

WHAT MAKES A GOOD GIS LAB EXERCISE?

Robert N. Martin

ABSTRACT

Having taught GIS for over fifteen years in a lecture/laboratory format, I have developed a series of laboratory exercise that provide a good lab experience for the students. By a good lab experience I mean a series of labs that challenge the students to think and labs that are related to each other. In this presentation I will discuss the major types of data and methods of analyses that students work with over a semester. In this process specific exercises will be discussed. It is hoped that feedback will allow the development of a better laboratory experience for the students.

STATEMENT OF PROBLEM

A basic problem faced by any instructor teaching GIS is to develop a series of laboratory exercises for students that are relevant and challenging. The laboratory exercise problem is compounded by the following:

- a. that for most of the students this is their first exposure to working with GIS;
- b. to develop a really challenging series of labs the students should have taken the entire lecture section of the course prior to entering the laboratory; and
- c. the laboratory exercises should be interesting to students from many different academic disciplines.

The software, data sets, and laboratory exercises I use in teaching *GEG 274 – Introduction to GIS* at Kutztown University have evolved over the last decade and one-half. From OSU-MAP to IDRISI to ArcView to the present ArcGIS software and extensions the capability of GIS software to analyses problems has grown tremendously. From relatively low resolution data sets covering limited areas to very high resolution data sets covering vast regions of this planet that may be downloaded with ease, the ability of GIS to visualize and interpret spatial distributions continues to expand. From a literal handful of geography majors at the beginning the student enrollment has expanded to sixty students in a lecture divided into three lab sections of twenty. These students major in geography, geology, biology, environmental science, marine science, anthropology, business, psychology and several other fields. From a course focused on building databases the *Introduction to GIS* course at Kutztown University has developed into one with a focus that is moving to methods of analyzing spatial data. The challenge in the lab is to continue to adapt, revise, and expand the type of exercises that provide the students with a sense of the capabilities of GIS systems.

What I will be presenting in this paper is how I have met these challenges. This is not the be all to end all set of GIS lab exercises but rather a guide to other who will be developing labs for their students. The focus therefore is on what constitutes a good laboratory experience based on my experiences in the classroom and lab working with students from many fields. In the following section of this paper, the general database model for the exercises will be discussed and the reasons for this type of model will be identified. After a discussion of the reasons for the selection of the study area, the basic data elements with the database will be identified as well as their reasons for inclusion. An overview of the exercises, methods of analyses, and how they are integrated will then

