

# **Applications of GIS Technology in Bio Sciences**

by

Naga Deepa Vuppaladadium

An Internship Report Presented in Partial Fulfillment  
of the Requirements for the Degree  
Master of Science

ARIZONA STATE UNIVERSITY

May 2003





## ABSTRACT

Most natural resource and healthcare issues have been found to have a significant spatial component, such as the distribution of land use in a watershed, proximity of human populations to a contaminated site or accident, regions affected by spread of infectious diseases or the degree of habitat fragmentation in an ecosystem. In order to accommodate this geographical aspect, a cartographic method namely Geographic Information System (GIS) technology has emerged to aid in visualizing, analyzing, and modeling data for management, and problem solving.

Long-term outcomes of connecting the two seemingly disparate fields of Bio-Informatics and GIS fields are better modeling, better monitoring of diseases such as rabies in wildlife populations; improvements in agricultural productivity; and reductions in the negative environmental impacts of some agricultural activities. As an intern in Center of Environmental Studies (CES) department in Arizona State University, I have learnt to implement and apply the GIS technology in several areas of agriculture, archeology, paleontology and ecology.

This paper gives an overview of GIS applications, the methodology and detailed description of various projects that I covered as an intern. Some of the several projects that I dealt with while working in CES, involved the study of enhancing the yield of crops by developing production models, cost analysis of flow of prehistoric-canals in Phoenix, generating environmental high hazard location maps and managing the spatial data in databases. In this paper I have presented a GIS case study to illustrate how Geographical Information Systems technique involves efficient integration of Mathematics, Biosciences, Computers, and Databases.



Written for and Dedicated to the

Divya Dampathi

**Srimathi Sitamma, Sri Ramayya.**

## TABLE OF CONTENTS

<b>CHAPTER 1.....</b>	<b>1</b>
<b>INTRODUCTION.....</b>	<b>1</b>
Background .....	1
Internship Requirements.....	1
Work Environment.....	2
Nature of Work and details.....	2
<b>CHAPTER 2.....</b>	<b>6</b>
<b>REVIEW OF LITERATURE.....</b>	<b>6</b>
Social and Environmental Aspects of Irrigation Systems.....	7
Overview of the Prehistoric Canals Systems.....	8
Geographical Information System (GIS).....	9
Applications of GIS.....	10
GIS Terms.....	10
Maps and Metadata.....	11
<b>CHAPTER 3.....</b>	<b>13</b>
<b>METHODOLGY AND ANALYSIS.....</b>	<b>13</b>
Category 1: Instructor for BIO 591 Course .....	15
Part1: Soils Analysis.....	15
<i>The Logistic Model for Crop Production.....</i>	<i>20</i>
Part 2 Canal Systems.....	23
<i>Steps for Developing a Geographical Model of Slope.....</i>	<i>24</i>
<i>Taylor's Series Expansion of Slope.....</i>	<i>28</i>
Category 2: Research Assistant.....	31
Part 1: Creating Maps.....	31
Part 2: Writing Metadata.....	33
<b>CHAPTER 4.....</b>	<b>36</b>
<b>A GIS CASE STUDY.....</b>	<b>36</b>
Case Study of The Development of Environmental Inequality in Phoenix, Arizona.....	36
Introduction.....	36
General Approach.....	36
Mathematical Concept behind the measurement of CHDI.....	38
Data Collection and Management.....	38
Spatial Methodology.....	39
Results and Conclusions.....	41
<b>CHAPTER 5.....</b>	<b>42</b>

<b>CONCLUSIONS AND FUTURE WORK.....</b>	<b>42</b>
Conclusions.....	42
Future Work.....	43
<b>LIST OF REFERENCES.....</b>	<b>45</b>

## TABLE OF FIGURES

<b>Figure 1: This figure diagrams the logical structure of GIS operation.....</b>	<b>14</b>
<b>Figure 2: Schematic representation of map data and associated attribute data.....</b>	<b>16</b>
<b>Figure 3: This map shows the zoomed portion of the soils boundary.....</b>	<b>17</b>
<b>Figure 4: Soil boundary map of Mexico.....</b>	<b>18</b>
<b>Figure 5: Yield Vs Soil Moisture.....</b>	<b>19</b>
<b>Figure 6: A Grid .....</b>	<b>25</b>
<b>Figure 7: Slope model of pre historical canals surrounding region.....</b>	<b>28</b>
<b>Figure 8: Land use transition of toxic tracts.....</b>	<b>32</b>
<b>Figure 9: Section of metadata design format used in CES department.....</b>	<b>35</b>
<b>Figure 10: Example of Cumulative Hazard Density Index (CHDI).....</b>	<b>38</b>
<b>Figure 11: Hazardous Sites in Phoenix.....</b>	<b>40</b>



# CHAPTER 1

## INTRODUCTION

This chapter gives an overview of my internship as a Graduate Research Assistant in Center of Environmental Studies and a brief description of internship requirements, the type of environment in which I was placed as an intern followed by the details of my work.

### **Background**

My background in Industrial Engineering has given me a strong foundation of mathematics, computers and technology. It has given me a base for developing a logical and analytical approach for problems. I was very keen to implement my skills and the knowledge gained in understanding and studying the agricultural and geographical aspects of Biosciences because India is vastly an agricultural country, unlike the USA. My aim was to establish the use of technology in the field of natural resources, and healthcare. Professional Masters in Computational Biosciences program has provided me with the basic knowledge in biology, and has allowed me to study the integration of technology in the field of biosciences. A semester long internship in the Center of Environmental Studies in Arizona State University, has proved to be very exciting for me, because I learnt to implement and integrate a spatial technology called GIS (geographical Information Systems) in various fields like health care, agriculture, archeology, paleontology and ecology to develop models for analysis.

### **Internship Requirements**

The basic requirements for the internship at CES were having an understanding of some or all of the following:

Programming: Visual Basic 6, SQL, Java 2

Software Packages: ArcGIS, ArcCatalog, ArcMap, ArcInfo, AdobePhotoshop, SPSS.

Databases: MS Access, SQL Server

Others: Knowledge of Mathematics, Biology, Ecology and geography.

My main job was to create models, generate maps and layouts, study the behavior of models and perform analysis as required, database management and sometimes, administrative work too.

### **Work Environment**

The work environment was excellent and it allowed me learn to several technologies. There were several books, Interactive CDs, help-files, and assistance was often available for me to perform my duties promptly. It gave me a wonderful opportunity to study all details at my own pace and my employers were very considerate in letting me do my tasks to the extent of my understanding. Interestingly many of the staff members and professors had a biology background in Life Sciences department at ASU and this has given me an opportunity to talk to them more about their work and area of research.

### **Nature of Work and details**

My work as an intern can be classified into two major categories and further divided into two parts under each category.

#### **Category 1**

Worked as an instructor in Arizona State University for *BIO591* Social Environmental Aspects of Irrigation Systems. I was given 10hrs per week to work as an instructor. My work can be grouped into two parts.

##### **Part-1**

**Description: Pre-historic analysis of canals systems for irrigation.**

Worked average 5 hours a week with additional hours spent at home for learning the new concepts.

My team consisted of four groups and we worked together on a project called "*Analysis of Prehistoric Canals Systems*". The four groups were:

1. Paleontologists
2. Archeologists
3. GIS specialist
4. Software analysts/programmers who knew GIS as well as basic details of plants.

I was under the last category of software programmers for GIS.

My task included the following:

- Writing and editing short scripts and programs in Visual Basic / Avenue Scripts for quick analysis of Arizona and Mexico aerial images.
- Mathematically analyzing the slopes costs to derive optimal values for flow direction of the canals.
- Coordinating with the other groups and learning about crops, plants and a little of archeology.
- Learning new technologies in GIS software in order to implement the different concepts of plants, irrigation and pre-historic data that dated back to 1100 AD.
- Develop watershed, flow direction, flow accumulation, elevation models for analysis.

## **Part-2**

**Description: Modeling the soils of Arizona and Mexico**

Worked average 5 hours a week with additional hours spent at home for learning the new concepts and literature. My team consisted of two groups:

- Paleontologists
- GIS specialists with computer and mathematical knowledge.

I was under the second category and my task included the following:

- Determining the regions in Arizona and Mexico where certain predefined soils are present and querying the datasets in various ways to get the required information like,
  - Types of soils
  - Moisture content of each soil type
  - Flooding frequency
  - The slope gradient of the ground, PH content.
  - Crops that can be grown in a given soil condition.
- Learning the production models for crops and plants like corn maize etc.
  - Soils contain moisture and moisture is required for plant growth. However the water evaporates. If the rate of evaporation is greater than the retention capacity of moisture by the soil, then the plant doesn't have water and would die. In Such cases it would be beneficial to know, how soon the water must be provided? How soon must it rain? If it rains in a particular interval of time, is it sufficient?
- Comparing Arizona and Mexico Soils using GIS technology.
- Creating digital maps by incorporating data on maps.

## **Category 2**

Assisted the environmentalists and software programmers in creating maps and metadata. I was given 10 hours per week to work. This category can be divided into 2 parts.

## **Part-1**

### **Description: Creating Maps.**

Worked approximately 5 hours on various projects, that required maps, layouts, and graphs for analysis.

Many environmentalists and sociologists today don't have much expertise in computers or mathematics. They need efficient programmers, statisticians, and occasionally good modelers who can represent their ideas diagrammatically on maps, layouts and even program their concepts for best results. So a GIS specialist/technician must have the basic knowledge of these various fields and must know how to integrate them suitably as per the requirements. As an intern I have created several maps, which aided the scientists at CES in several decision-making issues. The maps became a part of paper publications, power point presentations, thesis reports, CAP LTER Annual Symposium banners etc.

