ArcBASINS: A GIS Model for the BASINS Database

by

Kristina Antonia Schneider, B.S., B.A.

Thesis

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Master of Science in Engineering

The University of Texas at Austin

May, 2002

ArcBASINS: A GIS Model for the BASINS Database

Approved by
Supervising Committee:

____
David R. Maidment, Supervisor

____
Lynn E. Katz

____
Michael E. Barrett

Dedication

To Austin, Texas

Acknowledgements

First, I would like to thank the US Environmental Protection Agency and the US Geological Survey for providing the funding for my research. I would like to personally thank Russell Kinerson, Paul Cocca, and Mauro Di Luzio for answering my questions concerning the BASINS program. I would like to thank Steve Grise, from Environmental Systems Research Institute Inc., for teaching me the basics of Visio and data modeling.

At the University of Texas at Austin, I would like to thank Dr. David Maidment for his guidance and visions. My thanks goes to the Maidment Research group, past and present, for their insight and support. I also would like to thank Dr. Francisco Olivera, now of Texas A&M, and Dr. Desmond Lawler for providing me with invaluable advice, encouragement, and insights during my studies. At the University of Maryland, I would like to thank Dr. Richard McCuen for persuading me to purse my graduate degree.

My gratitude goes to my parents for their love and support through the years. I would not have made it as far without them. I would like to thank Mark for being the steadying force in my life for so long. Finally, I would like to thank my friends. To all my Austin folk, graduate school would not have been as fun without you. To Carolyn and Erin, thank you for listening and laughing all these years.

April 24, 2002

Abstract

ArcBASINS: A GIS Model for the BASINS Database

Kristina Antonia Schneider, M.S.E.

The University of Texas at Austin, 2002

Supervisor: David R. Maidment

The Better Assessment Science Integrating Point and Nonpoint Sources (BASINS) program, created by the EPA to aid in the TMDL process, provides an extensive database of environmental information. The object-oriented modeling techniques now available with Geographic Information Systems present new possibilities for modeling and organization of environmental data. This thesis explores the creation of a data model for the BASINS database, entitled ArcBASINS. The modeling techniques utilized, the two implementation methods devised for model, and some sample applications are presented. The result is a data structure that provides a standard format for the categorization and maintenance of both spatial and tabular environmental data.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

Environmental management evolves and changes as more knowledge and technology becomes available to the environmental community. One of the most powerful tools to emerge in the environmental field is Geographic Information Systems (GIS), which allows for the collection and analysis of environmental data. GIS fits perfectly into the environmental decision process described by Slaweckt (Slaweckt et al., 2000). He states that the decision process starts with observed data that supports the creation of information through modeling, the information evolves into knowledge through visualization and analysis, and finally the knowledge supports environmental decisions.

Within the steps of this environmental management procedure, innovation must be spurred on by ideas, the users, and technology (Figure 1.1). A more thorough analysis of these elements demonstrates that the thought process clarifies our understanding of the natural environment, the technology allows a straightforward representation of this understanding, and the users apply the knowledge and tools gained.

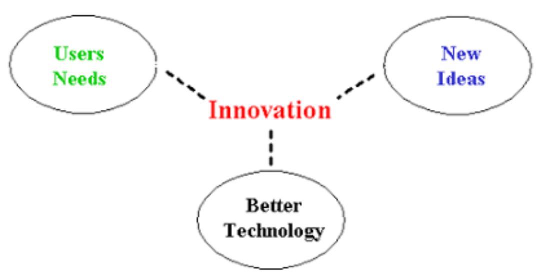


Figure 1.1: The Elements of Innovation.

All three elements are intrinsically linked and they are the drivers of change in the development of information systems. For example, a user's need may call for advancement of the technology and ideas used to represent the environment. One clear example of this is the Total Maximum Daily Load (TMDL) regulation. States and localities are now required to take into consideration all sources of pollution on a watershed level when allocating contaminant loads. The science and engineering communities' understanding and ability to model this all-inclusive view of the environment is incomplete. However, as research is focused on this topic, the community's knowledge base is growing. The Better Assessment Science Integrating Point and Nonpoint Sources (BASINS) program design by EPA's Office of Science and Technology is one of the new ideas spurred on by the TMDL requirement. This program for the first time brings together large amounts of environmental data and modeling capabilities in a single package. The platform of the program is Environmental Systems Research Institute's (ESRI) ArcView v 3.0 GIS program.

ESRI is the leading producer of GIS software and they have heavily influenced the market. Their first major GIS software product was ArcInfo, which was followed by the creation of ArcView. ArcView ver. 3 is still the most prevalent program in the GIS world, but in 2000 ESRI released their newest GIS program ArcGIS. This software system is a true step forward for the GIS industry and has prompted the investigation of possibilities for the implementation of BASINS in ArcGIS explored in this thesis. ArcGIS continues the integration of data and analysis first attempted by ArcView by allowing a truly object-oriented representation of geographic features. ArcGIS utilizes Object Oriented programming and data design to create and relate data within its structure. The blue lines representing a river on a GIS map may now possess attributes and behaviors that more accurately represent their function in the real world.

One of the ways to harness the capability of ArcGIS is to create data models, which use ArcGIS' ability to create and interact with custom objects. To this end, ESRI and the Center for Research in Water Resources at the University of Texas at Austin have formed a Consortium for GIS in Water Resources. This Consortium has created a data model to represent water resource data called the ArcGIS Hydro data model (ArcHydro). The data model's primary purpose is to bring the worlds of GIS and hydrology closer together. It provides a geospatial framework for describing water resources data as well as representing the connection between the data. The model establishes a GIS basis for hydrologic modeling. The ArcHydro data model has gone through numerous revisions since its conception in the summer of 1999 and is now structurally sound. The creation of the ArcHydro data model and the innovation in GIS technology presented by ArcGIS provides a mechanism to organize and attribute environmental data in such a manner that will enhance the environmental management process.

1.1 MOTIVATION

The BASINS program has three objectives: to facilitate the examination of environmental information; to support analysis of environmental systems; and to provide a framework for examining management alternatives (US EPA, 1998). The program brings together key environmental data and analysis components in one program to streamline the many steps involved in watershed based environmental modeling. The ArcView GIS is the instrument currently used to bring together the data and analysis.

The creation of ArcGIS has provided a springboard to enhance the execution of BASINS' objectives. By using the capabilities of ArcGIS to create and maintain data models, it is possible to provide a robust method for organizing and presenting the environmental information available in the BASINS database in the form of a data model. In terms of a database, a data model is a conceptual representation of the data structures that are required by the database (ITS, 2002). The model includes data objects and the associations between them. The model focuses on providing a framework that describes the data needed and how they should be organized.

The ArcHydro data model accomplishes this task for water resources data and it provides a basis for a BASINS data model. The BASINS system is organized around watersheds, which are an important part of the hydrologic system. In addition the BASINS program was designed to help in the implementation of TMDLs, which are water quality based, further justifying the use of ArcHydro as a foundation for the BASINS model. By modifying the ArcHydro data model to fit the needs of the BASINS dataset, the examination of environmental data will be made simpler for the users of BASINS. Data will be presented in a consistent manner for the entire country, thus easing communications between stakeholders.

The implementation of BASINS in ArcGIS is another reason to look into the creation of an ArcGIS BASINS data model. As mentioned, ArcView is the GIS software currently used by BASINS. ArcGIS is the future of GIS and BASINS should be updated to meet the current and future technological needs of its users. The conversion of BASINS into ArcGIS will require an extended process and the exploration of a BASINS data model is the first step in this direction.

1.2 OBJECTIVE AND SCOPE

The key question that underlies the research presented in this thesis is how can the BASINS program be implemented in ArcGIS. ArcGIS is the improved technology that has prompted an investigation into how this new technology can help and support the goal of the BASINS program. ArcGIS has stronger support for networks as compared to ArcView, allows for a fully relational database with permanent feature-to-feature relationships, and permits the use of object modeling for GIS data from ESRI's Arc Objects classes.

Therefore, the purpose of this research is to develop a data model that encompasses the data available within BASINS and provides a basis to fulfill the first objective of the BASINS program in ArcGIS. The model will contain all elements necessary to accurately portray and relate BASINS data. The ArcGIS BASINS Hydro data model is not a water quality model just like the ArcHydro model is not a hydrologic model. It represents a framework for storing environmental information in ArcGIS based on the BASINS database. Water quality models may be built upon this framework, but that is not the focus of this research.

This thesis presents the changes made to the ArcHydro model that are necessary to create a model suited specifically for the BASINS database. The creation of a schema for the implementation of the new data model is explored. Also, the actual implementation of the model for a particular study area is researched to ensure a complete understanding of how the model's structure is achieved. Finally, some hypothetical environmental questions are proposed to investigate the data model's ability to provide answers for environmental managers.

1.3 Study Area

The ArcGIS BASINS Hydro data model stores data from across the United States based on USGS's Hydrologic Cataloging Unit (HUC) system. A cataloging unit is a geographic area representing all of a surface drainage basin, a combination of drainage basins, or a distinct hydrologic feature. The data available from the BASINS database is packaged based on the HUC unit. Naturally, the model allows combinations of HUCs to create a larger dataset. In addition, data not native to the BASINS dataset may be imported into the model's framework.

The Austin-Travis Lakes cataloging unit, whose HUC code is 12090205, was chosen to serve as a study area for the application of the model. The study area is located in central Texas and extends from Burnet in the northwest corner to Austin in the southeast and covers about 1250 sq mi. The northwest corner of the watershed is rural, while the southeast portion contains the urban area of Austin. The Colorado River flows through the Highland Lake system, which is an integral part of the hydrology of the area. Figure 1.2 provides a view of the study area.

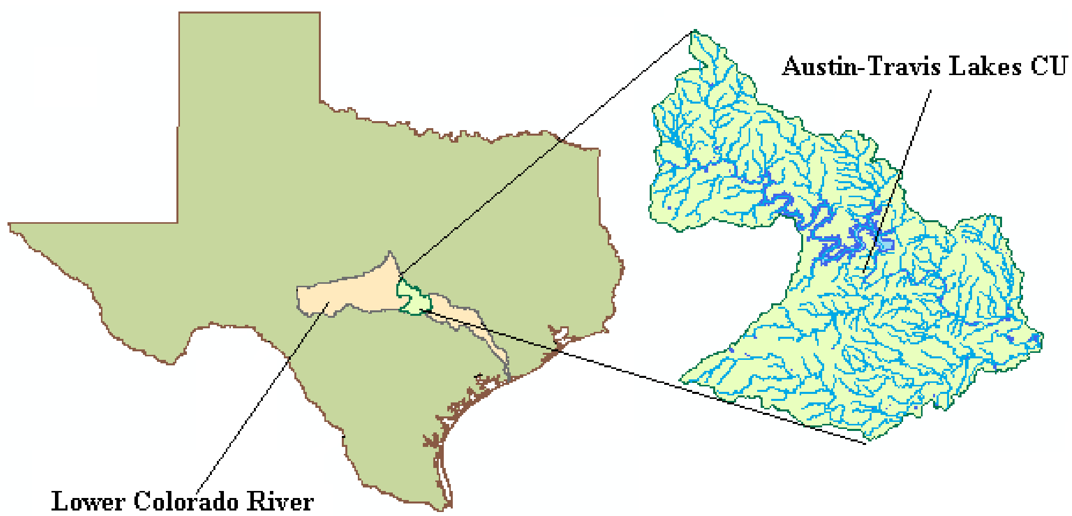


Figure 1.2: The Austin-Travis Lakes Cataloging Unit 12090205

1.4 THESIS OUTLINE

This paper contains six chapters. Chapter One presents an overview of the motivation, objectives, and scope for the development of the ArcGIS BASINS Hydro data model in addition to some background on Geographic Information Systems and the ArcGIS Hydro data model. The second chapter explores two national datasets, the National Hydrography Dataset (NHD) and the BASINS dataset, highlighting the two different levels of data organization. The chapter also explores current GIS data models and the environmental modeling community's call for data standardization. Chapter Three details some of the innovations in the design of the ArcGIS software that were necessary for the creation of the data model. Object oriented modeling is explained. The chapter then outlines the components of the ArcGIS Hydro data model, which provides the basis for the ArcGIS BASINS Hydro data model. The final section of the chapter details the components with their attributes and relationships between the components of the ArcGIS BASINS Hydro data model. The fourth chapter describes the two methods available for the implementation of the data model. The Single Geodatabase method places all BASINS data into one geodatabase. The Hybrid Geodatabase method creates separate databases for the spatial and tabular data available. Chapter Four also considers the role of time series data in the model. The manipulation of the ArcGIS Hydro Data Model, the creation of a model geodatabase, and the application of the model schema is chronicled for each version of implementation. Chapter Five examines the results of the model application: exploring a completed geodatabase, comparing the implementation of both methods, investigating the relationships of the model, describing a procedure for updating the data, and using the BASINS geodatabase to answer three hypothetical environmental questions. The sixth and final chapter draws conclusions, provides recommendations, and looks to future work that may be possible.

Chapter 2: Literature Review

Environmental modeling requires large amounts of data, which is necessary for producing and validating all models. Efficient management and manipulation of large spatial and temporal datasets are needed to apply dynamic watershed models. (DePinto et al., 1994). To aid in the organization of environmental data, numerous data models have been developed. These models can be quite simple, such as a line representing a river, or quite complicated, such as semantic models that turn relationships into mathematical expressions.

To gain insight into the process of data modeling and its application to the environmental field multiple papers are reviewed. First two established datasets, the National Hydrography Dataset (NHD) and the BASINS data package, are explored to give the reader a sense of currently established methods of data organization. Then, a general overview of some GIS data modeling techniques and methods is presented. Finally, the importance of data modeling and management in the field of environmental modeling is investigated.

2.1 ESTABLISHED DATASETS

2.1.1 The National Hydrography Dataset (NHD)

The newest national set of standardized data for hydrography is the National Hydrography Dataset. It is a national set of standardized data for hydrography (the mapping of rivers) developed by the US Environmental Protection Agency (EPA) and the United States Geological Survey (USGS). The NHD incorporates the EPA's River Reach File 3 system and the Digital Line

Graph hydrography data from the USGS. It catalogs the hydrographic features found on USGS topographic maps and is at the same time also a data model currently implemented for the US at 1:100,000 scale and for some areas with the US at 1:24,000 scale. The NHD describes naturally occurring and constructed bodies of water and their paths.

The NHD is currently distributed over the Internet in the NHDinARC format, which is the NHD presented in an ArcInfo format (http://nhd.usgs.gov/data.html). Data can be downloaded from the web in packages that are based on the USGS HUC unit (NHD, 2002). The organization of the NHD river system is broken up into two components, features and reaches, which are the basis for the NHD data model. The dataset is available for all of the United States except for some parts of Alaska.

Features represent hydrographic elements such as paths through water flows and bodies of water. Each feature is classified by a specific Feature Type, which has its own unique Feature Code. Features can be points, lines, or areas. An example of a point, line, or area feature respectively is a Gaging Point, a Stream/River, or a Lake/Pond. Groups of features form Reaches. The features also contain landmark data that is not part of the drainage system. These data are purely cartographic. The features are the raw material upon which the NHD is based (Davis, 2000). Figure 2.1 displays some feature types available in the Austin-Travis Lakes study area.

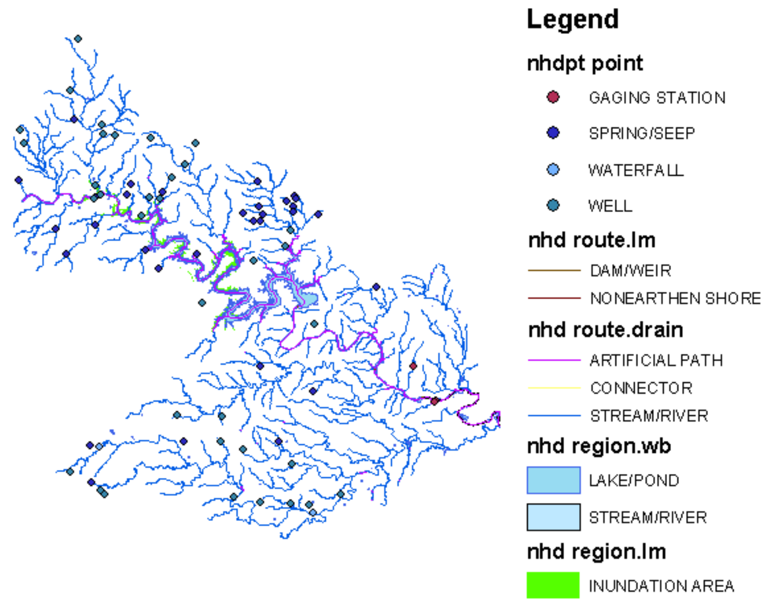


Figure 2.1: NHD Feature Types for HUC 12090205

NHD Reaches are composed of hydrographic entities that have similar hydrologic characteristics. Each reach also has a unique identifying reach code called Reach_Com_ID. Three types of reaches exist: transport reaches, coastline reaches, and waterbody reaches. Transport reaches represent the conveyance of water through the river network. Coastline reaches are located along the major coastlines of the United States and allow for the storage of information necessary to coastlines. Waterbody reaches consist of a set of one or more waterbody features that allow for the relation of additional information to the waterbody. The reaches provide a basis for linking water-related data to the drainage network. This is accomplished by using flow relationships, which relate individual transport and coastline reaches to each other. It is interesting to note that the geometry of the reaches is the same as the features, which is to be expected, given that reaches are derived from features. Figure 2.2 demonstrates this aspect of the NHD.

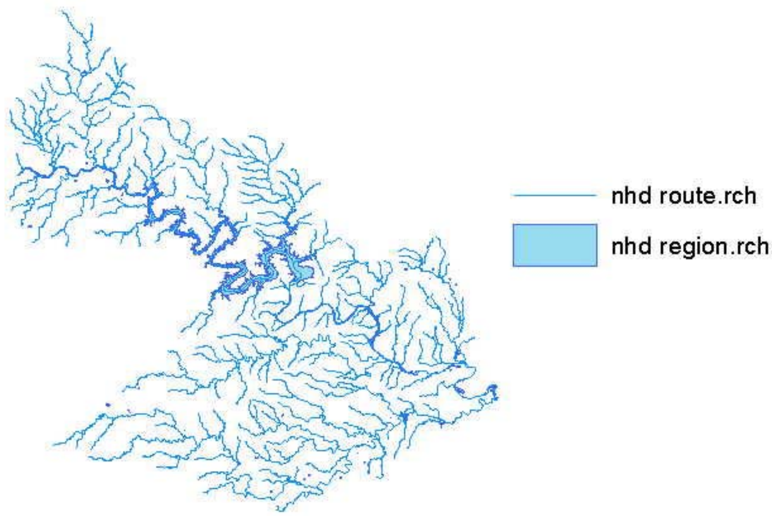


Figure 2.2: NHD Reaches for HUC 12090205

Another interesting feature of the NHD data model is the network flow table, which allows for a tabular tracing of the network. Network information is stored based on the reaches. The flow validation table contains an entry for each reach-to-reach-connection. The table indicates the "flows from" reach and the "flows to" reach for each entry in the table. "Underpasses", rare instances where water crosses each other but does not converge, are also contained in the table (Davis, 2000). The table contains seven attributes to describe the flow along the reaches. The attributes Com_ID_1 and COM_ID_2 list the Reach_Com_ID, which are individual identifiers of the reaches, for the preceding and following reaches for each reach within the flow table. The table also includes descriptions of the direction of flow, the ordering of the inflows and outflows along the interior of the second reach, and the difference in the stream levels between the reaches.

2.1.2 The BASINS Dataset

As described in the introduction, BASINS is an information and modeling package designed to aid in the water quality management process. One of the most important contributions of BASINS is the dataset that is distributed with the program. It was derived from multiple federal sources such as the EPA, the USGS, and the Bureau of the Census. The data was selected for its usefulness for environmental modeling and national availability (U.S. Environmental Protection Agency, 1998). The BASINS dataset is updated with each new version of the program, which happens every two to three years. In the BASINS 2.0 User's Manual the data is broken up into four types: base cartographic data, environmental background data, environmental monitoring data, and point sources/loading data. The following paragraphs and tables describe the data and their sources.

The *base cartographic data* contains information that is vital for the definition of study areas. These data include administrative boundaries, hydrologic boundaries, and major roads. The following table describes the base cartographic data.

BASINS Data Product	Source
Hydrologic Unit Boundaries	USGS
Major Roads	Federal Highway Administration
Populated Place Locations	USGS
Urbanized Areas	Bureau of the Census
State and County Boundaries	USGS
EPA Regions	USEPA

Table 2.1: Base Cartographic Data

The *environmental background data* is intended to support environmental analysis and watershed characterization (U.S. Environmental Protection Agency, 1998). Basic environmental data such as stream location and soil characteristics are located in this data group. The data is used with the models of the BASINS package. The environmental baseline is established by this type of BASINS data. Table 2.2 lists the data available in the environmental background data group.

BASINS Data Product	Source
Ecoregions	U.S. Environmental Protection Agency (USEPA)
National Water Quality Assessment (NAWQA) Study Unit Boundaries	USGS
State Soil and Geographic (STATSGO) Database	U.S. Department of Agriculture, Natural Resources Conservation Service
Managed Area Database	University of California, Santa Barbara
Reach File 1 (RF1)	USEPA
Reach File Version 3 (RF3)	USEPA
Digital Elevation Model (DEM)	USGS
Land Use and Land Cover	USGS

Table 2.2: Environmental Background Data

The *environmental monitoring data* is developed from existing national water quality databases. A few of the data sources had to be converted to geospatial data layers to facilitate their spatial applicability (U.S. Environmental Protection Agency, 1998). In addition, historical water quality trends can be determined from the tabular data that include information from 1970 to 1997. The data types in this group are shown in Table 2.3 below.

BASINS Data Product	Source
Water Quality Monitoring Stations and Data Summaries	USEPA
Bacteria Monitoring Stations and Data Summaries	USEPA
Water Quality Stations and Observation Data	USEPA
National Sediment Inventory (NSI) Stations and Database	USEPA
Listing of Fish and Wildlife Advisories	USEPA
Gage Sites	USGS
Weather Station Sites	National Oceanic and Atmospheric Administration (NOAA)
Drinking Water Supply (DWS) Sites	USEPA
Watershed Data Stations and Database	NOAA
Classified Shellfish Areas	NOAA

Table 2.3: Environmental Background Data

The final data type distributed by BASINS is *Point Source/Loading data*. These data deal exclusively with point source pollutant discharges. A large range of data categories is available; hazardous waste site loadings, air emissions, and traditional water point pollution sources. For the point pollutants, loads and location are estimated and the type of facility is described (U.S. Environmental Protection Agency, 1998). Table 2.4 catalogs the Point Source/Loading data available to BASINS users.

BASINS Data Product	Source
Permit Compliance System (PCS) Sites and Computed Annual Loading	USEPA
Industrial Facilities Discharge (IFD) Sites	USEPA
Toxic Release Inventory (TRI) Sites and Pollutant Release Data	USEPA
Superfund National Priority List Site	USEPA
Resource Conservation and Recovery Information System (RCRIS) Sites	USEPA
Minerals Availability System/Mineral Industry Location System (MAS/MILS)	U.S. Bureau of Mines

Table 2.4: Point Source/Loading Data

The data available with BASINS was presented as four major data types as in the BASINS 2.0 User's Manual. It is interesting to note that, when downloading the BASINS data from the web or the program, the data is not introduced to the user in this format. The data is presented in a data folder in no specific order, as demonstrated in Figure 2.3. This figure demonstrates the lack of data organization present when the BASINS data is originally viewed. The data files do not have the same name as mentioned above and one must refer to the BASINS manual to decipher the names of the files. For example Hydrologic Unit Boundaries are represented in the acc.shp and cat.shp shapefiles. In addition, when one views the data in the BASINS program, it is not presented in the fashion described above. It seems as if the organizational basis for a BASINS' data model exists but is not been implemented in BASINS.

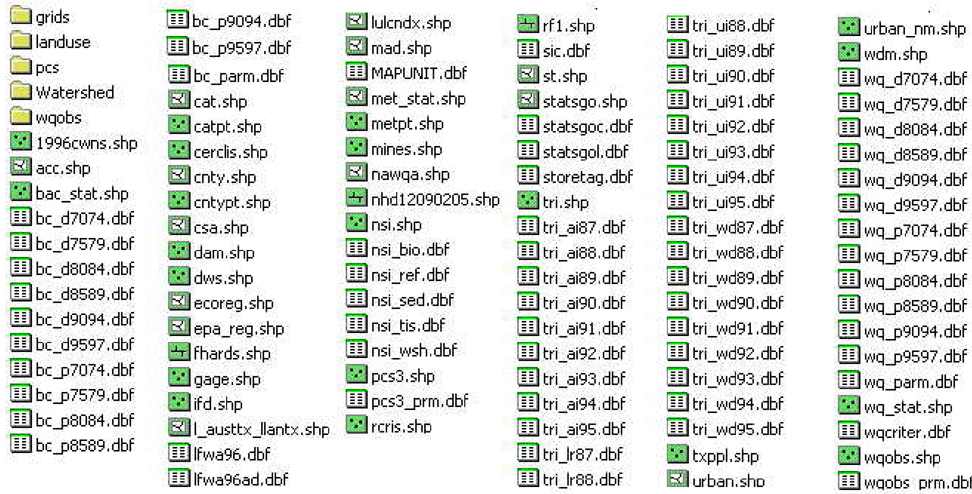


Figure 2.3: The BASINS data folder as viewed in ArcCatalog.

2.2 GIS DATA MODELING

One of the main features of GIS is to present data in both a spatial and tabular form, which creates the need for data management and modeling. Since there are generally two distinct modeling problems (the representation of the natural world and the representation of the man-made world) that GIS technology can manage, two basic data models have been created (Goodchild, 1993). Recently, GIS data modeling has begun to evolve with the implementation of ESRI's geodatabase model.

The two basic data models used in GIS are the field model and the object model. The field model represents the spatial variation of a single variable using a continuous surface (Goodchild, 1993). They represent *continuous space* and therefore are often used to display natural distributions. There are five commonly used field models: irregular point sampling, regular point sampling, contours, cell grid, and triangular nets. A Digital Elevation Model is an example of regular and irregular point sampling that is derived from topographic data and has been converted into a continuous surface. The contour model is compromised of contour lines on maps and contour polygons represent irregular shaped areas with the same elevation values. A TIN (Triangulated Irregular Network) is an example of a triangular net model, and remotely sensed imagery is an example of the cell grid model because a single value is attached to each grid cell (Goodchild, 1993).

The object model represents *discrete space* using points, lines, or areas. Often object models represent man-made creations. They capture abstract human concepts such as the "suburban" area of a city. Also, object models provide a mechanism for the depiction of scientific categorization such as the geomorphological representation of the land surface (Goodchild, 1993). The network model is an extension of the object model, allowing for the representation of continuous data variation over one dimension. Networks form a framework layout for connecting different locations with sets of lines joined at points (Maidment, 2001). The object and network models were represented previously by the *coverage data model*, which was introduced in the ArcInfo GIS software created by ESRI (Environmental Systems Research Inc.). In the coverage model, spatial data was attributed and related topologically.

The newest data model that has been created by ESRI is *the geodatabase data model*. It is an object-oriented model that is introduced in the new software ArcGIS (Zeiler, 1999). This model brings the worlds of field, object, and network modeling closer together. The basic form of the model contains a geodatabase, feature datasets, feature classes, and object classes. An example of a feature dataset would be a group of monitoring stations, where a feature class would be the rain gauge station locations and the object class the yearly rainfall data collected at all the stations. The model is packaged in a geodatabase, which contains the feature datasets. Feature datasets in turn contain all the feature classes in a model and the relationships among them within a common coordinate system (Zeiler, 1999). An object class is a nonspatial entity like a data table and feature classes are objects plus spatial coordinates. Figure 2.4 illustrates the basic framework of the geodatabase model.

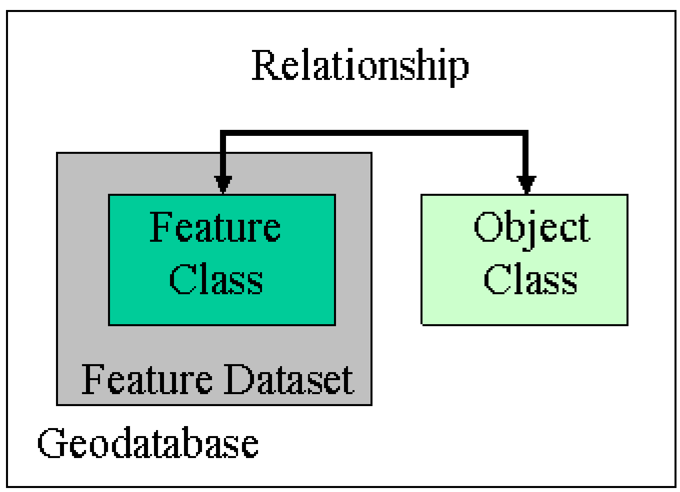


Figure 2.4: Basic Geodatabase Design

The packaging system of the geodatabase model allows users to modify the model for their specific needs or to create formalized geodatabase data models, which detail the geodatabase objects that are necessary for the implementation of a specific model. For example, the Arc Facilities Manager data model was created to provide a GIS data standard for the water and wastewater utility industry. Models such as this can then be applied to standard geodatabases, allowing for almost automatic data standardization.

2.3 Data Modeling and Environmental Modeling

As demonstrated by the two datasets presented previously, GIS data presents a wealth of information to environmental modelers. It provides modelers with an ideal platform for data inventory, mapping, and visualization thereby greatly assisting in model implementation (Sui, 1999). Also, GIS can reduce the many data integration issues that are associated with bringing together data from various sources as illustrated by the BASINS dataset. The merger of multi-scale data from ground-based and remote sensing sources is simplified with GIS (Steyaert, 1993).

However, when one brings multiple data sources together, as demonstrated with the BASINS dataset, there still exists a need to organize the data. Many authors have called for data standardization schemas in the environmental field. Eddy (1993) stated the need to create a global data and information system that makes environmental data available to all in the useable form. Sui and Maggio in their proposal for the new field of GIScience declared the need for new models of geographic concepts as one of the three core elements of the new field of GIScience (Sui, 1999). They stress the importance of developing new models that present key geographical concepts such as space and time in contexts useful for specific disciplines such as environmental modeling. Goodchild (1993) mentions the need for GIS data models to match the data format in environmental models to facilitate a greater integration of the two modeling worlds (Goodchild, 1993). Finally, Robinson and MacKay highlight the fact that modelers may be more comfortable with data models that are based on an environmental framework because of their familiarity with the terminology (Robinson, 1996).

2.4 CONCLUSIONS

Data management is one of the many benefits of using GIS to support the implementation of environmental models. Some datasets, such as the NHD, already contain data organization whereas others like BASINS provide a framework for data management but they do not implement it. In addition, GIS data models have evolved from the field model and object, which provides a context for data representation but does not allow for data interaction, to the new geodatabase model available with ArcGIS. The geodatabase model and its formalized data models present a new possibility for the environmental modeling world. The environmental modeling community is calling for data standardization and the geodatabase model may be a structure that allows the development of standardized data.

Chapter 3: Methodology

The creation of the ArcGIS Hydro Data model and its variations is possible because of the evolution of GIS technology associated with introduction of ArcGIS and the geodatabase model (Davis, 2000). The geodatabase model was briefly reviewed in Chapter 2. This chapter delves more deeply into the geodatabase model and the software that is needed to create customized data models within this framework. Next, the established ArcHydro data model is explained to provide the basis for the creation of the ArcBASINS data model, which is the final result of the research presented in this chapter.

3.1 Geodatabase Model

The current ESRI software, ArcGIS, has introduced an improved structure of the relational database management systems (RDBMS) to the world of GIS. Previous versions of ESRI's GIS software, ArcInfo ver. 7.0 and ArcView ver. 3, contained less sophisticated structures that have some of the functions of a RDBMS and require a feature's spatial information and attribute information to be stored in separate locations (Whiteaker, 2001). The software contained a flat file that is simply a listing or table with data plus a data dictionary. The table allows column values to have a special data type and header and rows to contain a data value. The result is stored in a binary data file that contains both the data and a dictionary specifying the field data types and titles. The new ArcGIS technology allows users to create classes that have an intimate link between an objects spatial representation because data is stored as a set of tables linked by relationships.

These relationships are associations between records in connected tables through values in key fields that the tables share. The user only sees a copy of the original data and the structure is more rigid and interconnected than the file structure used previously (Maidment, 2001).

3.1.1 Geodatabase Concepts

The basic representation of data in ArcGIS is a Feature, which can exist as a Simple Feature or a Network Feature. All points, lines, and polygons are represented by Simple Features. Network Features contain information to define network topology, which represents the connectivity between different point and line features. Simple or Complex Edges and Simple or Complex Junctions are Network Features. A Simple Edge or Junction is a feature that only corresponds to one network element in the logical network whereas a Complex Edge or Junction is a feature that corresponds to more than one network element. A Feature Class is then a collection of Features with each row in a Feature Class table representing an individual Feature (Whiteaker, 2001).

The Feature Dataset gathers all Feature Classes together that share a similar reference frame. The reference frame designates the datum, map projection, the coordinate range, and coordinate precision of the classes. The coordinate range is the domain of the x, y coordinates, the m- (measure along a feature) values, and z values. The coordinate precision basically describes the number of significant digits of the measurements within the feature dataset. For example a spatial reference with a precision of 1 will store integer values, while one with a precision of 100 will store two decimal places. The geometric network built from Network features can be created only within a Feature Dataset. Relationships, an association or link between two objects in a database, can also be created in the Feature Dataset or between Features and Objects in one or more geodatabases. Relationships are stored by Relationship Classes. The relationships are similar to those created in other RDBMS with a common identifier linking related rows in different tables (Whiteaker, 2001). The usefulness of ArcGIS' RDBMS is that spatial features can be methodically related to tabular data. The relational database that stores all relationships, geometric networks, non-spatial tables, Feature Classes and Feature Datasets is the Geodatabase. The Geodatabase may be opened in other RDBMS software such as MS Access and Oracle. Figure 3.1 below shows a simple geodatabase.

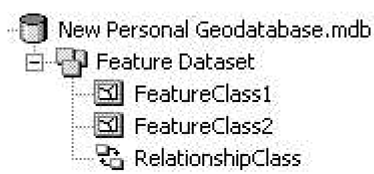


Figure 3.1: Simple Geodatabase Structure

3.1.2 Data Management Software

The program that up until now has been referred to as ArcGIS is actually composed of three programs, ArcInfo 8.1, ArcView 8.1 and ArcEditor 8.1, and with separate add-ons like Spatial Analyst and 3D Analyst. ArcGIS is the main GIS program used throughout this project. These programs are the latest releases of ESRI's GIS software. ArcInfo 8.1 is ESRI's enterprise GIS Software with the geodatabase creation capabilities (Whiteaker, 2001). A user of ArcView 8.1 can edit simple feature classes and shapefiles making it a lightweight auxiliary program for viewing and querying geodatabases (Davis, 2000).

The ArcGIS suite has three major programs, which are ArcCatalog, ArcMap, and ArcToolbox. ArcCatalog is the data management environment of ArcGIS. Here geodatabases can be created, modified, and refined (Zeiler, 1999). The ArcCatalog interface is quite similar to that of Windows Explorer and it allows the users the ability to view their data in three formats: Contents, Preview, and Metadata. The Contents View displays the geodatabases, as well as, coverages and shapefiles that are located in a folder on a disk. Preview allows the user to do what its name implies, which is preview the data in either a geographic or tabular form. Metadata simply displays the metadata associated with a GIS file such as the date the data was created and when it was last updated. Figure 3.2 through Figure 3.4 show an example of the Contents, Preview, and Metadata views of ArcCatalog.