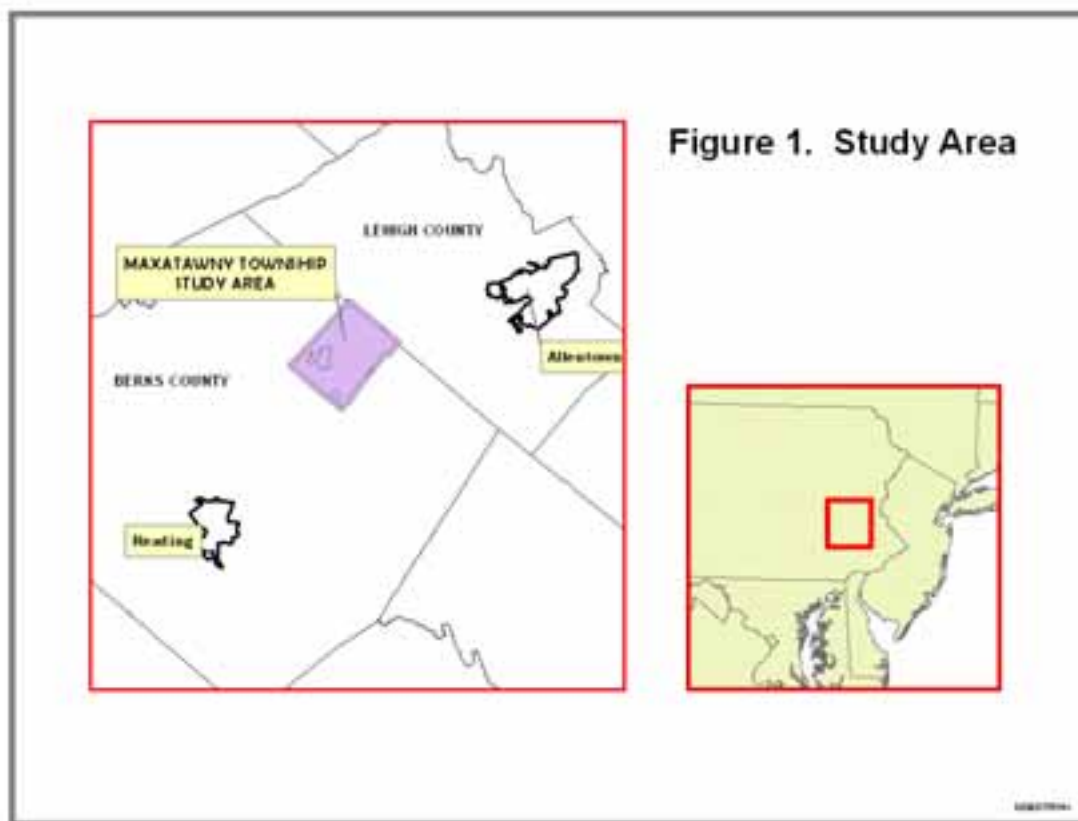
be discussed. The paper will conclude with future directions and new exercises being considered.

GENERAL DATA MODEL

A land-use database model is the general model used for the laboratory exercises. There are two major reasons for using this data model. First it provides for the inclusion of a wide variety of data and data types. This variety of data will provide, hopefully, a mix of data that will be of interest to the students or at least pique their curiosity. The second reason for this choice of data model is that it contains many elements used in planning agencies of local government units. Since many of the students will be taking internships with these types of agencies, familiarity with the types of data they will be using and methods of analysis will be helpful.

IDENTIFICATION OF STUDY AREA

The study area for the exercises is Maxatawny Township, Berks County, Pennsylvania, see Figure 1. The township covers 6,736 hectares or 16,644 acres in the northeastern portion of the county approximately halfway between the cities of Allentown and Reading.



There are several factors for the selection of this study area. First, Kutztown University is located within this township and most of the students attending the university are from within 75 miles of the campus. Since most of the students are taking the course during

their junior or senior year they are familiar with the area. In addition, because the study area is local and most students have some form of personal transportation, they can field check information with relative ease. Many students quickly realize that they do not know the area as well as they thought. The final reason for the selection of this study area is that four 7.5 minute topographic maps are needed to cover it. This highlights one of the basic “truths” of GIS that data is almost never found in a single map, a single projection, a single datum, or a single resolution.

TYPES OF DATA AND DATA SOURCES

As preparation for the lab exercises, in early course lectures I discuss the types of spatial information that may be included in a GIS database for a land-use planning, focusing on the types of analyses that can be done. A preliminary assignment is to find what data may be available on the web, specifically at state and local internet sites. The data feature classes used to build the database are listed in Table 1. As indicated in the table, there are a number of features classes that must be developed by the students.

TABLE 1 - DATABASE ELEMENTS

[Student Processed Data Shown by **]

Data Feature Class	Data Feature Class
Census Block Groups	State Roads**
Census Tracts	Streams
Contour Lines	Study Area Polygon**
Floodplains	Water Bodies
Geology	Water Sheds
Local Roads**	Wells
Minor Civil Divisions**	Wetlands
Property Parcels**	Woodlands
Point Elevations	Zip Code Zones
Quad Map Sheet Boundaries	
Railroads	DEM at Level 1 and Level 2**
Soils**	DOQQ
Springs	DRG

These data feature classes include all vector data types, point, arc, polygon, as well as raster data from a variety of sources and at various resolutions. The feature classes also cover a broad range of socio-economic, topographic, physical geography variables. To speed the development of the geodatabase, many of the feature classes have been projected and/or geographically transformed, if necessary, and clipped to the study area. The base map projection for the study area is UTM Zone 18N, NAD 1983.