Some of the maps include:

- Census data imposed on Maricopa County.
- Pollution data overlaid on Maricopa County.
- Health hazards and toxic data superimposed on the Arizona State.
- Land use change of Maricopa tracts over the periods 1970 – 2000.
- Floral collections map.

## **Part-2**

### **Description: Updating metadata on the website <http://www.gp2100.com/eatlas/>**

Worked approximately 5 hours a week to update metadata for e-Atlas website that consisted of various geo-databases and shapefiles.

Tools used for the writing the metadata are Xml-Spy, Stylus Studio, Arc Catalog.

A geo-database is a set of spatial data with information about images, its physical attributes and non-physical attributes. A geo-database converted into a set of visible points, lines or polygons is called a 'shapefile'. Researchers are often interested in knowing the details of the shapefiles and their contents information such as who generated the file and when, the purpose of its creation, spatial attributes involved, accuracy etc.

For this purpose, the shapefiles, and images are each described appropriately using an xml-predefined language. There were nearly 50 datasets and shapefiles that needed updates in which I was involved.

## **CHAPTER 2**

### **REVIEW OF LITERATURE**

The goal of this chapter is to provide a literature background of the various projects that I covered as an intern. The review has been divided into four sections. The first section includes an overview of the BIO591 course and the related project in soil modeling. The second section reports on prehistoric canal systems project. The third section focuses on

GIS concept and its applications. The final section gives a background of metadata, maps and its uses.

### **Social and Environmental Aspects of Irrigation Systems**

BIO 591 in Life Sciences Department of Arizona State University was a course offered in Spring 2003 for students who were interested in studying the environmental aspects and the social aspects involved in irrigation systems in Arizona and Mexico. Dr. Jason and his group of students were interested in developing a model of soils and crops that followed certain production patterns. They were also interested in doing a comparative study for the Mexico soils.

The research focuses on feedbacks between terrestrial ecosystems and environmental change. Many of these feedbacks are driven by changes in biogeochemical cycles, i.e. changes in the storage and movement of energy and nutrients in the Earth system. The interest is on how changes in land-use and atmospheric nitrogen deposition alter terrestrial biogeochemistry. For example:

- Land use - urban ecosystems: Urban ecosystems often receive large inputs of water and nitrogen that accelerate carbon and nutrient cycling through plants and soil microorganisms. It was found that fluxes of greenhouse gases from urban soils to the atmosphere are much higher than in adjacent agricultural and native grassland ecosystems.
- Land use - fire management: For the past century, fire suppression has been the dominant fire management tool in the Southwest, but recently ecological restoration (thinning and prescribed burning) is becoming a widely applied alternative. The interest is in how these forest management choices affect carbon storage and nutrient cycling.

- Nitrogen deposition: Human activities are dramatically increasing the amount of nitrogen in the biosphere. Dr. Jason Kaye and his group found that soils have a large capacity to store new nitrogen in stable organic pools, partially mitigating effects of nitrogen pollution on water quality and atmospheric greenhouse gas concentrations. Urban ecosystems will provide an interesting test of this idea because persistent, high nitrogen inputs may saturate stable organic nitrogen pools in soils.

### **Overview of the Prehistoric Canals Systems**

Recent archaeological evidence of Hohokam land-use practices provides an insight into how large populations can live sustainably in dry lands. The Hohokam did it from at least 500 to 1450 A.D. by being farmers and environmentalists at the same time [2]. As farmers, they adapted their subsistence strategies to specific environmental conditions. As environmentalists, they used resources conservatively and minimized their impact on the ecosystem on which they depended.

Hundreds of miles of prehistoric canals in the Salt, Gila, and Santa Cruz river basins are the most dramatic evidence of the Hohokam's sophisticated management of their crucial water resource. These canals are located in the Sonoran Desert of southern Arizona in the United States. The Hohokam used a complex network of canals to irrigate fields surrounding the area's permanent rivers. Further from the permanent water sources, farmers diverted seasonal flood waters to irrigate their fields, or used dams, terraces, rock walls, and rock mulches to concentrate and retain runoff in cultivated areas.

In addition to growing crops, such as corn, cotton, and tepary beans, the Hohokam also gathered a variety of plants from the surrounding desert. Many Hohokam land-use practices amounted to a sort of environmental protection plan. Hohokam irrigation systems provided a sustainable source of soil fertility. The irrigation canals and runoff catchment

systems harnessed the fertile sediment carried by the running water. The fine sediment added essential nutrients and improved the texture of soil for retaining water and nutrients. This sediment was both added to fields along with irrigation water, and manually spread on the fields during periodic cleaning of the irrigation canals.

Well-adapted strategies for producing and gathering a wide variety of crops in diverse conditions and zones provided the Hohokam a stable basis to assure subsistence and reduce risks in an unpredictable environment. If flooding from excessive rains destroyed some fields, or if drought prohibited some crops from maturing, the Hohokam had others to fall back upon. In order to sustain this diversity of strategies, the Hohokam had organized to spread enough people over various zones to implement land-use strategies and maintain canals, while not overpopulating any given area. Organization has allowed for sufficient specialization and cooperation to exploit the different zones while spreading the risks. By using agricultural practices that minimized destructive impact on their delicate environment, by conservative use of plant resources, and by having a built-in system for replenishing soil fertility, the Hohokam were able to sustain their civilization for more than 1,000 years in a marginal environment.

An ongoing research project by the Archeology department in Arizona State University is trying to determine the Hohokam civilization in relation to west central Mexico from where the Hohokams are believed to have originated. The study aims at understanding how the Mexican culture influenced the Hohokams in their dealing with agriculture and farming.

### **Geographical Information System (GIS)**

What is GIS? It is a particular form of *Information System* applied to geographical data. A *System* is a group of connected entities and activities, which interact for a common

purpose. An *Information System* is a set of processes, executed on raw data, to produce information, which will be useful in decision-making. An information system must have a full range of functions to achieve its purpose, including observation, measurement, description, explanation, forecasting, and decision-making [9].

A Geographic Information System uses geographically referenced data as well as non-spatial data and includes operations, which support spatial analysis. In GIS, the common purpose is decision-making, for managing use of land, resources, transportation, retailing, oceans or any spatially distributed entities. The connection between the elements of the system is geography, e.g. location, proximity, spatial distribution.

In this context GIS can be seen as a computer based system of hardware, software and procedures designed to support the capture.

### **Applications of GIS**

There are several areas where GIS can be applied. Some of the major areas of practical application of GIS Technology, in Natural Resource area are wildlife habitat, forest management, migration routes management, wild and scenic rivers preservation, floodplain management, Wetland preservation, Agricultural lands management, Health Care, Groundwater modeling and contamination tracking, Environmental impact [9].

### **GIS Terms**

Some important terms that one might come across in this paper are:

**Digital Elevation Model** (DEM) is a digital representation of a continuous variable over a two- dimensional surface by a regular array of z values (elevation) referenced to a common datum [7]. Digital elevation models are typically used to represent terrain relief.

**Geo-code** is the process of identifying the coordinates of a location given its address. For example, an address can be matched against a TIGER street network to determine the location of a home. Also referred to as address geo-coding.

**Grid** is a fully integrated grid (cell-based) geo-processing system for use with ArcInfo. GRID supports a Map Algebra spatial language that allows sophisticated spatial modeling and analysis. Also known as “raster”.

**Theme** is a user-defined perspective on a coverage, grid, tin or image geographic data set specified, if applicable, by a coverage name and feature class or data set name, attributes of interest, a data classification scheme, and theme-specific symbology for drawing.

**Shapefiles** are the native format for ArcView and store all of the necessary geometric, locational and attribute information of geographic features (points, lines or polygons).

### **Maps and Metadata**

**Maps:** What Are They? Maps depict, most commonly on a flat surface, the spatial organization of any part of the physical universe at any scale, and at the same time symbolize a wide variety of information, both static and dynamic. Whether it is a chart of outer space or a plan of a university campus, maps play an essential role in our everyday lives. They are used as windows to organize and examine data of various kinds that relate to location, to places. Basically there are two kinds of map categories. Base map and thematic maps. Base Maps depict fundamental information about the Earth's surface such as landforms and drainage. They also symbolize landmark features like roads, railways, populated places and buildings. Thematic Maps show the spatial distribution of a wide variety of qualitative and quantitative information. In fact, almost any subject that can be expressed as a geographical distribution can be mapped. By mapping thematic information and overlaying these layers of data in a GIS (Geographical Information System), we can

analyze human spatial activity. Knowledge gained from this technology allows us to detect problems and plan strategies to manage our environment more successfully. ArcMap is a convenient software package that can handle the various features concerning the development of maps and layouts [9].

### **Metadata**

Metadata or "*data about data*" describes the content, quality, condition, and other characteristics of data. The concept of metadata is familiar to most people who deal with geo-spatial data issues. A map legend is pure metadata. The legend contains information about the publisher of the map, the publication date, and the type of map, a description of the map, geo-spatial references, the map's scale and its accuracy, among many other things. Metadata are simply descriptive information applied to a digital geo-spatial file. Metadata utilize a common set of terms and definitions to document geo-spatial data.

Metadata helps people who use geo-spatial data find the data they need and determine how best to use it. Metadata benefit the data producing organization as well. As personnel change in an organization, undocumented data may lose their value. Later workers may have little understanding of the contents and uses for a digital database and may find they can't trust results generated from these data. Lack of knowledge about other organizations' data can lead to duplication of effort. The generation of metadata is an essential element of building a digital map database [9].