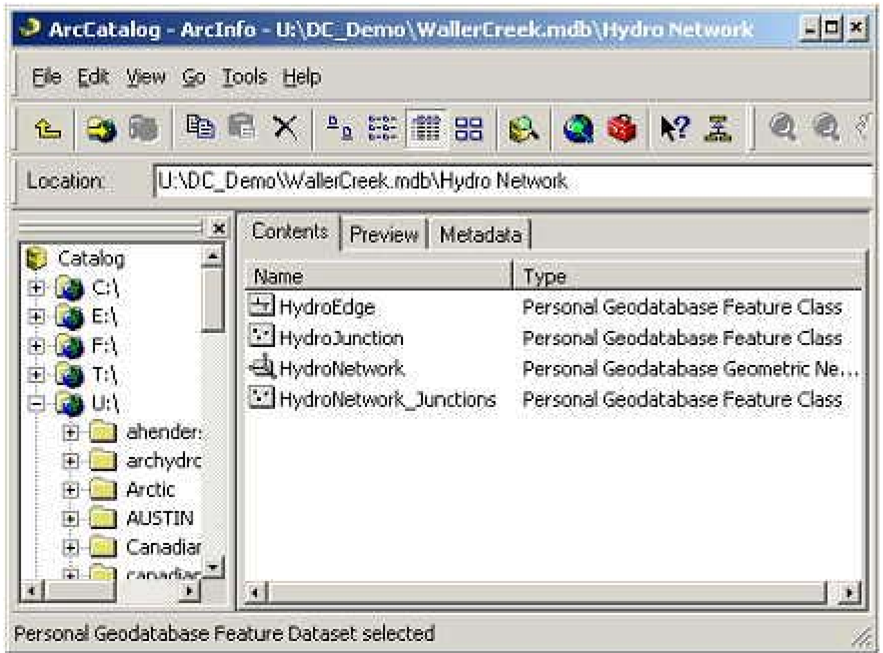


Figure 3.2: The Contents View in ArcCatalog.

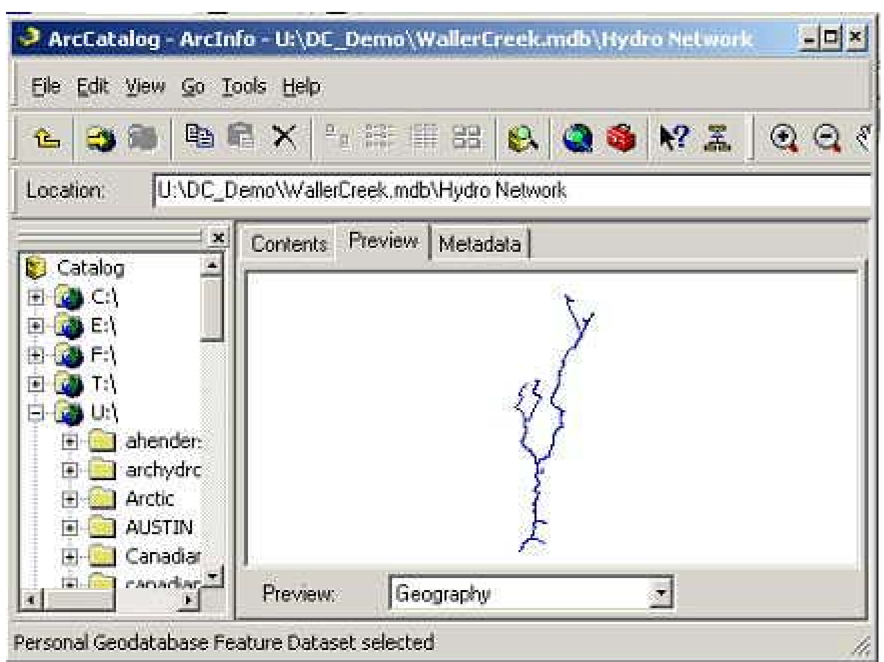


Figure 3.3: The Preview View in ArcCatalog.

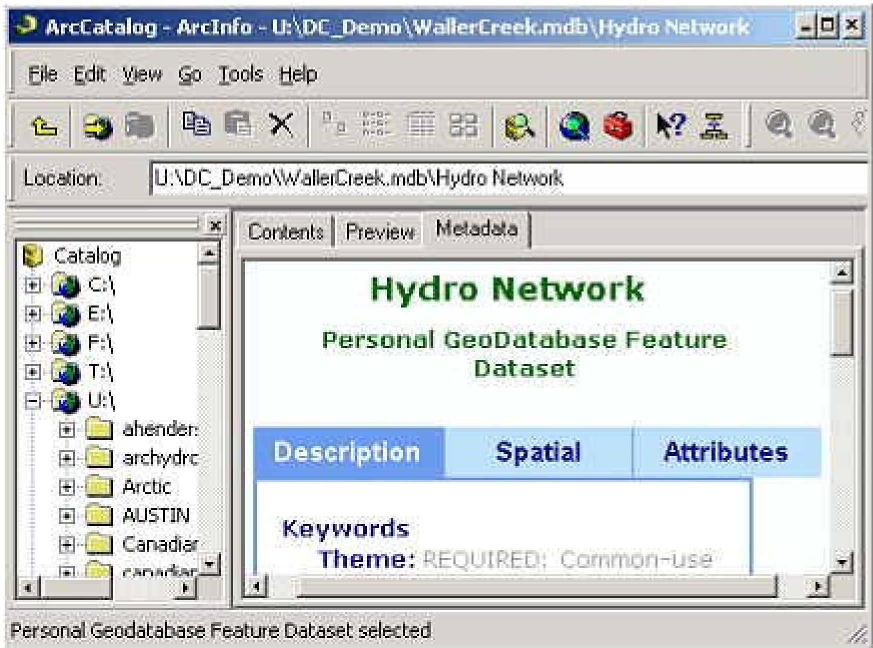


Figure 3.4: The Metadata View in ArcCatalog.

ArcMap is the program in the ArcGIS suite used for displaying and analyzing GIS data. Here, network tracing tools allow for navigation of geometric networks and editing tools are available to manipulate data. Tools can be added and removed from the interface using the Customize interface. Also, ArcMap contains functions for zooming, panning, and making printable maps (Whiteaker, 2001). Figure 3.5 provides a view of ArcMap's Graphical User Interface (GUI).

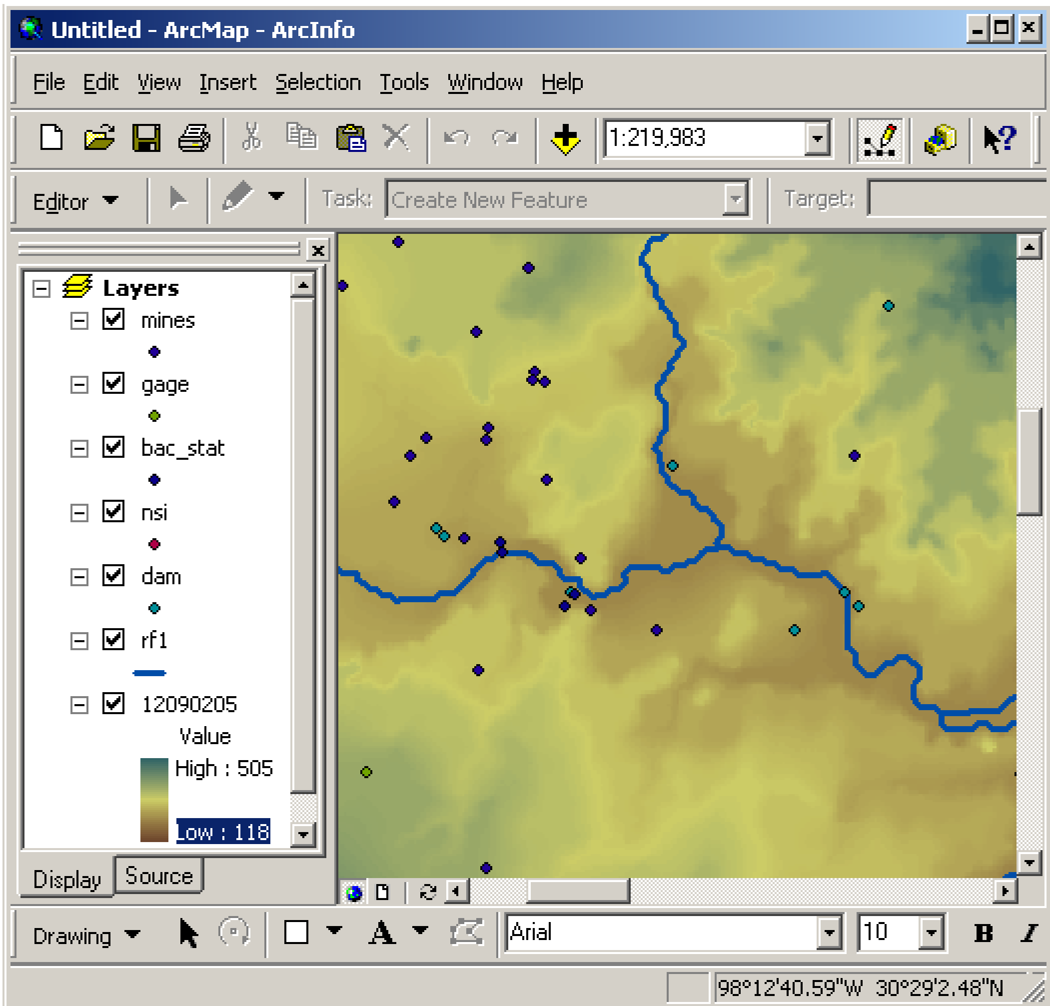


Figure 3.5: ArcMap GUI

ArcToolbox contains many commands that were available with earlier versions of ArcInfo but do not fit into the framework of ArcCatalog or ArcMap. For example, tools exist for projecting data and converting data between different GIS formats. Also, ArcToolbox provides algorithms for creating new data from existing GIS data such as the Centerline tool, which creates a centerline between a dual-line feature (Whiteaker, 2001).

The final database software program that is used in the completion of this research is MS Access. Access as mentioned before is another RDBMS program. Geodatabases can be viewed in Access and Access databases can also be viewed in ArcCatalog. Access is composed of seven major components: tables, relationships, queries, forms, reports, macros, and modules. Tables are the foundation of the database and they store all the data. Each table has a field name at the top of a column and each row presents a record. The Design view, shown in Figure 3.6, contains all the tables available with a geodatabase.

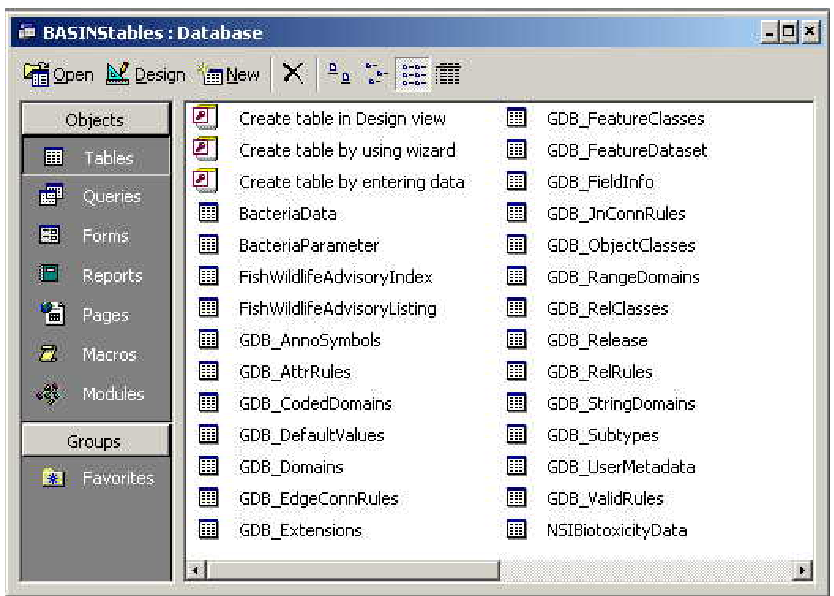


Figure 3.6: Design View in Access

Relationships bind tables together by associating elements that have the same values, with the primary key and foreign key linking the tables. Relationships can be one to many or one to one, where one primary key matches one or many foreign keys in the second table. Queries sort, group, filter, join tables, delete data, update data, and much more. They allow the user to organize the data in their databases in any way needed. Queries are executed by the Structured Query Language (SQL) and can be created in the Design View or by using a wizard. Forms are how data is entered into the database. Forms can be created by using a Form Wizard, as shown in Figure 3.7.

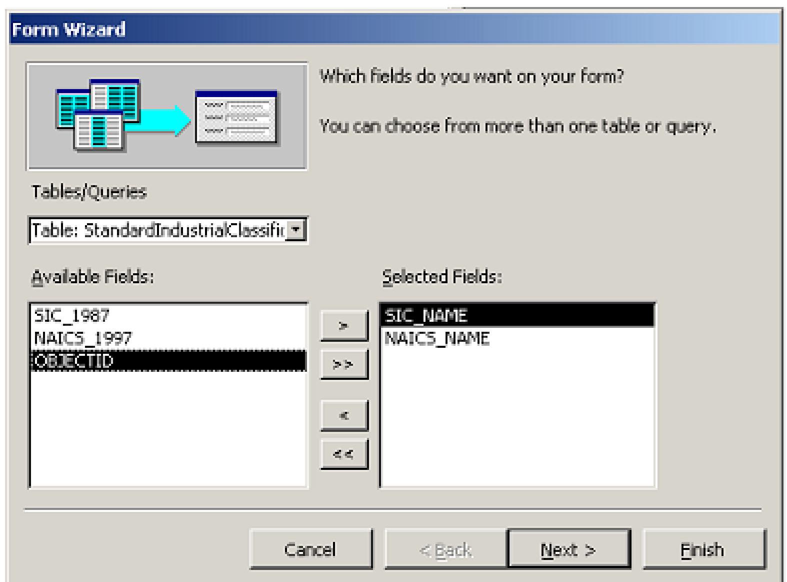


Figure 3.7: A Form Wizard in Access

Reports present the results of the manipulations of the data. They cannot be edited like forms but are intended to be an output to other software. The last components of Access, macros and models, behave much like their counterparts in any other COM compliant software. Macros allow for the automation of actions and modules are where VBA (Visual Basic of Applications) codes are written. Access is a good environment for preprocessing tables that will be placed in a geodatabase and it also provides a database structure for purely tabular data that can be related to ArcGIS geodatabases.

3.2 Object-Oriented Data Modeling

The ArcGIS software provides a greater integration of an object's spatial and attribute information. Object-oriented design is based on the concept of an object. Previously, data and functions have been separated in programming, but an object brings these two important parts of design together (Montlick, 2002).

3.2.1 Unified Modeling Language (UML)

An object, which contains both data and functions for the data, can be interpreted multiple ways, resulting in over 50 different object oriented methods in the early 1990s. In 1994, the authors of the three dominant systems came together and unified their models, creating in 1997 the Unified Modeling Language (UML) (Davis, 2000). UML is a standard language for writing software blueprints using object-oriented techniques (Booch, Rumbaugh, and Jacobson, 1999).

UML provides the user with nine different types of diagrams, ranging from class diagrams to deployment diagrams (Bezivin & Muller, 1998). ArcGIS uses class diagrams to create custom features. A combination of UML's class and package diagram is used to create the features of the Arc Hydro data model. A class diagram describes the static structure of a model and a package diagram, which is a collection of class diagrams, brings together related groups to minimize dependences (Smartdraw, 2002). Custom Features and Object Classes are presented as nodes in the static structure diagram and arcs exhibit the relationships between classes (Whiteaker, 2001). The package diagrams create the Feature Datasets of the geodatabase by bringing together related Features. Figure 3.8 below shows an example of a simple a class diagram.

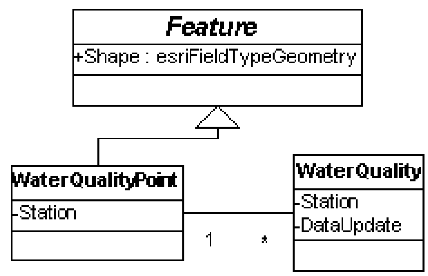


Figure 3.8: Sample UML Diagram

As demonstrated by the generalization (the line with an arrowhead) pointing from WaterQualityPoint to Feature in Figure 3.8, classes can inherit attributes from parent classes. The generalization indicates a relationship between the two classes. Feature is the parent class and WaterQualityPoint is a child class inheriting the Shape attribute. Feature is an abstract class, which stores the common attribute inherited by its child classes but it never has objects of its own. WaterQualityPoint is Feature class that possesses spatial coordinates and it has the attribute Station. WaterQuality is an Object class, which is related to WaterQualityPoint by a one to many relationship using common values of the attribute station to link related data. An Object Class represents a table in which objects are stored with an ObjectID and attributes. One entry in the WaterQualityPoint class may have many entries in the WaterQuality table to represent sample data collected on that station. This is an example of a noninheritance relationship called an association. The UML diagram shown above can be used to create a data model. A more formalized method of showing the relationships within a data model is a UML Analysis diagram. A UML Analysis diagram uses various fonts and color schemes to represent related objects and classes.

3.2.2 CASE Tools

To create the UML drawing and its corresponding data model, a Computer Aided Software Engineering (CASE) tool is utilized. In a CASE program such as Microsoft's Visio 2000, UML diagrams can readily be made. Class and Analysis diagrams are created detailing the elements of an object-oriented data model. The class diagram allows for the object model to be specified but it does not create the code that will inform ArcGIS which classes possess certain attributes or methods. A tool within the CASE program provides a mechanism to export the UML into a Microsoft Repository format, which is an industry wide standard. ArcCatalog reads the Repository using the Schema Creation Wizard. The schema created stores the custom Feature Classes created or it may be applied to existing data. If behaviors are to be added to the objects, code must be written in C++ detailing the actions of the methods. A DLL is then created to house the behavior's code that is linked to the Features in ArcGIS (Whiteaker, 2001).

3.3 THE ARC HYDRO DATA MODEL

The ArcHydro data model is a customized data model created for ArcGIS. As explained in the introduction, the purpose of ArcHydro is to support the data needs of hydrologists and hydrographers. ArcHydro is only intended for natural systems and does not support constructed water infrastructure. ArcHydro was designed in UML using Microsoft's Visio 2000 and as expected works in the ArcGIS software system. The ArcGIS Hydro Data model provides a structured framework for the management of hydrologic data.

The model is composed of Hydro Features, which describe the physical environment in which water flows, and Time Series, which describe the water quality and flow characteristics of the represented water. The Hydro Features each have their own unique identifier called HydroID. This identification allows for relationships to be created between the Hydro Features themselves and Hydro Features and Times Series data. The model is broken up into five components, four of which are represented as Feature Datasets in the geodatabase model. Figure 3.9 displays the five components of the ArcHydro data model. The Feature Datasets are Hydrography, Network, Drainage, and Channel. The fifth package represents Time Series data. The following sections give an overview of the ArcGIS Hydro Data model and Appendix A contains a full scale Analysis diagram of the model.

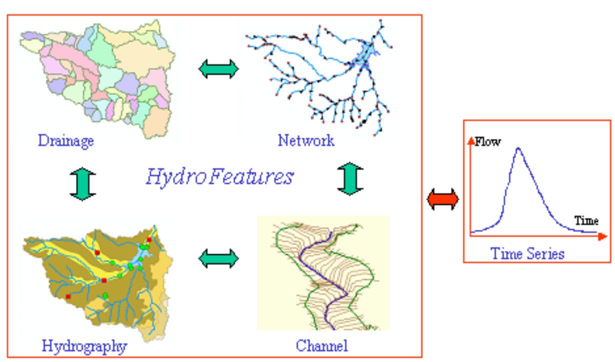


Figure 3.9: The Five Components of the ArcHydro Data Model

3.3.1 Hydrography

The Hydrography Feature Dataset contains the features, which represent the cartographic features of the surface water landscape. The parent class for all Hydro Features is located in this data package. HydroFeature is an abstract class that creates the HydroID and HydroCode, which are the two descriptive codes used in the model. The abstract class Hydrography contains the attributes Name and Ftype, which are necessary for all Hydrography classes. Table 3.x gives a complete listing of Hydrography's feature classes and their attributes and Figure 3.10 displays the Hydrography Analysis Diagram.

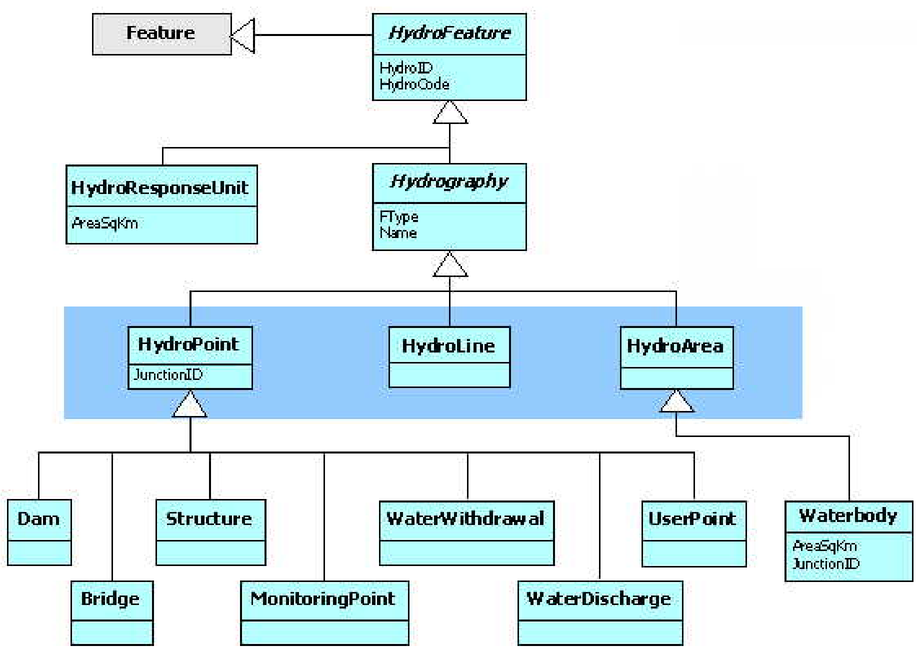


Figure 3.10: Hydrography Analysis Diagram

The cartographic data available in this dataset are stored in the HydroPoint, HydroLine, HydroArea, and Waterbody classes. These data are typical hydrography data such as streams, ponds, and marshes, which are derived from the "blue lines" on cartographic maps. HydroPoints are then further divided into seven more classes to allow for the representation of common tabular datasets. The point classes are Dam, Bridge, Structure, MonitoringPoint, WaterWithdrawal, WaterDischarge, and UserPoint, which are depicted for the Austin-Travis Lakes Study area in Figure 3.11. The final class contained in Hydrography is HydroResponseUnit. HydroResponseUnit represents data describing the hydrologic character of the landscape based on the vertical water balance. An example of a HydroResponseUnit is an area with a unique combination of soil and land use characteristics.

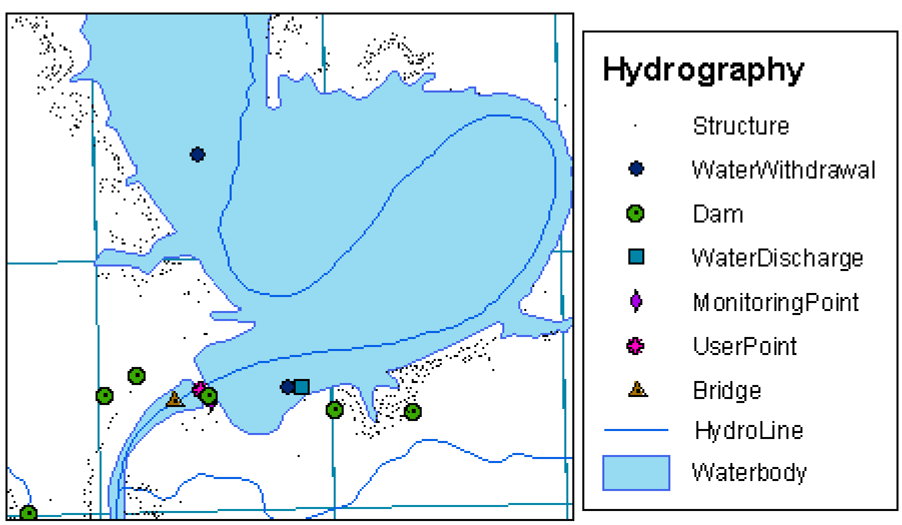


Figure 3.11: HydroPoints

Class (Type)	Attribute	Description
Dam (Point)		A structure creating a pond or reservoir storing water
Bridge (Point)		A structure carrying a road across a stream
Structure (Point)		Any other kind of water resources structure
MonitoringPoint (Point)		A measurement station or sampling point
WaterWithdrawal (Point)		Point of withdrawal of water
WaterDischarge (Point)		Point of discharge of water
UserPoint (Point)		Any other point of interest
HydroFeature (Abstract)		Abstract class with common attributes and methods for Hydro Features
	HydroID	Unique feature identifier in the geodatabase
	HydroCode	A permanent, public identifier of the Hydro Feature.
Hydrography (Abstract)		Abstract class containing attributes and methods that are unique for Hydrography
	FType	Type of geographic feature
	Name	Geographic name
HydroPoint (Point)		Point features from map hydrography and inventory sources
	JunctionID	Identifier for the corresponding junction on the network
HydroLine (Polyline)		Line features from map hydrography
HydroArea (Polygon)		Area features from map hydrography
Waterbody (Polygon)		An area of water
	AreaInSqKm	Area independent of map units
	JunctionID	Identifier for the junction at the outlet of the area
HydroResponseUnit (Polygon)		Any subdivision of the landscape used for surface water balance accounting
	AreaInSqKm	Area independent of map units

Table 3.1: Hydrography Classes

3.3.2 Network

The Network describes the connectivity of water flow through the landscape using a water resource network based on streams, rivers, and the centerlines of waterbodies. It is the heart of the model allowing for river flow and contaminant transport to be routed through the landscape. The Network contains two networks: HydroNetwork and a Schematic Network. Table 3.x gives a complete listing of the classes and their attributes available in the Network Feature Dataset. Also, the Analysis Diagram of the HydroNetwork is located in Figure 3.12 below.

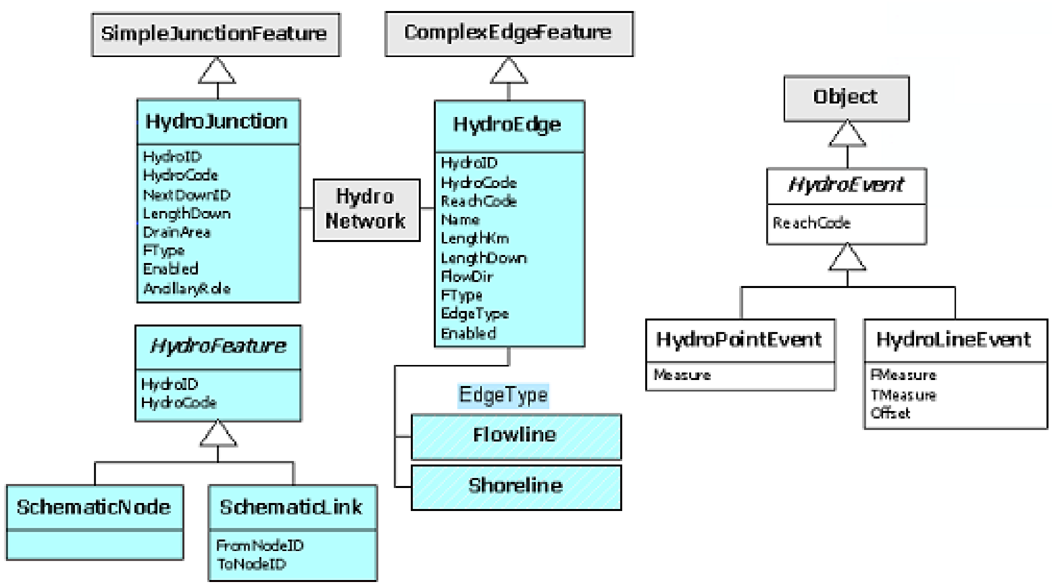


Figure 3.12: Network Analysis Diagram

The HydroNetwork is associated with ArcGIS's Utility Network Analysis system and is a geometric network. The main Hydro Features of HydroNetwork are HydroEdges and HydroJunctions. HydroJunctions connect DrainageAreas, Waterbodies, and HydroPoints to the river network. They are Simple Junction Features, which may be referred to by FType. The Ftype attribute categorizes the type of geographic feature represented by the HydroJunction. HydroJunctions may include any points that are of interest to the user for river network navigation. Generic Network Junctions are created by ArcGIS when the geometric network is created. They connect HydroEdges to each other. HydroEdges transport water along the network and can represent two categories of natural river systems: Flowlines and Shorelines. Flowlines represent single line streams, the centerlines of double lined streams, and the centerlines of waterbodies. Shorelines are the edges of double lined streams, the coastlines of oceans and seas, and the shorelines of lakes and ponds. Water does not flow along Shorelines but only along Flowlines. HydroEdges are of the ComplexEdgeFeature type to allow the placement of junctions along an edge without breaking the edge. The key attributes of HydroEdge are ReachCode, LengthDownstream, EdgeType, and FType. These attributes of HydroEdge as well as others are described in Table 3.2, which summaries the features of Network at the end of this section.

The other network available in Network is the Schematic Network, whose main purpose is to provide a simplified view of water flow through the natural environment and represent the connection of drainage areas to the river network. The Schematic Network is made from simple point and line features that are linked by the fact that the Schematic Links have attributes, FromNodeID and ToNodeID, that refer to the HydroID's of the Schematic Nodes at their end points. The Schematic Links and Nodes may be used to create a geometric network if necessary. SchematicLink and SchematicNode are the two Hydro Features that create the Schematic Network. SchematicLink is line feature of this simple network and it uses the straightforward ToNode – FromNode topology to describe its connectivity. Figure 3.13 shows an example of Austin-Travis Lakes' Schematic Network overlaid on top of the area's Digital Elevation Model along with Watersheds derived for the HUC.

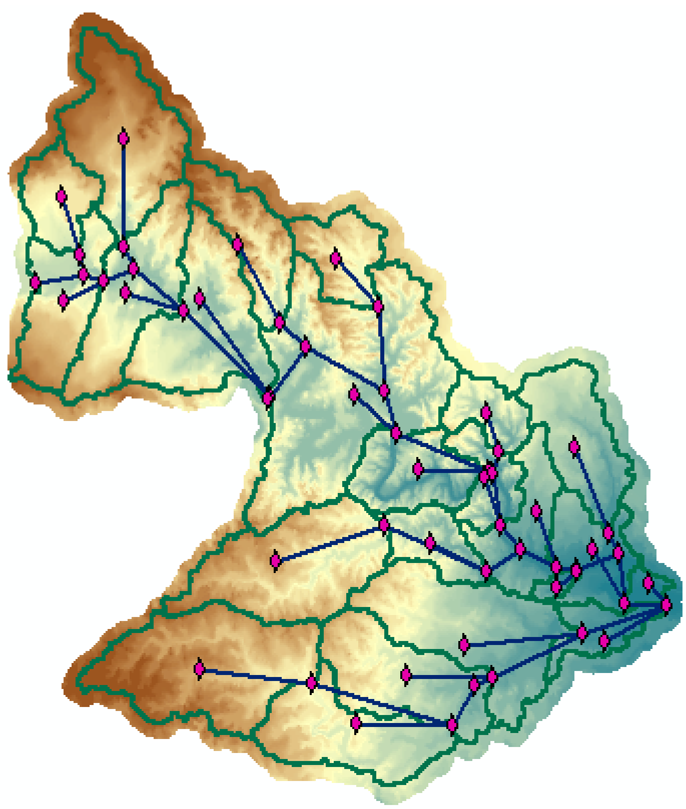


Figure 3.13: Schematic Network for HUC 12090205 with its DEM and Watersheds

The final element of the Network Feature Dataset is the HydroEvent abstract class and its child classes HydroPointEvent and HydroLineEvent. This group of features is designed to support linear referencing on the network and the creation of other Events. The HydroEvent class contains the ReachCode attribute that is related to the ReachCode attribute on HydroEdge.

Class (Type)	**Attribute**	**Description**
HydroEdge (Complex Edge)		Linear segments in the HydroNetwork
	HydroID	Unique feature identifier in the geodatabase
	HydroCode	A permanent, public identifier of the Hydro Feature.
	ReachCode	Reach identifier
	Name	Geographic name
	LengthKm	Length independent of map units
	LengthDown	Length along shortest path to a downstream reference location
	FlowDir	Labels flow direction
	FType	Type of geographic feature
	EdgeType	Type of HydroEdge: Flowline = 1or Shoreline = 2
HydroJunction (Simple Junction)		Junctions in the HydroNetwork, used for outlets, sinks, or other purposes
	HydroID	Unique feature identifier in the geodatabase
	HydroCode	A permanent, public identifier of the Hydro Feature.
	NextDownID	Identifier for next downstream feature in the HydroNetwork
	LengthDown	Length along shortest path to a downstream reference location
	DrainArea	The upstream drainage area to this junction.
	FType	Type of geographic feature

SchmaticNode (Point)		A point in a Schematic Network
SchmaticLink (Polyline)		Straight line connections between selected points in network
	FromNodeID	HydroID of SchmaticNode located upstream of link.
	ToNodeID	HydroID of SchmaticNode located downstream of link.
HydroEvent (Abstract)		Abstract class with common attributes and methods for events
	ReachCode	Identifier of linear referencing segment, maps to ReachCode on HydroEdge
HydroPointEvent (Object)		A point event
	Measure	Measure location of an event
HydroLineEvent (Object)		A line event
	FMeasure	Measure location of the start of the line event
	TMeasure	Measure location of the end of the line event
	Offset	Offset distance that the event is displayed from the HydroEdge

Table 3.2: Network Attributes

3.3.3 Drainage

The data contained in the Drainage Feature dataset are derived from elevation based drainage patterns. The definition of a drainage area is the area of the land contributing drainage to a particular feature (could be a point, line, or area). Digital Elevation Models (DEMS) allow for the simple calculation of a drainage area for a given point in GIS. DEMS have become vital to hydrologic calculations in the GIS field and for this reason the Drainage Feature dataset was derived. Figure 3. 13 shows the DEM for the Austin Travis Lakes study area and some of its derived drainage areas.

HydroFeature is the abstract parent class of all the classes in this feature dataset as it is for most classes in the model. DrainageFeature is the abstract class that provides the DrainID for all features in the feature dataset. The feature dataset also contains a point, line, and area class to house these specific geometries for DEM/TIN derived data. Under the DrainageArea class are three more specific types of drainage areas. They are Catchment, Watershed, and BASIN. Figure 3.14 illustrates the Analysis diagram for Drainage, and Table 3.3describes these areas in more details as well as the other features in the Drainage Feature Dataset.

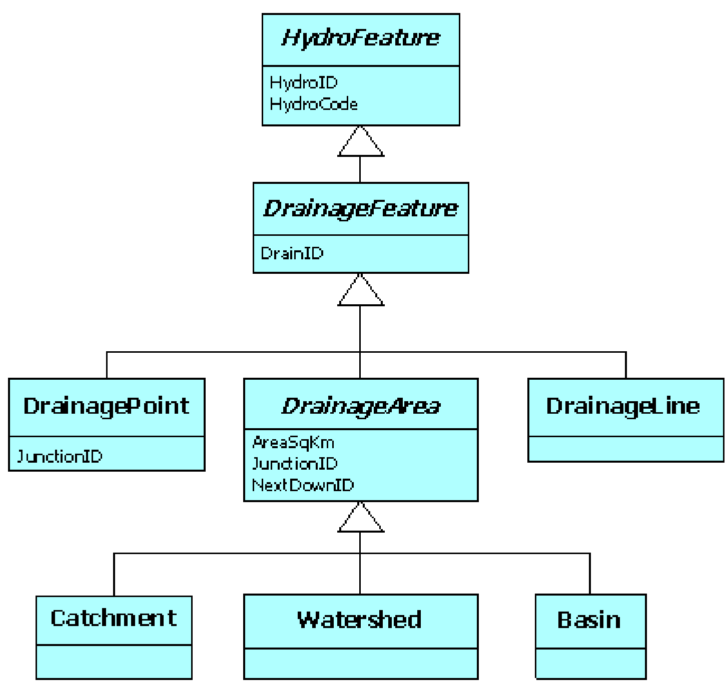


Figure 3.14: Analysis Diagram for the Drainage Feature Dataset

Class (Type)	**Attribute**	**Description**
HydroFeature (Abstract)		Abstract class with common attributes and methods for Hydro Features
	HydroID	Unique feature identifier in the geodatabase
	HydroCode	A permanent, public identifier of the Hydro Feature.
DrainageFeature (Abstract)		Abstract class for drainage system features
	DrainID	Link between point, line and area features of a drainage system, such as GridCode, Pfaffstetter number, or HUC number
DrainagePoint (Point)		Point at the center of a DEM cell on a drainage path, usually seed point location for drainage area delineation
	JunctionID	Identifier for the HydroJunction that corresponds to the drainage point
DrainageLine (Polyline)		Line through the centers of the DEM cells on a drainage path
DrainageArea (Abstract)		Abstract class for common drainage area attributes
	AreaSqKm	Drainage area independent of map units
	NextDownID	Identifier of next downstream area in this drainage area Feature Class
	JunctionID	Identifier for the HydroJunction at the outlet of the area
Catchment (Polygon)		An elementary drainage area produced by a uniform process of landscape subdivision
Watershed (Polygon)		Any subdivision of the landscape into drainage areas
Basin (Polygon)		A set of standardized drainage areas for data archiving and delivery

Table 3.3: Drainage Classes

3.3.4 Channel

The channel feature dataset describes the three dimensional nature of the riverbeds. This information is necessary for the study of the river geomorphology, ecology, and flood inundation properties. The geometry of the river channel and the adjacent floodplain are shown as combination of lines and points. ProfileLines are drawn parallel to the flow of the river and can represent the thalwag or banklines of the river. CrossSections are drawn perpendicular to the channel and they provide information about the elevation variation of the channel bed. ProfileLine and CrossSection are of the 3D PolylineM geometry type meaning they contain (x, y) spatial location, elevation, and measured values. CrossSectionPoint is an Object class that stores the cross-sectional data associated with the numerous cross section points on the CrossSection feature. All features in the class inherit attributes from HydroFeature and ChannelFeature. The table and figure below describe Channel's classes.

