>NPS – sediment, nutrients</td>
<td>Comparison of coupled model results to measured runoff values.</td>
</tr>
<tr>
<td>Gao et al., 1993</td>
<td>GRASS</td>
<td>NASA-EOS distributed rainfall-runoff model</td>
<td>Lucky Hills-104 watershed, AZ; 4.4 ha</td>
<td>NPS – sediment</td>
<td>Coupling of GIS to a physically based distributed model.</td>
</tr>
<tr>
<td>Garnier et al., 1998</td>
<td>IDRISI and GRASS</td>
<td>GLEAMS NPS model</td>
<td>River Chiana, Tuscany, Italy;</td>
<td>NPS – Nitrate</td>
<td>Development of an inexpensive model to support decisions about animal waste disposal.</td>
</tr>
<tr>
<td>Harris et al., 1993</td>
<td>Arc/Info</td>
<td>Coupled, fluid, energy, and solute transport (CFEST)</td>
<td>San Gabriel Basin, CA; 170 mi²</td>
<td>NPS – VOC</td>
<td>Use of numerical models to understand groundwater flows.</td>
</tr>
<tr>
<td>Kim and Ventura, 1993</td>
<td>Arc/Info</td>
<td>Source Loading &amp; Management Model (SLAMM)</td>
<td>Kinnickinnic River watershed; Milwaukee, WI; 27 mi²</td>
<td>NPS – heavy metals, nutrients, sediment</td>
<td>Establishment of effective urban BMPs based on model outputs of pollution.</td>
</tr>
<tr>
<td>Kim et al., 1993</td>
<td>Arc/Info</td>
<td>SLAMM</td>
<td>City of Beaver Dam, WI; 4.78 mi²</td>
<td>NPS – heavy metals, nutrients, sediment</td>
<td>Use of GIS and local NPS model to help city meet EPA stormwater regs.</td>
</tr>
<tr>
<td>Reference</td>
<td>GIS</td>
<td>Environmental Model(s)</td>
<td>Area</td>
<td>Problem</td>
<td>Focus of Study</td>
</tr>
<tr>
<td>----------------------------</td>
<td>-----------</td>
<td>------------------------</td>
<td>-------------------------------------------</td>
<td>---------------------------------------</td>
<td>-------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Krummel et al., 1996</td>
<td>Arc/Info</td>
<td>ANSWERS</td>
<td>Cadiz Township in Green County, WI, 10,000 ha</td>
<td>Sediment runoff in the river basin</td>
<td>Analysis of the effects of land cover changes on hydrologic parameters.</td>
</tr>
<tr>
<td>Liao and Tim, 1997</td>
<td>Arc/Info</td>
<td>AGNPS</td>
<td>Westlake watershed, IA; 6,500 acres</td>
<td>Ag. Runoff and pesticide pollution</td>
<td>Close linkage of a distributed NPS model to a GIS interface.</td>
</tr>
<tr>
<td>Srinivasan and Arnold, 1994</td>
<td>GRASS</td>
<td>SWAT</td>
<td>Seco Creek basin, TX; 114 km²</td>
<td>NPS – sediment, pesticide, nutrients</td>
<td>Development of a continuous-time, distrib.-parameter tool to manage basin-scale runoff.</td>
</tr>
<tr>
<td>Srinivasan and Engel, 1994</td>
<td>GRASS</td>
<td>AGNPS</td>
<td>Seco Creek basin, TX; 114 km²</td>
<td>NPS – sediment, nutrients, COD</td>
<td>Development of a spatial decision support system for NPS pollution from agricultural lands.</td>
</tr>
<tr>
<td>Tim and Jolly, 1994</td>
<td>Arc/Info</td>
<td>AGNPS</td>
<td>Bluegrass watershed, IA; 1033 acres</td>
<td>NPS – sediment</td>
<td>Application of GIS to model runoff to streams in an ag. Watershed.</td>
</tr>
<tr>
<td>USEPA, 1999</td>
<td>ArcView</td>
<td>HSPF, QUAL2E, and TOXIRoute</td>
<td>Any major USGS Cataloging Unit watershed</td>
<td>NPS loads, eutrophic., and toxics</td>
<td>Development of TMDLs and water quality simulations</td>
</tr>
</tbody>
</table>
Table 2. Selected Studies Involving Holistic GIS/Environmental Models.

<table>
<thead>
<tr>
<th>Reference</th>
<th>Software for GIS/Env. Model</th>
<th>Area</th>
<th>Problem</th>
<th>Study/Model Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adamus and Martinus, 1995</td>
<td>Pollution Load Screening Model (PLSM) – built in Arc/Info</td>
<td>Lower St. John’s River basin, FL; 2,754 mi²</td>
<td>NPS pollution – sediment, nutrients, TSS, BOD, lead, and zinc</td>
<td>Estimation of annual NPS pollutant loads to surface waters for regulatory prioritization.</td>
</tr>
<tr>
<td>Bennett et al., 1996</td>
<td>Object oriented programming environment</td>
<td>Hypothetical watershed</td>
<td>Hydrologic flow and sediment and debris tracking</td>
<td>Demonstration of a flexible programming environment for building watershed models.</td>
</tr>
<tr>
<td>Fraser, Barten, and Pinney, 1998</td>
<td>SEDMOD – built in Arc/Info with specialized applications</td>
<td>Saw Kill watershed, NY; 12 subwatersheds from 1.5 to 50 km²</td>
<td>NPS pollution – fecal coliform</td>
<td>Development of a model that uses available spatial data to help regulators prioritize sites for NPS pollution control.</td>
</tr>
<tr>
<td>Heidtke and Auer, 1993</td>
<td>Regional GIS from the Cayuga County Planning Board with USLE and empirical loading functions</td>
<td>Oswasco Lake watershed, NY</td>
<td>NPS pollution – sediment, phosphorus</td>
<td>Development of model that predicts NPS pollution as a result of land use changes without extensive input data requirements.</td>
</tr>
<tr>
<td>Mattikalli and Richards, 1996</td>
<td>Arc/Info with an export coefficients model</td>
<td>River Glen watershed; S. Lincolnshire, England; 330 km²</td>
<td>NPS pollution – nitrogen</td>
<td>Use of GIS and remote sensing to predict the effects of land use on solute transportation in runoff.</td>
</tr>
<tr>
<td>Pilesjo, 1992</td>
<td>Arc/Info with USLE adapted for Ethiopian conditions</td>
<td>Shewa Province in central Ethiopia; 600 km²</td>
<td>NPS pollution – sediment</td>
<td>Use of remote sensing and GIS for modeling soil erosion in semi-arid regions.</td>
</tr>
<tr>
<td>Reitsma, 1996</td>
<td>Object oriented River Simulation System (RSS)</td>
<td>Variable – handles river basins of any size</td>
<td>River basin management</td>
<td>Use of object oriented programming and GIS topology to simulate river and reservoir hydrology.</td>
</tr>
<tr>
<td>Rifai, Newell, and Bedient, 1993</td>
<td>Arc/Info</td>
<td>Galveston Bay system, TX; 1,430 km²</td>
<td>NPS pollution – TSS, N, P, BOD, fecal coliform, heavy metals, oils, SOC</td>
<td>Quantification and mapping of eight priority pollutants for water quality characterization</td>
</tr>
<tr>
<td>Robinson and Ragan, 1993</td>
<td>Microsoft BASIC Professional Development System 7.1</td>
<td>Kensington Quad; Montgo. County, MD; 37,000 acres</td>
<td>NPS pollution -- N, P, Zn, Pb, BOD, and sediment</td>
<td>Development of a special purpose GIS using a NPS model developed by the No. VA Planning District</td>
</tr>
<tr>
<td>TNRCC and ESRI, 1997</td>
<td>ArcView, Avenue Programming Language</td>
<td>Any USGS DEM file</td>
<td>Any</td>
<td>Development of watershed and subwatershed boundaries.</td>
</tr>
</tbody>
</table>
3.0 Study Area

3.1 Main Study Area

The study area for this research project is the 1,480-square-km Upper Roanoke River Watershed, with its outlet point set just downstream of the confluence of the Upper Roanoke River and Back Creek (Figure 4). This watershed lies in southwestern Virginia within zone 17 of the Universal Transverse Mercator (UTM) coordinate system, reaching into four counties and two cities – Roanoke, Montgomery, Floyd, and Botetourt Counties, and the cities of Roanoke and Salem (Figure 5).

Land uses in the area range from the urban streets of Roanoke to the heavily forested slopes of Jefferson National Forest. Hydrologic conditions within the watershed are as diverse as land uses. Elevation varies from 237 to 1,197 meters above sea level (Figure 6). Each of the models from the fields of hydrology, economics, and ecology is applied to this study area or some portion of it.

Figure 4. Upper Roanoke River Watershed. Note: This and all other maps are projected in the UTM Zone 17 coordinate system with the NAD 1927 datum.
Figure 5. Location of the Upper Roanoke River Watershed.

Figure 6. Physical Characteristics of the Upper Roanoke River Watershed.
3.2 Study Area for the Prototype Software

The study area for the prototype software that integrates the discipline-specific models is the 145-square-km Back Creek Subwatershed in the southeastern corner of the major study area (Figure 7). This subwatershed has the same outlet point as the Upper Roanoke River Watershed. Its variety of land uses and elevations is representative of the overall study area (Figure 8).

Figure 7. Location of the Back Creek Subwatershed.

Figure 8. Physical Characteristics of the Back Creek Subwatershed.
4.0 Methods

The methods used to achieve GIS integration in this project fit into three major categories:

1. Data selection and development – data collection efforts and procedures for enhancing spatial data.
2. Aid in development of discipline-specific models – Use of GIS functions to provide input and calculations necessary for discipline-specific models.
3. Software interface construction – Use of GIS and programming tools in tandem to produce a functional interface for DesktopL2W, the software modeling environment.

4.1 Data Selection and Development

A key step in building a GIS is selecting and formatting spatial data layers. This step is especially important when those layers, along with their underlying attribute data, provide vital input to modeling tools.

4.1.1 Development of Key Base Data Layers

Selection of data for a GIS that integrates various watershed models requires careful inspection of the needs of the models, but there are some data sets that are necessary for almost any analysis of a watershed. These data sets are referred to as base data layers, and they include DEMs, DLG streams, watershed boundaries, and land use.

DEM Mosaic

DEMAs provide both watershed delineation capabilities and input values to the discipline-specific models, so it is very important to obtain a full coverage of elevation data. The overall study area, the Upper Roanoke River Watershed, is covered by 24 individual DEMs of thirty-meter resolution at a scale of 1:24,000 (Figure 9). The 24 DEMs are joined in a seamless mosaic to provide continuously varying elevation data over the whole study area. For information on the construction of this mosaic, see Appendix A.
DLG Network

DLG streams, like DEMs, play a role in watershed delineation, but they also provide accurate locations of waterways, which are essential for various model inputs and map displays. The Upper Roanoke River Watershed is covered by six DLG files that are based on 1:100,000 scale USGS topographic quadrangles (Figure 10). The streams from each quadrangle are linked to form an unbroken network. For more information on the DLG stream network, see Appendix A.
Watershed Boundaries

Watershed boundaries are a fundamental base data layer because they define the study area. In evaluating the impacts of land use changes on the watershed, the watershed in question must be clearly defined. In addition, models such as HSPF and the fisheries health model require subsets of the main watershed, or subwatersheds, as units of analysis (Figure 11). The Upper Roanoke River Watershed, Back Creek Subwatershed, and other subwatersheds in the study area are created using the GIS-based Watershed Delineator (TNRCC, 1997 and Maidment, 1996) extension for ESRI’s ArcView. The Watershed Delineator applies a method for establishing
flow direction and flow accumulation to DEMs with inlaid stream lines. For more information on the production of watershed boundaries, see Appendix A.