The FGDC " The Federal Geographic Data Committee " was developed to affect a common set of terminology and definitions for the documentation of digital geo-spatial data. XML-Spy and Stylus Studio software packages are usually used to develop metadata. Further details of metadata implementation in projects are covered in chapter 3.

## CHAPTER 3

### METHODOLOGY AND ANALYSIS

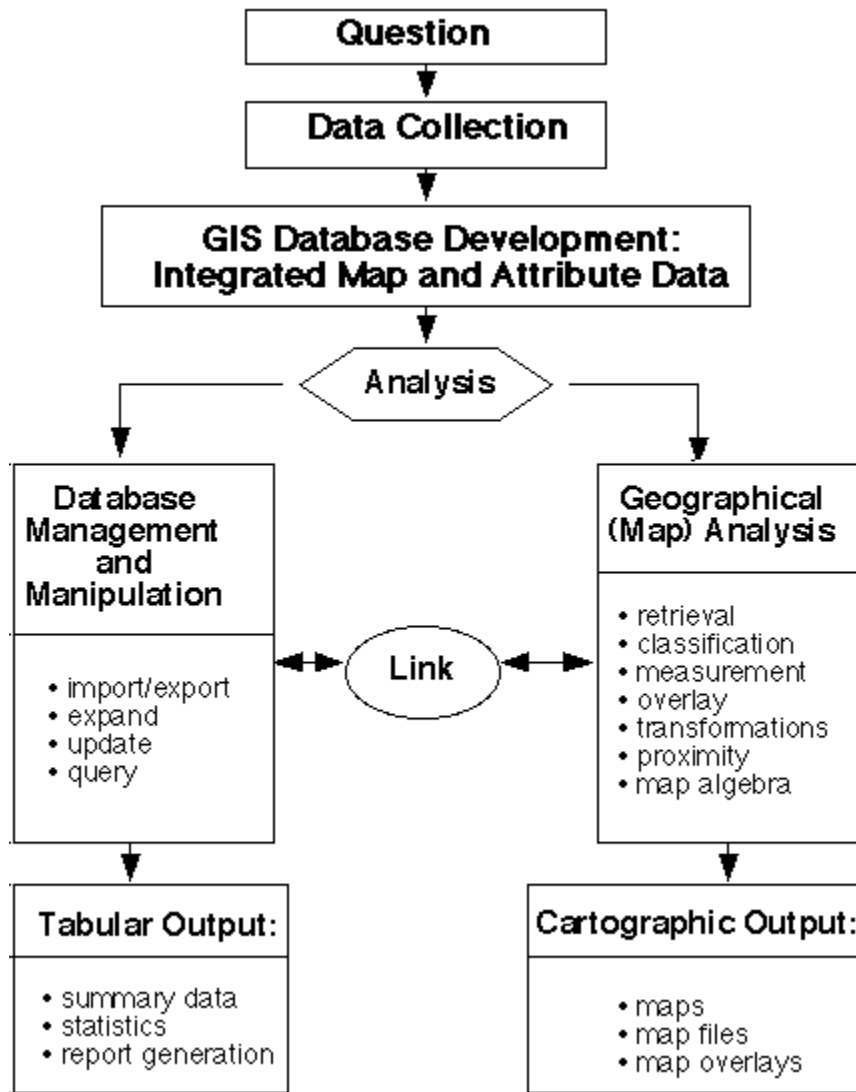
Map data used by GIS are collected from existing maps, aerial photos, satellites, and other sources. A digitizer or similar device is used to convert compiled map data to a digital form in order to make it computer compatible. This transformation allows the storage, retrieval, and analysis of the mapped data to be performed by the computer. What sets GIS apart from even the most sophisticated mapping system is its power to analyze data and to present the results of that analysis as useful information to assist decision makers. This chapter examines the methodology of using GIS for several projects related to various fields like agriculture, archeology etc.

The chapter is divided into two categories based on my work as an intern in the department of CES ASU. Each category is further divided into two parts to describe the approach involved in the projects. In order to get a better understanding of the nature of work, each part has been explained with an example.

#### **Implementing GIS**

Before getting into the details of the project, it is essential to know how GIS works. Use of GIS is driven by the need to answer geographical or spatial questions. The ensuing data collection, database analysis, and output are in response to those questions (Figure 1). The purpose of collecting data for a GIS is (a) to inventory a geographically defined area (for example, to locate all state-owned parcels within a particular area) or (b) to test hypotheses and build models [4]. Though the data is initially collected for one of these reasons, data collected to answer one set of questions is frequently used in subsequent analyses to answer questions that were not anticipated at the time the data was collected.

*For example, after capturing detailed soils data to determine what crops can be grown, the data may be useful for answering questions about septic tank suitability.*



**Figure 1:** This figure diagrams the logical structure of GIS operation.

An initial geographic question drives the data collection. Commonly data collection is done by either manually digitizing data from appropriate maps or by importing tabular data, which is referenced with geo-referenced coordinates that define the geographic position of the tabular (attribute) data. The data is then converted to a digital format and is entered into a database where it is stored as integrated map and attribute data. The

analytical process then manipulates both map and attribute data through the linkages that GIS establishes between them. Two types of output are common: tabular (summary data, statistics, reports) and cartographic (maps, map files, and map overlays). The following sections delineate the basic methodology of GIS technology applied in the projects.

### **Category 1: Instructor for BIO 591 Course**

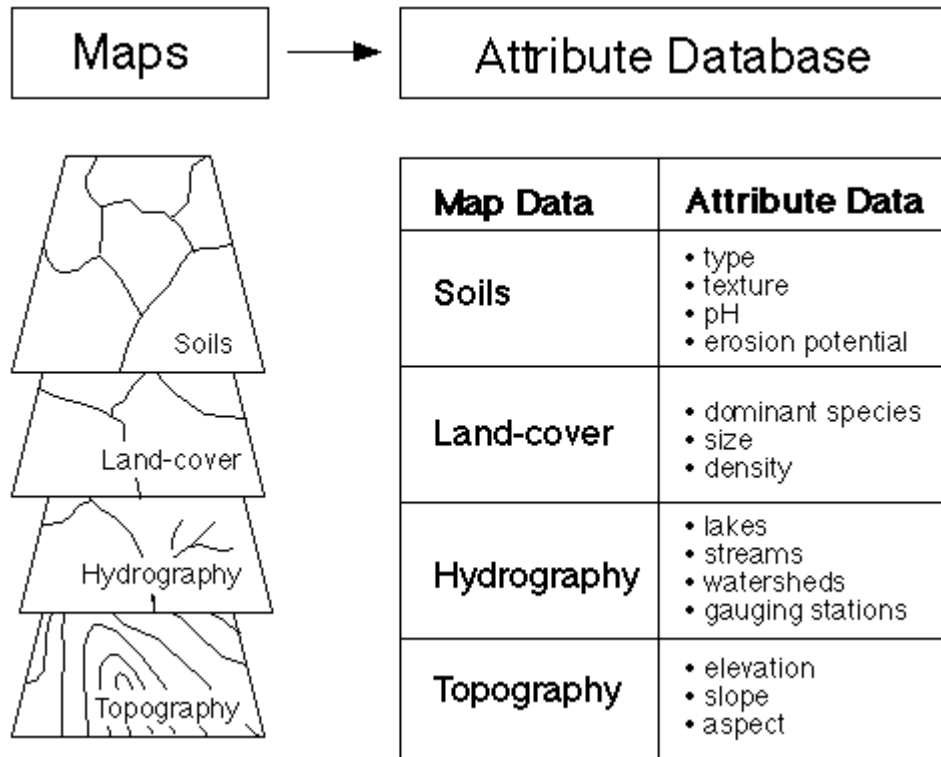
Worked as an instructor in Arizona State University for *BIO591*- Social Environmental Aspects of Irrigation Systems.

#### **Part1: Soils Analysis**

As an instructor I was involved in helping students of BIO591 to analyze various soil data and build models for the same using GIS.

Initially students collected the soil data from fields, websites, laboratories and other resources. The data is formatted and stored as datasets in GIS database. Each dataset typically consists of several attributes like soil name, soil type, soil key, soil texture, slope of the ground, moisture content, PH content, flooding frequency, hydrology, etc.

The GIS database contains both map data (depicting location of geographical objects) and attribute data (describing physical characteristics of each object). Physical characteristics (such as soil type, soil texture, PH content, moisture content) and/or non-physical characteristics (such as estimated soil code, key) are examples of attribute data that could be contained in a GIS used to analyze agriculture related problems. During a GIS analysis, site (map) data is linked with situation (attribute) data for each mapped timber stand [5]. It is this link, which is automatically performed by the GIS software, which gives GIS its analytical power. The relationship between map data and its associated attribute data are shown in Figure 2 for typical natural resource map.