EXERCISES AND METHODS OF ANALYSIS

The fifteen lab sessions are divided into three parts. An introduction to the ArcGIS software takes the first two lab sessions. These two lab sessions primarily focus on the components of ArcCatalog and ArcMap including ArcToolbox. The students are encouraged to work through several ArcGIS tutorials stored on their computers outside of the laboratory period. These tutorials are *Using ArcCatalog*, *Using ArcMap*, and

Getting Started with ArcGIS tutorials. I have also found it very helpful in the long run to introduce the students to built-in help function to answer their questions.

The second part of the laboratory periods consists of twelve lab sessions working with the land-use planning database model. During the twelve lab sessions the students work on the eight exercises listed in Table 2. During last laboratory period the students take their laboratory final exam, which constitutes the third part of the lab experience.

TABLE 2 – GIS LABORATORY EXERCISES

Building a GIS Database - I
Building a GIS Database - II
Organizing the Geodatabase and Error Checking
Working with DEM data
TIN versus DEM
Geoprocessing – Dissolve, Intersection and Union
Spatial Statistics
Hydrologic Modeling

In the following sections each of the exercises listed in Table 2 will be discussed briefly. The objectives of the exercises and how these objectives are accomplished will be discussed. The discussion will also look at the relationship between the exercises as well as the elements that I believe make it a good laboratory experience.

EXERCISE 1

The '***Building a GIS Database - I***' exercise introduces the students to the data files in Table 1 that they will be working with over the next two and one-half months. In this exercise the students create and begin to organize a personal geodatabase with feature datasets and feature classes using *ArcCatalog*. After creating the personal geodatabase, the students create three feature datasets to organize the data that will be used during the term. The feature datasets are **BaseStudyArea**, **LandCover**, and **Transportation**, see Table 3 for more detail.

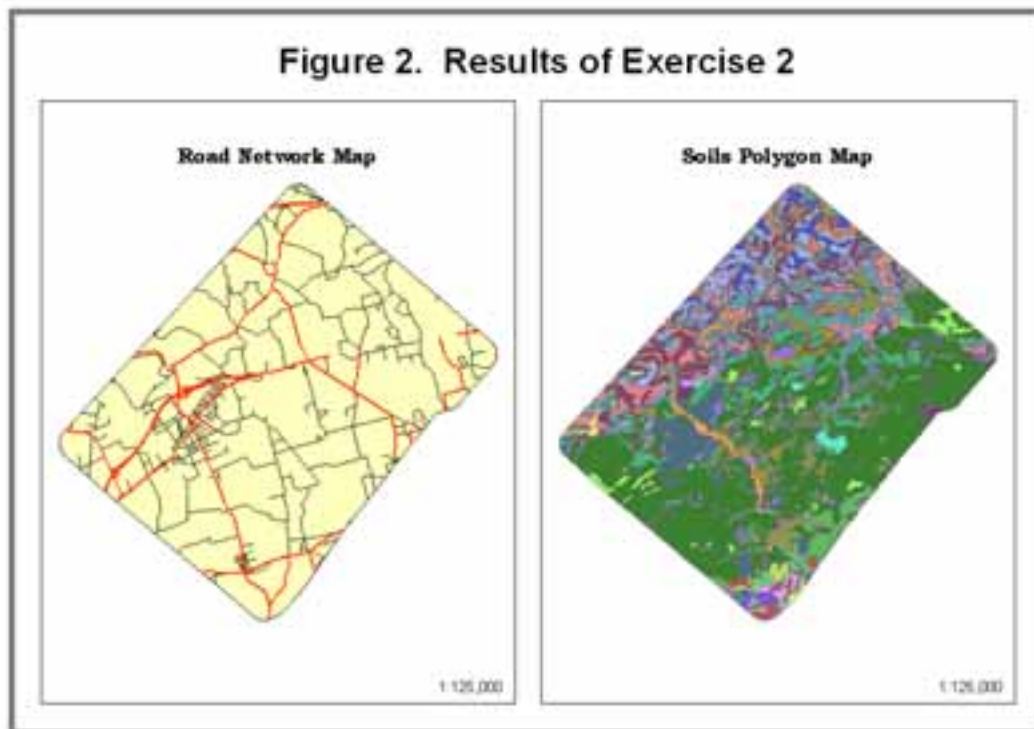
Using the state-wide municipalities' shapefile in *ArcMap*, the students select Maxatawny Township from the attribute table and then select the surrounding municipalities by location. These selected municipalities are exported as a shapefile into a temporary folder. Using *ArcToolbox* within *ArcCatalog* the students project the file into UTM coordinates from geographic coordinates. The projected political units are brought into *ArcMap*. Having selected Maxatawny Township and the boroughs within the township, the students create a 500 meter buffer around the selected objects. This is the ***Study_Area*** cookie cutter that they students will be using in the remaining exercises either for clipping features or providing a reference frame for the analyses. This is stored as a feature class within **BaseStudyArea** dataset of the geodatabase. It is also immediately used by the students to clip the political units into a feature class within the geodatabase called ***MinorCivDiv***, minor civil divisions.