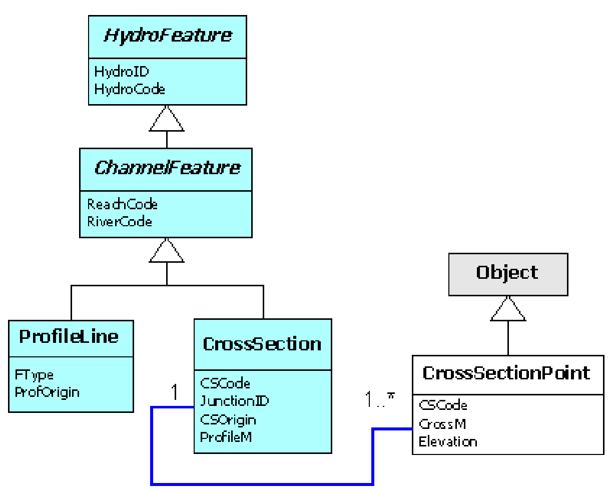


Figure 3.15: Channel Feature Class Analysis Diagram

Class (Type)	Attribute	Description
HydroFeature (Abstract)		Abstract class with common attributes and methods for Hydro Features
	HydroID	Unique feature identifier in the geodatabase
	HydroCode	A permanent, public identifier of the Hydro Feature.
ChannelFeature (Abstract)		Abstract class for common channel attributes
	ReachCode	Identifier of linear referencing segment, analogous to ReachID on HydroEdge
	RiverCode	Identifier of linear referencing segment, usually corresponds to named rivers
ProfileLine (3D Polyline)		Longitudinal profile of the channel
	FType	Type of geographic feature
	ProfOrigin	Description of origin of profile line data
CrossSection (3D Polyline)		Transverse section of a channel
	CSCode	CrossSection identifier
	CSOrigin	Description of origin of cross section data
	ProfileM	Location of the CrossSection on ProfileLine's measure system
	JunctionID	Identifier for the junction at the outlet of the area
CrossSectionPoint (Object)		Non-spatial cross-section data
	CSCode	Identifier of the corresponding CrossSection feature
	CrossM	CrossSection measure point location
	Elevation	Elevation of CrossSection point above mean sea level

Table 3.4: Channel Classes

3.3.5 Time Series

The Time Series package allows for the representation of the feature classes' attributes. For example, the discharge and water quality parameters of a river are shown in this package. Time Series contains two related tables. Time Series contains the actual time varying data associated with specific features in the model. The data is related to the feature based on the FeatureID attribute, which is the HydroID of the intended feature. The other attributes of TimeSeries allow for the representation of the type of time series data, the date and time the data was taken, and the actual record. TimeSeries is related to TSType through the TSType and TSTypeID fields respectively. The TSType table allows for a more complete description of the time series type. This table stores information about the variable that is recorded, its units, the data type, and other information. Figure 3.x below gives a pictorial explanation of the Time Series package.

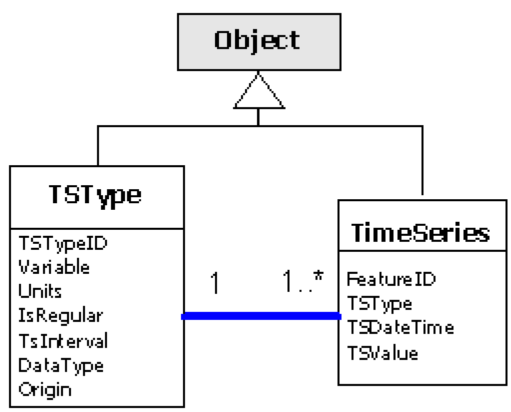
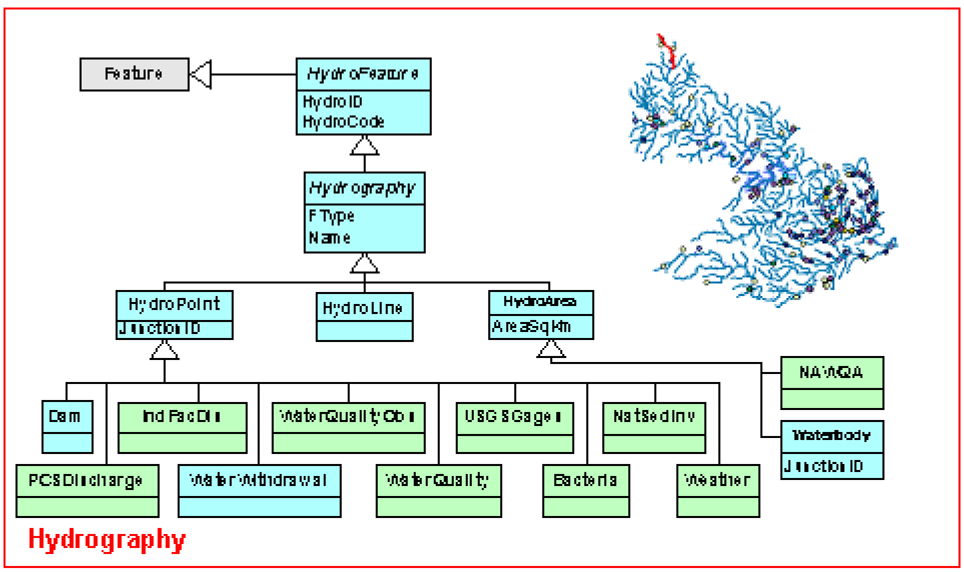
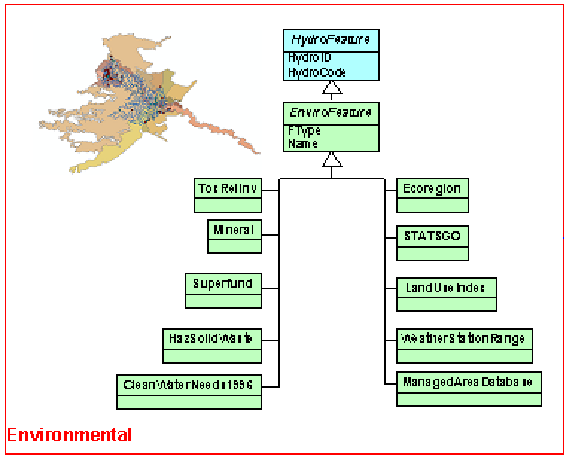


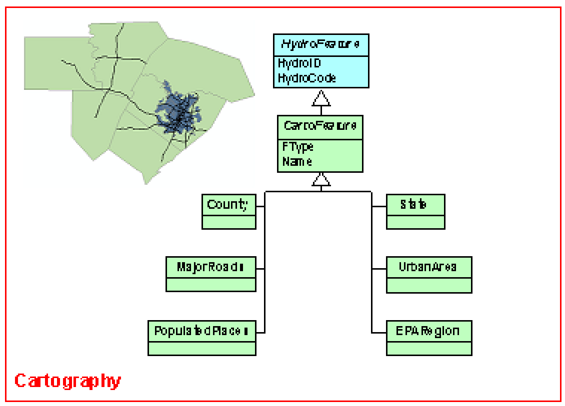
Figure 3.16: Time Series Package Analysis Diagram

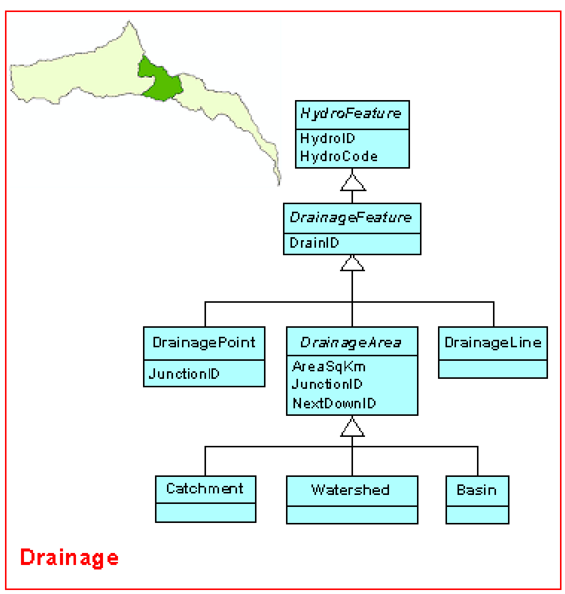
3.4 THE ARCGIS BASINS HYDRO DATA MODEL

As mention previously, the data available from the BASINS GIS package is a powerful tool but it is presented to the user in an unorganized fashion. To aid the users in application of the data and to help usher the BASINS program into the realm of ArcGIS technology, an ArcBASINS data model based on the ArcGIS Hydro Model has been created. This model combines the worlds of environmental and water resource data, using the data model described in the BASINS version 2.0 Manual as a guide. Figure 3.17 shows the spatial data packages of the ArcGIS BASINS Hydro Data Model and Appendix B contains a full version of the model.




HydroFeature
HydroID
HydroCode
EnviroFeature
FType
Name
ToxRelInv
Mineral
Superfund
HazSolidWaste
CleanWaterNeeds1996
Ecoregion
STATSGO
LandUseIndex
WeatherStationRange
ManagedAreaDatabase
Environmental


HydroFeature
HydroID
HydroCode
CartoFeature
FType
Name
County
MajorRoads
PopulatedPlaces
State
UrbanArea
EPARegion
Cartography


HydroFeature
HydroID
HydroCode
DrainageFeature
DrainID
DrainagePoint
JunctionID
DrainageArea
AreaSqKm
JunctionID
NextDownID
DrainageLine
Catchment
Watershed
Basin
Drainage

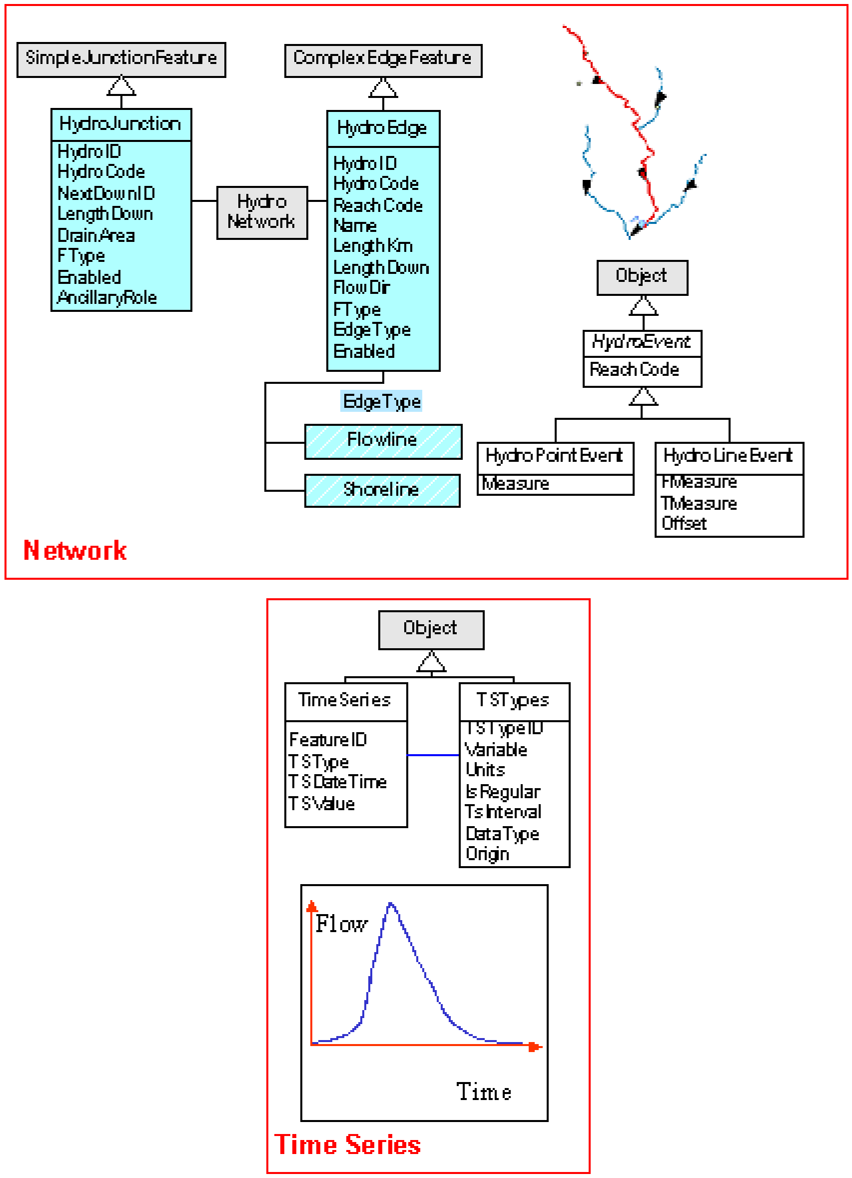


Figure 3.17: ArcGIS BASINS Hydro Data Model Feature Datasets

This section explores the creation of the model by first detailing the feature datasets and feature classes that were deleted from the original Arc Hydro model. Some datasets, as well as, object tables had to be added to completely capture the information available in BASINS. Finally, the relationships necessary to the full implementation of the BASINS data management scheme are explained.

3.4.1 Original Feature Datasets

The original feature datasets of the Arc Hydro model provide an excellent starting point for the creation of the Arc BASINS data model. Of the five data packages of Arc Hydro, four packages were retained. The one package that was deleted was the Channel feature dataset, which deals with three-dimensional nature of river channels. No such data is available through the BASINS dataset and therefore it was decided to eliminate this feature dataset.

The four preserved data packages are Time Series, Drainage, Network, and Hydrography. Time Series and Drainage are not altered in any fashion in comparison with the ArcHydro Data model. Network and Hydrography are altered to match the needs of the BASINS database. The feature classes SchematicNode and SchematicLink are removed from the Network feature dataset. As with the Channel feature dataset, the data available within BASINS does not contain data that matches with the purpose of SchematicNode and SchematicLink and therefore these classes were deleted from the model. Figure 3.18 allows for a further comparison of the original datasets in the ArcHydro and ArcBASINS models.

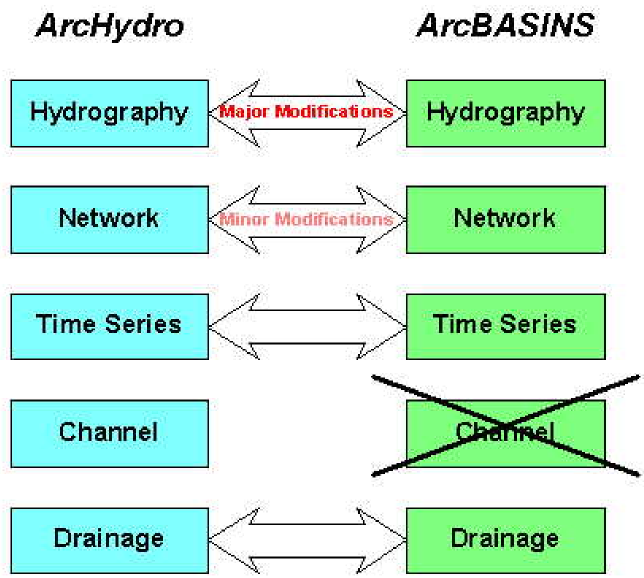


Figure 3.18: A Comparison of the ArcHydro and ArcBASINS models

As shown in Figure 3.x, most changes have occurred to the Hydrography data package. Multiple classes were removed because of their absence from the BASINS dataset. HydroResponseUnit was deleted from the model. The removed HydroPoint child classes are Structure, Bridge, UserPoint, WaterDischarge and MonitoringPoint. The feature classes USGSGage, NatSedInv, Bacteria, WaterQuality, WaterQualityObs, and Weather are the replacements for MonitoringPoint. WaterDischarge is split into two classes IndFacDis and PCSDischarge. One class, NAWQA, was added under HydroArea. Table 3.5 documents the data origins of the new feature classes and Figure 3.19 shows the altered Hydrography Analysis diagram. The green boxes in Figure 3.19 represent classes added because of the BASINS dataset and blue boxes are from the original ArcHydro model.

Feature Class	BASINS Data Origin
Dam	National Inventory of Dams
IndFacDis	Industrial Facilities Discharge (IFD) Sites
PCSDischarge	Permit Compliance System (PCS) Sites and Computed Annual Loading
WaterWithdrawal	Drinking Water Supply (DWS) Sites
WaterQuality	Water Quality Monitoring Stations and Data Summaries
WaterQualityObs	Water Quality Stations and Observation Data
USGSGages	Gage Sites
NatSedInv	National Sediment Inventory (NSI) Stations and Database
Bacteria	Bacteria Monitoring Stations and Data Summaries
Weather	Weather Station Sites
NAWQA	National Water Quality Assessment (NAWQA) Study Unit Boundaries

Table 3.5: New Hydrography Feature Class BASINS Data

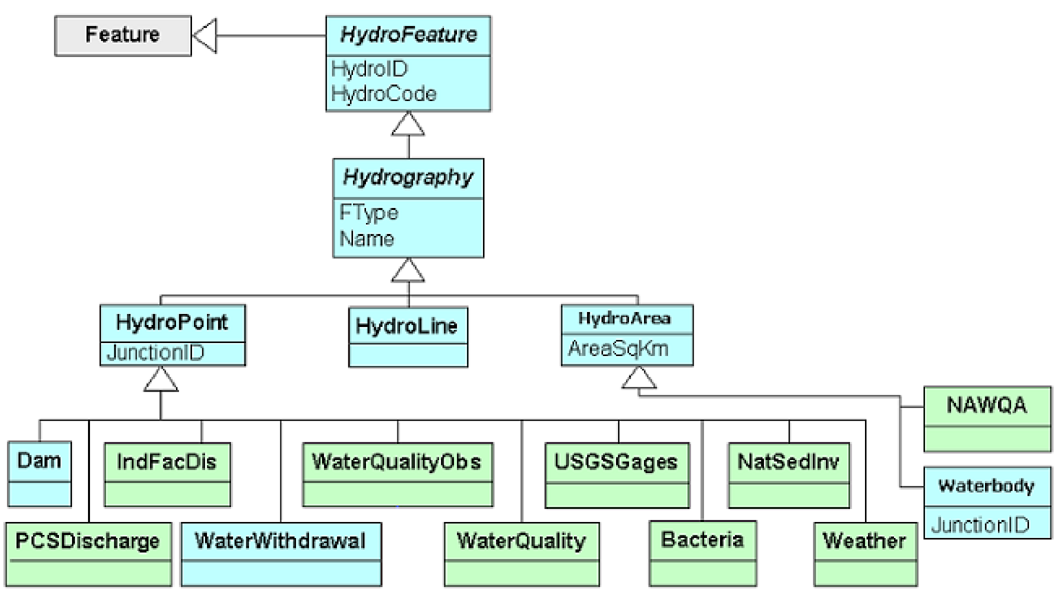


Figure 3.19: ArcBASINS' Hydrography Feature Dataset

3.4.2 Added Feature Datasets

The original structure of the ArcHydro data model was not sufficient to capture the true essence of the data package provided by BASINS. Therefore, it was decided that two new feature datasets should be added to the model. These packages are inspired by the data design described in the BASINS 2.0 Manual.

All environmentally relevant data that could not be accommodated in another feature dataset is placed in the Environmental Feature dataset. This package contains information that is vital to the exploration of ecological stability of a study area but is not directly related to the hydrography of the system. For example soil and land use data, which are located in this feature dataset, are important in determining the impact of land use on a study area's environmental stability, but it is not dependent on the location of waterbodies. All features are child classes of the abstract classes HydroFeature and then EnviroFeature. EnviroFeature provides the attributes Ftype and Name to the feature classes in the Environmental Feature Dataset. The point files located in this feature dataset are ToxRelInv, Superfund, Mineral, HazSolidWaste, and 1996CleanWaterNeeds. ToxRelInv and Superfund are derived from the Toxic Release Inventory and Superfund National Priority List respectively. Mineral represents data available from the US Bureau of Mines as the Minerals Availability System/Mineral Industry Location. The environmental polygon files include Ecoregion, STATSGO, LandUseIndex, WeatherStationRange, and ManagedAreaDatabase. Figure 3.20 presents the analysis diagram of the Environmental feature dataset for clarification.

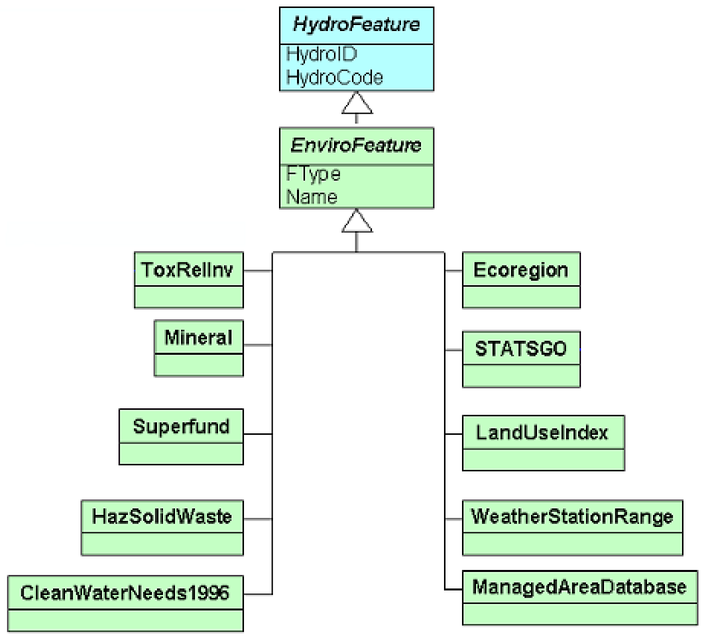


Figure 3.20: Analysis Diagram for ArcBASINS Environmental Feature Dataset

The Cartography package was created to house cartographic data, which show important political boundaries and man-made features of a study area. This dataset houses the boundaries of the county and state of the study area in the County and State feature classes respectively. UrbanArea and PopulatedPlaces are polygon and point files that inform the user what cities may be located near a study area. The roads are presented in the MajorRoad Polyline file and the administrative EPA Region is located in the EPARegion class. As usual, useful attributes are located in the HydroFeature and CartoFeature abstract classes. The data in this feature dataset are vital for placing the environmental issues into a political context spatially. Figure 3.21 shows an example of Cartography data from the Austin-Travis Lakes study area, while Figure 3.22 displays the analysis diagram for the Cartography Feature Dataset.

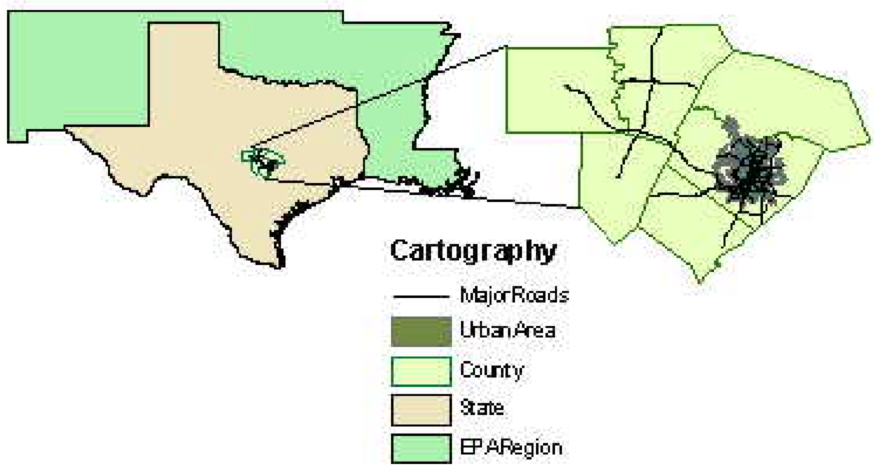


Figure 3.21: Cartography of Austin-Travis Lakes study area.

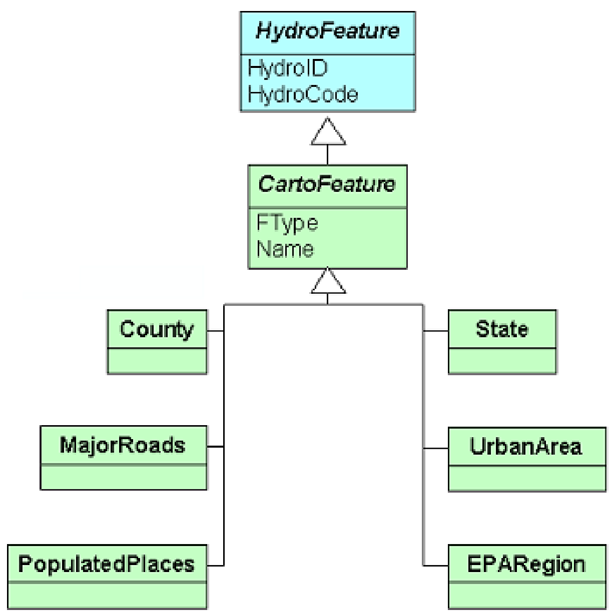


Figure 3.22: Cartography Analysis Diagram

3.4.3 BASINS Tables

In addition to the abundance of data presented by the spatial data available through BASINS, there are multiple tables containing large amounts of data. Some data is monitoring data and other data characterizes the features of the landscape. The tables add a total of nine object data packages to the model.

Four of the packages are altered from their original state in the BASINS database available off the web. These data groups contain multiple years of records with either a single year of record or a sequence of years presented in individual tables. For example the TRI (Toxic Release Inventory) data package contains data tables that each represent a single year of data for the years 1987 to 1995 for chemical air emissions, land releases, underground injections, water discharges. The data package also contains TRI data from POTW (Publicly Owned Treatment Works) sites for years 1991-1995. Finally, the TRI package has a table to explain the parameters described in its tables. This adds up to a total of 41 tables just for the TRI package. In Arc BASINS, the yearly tables for chemical air emissions, land releases, underground injections, water discharges and POTW releases are merged into one table for each topic with an added attribute, YearIndex, that designates the year of the record. This process condenses the tables in the package to a total of six tables. Yearly tables for the Bacteria Data, Permit Compliance System Computed Loading, and Water Quality Monitoring Data Summaries are also merged into one with the added YearIndex attribute. This editing of the tables increases the size of the tables but it streamlines the structure of the ArcBASINS model.

Five of the object packages are constructed in ArcBASINS in the same fashion as they are built in the BASINS data management system. The reason for this is that these tables do not contain data that is presented in a yearly fashion. Therefore, it was acceptable to implement the tables as they are described in the BASINS 2.0 Manual. The data packages, which are in their original BASINS format, are Fish & Wildlife Advisories, Lookup Tables, State Soil and Geographic Database, Water Quality Station Parameters, and the National Sediment Inventory Database. Table 3.6 clarifies the tables in the ArcBASINS model by noting the original BASINS table, describing the attributes assigned to the table in the model, and describing the purpose of the table.

Data Package	Table	Attributes Assigned	Original tables	Description
Fish & Wildlife Advisories				State reporting of locations with advisories for fishing
	Fish Wildlife Advisory Index		Fish Wildlife Advisory (1996) Index	
	Fish Wildlife Advisory Listing		Fish Wildlife Advisory (1996) Listing	
Lookup Tables				General lookup data tables
	Water Quality Criteria Table		Water Quality Criteria Table	Water quality criteria for various chemicals such as the MCL
	State Agency Code		State Agency Code	Contact information for state environmental agencies
	Standard Industrial Classification		Standard Industrial Classification	Listing of Standard Industrial Classification codes.
Water Quality Station Parameters				Observation-level water quality monitoring data parameters
	WaterQuality Obs Parameter		Water Quality Observation Parameter Table	
National Sediment Inventory Database				Sediment chemistry, tissue residue, and benthic abundance monitoring data for sediments
	NatSedInvData	Station		Abstract object class that provides attributes to child classes to facilitate relationships
	NSI Biotoxicity Data		NSI Biotoxicity Data	Percent mortality of test species

Data Package	**Table**	**Attributes Assigned**	**Original tables**	**Description**
	NSITissue ResidueData		NSI Tissue Residue Data	Tissue levels of chemical in chosen species
	NSISediment Chemistry Data		NSI Sediment Chemistry Data	Sediment's chemical features
	NSI Reference Values		NSI Reference Table	Information for sediment chemicals
	NSI Watershed Summary Data		NSI Watershed	NSI created watershed information such as number of NSI stations
Bacteria Data				Stastical summaries of bacteria monitoring
	BacteriaData	Station Parameter YearIndex	Bacteria Data 70-74, 75-79, 80-84, 85-89, 90-94, 95-97	Parameter specific statistics computed by station
	Bacteria Parameter	Parameter ID	Bacteria Parameter Table	Bacteria parameter characteristics
Water Quality Monitoring Data Summaries				Stastical summaries of water quality monitoring for physical and chemical-related parameters
	Water Quality Data	Station Parameter YearIndex	Water Quality Data 70-74, 75-79, 80-84, 85-89, 90-94, 95-97	Parameter specific statistics computed by station
	Water Quality Parameter	Parameter ID	Water Quality Parameter Table	Water quality parameter characteristics

Data Package	Table	Attributes Assigned	Original tables	Description
Permit Compliance System Computed Loading				NPDES permit-holding facility information
	Permitted Discharges	NPDES Parameter YearIndex	Permitted Discharges 1991-1996	Parameter specific loadings to surface waters computed using EPA Effluent Decision Support System
	Permitted Discharges Parameter	Parameter ID	Permitted Discharges Parameter Table	PCS chemical parameter characteristics
	PCSCode Description		Permitted Discharges Code	Description of PCS codes
State Soil and Geographic Database				National soil information
	Soil Component Data	MUID	Soil Component Data	Data for each specific soil component
	SoilLayerData	MUID	Soil Layer Data	Data for the different layers of the soil components
	StatsgoMapUnit	MUID	Table not included in BASINS database	Pulled from national Statsgo database
Toxic Release Inventory Sites				Facility information for TRI public data
	TRIData	TRIID TRIChemID YearIndex		Abstract object class that provides attributes to child classes to facilitate relationships

Data Package	Table	Attributes Assigned	Original tables	Description
	TRIAir		TRI Air Emissions Data 1987-1995	Chemical air emissions from TRI sites
	TRILand		TRI Land Release Data 1987-1995	Chemical land releases from TRI sites
	TRI Undergrd		TRI Underground Injection Data 1987-1995	Chemical underground injections from TRI sites
	TRIDischarge		TRI Water Release Data 1987-1995	Chemical water discharges from TRI sites
	TRIPOTW		TRI POTW Data 1991-1995	Chemical releases from POTWS
	TRI Parameter		TRI Parameter Data	TRI parameter characteristics

Table 3.6: Arc BASINS Table Description

3.4.4 Relationships

One of the most important parts of the BASINS database is the data provided by the monitoring and reference tables. These tables contain large amounts of information that are vital to understanding the environmental status of a study area. The database available from the EPA incorporates some mechanisms for the relation of the tables to one another as well as to spatial data. For example, one can determine which Bacteria station points are related to certain records in the Bacteria Data tables. The Bacteria Data tables are then further related to the Bacteria Parameter table, which explains which parameter is measured in a specific record of the Bacteria Data table.

To bring these associations into the formalized structure of the ArcBASINS model, relationship classes are created that mirrored the associations in the original database. A relationship class stores the associations between the different feature classes and tables in the model. The Hydrography and Environmental feature datasets contain the feature classes to which tables are related. In Hydrograph, the point classes NatSedInv, WaterQualityObs, PCSDischarge, and Bacteria are associated with a data table. Some of these tables are then related to another table, which clarifies the data even more. STATSGO and ToxRelInv are the classes in the Environmental feature dataset related to an object. ToxRelInv is associated with five tables: TRIAir, TRILand, TRIUndergrd, TRIDischarge, and TRIPOTW. Figure 3.23 below demonstrates the relationships in the model graphically with the green boxes and white boxes representing feature classes and tables respectively and Table 3.7 displays the attributes which support the relationships.

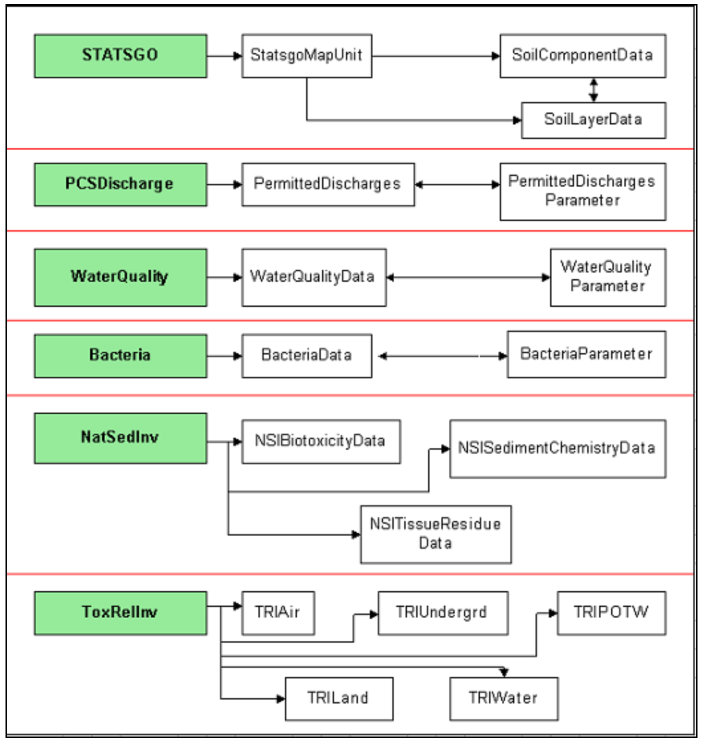


Figure 3.23: Relationships among Feature Classes and Tables in ArcBASINS

Origin Object	Attribute	Destination Object	Attribute
Bacteria	Station	BacteriaData	Station
BacteriaData	Parameter	ParameterID	Parameter ID
PCSDischarge	NPDES	PermittedDischarges	NPDES
PermittedDischarges	Parameter	PermittedDischarge Parameter	ParameterID
NatSedInv	Station	NSIBiotoxicityData	Station
NatSedInv	Station	NSITissueResidue Data	Station

NatSedInv	Station	NSISediment ChemistryData	Station
STATSGO	MUID	StatsgoMapUnit	MUID
StatsgoMapUnit	MUID	SoilComponentData	MUID
SoilComponent Data	MUID	SoilLayerData	MUID
ToxRelInv	TRIID	TRIAir	TRIID
ToxRelInv	TRIID	TRILand	TRIID
ToxRelInv	TRIID	TRIUndergrd	TRIID
ToxRelInv	TRIID	TRIWater	TRIID
ToxRelInv	TRIID	TRIPOTW	TRIID
TRIAir	TRChemID	TRIParameter	ChemID
TRILand	TRChemID	TRIParameter	ChemID
TRIUndergrd	TRChemID	TRIParameter	ChemID
TRIWater	TRChemID	TRIParameter	ChemID
TRIPOTW	TRChemID	TRIParameter	ChemID

Table 3.7: Relationship Completing Attributes

3.4.5 Summary of Changes

The ArcBASINS model creates a data structure in which the remarkable database presented by BASINS may be incorporated into the new ArcGIS geodatabase format. The ArcHydro model provides the basis for the model and four of the original data packages are retained in the new model. The addition of two new feature datasets, Cartography and Environmental, and multiple data object tables complete the ArcBASINS model. Finally, multiple relationship classes allow for straightforward associations among the many feature classes and objects of the model.