Figure 11. Division of Back Creek Subwatershed into Ten Smaller Subwatersheds.

Land Use

Land use for the study area (Figure 12) is a clipped portion of the draft land cover grid of Virginia produced through the Virginia Gap Analysis Project (VAGAP). VAGAP interprets fourteen LANDSAT Thematic Mapper scenes, using several methods to create the grid. The grid, which has a pixel resolution of thirty meters, is georeferenced in UTM zone 17 coordinates with NAD 1927 as the datum. Major classes of pixel values include deciduous forest, coniferous forest, mixed forest, herbaceous mostly agriculture, open water, and disturbed/lack of vegetation.
In addition to DEMs, DLGs, watershed boundaries, and land use, the GIS for this study includes base data such as flood plains from the Federal Emergency Management Agency (FEMA), land ownership/tax parcels from Roanoke County, soils from the Natural Resource Conservation Service (NRCS), and TIGER files from the U.S. Census. Some of the base data layers also play a role in creating derived layers that act like base data. For example, the DEM mosaic is used to derive a slope grid, which in turn provides input data for discipline-specific models. For more information on base and derived data layers, see the metadata tables (Table 9-Table 15) in Section 5.1.

4.1.2 Construction of Layers for the Software

A major result of this study of GIS integration is a functional watershed analysis software tool called DesktopL2W, which stands for the desktop version of the Landscapes to Waterscapes problem solving environment. A problem solving environment is a form of decision support system that provides access to various software components in a way that allows the user to think in terms of the higher-level problem domain rather than in terms of the lower-level simulation details (Rubin and Shaffer, 1999). A critical process in building such a GIS-based tool is development of spatial data layers. The layers are necessary for two main purposes:

- They allow the user to have access to a visual interface for simulating and understanding land use changes in the watershed; and
• They allow for data storage and calculations that feed discipline-specific modeling efforts.

DesktopL2W uses a number of these spatial data layers. Some are absolutely essential to its operation, while others may be accessed at the user’s discretion to provide more in-depth information. The two most essential layers, the development grid and the development tracts, are discussed in detail. Other layers are summarized in Section 5.1 (Table 16).

Development Grid

The development grid is the template upon which the user alters the landscape by adding development tracts (Figure 13). The grid layer for Back Creek Subwatershed consists of 1,744 squares, each of which is 300 meters by 300 meters (nine hectares). The spatial structure of the layer was constructed by the following five processes:

1. Exportation of the Back Creek Subwatershed shapefile to a .dxf AutoCAD file using ArcView’s DXF Conversion Extension;
2. Delineation of a nine-hectare square in AutoCAD;
3. Replication of the square (with subsequent end-to-end placement) enough times to build a grid of squares to cover the area of the subwatershed;
4. Importation of the grid to ArcView using the CAD Reader Extension; and
5. Selection and deletion of those squares that do not intersect the subwatershed.

![Development Grid](image_url)

*Figure 13. Development Grid.*
The grid not only sets the size of the tracts, but it also suggests which tracts are available for development. The uncolored squares are developable, while the colored squares are considered undevelopable. The distinction is based on a raster overlay of four spatial data layers: slope, land use, preservation status, and flood plain location. In the overlay, the pixels in each of the four layers are reclassified with a value of either 0 or 1. Which value is assigned depends on whether the original value meets the criteria for developability. For example, pixels with average slopes of less than 20% are developable and are assigned a 0, while those over 20% are not developable and are assigned a 1. Each of the other layers is reclassified in a similar method (Table 3). If any one of the four layers for a pixel is equal to one, then the pixel is not developable. Overall developability for a particular pixel, therefore, is achieved by summing the values for each of the four layers. If they sum to zero, then the pixel is developable; if the sum exceeds zero, then it is not developable (Figure 14).

Table 3. Method of Raster Overlay for Determining Developability.

<table>
<thead>
<tr>
<th>GIS Layer</th>
<th>Criteria</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Slope</td>
<td>&gt;20%</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>&lt;20%</td>
<td>0</td>
</tr>
<tr>
<td>Land Use</td>
<td>Disturbed and Water</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Forest and Herb/Ag</td>
<td>0</td>
</tr>
<tr>
<td>Preservation Status</td>
<td>Preserved</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Unpreserved</td>
<td>0</td>
</tr>
<tr>
<td>Flood Plain Location</td>
<td>Inside Flood Plain</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Outside Flood Plain</td>
<td>0</td>
</tr>
<tr>
<td>Raster Overlay</td>
<td>Sum of Values:</td>
<td>0 = Developable</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&gt;0 = Undevelopable</td>
</tr>
</tbody>
</table>
Development Tracts

The development tracts layer is a shapefile that is written by DesktopL2W after the user specifies a development scenario for the landscape. The user has the option of placing each of the following five types of development tract on the landscape:

- Low density residential;
- Mid density conventional residential;
- Mid density cluster residential;
- High density residential; and
- Commercial.

The residential tracts that the user places on the landscape take their attributes from standard templates (Arendt, 1994; Center for Watershed Protection, 1998; and Stephenson, 1999). Each template tract accommodates fifty residential dwelling units and 125 people, and possesses a unique combination of land uses (Table 4 - Table 7). Commercial tracts derive attributes from a sample commercial area in the town of Blacksburg, VA (Figure 15). In addition, commercial tracts are assumed to house zero people and have the same number of parcels as high density residential tracts. The attribute data for each type of development tract is rescaled to an area of 9 hectares (22.2 acres) to fit the resolution of the development grid.
Table 4. Characteristics of Low Density Residential Tracts.

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Size of Tract</td>
<td>150 acres</td>
</tr>
<tr>
<td>Percent Impervious</td>
<td>10 %</td>
</tr>
<tr>
<td>Percent Lawn/Agriculture</td>
<td>15 %</td>
</tr>
<tr>
<td>Percent Lawn or Forest (Mixed)</td>
<td>70 %</td>
</tr>
<tr>
<td>Percent Forest</td>
<td>5 %</td>
</tr>
<tr>
<td>Household Effluent Disposal</td>
<td>Septic</td>
</tr>
</tbody>
</table>

Table 5. Characteristics of Mid Density Conventional Residential Tracts.

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Size of Tract</td>
<td>21 acres</td>
</tr>
<tr>
<td>Percent Impervious</td>
<td>27 %</td>
</tr>
<tr>
<td>Percent Lawn/Agriculture</td>
<td>50 %</td>
</tr>
<tr>
<td>Percent Lawn or Forest (Mixed)</td>
<td>23 %</td>
</tr>
<tr>
<td>Household Effluent Disposal</td>
<td>Sewer</td>
</tr>
</tbody>
</table>

Table 6. Characteristics of Mid Density Cluster Residential Tracts.

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Size of Tract</td>
<td>21 acres</td>
</tr>
<tr>
<td>Percent Impervious</td>
<td>20 %</td>
</tr>
<tr>
<td>Percent Lawn/Agriculture</td>
<td>35 %</td>
</tr>
<tr>
<td>Percent Lawn or Forest (Mixed)</td>
<td>45 %</td>
</tr>
<tr>
<td>Household Effluent Disposal</td>
<td>Sewer</td>
</tr>
</tbody>
</table>

Table 7. Characteristics of High Density Residential Tracts.

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Size of Tract</td>
<td>12 acres</td>
</tr>
<tr>
<td>Percent Impervious</td>
<td>35 %</td>
</tr>
<tr>
<td>Percent Lawn/Agriculture</td>
<td>60 %</td>
</tr>
<tr>
<td>Percent Forest</td>
<td>5 %</td>
</tr>
<tr>
<td>Household Effluent Disposal</td>
<td>Sewer</td>
</tr>
</tbody>
</table>
4.2 GIS Methods Used to Develop Discipline-Specific Models

Prior to building the final software product, GIS performs an integrating function by tabulating data and calculating inputs for each watershed model. By using GIS methods that build upon base data sets, the hydrology, economics, and fisheries models adopt a degree of common ground spatially.

4.2.1 GIS Methods for Hydrology

The surface hydrology group of the project employed HSPF (Hydrologic Simulation Program Fortran) for simulating hydrologic responses to land use change. HSPF is one of the most comprehensive hydrologic models used to simulate water quantity and quality processes that occur as a result of changes in land use or meteorological inputs. GIS data layers, such as DEMs, DLG streams, and land use, were used to provide physical watershed data for HSPF.
A fundamental input to HSPF when modeling a watershed is the physical characteristics of the land segments that make up the watershed. For the HSPF application, the watershed is divided into hydrologically homogeneous land segments based on the drainage patterns of various stream reaches. HSPF requires travel times through reaches that approximate the simulation time step (the time step defines the unit of time used in an HSPF simulation to calculate hydrologic response – in this case, the time step is one hour). The following methods were used to calculate travel times and delineate ten land segments for the Back Creek Subwatershed (Chanat, 1999, Figure 11):

1. Schematization of the stream network into reaches by selection of main stem, permanent channels.

2. Approximation of appropriate reach length to realize a travel time around one hour.

3. Demarcation of the most upstream reach with an outlet point placed downstream at a distance equal to the approximated reach length.

4. Calculation of average slope for the reach.

5. Estimation of hydraulic radius, velocity, and travel time for the reach based on cross-sectional data and Manning’s formula.

6. Adjustment of reach length if necessary to realize a travel time of about one hour.

7. Repetition of steps 3 through 6 until the whole watershed is divided into reaches with outlet points.

8. Use of the TNRCC Watershed Delineator to define subwatersheds from the outlet points.

9. Spatial subtraction (Figure 16) of upstream subwatersheds from downstream ones to assign land segments. Such spatial subtraction entails subtracting the area of one polygon from that of a larger polygon to obtain a separate polygon which is a subset of the larger polygon. In the figure, polygon numbers 1, 2, and 3 represent subwatersheds that are upstream of polygon number 4. Subtraction of polygons 1, 2, and 3 from the entire polygon results in polygon 4 as a separate entity (land segment or subwatershed number 4).
Figure 16. Spatial Subtraction.

The land segments created by this process are the units of spatial resolution for the HSPF model. Inputs to the model are, therefore, based on the land segments, and results are aggregated for each land segment. A major function of the GIS in aiding the hydrologic simulation is to determine the proportion of each type of land use for the land segments. Land use data are reclassified from the original VAGAP data set into five summative categories, forest, herbaceous/agriculture, open water, disturbed, and mixed (Figure 17). The land uses in this reclassified grid layer are then tabulated for each land segment (Table 8).
Beyond development of land segments and their associated land uses, other GIS functions were used in supporting the surface hydrology model as follows:

- Precipitation data (hourly record) from the IFLOWS rain gage network at the Blacksburg office of the National Weather Service were used by the hydrology group in calibrating the model. The rain gage locations were added to the GIS database using their latitude and longitude coordinates, and the range of influence of each gage was found by
delineating Theissen polygons. The polygons show four gages that influence Back Creek Subwatershed (Figure 18).