This makes a good lab experience because the students are starting from scratch and the results of this exercise will have impacts on further analyses. They have to create the study area by selecting features through the use of attributes tables and selection by location. They have to project the file into the appropriate projection and use the result to create a buffered study area file that will be used to clip the political units. If the selection is done wrong, the study area will be incorrect as will be the minor civil divisions.

EXERCISE 2

The second exercise, *Building a GIS Database – II*, reinforces the tasks learned in the first exercise, specifically projecting data and clipping files to the study area. In addition it expands the skill set to new activities of joining geographic objects with attribute data, merging geographic data sets, creating new attribute fields and calculating values of attributes.

After opening *ArcMap* the students add the projected *Study_Area* feature class. This sets the projection for the data frame, which is one of the learning points of the exercise. The road networks files, which are in geographic coordinates, are added to the map and clipped to the study area. These clipped files are saved as feature classes within the **Transportation** dataset within the geodatabase. These files will be used in the error checking exercise which follows as well as reference files in maps.



Starting with a new map, the *Study_Area* feature class is added along with the soil files from the SSURGO database for Berks and Lehigh counties. These files will be used to make the students familiar with the use of joins and merges, as well as creating attribute fields and calculating values. The county soil polygon files are clipped by the students to the study area and the polygons are joined with map unit information tables as well as crop yield data tables. These respective county files are then merged into a single *Soils* feature class within the **LandCover** dataset. Attribute fields for hectares and acres are added and calculated. Based on the calculated values students must answer a series of questions about the soils within the study area.

As mentioned this exercise reinforces the students experience with map projection and clipping. By using the join function, merging and adding attributes their skill set is improved. These are skills they will need in later exercises and in the ‘real world’.

EXERCISE 3

‘Organizing the Geodatabase and Error Checking’ exercise has two major objectives. First, the students must complete the building of the geodatabase for the study area. The second objective is to make the students familiar with the importance of error checking.

The first objective is accomplished by having the students import the most of the features listed in Table 1 into the feature datasets listed in Table 3 using the appropriate file names. The exception to this importation process are the study area, minor civil divisions, soils, and the local and state roads features that the students have already created in Exercise 1 or 2. The property parcels and assorted raster files are developed or used in later exercises.