**Figure 2:** Schematic representation of map data and associated attribute data.

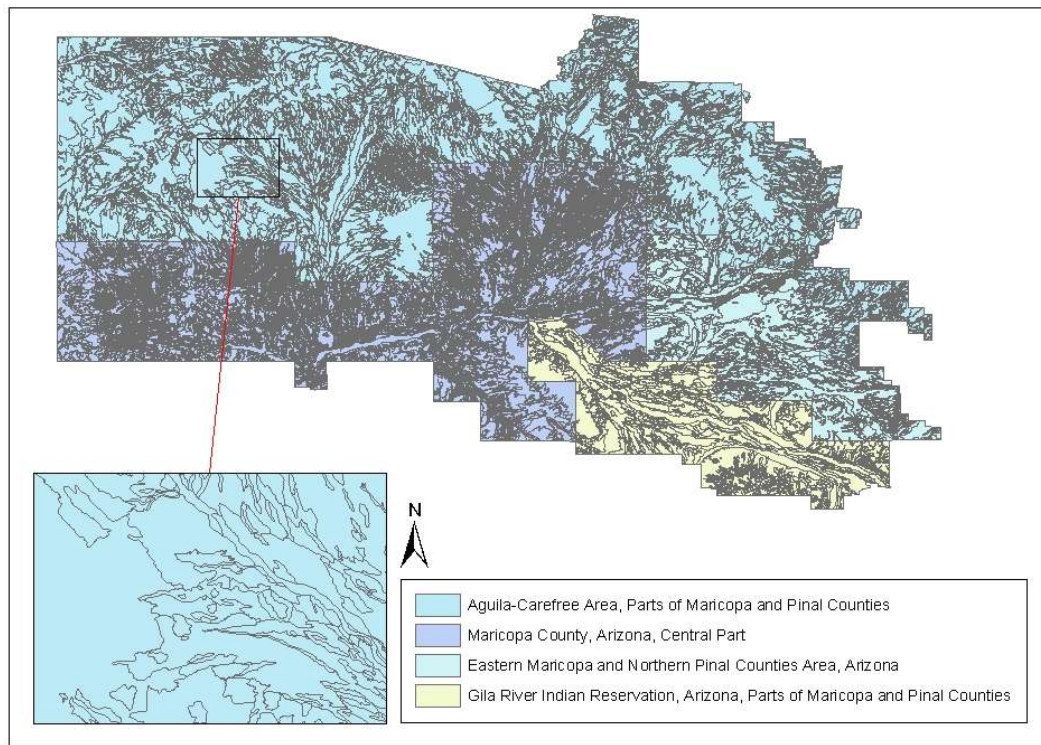
Once the above-described method is established, the next step is to query the database.

To query the database, logical expressions that impose limits or conditions on the database search are defined. These logical expressions specify which geographical objects are to be included in the analysis and/or how that data is to be analyzed. A subset of the database is produced. Some logical expressions are simple and require only one condition while others are very complex and contain multiple conditions. Some of the queries performed are:

- Name the soils whose slope gradient was greater than 40%?
- What is the extent of soil types in Phoenix region? Vaules in Sq Kms. (See Figure 3)
- In a buffer zone of 1-mile radius around the city of Phoenix what are the various soils found?
- What are the various soils textures found in the region of Zacatacus in Mexico? (See Figure 4)

- Are any definite interesting patterns observed in the data for the regions of Mexico and Arizona?

These questions prompt the DBMS to perform functions such as calculating density, generating descriptive statistical summaries, and translating between measurement systems (e.g., converting area defined in square meters into acres or square miles). These types of operations are common when the database contains numeric data and a statistical analysis of the data is required [12].



**Figure 3:** This map shows the zoomed portion of the soils boundary.

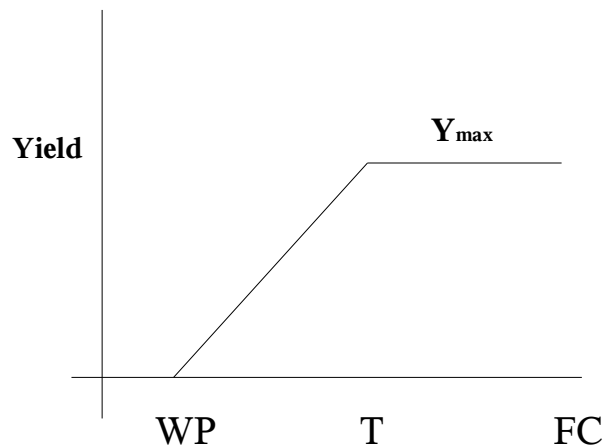


- Keep a running tally of average SM ( $\langle SM \rangle$ ) over the entire 120-day (or given time period) growing period.
- If SM from 0 days to 30 days (making this up, maybe the cut-off is 15 days) is below some critical value (at or near Wilting Point (WP)) then the crop is killed.
- If rainfall from 100-140 days is  $>$  some value, the harvest rots.
- OTHERWISE ( $Y = \text{Yield}$ )

$$\text{If } \langle SM \rangle > T, Y = Y_{\max}$$

$$\text{If } \langle SM \rangle \leq WP, Y = 0$$

Otherwise,  $Y = Y_{\max} * (\langle SM \rangle - WP) / (FC - WP)$  (If we decide to go with the linear version —see Figure 5 and refer the ‘Logistic Model for Crop Production’ for further analysis of the Figure 5).



**Figure 5:** Yield Vs Soil Moisture

Developing production models for soils and crops like these help us determine the following points as discussed by Dr. Jason Kaye.

- (1) What is starting soil moisture (SM)?
  - a. For irrigation with enough water, SM (0) (where 0 is planting time) what would be the Field Capacity (FC).

(2) What is the appropriate  $\Delta t$ ?

- a. Rainfall comes in 1-month resolution, so we may want to pick 1 month. On the other hand, we may want to pick  $\Delta t$  to mimic a time period over which the plant would die if  $SM = WP$  over that whole time period.

(3) Note that if the fields are flooded, evaporation just comes from the surface water and the soils should be considered to be at FC. So we may want an equation in there for irrigated systems—as long as the field is flooded,  $SM = FC$ .

(4) The  $t = 0$  time may itself change. What kind of flexibility is there in the planting date for corn? Could one wait for the rain (or at least the clouds)?

### **The Logistic Model for Crop Production**

Let Soil Moisture =  $s$

Rainfall =  $r$ , Irrigation =  $i$  and Evaporation =  $e$

We have the rate of change of  $s$  as,

$$s(t + dt) - s(t)$$

Let, in time  $dt$ ,

$$\text{rainfall} = r s(t) dt$$

$$\text{irrigation} = i s(t) dt$$

$$\text{evaporation} = e s(t) dt$$

The change of  $s$  in time  $dt$  depends on rainfall, irrigation and evaporation. So we have,

$$s(t + dt) - s(t) = r s(t) dt + i s(t) dt - e s(t) dt$$

Substituting  $r = n - i + e$  we have

$$s(t + dt) - s(t) = s(t) dt n$$

$$\frac{s(t + dt) - s(t)}{dt} = s(t) n$$

Applying Limits  $dt \rightarrow 0$  we get,

$$\frac{d}{dt} s(t) = s(t) n$$

$$\frac{\frac{d}{dt} s(t)}{s(t)} = n$$

Integrating on both sides we get,

$$\int \frac{1}{s} ds = \int n dt + C$$

$$\ln(s) = n t + C$$

The solution yields an exponential function, where  $e^C$  is a constant =  $s_0$ .

$$s = e^{nt} e^C$$

$$s = s_0 e^{nt}$$

This is an *exponential model*. While soil moisture may manifest such increase of moisture content patterns during periods of short duration, it is clear that there must be factors which attenuate increase in the long run. One model that takes these factors into account is the Verhulst-Pearl logistic equation,  $dN / dt = rN (1 - N / K)$ .

Since the resources for moisture are generally limited, it is reasonable to assume that as moisture content in the soil increases and it approaches the carrying capacity which is in our case the field capacity  $K$ , of its environment, the rate of rainfall reduces, irrigation is not continued or irrigation rate is gradually reduced, and the evaporation rate begins to increase. Hence we can re-define these terms as follows:

$$r = r_0 - k_r s$$

Where  $r_0$  is the initial rate of rainfall and  $k_r$  is the rate at which the rainfall declines with an increase in soil moisture.

$$i = i_0 - k_i s$$

Where  $i_0$  is the initial rate of irrigation and  $k_i$  is the rate at which the irrigation reduces with an increase in soil moisture.

$$e = e_0 + k_e s$$

Where  $e_0$  is the initial rate of evaporation and  $k_e$  is the rate at which the evaporation increases with an increase in soil moisture.

In other words, the rainfall rate, irrigation rate and evaporation rates can be interpreted as linearly related to the amount of soil moisture.  $s$  will stabilize, or reach a fixed point, if

$$r + i = e$$

Or,

$$r_0 - k_r s + i_0 - k_i s = e_0 + k_e s$$

Solving for  $s$ , we get a carrying capacity  $K$

$$K = \frac{r_0 + i_0 - e_0}{k_i + k_r + k_e}$$

Let  $c = r_0 + i_0 - e_0$ . Then,

$$k_i + k_r + k_e = \frac{c}{K}$$

Coming back to our original differential equation, we can modify the rate of growth in terms of  $r$ ,  $i$  and  $e$ , and perform some algebraic simplification as follows,

$$\frac{d}{dt} s(t) = s(t) n$$

Substituting  $n = r + i - e$  we get,

$$\frac{d}{dt} s(t) = s(t) (r + i - e)$$

$$\frac{d}{dt} s(t) = s(t) (r_0 - k_r s + i_0 - k_i s - e_0 - k_e s)$$

$$\frac{d}{dt} s(t) = s(t) (r_0 + i_0 - e_0 - s(t) (k_r + k_i - k_e))$$

From earlier calculation we have  $k_i + k_r + k_e = \frac{c}{K}$  and  $c = r_0 + i_0 - e_0$ .

$$\frac{d}{dt} s(t) = s(t) \left( c - \frac{s c}{K} \right)$$

Re-arranging the above equation we get,

$$\frac{d}{dt} s(t) = s(t) c \left( 1 - \frac{s}{K} \right)$$

Where  $s$  is Soil Moisture and  $K$  is the Field Capacity. Thus, the net rate of growth in the logistic equation is continually changing. If  $s$  is small relative to its carrying capacity,  $K$ , the rate will be near to the constant rate,  $c$ , of the exponential model. As  $s$  nears  $K$ , the rate will shrink toward 0. Theoretically, when  $s$  reaches  $K$  a situation of equilibrium will be reached and the moisture content will be stable. If  $s$  were to somehow exceed  $K$ , the rate would become negative and the moisture would decrease toward the carrying capacity.