- The GIS database also displays the location of stream gages and stores data used for calibration for each one (Figure 18).
- For more information on GIS data layers and methods used for the hydrologic model, see Section 5.1 (Table 10-Table 16).

![Figure 18. Gages, Thiessen Polygons, and Other Hydrologic Entities in the Back Creek Area.](image)

4.2.2 GIS Methods for Economics

The economics model estimates the effects of land use changes on land values, public expenditures, and tax revenues. The model specifies a complex regression equation, which requires a great deal of input from the GIS and its associated spatial data layers. Variables related to land ownership parcels that are utilized by the economics model include parcel size, coordinates, elevation, taxonomic soil order, population density, distance to landscape features, and road attachment (Kaltsas and Bosch, 1999). These variables are all attributes of the ownership parcels for Roanoke County (Figure 19). Data layers, beyond the Roanoke County parcels, involved in specifying these variables include USGS DEMs, NRCS STATSGO soils, U.S. Census TIGER files, and USGS DLG roads.
Application of various GIS functions to the data layers provides the values of the variables needed for the economic land value regression (Figure 20). Standard GIS algorithms calculate the area and UTM coordinates of each parcel from the parcels layer itself. A GIS summarization process is employed to find mean elevation and predominant soil type for each parcel (for more information on preparation of the soil data, see Appendix A). Population density is calculated for each Census block group, and the block groups layer is converted to a raster. A summarization process is again employed to obtain mean block group population density for each parcel. Distance calculation algorithms provide a raster of distances from a particular landscape feature, such as a mall or the city center (Figure 21). The average value of this raster’s pixels that fall within each parcel provides the distance from that parcel to the feature. The Boolean variable of road attachment (true if a major road is attached to parcel, and false if it is
not) is assigned for each parcel simply by selecting those parcels that touch a major road, and updating their attribute data accordingly (for more information on preparation of the road data, see Appendix A).

Another economic modeling task that uses GIS functions is the calculation public expenditures associated with development of new tracts. The public expenditures that are evaluated are sewer and water line construction and connection costs. The GIS uses range finding algorithms to estimate the distance from potential tracts to existing sewer and water lines. Distance values are a key input in determining the costs to developers and the public for hooking new developments to utility lines (Speir, Bosch, and Stephenson, 2000).
4.2.3 GIS Methods for Fisheries

The fisheries model relates the land use in zones of influence to the health of fish populations in those areas. A statistic called the mean metric score rates the quality of health for fish populations. The mean metric score represents the estimated integrity of the fish community at each site by examining its taxonomic structure and ecological function. Streams with high integrity contain a large number of native species as well as species with specialized ecological characteristics. Streams with low integrity contain few native species, and the species that are present are tolerant of degraded conditions. The mean metric score combines values for 13 different metrics commonly used in applications of the Index of Biotic Integrity, such as number of native species and proportion of tolerant individuals. The actual score for a zone of influence can range from 0 (worst) to 1 (best) (Stancil and Orth, 2000).

To collect data for calculating mean metric scores, the fisheries team sampled 43 sites in the Upper Roanoke River Watershed. The sites are incorporated into the GIS using coordinates from maps and GPS units. The TNRCC Watershed Delineator provides the watershed boundaries upstream of these sample sites (Figure 22). The watersheds themselves, along with smaller portions of the watersheds, serve as the zones of influence in correlating land use characteristics to the mean metric score.
The GIS delineates a variety of areas to serve as zones of influence in analyzing the fisheries data. There are a total of 28 types of influence zones that range in size from a site’s entire watershed down to a piece of a stream buffer within that watershed. The following steps are used to construct the zones of influence:

- Watersheds for fish sampling sites are divided into three groups based on size (large, medium, and small).
- Streams are buffered at 30, 60, 90, and 150 meters.
- Circles of influence for each size of watershed are created by buffering the sample points within each watershed. Circles are defined to approximate 5, 10, 20, and 40 percent of the size of each category of watershed. In addition, for the large watersheds, circles of 2.5 percent are developed. Each circle is moved to a location approximately tangent to the sample point and containing as much of the watershed/stream habitat as possible (Figure 23).
• Stream buffers, circles of influence, and watersheds are all intersected to create zones of influence. Each combination of stream buffer distance and circle size is applied to develop a specifically sized zone of influence (Figure 24).

• Land uses (forest, agricultural/herbaceous, open water, disturbed, and mixed) are tabulated for each zone of influence much like the tabulations for the hydrology model (Table 8).

Figure 23. Circles of Influence (20% Size) Used to Determine Zones of Influence for Each Category of Watershed.
4.3 Construction of the GIS Interface and Modeling Systems for DesktopL2W

Construction of DesktopL2W, the resultant watershed analysis software tool, involves four major development processes: (1) design of software architecture, (2) selection and preparation of a programming environment, (3) use of control objects and programming code to mold custom interface features, and (4) insertion of model/simulation capabilities.
4.3.1 Design of Software Architecture

The architecture of DesktopL2W follows the template of a coupled version of a GIS-based model (Figure 25). The GIS provides an interface for building scenarios, stores data about the scenarios and the landscape, communicates those scenarios to discipline-specific models, and helps compile and display model results. Each of the modeling functions is contained in a module that is set aside from, but communicates with the GIS warehouse.

![Figure 25. Architecture for DesktopL2W.](image)

4.3.2 Programming Environment

Three commercial software packages are used in development of DesktopL2W. These packages are Microsoft Visual Basic 6.0 (Professional Edition), ESRI ArcView 3.1, and ESRI MapObjects 2.0. Visual Basic (an object oriented programming language) is used in development of the user interface, transfer of data from user inputs to models, and aid in specifying and running simulations. ArcView is used in preparation of base data layers, which are utilized and presented in DesktopL2W. MapObjects (a control that is added to Visual Basic) provides map display functions, GIS functions, and database management.
4.3.3 Design of Custom Interface Features

The controls available in the programming environment are used to create forms that serve as the skeletal structure for the DesktopL2W interface. The controls are part of the Visual Basic and MapObjects packages, and provide objects such as maps, input boxes, tabular displays, and command buttons with which the user interacts. The forms also contain some of the program code that runs the software. The majority of the code, however, is contained in a set of six modules. The modules and their functions are described below:

- **BasicOperations** – Controls fundamental GIS functions on the map, such as zooming and panning, and conducts basic file operations like exiting the software.
- **AddDevelopmentsOperations** – Handles processes associated with setting up a scenario of development tracts within the watershed, including drawing them on the map, creating a shapefile and database of development tracts, editing a collection of development tracts, and recognizing a cluster of tracts as a single development feature.
- **GlobalVariables** – Declares variables that are used in multiple modules and forms.
- **HydroOperations** – Develops input variables and the pass file necessary to run an HSPF simulation, and populates a results form.
- **EconOperations** – Accesses database variables for the economics model, calculates certain variables (such as distance from tracts to sewer and water lines), populates a user input form, plugs the variables into the regression equation, and produces results in a Microsoft Excel spreadsheet.
- **FisheriesOperations** – Accesses variables in a database, runs a regression to estimate fisheries health, and provides a form with map-based visual results.

A schematic of the software design and a full copy of all the program code, which contains much plain-text commentary to describe the purposes of various blocks of code, is available in Appendix B.

4.3.4 Development of Model/Simulation Capabilities

The purpose of designing a custom GIS interface is to provide the ability to analyze a watershed by allowing a user to access advanced modeling tools. Incorporation of modeling/simulation capabilities into the interface, therefore, is a fundamental step in creating a tool for integrated analysis. DesktopL2W currently incorporates three models, surface hydrology, economics, and fisheries. Two categories of abilities are necessary to set up or run these models: provision of data related to spatial and physical characteristics of the landscape, and communication of input data that the user specifies to the appropriate model. The three-point strategy to use these abilities is (1) store baseline data in the development grid layer, (2) calculate spatially oriented values using GIS functions on base data layers, and (3) store scenario data in the development tracts layer. The development grid layer stores data such as current land use and baseline parcel values that are necessary to set up model runs. The GIS calculates additional values as needed (when a particular model is chosen), and both of these sets of values are transferred to development tracts when the user inputs a scenario. The development tracts layer winds up storing a mass of data, including area, type of development tract, UTM coordinates, pre-
development land use, post-development land use, containment within a subwatershed/land segment, baseline parcel value, post-development parcel value, distance to sewer and water lines, and membership in a cluster of tracts. Each model accesses any data required from this layer. More detailed information on the preparation of data for these layers is provided in a sample from the economic land values model in Appendix A.
5.0 **RESULTS**

Three major outcomes of this study are: (1) construction of a data library for interdisciplinary watershed analysis; (2) collection of a set of procedures for accomplishing GIS-based integration; and (3) development of DesktopL2W software that integrates various modeling procedures.

5.1 **Data Library**

Any working GIS, especially one aimed at integrating disparate models, requires a collection of data layers. These layers form a library that is a valuable and repeatedly useable tool (in fact, the layers have already been shared with several other GIS-oriented research projects). It is important to catalog the layers in the library and specify metadata, or data about the layers. The data library for this research effort can be classified into two major categories: general data layers for the Upper Roanoke River Watershed (Table 9-Table 15), and layers created explicitly for DesktopL2W (Table 16).

While the data library itself is important, the rationales for selecting members of that library are equally important. In selecting base data, there are several key issues to consider, such as data availability, needs of the models to be integrated, spatial resolution, and computing capability and storage space.

As unfortunate as it may seem, data availability is a central issue in compiling the data library. Development of the GIS and working software is a fairly intensive effort even after base data sets are available. Efforts intensify greatly in relation to the complexity of the discipline-specific models, so as models become more elaborate, a GIS analyst must spend more time revamping data and producing useful tools for the models and the integration process. The ability, therefore, to access readily available and cheap or free base data often determines what kinds of base data will be used for the project.

The needs of the models, naturally, also play a central role in selecting data sets. For example, the economic model requires population-related variables. U.S. Census TIGER data sets provide a way to connect population data to spatial entities, so they are a necessary addition to the data library.

Finally, spatial resolution, which is discussed in more detail in Chapter six, is an important factor in selecting data. Many data sets, such as USGS DEMs, are available in several resolutions. Which one is appropriate depends on the level of detail required for input for the models. In addition, the more detailed (higher resolution) a data set is, the more space it occupies on a computer storage system. Limits on space, therefore, may determine the appropriate resolution and data set. Similarly, as resolution increases, the computing time and power necessary to alter, enhance, or otherwise work with spatial data increases, so computer speed may also govern which data sets are added to a project library.
Table 9. Metadata for Elevation-Related Layers (Upper Roanoke River Watershed).