To accomplish the second objective the students need to import an ASCII file into a raster file within the geodatabase and define the projection for the raster file. The file that is imported is the clipped mosaicked digital raster graphic (topographic) map for the study area. Using the USGS DRG color palette the students adjust the colors of *DRGclip* and saves the results as a layer file within the geodatabase. Then using the *DRGclip* as a background image the students add each of the feature classes that were imported and determines if the DRG would be an appropriate base for error checking. If the DRG is appropriate for checking, the students are to identify any errors of commission or omission. In addition, they are to identify additional features that could be added based on the DRG file. This error checking process is then repeated using the digital orthophotos quarter quadrangles, DOQQ’s, for the study area. While the majority of the students have worked with topographic maps and are familiar with air photos, they soon realize that features such as wells, census units, watersheds, etc. do not appear on these sources.

In addition to practicing organizing the geodatabase into logical datasets, the major learning outcome of this exercise was for the students to realize that you should not assume that all the data that you import from data providers is correct or up-to-date.

TABLE 3 – GEODATABASE ORGANIZATION

BaseStudyArea	LandCover	Transportation
BlockGroups	Contours	LocalRoads
CensusTracts	Floodplains	Railroads
MinorCivDiv	Geology	StateRoads
QuadMapSheets	PointElevations	StRdAdmin
StudyArea	Soils	StRdTrafic
ZipcodeZones	Springs	
	Streams	
	WaterBodies	
	WaterSheds	
	Wells	
	Wetlands	
	Woodlands	

EXERCISE 4

In '*Working with DEM Data*' the objectives are to give the students experience working with digital elevation models by using a mask to 'clip' raster files, and to create hill shade and slope models of the landscape. They are also to note the differences in working with level 2 DEM data versus level 1 DEM data.

Working with digital elevation models that have been mosaicked and projected to the study area projection, the students have to extract the level 1 and level 2 DEM data for Maxatawny Township using *Study_Area* as a mask. Since the students have already done clipping using polygons on feature classes, this is an easy transition to a different data format. The students must record a variety of measures comparing the differences between 10-meter versus 30-meter DEM files as they extract the data, create hill shade and slope files. Specifically, they must record the number of rows and columns, minimum and maximum elevations, and the elapsed time to do the analyses.

This exercise provides the initial experience for the students to do analysis with raster data rather than just viewing the data as images. The raster models developed in this exercise will be used in Exercises 5 and 8.