## **Part 2 Canal Systems.**

### **Modeled the pre-historic canals systems for irrigation.**

The canals that were discovered in the Phoenix Basin are helping settle questions about the influence of Mexican cultures on the development of irrigation technology in the U.S. Southwest [2]. Southwestern cultures farmed crops that derived from Mexico, including maize, squash and beans, and it was a logical inference that knowledge of canals and irrigation also migrated northward from more advanced Mexican civilizations. The methodology and analysis for this project can be described in the following steps.

## **Steps for Developing a Geographical Model of Slope**

### **Step 1: Problem Statement**

- Develop a geographic model depicting the regions of pre-historic canals in Phoenix Basin to analyze the irrigation system.
- Understand the slope and aspect ratio to determine how the ‘Hohokam’ people learned to control floods, irrigate their lands, grow crops etc.
- Study the relation between the two cultures and find out if the Hohokam were originally people from Mexico.

### **Step 2: Problem Breakdown**

The problem can be broken down into a series of objectives:

- Develop a basic phoenix map showing the pre-historic canals.
- Determine the various statistics that are required for analyzing the canals.
- Project these statistics on the map.
- Develop and query similar regions in Mexico.
- Analyze the data.

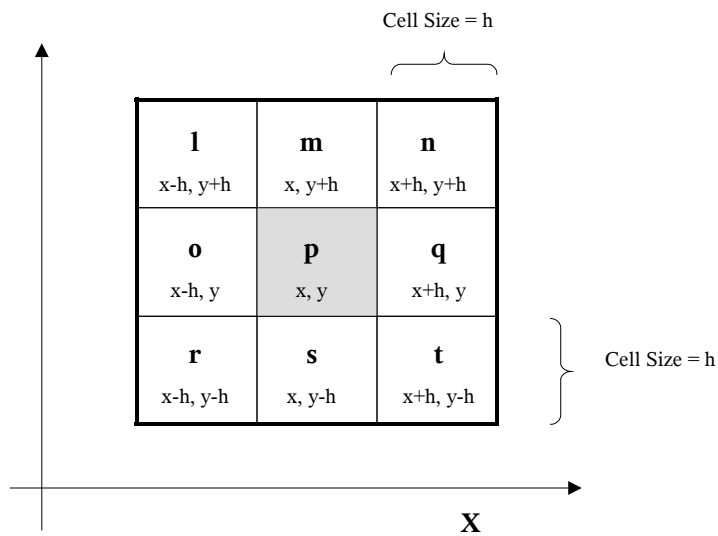
### **Step 3: Exploring Input Datasets**

This is a very important step in modeling the irrigation canal system. The pre-historic canals that were discovered during excavation in South West Phoenix Basin have their origin in the Salt River. The slope of the riverbed was found to be an important criterion for the researchers to determine the crops grown, irrigation and flood control methods of the Hohokam.

Logically inferring, some combinations of slope and soil type represent high risk for erosion, and knowing where these conditions are, enables risk assessment of crops. For instance, rocky or sandy soil on a relatively steep slope could represent flash-flooding

dangers to either crop or grazing agricultural activities during rainstorms. The following discussion helps in the analysis of slope.

A grid file is a raster image consisting of several pixels called cells. Each cell carries some information and this is called as ‘pixel value’ or ‘cell value’. DEM images are actually grid images with each cell carrying the information of elevation of the land. Hence DEMs are called Digital Elevation Models. In this project, the Arizona DEM image is used for the analysis of slope, and cost paths.



**Figure 6: A Grid**

**Slope Analysis:**

Slope identifies the maximum rate of change in value from each cell to its neighbors. An output slope grid can be calculated as percent slope or degree of slope. The slope for the cell is calculated from the 3x3 neighborhood using the average maximum technique [4]. The actual algorithm that is used to calculate slope is:

$$\sqrt{\left(\frac{dz}{dx}^2 + \frac{dz}{dy}^2\right)}$$

The deltas are calculated using a 3x3-roving window, where  $l$  through  $t$  in Figure 6 represents the values in the window above. The derivation of the following terms is discussed in the section-‘Taylor Series Expansion of Slopes’.

$$\frac{dz}{dx} = \frac{(n + 2q + t) - (l + 2o + r)}{8h}$$

$$\frac{dz}{dy} = \frac{(l + 2m + n) - (r + 2s + t)}{8h}$$

Where  $h$  is the cell size.

The use of a z-Factor is essential for correct slope calculations when the surface z-units are expressed in units, which are different from the ground units.

After the slope calculation the next part is the union process, for which polygon themes for slope and soil type data are needed. The boundaries of slope (areas that have the same range of slope angle with respect to elevation change) are used as the input theme. The soil type data is used as the overlay or Union theme. The result is a new theme that contains the spatial combination of information with attribute data, permitting evaluation of erosion potential.

Once we compute the slope, the subsequent calculations are cost path, and contour lines. Cost path tells us the cost of traveling from one cell to another given the elevation constraint. This is important since it gives an overview of how difficult it could be to deviate a canal from its original path as a process of flood control.

In order to calculate the cost path, the ArcMap software requires a cost surface, which is used to compute the accumulated cost for traveling from a given point (source cell) to any part of the land (any other cell). So it basically needs two input datasets. One is the cost surface dataset and another is a point theme, which contains point(s) that act as source cells for the calculation of accumulated costs for all the remaining cells in the

topography of the DEM image [7]. From the resulting cost path 'shapefiles', the contour lines are derived to show the regions of similar zones.

#### **Step 4: Perform analysis**

Spatial Analyst Extension in ArcMap software provides a wide variety of tools to serve this purpose. In this project I used the following tools:

**Programming:** Visual Basic 6.0, SQL.

**Databases:** MS Access Database, MS Excel.

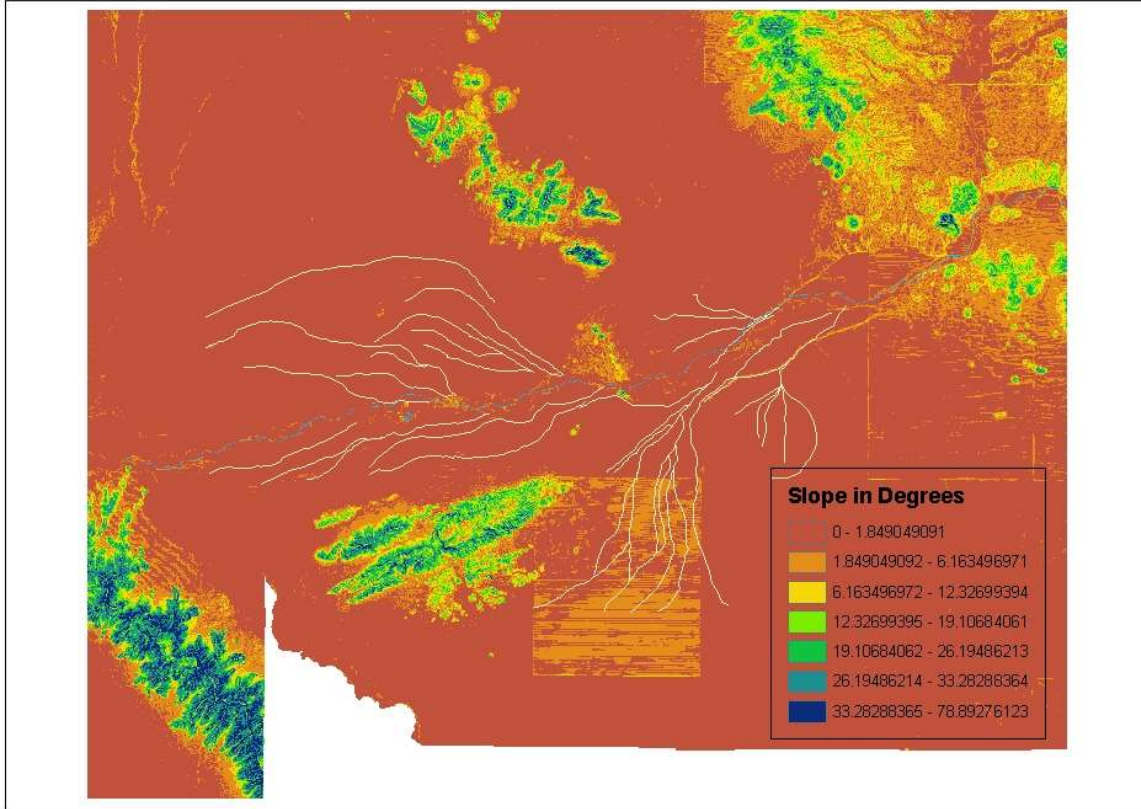
**Software Packages:** ArcGIS, ArcView, ArcInfo, ArcMap, ArcCatalog, ArcSDE, Adobe Photoshop, Microsoft office suite.

#### **Step 5: Verify the model's result**

Quality Assurance should be done on each of the maps to make sure they depict the required information. Projecting the maps in proper units require consistent units for all features like elevation values, x-coordinates, and y-coordinates.

#### **Step 6: Implement the result**

Once the spatial problem is analyzed and solved, the final output needs to be generated. This output is eventually used for decision-making. A temporary slope model is as shown in Figure 7.