<table>
<thead>
<tr>
<th>Layer(s)/File Name(s)</th>
<th>Contents</th>
<th>File Type</th>
<th>Original Source</th>
<th>Intermediate Source(s)</th>
<th>Acquisition Method</th>
<th>Projection/Scale/Accuracy</th>
<th>Classification Scheme</th>
</tr>
</thead>
<tbody>
<tr>
<td>grid**** folder</td>
<td>DEM for each 7.5' quad</td>
<td>.DEM Raster</td>
<td>USGS web site</td>
<td>None</td>
<td>Downloaded from web, unzipped, renamed, converted with SDTS, imported to ArcView</td>
<td>UTM zone 17 (1927), USGS 1:24,000 scale topographic quadrangle map series, 30 m resolution.</td>
<td>Elevation values in meters (quads given in ft. were transformed with Map Calculator)</td>
</tr>
<tr>
<td>Mosaic folder</td>
<td>Watershed elevation model, mosaic of 7.5' quads</td>
<td>.DEM Raster</td>
<td>USGS web site</td>
<td>None</td>
<td>Individual DEMs mosaiced with Spatial Tools (Mosaic and Boundary Clean commands). Boundary Clean parameters: #Grid_sorttype_nosort, true.</td>
<td>Same as grid****</td>
<td>Elevation values in meters</td>
</tr>
<tr>
<td>Hillshade folder</td>
<td>Watershed hillshade model, mosaic of 7.5' quads</td>
<td>Arc Grid</td>
<td>USGS web site</td>
<td>DEM mosaic</td>
<td>Calculated from mosaic using Spatial Analyst (Compute Hillshade command)</td>
<td>UTM zone 17 (1927), USGS 1:24,000 scale topographic quadrangle map series, 30 m resolution.</td>
<td>Hillshade values in degrees</td>
</tr>
<tr>
<td>Tin folder</td>
<td>Triangulated irregular network (surface model) of the mosaic of 7.5' quads</td>
<td>Arc TIN</td>
<td>USGS web site</td>
<td>DEM mosaic</td>
<td>Calculated from mosaic using Spatial Analyst (Convert Grid to TIN command)</td>
<td>UTM zone 17 (1927), USGS 1:24,000 scale topographic quadrangle map series, z-value tolerance = 50.2.</td>
<td>Elevation values in meters</td>
</tr>
<tr>
<td>Slopedeg folder</td>
<td>Slope in degrees for the watershed area</td>
<td>Arc Grid</td>
<td>USGS DEMs</td>
<td>DEM mosaic</td>
<td>Calculated from mosaic using Spatial Analyst (Derive Slope command)</td>
<td>UTM zone 17 (1927), USGS 1:24,000 scale topographic quadrangle map series, 30 m resolution.</td>
<td>Slope value in degrees</td>
</tr>
<tr>
<td>Slopepect folder</td>
<td>Slope in percentage for the watershed area</td>
<td>Arc Grid</td>
<td>USGS DEMs</td>
<td>Slopedeg</td>
<td>Calculated from slopedeg by using the Map Calculator to convert degrees to percentage</td>
<td>UTM zone 17 (1927), USGS 1:24,000 scale topographic quadrangle map series, 30 m resolution.</td>
<td>Slope value in percentage</td>
</tr>
<tr>
<td>Hypsography.shp</td>
<td>DLG topographic lines for the watershed</td>
<td>ArcView Vector</td>
<td>USGS web site</td>
<td>None</td>
<td>Downloaded from web, unzipped, converted with DLGLX149 to CAD files, imported to AV, merged, converted to shapefile, and node-connected</td>
<td>UTM zone 17 (1927), USGS 1:100,000 scale topographic quadrangle map series, originally each 30'-by-60' quad is divided into 15' tiles.</td>
<td>None</td>
</tr>
</tbody>
</table>
Table 10. Metadata for Stream Layers (Upper Roanoke River Watershed).

<table>
<thead>
<tr>
<th>Layer(s)/File Name(s)</th>
<th>Contents</th>
<th>File Type</th>
<th>Original Source</th>
<th>Intermediate Source(s)</th>
<th>Acquisition Method</th>
<th>Projection/Scale/Accuracy</th>
<th>Classification Scheme</th>
</tr>
</thead>
<tbody>
<tr>
<td>streams.shp</td>
<td>DLG stream network for the watershed</td>
<td>ArcView Vector</td>
<td>USGS web site</td>
<td>None</td>
<td>Downloaded from web, unzipped, converted with DLGLX149 to CAD files, imported to AV, merged, converted to shapefile, and node-connected. Some additional reaches added based on EPA RF3 data.</td>
<td>UTM zone 17 (1927), USGS 1:100,000 scale topographic quadrangle map series, originally each 30'-by-60' quad is divided into 15' tiles.</td>
<td>None</td>
</tr>
<tr>
<td>va_river</td>
<td>VA rivers, streams, and lakes</td>
<td>ArcView Vector</td>
<td>USGS DLGs</td>
<td>VT FWIE</td>
<td>Provided on disk by FWIE</td>
<td>UTM zones (1927)</td>
<td>None</td>
</tr>
<tr>
<td>Tgr51***lkH.shp</td>
<td>Hydrology for each county and city in the watershed</td>
<td>ArcView Vector</td>
<td>U.S. Census TIGER 95 Files</td>
<td>None</td>
<td>Downloaded from TIGER CD-ROM, converted to ArcView with Tiger2shape software (Geography Lab)</td>
<td>Unprojected (decimal degrees), based on USGS 1:100,000 scale topographic quadrangle map series.</td>
<td>None</td>
</tr>
<tr>
<td>03010101.shp</td>
<td>Hydrology for the entire Roanoke Riv. Watershed</td>
<td>ArcView Vector</td>
<td>USEPA Basins web site (RF3 data)</td>
<td>None</td>
<td>Downloaded from the web</td>
<td>UTM zone 17 (1927), USGS 1:100,000 scale topographic quadrangle map series</td>
<td>None</td>
</tr>
<tr>
<td>mergestreams.shp</td>
<td>Merged streams for each county in the watershed</td>
<td>ArcView Vector</td>
<td>USGS DLG data; major rivers and water bodies conflated to SPOT.</td>
<td>VA Dept. of Trans., County Map Series CD Enhanced Package</td>
<td>Desired county data extracted from CD-ROM, stream coverages from each county merged into one theme</td>
<td>Lambert Conformal Conic, NAD 83, Cent. Merid. -79.5, Lat. Of Orig. 36.0, 1st Std. Par. 37.0, 2nd Std. Par. 39.5, Conflated to 1:24,000 SPOT Imagery</td>
<td>Legend of stream types provided by VDOT</td>
</tr>
<tr>
<td>Mergewater.shp</td>
<td>Merged bodies of open water for each county in the watershed</td>
<td>ArcView Vector</td>
<td>USGS DLG data; major rivers and water bodies conflated to SPOT.</td>
<td>VA Dept. of Trans., County Map Series CD Enhanced Package</td>
<td>Desired county data extracted from CD-ROM, stream coverages from each county merged into one theme</td>
<td>Lambert Conformal Conic, NAD 83, Cent. Merid. -79.5, Lat. Of Orig. 36.0, 1st Std. Par. 37.0, 2nd Std. Par. 39.5, Conflated to 1:24,000 SPOT Imagery</td>
<td>Legend water body types provided by VDOT</td>
</tr>
</tbody>
</table>
Table 11. Metadata for Transportation Layers (Upper Roanoke River Watershed).

<table>
<thead>
<tr>
<th>Layer(s)/File Name(s)</th>
<th>Contents</th>
<th>File Type</th>
<th>Original Source</th>
<th>Intermediate Source(s)</th>
<th>Acquisition Method</th>
<th>Projection/Scale/Accuracy</th>
<th>Classification Scheme</th>
</tr>
</thead>
<tbody>
<tr>
<td>roads.shp</td>
<td>Detailed DLG road network for the watershed</td>
<td>ArcView Vector</td>
<td>USGS web site</td>
<td>None</td>
<td>Downloaded from web, unzipped, converted with DLGLX149 to CAD files, imported to AV, merged, converted to shapefile, and node-connected</td>
<td>UTM zone 17 (1927), USGS 1:100,000 scale topographic quadrangle map series, originally each 30'-by-60' quad is divided into 15' tiles.</td>
<td>Road size/type; e.g., class 4 road or street</td>
</tr>
<tr>
<td>va_roads</td>
<td>Major roads for the state of Virginia</td>
<td>ArcView Vector</td>
<td>USGS DLGs</td>
<td>VT FWIE</td>
<td>Provided on disk by FWIE</td>
<td>UTM zones (1927)</td>
<td>None</td>
</tr>
<tr>
<td>Tgr51***lkA.shp</td>
<td>Detailed road network for each county and city in the watershed</td>
<td>ArcView Vector</td>
<td>U.S. Census TIGER 95 Files</td>
<td>None</td>
<td>Downloaded from TIGER CD-ROM, converted to ArcView with Tiger2shape software (Geography Lab)</td>
<td>Unprojected (decimal degrees), based on USGS 1:100,000 scale topographic quadrangle map series.</td>
<td>None</td>
</tr>
<tr>
<td>mergeroads.shp</td>
<td>Merged road network for each county in the watershed</td>
<td>ArcView Vector</td>
<td>TIGER data (1:100,000); Conflated to SPOT Imagery</td>
<td>VA Dept. of Trans., County Map Series CD Enhanced Package</td>
<td>Desired county data extracted from CD-ROM, road coverages from each county merged into one theme</td>
<td>Lambert Conformal Conic, NAD 83, Cent. Merid. -79.5, Lat. Of Orig. 36.0, 1st Std. Par. 37.0, 2nd Std. Par. 39.5, Conflated to 1:24,000 SPOT Imagery</td>
<td>Legend provided by VDOT</td>
</tr>
<tr>
<td>railroads.shp</td>
<td>Railroad network for the watershed area</td>
<td>ArcView Vector, DLG</td>
<td>USGS web site</td>
<td>None</td>
<td>Downloaded from web, unzipped, converted with DLGLX149 to CAD files, imported to AV, merged, converted to shapefile, and node-connected</td>
<td>UTM zone 17 (1927), USGS 1:100,000 scale topographic quadrangle map series, originally each 30'-by-60' quad is divided into 15' tiles.</td>
<td>Rail line type</td>
</tr>
<tr>
<td>Tgr51***lkB.shp</td>
<td>Railroads for each county and city in the watershed</td>
<td>ArcView Vector</td>
<td>U.S. Census TIGER 95 Files</td>
<td>None</td>
<td>Downloaded from TIGER CD-ROM, converted to ArcView with Tiger2shape software (Geography Lab)</td>
<td>Unprojected (decimal degrees), based on USGS 1:100,000 scale topographic quadrangle map series.</td>
<td>None</td>
</tr>
<tr>
<td>mergeutilities.shp</td>
<td>Railroads, power lines, and bridges for each county in the watershed</td>
<td>ArcView Vector</td>
<td>TIGER data (1:100,000); Conflated to SPOT Imagery</td>
<td>VA Dept. of Trans., County Map Series CD Enhanced Package</td>
<td>Desired county data extracted from CD-ROM, utility coverages from each county merged into one theme</td>
<td>Lambert Conformal Conic, NAD 83, Cent. Merid. -79.5, Lat. Of Orig. 36.0, 1st Std. Par. 37.0, 2nd Std. Par. 39.5, Conflated to 1:24,000 SPOT Imagery</td>
<td>Legend provided by VDOT</td>
</tr>
</tbody>
</table>
Table 11, Continued.