EXERCISE 5

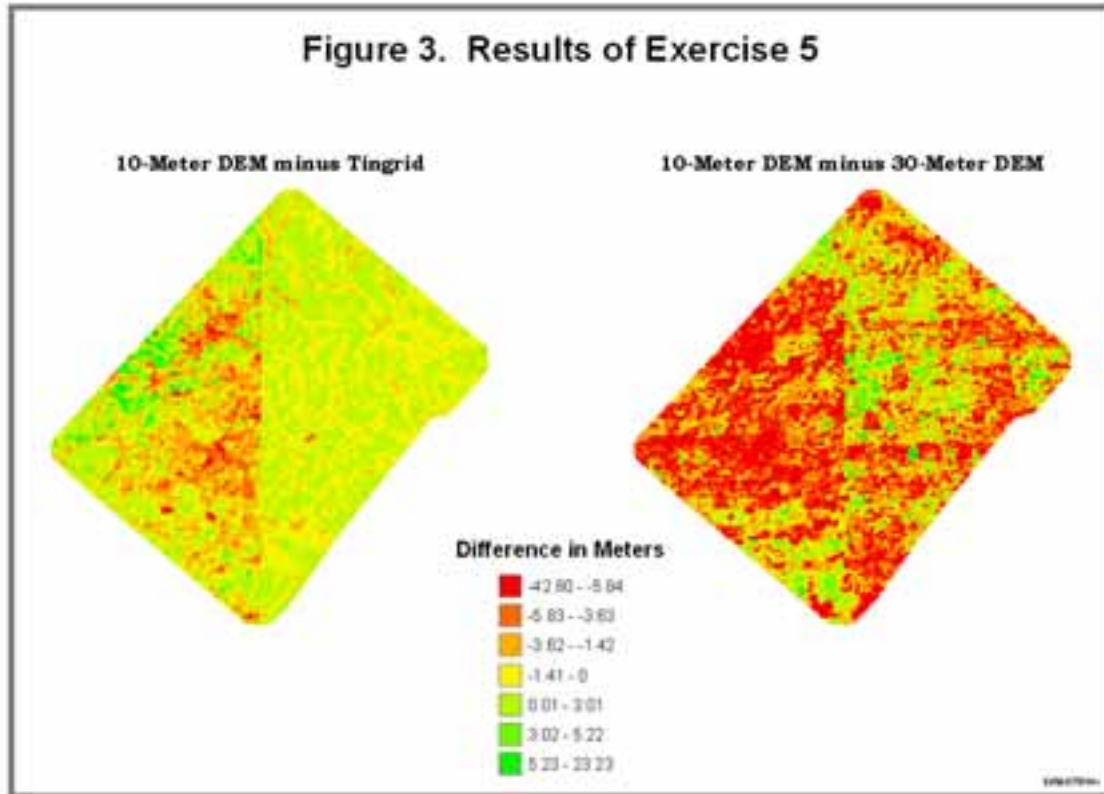
The '*TIN versus DEM Data*' exercise has the students comparing the digital elevation models they created in Exercise 4 with a triangulated irregular network developed in this exercise. One of the major objectives is for the students to realize how different models of the same landscape can have different results.

Using the *Contours*, *PointElevations*, and *Streams* features classes the students placed in the geodatabase, they create a triangulated irregular network, *TIN10m*, to represent the surface of the study area. An attribute field has to be added to the *PointElevations* feature class to store the elevation in meters. In addition, a missing value for *PointElevations* has to be determined by using the *DRGclip* created in Exercise 3. Once

the TIN is created it is converted to a raster, *Tingrid*, with the same resolution as the *Maxt10m* file created in Exercise 4. Map Algebra is then used for the following operations.

$$Tin10Diff = [Maxt10m] - [Tingrid]$$

$$Maxt10_30 = [Maxt10m] - [Maxt30m]$$



The students map *Tin10Diff* and *Maxt10_30* using the same standard deviations classification scheme and have to explain the possible reasons for the variations in the differences in elevation between the various models. By comparing the maps the students can see the objective visually.

EXERCISE 6

In Exercise 6, '*Geoprocessing – Dissolve, Intersection and Union*' the impacts of different types of geoprocessing are introduced to the students. Again, the students have to join attribute tables, add attribute fields, and calculate values as was done in previous exercises.