**Figure 7:** Slope model of pre historical canals surrounding region.

Each color in the map represents a different slope value. For example, the red color indicates the least range of slope which is 0 – 1.85 degrees. The regions where the blue is prominent are the regions of highest slope, which are the mountains. The Salt River is flowing right across the map and the canals are originating from the Salt River.

The next section describes how the mathematical model for the slope was achieved using Taylor Series.

### Taylor's Series Expansion of Slope

By Taylor's series expansion we have,

$$\begin{aligned}
 f(x+h, y+k) &= f(x, y) + \left( h \frac{\partial f}{\partial x} + k \frac{\partial f}{\partial y} \right) + \frac{1}{2!} \left( h^2 \frac{\partial^2 f}{\partial x^2} + 2hk \frac{\partial^2 f}{\partial x \partial y} + k^2 \frac{\partial^2 f}{\partial y^2} \right) + \dots \\
 &= \sum_{p=0}^{N-1} \frac{1}{p!} \left( h \frac{\partial}{\partial x} + k \frac{\partial}{\partial y} \right)^p f(x, y) + R_N
 \end{aligned}$$

Using Taylor's series to expand each of the eight surrounding nodes around the node  $p$  in Figure 6, we get the following eight equations:

$$f(x+h, y) = f(x, y) + h \left( \frac{\partial}{\partial x} f \right) + \frac{1}{2} h^2 \left( \frac{\partial}{\partial x} f \right)^2 + O(h)^3 \quad 1$$

$$f(x-h, y) = f(x, y) - h \left( \frac{\partial}{\partial x} f \right) + \frac{1}{2} h^2 \left( \frac{\partial}{\partial x} f \right)^2 + O(h)^3 \quad 2$$

$$f(x, y+h) = f(x, y) + h \left( \frac{\partial}{\partial y} f \right) + \frac{1}{2} h^2 \left( \frac{\partial}{\partial y} f \right)^2 + O(h)^3 \quad 3$$

$$f(x, y-h) = f(x, y) - h \left( \frac{\partial}{\partial y} f \right) + \frac{1}{2} h^2 \left( \frac{\partial}{\partial y} f \right)^2 + O(h)^3 \quad 4$$

$$f(x-h, y+h) = f(x, y) + h \left( -\left( \frac{\partial}{\partial x} f \right) + \left( \frac{\partial}{\partial y} f \right) \right) + \frac{1}{2} h^2 \left( -\left( \frac{\partial}{\partial x} f \right) + \left( \frac{\partial}{\partial y} f \right) \right)^2 + O(h)^3 \quad 5$$

$$f(x+h, y-h) = f(x, y) + h \left( \left( \frac{\partial}{\partial x} f \right) - \left( \frac{\partial}{\partial y} f \right) \right) + \frac{1}{2} h^2 \left( \left( \frac{\partial}{\partial x} f \right) - \left( \frac{\partial}{\partial y} f \right) \right)^2 + O(h)^3 \quad 6$$

$$f(x+h, y+h) = f(x, y) + h \left( \left( \frac{\partial}{\partial x} f \right) + \left( \frac{\partial}{\partial y} f \right) \right) + \frac{1}{2} h^2 \left( \left( \frac{\partial}{\partial x} f \right) + \left( \frac{\partial}{\partial y} f \right) \right)^2 + O(h)^3 \quad 7$$

$$f(x-h, y-h) = f(x, y) + h \left( -\left( \frac{\partial}{\partial x} f \right) - \left( \frac{\partial}{\partial y} f \right) \right) + \frac{1}{2} h^2 \left( -\left( \frac{\partial}{\partial x} f \right) - \left( \frac{\partial}{\partial y} f \right) \right)^2 + O(h)^3 \quad 8$$

Subtracting equation 2 from equation 1, we get equation 9

$$f(x+h, y) - f(x-h, y) = 2 h \left( \frac{\partial}{\partial x} f \right)$$

$$\frac{\partial}{\partial x} f = \frac{1}{2} \frac{f(x+h, y) - f(x-h, y)}{h} \quad 9$$

Subtracting equation 4 from equation 3, we get equation 10

$$f(x, y+h) - f(x, y-h) = 2 h \left( \frac{\partial}{\partial y} f \right)$$

$$\frac{\partial}{\partial y} f = \frac{1}{2} \frac{f(x, y+h) - f(x, y-h)}{h} \quad 10$$

Subtracting equation 5 from equation 6, we get equation 11

$$f(x+h, y-h) - f(x-h, y+h) = 2h \left( \left( \frac{\partial}{\partial x} f \right) - \left( \frac{\partial}{\partial y} f \right) \right) \quad 11$$

Subtracting equation 8 from equation 7, we get equation 12

$$f(x+h, y+h) - f(x-h, y-h) = 2h \left( \left( \frac{\partial}{\partial x} f \right) + \left( \frac{\partial}{\partial y} f \right) + 2h \left( \frac{\partial}{\partial x} f \right) \left( \frac{\partial}{\partial y} f \right) \right) \quad 12$$

Adding equation 11 and equation 12, we get equation 13.

$$\begin{aligned} f(x+h, y-h) - f(x-h, y+h) + f(x+h, y+h) - f(x-h, y-h) &= \\ 4h \left( \frac{\partial}{\partial x} f \right) \left( 1 + h \left( \frac{\partial}{\partial y} f \right) \right) & \\ \frac{\partial}{\partial x} f = \frac{1}{4} \frac{f(x+h, y-h) - f(x-h, y+h) + f(x+h, y+h) - f(x-h, y-h)}{h} & \end{aligned} \quad 13$$

Subtracting equation 11 from equation 12, we get equation 14.

$$\begin{aligned} f(x+h, y+h) - f(x-h, y-h) - f(x+h, y-h) + f(x-h, y+h) &= \\ 4h \left( \frac{\partial}{\partial y} f \right) \left( 1 + h \left( \frac{\partial}{\partial x} f \right) \right) & \\ \frac{\partial}{\partial y} f = \frac{1}{4} \frac{f(x+h, y+h) - f(x-h, y-h) - f(x+h, y-h) + f(x-h, y+h)}{h} & \end{aligned} \quad 14$$

Adding equation 9 and equation 13, we obtain equation 15 in terms of  $\frac{\partial}{\partial x} f$

$$\begin{aligned} 2 \left( \frac{\partial}{\partial x} f \right) &= \frac{1}{4} (f(x+h, y-h) - f(x-h, y+h) + f(x+h, y+h) - f(x-h, y-h) \\ &\quad + 2f(x+h, y) - 2f(x-h, y))/h \\ \frac{\partial}{\partial x} f &= \frac{1}{8} (f(x+h, y-h) - f(x-h, y+h) + f(x+h, y+h) - f(x-h, y-h) \\ &\quad + 2f(x+h, y) - 2f(x-h, y))/h \end{aligned} \quad 15$$

Substituting the nodes values from  $a$  through  $i$  in equation 15, we get  $\frac{\partial}{\partial x} f$  as,

$$\frac{\partial}{\partial x} f = \frac{(n+2q+t) - (l+2o+r)}{8h}$$

Adding equation 10 and equation 14, we get equation 16 in terms of  $\frac{\partial}{\partial y} f$

$$\begin{aligned} 2 \left( \frac{\partial}{\partial y} f \right) &= \frac{1}{4} (-f(x+h, y+h) + f(x-h, y-h) + f(x+h, y-h) - f(x-h, y+h) \\ &\quad - 2f(x, y+h) + 2f(x, y-h))/h \end{aligned} \quad 16$$

Substituting the nodes values from  $a$  through  $i$  in equation 16, we get  $\frac{\partial}{\partial y} f$  as,

$$\frac{\partial}{\partial y} f = \frac{(l + 2m + n) - (r + 2s + t)}{8h}$$

## **Category 2: Research Assistant.**

My work in this category included writing metadata and creating maps and layouts for various datasets. This section again contains two parts and the methodology of the tasks is described briefly below.

### **Part 1: Creating Maps.**

As described earlier, there are several maps that were needed for the purpose of analysis. A simple procedure for map creation is outlined in steps 1 through 5.

ArcMap™ software package provides two different ways to view a map. One is Data view and second is a Layout view. Each view allows us look at and interact with the map in a specific way. When we want to browse the geographic data on a map, we choose data view. It is not surrounded by any border, lines and not bounded in any framework. Layout view should be used when we preparing a map to hang on the wall, put in a report, or publish on the Web.

#### **Step 1**

The shapefiles that are required to be projected in the form of a map are first brought together in a frame. The shapefile may include, images (Ex: Aerial image), polygons (Ex: Arizona Counties), points (Ex: Arizona Cities), or lines (Ex: Railways).