<table>
<thead>
<tr>
<th>Layer(s)/ File Name(s)</th>
<th>Contents</th>
<th>File Type</th>
<th>Original Source</th>
<th>Intermediate Source(s)</th>
<th>Acquisition Method</th>
<th>Projection/Scale/ Accuracy</th>
<th>Classification Scheme</th>
</tr>
</thead>
<tbody>
<tr>
<td>misctrans.shp</td>
<td>DLG network of pipelines, power lines, etc. for the watershed area</td>
<td>ArcView Vector, DLG</td>
<td>USGS web site</td>
<td>None</td>
<td>Downloaded from web, unzipped, converted with DLGLX149 to CAD files, Imported to AV, merged, converted to shapefile, and node-connected</td>
<td>UTM zone 17 (1927), USGS 1:100,000 scale topographic quadrangle map series, originally each 30'-by-60' quad is divided into 15' tiles.</td>
<td>Type of transportation line</td>
</tr>
<tr>
<td>mergebridges.shp</td>
<td>Roadway bridges for each county in the watershed</td>
<td>ArcView Vector</td>
<td>TIGER data (1:100,000); Conflated to SPOT Imagery</td>
<td>VA Dept. of Trans., County Map Series CD Enhanced Package</td>
<td>Desired county data extracted from CD-ROM, bridge coverages from each county merged into one theme</td>
<td>Lambert Conformal Conic, NAD 83, Cent. Merid. -79.5, Lat. Of Orig. 36.0, 1st Std. Par. 37.0, 2nd Std. Par. 39.5, Conflated to 1:24,000 SPOT Imagery</td>
<td>Legend provided by VDOT</td>
</tr>
</tbody>
</table>

Table 12. Metadata for Boundaries (Political and Watershed -- Upper Roanoke River Watershed).

<table>
<thead>
<tr>
<th>Layer(s)/ File Name(s)</th>
<th>Contents</th>
<th>File Type</th>
<th>Original Source</th>
<th>Intermediate Source(s)</th>
<th>Acquisition Method</th>
<th>Projection/Scale/ Accuracy</th>
<th>Classification Scheme</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uproanws.shp</td>
<td>Shapefile of the Upper Roanoke River Watershed</td>
<td>ArcView Vector</td>
<td>Calculated from DEM Mosaic burned with DLG streams</td>
<td>None</td>
<td>TNRCC Watershed Delineator, default settings</td>
<td>Based on UTM zone 17 (1927), USGS 1:24,000 scale topographic quadrangle map series, 30 m resolution.</td>
<td>None</td>
</tr>
<tr>
<td>Fishshed.shp</td>
<td>Subwatersheds that are upstream of fish sampling sites</td>
<td>ArcView Vector</td>
<td>Calculated from DEM Mosaic burned with DLG streams</td>
<td>None</td>
<td>TNRCC Watershed Delineator, default settings</td>
<td>Based on UTM zone 17 (1927), USGS 1:24,000 scale topographic quadrangle map series, 30 m resolution.</td>
<td>None</td>
</tr>
<tr>
<td>Lgefishshed.shp</td>
<td>Large size subwatersheds that are upstream of fish sampling sites</td>
<td>ArcView Vector</td>
<td>Parsed from Fishshed.shp</td>
<td>None</td>
<td>None</td>
<td>Based on UTM zone 17 (1927), USGS 1:24,000 scale topographic quadrangle map series, 30 m resolution.</td>
<td>None</td>
</tr>
<tr>
<td>Medfishshed.shp</td>
<td>Medium size subwatersheds that are upstream of fish sampling sites</td>
<td>ArcView Vector</td>
<td>Parsed from Fishshed.shp</td>
<td>None</td>
<td>None</td>
<td>Based on UTM zone 17 (1927), USGS 1:24,000 scale topographic quadrangle map series, 30 m resolution.</td>
<td>None</td>
</tr>
<tr>
<td>Layer(s)/File Name(s)</td>
<td>Contents</td>
<td>File Type</td>
<td>Original Source</td>
<td>Intermediate Source(s)</td>
<td>Acquisition Method</td>
<td>Projection/Scale/Accuracy</td>
<td>Classification Scheme</td>
</tr>
<tr>
<td>----------------------</td>
<td>----------</td>
<td>-----------</td>
<td>-----------------</td>
<td>------------------------</td>
<td>--------------------</td>
<td>-------------------------</td>
<td>-----------------------</td>
</tr>
<tr>
<td>Smfishshed.shp</td>
<td>Small size subwatersheds that are upstream of fish sampling sites</td>
<td>ArcView Vector</td>
<td>Parsed from Fishshed.shp</td>
<td>None</td>
<td>None</td>
<td>Based on UTM zone 17 (1927), USGS 1:24,000 scale topographic quadrangle map series, 30 m resolution.</td>
<td>None</td>
</tr>
<tr>
<td>Streambuffer**m.shp</td>
<td>Buffers of the streams at distances ranging from 30 to 150 meters</td>
<td>ArcView Vector</td>
<td>Calculated from USGS DLG streams.shp</td>
<td>None</td>
<td>ArcView Analysis – Create buffers</td>
<td>Based on UTM zone 17 (NAD 1927), 1:100,000 scales.</td>
<td>None</td>
</tr>
<tr>
<td>Circle***.shp</td>
<td>Circle of influence around fish sampling sites, from 5 to 60 pct. of watershed area</td>
<td>ArcView Vector</td>
<td>None – original theme created in ArcView</td>
<td>None</td>
<td>Buffered sample points at appropriate distances and repositioned circles.</td>
<td>UTM zone 17 (NAD 1927)</td>
<td>None</td>
</tr>
<tr>
<td>InfluenceZone***.shp</td>
<td>Zones of influence around fish sampling sites for land use analysis</td>
<td>ArcView Vector</td>
<td>Streambuffer<strong>m, Fishshed, and Circle</strong>* shapefiles</td>
<td>None</td>
<td>Created buffers around streams and clipped them by watersheds and circles of influence.</td>
<td>UTM zone 17 (NAD 1927)</td>
<td>None</td>
</tr>
<tr>
<td>va_cnty</td>
<td>VA county boundaries</td>
<td>ArcView Vector, DLG</td>
<td>USGS DLGs</td>
<td>VT FWIE</td>
<td>Provided on disk by FWIE</td>
<td>UTM zones (NAD 1927)</td>
<td>None</td>
</tr>
<tr>
<td>cntyclip.shp</td>
<td>Boundaries for the counties and cities in the watershed area</td>
<td>ArcView Vector, DLG</td>
<td>USGS DLGs</td>
<td>VT FWIE</td>
<td>Extracted from va_cnty after conversion to shapefile</td>
<td>UTM zone 17 (NAD 1927)</td>
<td>None</td>
</tr>
<tr>
<td>Tgr51***cty.shp</td>
<td>Borders for each county or city in the watershed</td>
<td>ArcView Vector</td>
<td>U.S. Census TIGER 95 Files</td>
<td>None</td>
<td></td>
<td>Unprojected (decimal degrees), based on USGS 1:100,000 scale topographic quadrangle map series.</td>
<td>None</td>
</tr>
<tr>
<td>Tgr51***grp.shp</td>
<td>Census tracts for each county or city in the watershed</td>
<td>ArcView Vector</td>
<td>U.S. Census TIGER 95 Files</td>
<td>None</td>
<td></td>
<td>Unprojected (decimal degrees), based on USGS 1:100,000 scale topographic quadrangle map series.</td>
<td>None</td>
</tr>
<tr>
<td>Tgr51***blk.shp</td>
<td>Block groups for each county or city in the watershed</td>
<td>ArcView Vector</td>
<td>U.S. Census TIGER 95 Files</td>
<td>None</td>
<td></td>
<td>Unprojected (decimal degrees), based on USGS 1:100,000 scale topographic quadrangle map series.</td>
<td>None</td>
</tr>
</tbody>
</table>
### Table 12, Continued.

<table>
<thead>
<tr>
<th>Layer(s)/File Name(s)</th>
<th>Contents</th>
<th>File Type</th>
<th>Original Source</th>
<th>Intermediate Source(s)</th>
<th>Acquisition Method</th>
<th>Projection/Scale/Accuracy</th>
<th>Classification Scheme</th>
</tr>
</thead>
<tbody>
<tr>
<td>data/mergeboundary.shp</td>
<td>County, city, and more detailed level boundaries for each county in the watershed</td>
<td>ArcView Vector</td>
<td>TIGER data (1:100,000) and DLG public lands; water boundaries conflated to SPOT Imagery; municipal, forest and park boundaries based on County Map Mylars.</td>
<td>VA Dept. of Trans., County Map Series CD Enhanced Package</td>
<td>Desired county data extracted from CD-ROM, coverages from each county merged into one theme</td>
<td>Lambert Conformal Conic, NAD 83, Cent. Merid. -79.5, Lat. Of Orig. 36.0, 1st Std. Par. 37.0, 2nd Std. Par. 39.5, Conflated to 1:24,000 SPOT Imagery</td>
<td>Legend provided by VDOT</td>
</tr>
<tr>
<td>data/boundaries.shp</td>
<td>County, city, and more detailed level boundaries for the watershed area</td>
<td>ArcView Vector, DLG</td>
<td>USGS web site</td>
<td>None</td>
<td>Downloaded from web, unzipped, converted with DLGLX149 to CAD files, imported to AV</td>
<td>UTM zone 17 (1927), USGS 1:100,000 scale topographic quadrangle map series, each 30'-by-60' quad is divided into 15' tiles.</td>
<td>None</td>
</tr>
<tr>
<td>data/Q3***</td>
<td>Q3 data, such as 100-year flood plains for cities and counties in VA</td>
<td>ArcView Vector</td>
<td>Scanned versions of FEMA's paper flood insurance maps</td>
<td>Fifth Planning District</td>
<td>Desired county data extracted from CD-ROM and imported to ArcView</td>
<td>Unprojected (decimal degrees), 1:24,000 scale</td>
<td>Can be classified by many flood-related attributes</td>
</tr>
</tbody>
</table>

### Table 13. Metadata for Land Use and Parcels (Upper Roanoke River Watershed).

<table>
<thead>
<tr>
<th>Layer(s)/File Name(s)</th>
<th>Contents</th>
<th>File Type</th>
<th>Original Source</th>
<th>Intermediate Source(s)</th>
<th>Acquisition Method</th>
<th>Projection/Scale/Accuracy</th>
<th>Classification Scheme</th>
</tr>
</thead>
<tbody>
<tr>
<td>data/Vagap2</td>
<td>Specific land use for the area of the Upper Roanoke River Watershed</td>
<td>Arc Grid Raster</td>
<td>VT FWIE Virginia GAP Analysis – LANDSAT Thematic Mapper (Interpretation)</td>
<td>None</td>
<td>Provided on disk, along with legend and classification files</td>
<td>UTM zone 17 (NAD27), 30x30 meter resolution, est. 80% accuracy</td>
<td>28 classes with specific breakdowns within broad categories like forest and disturbed</td>
</tr>
<tr>
<td>data/Reclass</td>
<td>Broad land use for the area of the Upper Roanoke River Watershed</td>
<td>Arc Grid Raster</td>
<td>VT FWIE Virginia GAP Analysis – LANDSAT Thematic Mapper (Interpretation)</td>
<td>None</td>
<td>Reclassified version of Vagap2</td>
<td>UTM zone 17 (NAD27), 30x30 meter resolution, more accurate than Vagap2</td>
<td>10 = forest, 30 = herb/ag, 40 = open water, 50 = disturbed, and 99 = mixed/forest edge</td>
</tr>
</tbody>
</table>
Table 13, Continued.