Chapter 4: Procedure of Application

The ArcGIS BASINS Hydro data model can be implemented using two methods. The first method modifies the original ArcHydro model by adding all aspects of the ArcBASINS model to the model's UML and schema. That means all tables and feature classes exist in the ArcGIS schema and its relationships are created automatically when the schema is applied. Method One is labeled as the Single Geodatabase method of the implementation for the ArcBASINS model.

The second method, the Hybrid Geodatabase method of the implementation procedure separates the spatial and tabular data available from BASINS into a geodatabase for the geospatial information and separate related database for the purely tabular data. All spatial data is modeled using UML and when the schema is applied, the BASINS feature datasets and the original tables of the ArcHydro model are created or updated. A separate database is created in MS Access, which contains the tabular BASINS data packages of the model. The tables are then linked to the geodatabase in Access, which allows for a looser coupling of the data. Finally, the relationships of the model are created individually within ArcCatalog. This implementation allows the tabular data to be updated without interfering with the geospatial information.

4.1 Implementation of BASINS Model: Single Geodatabase Method

4.1.1 Schema Creation

The schema containing the ArcBASINS data model is created in Visio 2000, a CASE Tool. The model is based on the ArcHydro data model, therefore the model must only be modified from its original structure. It is in the initial phase of the model creation that feature datasets and classes are deleted and added to the model, relationships are established, and model consistency is explored. The following sections will explain the steps involved in creating the ArcBASINS model schema.

4.1.1.1 Model Modification

To begin adaptation of the ArcHydro model, the first step is to open the Visio 2000 program file that contains the ArcHydro UML. The screen includes two windows, the UML Navigator and the model view (Figure 4.1 and Figure 4.2). The UML Navigator provides a tree view of the model and allows navigation from one static structure diagram to another. The model view contains the actual static structure of the model on multiple pages, which permits straightforward examination of the model. On the left side of the model view UML stencils are available for the different types of UML models. Only the UML Static Structure Stencils are used to create the ArcBASINS model. It is also important to note that any additions made in the model view will be reflected in the UML Navigator but deletions must be made in the Navigator in order to register in the model.

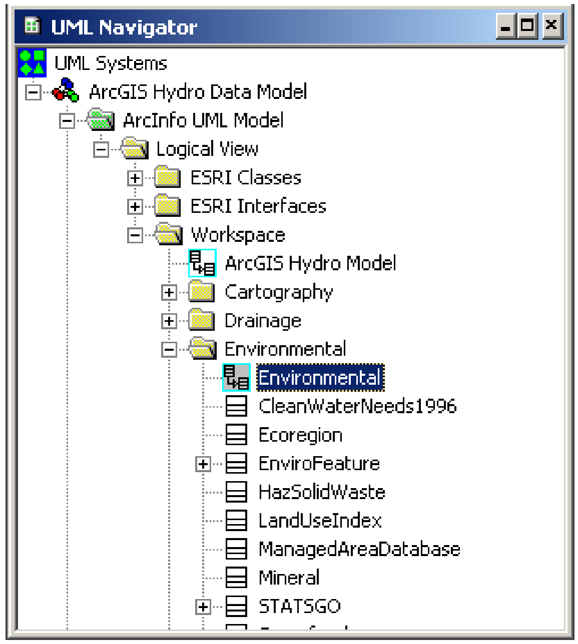


Figure 4.1: Visio 2000 Navigator View

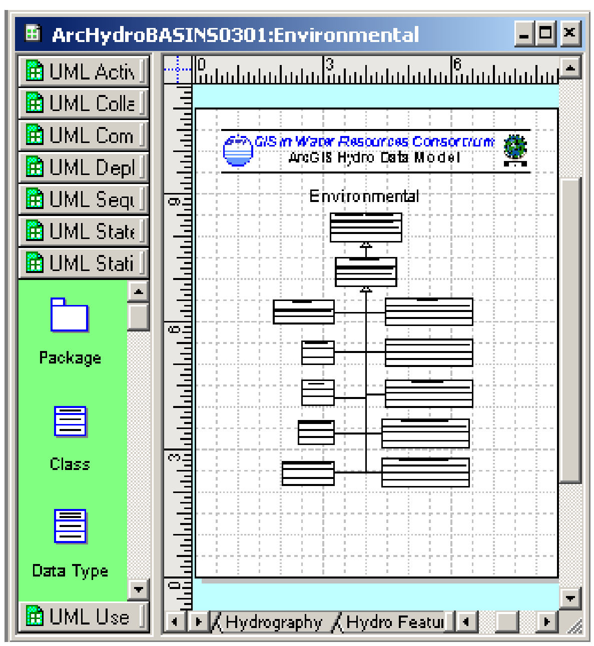


Figure 4.2: Visio 2000 Model View

The Channel data package is the first item removed from the ArcHydro model. To remove the package, right click on the package in the UML Navigator screen and select "Delete" as demonstrated in Figure 4.3. The package disappears from both screens. The process of removing elements of the model is continued until all of the features and tables listed in Table 4.1 are eliminated from the model. In addition all generalizations and related classes that are left incomplete after deleting the listed elements must be removed or corrected.

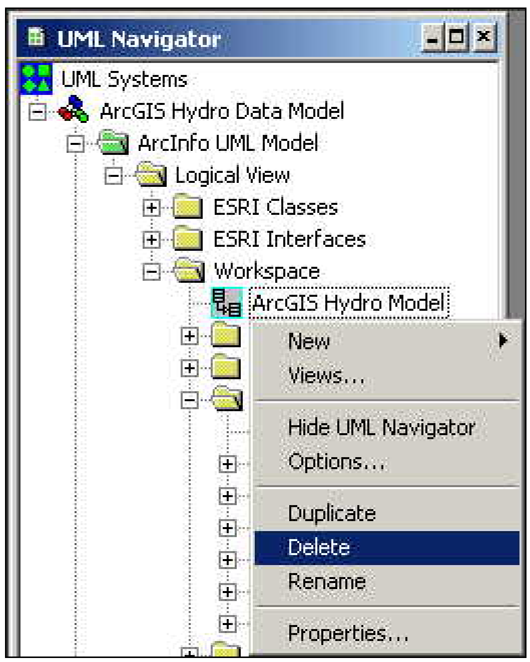


Figure 4.3: Removing Elements from the ArcHydro Model

Data Package	Feature Class or Object Removed
Channel	
	All Objects and Classes
Network	
	SchematicLink
	SchematicNode
Hydrography	
	HydroResponseUnit
	Structure
	Bridge
	WaterDischarge
	MonitoringPoint
	UserPoint

Table 4.1: Feature Classes and Objects Removed from ArcHydro to form ArcBASINS.

The next step is to add the new data packages, Environmental and Cartography, to the model. To insert a data package, select the ArcGIS Hydro Model static structure diagram and drag a package object from the UML Static Structure stencils (Figure 4.4). The package and a static structure page appear in the Navigation window. Create two packages and rename them "Environmental" and "Cartography."

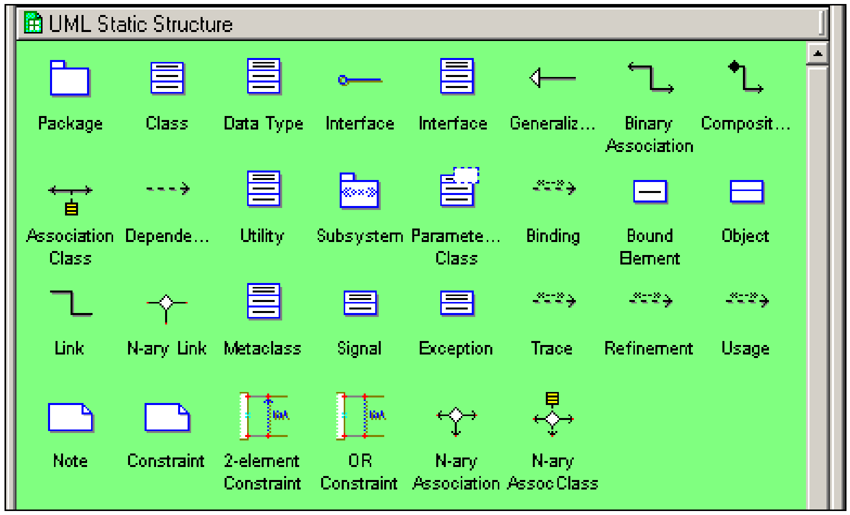


Figure 4.4: UML Static Structure Stencils

After the new data packages are included in the model, the feature classes and objects in these Feature Datasets are added. The class HydroFeature, which is present in all data packages, is placed in the new static structure diagrams by dragging the class to the new diagram from the Navigator View. HydroFeature is placed in both Cartography and Environment. New feature classes can be created either by copying an existing feature class and modifying it or by dragging a class element from the UML Static Structure stencil. An example of a class that serves as a good base class is the abstract class Hydrography. Hydrography specifies the attributes Name and Ftype, which the classes EnviroFeature and CartoFeature also create. Remember an abstract class creates common attributes shared by the classes inheriting from it but does not have objects of its own. To create EnviroFeature from Hydrography, right-click on Hydrography in the model view and select Copy. Navigate to the Environmental static structure page, right-click on the page, and choose Paste. Double click on the new version of Hydrography and in the UML Class Properties screen to change the name of the class to EnviroFeature.

Once the EnviroFeature is created, its child classes may be generated. The classes are best created by dragging the class stencil onto the static structure diagram. Once the class is created, it must be modified to match the requirements of the data model; the STATSGO class will be used to illustrate this process. Double click on the new class, changing the name to STATSGO and adding the attribute MUID. To add the attribute, navigate to the Attribute tab in the UML Class Properties screen and click "New". Type in the attribute name and specify its Type, which is esriFieldTypeString. Finally, the class's geometry must be declared. The default geometry type is point and STATSGO is a polygon. Select the Tagged Values tab and click "New", setting the Tag as GeometryType and Value as esriGeometryPolygon. Figure 4.5 that chronicles the process of creating a new feature class followed by Figure 4.6 displays the completed STATSGO class.

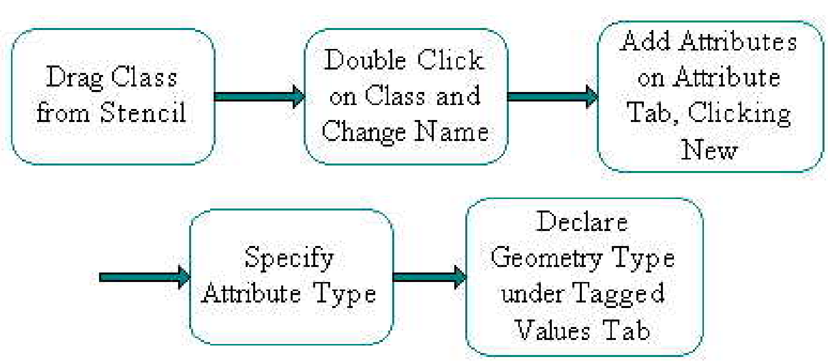


Figure 4.5: The Process for Creating a New Feature Class in Visio 2000

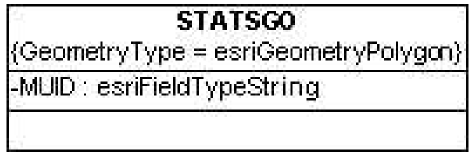


Figure 4.6: Complete STATSGO Feature Class

The Environmental and Cartography datasets are nearly complete when all required classes have been created. To complete the data packages, insert generalizations which inform the classes how they are related to one another.

Hydrography is the final data package where feature classes must be included. A number of classes were deleted previously and now their replacements must be inserted. Following the formula of the Arc BASINS model described in Chapter 3, the classes are created as described for the STATSGO feature class. To finalize the insertion of the new classes, all generalizations are updated in the Feature Dataset.

To complete the modification of the model, the tables described in Chapter 3 are included in the model. Tables are objects in the UML diagram. Create a new static structure page under the Objects data package and rename it as BASINS Tables 1, since two pages will be needed to contain all of the tables provided by the model. Drag the Object class from the ESRI Classes package on to the diagram. Figure 4.7 shows the location of the ESRI Object class within the UML Navigator frame.

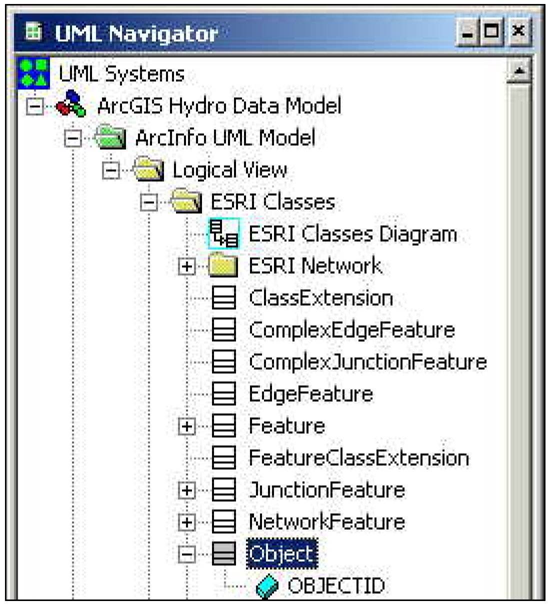


Figure 4.7: ESRI Object Location

All tables will inherit the ObjectID attribute from Object. Subsequently the tables are created from the Class stencil by dragging the stencil from the left side of the model view onto the static structure diagram. The object's names and attributes are modified to fit the ArcBASINS model. Figures 4.8 and 4.9 show the process of creating tables for ArcBASINS and the completed UML diagram for the Bacteria Data package respectively.

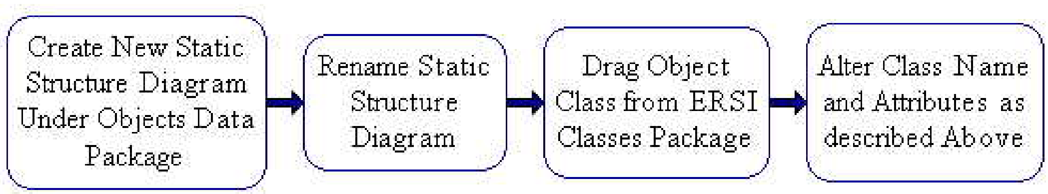


Figure 4.8: Creating Tables for ArcBASINS.

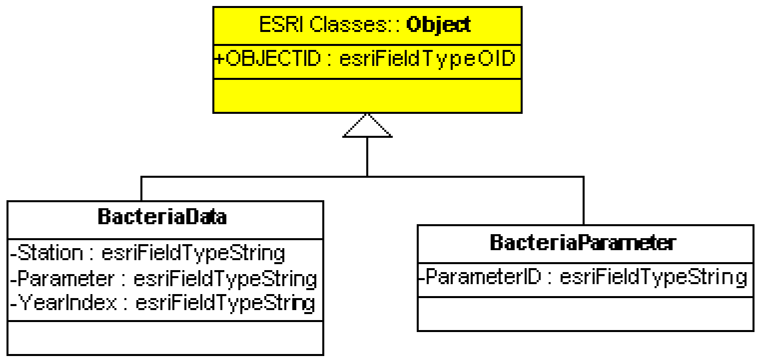


Figure 4.9: Completed Bacteria Data Package

4.1.1.2 Creating Relationships

The relationships established in the ArcBASINS model provide the integration between the Feature Classes and Tables. Relationships are created within the Relationship data package using existing classes and objects. A new static structure diagram for the relationships is added to the model with the tables and classes that are involved in the desired relationship. The binary association class stencil is used to connect the elements of the model. The overall procedure for creating relationships in the ArcBASINS schema is presented in Figure 4.10.

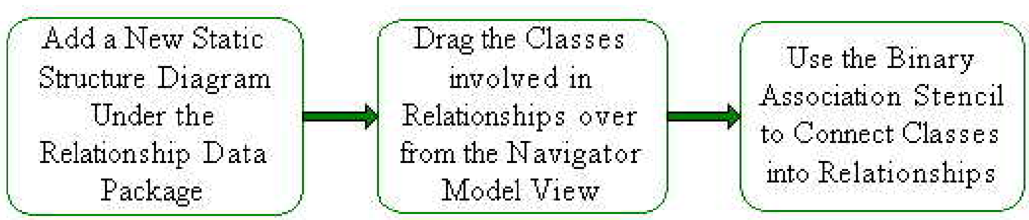


Figure 4.10: The Process for Creation Relationships in ArcBASINS

As an example, the process of the creating the relationships associated with the Bacteria Data package is explained. The first step is to drag the Bacteria feature class, the BacteriaData table, and BacteriaParameter table onto the static structure diagram. Two binary associations are used to connect Bacteria to BacteriaData and BacteriaData to BacteriaParameter, one for each relationship. Figure 4.11 provides an example of a binary association.

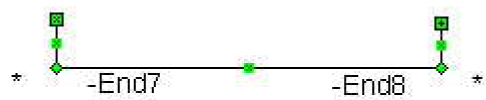


Figure 4.11: A Binary Association in Visio 2000

Double-clicking on the association between Bacteria and BacteriaData the UML Associations Properties screen appears. The name of the association is changed to BacteriaHasData. End 7, which is the starting end of the relationship, is changed to Bacteria and given the multiplicity of 1. End 8 has the multiplicity of many and the name BacteriaData. This defines the relationship that one entry in the Bacteria feature class can be related to many BacteriaData entries. Figure 4.12 displays the screen where this information is specified.

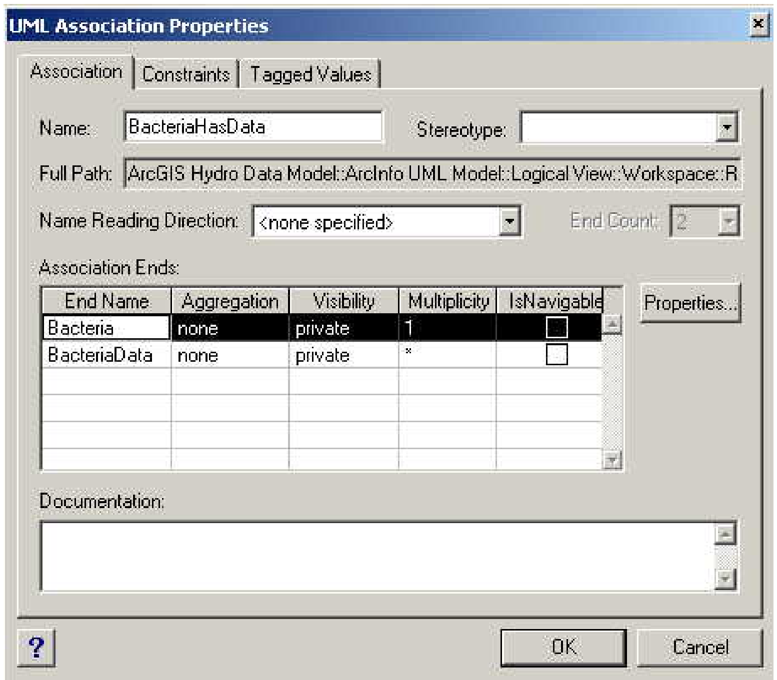


Figure 4.12: Association Screen of UML Association Properties

To finalize the relationship the attributes, which define the relationship, must be specified. Navigate to the Tagged Values tab and add two new Tagged Values. One has the tag name OriginForeignKey with a value FeatureID, and the other the tag name OriginPrimaryKey with the same value. Figure 4.13 displays the Tagged Values Screen. The OriginPrimaryKey and OriginForeignKey tagged values provide the defining attribute for Bacteria and BacteriaData respectively. The process is repeated for all relationships in the model. Figure 4.14 shows the UML for the relationships in the Bacteria Data package.

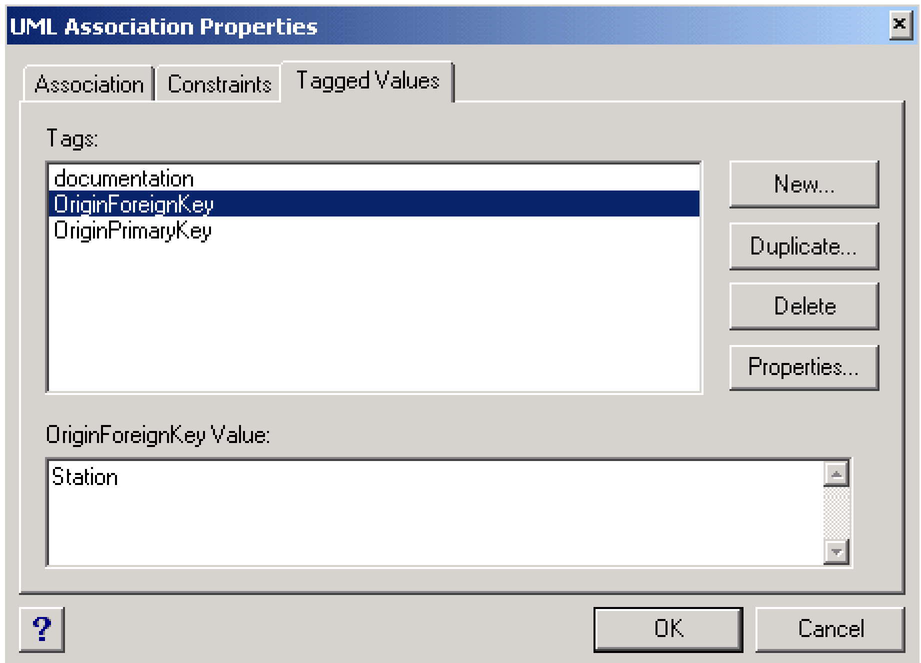


Figure 4.13: Creating Defining Attributes for Relationships

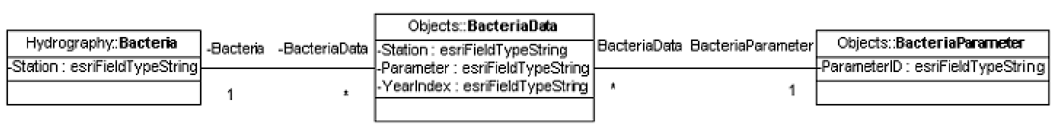


Figure 4.14: Relationships in the Bacteria Data Package

4.1.1.3 Exporting the Model

Once the model modifications and relationships are complete, the model repository containing the data model requirements may be exported from Visio 2000. To export the model, navigate to UML => Repository => Export, which causes the Connect to MS Repository window to appear. In the new window declare the name of the repository and where it should be stored, then click OK.

The Repository Export will require a model to be selected. Only one model, the ArcGIS Hydro Data model, is available for selection. Once the model is selected, the UML Add-on window appears and the model is exported. The Add-on window contains a progress bar. When the repository export is complete, the program screen again contains only the UML Navigator and model screens.

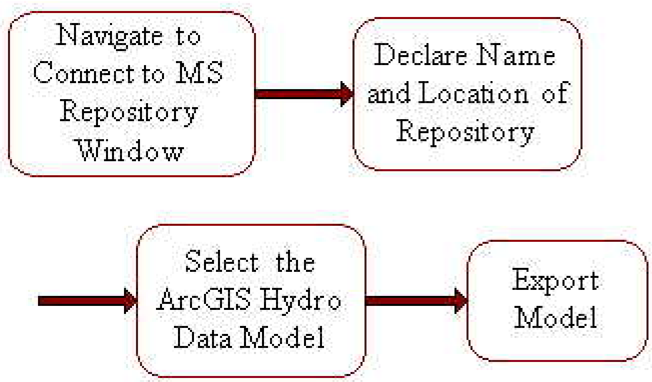


Figure 4.15: Exporting the Data Model

4.1.1.4 Checking the Semantics

The final step in creating the schema is to check if the model semantics are correct. An ESRI Macro folder can be added to Visio 2000 that contains a semantics checker, which is designed to test if there are any model inconsistencies with the ArcGIS data model format. Figure 4.16 illustrates where the Semantic Checker can be found in Visio.

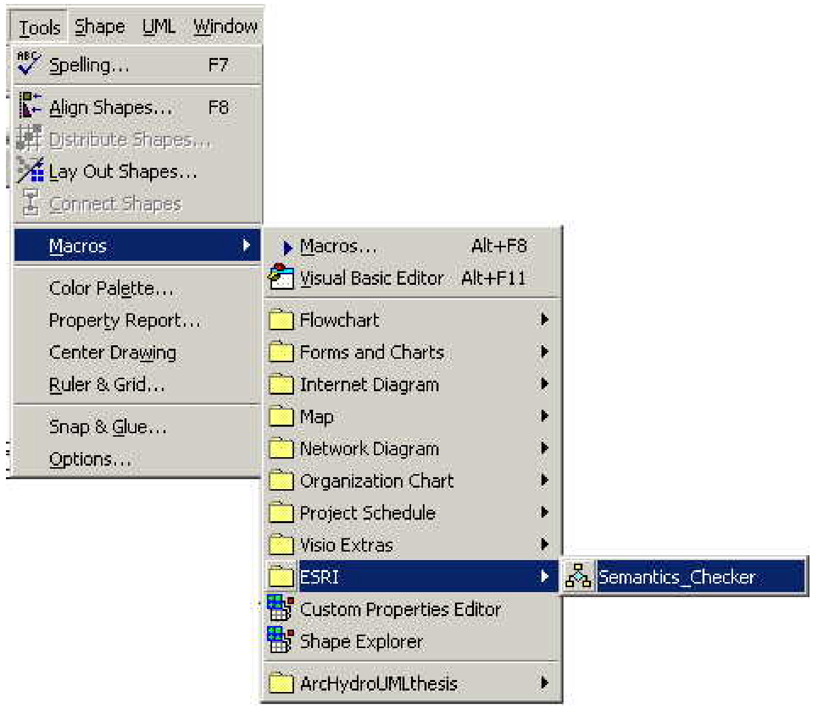


Figure 4.16: Semantics_Checker Location

Once the Semantics Checker is selected, browse to the location of the repository and list the models in the repository. Choose the ArcGIS Hydro Data Model and press the Check button. If an error is found, a screen such as the one displayed in Figure 4.17 is displayed. If no errors are found, a small screen appears declaring this fact and the model is ready for application.

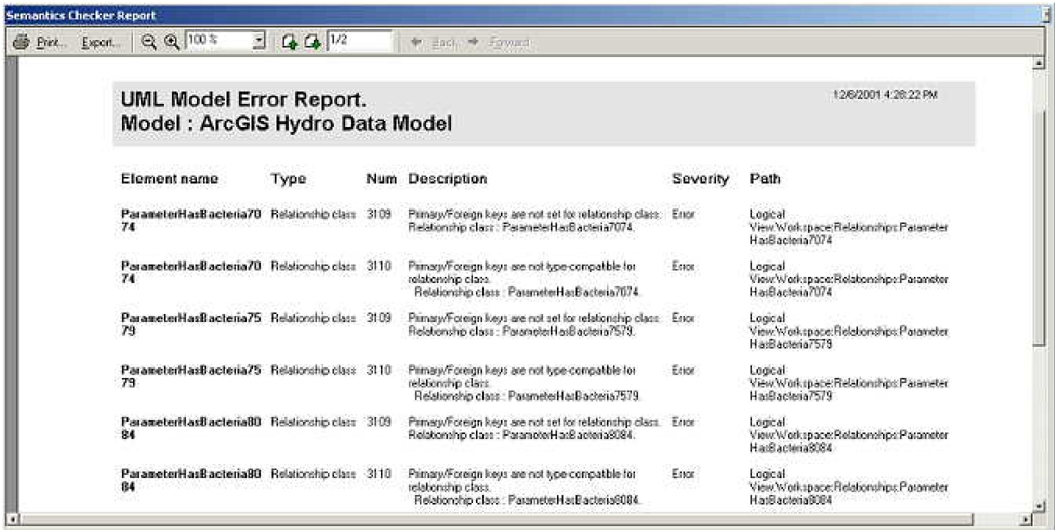


Figure 4.17: Semantics Checker Report

4.1.2 Geodatabase Creation

The creation of the geodatabase requires two steps. The first is to preprocess the data so that it conforms to the format required by the data model. This eases the implementation of the model since some of the data from the data BASINS database are in a format that is a bit removed from the Arc BASINS model format. Once the data are in the correct format, the second step is to load it into a geodatabase to which the model schema will be applied.

4.1.2.1 Data Preprocessing

The only data that must be preprocessed from their format in the BASINS database are the data tables. In Chapter 3, the available data tables and their BASINS source are explained. No data management is necessary for the tables in the Fish & Wildlife Advisories, Lookup Tables, Water Quality Station

Parameters, National Sediment Inventory Database and State Soil and Geographic Database packages. They may simply be imported into the geodatabase.

However, in the Bacteria Data, Water Quality Data Summaries, Permit Compliance System Computed Loading, and Toxic Inventory Release Sites data packages the format of some of their input sources must be altered. As mentioned in Chapter 3, the records in these packages are available from the BASINS database as yearly or multiyear tables. To simplify of the model these yearly tables are combined into one table for their respected themes. The combination of the Bacteria Data tables for 1970-1997 into one Bacteria Data table is shown as an example of this process.

Two programs are necessary to manipulate the tables, i.e. Microsoft Access and Microsoft Excel. Access is used to merge the tables together, as well as, to add the YearIndex attribute. The YearIndex attribute allows the user to identify the original table from which a record was taken. Excel is used to speed up the process of filling in the values for the YearIndex attribute. This attribute may be modified in Access but each record must be changed individually; this is quite inefficient when the table contains over 25 records, which all tables do.

To begin, open Access and select the "New" button. A screen will appear where the database will be created. A name is give to the database and Create is selected. The Bacteria database is called Bacteria. Once the new database is created, the original tables must be added. This is done by clicking File => Get External Data => Import. Navigate to the folder, which contains the files for the

Bacteria Monitoring Stations and Data Summaries, and select here the first file bc_d7074.dbf as illustrated in Figure 4.18.

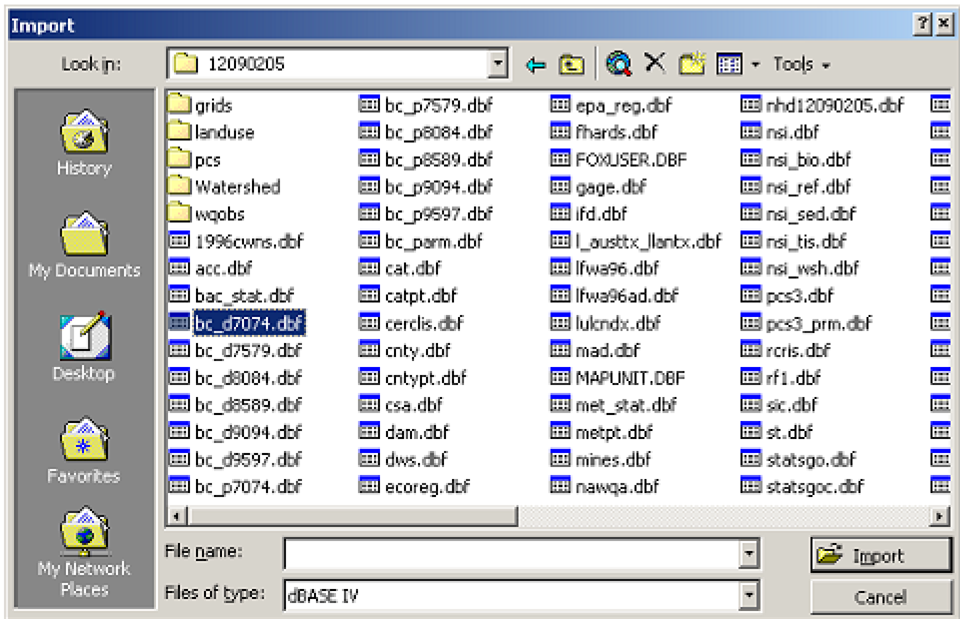


Figure 4.18: Selecting Import Tables in MS Access

Press Import and repeat this process for the remaining five bc_dxxxx.dbf files. Close the import window and open the bc_d7074 tables. The next step is to add the YearIndex attribute. Do this by selecting Insert => Column and name the column YearIndex. Place the value 7074, which represents the years recorded in the table, in the first record of the table.

Now the table is exported to Excel to ease data entry. Select Tools => Office Links => Analyze It with MS Excel. This opens up the table in MS Excel, where the value of 7074 for the attribute YearIndex is copied into each of the records. Save the table edited in Excel in the dBASE IV format. Repeat this procedure for the remaining five tables in the database. Now return to MS Access and delete the tables that are located in the database. The current tables do not contain the changes that were made in Excel. Therefore, the original tables will be deleted and the updated tables will be imported. Once all tables are deleted, follow the table import procedure described above until all updated tables are loaded into the database.

The final step in the table preparation is to merge the six tables into one. Right-click on the bc_d7074 table and select Copy. Then right-click on the white region of the database management screen and select Paste. The following screen will appear.

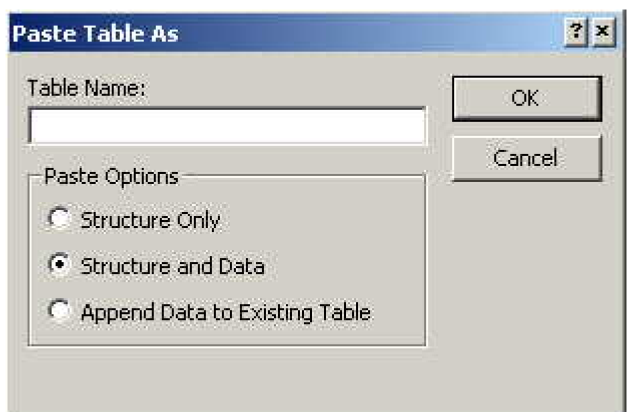


Figure 4.19: Pasting Tables in MS Access

Use the default toggle, Structure and Data, name the new table Bacteria, and click OK. This creates a new table Bacteria where all tables are merged. Now, copy the bc_d7579 table and right-click to paste the table. This time the Append Data to Existing Table toggle is selected and the table name Bacteria is specified again.

Click OK. Open the Bacteria table and scroll down to the bottom of the table. It is apparent that the information from bc_d7579 has been added to the table. Repeat this process for the remaining five tables. The final step is to export the Bacteria table into the dBASE IV format so that it may be imported in the geodatabase. Right-click on the Bacteria table and select Export. Choose the dBASE format and press Save. The Bacteria tables have been merged into one with the additional attribute YearIndex; the table is ready to be imported on the BASINS geodatabase.

4.1.2.2 Creating the Geodatabase

The creation of the model geodatabase is the second step in the ArcBASINS data model process. Here the shapefiles that are in the BASINS database are imported into the ArcCatalog geodatabase format, which is the basis for the application of the model. The data are loaded into feature classes and feature datasets that have the same name as required by the ArcBASINS data model.

To start creating the geodatabase, open ArcCatalog from the Start Menu and navigate to the folder in which the geodatabase will be located. Right-click on the folder and select New => Personal Geodatabase. A new geodatabase will be created and the next step is to rename the geodatabase. Once the geodatabase is created, data may be entered; this is usually done feature dataset by feature dataset. The first feature dataset that is created is Hydrography. Right-click on the geodatabase you created and click Import => Shapefile to Geodatabase. This tool facilitates the importation of all shapefiles into the geodatabase. Navigate to the input shapefile for the Weather feature class, which is metpt.shp. The first shapefile is imported into the geodatabase for a feature dataset; it establishes the spatial extent of the feature dataset. Therefore, the shapefile with the largest spatial extent should be imported first to allow all feature classes to exist in the feature dataset. The Weather class of Hydrography has the greatest spatial extent for the Austin-Travis Lakes study area. Once the shapefile is selected, the name of the class is specified as Weather and the name of the feature dataset is written in the "Select an existing feature dataset or enter a new one" dialog box. Figure 4.20 below displays the shapefile dialog box.