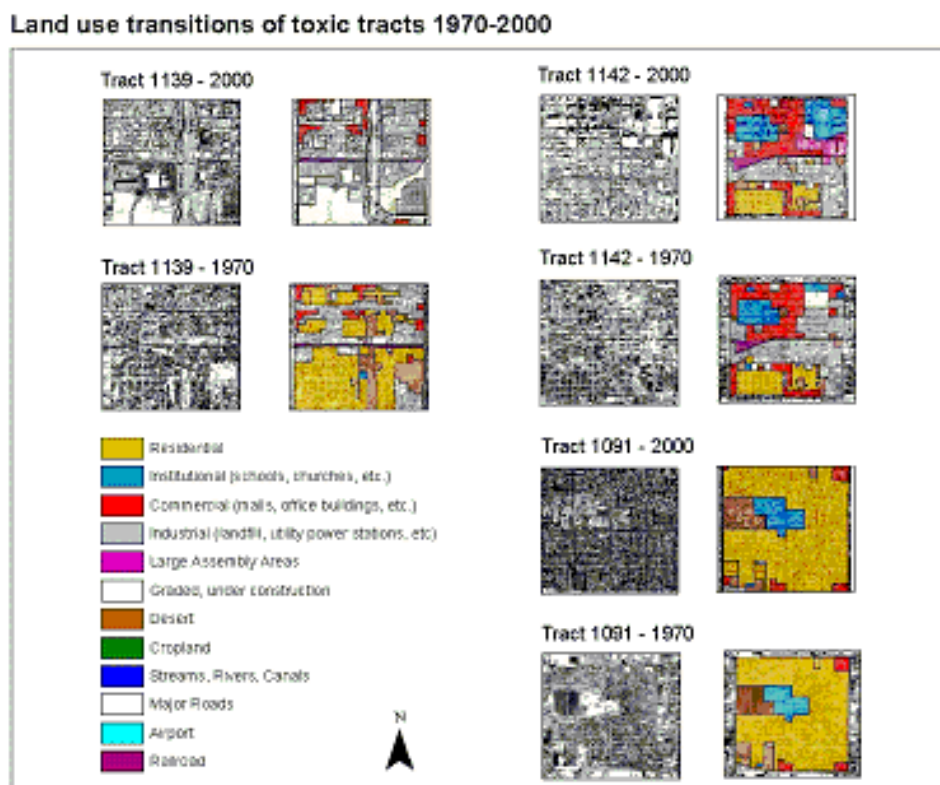
#### **Step 2**

A layout is designed with required page set up where all the elements will be projected. Ex: 8 ½ X 11 inch, 1-inch margin, black border etc. On a layout one can add

several features like data-frames, tables, images, objects and graphs. A layout can always be altered and made suitable for printing. The elements like legend, north arrow, title, description, labels are added to enhance the layouts. Map Locators can be added to zoom into specific sections and show it in a different view.

### Step 3

This final map can now be printed, published and data can be super imposed for queries. See Figure 8.



**Figure 8:** Land use transition of toxic tracts.

Apart from being highly useful for pictorial representation, maps are necessary to depict and make comparisons with various topographical images for better understanding. It helps us in making decisions, which is not otherwise possible. For instance, consider the map (Figure 8) generated in ArcMap [11].

The Figure 8 shows a land use distribution and transition of tract 1142, tract 1091 and tract 1139 in Maricopa County over a period of 30 years. It is clear from the map that there is a significant land use transition from residential (shown Yellow) to industrial (in Gray) in a period of 30 years for the tract 1139. Such vital decisions and several more such analysis would not have been possible with just data in tabular format. It's the ability to make '*spatial distribution*' of the data that makes GIS so important.

## **Part 2: Writing Metadata**

Metadata encompass the information about a spatial data set. A metadata document should comply with the Federal Geographic Data Committee (FGDC) standards, and should be available with the spatial data set. A metadata is a very useful feature of GIS because it helps maintain the details of any digital image created for the purpose of study.

Metadata for all shapefiles created were written using XML-Spy and Stylus Studio software packages. The procedure can be briefly described in 3 steps.

- First an XML file is generated using ArcCatalog based on the required format. The ArcCatalog consists of metadata conversion tool in FGDC format.
- Next the XML is saved into a pre-defined EML-format using Stylus Studio.
  - Ecological Metadata Language (EML) is a metadata standard developed by the ecology discipline and for the ecology discipline. EML is implemented as a series of XML document types that can be used in a modular and extensible manner to document ecological data. Each EML module is designed to describe one logical part of the total metadata that should be included with any ecological dataset.
- Finally the EMLs are next edited in xml-spy suitably to add additional metadata.

The following are a set of question the metadata must answer:

**Identification information:** What is the name of the data set? Who developed the data set? What geographic area does it cover? What themes of information does it include? How current are the data? Are there restrictions on accessing or using the data?

**Data quality information:** How good are the data? What is the positional and attribute accuracy? What data were used to create the data set, and what processes were applied to these sources?

**Spatial data organization information:** What spatial data model was used to encode the spatial data? How many spatial objects are there? Are methods other than coordinates, such as street addresses, used to encode locations?

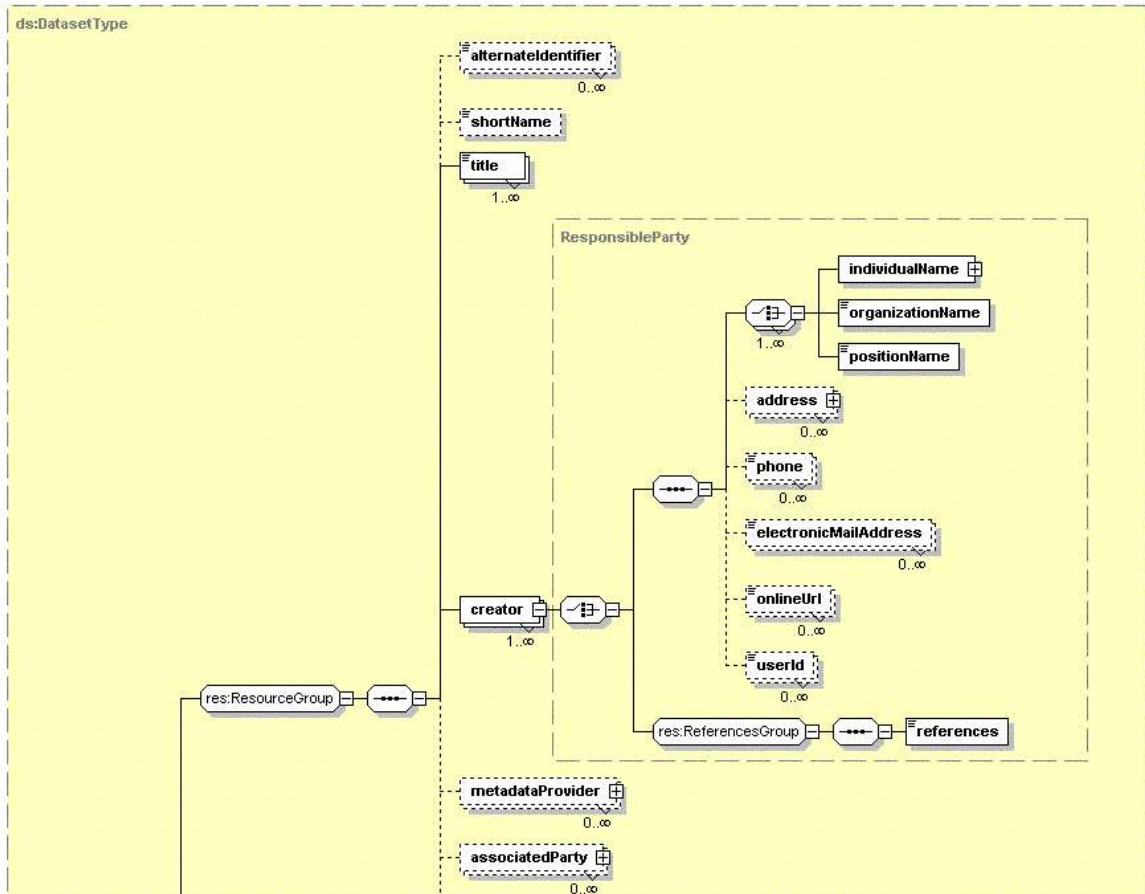
**Spatial reference information:** Are co-ordinate locations encoded using longitude and latitude? Is a map projection or grid system, such as the State Plane Coordinate System, used? What parameters should be used to convert the data to another coordinate system?

**Entity and attribute information:** What geographic information (roads, houses, elevation, temperature, etc.) is included? How is this information encoded? Were codes used? What do the codes mean?

**Distribution information:** From whom can I obtain the data? What formats are available? What media are available? Are the data available online? What is the price of the data?

**Metadata reference information:** When were the metadata compiled and by whom?

A typical metadata template model looks as shown in Figure 9.



**Figure 9:** Section of metadata design format used in CES department.

## CHAPTER 4

### A GIS CASE STUDY

So far, a detailed overview of my work as a GIS Research Assistant in Center of Environmental Studies Department at Arizona State University has been given. Now, in this chapter I would like to pick a case study and apply the GIS concepts, simultaneously integrating the IT aspects to finally develop a model. This case study was part of my internship in CES and I was primarily involved in generating maps and data collection.

#### **Case Study of The Development of Environmental Inequality in Phoenix, Arizona**

##### **Introduction**

In this case study the development of environmental inequalities in the Phoenix, Arizona metropolitan area shall be examined [1]. Four hazard types are utilized in order to identify environmental inequities. These are: Toxic Release Inventory facilities, Large Quantity Generators of hazardous waste, Treatment, Storage and Disposal Facilities, and toxic contamination sites identified by the EPA. A spatial methodology that provides a cumulative hazard density index (CHDI) is employed to quantify proximity-based risk burdens for each census tract in the metropolitan region. After identifying tracts that exhibit the highest concentrations of hazards, case studies are developed in order to characterize the selected tracts and adjacent areas.

The focus is on identifying historical-geographic changes in land uses, demographic composition of neighborhoods and the silting/abandonment of hazardous industrial facilities that have resulted in pronounced environmental inequities in the metropolitan area.

##### **General Approach**

In what follows, we can trace the production of environmental inequities in Phoenix by examining the succession of land uses and zoning, demographic changes, and the

development of industrialized nodes within the city. The approach is to apply a two-stage strategy.