<table>
<thead>
<tr>
<th>Layer(s)/File Name(s)</th>
<th>Contents</th>
<th>File Type</th>
<th>Original Source</th>
<th>Intermediate Source(s)</th>
<th>Acquisition Method</th>
<th>Projection/Scale/Accuracy</th>
<th>Classification Scheme</th>
</tr>
</thead>
<tbody>
<tr>
<td>utmparcels.shp</td>
<td>Preliminary Version of Roanoke County Tax Parcels</td>
<td>ArcView</td>
<td>Roanoke County cadastral maps digitized in AutoCAD</td>
<td>Fifth Planning District – files converted to ArcView</td>
<td>Provided on disk in 6 files – State Plane NAD27 (VA South, feet), converted to UTM</td>
<td>UTM zone 17 (NAD27)</td>
<td>None</td>
</tr>
<tr>
<td>Preservedland.shp</td>
<td>National forests and other preserved lands in the Upper Roanoke Area</td>
<td>ArcView</td>
<td>TIGER data (1:100,000); Conflated to SPOT Imagery</td>
<td>VA Dept. of Trans., County Map Series CD Enhanced Package</td>
<td>Desired county data extracted from CD-ROM, preserved land coverages from each county merged into one theme and reprojected</td>
<td>UTM zone 17 (NAD27)</td>
<td>None</td>
</tr>
<tr>
<td>Developable.shp</td>
<td>Regions within watersheds that are open to additional development</td>
<td>Arc Grid</td>
<td>Combined USGS DEM, FEMA Q3, PreservedLand.shp, VAGAP land use, and WS boundaries</td>
<td>None</td>
<td>Created 0 or 1 raster model for each filter layer and summed to obtain developable cells.</td>
<td>UTM zone 17 (NAD27), 30x30 meter resolution</td>
<td>0 = can't be developed 1 = can be developed</td>
</tr>
</tbody>
</table>

Table 14. Metadata for Soils (Upper Roanoke River Watershed).

<table>
<thead>
<tr>
<th>Layer(s)/File Name(s)</th>
<th>Contents</th>
<th>File Type</th>
<th>Original Source</th>
<th>Intermediate Source(s)</th>
<th>Acquisition Method</th>
<th>Projection/Scale/Accuracy</th>
<th>Classification Scheme</th>
</tr>
</thead>
<tbody>
<tr>
<td>utmsoil.shp</td>
<td>STATSGO soil polygons for the state of Virginia</td>
<td>ArcView</td>
<td>USDA NRCS Web Site</td>
<td>None</td>
<td>Downloaded in ArcView format, reprojected, and saved as a shapefile</td>
<td>Originally Albers conic equal area, converted to UTM zone 17 (NAD 27), 1:250,000</td>
<td>Any variable attached to map unit ID</td>
</tr>
<tr>
<td>Soilordfolder</td>
<td>Taxonomic soil order for the state of Virginia</td>
<td>Arc Grid</td>
<td>USDA STATSGO</td>
<td>None</td>
<td>Attached soil order data to the map unit ID</td>
<td>UTM zone 17 (NAD 1927)</td>
<td>Unique value for each order, e.g., 1 = alfisols</td>
</tr>
<tr>
<td>Va161_afolder</td>
<td>SSURGO soil polygons for Roanoke County</td>
<td>ArcView</td>
<td>USDA NRCS Web Site</td>
<td>None</td>
<td>Downloaded in ArcView format, unprojected -- decimal degrees</td>
<td>Unprojected -- decimal degrees</td>
<td>Any variable attached to map unit symbol</td>
</tr>
</tbody>
</table>

52
Table 15. Metadata for Point Locations (Upper Roanoke River Watershed).

<table>
<thead>
<tr>
<th>Layer(s)/File Name(s)</th>
<th>Contents</th>
<th>File Type</th>
<th>Original Source</th>
<th>Intermediate Source(s)</th>
<th>Acquisition Method</th>
<th>Projection/Scale/Accuracy</th>
<th>Classification Scheme</th>
</tr>
</thead>
<tbody>
<tr>
<td>Samplepnts.shp</td>
<td>Fish sampling sites for the watershed</td>
<td>ArcView</td>
<td>Fisheries modelers, from topo map estimates and GPS coordinates</td>
<td>None</td>
<td>Event them from Excel lat/long table</td>
<td>UTM Zone 17 (NAD27)</td>
<td>None</td>
</tr>
<tr>
<td>Samplepntslg.shp</td>
<td>Fish sampling sites for large subwatersheds</td>
<td>ArcView</td>
<td>Fisheries modelers, from topo map estimates and GPS coordinates</td>
<td>None</td>
<td>Parsed from Samplepnts.shp</td>
<td>UTM Zone 17 (NAD27)</td>
<td>None</td>
</tr>
<tr>
<td>Samplepntsmd.shp</td>
<td>Fish sampling sites for medium subwatersheds</td>
<td>ArcView</td>
<td>Fisheries modelers, from topo map estimates and GPS coordinates</td>
<td>None</td>
<td>Parsed from Samplepnts.shp</td>
<td>UTM Zone 17 (NAD27)</td>
<td>None</td>
</tr>
<tr>
<td>Samplepntssm.shp</td>
<td>Fish sampling sites for small subwatersheds</td>
<td>ArcView</td>
<td>Fisheries modelers, from topo map estimates and GPS coordinates</td>
<td>None</td>
<td>Parsed from Samplepnts.shp</td>
<td>UTM Zone 17 (NAD27)</td>
<td>None</td>
</tr>
<tr>
<td>UtmGage.shp</td>
<td>Stream gage sites for the watershed</td>
<td>ArcView</td>
<td>USGS web site with stream gage site information</td>
<td>Hydrology modelers</td>
<td>Event them from Excel lat/long table</td>
<td>UTM Zone 17 (NAD27)</td>
<td>None</td>
</tr>
</tbody>
</table>
### Table 16. Metadata for DesktopL2W Layers.

<table>
<thead>
<tr>
<th>Layer(s)/File Name(s)</th>
<th>Contents</th>
<th>Source</th>
<th>Acquisition Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>streams.shp</td>
<td>Streams for the Back Creek Subwatershed</td>
<td>USGS DLG streams</td>
<td>Clipping from larger DLG streams file using the watershed boundary</td>
</tr>
<tr>
<td>subwatersheds.shp</td>
<td>Back Creek Subwatershed broken into 10 subwatersheds/land segments which serve as units of analysis for the hydrology and fisheries models</td>
<td>USGS DEMs burned with DLGs</td>
<td>TNRCC Watershed Delineator with spatial subtraction.</td>
</tr>
<tr>
<td>grid_pattern.shp</td>
<td>9-hectare grid cells that cover the Back Creek Subwatershed, complete with physical and economic attribute data</td>
<td>Original AutoCAD drawing</td>
<td></td>
</tr>
<tr>
<td>aacombined.shp</td>
<td>User-specified development tracts (all types combined) with modeling data attached</td>
<td>User input and specifications in DesktopL2W</td>
<td>Delineation of square areas to match the development grid</td>
</tr>
<tr>
<td>land_use.shp</td>
<td>Major category land uses for the Back Creek Subwatershed</td>
<td>VT FWIE Virginia GAP Analysis -- LANDSAT TM</td>
<td></td>
</tr>
<tr>
<td>ap_roan_ws.shp</td>
<td>Boundary of the Upper Roanoke River Watershed</td>
<td>USGS DEMs burned with DLGs</td>
<td>TNRCC Watershed Delineator</td>
</tr>
<tr>
<td>tinker_ws.shp</td>
<td>Boundary of the Tinker Creek Subwatershed</td>
<td>USGS DEMs burned with DLGs</td>
<td>TNRCC Watershed Delineator</td>
</tr>
<tr>
<td>rain_gages.shp</td>
<td>Rain gage sites for the area</td>
<td>Coordinates from USGS web site</td>
<td>Conversion of coordinates to an event theme and re-conversion to a shapefile</td>
</tr>
<tr>
<td>thiessen.shp</td>
<td>Thiessen polygons (areas of precipitation influence) based on the rain gages</td>
<td>Rain_gages.shp</td>
<td>Use of ArcView extension &quot;Create Thiessen Polygons&quot;</td>
</tr>
<tr>
<td>reaches.shp</td>
<td>Main stem streams in Back Creek Subwatershed</td>
<td>Streams.shp (USGS DLG streams)</td>
<td>Removal of minor streams from streams.shp</td>
</tr>
<tr>
<td>outlet_points.shp</td>
<td>Outlets points for streams leaving each of the ten land segments in subwatersheds.shp</td>
<td>Original ArcView theme</td>
<td>Point digitization at downstream intersection of reaches and subwatershed boundaries</td>
</tr>
<tr>
<td>stream_gages.shp</td>
<td>Stream gage sites for the area</td>
<td>Coordinates from USGS web site</td>
<td>Conversion of coordinates to an event theme and re-conversion to a shapefile</td>
</tr>
<tr>
<td>roads.shp</td>
<td>Roads for the Back Creek Subwatershed</td>
<td>USGS DLG roads</td>
<td>Clipping from larger DLG roads file using the watershed boundary</td>
</tr>
<tr>
<td>Sewertest.shp</td>
<td>Simulated main stem sewer lines for the Back Creek area</td>
<td>None – original ArcView theme</td>
<td>Preliminary (testing purposes only) line digitization</td>
</tr>
<tr>
<td>Watertest.shp</td>
<td>Simulated main stem water lines for the Back Creek area</td>
<td>None – original ArcView theme</td>
<td>Preliminary (testing purposes only) line digitization that follows sewer lines</td>
</tr>
<tr>
<td>fish_sites.shp</td>
<td>Fish sampling sites throughout the Upper Roanoke River Watershed</td>
<td>Topo map estimates and GPS coordinates</td>
<td>Conversion of coordinates to an event theme and re-conversion to a shapefile</td>
</tr>
<tr>
<td>fish_outlets.shp</td>
<td>Polygonal versions of outlet_points.shp to display the status of fish health on a results map</td>
<td>outlet_points.shp</td>
<td>Buffering of outlet points</td>
</tr>
</tbody>
</table>

Notes: All layers in this table are ArcView vector files, which are projected in UTM zone 17 (NAD 1927).
5.2 Set of Procedures for GIS Integration

The approach taken to build integrated watershed analysis capabilities is dependent on the types of models that are being included in the system, but examples from this project may be used for a variety of models and provide a generalized template for what can be done. The following steps outline a broad methodology for how to integrate models into a GIS:

- Initiate a discussion with the modeling team to establish the purpose of the integrated system, the general capabilities of involved models, and potential levels of interaction among models.

- Select a GIS software package for collecting, manipulating, and displaying data. The selection should be made with an eye toward compatibility, ease of use, availability of built-in analysis functions, and ability to use its spatial data in the final software tool.