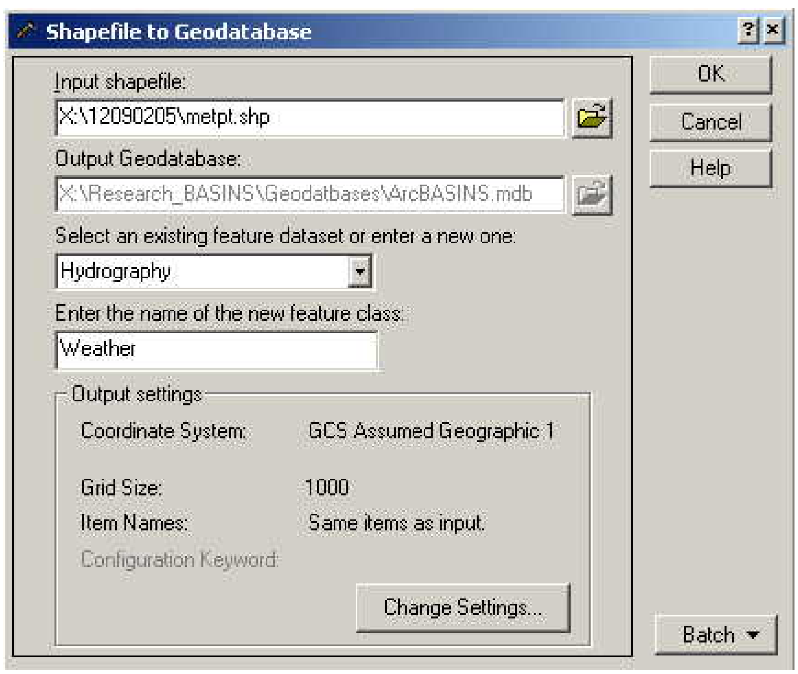


Figure 4.20: Inputting Shapefiles into a Geodatabase

Click OK and repeat this process for the remaining feature datasets and classes in the ArcBASINS data model. It is important to note that, once a feature dataset is created, it may be selected from the appropriate dialog box. Appendix C contains information on which BASINS shapefiles are used to create to each ArcBASINS feature class.

The creation of the geometric network, located in the Network feature dataset, is accomplished by running the Geometric Network Wizard, which goes through the process of creating a network step by step. To initiate the wizard, right-click on the Network dataset and select New => Geometric Network. Follow the instructions for each step, naming the network as HydroNetwork and selecting HydroEdge and HydroJunctions as the features from which to create the network. Also, note that the edges must be specified as complex edges since the data model calls for a network with complex edges. None of the features need to be snapped, but the HydroJunctions must be indicated as sinks in the network, otherwise flow may not be routed through the system. Finally, no weights will be put on the network. Once the wizard is run, Network has two new features: HydroNetwork and HydroNetwork_Junctions. The HydroNetwork_Junctions may be empty but must be present for the network to exist.

Importing tables into the data model is very similar to importing shapefiles. To import the tables right-click on the geodatabase and click Import => Table to Geodatabase. The interface is similar to the Shapefile to geodatabase interface. For example to retrieve the TRIAir table, navigate to the table in the BASINS database source folder and select it. Give the table the output name of

TRIAir as specified by the model; the dialog box looks like Figure 4.21. Click OK.

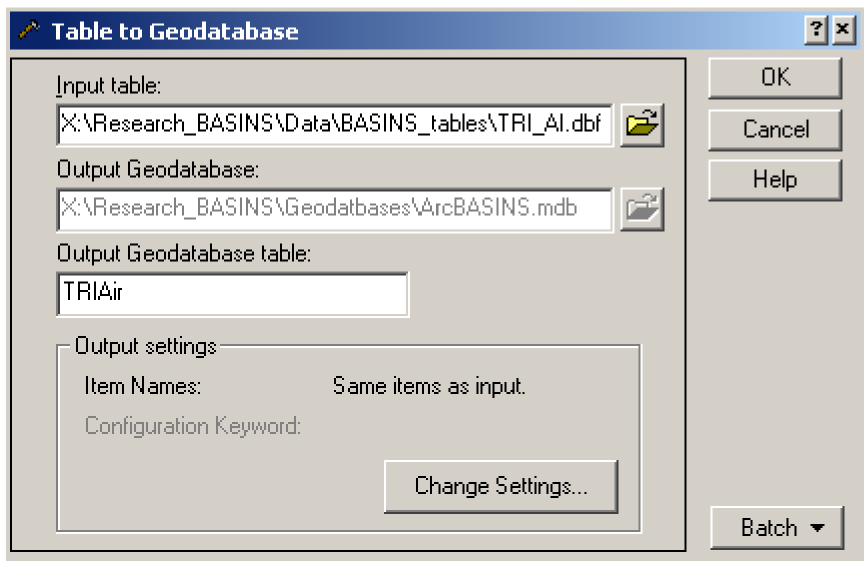


Figure 4.21: Importing Tables in ArcCatalog

This procedure is followed for all the tables in the data model. Once all tables and feature classes have been imported, the geodatabase is completed; Figure 4.22 shows a view in ArcCatalog of the completed ArcBASINS geodatabase before schema application.

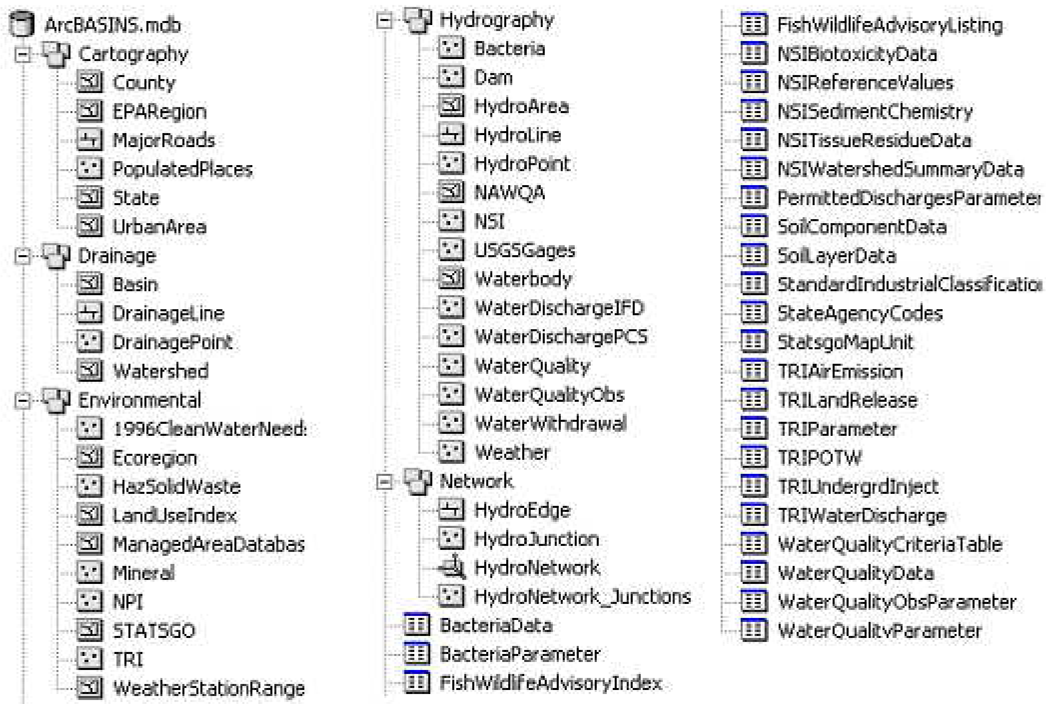


Figure 4.22: The ArcBASINS Single Geodatabase Method before Schema Application

4.1.3 Application of the ArcGIS BASINS Hydro Data Model

Now that the geodatabase is completed, it is ready for application of ArcBASINS data model. This phase of implementation is a two-step process that calls for schema application and a final stage of data editing to allow for the instantiation or invocation of the relationships created by the model. Application of the data model is accomplished in ArcCatalog, while the data is edited in ArcMap.

4.1.3.1 Applying the Schema

To apply the schema, another wizard, like as the Geometric Network Wizard, must be used. The Case Schema Wizard is a tool that usually must be added to the toolbar the first time that it is used. This is accomplished by selecting the Customize option under the Tools menu. Under the Commands tab select Case Tools, and drag the Schema Wizard command to the toolbar as illustrated in Figure 4.23.

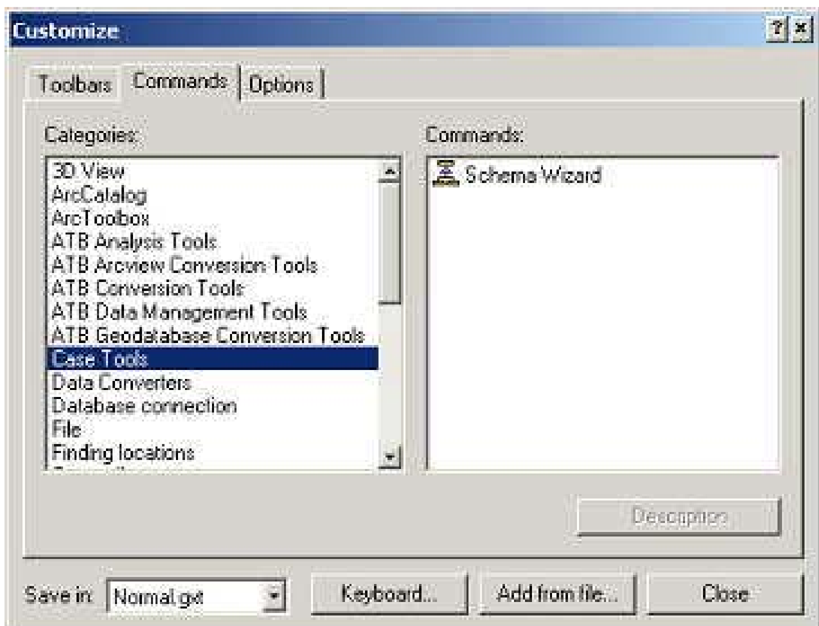


Figure 4.23: Tools Customize Window

To activate the Schema Wizard, select the geodatabase containing the preformatted BASINS data and press the Schema Wizard button. The wizard will appear. Click Next and on the following screen navigate to the folder that contains the repository file that was created in Visio. Figure 4.24 illustrates this view.

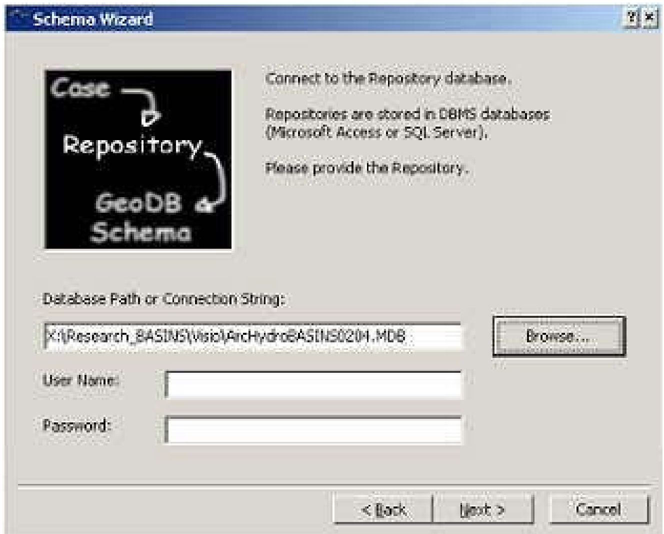


Figure 4.24: Selecting the Repository in the Schema Wizard

The next step is to select the object model; there is only one option, ArcGIS Hydro Data Model : : ArcInfo UML Model. The repository is loaded and the screen shown in Figure 4.25 appears.

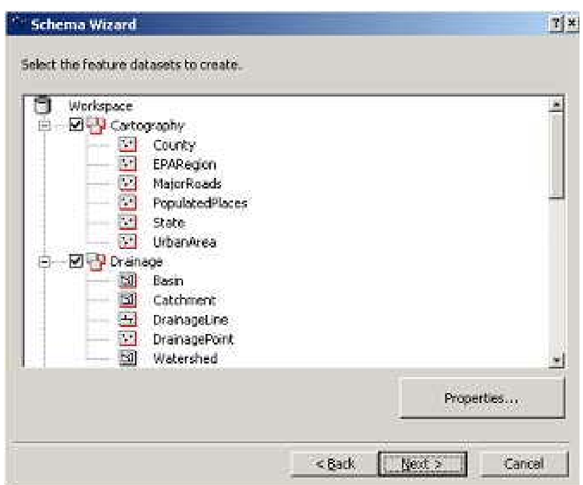


Figure 4.25: Feature and attribute creation

The feature classes outlined in red are classes that already exist in the geodatabase and have the same name as the feature classes in ArcBASINS model. Each class is selected and the Properties button is clicked to determine if all necessary attributes have been created. The attributes of a class may be examined under the Exists tab of the Properties dialog box, as displayed for Bacteria in Figure 4.26.

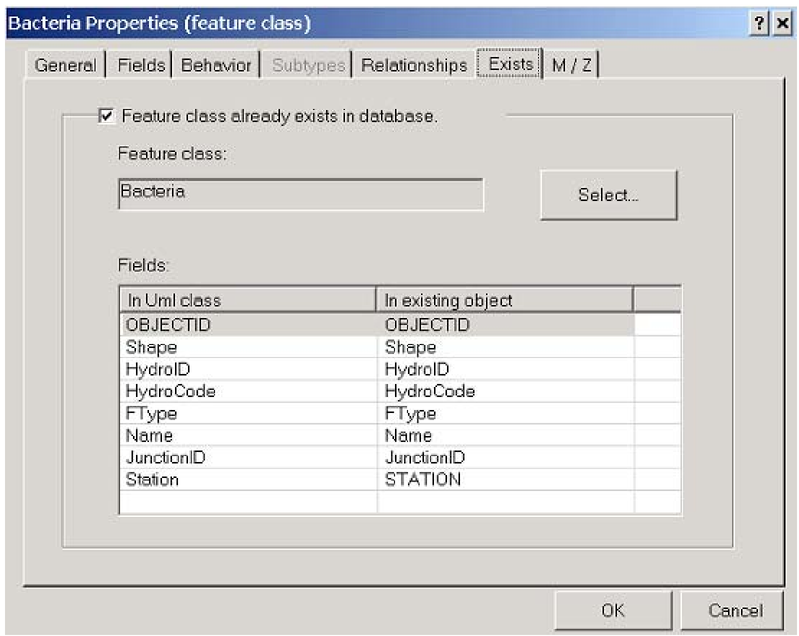


Figure 4.26: Exists Tab

If this is not done, the Wizard will not allow the process to proceed to the next step and a warning will appear, detailing the class, which must be inspected. The final screen that will appear is a summary screen, as shown in Figure 4.27. The summary screen details the elements of the model created.

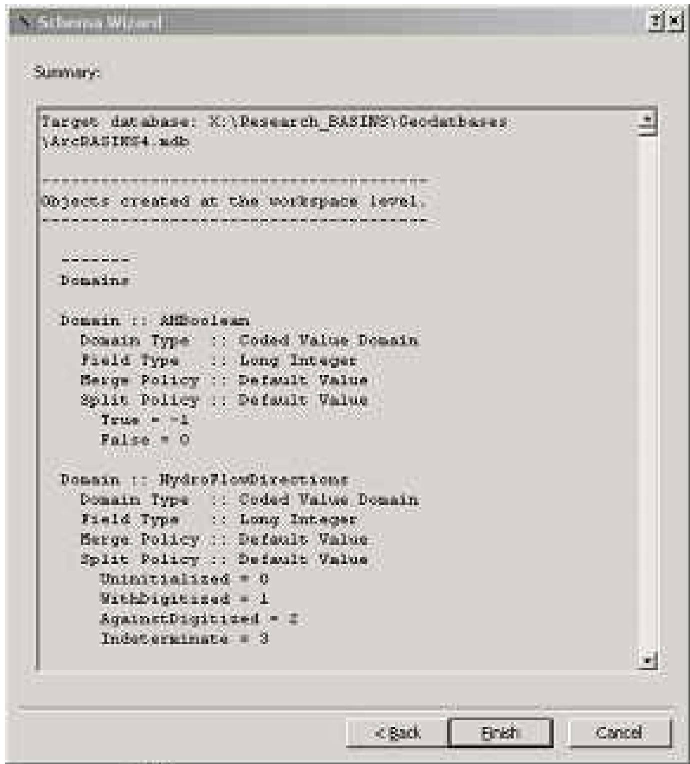


Figure 4.27: Schema Summary Screen

The schema now has been applied and only some data entry is needed to finalize the implementation of the ArcBASINS model.

4.1.3.2 Completing the Relationships

After applying the model, the final step in the implementation of the ArcBASINS model is to fill in the attributes of the features and tables created by the schema application. The new attributes such as HydroID and ParameterID contain a null value, which does not allow for the instantiation of the relationships created by the model. This makes sense since an object must know how it is related to another object by its common attribute. One method to fill the needed attributes is to apply the ArcHydro Tools, which were created at the Center for Research in Water Resources (CRWR) and ESRI to automate this process. As another option simple data management can be used to complete the Tables.

The application of the ArcHydro Tools is completed in ArcMap. To start the process, ArcMap is opened and the completed geodatabase is added to the Map document. The ArcHydro Tools are available from the CRWR website (www.crwr.utexas.edu) with instructions for installation. Once the ArcHydro Tools are installed, they can be added to a Map document like any toolbar by navigating to the Tools button and then to Customize. The ArcHydro Tools menu contains multiple tools, but for the initial application of the ArcBASINS model only AssignHydroID is of interest. The AssignHydroID assigns a value to each HydroID attribute in the model. To apply the AssignHydroID tool, the document must be in edit mode, which is entered by clicking on Edit => Start Editing. Then simply select the AssignHydroID button. The program will calculate a HydroID. The calculation may take some time since the ArcBASINS dataset is large. However, once the status bar has disappeared from the lower left-hand corner of the Map document, the calculation is complete.

The other attributes that must be entered in the model are the relationship instantiating attributes such as BacteriaData's Parameter. This data is already contained in the table but under the original name PARAMETR. The process of filling the Parameter attribute is quite simple because the values of PARAMETR are copied into the column for Parameter. To copy data in an ArcMap table, the document must again be in editing mode. The Parameter field is selected by right-clicking on the column heading and the Calculate Values option is chosen. In the Calculate Values dialog box the PARAMETR field is selected to equal Parameter, as shown in Figure 4.28, and OK is selected. Save the edits to make sure that no data calculations are lost. This process of data transfer is repeated until all relationship attributes are filled.

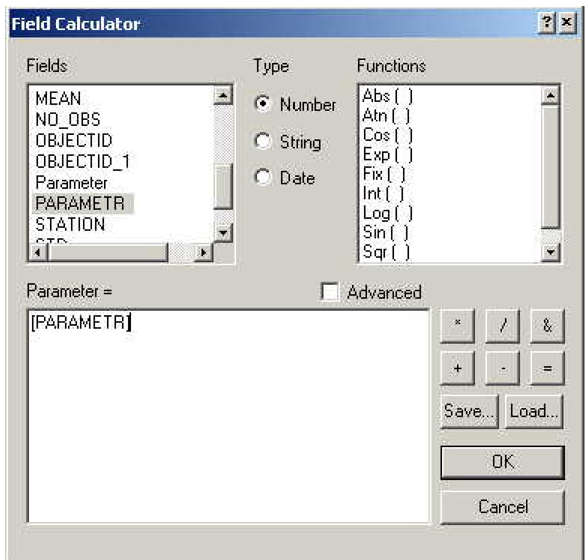


Figure 4.28: ArcMap's Field Calculator

4.2 Implementation of BASINS Model: Hybrid Geodatabase Method

4.2.1 Schema Creation

Creating the schema to be used in the Hybrid Geodatabase method of Implementation of the BASINS model is very similar to the schema creation for the Single Geodatabase. The feature datasets Hydrography, Network, Drainage, Environmental, and Cartography are created as described in section 4.1.1.

Drainage is not altered from its original state in the ArcHydro model and only Schematic Node and Schematic Link are removed from the Network data package. Hydrography is amended significantly from its form in ArcHydro by deleting some existing feature classes and adding new classes. The new data packages Environmental and Cartography are inserted as described in the Model Modification section; the objects in the Time Series group are left untouched.

The most important and noteworthy change from the Single Geodatabase method is the removal of BASINS tables from the schema. Since all BASINS tables will be included in a separate Access database, they do not have to be included within the ArcBASINS schema. The elimination of the tables streamlines the schema. Only feature classes and the important Times Series objects are included in the schema. Since the tables are removed from the schema, all relationship classes based on the BASINS database are eliminated. However, they are recreated in ArcCatalog, which allows the important relationships between spatial and tabular data to continue. Figure 4.29 illustrates the major differences between the implementation versions in regard to the Object and Relationships data packages.

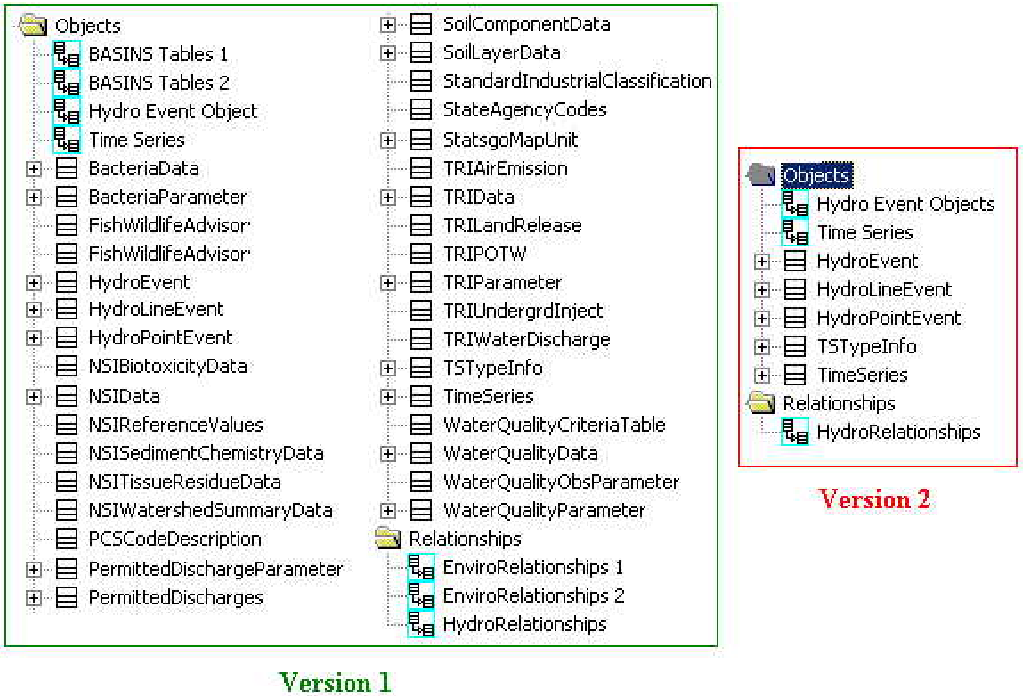


Figure 4.29: Single Geodatabase and Hybrid Geodatabase's Object and Relationship Data Packages.

4.2.2 Database Creation

The Hybrid Geodatabase method requires two databases, a geodatabase and an Access database. The geodatabase contains all spatial data and the Access database brings together the tables of the model. This method allows for a more robust manipulation of the BASINS tables within the Access framework.

4.2.2.1 Geodatabase Creation

The creation of the BASINS geodatabase for the Hybrid Geodatabase method follows the same process as described for Single Geodatabase method. Begin by opening ArcCatalog and creating a new geodatabase. Create each feature dataset separately, using the Shapefile to Geodatabase interface. It is important to remember to first enter the shapefile that has the largest spatial extent, since it creates the spatial extent of the feature dataset. More than one shapefile associated with a feature dataset may be loaded at a time in the Shapefile to Geodatabase, which can save time in preparing the database. Finally, once all feature datasets are created, the geometric network must be generated. Figure 4.30 presents the BASINS geodatabase for Hybrid Geodatabase Method.

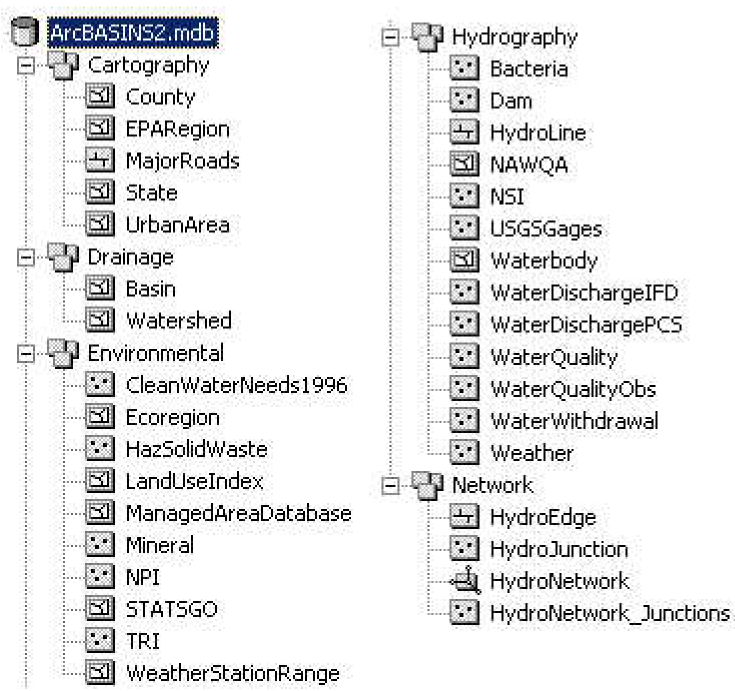


Figure 4.30: The ArcBASINS Hybrid Geodatabase before Schema Application

4.2.2.2 BASINS Table Database Creation

The BASINS Table Access database contains both preprocessed data and pure BASINS data. As mentioned previously, the tables in the Fish & Wildlife

Advisories, Lookup Tables, Water Quality Station Parameters, National Sediment Inventory Database and State Soil and Geographic Database packages are imported in their original BASINS format into the new database. The Bacteria Data, Water Quality Data Summaries, Permit Compliance System Computed Loading, and Toxic Inventory Release Sites data packages are reformatted so that their multiyear tables are combined into one single table per package as described in the data model. The new tables are created in Access and Excel, using the procedure outlined in the Data Preprocessing section of this chapter. Once the tables have been modified, they may be imported into a new Access database which will contains the BASINS tables.

This first step in creating the BASINS table database is to create a new Access database, which will be called BASINStable. Import the preprocessed tables by selecting File => Get External Data => Import and displaying Microsoft Access tables in the "Files of type" dialog. Since each data package table was created in its own Access database, the importation process will be the same. For example, select the database in which BacteriaData table was created and select the merged table. If necessary, rename the table as BacteriaData. Repeat this process for all preprocessed tables until all have been entered into the BASINS Access database. The remaining tables are also imported using the Import window, however, the "Files of type" dialog is set to dBASE IV. Once the tables have been imported an- attribute OID, an internal ArcGIS attribute, must be added to the BASINS' tables. If this attribute is not added, ArcCatalog will not recognize the table and relationships may not be made between the tables and feature classes. Also, all table attributes as specified by model must be added to the tables or the table relationships may not be created as directed. Finally, rename the tables so that they match the nomenclature of the data model, and the BASINS Access table should appear as shown in Figure 4.31. It is important to note that the data available for the Austin Travis Lakes study area does not contain a complete Permit Compliance System Computed Loading data package and was therefore emitted from the database.

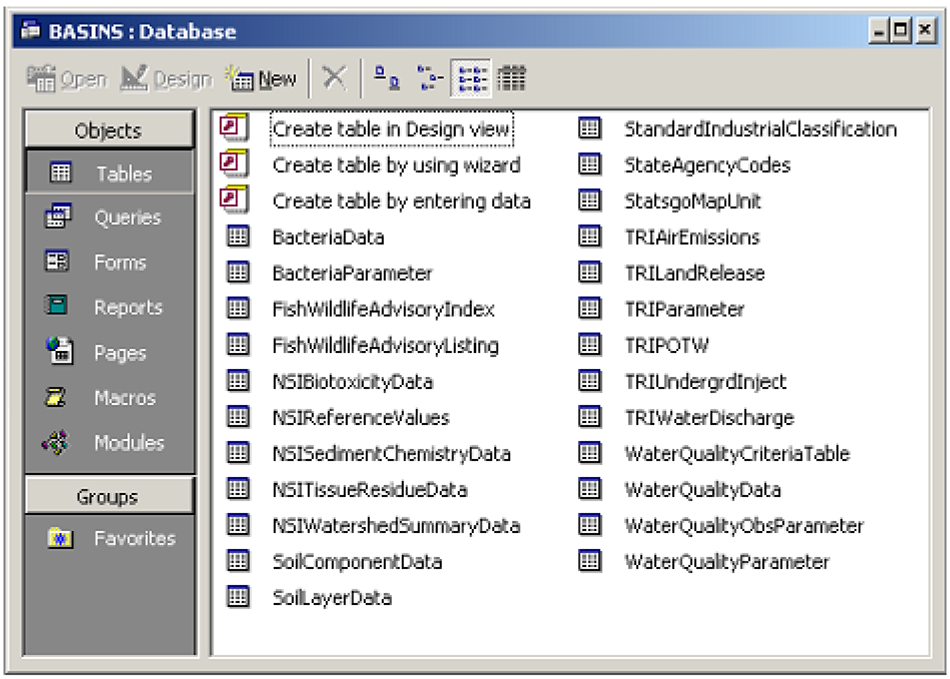


Figure 4.31: The BASINStable Access Database for Austin Travis Lakes Study Area.

After all tables have been added to the database the relationships between the tables are created. To start creating relationships select Tools => Relationships and a Relationship window. Once in this window the needed tables are added by selecting Relationship => Show Table. In the Show Table dialog select the desired tables, as shown in Figure 4.32, and press Add.

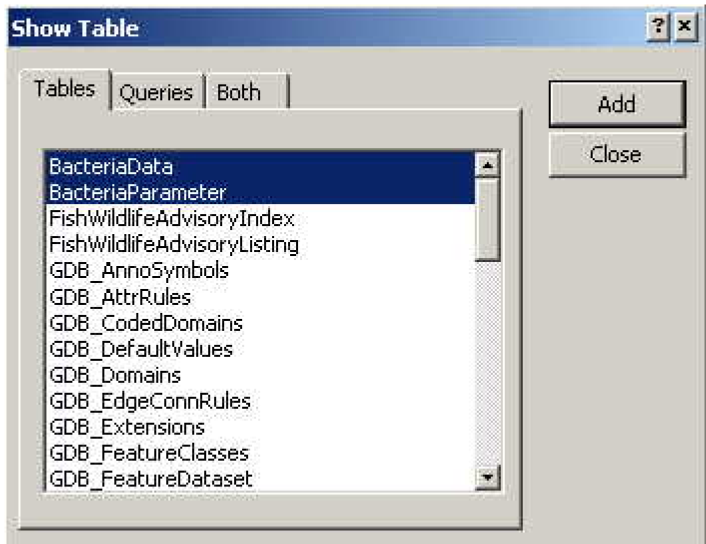


Figure 4.32: Adding Tables to the Relationship Page in Access.

Once the tables are added to the Relationship page, the relationship is established by selecting the attributes on which the relationship is based. For the example for the relationship between Bacteria Data and BacteriaParameter, the attributes Parameter and Parameter ID are selected respectively. It is important to first select Parameter followed by ParameterID to insure the relationship is created correctly. When both attributes are selected the Edit Relationships screen will appear as demonstrated by Figure 4.33. At this time it is possible to change the attributes to be used for the relationship if an error was made earlier. Once, the attributes have been verified press the Create button and a new relationship is created. One important note about relationships in Access is that there can be more than one relationship between two tables if necessary. The Relationship screen will appear like Figure 4.34 when all table relationships are created.

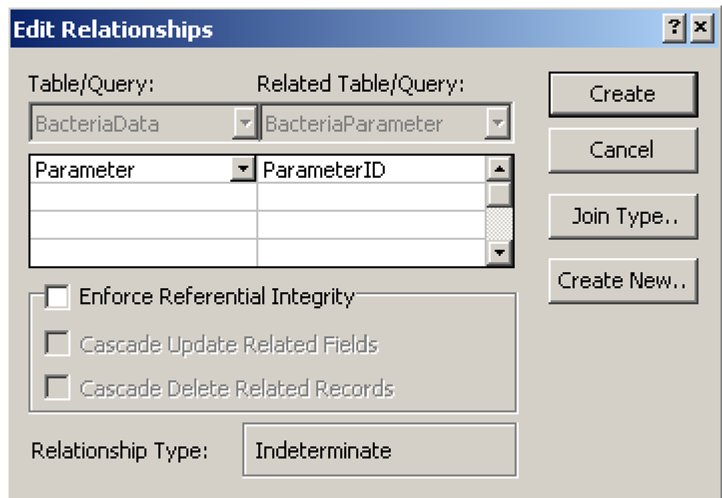


Figure 4.33: Edit Relationships Screen in Access.

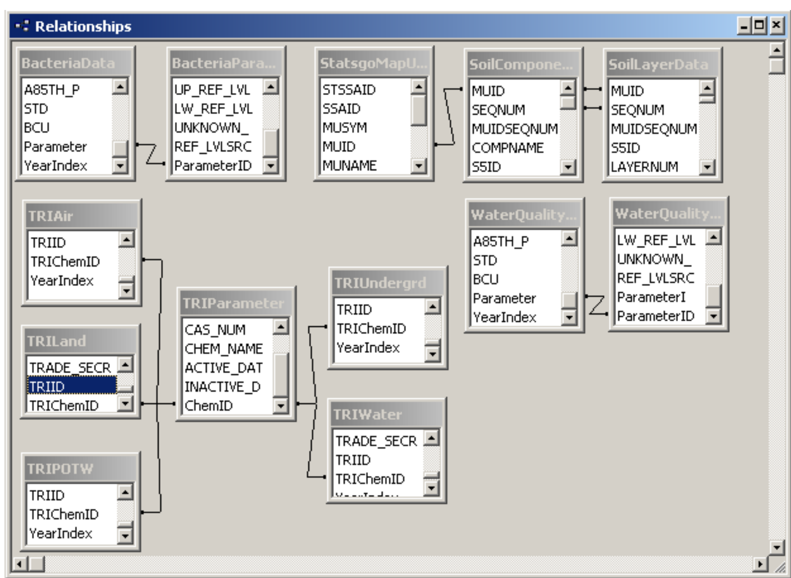


Figure 4.34: Relationships within BASINStables database.

4.2.3 Applying the Arc BASINS Hydro Data Model

The application of the ArcBASINS model for Hybrid Geodatabase implementation requires the application of the schema and filling in all appropriate attributes. Applying the schema requires the Schema Wizard and follows the same procedure explained in section 4.1.3.1. The only difference is that the schema specific to the Hybrid Geodatabase will be used. Once the schema is applied, the tables that are an element of the original ArcHydro model are included in the model geodatabase. The attributes are populated following the method described in section 4.1.3.2. It is important to make sure that all attributes that are necessary for relationships are complete. Once this has been verified the relationships can be created between the two databases.

4.2.4 Creating Relationships

The relationships in the Hybrid Geodatabase are created by the user and are not part of the automated process of applying the schema. The process is a two step procedure. The BASINS tables are first linked to the geodatabase in Access and then the relationships are created in ArcCatalog.

To link the BASINS tables to the ArcBASINS geodatabase, the geodatabase is opened in Access. Open the geodatabase in Access. Navigate from File => Get External Data => Link tables. Navigate to the BASINS Tables Access database. The Link Tables interface will appear and press the Select All button. Figure 4.35 shows the Link Tables interface. All tables in the BASINS Table database is linked to the geodatabase.