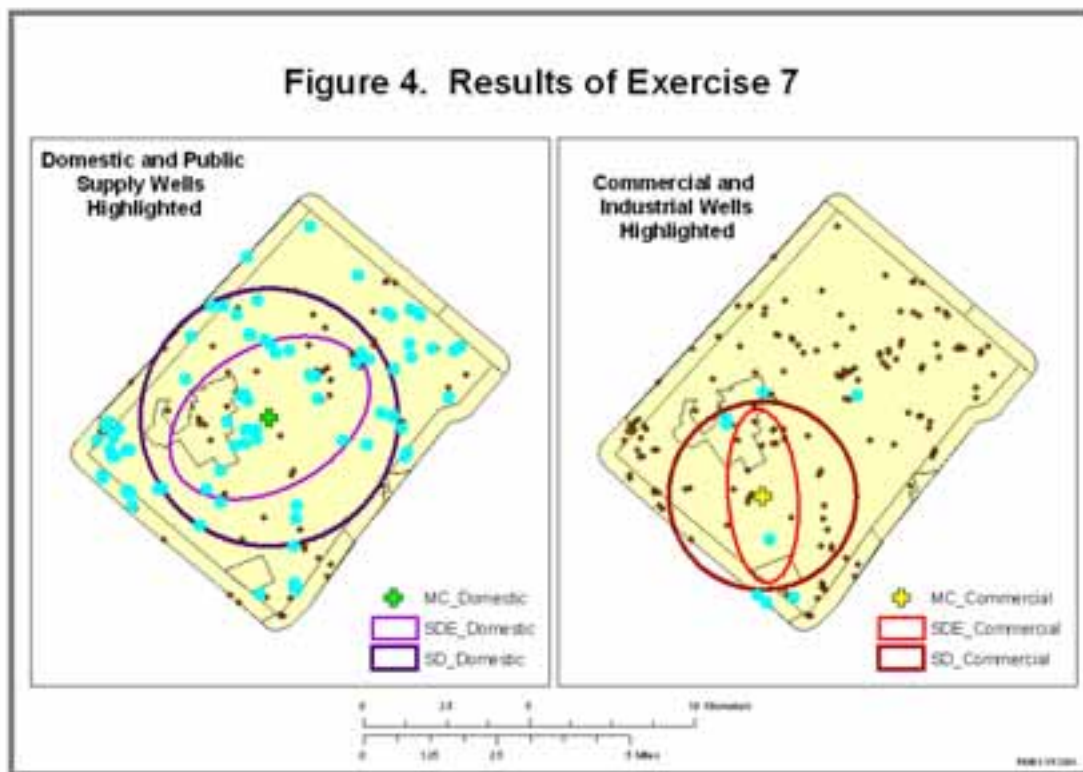
The students import a projected and clipped file of tax parcels for the study area and create a feature class within the **LandCover** dataset of the geodatabase. A dissolve operation is done based on the Property Identification Number, PIN. This unifies tax

parcels that are split along minor civil division boundaries. The dissolved tax parcels are then joined with the original attribute data using the PIN to attach all attributes. The students then calculate the acres, hectares, and value per acre. After intersecting the floodplains with the parcels the students must calculate the values of the parcels within the floodplains. Then a union of the parcels and floodplains is produced by the students and compared to the intersection of the features.

This exercise brings home the difference between doing an intersection, a Boolean AND, versus doing a union, a Boolean OR, on data features. It also introduces the students to the process of prorating data to use in future calculations.

EXERCISE 7

'Spatial Statistics' are the major focus of Exercise 7. In previous exercises the students have used standard statistics to note the mean and standard deviation of an attribute. This exercise shows the importance of spatial statistics, i.e., the importance of location in the description of a distribution.