- Estimate approximate watershed size and assess potential model inputs to determine likely spatial resolution needs (fine vs. coarse).

- Assemble available base spatial data layers. These layers include both physical features (elevation, hydrography, soils, land use, geology) and socioeconomic features (city and county boundaries, ownership parcels, roads). This process involves downloading files from federal and state government sources, and contacting local groups like county offices, metropolitan planning organizations, college departments, libraries, and other holders of GIS data.

- Work with modelers of each discipline to (1) explain the base spatial data and what GIS can accomplish with these layers, and (2) ascertain what sorts of data are needed and available for modeling. If modelers are developing data sets, the GIS analyst should help fashion them so they can be seamlessly added to the GIS (e.g., the GIS analyst can help set up appropriate key fields for data sets that will be attached to a spatial layer).

- Collect, organize, and enhance both spatial and tabular data sets necessary for the modeling processes. This step involves close collaboration with modelers to understand needs. It may be required that the GIS analyst and modelers sit down together in front of the computer to realize the best use of GIS functions and produce the most effective data sets. Development of all three models in DesktopL2W required personal collaboration to create certain input data sets. In most cases, preparation of data sets involves a repetitious effort, so there can be a first-time collaborative effort, with the GIS analyst completing the rest.

- After provision of input data to modelers, develop an interface for the custom software. The interface should be as easy to use as possible for the intended audience, but also must provide the necessary data to serve as input to each of the models. Knowledge of the models and what they require is crucial for this step. The interface should be somewhat flexible or universal in nature, so additional models can be added in the future.
As modelers finish their specification processes, add their models to the GIS-based software. In adding a model, the first decision is whether the model is an external model (like the surface hydrology model) or an internal model (such as the economics and fisheries models). If it is external, meaning that its equations and simulations are run by a stand-alone program outside of the GIS-based software, the GIS analyst must set up necessary routines to create a pass file with necessary input from the GIS, invoke the model program, run a simulation, and return results for the user to view (Figure 26). If it is internal, the analyst must build all of the input and output screens directly into the software. This step requires building menus, forms, and other features into the software to allow the user access to the models.

- Finalize and debug the custom GIS-based software tool.

---

**Figure 26. Example of GIS Communication with the External HSPF Model.**

5.3 Desktop L2W Software

DesktopL2W is a prototype, GIS-based software program that allows a user to input land use changes within a watershed, and run (or at least produce input data for) hydrologic, economic, and fisheries simulations. It is intended to serve as a planning tool to help watershed managers and policy makers craft sound decisions about development within the watershed. The intended audience can use DesktopL2W without advanced knowledge of these disciplines and their
respective modeling techniques. The main purpose of the software is to demonstrate how GIS can integrate modeling efforts from diverse fields.

5.3.1 Outline of DesktopL2W Functionality

The DesktopL2W interface allows the user to create informed development scenarios and understand the results of model simulations. The point is to let the user take full advantage of the expertise contained in the discipline-specific models. The interface accomplishes this task by providing advice on which areas to develop, permitting new and unique settlement patterns on the landscape, and supplying opportunities to compare simulation results to current settings (Figure 27). In addition, the user has access to help files and status boxes that explain each facet of the software.

![Figure 27. Outline of DesktopL2W Functionality.](image)

5.3.2 Land Use Scenario Input

Upon start up, DesktopL2W displays the scenario input screen (Figure 28). The screen consists of an interactive map and legend. The legend shows which data sets are currently depicted in the map, and allows the user to hide layers and determine the order in which they are drawn. Three layers (streams, Back Creek Watershed with ten subwatersheds/land segments, and the development grid) automatically load into the map.

The user drops various types of development tracts on the landscape by clicking the button of the desired tract type and then clicking one of the grid squares on the map. The color of a grid square suggests whether it is suitable for holding a new tract. When the user clicks a tract onto
the map, the tract is displayed in a color that corresponds to its type. If the user makes a mistake in placing tracts on the map, they can be removed either individually or all at once.

Development tracts can be selected in different combinations and locations to create a land use change scenario for the watershed. For example, a development scenario strategy could be to add different types of development tracts to a subwatershed until certain hydrologic or biological thresholds are exceeded. Such a strategy permits the project to investigate how many tracts of each type could be placed in a subwatershed without exceeding critical environmental or economic thresholds.

Figure 28. DesktopL2W Scenario Input Screen.
In addition to dropping development tracts, the user can interact with the map in other ways. The input screen offers standard GIS display functions like zooming, panning, and map printing. It also allows the user to add more data sets to the map (Figure 29) to aid in deciding where to place tracts. For example, the user can view layers like roads or current land use to make realistic decisions about where development might occur in the future (Figure 30).

![Add Data to the Map](image)

*Figure 29. DesktopL2W Screen for Adding Data Layers.*
Figure 30. DesktopL2W Input Screen with Land Use Layer Added.

A few functions on the input screen are available in the drop-down menus, but the toolbar buttons on the top of the map screen are the main way to access the functions. Functions of various buttons on the map screen are described below:

- **Full Extent**: Zooms the map to its full extent to show all of the spatial data layers
- **Pan**: Pans or drags the map when it is zoomed in closer than its full extent
- **Zoom In**: Zooms the map in on a rectangle as drawn by the user
- **Zoom Out**: Zooms the map out 2x with each click until the full extent is reached
Add Data  Provides the option of adding additional spatial data to the map
Print Map  Prints a copy of the map to the default printer
Low Dens  Drops a low density residential tract on the map
Mid Conv  Drops a mid density conventional tract on the map
Mid Clust  Drops a mid density cluster tract on the map
High Dens  Drops a high density tract on the map
Commerc  Drops a commercial tract on the map
Remove  Removes a tract when the user clicks on a tract that has been placed on the map
Clear All  Removes all of the tracts that have been placed on the map
Quit  Exits the program

5.3.3 Model Selection and Simulation Results

Upon entry of a “what if” scenario, the user clicks the “Accept Input and Run Model(s)” button on the main input screen to access the modeling applications. This selection signals DesktopL2W to display a model selection screen (Figure 31) that lets the user choose models to run. Depending on the choices made by the user, discipline-specific models run and provide results. Simulation results are all displayed in relation to baseline values so that the user is fully aware of changes that arise from the input scenario.
5.3.4 Sample Run and Results

A sample run of DesktopL2W software is described as follows. In the sample scenario, a number of commercial tracts are placed in the most northwestern subwatershed (land segment 1), and a variety of residential tracts are placed in the most southwestern subwatershed (land segment 2, Figure 32). After clicking the “Accept Input and Run Model(s)” button, all models are selected (Figure 31). The hydrology, economics, and fisheries modules all run various algorithms to produce results. The hydrology model outputs a form (along with pass files for HSPF) that displays the land use portions for each subwatershed (Figure 33). The economics model, after receiving further user input (Figure 34), produces a Microsoft Excel spreadsheet that shows changes in land values, tax revenues, and government expenditures (Figure 35). Finally, the fisheries model outputs a form that shows the relative quality of each outlet point for sustaining healthy fish populations (Figure 36).
Figure 32. Input Scenario for the Sample Run.
The surface hydrology output form shows that the percentages of disturbed and mixed land areas have increased at the expense of forested and herbaceous lands in land segment one. Less pronounced, but similar, changes can be seen for land segment two. These alterations, which can be input to a calibrated HSPF simulation, would produce associated changes in predicted flow statistics and other hydrologic phenomena.
Figure 34. User Inputs for the Economics Model in the Sample Run.
Figure 35. Economic Results for the Sample Run.

The economic output for the sample DesktopL2W run shows increases in the amount of land occupied by development tracts. These increases result in new development costs for construction of homes and infrastructure, as well as increased tract values. Changes in assessment values for the newly developed tracts are also reported along with tax revenues that can be collected on those assessments. Finally public sewer and water supply costs are costs are displayed to give a planner knowledge of the likely expenses that accompany the new developments. These results allow the user to assess some of the direct economic benefits and costs of adding development tracts to the watershed.
Figure 36. Fisheries Results for the Sample Run.

The results of the fisheries model in the sample run show that new development is predicted to be especially detrimental to fish health in land segment one. The mean metric score for land segment one drops below the threshold for poor habitat as a result of the addition of disturbed land use. The new developments in land segments one and two in this scenario also adversely affect the mean metric scores in all other land segments, as impacts are carried downstream. These impacts, however, are not serious enough to raise warnings (red colored outlet points) for those land segments.
6.0 DISCUSSION AND CONCLUSIONS

While the data library, set of procedures for GIS-based integration, and DesktopL2W software design provide a useful template for how to analyze a watershed, there are several technical issues that require further discussion.

6.1 Technical Considerations in GIS-Based Integration

A GIS analyst has numerous choices to make along the way in developing an integrative analysis environment. Some of these choices involve raster/vector modes of GIS, spatial resolution, and temporal resolution.

6.1.1 Raster/Vector Modes

Geographic information systems typically rely on two methods of representing space, raster and vector. In the raster mode, the GIS uses regular grid cells to quantize space, while in vector mode, the GIS uses points, lines, and polygons to represent geographic entities. The difference between the two is readily apparent in how each represents polygonal areas (Figure 37). Another less visible difference between the two is their attribute data storage abilities. Raster data sets typically store one piece of data per cell (if more than one piece of data is required, then multiple raster data sets are created). In contrast, vector data sets can store any amount of attribute data attached to its points, lines, or polygons. Both raster and vector modes are used for base and derived data (Table 9-Table 15), as each mode has advantages in representing certain forms of geographic space.

![Figure 37. Raster Versus Vector Representations of Polygonal Regions.](image)

The mode may be given by a source of data (e.g., USGS DEMs come in raster format and DLGs come in vector format), but the GIS analyst decides the mode to be used in the custom integrative software. What mode is picked depends on the needs of models and the style of the interface, but of greater importance are the processing capabilities of the programming environment. In the DesktopL2W software, there are several competing factors that play a role...
in determining the mode of quantizing space. The models tend to rely on input data that are attached to discrete shapes, such as watershed boundaries, so the vector mode is more appropriate for developing model input data. The style of interface for inputting development scenarios in DesktopL2W, however, clamors for the use of raster mode. In laying development tracts on a mapped landscape, the use of regular grid cells is the simplest, most intuitive method. Further promoting the use of the vector mode is the MapObjects programming environment. This environment is capable of displaying both raster and vector data, but all of its functional algorithms (e.g., distance finding, attribute data storage, and rendering/classification) apply only to vector data. This characteristic of the programming environment clearly seals the victory in favor of vector data.

DesktopL2W uses vector data, but manages to look like a raster display by employing a false raster for the development grid. This false raster is composed of a number of cells, each of which is actually a vector square (Figure 38). This technique has the advantage of possessing the simplicity of a raster display, while storing extensive attribute data that is accessed by the models in the software.

![False Raster](image)

Figure 38. False Raster.
6.1.2 Spatial Resolution

Spatial resolution in a GIS environment defines the level of detail at which geographic space is discretized. Fine spatial resolution means that the GIS and its data layers contain more detail, while coarse resolution means less detail. It is assumed that fine-resolution data sets are preferable to coarse data for a model because they more closely approximate reality. The issue of selecting spatial resolution continually appears throughout the process of building an integrative GIS. It is important in selecting base data, storing and manipulating data layers, and developing an interface for models.