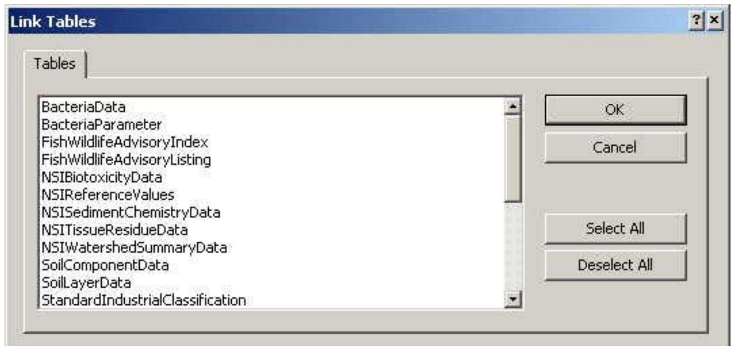


Figure 4.35: Link Tables Interface

The next step is to register the model's relationships within ArcCatalog. Open ArcCatalog and navigate to the ArcBASINS geodatabase. The tables that were absent from the schema appear in ArcCatalog. If one tries to create a relationship class between the geodatabase features and a table from the BASINS database with out registering the database, an error appears that states a table may not participate unless it is registered with the geodatabase. To register a table, right-click on the table. One of the options available is Register with Geodatabase; select this option. Once this selection is made for all BASINS tables, the geodatabase is ready for the creation of relationships.

Relationship classes are created in a similar fashion to Feature Classes. Right-click on the geodatabase containing the BASINS data and select New => Relationship class. The New Relationship Class interface appears. The relationship between the Bacteria Feature class and the BacteriaData table is created as an example. Type in the name of the relationship based on the nomenclature of the data model. Select the Bacteria class under Hydrography as the origin table and the BacteriaData table as the destination table, as illustrated in Figure 4.36.

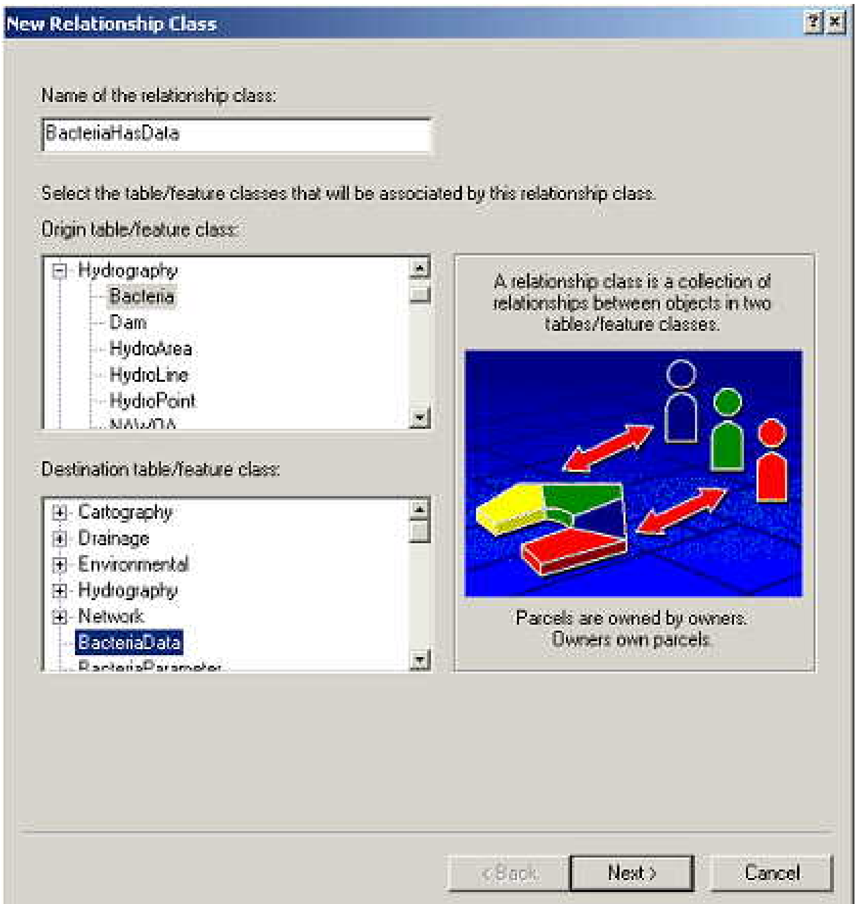


Figure 4.36: The First Screen of the New Relationship Class Interface

Press Next and toggle the simple relationship option. Proceeding onto the next screen, keep all default selections. The fourth screen asks if the relationship will be one to one, one to many, or many to many, as shown in Figure 4.37. The relationship is one to many.

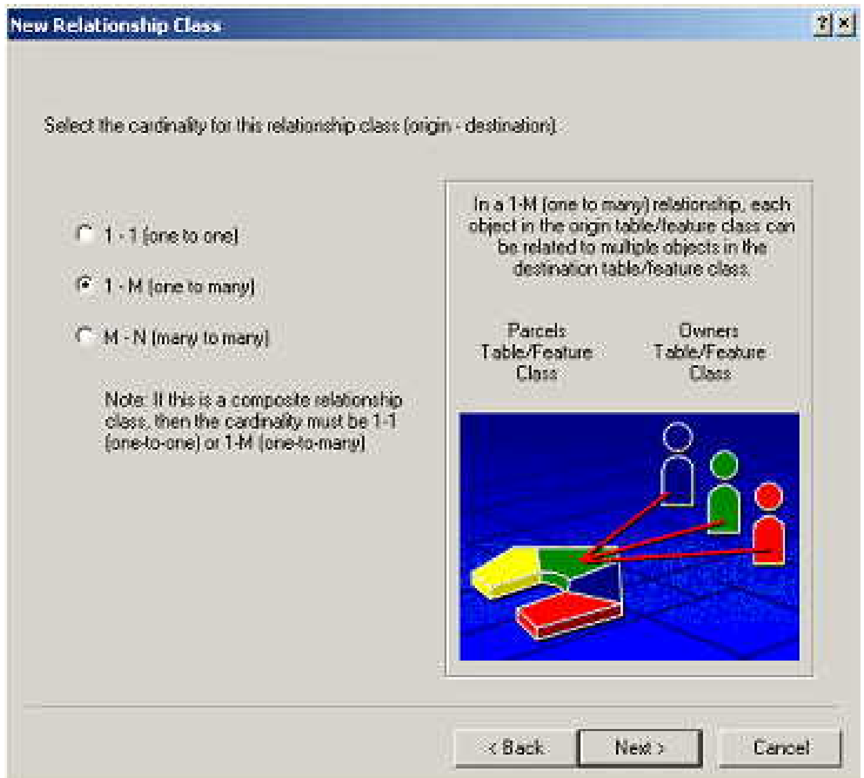


Figure 4.37: Cardinality for a Relationship Class.

Press Next and instruct the interface that no new attributes need to be added to the relationship. On the following screen specify the relationship's keys. The attribute Station is the key fields for both Bacteria and BacteriaData, as shown in Figure 4.38. Press Next and select Finish. The relationship class BacteriaHasData is added to the geodatabase. Appendix D contains all relationships located within the model.

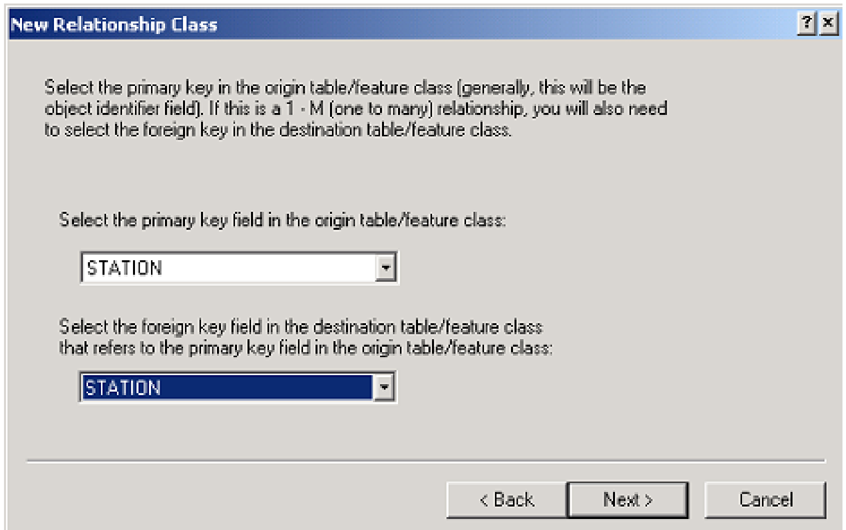


Figure 4.38: Specifying Relationship Keys in ArcCatalog.

Once all relationships have been added to the geodatabase, it should look like the geodatabase created using the Single Geodatabase method. The major difference with the Hybrid geodatabase is that its tables are only linked to the geodatabase, and therefore, can be updated separately from the database.

4.3 Time Series

One component of the model that must be considered separately from the implementation of both the Single and Hybrid Geodatabase methods is the Time Series component. Time Series is an important part of a hydrological and water quality model. The flow in a river and rainfall over a study are vital time varying elements that must be represented. The BASINS database itself does not come with flow or rain data that is presented in a tabular manner. The HSPF program that accompanies BASINS requires a Watershed Data Management (WDM) file that contains time series data for all meteorological parameters. The file is binary and can be created with a Utility available with BASINS 3.0 called WDMUtil. To utilize the TimeSeries table available with ArcBASINS, data must be imported from another source, which may not cause a problem for some projects that have local flow or rainfall data available. If such data is not available, CRWR (www.crwr.utexas.edu) has developed a tool that downloads flow from the USGS website and populates the TimeSeries table. The following section will explain this tool further.

4.3.1 NWIS Tool and TimeSeries

The NWIS tool retrieves data from the USGS Internet site and constructs a TimeSeries table for selected gage stations. NWIS stands for National Water Information System. This system contains surface water, groundwater, and water quality information for USGS stream gages in the United States, with records dating back as far as 100 years for some stations. The data is obtained by specifying parameters such as the gage number and years of record desired. It is exciting to note that some stations now provide real time data, whereas previously only historical data was available. The NWIS tool currently downloads historical daily stream flow information.

To use the NWIS tool, the first step is to load the toolbar into ArcMap. Go to Select Tools => Customize and under the Commands tab press the Add from File button, which allows you to select the NWIS DLL. Figure 4.39 demonstrates how to add a dll from a file. A dll is a packaging mechanism for Visual Basic code. Many tools for ArcGIS are packaged as dlls.

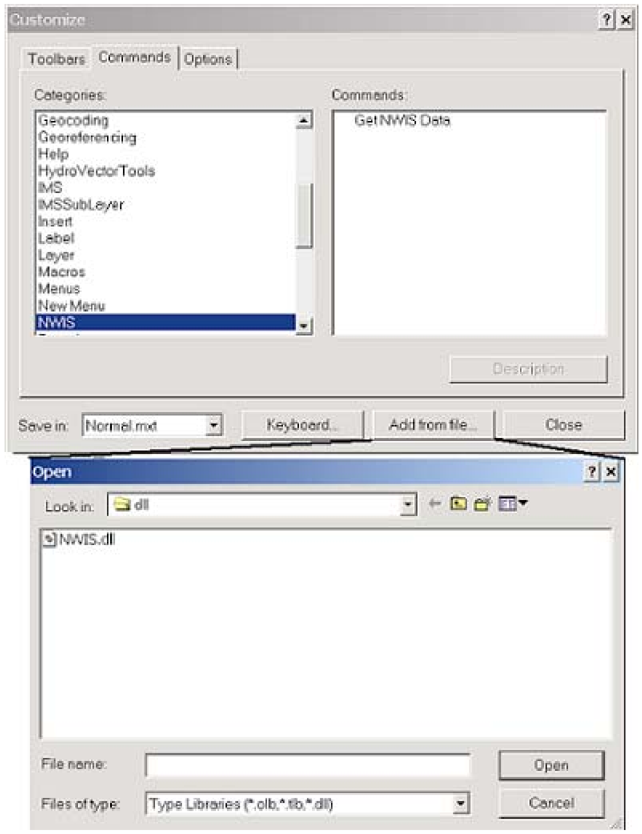
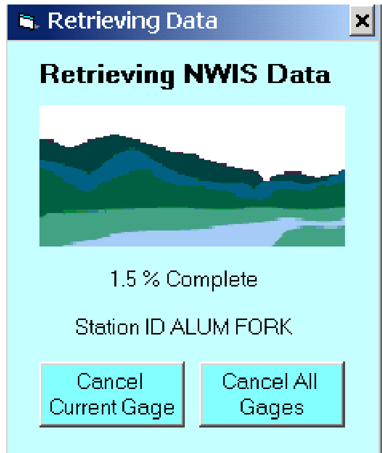


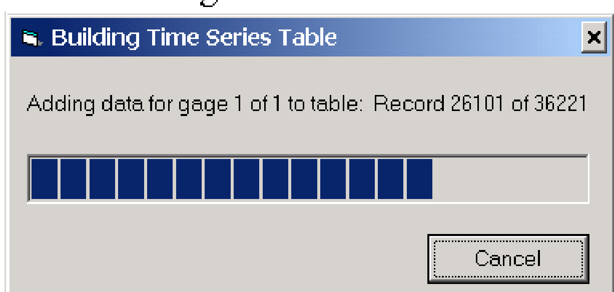
Figure 4.39: Importing a DLL into ArcMap

Once the tool has been added to the ArcMap document, open the attribute table for the USGSGages feature class and make sure that the USGS Gage ID is available in the format used on the USGS website. The ID should be a number that is 8 – digits long. The USGSGages feature class created from the gages point file available from the BASINS database does not have this ID isolated, it is presented in the ACCY field as USGS plus the 8 digit code. Therefore the ID must be isolated using some data management techniques. The method utilized for this example is to create a new field called LocationID. Then using a VBA script written by Tim Whiteaker, from CRWR, the USGS Gage ID is isolated. Appendix F contains this code for reference.

The next step is to select a gaging station of interest. For this example, all of the USGS gaging stations in the Austin-Travis Lakes Cataloging Unit are selected. Then the Get NWIS Data button is pressed. The screen that appears asks for the Gage Station Layer, GageID field, and the FeatureID field, which are USGSGages, LocationID, and FeatureID respectively for the example. With these items selected, press OK. The next screen requests the period of record retrieval, enter the dates of interest, selecting data for the last thirty years with the dates 01/01/1970 to 12/31/2001. After pressing OK, the next screen asks for the TimeSeries table. Once this table is selected a status screen that tracks the progress of the data download will appear followed by another status screen for the creation of the TimeSeries table for each station. Figure 4.40 displays these two status screens. The creation of the table will take a little more than twenty minutes.



Retrieving NWIS Data Status Bar



Building Time Series Table Status Bar

Figure 4.40: NWIS Tool Status Bars

The TimeSeries table now contains data for the selected USGS Gage sites for 30 years. It is possible to create a hydrograph for a selected station in Excel. A hydrograph is constructed for the period of retrieval for the station below Lake Travis that has the ID 08154510 in Excel. Start Excel from the start menu. To retrieve the data from the geodatabase, a query must be performed on the geodatabase. Select Data => Get External Data => New Database Query. In the window that appears select MS Access Databases and press OK. Then the

ArcBASINS geodatabase is selected. The Query Wizard will ask for the table and attributes required, Figure 4.41 displays these selections.

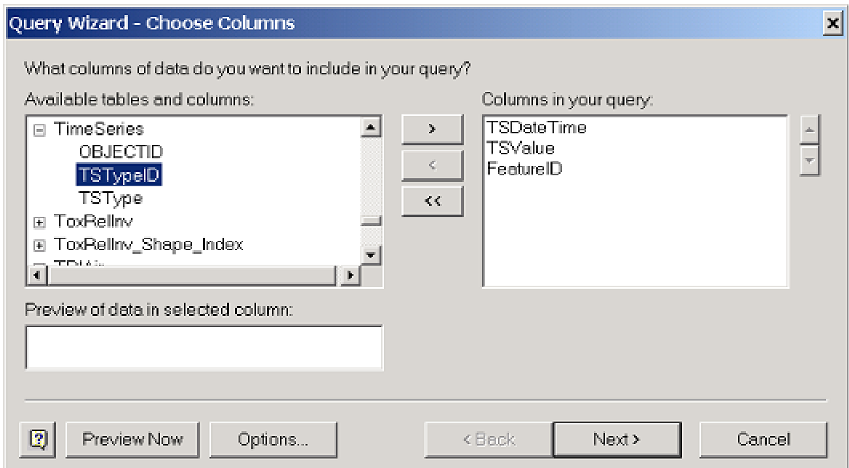


Figure 4.41: TimeSeries Selections in Excel's Query Wizard.

The next screen allows the data to be filtered. Filter the data using the Feature ID for the station as shown in Figure 4.42.

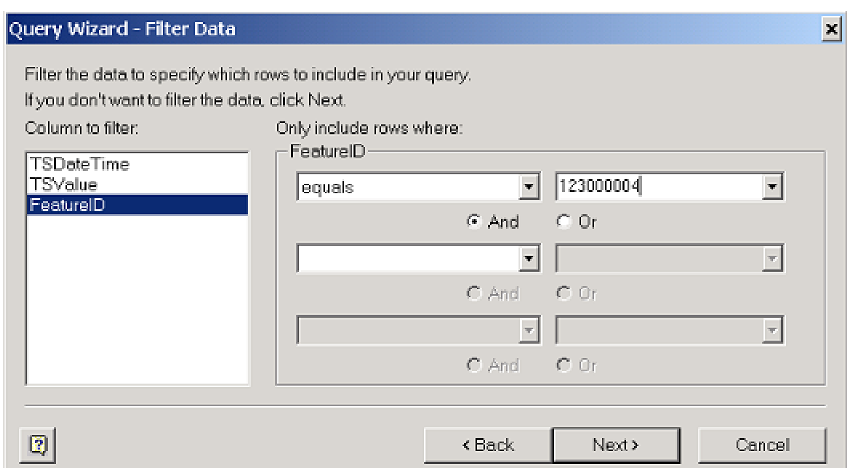


Figure 4.42: Filter Selections in Excel's Query Wizard.

The next screens will ask if the data should be sorted, which is ignored for this example. In the final screen select the first toggle, which brings the query back into the Excel and places the data in the A1 cell. The data will appear as in Figure 4.43.

	A	B	C
1	**TSDateTime**	**TSValue**	**FeatureID**
2	10/1/1974 0:00	5410	123000004
3	10/2/1974 0:00	5420	123000004
4	10/3/1974 0:00	5530	123000004
5	10/4/1974 0:00	5530	123000004
6	10/5/1974 0:00	5530	123000004
7	10/6/1974 0:00	5530	123000004
8	10/7/1974 0:00	5530	123000004
9	10/8/1974 0:00	5530	123000004

Figure 4.43: A Portion of the TimeSeries Table in Excel.

The final step will be to create the hydrograph, which is simply done by graphing the data in the TSDateTime and TSValue columns. Figure 4.44 displays the hydrograph for the selected station between the years 1974 and 1990. This is smaller than the retrieval request but it only indicates that the station was only active for the above listed years.

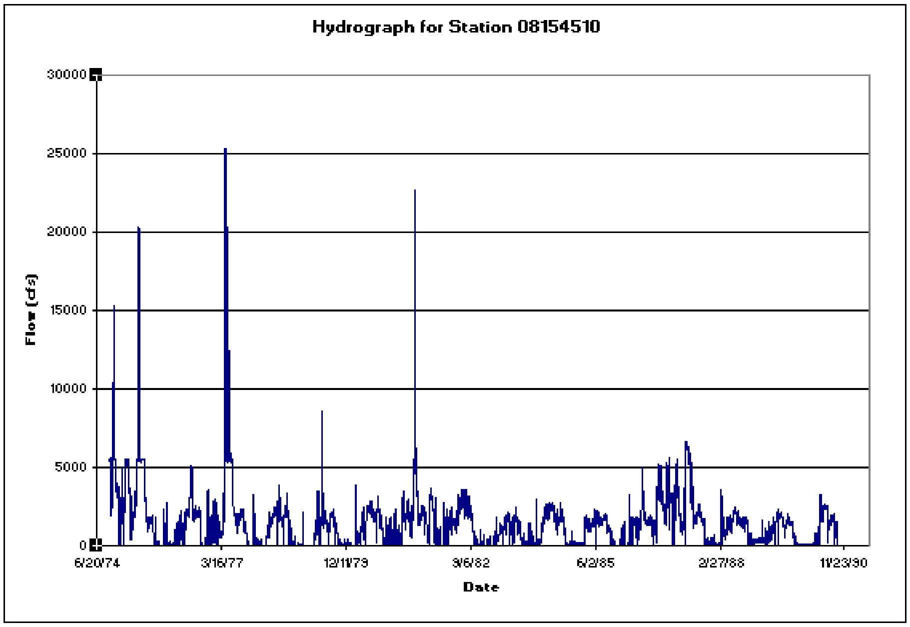


Figure 4.44: Hydrograph for Station 0815800.

The NWIS tool is a powerful mechanism for retrieving historical flow data from the USGS website. It allows users to pinpoint the stations and years of record for which data is needed. When the TimeSeries is populated, it may be used for hydrologic and water quality analysis. The NWIS tool is an example of a tool that may be used to fill the TimeSeries table since the BASINS database does not come with readily accessible time series data.

Chapter 5: Results

This chapter explores the final product of the application of the ArcGIS BASINS Hydro data model. The completed geodatabase for the Austin Travis Lakes study area is examined, highlighting some differences between the ArcBASINS model as described earlier and its actual implementation. The two procedures for the implementation of the model are compared to determine the pros and cons of each method and the methodology for the exploration of the relationships is explained. A method to update data is also described. Finally, three hypothetical examples are presented to demonstrate the functionality of the ArcBASINS model.

5.1 COMPLETED GEODATABASE OVERVIEW

The Arc BASINS geodatabase presents the results of applying the ArcGIS BASINS Hydro data model schema to the Austin-Travis Lakes study area. The data is organized into five separate feature datasets, which contain data related to one another as established by the data model. The feature classes carry the nomenclature of the model and each has the required attributes. A geometric network has been created to allow for the navigation of the dataset topology. Multiple tables as well as relationship classes are available to the user to facilitate environmental examination of the study area. Figure 5.1 contains a view of the BASINS geodatabase in ArcCatalog.

Figure 5.1: Arc BASINS Single and Hybrid Geodatabase after Model Application

The geodatabase looks the same for both methods of the model implementation since both produce the same final product. For the Hybrid Geodatabase Method the tables appear as if they are located in geodatabase even though they are actually stored in their own Access database. In addition, it is important to note that some classes are empty for the study area. The NSI Biotoxicity, Permitted Discharge Parameter, and Permitted Discharge tables are empty since the data available from BASINS source database for the example area does not include this information. Situations like this often occur when applying the model because it is created to contain all possible data types that may be obtained from the BASINS source database. This database was created from various governmental agency sources and so not all data sets are complete. Also, there are some classes such as Catchment and Drainage Line that are important to the ArcHydro data model but are not provided by BASINS. It was decided to retain these classes because of their importance in presenting a complete picture of a region's hydrologic character. The user of the model has the option to fill or delete any of the empty classes present in the ArcBASINS model.

5.2 A Comparison of the Implementation Methods

The results of the two implementation versions in ArcGIS are the same, a geodatabase formatted and related as described by the ArcBASINS model. The major differences between the creation methods are related to the actual implementation and the maintenance of the databases after completion. The Single Geodatabase method is applied more rapidly since its schema contains all classes of the model: feature classes, relationship classes and tables. Once the data is arranged in the geodatabase, the schema is merely applied and the geodatabase fits the bounds of the ArcBASINS model. Relationships are created automatically and only a verification of the relationship determining attributes is necessary. However, the geodatabase for the Single Geodatabase method is very bulky and large. All data updates must be made in ArcGIS, which is fine for the geospatial data. Updating tabular data in ArcGIS can at times be cumbersome since the program was not optimized for this function.

The Hybrid Geodatabase method was conceived to ease access to the tabular data available with the BASINS database. The tables contain large amounts of important data that should be updated on a regular basis. By placing the tables within an Access database their management is simplified since Access is especially designed to deal with large amounts of tabular data. Therefore, tables may be updated separately from the geodatabase but still register within the spatial context of the model. The schema for the Hybrid Geodatabase method is also more streamlined and requires less adjustment when elements of the model are updated. The drawback of the Hybrid Geodatabase method is that the relationships must all be created by hand, which can be a long process and prone to human error.

The Single Geodatabase method is a compact process that yields a sizeable schema and geodatabase that may be hard to work with. In addition, the maintenance of tabular data may be hampered by the geodatabase structure. The Hybrid Geodatabase method remedies this problem, creating a more streamlined schema that allows for the maintenance of tabular data. However, the Hybrid Geodatabase method requires more hands on manipulation of the geodatabase that may cause confusion. The tables and relationship classes are the two elements of the model that motivate the use of either of the two methods.

5.3 RELATIONSHIPS IN BASINS

The relationships stipulated by the ArcBASINS model allow for straightforward navigation of the many tables and classes of the model. ArcCatalog allows for the management of relationship classes, while ArcMap allows for the instantiation of the relationships. Relationships do not appear in the navigation window of ArcMap but are invoked when needed in the attribute table of the desired class. To invoke a relationship, the attribute table of the selected classes must be open. The Options button that is at the bottom of the attribute table contains the Related Tables option, which displays the related tables. Figure 5.2 shows the relationships available for the Bacteria Feature class. The relationships are listed by the name of relationship given in the ArcBASINS model. If features have been selected in the attribute table and the relationship is invoked, only the features related to the selected elements are shown. In this manner all related tables can be accessed.

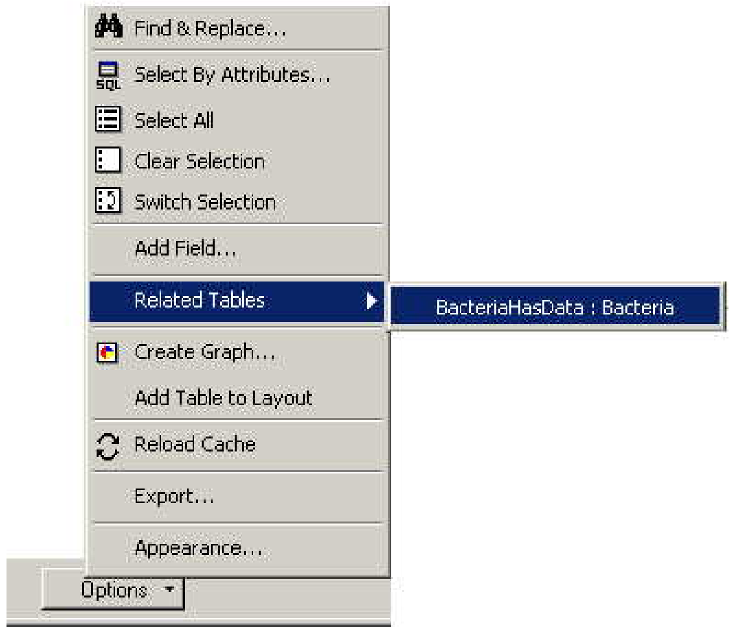


Figure 5.2: Bacteria's Related Tables

5.4 UPDATING DATA SOURCES

Once the ArcBASINS geodatabase is created, it will be necessary to maintain the data. Data may be added to the geodatabase and the schema reapplied, where the attributes associated with the ArcBASINS model will be inserted to the new data. It is also possible to add data to an existing feature classes. The Load Data tool, available with ArcMap, allows data to be added to a feature class.

To use the Load Data tool open ArcMap and navigate to the Customize option under the Tools heading. The tool must be added to the Map document. In the Customize screen under the Commands tab select the Data Converters Category. The Load Data tool will appear, drag the tool to the main toolbar. Start an editing session, since the tool may only be used in an editing session. Select the feature class to which data will be adjoined and press the Load Data button. An Object Loader screen appears and the shapefile or feature class that contains the new data is selected. Press Add and the screen appears as in Figure 5.3. In the following two screens, it is possible to specify the source field for each field in the existing feature class and decide to import all features from the source data or just a select few. The fourth screen of the Load Data tool prescribes the snapping and validation options for the added data. The fifth and final screen summaries the choices made previously. Press the Finish button and the objects will be loaded. It important to note that the snapping environment under the Editor toolbar must be set before the Load Data tool is used if the items will be snapped.

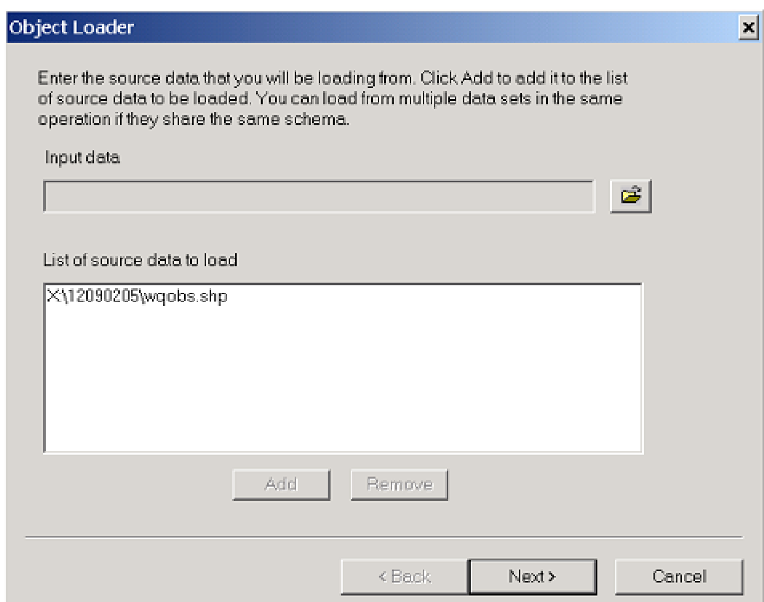


Figure 5.3: Load Object's Object Loader Screen.

To add data in MS Access, open the table to which data will be added. Select Insert => New Record and a new record will be added to the table. Then just copy the new data into the table, using Copy and Paste commands. It also possible to keep the tables separate and perform a query that will contain both the old and the new data. A query is created in the Query View using the query wizard; note that a relationship must exist between the two tables for a query to performed on them simultaneously.

5.5 Application of the BASINS Geodatabase

With the creation of the BASINS geodatabase, an organized structure of environmental data has been created. The collection and relation of the data allows for examination of the environmental condition of an area. The Austin-Travis Lakes study area is used to demonstrate how the BASINS geodatabase and ArcGIS can be used to answer some simple but relevant environmental status questions. Three hypothetical environmental questions are posed and answered using the BASINS geodatabase.

5.5.1 Bacteria Measurement Sites

The first hypothetical situation occurs in the City of Marble Falls, which is located on the Colorado River in the middle of the Highland Lakes chain. Figure 5.4 shows the city of Marble Falls in relation to Austin, TX. The city has a population of five thousand and is growing, as is much of Central Texas.

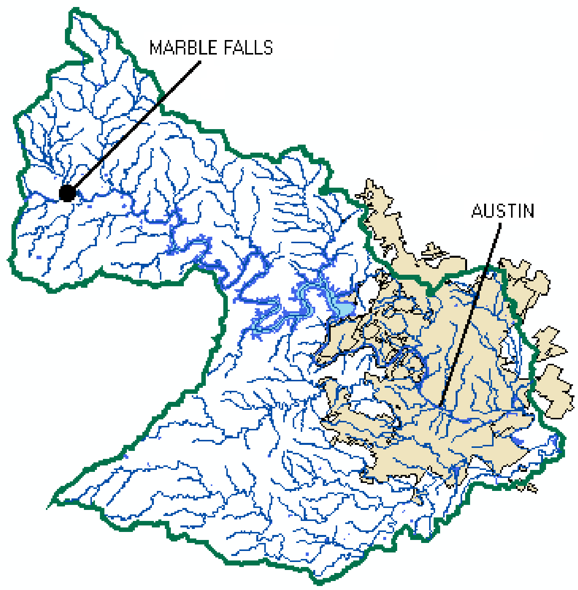


Figure 5.4: The City of Marble Falls

The city has recently become aware of some cattle ranches upstream of the city that have increased rapidly in size over the last few years. The Water and Wastewater Utility is worried about the quality of the water near the water intake for the city's public drinking water supply system and would like to know if a monitoring site already exists for bacteria near the intake. This question is answered quite simply using the BASINS geodatabase.

In ArcMap, open the attribute table of the WaterWithdrawal point class. Select all points located in the city of Marble Falls. Next, open the Selection by Location interface available under the Selection menu. Select from the Bacteria class, all features that are located with in 750 feet of the selected WaterWithdrawal features as shown in Figure 5.5.

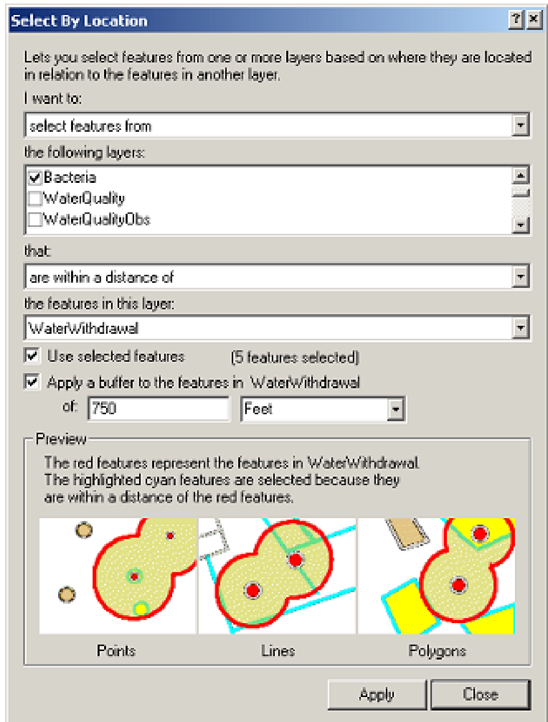


Figure 5.5: Selection by Location for Bacteria Stations.

Opening the Bacteria attribute table, only one station has been selected. The station has the station number 12320 and is located on Lake Marble Falls on the US 281 Bridge. The Utility Manger could now invoke the relationship to Bacteria Data from the Bacteria point class and determine how many years of record are available for this station. The station has six records available with the last record associated with the 8084 Year Index, meaning the station may have not been utilized in almost twenty years or data has not been reported.

5.5.2 Air Quality in Austin

The second hypothetical situation takes place in Austin. The City of Austin is worried about the air quality in the Austin region since the city may be in danger of violating new ozone regulations. The city's environmental department has decided to look at sources of air pollution near the city and analyze what sources may be a potential polluters.

The BASINS geodatabase is a good place to start because it contains Toxic Release Inventory Report information. The database contains a ToxRelInv point file and multiple TRI emission tables. One of the tables is TRI Air, which reports on toxic emissions to the air. In ArcMap, select the TRI sites that are in Travis and Williamson Counties, where Austin is located. This selection is created by using the Selection by Location interface and using the "completely contained within" command. Open the ToxRelInv attribute table and invoke the relationship to the TRIAir table as shown in Figure 5.6. The TRIAir table opens with 1338 out of 1409 records selected.

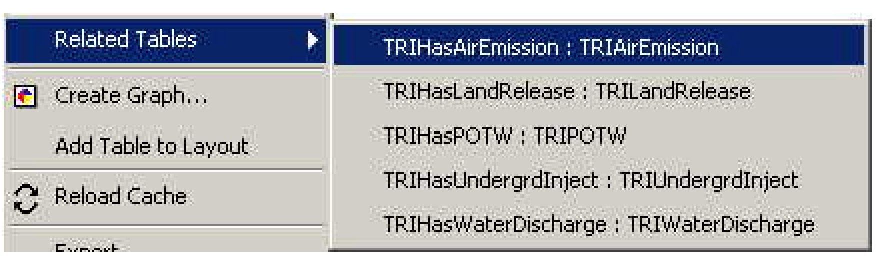


Figure 5.6: ToxRelInv has AirEmission

The relationship from the TRI Air table to the TRI Parameter table can be invoked, displaying the chemicals that are reported as released to the air. A large range of chemicals is presented.

The City of Austin would like to determine which TRI sites are emitting lead so that they may keep a closer eye on these facilities. Relationships and selections may also be used to answer this question. The first step is to determine the TRIChemID for lead using the relationship between the TRI Air and TRI Parameter tables. The code is N420. In the TRI Air table use the Select by Attribute function under the Option button to isolate from the currently selected TRI Air records those which contain lead as the chemical of interest. Figure 5.7 displays the entries for this dialog. The important item to note is that the selection method is Select from current selection. This means that from the ToxRelInv sites reporting for all of Travis and Williamson County the records reporting lead emissions will isolated. Then invoking the relationship between the TRI Air table and the TRI point file isolates one ToxRelInv site. Upon inspection the site is determined to be a Texas Instruments Inc plant located on Research Boulevard in North Austin. Figure 5.8 illustrates the location of the Texas Instruments plant and other ToxRelInv sites within the Austin's city limits.