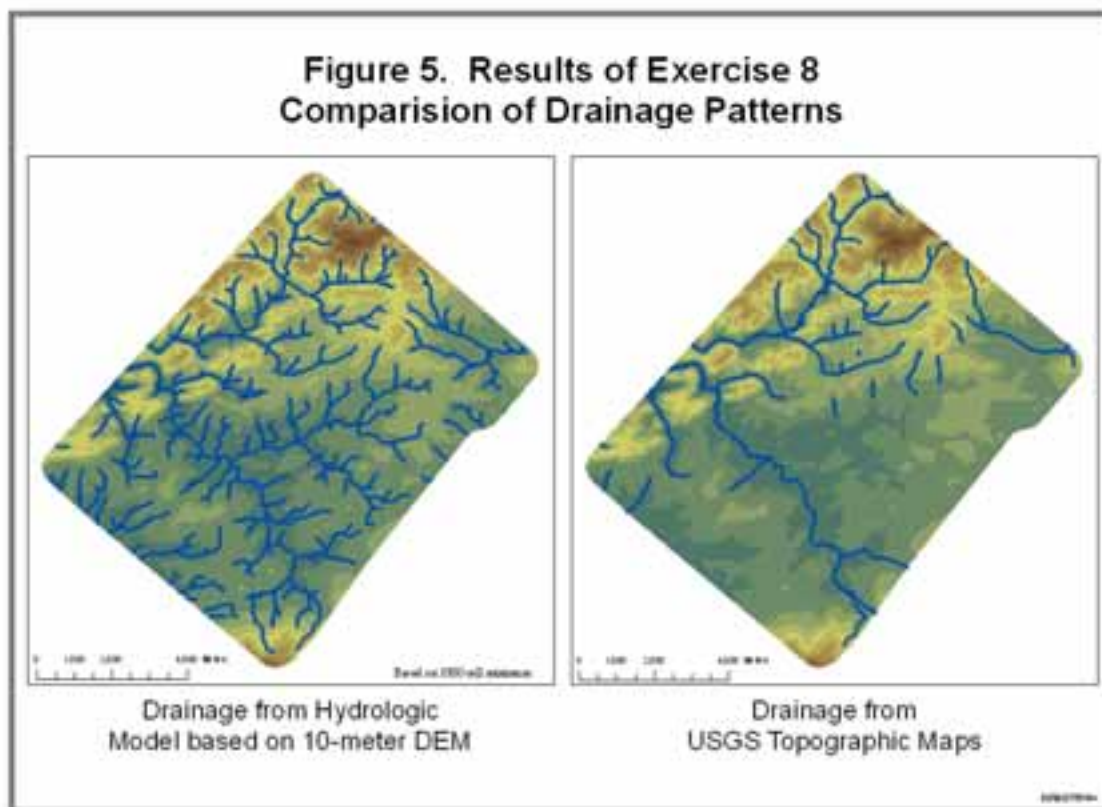
The students have to use measures of spatial central tendency to describe spatial distributions of point and arc features. Specifically they measure the mean center, standard distance and standard deviational ellipse for two sets of water wells data for the study area, domestic/public-use versus industrial/commercial. Additionally they have to

calculate and compare the linear directional mean for all streams within the study area versus the statistic for a single stream network within the study area.

The reason that is exercise is 'good' is that most of the students have had a standard statistics course but they have never had to measure a spatial distribution. This shows the students to power of spatial analysis based on readily available data.

EXERCISE 8

'Hydrologic Modeling' is the last laboratory exercise. Using a 10-meter digital elevation model, the *Maxt10m* created in Exercise 4 or the *Tingrid* created in Exercise 5, the students build a hydrologic model for Maxatawny Township. The students have to fill sinks, determine flow directions and accumulations, create watersheds, and model streams based on flow accumulations. They then compare the hydrologic model stream patterns with the topographic map stream patterns and provide reasons for the differences that exist, see Figure 5.



This exercise provides a good example of how the geology of the region can affect the surface pattern of streams. Specifically, the geology of the center portion of the study area is limestone and has numerous sinkholes which are a constant problem within the area. The students always seem to enjoy this project. From an abstract 'grid' of elevation they can create streams and watersheds that are relevant to analyzing run-off of chemicals from farm lands or industrial sites.

CONCLUSION AND FUTURE DEVELOPMENTS

This paper has attempted to show how an integrated data set for a single study area with a set of exercises can create an interesting and challenging set of experiences in a GIS laboratory setting. There are several areas in which I am developing additional exercises. An exercise using Network Analyst is being developed which will use the road network of the township to look at alternative routes to campus. A variation on this will be a research project for the students to use campus sidewalks to identify handicap access paths across the campus. Another exercise will be introduced to expand on the spatial statistics introduced in Exercise 7 and will focus on pattern analysis using a variety of techniques and statistics.

AUTHOR IDENTIFICATION INFORMATION

Robert N. Martin, Ph.D.
Professor and Chair
Department of Geography
Kutztown University
Kutztown, PA 19530
Phone: 610-683-4365
Fax: 610-683-4941
Email: martin@kutztown.edu