The first point at which an analyst considers spatial resolution is in the selection of base data. The size and complexity of the study area are key factors. For a relatively small study area, like the Back Creek Subwatershed, detailed data sets are appropriate. As study areas increase in size, however, storage and processing of spatial data can become cumbersome. For example, in working with DEMs that have thirty-meter cells, the number of cells required to cover the Back Creek Subwatershed is in excess of 160 thousand. The number of cells increases to over 1.6 million (in 24 DEMs) for the Upper Roanoke River Watershed. The size of each of the 24 DEMs is approximately 300 kilobytes. Expansion of a study area to a much larger area, therefore, would require vast storage space for data sets, as well as extreme patience with long processing times for managing them.

Perhaps a more important issue with respect to spatial resolution that arises from this GIS integration project is the resolution needed to work with multiple models. Just as GIS software uses specific methods (raster and vector) to discretize space, models related to water, economics, and ecology have ways of discretizing space. The method used by each model from these fields in DesktopL2W is provided below:

- The HSPF hydrology model uses the land segment as the base unit of analysis. The size of the land segment is determined by the travel time of a drop of water within the land segment’s main reach. In assessing the effects of land use changes to hydrologic conditions, inputs are aggregated for each land segment; i.e., if a commercial tract is added to a land segment, it does not matter where in the land segment it is added, it simply matters that it is in one particular land segment.

- The land values/revenues/expenditures economics model uses the ownership parcel as the base unit of analysis. This level of resolution is much finer than the land segment.

- The fisheries health model, like the hydrology model, uses subwatersheds as its unit of analysis. Note: a much finer resolution is possible for the fisheries model (using zones of influence as the unit of analysis), but the subwatersheds were chosen for demonstration purposes.

The data sets used in the integrative GIS must be able to supply information to the models, while satisfying the resolution requirements of those models. Experience in developing data for the fisheries model provides a good example of how this can be problematic. The model attempts to explain the correlation between the amount of disturbed land and the health of the fish population in a zone of influence. This zone of influence consists of a stream buffer. In the original design of the model, stream buffers were set at thirty meters. The spatial resolution of
the land use data is thirty meters (LANDSAT TM data has 30m x 30m cells). The land use data, therefore is coarse compared to the zone of influence – the zone holds only one cell across its width, and it consists of a string of single cells or portions of cells. With this discrepancy in spatial resolution, an attempt to aggregate land use types in the thirty-meter buffer area could be subject to large errors. Any errors in the land use data would be exaggerated due to the small number of cells per zone of influence. Either the zones of influence or the land use data set, therefore, must be substituted in order to maintain an accurate model. Since the land use data set is the best available, the zones of influence require alteration.

If a single interface for inputting land use changes is desired, then close attention must be paid to setting its spatial resolution, so that the input data can successfully feed all the models. Further considerations constrain the resolution of the interface. The user must be able to work with the interface easily and the unit of analysis must be realistic. The example of the DesktopL2W scenario input screen (Figure 28) describes each of these concerns:

- The unit of analysis for the scenario input screen is the development tract.
- The size of the development tract is much smaller than the size of the land segment or subwatershed, so data from the tracts can be aggregated to form input for the hydrology and fisheries models. Each tract type also contains a specific number and size of parcels, so the tracts can be disaggregated to form input for the economics model.
- The tract size (nine hectares, 300 meters per side) is sufficiently large to allow the user to view tracts clearly over all of Back Creek Subwatershed, but small enough to allow detailed input scenarios.
- The tract size is realistic in that a developer could conceivably develop a nine-hectare tract.
- A bonus of this tract size is that it easily aggregates thirty-meter land use and elevation data, as 100 cells of these data sets equal the size of one tract.

Even with careful attention to the issue of spatial resolution and compatibility, the GIS analyst faces challenges. For example the development grid in DesktopL2W has some cells that cross borders between land segments. This situation leads to controversy when deciding which land segment contains one of these cells. The decision affects inputs to the hydrology model if the user decides to develop a border-straddling cell. DesktopL2W simply assigns the whole cell to the land segment that contains the majority of it. Other solutions, such as splitting cells or using weighted averages of attribute data, are also possible.

6.1.3 Temporal Resolution

Another issue for the GIS analyst to consider in building an integrative system is temporal resolution, which is the time-oriented analog to spatial resolution. GIS is notorious for its lack of ability with respect to managing temporal resolution. Most GIS data sets are snapshots of geographic features. For example, land use, which is perhaps the most important input in DesktopL2W, is based on LANDSAT scenes that were acquired in 1996. Over time, especially as urbanization and suburbanization continue, the baseline conditions of the watershed will change and the model predictions will become less likely to approximate reality. The GIS
analyst should consider, therefore, the ease with which updated data sets can replace outdated ones. Such *swapability* of data, with full modularity of data sets could prove to be a very useful feature.

This condition of changing baseline data applies as strongly to individual models as to land use and other base data sets. For instance, inflation and fluctuating land prices could cause changes to the regression coefficients that form the basis of the economics model. In some cases, replacement of old data sets may not be necessary to account for problems with temporal resolution. For example, the hydrology model depends on precipitation for input. Precipitation data are available in files that contain the whole cycle of precipitation for a year. If the user of the model has the choice of selecting from several files that are representative of many precipitation conditions, then future precipitation activities can be covered by one of the files.

### 6.2 Achievement of Objectives

This research effort reaffirms the utility of GIS for integrated watershed analysis. Five characteristics of GIS account for this utility: (1) spatial functions directly apply to the physical landscape of the watershed, (2) spatial data provides a common source of base data for all models, (3) GIS easily handles the data manipulations and calculations required by models, (4) GIS serves as a warehouse which communicates data and results among models, and (5) GIS provides a common interface for both inputs and display of model results.

The DesktopL2W software tool takes advantage of these five characteristics to form an integrative system. It assembles disparate watershed modeling approaches in a visual setting. Construction of this tool required careful selection and enhancement of spatial data, close collaboration with modeling experts, and attention to the limitations of the GIS environment. Overall, the software and procedures for developing it serve as a template for development of more elaborate software models to meet the needs of EPA’s *Watershed Approach*.

### 6.3 Future Work

The present study underscores a significant potential for further development/improvement of the software. For example, sewer distance is one of the few variables that the GIS calculates in real-time. It breaks the sewer network into lines, calculates the distance from each development tract to each line and stores the minimum distance. This is valuable because the GIS creates attribute data based on the spatial relationship of the layers. In contrast much of the attribute data for the development tracts is derived from the development grid that is used to define where developments can be placed. All of the values for the land value analysis are pre-calculated and stored in the grid. The development tracts then pick up the data based on the grid squares they cover. This methodology requires that the grid pattern be loaded with a great deal of data in order for the GIS to function properly. A key to reproducibility is the ease with which a GIS designer can obtain and configure data sets. In the future, it would be preferable to avoid the detailed data setup of the development grid, and, like the sewer distance, calculate attributes on the fly. Another sample improvement is replacement of software parts with more universal pieces; e.g., the use of a Microsoft Excel spreadsheet for economic results could be substituted with a more universal control so that the user does not need Excel.
There are also numerous possibilities for making DesktopL2W more realistic. The reality of the input scenarios that the user creates is tied directly to the availability of high quality land use data and information about development tracts. Since these values are of the best available quality in the mode, most of the improvements center around data for models and the models themselves. For example, distances to sewer and water lines are currently calculated as straight-line distances. They could more realistically be calculated as lengths along a network of roads. As another example, far more advanced models based on the work of the fisheries modelers could be implemented – zones of influence could be used instead of subwatersheds as the unit of analysis so that proximity of development tracts to streams makes a difference.

As for efficiency, DesktopL2W could benefit from some remodeling undertaken by a computer scientist. Some of the algorithms, notably the one that determines whether a tract belongs to a group of tracts, can be rewritten so they run more quickly.

A key method of increasing accessibility is to make the modeling environment available for users on the world wide web. Researchers are developing this capability out of the functions contained in DesktopL2W. The transfer of functions to the web uses a Java-based interface with ESRI’s Internet Map Server providing a link to a MapObjects display. Progress on this effort can be monitored at the web site http://landscapes.ce.vt.edu/.
GLOSSARY

Buffering – The process of creating areas of calculated distance from a point, line, or area object (DeMers, 1997).

DesktopL2W – A prototype, GIS-based software program that allows a user to input land use changes within a watershed, and run (or at least produce input data for) hydrologic, economic, and fisheries simulations.

Digital Elevation Model (DEM) – Digital model of landform data represented as point elevation values (DeMers, 1997).

Digital Line Graph (DLG) – Digital representation of the graphics contained on USGS topographic maps (DeMers, 1997).

Geographic Information System (GIS) – Computer technology that allows management, analysis, and display of spatial data, which has attached attribute data.

Global Positioning System (GPS) – A satellite-based device that records locational (X,Y,Z) and ancillary data for portions of the earth (DeMers, 1997).

Hydrologic Simulation Program Fortran (HSPF) – A comprehensive hydrologic model used to simulate water quantity and quality processes that occur in a watershed as a result of changes in land use or meteorological inputs.

Hypsography – The use of contour lines to depict a continuously varying surface, such as elevation on a landscape.

LANDSAT Thematic Mapper – A satellite system which produces imagery of the earth’s surface in scenes that can be interpreted and classified into land use categories.

Mean Metric Score (MMS) – A scoring system ranging from 0 to 1 that represents the estimated integrity of a fish community by examination of its taxonomic structure and ecological function.

North American Datum of 1927 (NAD 1927) – A specific set of coordinate values that serve as a reference/base for mapping.

Raster – A form of GIS graphic data structure that quantizes space into a series of uniformly shaped (most commonly squares) cells (DeMers, 1997).

Spatial Data Transfer Standard (SDTS) – A protocol for the structure of spatial data sets meant to make the data sets transferable from one platform to another.

Spatial Resolution – The ability of a GIS or model to display increments in space. The smaller the units it can handle, the better its resolution. In a raster GIS, spatial resolution is represented by the size of a single grid cell.
Temporal Resolution – The level of detail at which a GIS or model is able to represent changes that occur over time.

Thiessen Polygons – A set of polygonal regions around point objects that is defined mathematically by dividing the space between each point and connecting these distances with straight lines (DeMers, 1997).

Topologically Integrated Geographic Encoding and Referencing System (TIGER) – A set of spatial data layers developed by the U.S. Census which enables attachment of tabular Census data.

Vector – A geographic data structure that represents the points, lines, and areas of geographic space by exact X and Y coordinates (DeMers, 1997).

Universal Transverse Mercator (UTM) – A metric grid system of the earth that has 60 zones used for mapping applications that require precise positioning.
REFERENCES


Chanat, Jeffrey, 1999, Personal Communication.


Kaltsas, Ioannis and Darrell Bosch, 1999, Personal Communication.


Rubin, Erik and Clifford A. Shaffer, 1999, Personal Communication.

Speir, Cameron, Darrell Bosch, and Kurt Stephenson, 2000, Personal Communication.


Stancil, Vann and Donald Orth, 2000, Personal Communication.