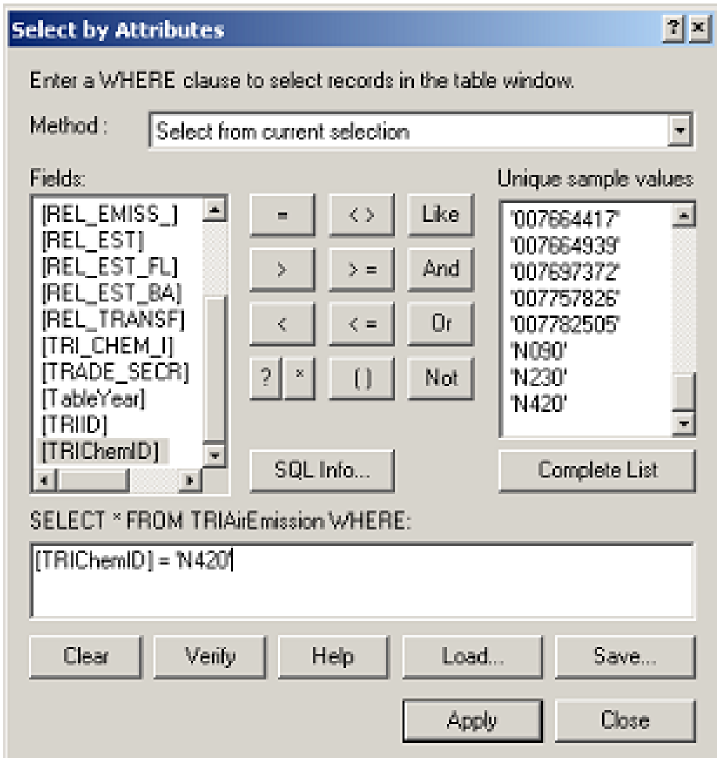


Figure 5.7: Selecting TRIAirEmission Records

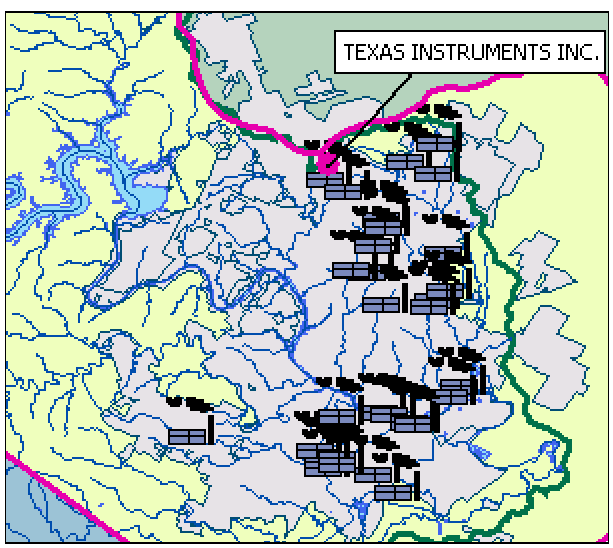


Figure 5.8: Selected ToxRelInv Site from Search

5.5.3 Highway Runoff

The final hypothetical situation looks at a problem that may be of concern for the Texas' Department of Transportation (TXDOT). TXDOT is in the process of estimating the affect of highway runoff on the water quality levels in streams located near major roads in Austin. Multiple constituents are being analyzed. Zinc, a common element in tires, is one of chemicals that after rainstorms can be measured at elevated levels in waterbodies. TXDOT would like to determine which of the water quality measurement stations located around Austin has the highest zinc concentration of record and if this concentration is within the limits set by the federal government's water quality criteria.

The BASINS geodatabase may be used to determine which water quality stations are near major freeways in Austin, as well as the zinc concentration records for these sites. The first step is to determine from the Major Roads feature class, which roads could be classified as major thoroughfares. Major Roads attributes include the FCLASS attribute, which identifies the assigned functional class of each feature. The functional class corresponds to the assigned classification from the 1992 Functional Reclassification by State agencies (FHWA, 2002). The FCLASS value of twelve relates to the classification of Urban Freeway or Expressway. In the Major Road class, all features with FCLASS = 12 are selected.

Using Select by Location, the Water Quality stations that are within 1000 feet of the selected roads are selected. A distance of 1000 feet is used since it is estimated that these stations pick up any changes in water quality caused by road runoff. Now, some Water Quality stations do not intersect the stream network used in the geodatabase. Many possible explanations for this data inaccuracy are available, but the most obvious is that the river network and water quality stations come from two different data sources. Therefore the next step is to determine which of the selected Water Quality stations is close to a 1:100,000-scale stream. The Select by Location dialog is useful for this task. Also, the stations that are apparently not located near a stream can simply be removed by deselecting them.

Once the Water Quality stations of interest have been determined, their relationship to the Water Quality Data table is invoked. From this table, the relationship to the Water Quality Parameter is activated to determine that the ParameterID for zinc is 01090 and if any of the stations have measurements for zinc. When determining which records carry the Parameter attribute for zinc in the Water Quality Data table, twenty of over two thousand records are selected. To determine if the records are within in regulatory limits, the Water Quality Criteria table is consulted and it is determined the Fresh Chronic criteria for zinc is 110 μg/L. All records are well below this limit. The station with the highest recorded mean value of 23 μg/L is shown in Figure 5.9 as the highlighted Water Quality point. It is interesting to note that this station is located near the Capital of Texas Highway and Route 1 intersection, which is one of the busiest intersections in Austin.

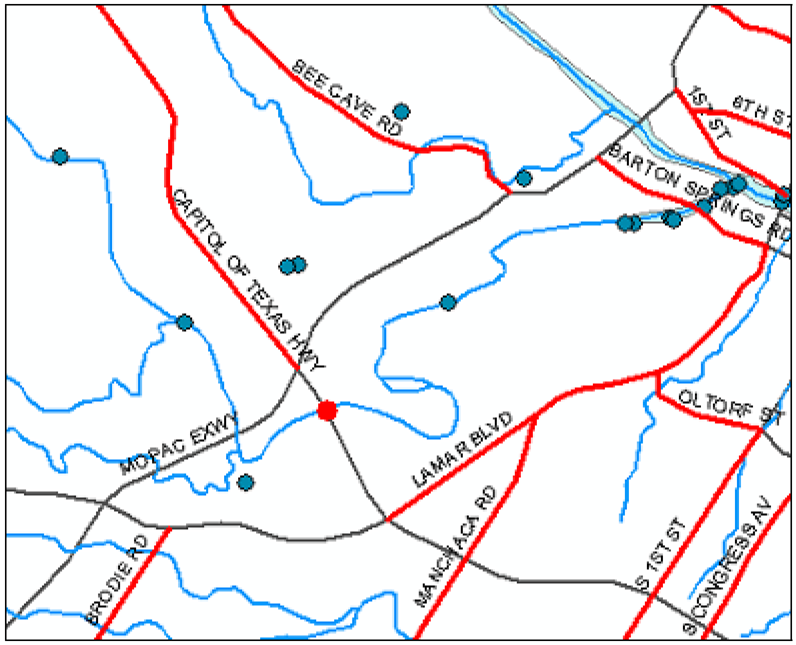


Figure 5.9: Selected Water Quality Station

5.5.4 Summary

The BASINS geodatabase, to which the ArcBASINS model has been applied, was utilized in the previous three examples to answer environmental policy questions. By providing a standardized framework for the organization and nomenclature of environmental data, the ArcBASINS model establishes a protocol for the analysis of environmental issues. If all three agencies mentioned in the examples used the ArcBASINS model as the foundation for the collection of their environmental data, they could communicate the results of their investigations to one another without confusion. The model is a foundation for the presentation of environmental information.

Chapter 6: Conclusions

This thesis chronicles the conversion of the ArcGIS Hydro data model into the ArcGIS BASINS Hydro data model (ArcBASINS). The ArcBASINS model establishes a basis for the adaptation of the BASINS program for the change from the ArcView GIS setting to the newer ArcGIS architecture. In addition, the model provides a new mechanism to aid in the realization of one of BASINS objectives, to facilitate the examination of environmental information.

6.1 Overview

The ArcGIS Hydro data model was the first user-created model produced to work with ArcGIS (Davis, 2000). The ArcHydro model creates a systematic method to organize water resources data while supporting hydrologic and hydraulic analysis. The model is based on the geodatabase method of presenting spatial data. A geodatabase is a store of geographic data. The model contains separate feature datasets for various related data types establishing a data schema that groups data based on its purpose. For example, the Drainage feature dataset includes data that has been derived from Digital Elevation Models. The Network feature dataset brings the other four components of the model together with a geometric network that allows for the navigation of the hydrologic features of the dataset. Other features may be linked to a network to allow for a complete integration of the dataset. The innovation in GIS technology presented by ArcGIS and this technology's implementation in the creation of the ArcHydro data model provided the impetus for the creation of the ArcBASINS data model.

The BASINS dataset is one of the most complete environmental management databases available in the United States. Data from various Federal agencies is brought together in one place providing place-based access to a diverse mixture of environmental data. After a comparison between the BASINS database and other national datasets, as well as GIS data models, it was determined that the present organization of the BASINS database in ArcView 3.2 is not readily apparent to the user. The ArcBASINS data model provides a structure for the storage, retrieval, and presentation of the BASINS database.

ArcBASINS was created from the ArcHydro model. Four of ArcHydro's five data components were retained: Hydrography, Drainage, Network, and Time Series. The Channel feature dataset was the only one removed. The BASINS database does not contain any data that corresponds to the purpose of the Channel feature dataset. The lack of corresponding data also mandated the removal of various feature classes throughout the original model with the Hydrography feature dataset being the most affected. Two feature datasets, as well as numerous feature classes, were added to the ArcBASINS model to hold the data objects that are present in the BASINS database but not in ArcHydro. Environmental and Cartography are the added feature datasets. The Environmental dataset contains data that is related to the environmental health of a region but is not directly linked to the hydrography of a study area such as Superfund site information. The information describing the political and man-made features of a region is placed in the Cartography dataset. The ArcBASINS Hydrography dataset contains several new feature classes that are added to flesh out the representation of Hydro Points within the model to accommodate more data from BASINS.

The original ArcHydro model contains only a few tables and relationships, which is consistent with the philosophy behind its creation to present a clean, simple hydrologic model that encompasses only the essentials. The ArcBASINS model, however, represents the entire data collection of the BASINS database, whose power is based on the information it provides. With the tables in place, the creation of relationships between the spatial and tabular data was the next logical step. Relationships use key fields with in the tables to create an integrated link between the tables. The fields for the relationships exist in the current BASINS database but are implemented in a more streamline fashion in ArcGIS. The ArcBASINS model contains twenty-two relationships which are built between tables as well as tabular and spatial data. The relationships allow the user to relate data and query based on their specified requirements.

The result of applying the ArcBASINS model is a geodatabase that contains the data available from the BASINS database as well as multiple relationships that facilitate the examination of the data. It was observed that some classes created by the ArcBASINS model were blank when the model was applied to the Austin-Travis Lakes study area. This is to be expected because while the data offered by BASINS is substantial, it is not uniform throughout the nation due to different data collection methods in different regions. The Austin-Travis geodatabase then was used in this research to answer three hypothetical environmental management questions. The first question looks into the location of a Bacteria monitoring point near the water intake of the City of Marble Falls, Texas. The second question explores air quality issues in Travis County and the third question investigates the affect of Major Highways on water quality in with Austin, Texas. The geodatabase offered a quick method to locate the information required and provided answers. The ArcBASINS model establishes a standardized method for the presentation of environmental data that can ease communications between different agencies and organizations involved in the management of the environment.

6.2 RECOMMENDATIONS

Having developed a method to model the data available with the BASINS program and having explored the possibilities of BASINS in ArcView, this study has resulted in a few recommendations concerning the nature of data available with BASINS and the manner in which it should be presented to the public. The first group of recommendations and conclusions relates to the data available from BASINS. As seen in Chapter 2, the shapefiles that are available from BASINS do not possess intuitive names. For example the fhards.shp file does not have a name that could be recognized without looking at the BASINS manual. The name MajorRoads is much more insightful. It is advisable to change the names of the shapefiles to ones that match the names given in the ArcBASINS model. In this way users who do not have ArcGIS would still benefit from the organization and terminology of the ArcBASINS data model. Users of ArcGIS will be able to import the data from the ArcBASINS database following the layout of the data model. Also, the data would be more user friendly for the public since it will be apparent what each file represents. In regard to data organization, the ArcHydro HydroID attribute is a method for the unique categorization of the BASINS data. Currently the data does not possess a unique identifier associated with BASINS, it only carries the identifiers of the source agency. Use of the ArcHydro HydroID would allow the BASINS program to bring the data into a closer union with program by providing an identification number upon which a new ArcGIS BASINS could be developed.

Once the data is presented in the ArcBASINS method, updating and maintaining the data will be much simpler than the current process. Currently, BASINS data is only updated every two to three years when a new version of the program is released. In the current release, the most recent data is from 1995 or 1997 depending on the data product. Thus, this data is five to seven years old. In the world of environmental modeling it is imperative to have recent data so that the actual condition of a waterbody may be assessed. Updating data will be relatively straightforward in the new ArcBASINS framework when using the Load Data tool in ArcGIS and appending data to a table in Access as described in Chapter 5. Also, new tabular data can be placed in its own table and linked to the original data table in the model in essence creating a chain of linked tables. Users can also update their databases with tools like the NWIS tool, which downloads USGS historical flow data directly off the web into a table with ArcGIS. This would mean the EPA would need to create a water quality data web page like that of the USGS. The user would not necessarily need access to the real-time data, but they would be able to receive data that is only a year old. The EPA could create partnerships with other agencies to allow for EPA to be a portal to their environmental data on the web.

The final recommendation deals with the implementation method that would best serve the purposes of BASINS users. Two methods were investigated: the Single Geodatabase method and the Hybrid Geodatabase method. The Single Geodatabase method places all data, spatial and tabular data, within one geodatabase. The Hybrid Geodatabase method separates the spatial and tabular data available from BASINS into a geodatabase for the geospatial information and a separate related database for the purely tabular data. The Single Geodatabase method creates a large unyielding geodatabase that creates hassles in the maintenance of tabular data. ArcGIS is not designed for tabular data but rather for geospatial data. The Hybrid Geodatabase method eliminates the problems associated with tabular data maintenance. However, users would need to have both ArcGIS and MS Access to work with the data. This may not be a big problem since MS Access is usually a product that is part of the Microsoft's Office package of programs. The Hybrid Geodatabase method would require a higher level of computer sophistication of its users. Where the Single Geodatabase method only requires knowledge of one program, the Hybrid method requires knowledge of two programs. However if a user could learn and utilize the many water quality models available with BASINS, it could be assumed that they could develop a command of ArcGIS and MS Access. Thus, it wise to offer the users of BASINS a choice of geodatabase setups to allow them the greatest flexibility.

6.3 Future Work

The ArcBASINS model establishes a sound basis for the transfer of the BASINS system into the new ArcGIS environment. The data model prepares a framework for access to the data for the user and programs. The next step will be to bring the programs of BASINS into the framework of ArcGIS. The COM compliant nature of ArcGIS makes this process simpler.

The first components of BASINS that can be brought into ArcGIS are the Assessment Tools available with the program. BASINS currently provides the users with TARGET, ASSESS, and Data Mining features. TARGET performs a broad assessment of the region while Assess looks at the data of an individual watershed. The queries created by these programs can be created with the use of the VB for Applications available with ArcGIS. Data Mining builds dynamic links between the map interface and their related tables. This process is already accomplished with the relationships of the data model. However, an ArcBASINS toolset could be produced that will activate the relationships automatically for the user.

A greater programming effort will be needed to convert the Data Extensions, Delineation Tools, Utilities, and Water Characterization Reports. The present Data Extensions' Theme Manager and BASINS import tool will no longer be necessary because of the ArcBASINS data model. All data will already be within the format prescribed by the model, and therefore, will not need to be converted to fit the confines of the models within the program. The other tools available from Data Extensions must be converted to work in the ArcGIS environment. The Delineation Tools available with BASINS have for the most part already been created by ESRI as toolbars in ArcGIS. They will only have to be converted to match any more specific needs of BASINS. Finally, the Utilities and Water Characterization Reports mechanism will need to be reprogrammed from Arc View's Avenue language into the COM-compliant languages available to ArcGIS such as Visual Basic.

The final effort for the conversion of BASINS into ArcGIS will be the conversion of the in-stream water quality models: Qual-2E, HSPF, SWAT, and PLOAD. An interface may be written for Qual-2E and HSPF, so that the models may run outside the context of ArcGIS. Also, PLOAD's Avenue script may be converted into VB to allow it to run in ArcGIS. SWAT will most likely be the hardest program to convert into the new GIS architecture. The model is currently based on raster calculations and rasters cannot be contained in feature datasets. An innovation in the structure of ArcGIS may be necessary for a smooth transition for SWAT.

6.3 SUMMARY

The process of converting ArcHydro to fit the needs of ArcBASINS has yielded some noteworthy observations about the process of converting one data model to another. The ArcHydro model is a good starting point for any data model related to the natural water systems. Much thought and time has been spent considering the different elements necessary for a complete representation of natural waters. It would be impractical to ignore these findings and therefore the ArcHydro data model should be thought of as a basis for data models centered on hydrologic systems. The modification of the model must take into consideration all elements of the source database. The first implementation of ArcBASINS did not contain the BASINS tables, which hampered the true representation of all the data available to the user. Some feature classes were added to the model that did not support a streamlined representation of the database, and therefore, had to be removed. Also, relationships were not considered until later in the development process. The relationships are at the heart of the models applicability and are intrinsic to its utilization. Finally, the organization of a data model should aid in the evaluation of the data. This concept seems quite apparent, but when dealing with large databases it can sometimes be lost in the charge to present all data available.

Bringing the BASINS program into world of ArcGIS will be a big task and the creation of the ArcBASINS model is the first step in this direction. This will aid in the completion of TMDLs, as dictated by the BASINS program. The ArcHydro model is used as a basis for the model and it provides a solid structure on which to base the model. The ArcBASINS model is another phase in the scientific and engineering communities continuing quest to represent the natural world.

Appendix A: ArcGIS Hydro Data Model Analysis Diagram

Appendix B: ArcGIS Basins Hydro Data Model Analysis Diagram

Appendix C: Data Dictionary

A note on the meaning of a Feature Class or Object's color: Abstract Classes are Blue and User Defined classes are Green. These classes do not have a specific BASINS data origin and therefore none is displayed.

Data Package	Feature Class or Object	BASINS Data Product	shapefile or dBase table
Hydrography			
	HydroFeature		
	Hydrography		
	HydroPoint		
	HydroLine	NHD, RF1 or RF3	rf1.dbf, rf1.shp, rf1.shx or (cu).dbf, (cu).shp, (cu).shx or nhd drain coverage
	Dam	Dam Locations	dam.dbf dam.shp dam.shx
	IndDisFac	Industrial Facilities Discharge (IFD) Sites	ifd.dbf ifd.shp ifd.shx
	PCSDischarge	Permit Compliance System (PCS) Sites and Computed Annual Loadings	pcs.dbf pcs.shp pcs.shx
	WaterQualityObs	Water Quality Stations and Observation Data	wqobs.dbf wqobs.shp wqobs.shx
	WaterWithdrawal	Drinking Water Supply (DWS) Sites	dws.dbf dws.shp dws.shx
	USGSGages	Gage Sites	gage.dbf gage.shp gage.shx

Data Package	Feature Class or Object	BASINS Data Product	shapefile or dBase table
Hydrography			
	WaterQuality	Water Quality Monitoring Stations and Data Summaries	wq_stat.dbf wq_stat.shp wq_stat.shx
	NatSedInv	National Sediment Inventory (NSI) Stations and Database	nsi.dbf nsi.shp nsi.shx
	Bacteria	Bacteria Monitoring Stations and Data Summaries	bac_stat.dbf bac_stat.shp bac_stat.shx
	Weather	Weather Station Sites	metpt.dbf metpt.shp metpt.shx
	HydroArea		
	Waterbody		
	NAWQA	National Water Quality Assessment (NAWQA) Study Unit Boundaries	nawqa.dbf nawqa.shp nawqa.shx
Network			
	HydroJunction		
	HydroEdge	NHD, RF1 or RF3	rf1.dbf, rf1.shp, rf1.shx or (cu).dbf, (cu).shp, (cu).shx or nhd drain coverage
	HydroEvent		
	HydroPointEvent		
	HydroLineEvent		
Time Series			
	Time Series		
	TSType		

Data Package	Feature Class or Object	BASINS Data Product	shapefile or dBase table
Environmental			
	HydroFeature		
	EnviroFeature		
	ToxRelInv	Toxic Release Inventory (TRI) Sites and Pollutant Release Data	tri.dbf tri.shp tri.shx
	Superfund	Superfund National Priority List Site	cerclis.dbf cerclis.shp cercils.shx
	Mineral	Minerals Availability System/Mineral Industry Location System (MAS/MILS)	mines.dbf mines.shp mines.shx
	HazSolidWaste	Resource Conservation and Recovery Information System (RCRIS) Sites	rcris.dbf rcris.shp rcris.shx
	CleanWaterNeeds 1996	1996 Clean Water Needs Survey	1996cwns.dbf 1996cwns.shp 1996cwns.shx
	Ecoregion	EPA Ecoregion	ecoreg.dbf ecoreg.shp ecoreg.shx
	STATSGO	State Soil and Geographic (STATSGO) Database	statsgo.dbf statsgo.shp statsgo.shx
	LandUseIndex	Land Use and Land Cover	lulcndx.dbf lulcndx.shp lulcndx.shx
	WeatherStation Range	Weather Station Area	met_stat.dbf met_stat.shp met_stat.shx
	ManagedArea Database	Managed Area Database	mad.dbf mad.shp mad.shx

Data Package	Feature Class or Object	BASINS Data Product	shapefile or dBase table
Cartography			
	HydroFeature		
	CartoFeature		
	County	State and County Boundaries	cnty.dbf cnty.shp cnty.shx
	MajorRoads	Major Roads	fhards.dbf fhards.shp fhards.shx
	PopulatedPlaces	Populated Place Locations	ppl.dbf ppl.shp ppl.shx
	State	State and County Boundaries	st.dbf st.shp st.shx
	UrbanArea	Urbanized Areas	urban.dbf urban.shp urban.shx
	EPARegion	EPA Regions	epa_reg.dbf epa_reg.shp epa_reg.shx
Drainage			
	HydroFeature		
	DrainageFeature		
	DrainagePoint		
	DrainagLine		
	DrainageArea		
	Catchment		
	Watershed	Hydrologic Unit Boundaries	acc.dbf acc.shp acc.shx
	Basin	Hydrologic Unit Boundaries	cat.dbf cat.shp cat.shx

Data Package	Feature Class or Object	BASINS Data Product	shapefile or dBase table
Fish & Wildlife Advisories			
	Fish Wildlife Advisory Index	Fish Wildlife Advisory (1996) Index	ifwa96.dbf
	Fish Wildlife Advisory Listing	Fish Wildlife Advisory (1996) Listing	ifwa96ad.dbf
Lookup Tables			
	Water Quality Criteria Table	Water Quality Criteria Table	wqcriter.dbf
	State Agency Code	State Agency Code	storetag.dbf
	Standard Industrial Classification	Standard Industrial Classification	sic.dbf
Water Quality Station Parameters			
	WaterQuality Obs Parameter	Water Quality Stations and Observation Data	wq_par.dbf
National Sediment Inventory Database		National Sediment Inventory (NSI) Stations and Database	
	NSIData		
	NSIBiotoxicity Data		nsi_bio.dbf
	NSITissue ResidueData		nsi_tis.dbf
	NSISediment Chemistry Data		nsi_sed.dbf
	NSI Reference Values		nsi_ref.dbf
	NSI Watershed Summary Data		nsi_wsh.dbf

Data Package	Feature Class or Object	BASINS Data Product	shapefile or dBase table
Bacteria Data		Bacteria Monitoring Stations and Data Summaries	
	BacteriaData		bc_d7074.dbf bc_d7579.dbf bc_d8084.dbf bc_d8589.dbf bc_d9094.dbf bc_d9597.dbf
	Bacteria Parameter		bc_parm.dbf
Water Quality Monitoring Data Summaries		Water Quality Monitoring Stations and Data Summaries	
	WaterQualityData		wq_d7074.dbf wq_d7579.dbf wq_d8084.dbf wq_d8589.dbf wq_d9094.dbf wq_d9597.dbf
	WaterQuality Parameter		wq_parm.dbf
Permit Compliance System Computed Loading		Permit Compliance System (PCS) Sites and Computed Annual Loading	
	Permitted Discharges		pcs (data format is currently being revised)
	Permitted Discharges Parameter		pcs_prm.dbf
	PCSCode Description		pcs_code.dbf
State Soil and Geographic Database		State Soil and Geographic (STATSGO) Database	
	Soil Component Data		statsgoc.dbf
	Soil LayerData		statsgol.dbf

	StatsgoMapUnit		
Data Package	**Feature Class or Object**	**BASINS Data Product**	**shapefile or dBase table**
Toxic Release Inventory Sites		Toxic Release Inventory (TRI) Sites and Pollutant Release Data	
	TRIData		
	TRIAir		tri_ai87.dbf tri_ai88.dbf tri_ai89.dbf tri_ai90,dbf tri_ai91.dbf tri_ai92.dbf tri_ai93.dbf tri_ai94.dbf tri_ai95.dbf
	TRILand		tri_lr87.dbf tri_lr88.dbf tri_lr89.dbf tri_lr90,dbf tri_lr91.dbf tri_lr92.dbf tri_lr93.dbf tri_lr94.dbf tri_lr95.dbf
	TRIUndergrd		tri_ui87.dbf tri_ui88.dbf tri_ui89.dbf tri_ui90,dbf tri_ui91.dbf tri_ui92.dbf tri_ui93.dbf tri_ui94.dbf tri_ui95.dbf
	TRIWater		tri_wd87.dbf tri_wd88.dbf tri_wd89.dbf tri_wd90,dbf tri_wd91.dbf tri_wd92.dbf tri_wd93.dbf tri_wd94.dbf tri_wd95.dbf

Data Package	Feature Class or Object	BASINS Data Product	shapefile or dBase table
	TRIPOTW		tri_pw91.dbf tri_pw92.dbf tri_pw93.dbf tri_pw94.dbf tri_pw95.dbf
	TRI Parameter		tri_parm.dbf

Table C.1: Feature Class Shapefile Sources

Appendix D: Relationship Keys

The Origin Class is the class from which the relationship originates. The Foreign Class is the destination of the relationship. Relationships can be activated from either the Origin or the Foreign Class.

Origin Class	Origin Key	Foreign Class	Foreign Key
Bacteria	Station	BacteriaData	Station
BacteriaData	Parameter	BacteriaParameter	ParameterID
NatSedInv	Station	NSIBiotoxicityData	Station
NatSedInv	Station	NSITissue ResidueData	Station
NatSedInv	Station	NSISediment ChemistryData	Station
PCSDischarge	NPDES	PermittedDischarges	NPDES
PermittedDischarges	Parameter	PermittedDischarges Parameters	ParameterID
STATSGO	MUID	StatsgoMapUnit	MUID
StatsgoMapUnit	MUID	SoilComponentData	MUID
SoilComponentData	MUID	SoilLayerData	MUID
ToxRelInv	TRIID	TRIAir	TRIID
ToxRelInv	TRIID	TRILand	TRIID
ToxRelInv	TRIID	TRIUndergrd	TRIID
ToxRelInv	TRIID	TRIWater	TRIID
ToxRelInv	TRIID	TRIPOTW	TRIID
TRIAir	TRIChemID	TRIParameter	ChemID
TRILand	TRIChemID	TRIParameter	ChemID
TRIUndergrd	TRIChemID	TRIParameter	ChemID
TRIWater	TRIChemID	TRIParameter	ChemID
TRIPOTW	TRIChemID	TRIParameter	ChemID

Table D.1: Relationship Keys

Appendix E: Applying the ArcGIS Hydro Data Model, Part 1

CE 394K GIS in Water Resources
University of Texas at Austin
Fall 2001

Prepared by: Reem Zoun, Kristina Schneider, Tim Whiteaker, and David Maidment

INTRODUCTION TO THE ARCGIS HYDRO DATA MODEL AND ARC HYDRO LITE

The ArcGIS Hydro Data Model (Arc Hydro) is an ArcGIS geodatabase model. It provides a standardized framework into which various types of water resources data can be loaded. In this manner the data forms an integrated water resources modeling and mapping database.

A geodatabase model is generated in a series of steps, beginning with the definition of classes and attributes in a Unified Modeling Language (UML) diagram created in the Visio 2000 drawing system. The second step is to export the diagram to a Microsoft repository format, which is an equivalent tabular structure, or schema, for loading into Microsoft Access (personal geodatabase) or other relational data servers (enterprise geodatabase). Finally, the data is imported into the Arc Hydro format by applying the schema to an ArcGIS geodatabase. Additional instructions for generating a schema in ArcCatalog can be found in the ArcGIS help files and in the book, ArcObjects Developer's Guide (shipped with ArcGIS).

The ArcGIS Hydro Data Model stores data in four feature datasets, each corresponding to one of the main domains of the UML analysis diagram: **Hydrography** (map hydrography and associated data inventories), **Drainage** (drainage areas derived from digital elevation models or manually digitized), **Channel** (3-D profile and cross-section representation of stream channels), and **Network** (a geometric network representation of the connectivity of the surface water features of the landscape). Associated with these four feature datasets are a set of object tables for additional information, such as events defined on the river network, and time series of monitoring data.

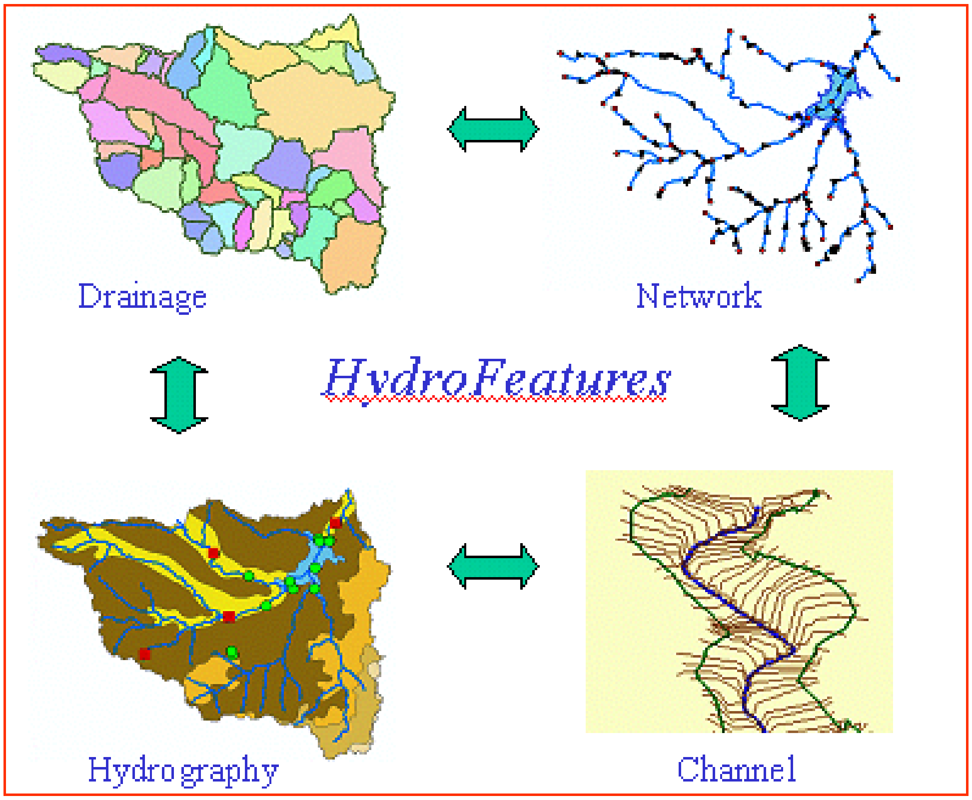


Figure E.1: HydroFeatures of the ArcHydro Model

To apply Arc Hydro, you simply apply the **schema**. One of the goals of this exercise is to apply a schema to a dataset. However, the Arc Hydro model, with four feature datasets, is a bit complicated if just used to demonstrate how to apply a schema.

The **Arc Hydro Lite** schema was created to provide a slimmed down version of the Arc Hydro model to provide practice in applying data models. Arc

Hydro attempts to capture the majority of water resources data available, while Arc Hydro Lite's goal is to represent the feature classes and relationships that will be used most often by users.

Arc Hydro Lite consists of one feature dataset called ArcHydro. The feature dataset contains only five feature classes: **HydroPoint**, **Waterbody**, **Watershed**, **HydroEdge**, and **HydroJunction**. HydroPoint represents point features from map hydrography and inventory sources. Waterbody presents area features from map hydrography. Watershed is a polygon feature class, which contains any subdivision of the landscape into drainage areas. HydroEdge and HydroJunction form a geometric network called HydroNetwork. The UML diagram below shows the relationships that create the schema.

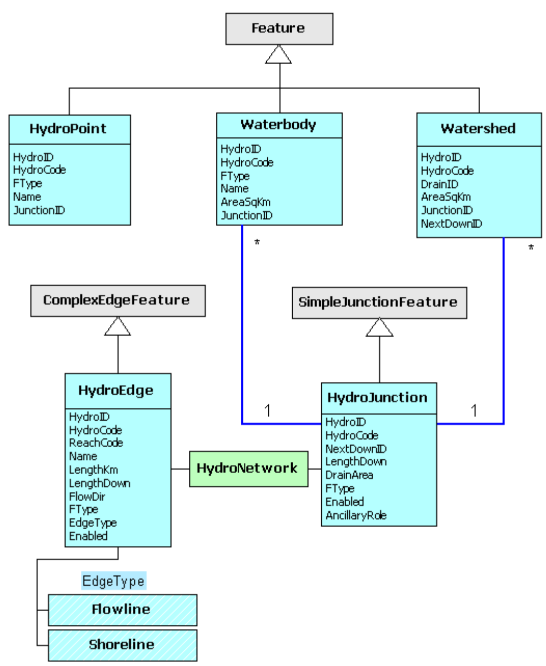


Figure E.2: Arc Hydro Lite Structure

OBJECTIVES OF THE EXERCISE

- To take regularly available geospatial data for hydrology and prepare it in the format needed for inclusion in a data model

- To apply the Arc Hydro Lite schema to these data.

- To run a set of Arc Hydro tools to fill in the attributes contained in the schema.

In this exercise, you'll learn editing skills for linear and aerial features, and also how to create new lines and polygons as part of existing feature classes.

COMPUTER AND DATA REQUIREMENTS

- ArcGIS 8.1
- 35 MB of disk space

Data Files:

ArcHydroLite.mdb	MS Repository for Arc Hydro
Albstat Shapefile	Point shapefile which contains gaging stations
Rf1guad Shapefile	Line shapefile showing the path of the river.
Hucguad Shapefile	Polygon shapefile representing watersheds derived for the Guadalupe basin.

Table E.1: Data files for Exercise.

These files are attached to this exercise as **ArcHydro.zip**.

PROCEDURE FOR THE ASSIGNMENT

1. View your data in ArcMap

(1) Open ArcMap.

(2) Navigate to the directory with your data. Add the **Albstat**, **Rf1guad** and **Hucguad** shapefiles to the data frame. Explore the dataset by looking at the attributes for each class and visualizing them separately. These are the same files that you prepared for the Guadalupe Basin in exercise 2 of the course.

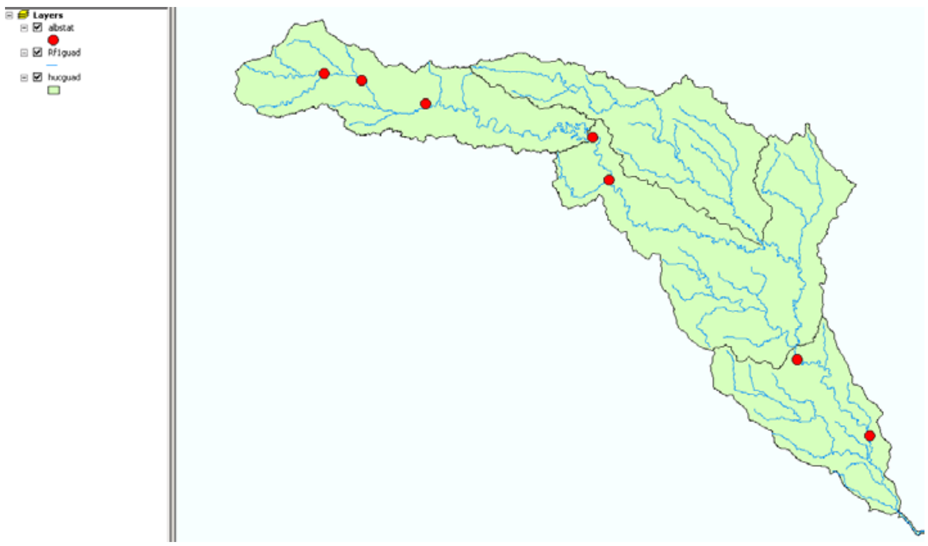


Figure E.3: Monitoring Points in Guadalupe Basin

(3) Close ArcMap, you don't need to save the data.

2. Prepare your Data for Schema Application.

The Arc Hydro Lite schema contains 5 feature classes and at present we have data for only 3 of them: HydroPoint (Albstat.shp), Watershed (Hucguad.shp) and HydroEdge (Rf1guad.shp). In this part of the exercise, we are going to create the data for the other two feature classes, Waterbody and HydroJunction. The first thing we'll do is to create a centerline through Canyon Lake

A. Create Centerline

(1) Open ArcMap and add 'Rf1Guad' shapefile to your data frame.

(2) Zoom into the area where the Guadalupe River goes around Canyon Lake to show the shape of the reservoir. We are going to create a centerline through Canyon Lake to form a complete network for Guadalupe River.

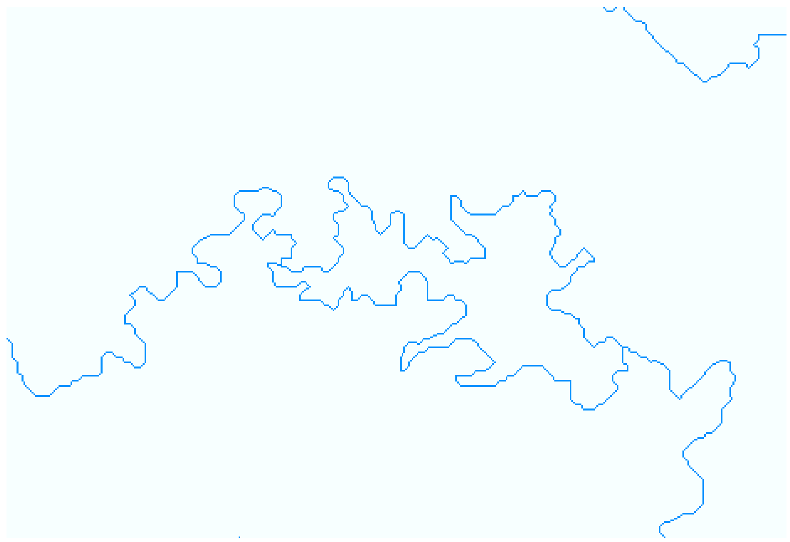


Figure E.4: Canyon Lake

(3) Go to **Tools/Editor Toolbar** and the Editor toolbar will appear. On the Editor toolbar go to Editor/Start Editing.

(4) Go to **Editor/Snapping** and the Snapping Environment dialog will appear. Click all the options on for Rf1guad. It should appear as the dialog below after you have turned them on.

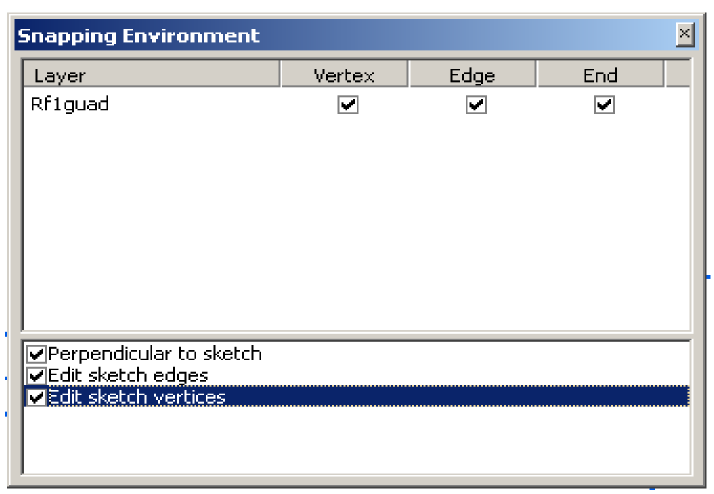


Figure E.5: Snapping Screen

(5) Close the Snapping Environment dialog.

(6) Zoom into the Canyon Lake area and click on the Create New Feature icon on the Editor Toolbar (the little pencil).

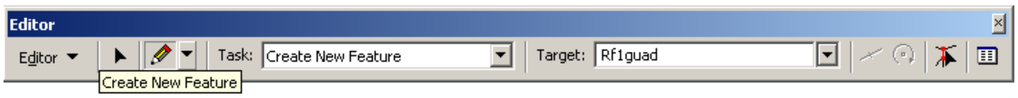


Figure E.6: Editor Toolbar

(7) Snap at the intersection of Guadalupe River and the reservoir and the continue clicking through the middle of the reservoir to crate a centerline. When you reach the end, double click. You should end up with a centerline in the middle of Canyon Reservoir.

(8) Use the **Editor/Save Edits** to save the edits that you've made. If you don't like the centerline you created, use **Editor/Stop Editing** to terminate the editing session and do not Save Edits, and then restart the Edit session and redo the centerline. To modify an existing feature without retracing it, use Task: **Modify feature** in the toolbar shown above and use the  next to the Create New Feature icon to select and edit the features you want to alter.

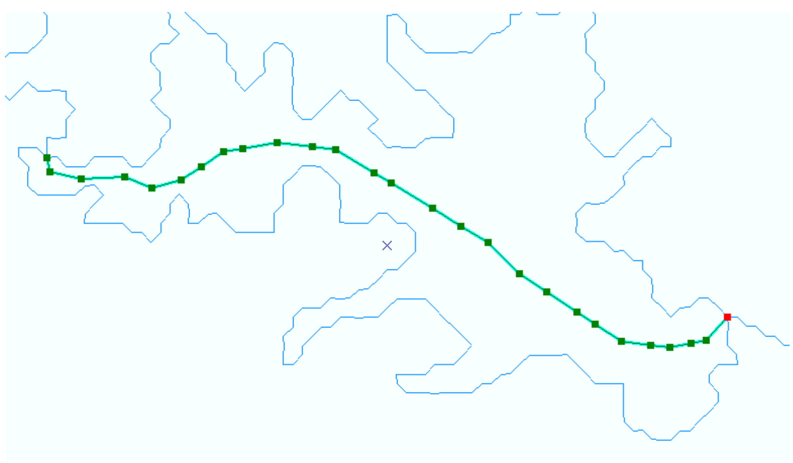
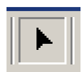


Figure E.7: Editing Centerline

B. Create Waterbody

Now we are going to take the lines that trace the shoreline of Canyon Lake and make them into a polygon to form a Waterbody. Make sure you have an active Edit session on

(9) Click on the Edit  button on the editor toolbar. Select the shorelines of the Canyon Lake by clicking on them using shift key. Use the **Editor/Merge** command to merge the two shorelines into a single feature.

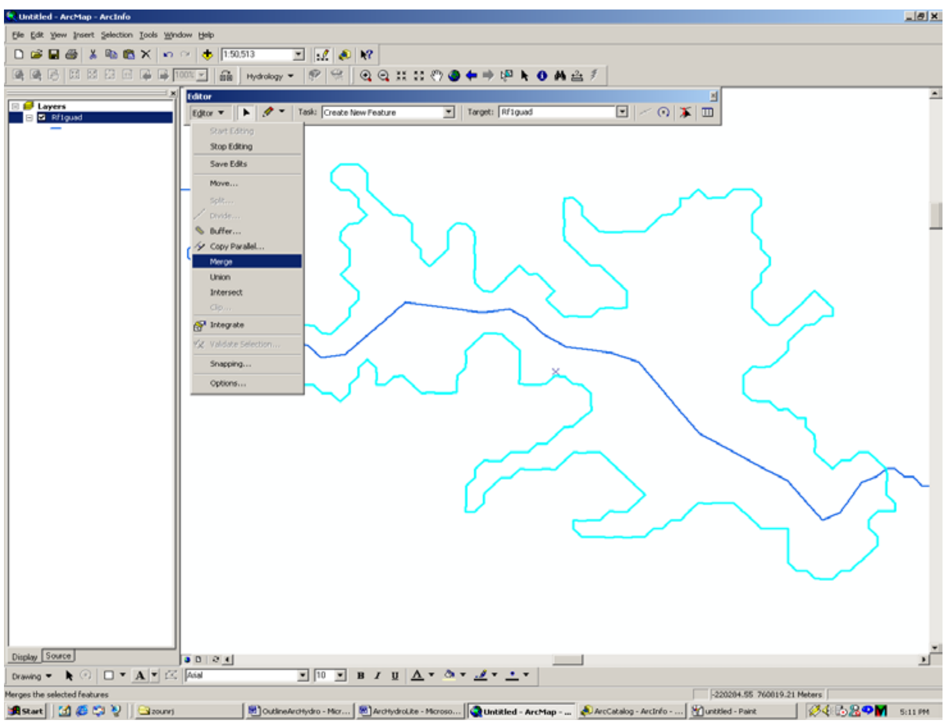


Figure E.8: Merging Files

(10) Go to **Editor/Save Edits** to save your editing and then Stop **Editing**.

(11) Select the merged waterbody. We are going export the selected data to create a new shapefile.

(12) Right click on Rf1guad and go to **Data/Export Data**. Export selected features and call the output file **Waterbody**. Add the exported data as a map to the layer.

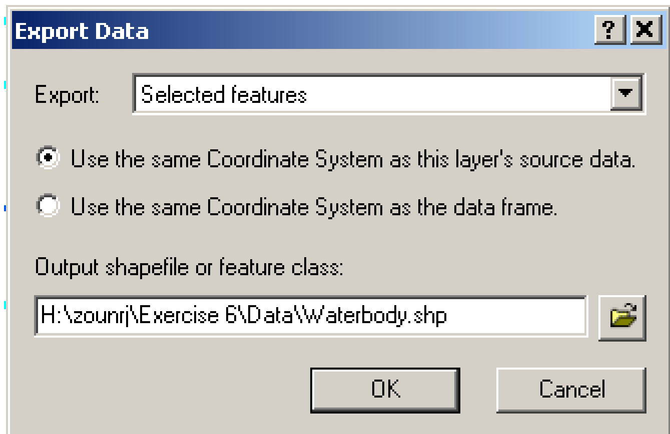


Figure E.9: Exporting Data

(13) Clear Selected features and turn off Waterbody.

(14) Go to **Editor/Start Editing**. Set the **Task** as **Modify feature** and Target as **Rf1guad** on the Editor Dialog box. Select the canyon reservoir shoreline and delete it. This creates a simple network through the lake rather than parallel paths around its edges.

(15) Go to **Edit/Save Edits** and **Stop Editing**. Exit ArcMap. You don't need to save changes.

(16) Now we are going to convert the shapefile Waterbody into coverage. We need to do this in ArcInfo. Go to ArcGIS/ArcInfo Workstation/Arc to start Arc.

(17) (You will need to execute the following sequence of Arc commands [without the explanations in square brackets]:

w [to see the set workspace]
w h:\….\data [change to the workspace where you have your Waterbody shape file]
w [check if you are in the right workspace]

shapearc waterbody waterbody_cov poly [change the shapefile Waterbody to a Arc/info coverage named waterbody_cov]
build waterbody_cov poly [create polygon attribute table]
clean waterbody_cov waterbody_cov # # poly [make sure that all the lines are properly closed onto one another in the polygon]

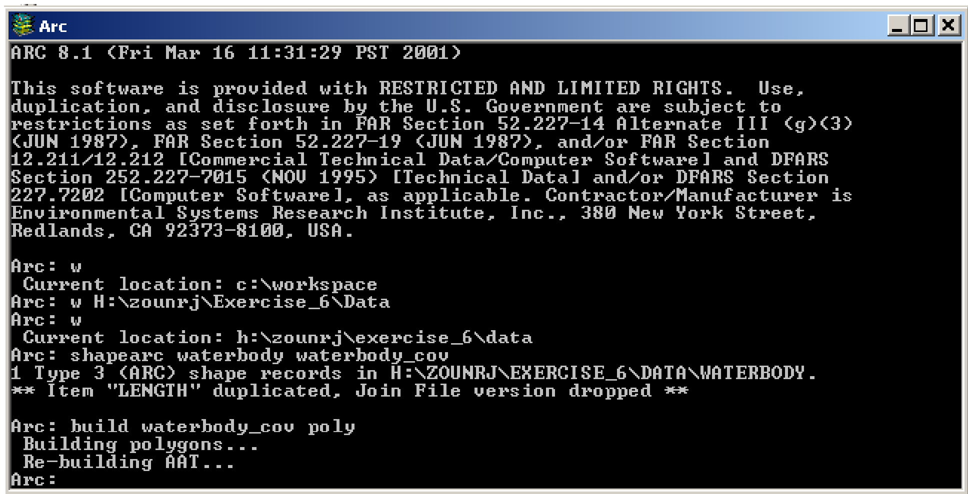


Figure E.10: Building Polygons in Arc

(18) Type ‘quit’ to exit Arc.

If you go back to ArcMap and add the newly created polygon coverage to the display, you’ll see that you have a polygon instead of the line that you had before:

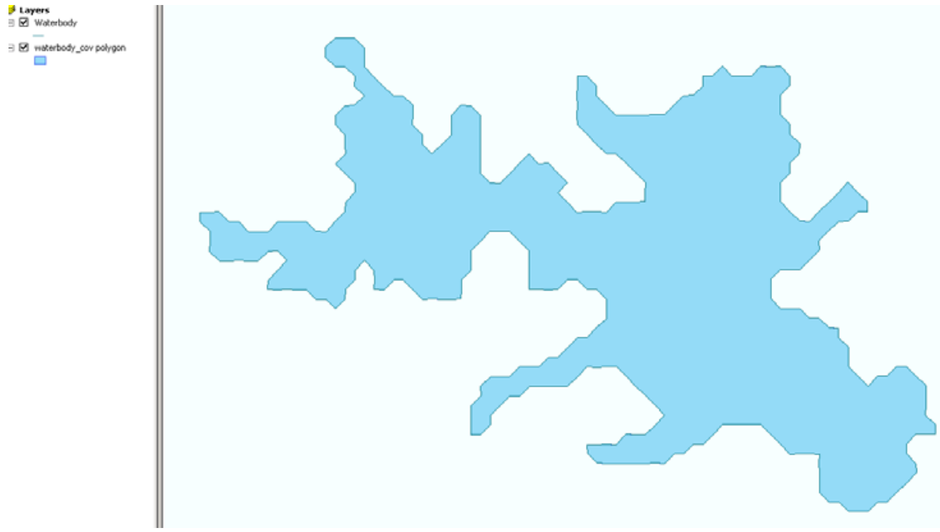


Figure E.11: Lake Polygon

C. Create network junctions

Now, we are going to create junctions for our network using the geometric network builder.

(1) Go to **Arc Catalog**, create a new geodatabase called **Guadalupe.mdb**, and import into it the shape file **hucguad.shp**, naming the feature dataset Guadalupe, and then import **rf1guad.shp** into this feature dataset. If you go to the Properties of the Guadalupe feature dataset, you'll see that it has the Albers projection of the Guadalupe data that you used in exercise 2 earlier in the course. The hucguad.shp file was imported here to make sure that the extent of the feature dataset was large enough to cover all the region of interest. It's really the rf1guad river reaches that we are interested in working with.

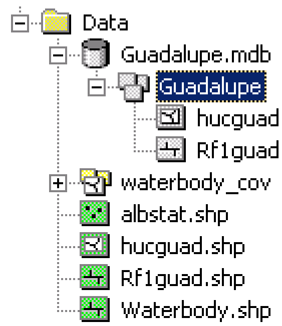


Figure E.12: Guadalupe Geodatabase

In the Guadalupe feature dataset, Create a new geometric network, Building with Existing Features, Select **Rf1guad** as the feature class and **Guadalupe_Net** as the network to be created. Do not use Complex Edges in the network. Do not snap features. Do not assign weights to the network.

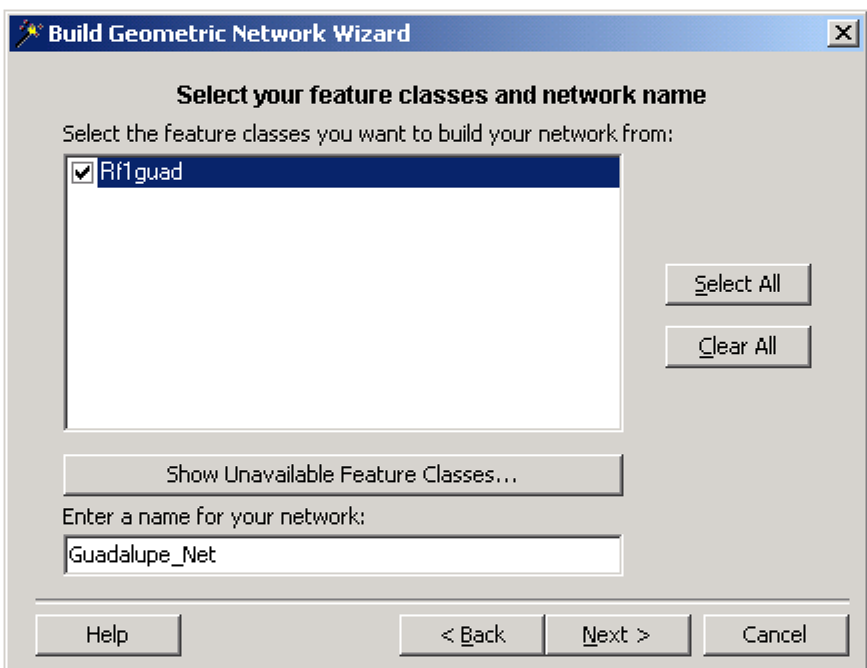


Figure E.13: Geometric Wizard for ArcHydro Lite

Once you've got the network created, you'll see that you've added a new feature class called **Guadalupe_Net_Junctions**, which is created during the network building process to link the lines in the Guadalupe network. These are called **ESRI Generic Junctions**. We are now going to transform them into Junctions for use in the Arc Hydro data model.

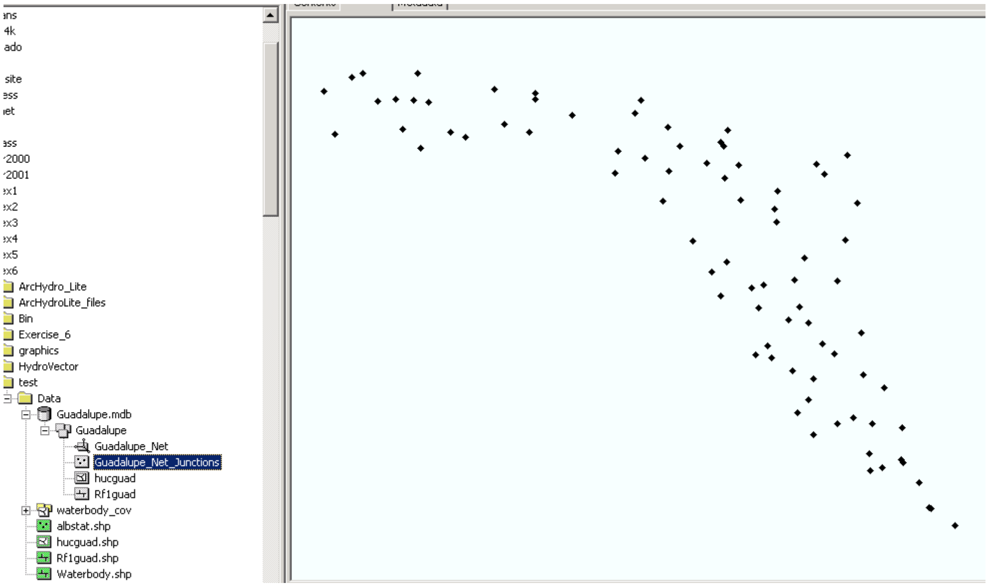


Figure E.14: Guadalupe Net Junctions

In Arc Catalog, right click on **Guadalupe_Net_Junctions** and use Export/Geodatabase to Shapefile to create a new shapefile called **Junctions.shp**

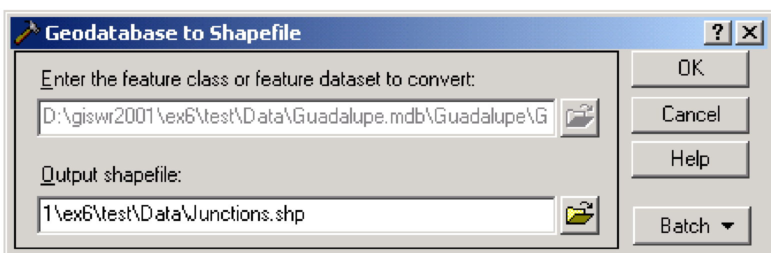


Figure E.15: Exporting Junctions

D. Create Geodatabase and Import Data

When applying the Arc Hydro schema, it works best when you have a geodatabase with feature classes already in it with the correct class names. So lets create that now.

(1) Open ArcCatalog. Right click on the data folder, press **New/Personal Geodatabase**. Call the new geodatabase **ArcHydro**.

(2) Now, you will import your shapefiles into the geodatabase. Right-click on your geodatabase and press **Import/Shapefile to Geodatabase**.

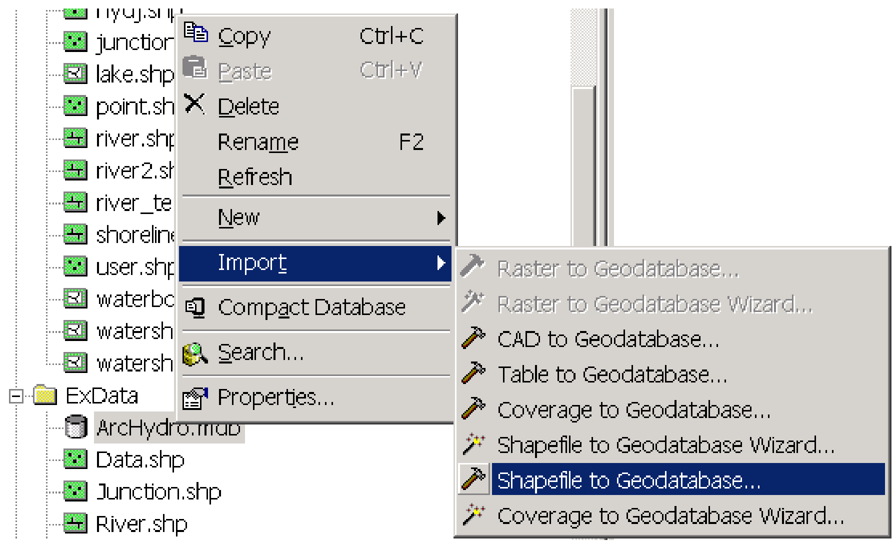


Figure E.16: Importing Shapefiles to Geodatabase

You will be importing all your shapefiles. First, you must navigate your data folder in the Input shapefile box. Choose the hucguad shapefile, since this file has the largest extent. Type in ArcHydro as the name for your new feature dataset and Watershed as the name of new feature class. You will repeat the above process until the remaining shapefiles are added. It is important to note that each feature class should have the name assigned to it in Arc Hydro Lite. See the table below for the corresponding information. *Do not import the waterbody.shp file*. We'll do that next using the Coverage form we created earlier.

Shapefile	Arc Hydro Lite Feature Class
hucguad	Watershed
albstat	HydroPoint
Junction	HydroJunction
Rf1guad	HydroEdge

Table E.2: Shapefile Feature Names

(2) We will import the coverage file **Waterbody_cov** to the Geodatabase. Right-click on your geodatabase and press **Import/Coverage to Geodatabase.**

(3) Navigate your data folder in the Input Coverage box. Choose the **waterbody_cov** coverage file as input coverage. Go inside the Waterbody_cov folder and select **polygon** as an existing feature class in the coverage and ArcHydro as feature dataset. Name the new feature class **Waterbody**.

The input screen will look like this when you are done selecting your data:

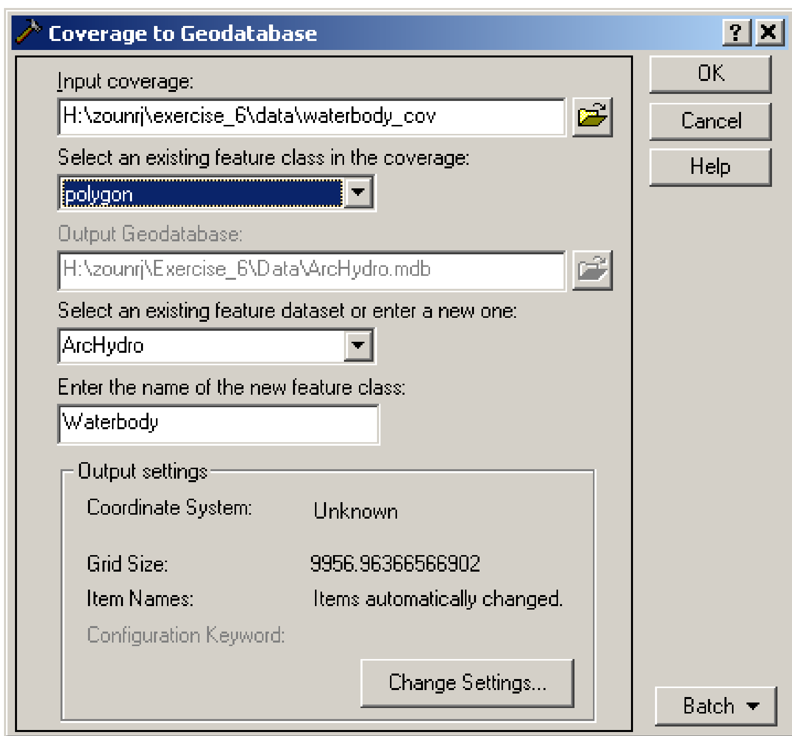


Figure E.17: Importing Coverages to the Geodatabase

(4) Press **OK**.

E. Create Geometric Network

(1) The next step is to create your geometric network, HydroNetwork. Right-click on the feature dataset Arc Hydro and press **New/Geometric Network**. Press **Next** on the first screen and then select **Build a geometric network from existing features** on the second screen. In third screen, select the HydroJunction and HydroEdge files to be part of the network.

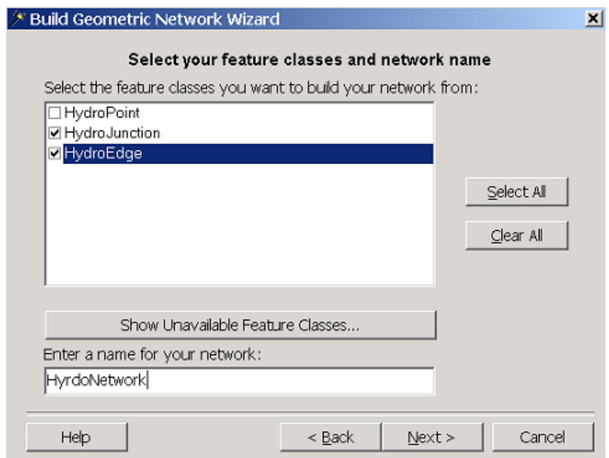


Figure E.18: Building Geodatabase Network

Also, name the network **HydroNetwork**. This is an important step since the schema will only accept a network with the same name as in the model description. Click **No** to enable all features in the network on the fourth screen. Click **Next**. You will be asked if your network has complex edges, select **Yes** and HydroEdge. Click **Next**. In the sixth screen, select **No** since your features do not need to be snapped and click **Next**. You will be asked if your network will have sources or sinks, in the seventh screen. Click **Yes** and select **HydroJunction**.

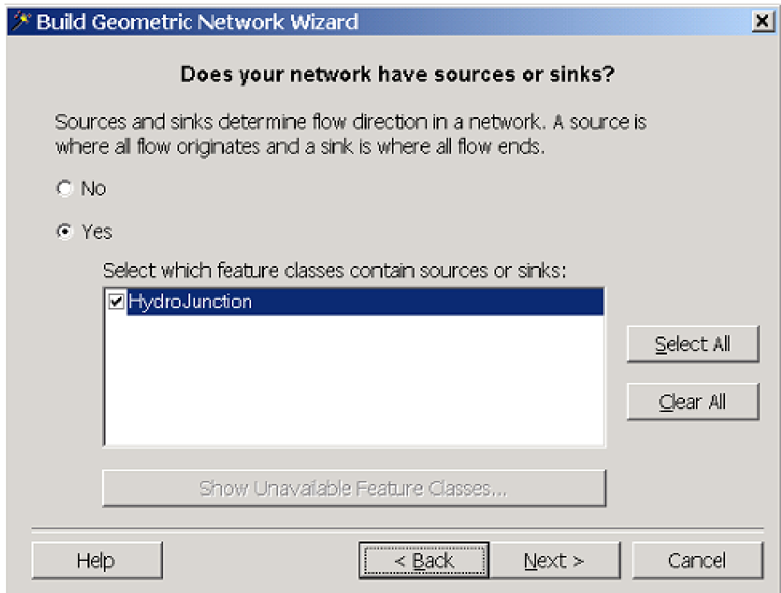


Figure E.20: Source or Sink Network Screen

Press **Next**. You will be asked if you would like to assign weights to the network, answer **No** and click **Next**. Finally, press **Finish** to create the geometric network.

Your geodatabase should look like this.

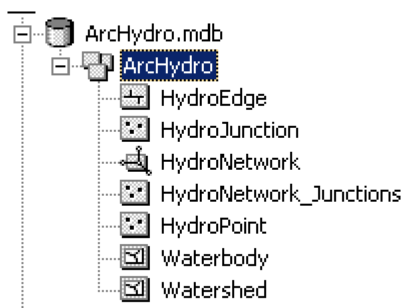


Figure E.21: Geodatabase with Network

In summary, the important factors to remember when preparing your data for schema application are the following:

1. Assign the appropriate Arc Hydro Lite name to your feature datasets, feature classes, and feature attributes allows for automatic recognition by the schema creation wizard.

2. Create the network (HydroNetwork) with HydroEdge and HydroJunction. HydroEdge will be a complex edge and HydroJunctions will contain sinks. Each point in the HydroJunction does not have to be a sink but all HydroJunctions will have the option of becoming a sink.

Ok, super duper, you've now created your input data sets!

In the next part of the exercise, we'll apply the Arc Hydro Lite schema and the Arc Hydro tools.

Appendix F: VBA Code for Isolating USGS Gage IDs

Written by Tim Whiteaker, CRWR

This code is hardwired for the current table structure of the USGSGages feature class and the ArcMap document to which it was applied.

```
(General)                                    AssignGageID
Option Explicit

Public Sub AssignGageID()
  Dim pDoc As IMxDocument
  Dim pMap As IMap
  Dim pFLayer As IFeatureLayer
  Dim pFCursor As IFeatureCursor
  Dim pFeat As IFeature
  Dim pFields As IFields

  Set pDoc = ThisDocument
  Set pMap = pDoc.FocusMap
  Set pFLayer = pMap.Layer(9)
  Set pFCursor = pFLayer.Search(Nothing, False)
  Set pFeat = pFCursor.NextFeature

  Dim pEditor As IEditor
  Dim pUID As New UID
  pUID = "esriCore.Editor"
  Set pEditor = Application.FindExtensionByCLSID(pUID)
  If pEditor.EditState <> esriStateEditing Then
    MsgBox "Please start editing first"
    Exit Sub
  End If

  pEditor.StartOperation

  Dim sNewVal As String, sOldVal As String
  Dim lOldFld As Long, lNewFld As Long
  lOldFld = 6
  lNewFld = pFLayer.FeatureClass.Fields.FieldCount - 1

  Do Until pFeat Is Nothing
    sOldVal = pFeat.Value(lOldFld)
    sNewVal = Right(sOldVal, Len(sOldVal) - 4)
    pFeat.Value(lNewFld) = sNewVal
    pFeat.Store
    Set pFeat = pFCursor.NextFeature
  Loop


  pEditor.StopOperation "yay"
End Sub
```

Figure F.1: Code for Gage Isolation

Bibliography

Bezivin, J. and P.A. Muller, 1998. *The Unified Modeling Language <<UML>>'98: Beyond the Notation*. Berlin: Springer-Verlag.

Booch, Grady, James Rumbaugh, and Ivar Jacobson, 1999. *The Unified Modeling Language User Guide*. Reading: Addison-Wesley Longman, Inc.

Davis, K. M. and D. R. Maidment, 2000. *Object-Oriented Modeling of Rivers and Waters in Geographic Information Systems.* CRWR Online Report 00-07. Internet Site: http://www.crwr.utexas.edu/reports/pdf/2000/rpt00-7.pdf.

DePinto, J.V. et al, 1994. An Approach for Integrating GIS and Watershed Analysis Models. *Microcomputers in Civil Engineering*, 9, 251-262.

Federal Highway Adminstration (FHWA) 2002. *The Arc Attribute Table*. Internet Site: http://www.fhwa.dot.gov////planning/nhpn/docs/aat.html.

Goodchild, M.F., and K.K. Kemp, eds., 1990. *NCGIA Core Curriculum in GIS.* National Center for Geographic Information and Analysis, University of California, Santa Barbara CA. Internet Site: http://www.geog.ubc.ca/courses/klink/gis.notes/ncgia/u23.html#UNIT23.

ITS UT Austin, 2002. *ITS Database Services*. Internet Site: http://www.utexas.edu/cc/database/datamodeling/dm/intro.html.

Maidment, D. R. 2001. *GIS Hydro 2001*. http://www.crwr.utexas.edu/gis/gishydro01/GISHydro2001.htm.

Maidment, D.R. 2002, New Tools for Applying GIS in Water Resources (Draft).

Montick, T., 1995. *What is Object Oriented Software?* Internet Site: http://catalog.com/softinfo /objects.html

USGS, 2000. *NHD FTP Site*. Internet Site: ftp://edc.usgs.gov/pub/data/nhd/fod_cache/228100/arc/.

U.S. Environmental Protection Agency, 1998. *Better Assessment Science Integrating Point and Nonpoint Sources, BASINS Version 2.0: Users*

Manuel. U.S. Environmental Protection Agency, Office of Science and Technology, Washington, D.C.

Whiteaker, T, 2001. *A Prototype Toolset for the ArcGIS Hydro Data Model*. Thesis. The University of Texas at Austin.

Zeiler, M., 1999. *Modeling our World: The ESRI Guide to Geodatabase Design.* Redlands: Environmental Systems Research Institute, Inc.

Smartdraw, 2002. *How to Draw UML Diagrams*. Internet Site: http://www.smartdraw.com/resources/centers/uml/uml.htm.

Vita

Kristina Antonia Schneider was born on November 18, 1977 in Washington, D.C. to Anna Schneider and Hans Alfred Schneider. After graduating in 1995 from Paint Branch High School in Burtonsville, Maryland, she entered the University of Maryland in College Park, Maryland. She studied abroad in London, England during the spring of 1997. In the spring of 2000, she studied at the Sprachinstitut in Tübingen, Germany followed by an internship in Weinheim, Germany. She received her Bachelor of Science in Civil Engineering graduating Magna Cum Laude in December 1999. In May 2000, she received her Bachelor of Arts in German Language. In August 2000, she entered the Graduate School at the University of Texas.

Permanent address: 2525 Countryside Dr.
Silver Spring, MD 20905

This dissertation was typed by the author.