First, the current distributions of four point-source hazards at the census tract level are mapped. These hazards include:

- Toxic Release Inventory (TRI) facilities
- Toxic contamination sites identified under CERCLIS (including Superfund sites)<sup>2</sup>
- Large Quantity Generators of hazardous wastes (LQG), and
- Treatment Storage Disposal facilities (TSDF).

This mapping is the basis for a spatial methodology that allows us to calculate a cumulative hazard density index (CHDI). This is a quantitative measure of the aggregate risk burdens each tract bears as a result of proximity to multiple hazard sources. This allows us to identify and rank tracts, producing a hierarchy of risk burdens linked to tract socio-demographic measures.

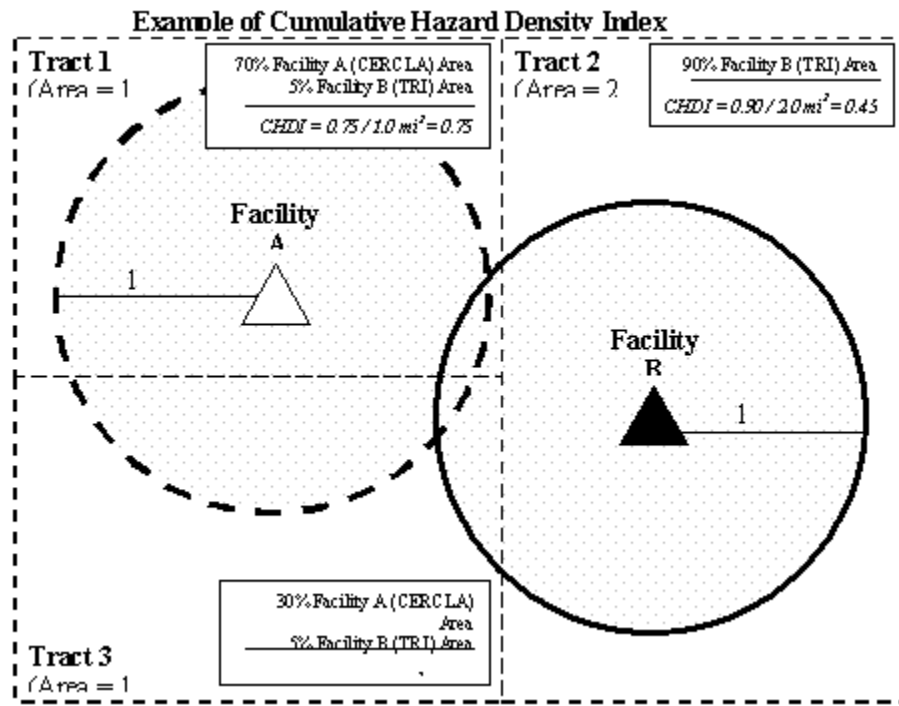
The second element of the analysis selects a group of census tracts with the highest CHDI scores for further examination. Tracts with disproportionate risk burdens are investigated using available data for changing demographic composition, land use patterns, zoning, and facility citing dates. Based on preliminary analysis of the ten tracts with the highest hazard density scores, we develop a typology of three developmental pathways to becoming a ‘toxic tract.’ Three tracts and their surrounding areas that each exemplifies one of the three pathways to environmental inequality are selected for more detailed historical analysis.

---

<sup>2</sup>CERCLA– the Comprehensive Environmental Response, Compensation, and Liability Act provides the legal mandates for the investigation and regulation of a wide range of toxic contamination sites. CERCLIS, the Information System maintained under the act, lists heavily contaminated sites (toxic waste dumps, abandoned industrial sites, mines, severely polluted federal facilities etc).

## Mathematical Concept behind the measurement of CHDI

Each hazardous facility releases certain quantity of toxic substance either into the atmosphere or on the land or in water. The population residing in the surrounding regions feels the effect of the hazard. It is assumed that the effect is shown in a fixed radius of 1 sq km or 1 sq mile surrounding each facility. The following Figure 10 clearly explains the methodology of CHDI calculations. This way the CHDI is calculated for each census tract.



**Figure 10:** Example of Cumulative Hazard Density Index (CHDI)

## Data Collection and Management

Census data for the years 1970 to 2000 provide information to document the patterns of socio-demographic change in the tracts in question. Historic data documenting aerial proportions of each tract given to industrial, residential, commercial, institutional, and other uses provides evidence for changes in land uses over time. Other data, obtained from CERCLIS, is used to locate and date contamination events and sources of contamination.

Business directories are used to date the establishment of selected industrial facilities (TRI and LQG) and waste handling facilities (TSDFs) in the tracts. The analysis also considers tract specific changes in relation to urban developments at other spatial and temporal scales.

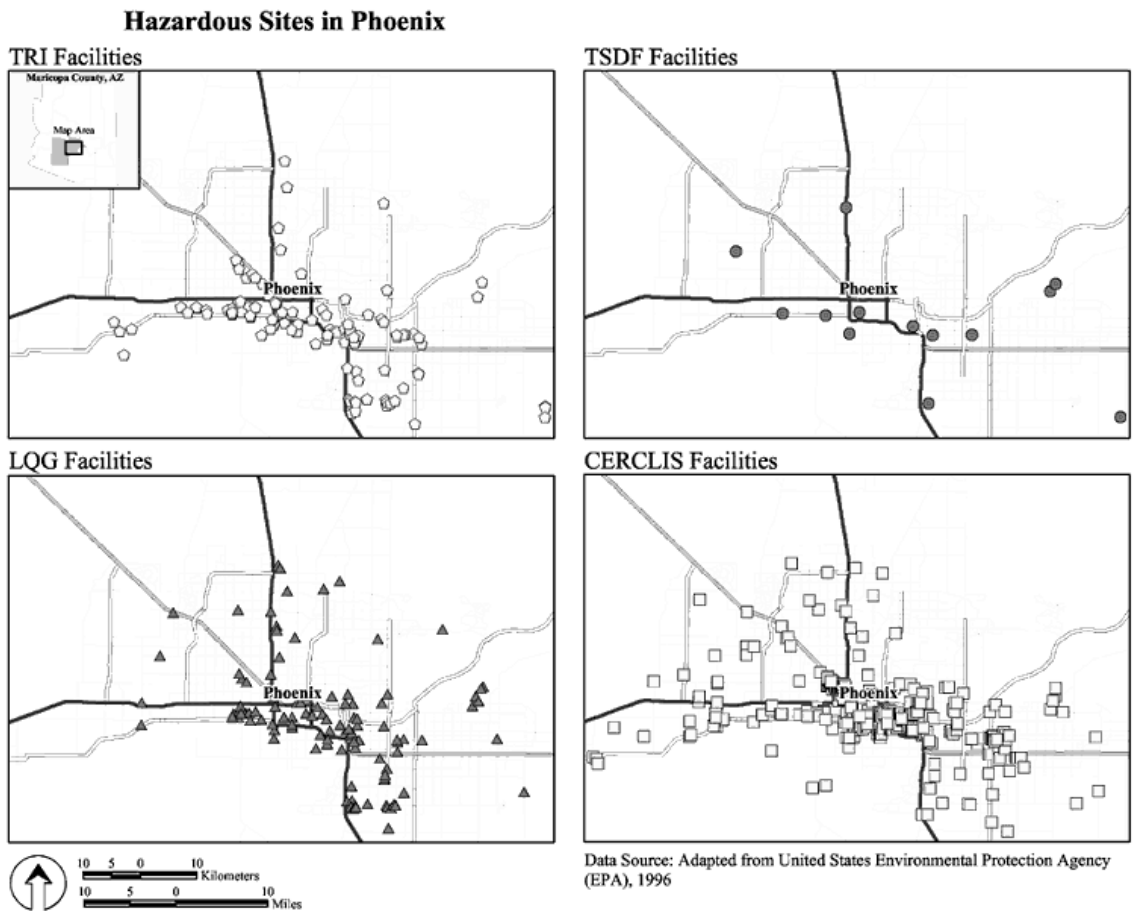
Often the data collected is obtained in a raw format. The data requires ‘cleaning’ and ‘processing’ before using it for analysis. MS Access database and SQL Query Language was extensively used to perform these tasks. Apart this, SPSS software was used for advanced calculations to determine correlations, ranks, percentage errors etc. An update of quality control of the data was maintained in a separate MS Excel sheet. Data transforming, data formatting, data reduction, sorting, ordering etc were some of the several functions that were carried out to use it for spatial analysis. This data will be next added with spatial information like co-ordinates, areas, perimeters, and index for each polygon with the help of GIS software packages. A database that carries this information of images and shapes is called a ‘geo-database’. Maintaining and managing a geo-database is a separate task altogether.

### **Spatial Methodology**

After gathering data from various sources and conceiving a mathematical concept the next step is to implement them for obtaining results. ArcGIS software provides several techniques to project the gathered data on a map and perform the required calculations. After obtaining the hazardous facilities location information, a geo-coding extension in ArcMap software helped in exactly placing the facilities as a point on a street layer map and generated a point theme. This process of locating addresses and projecting them on a map is called geo-coding. So more than 547 points were visible on the map after geo-coding the addresses. These 547 points represent the facilities and each of the facility was either a

TRI, TSD, CERLA or LQG facility. The next step is to determine the cumulative effect of the toxic release in each of the tracts.

Using the same software a radius of 1 Sq mile buffer was drawn around each facility (point) and the percentage of this area falling under each tract was noted in the database. With the help of ArcMap-calculators and editors the accumulated CHDI Score is computed for each tract as discussed above in Figure9 and ranked in the order of highest scores. This dataset which now consists of CHDI Score for each tract is superimposed on the point theme developed above to finally give digital maps as shown in Figure11. The information was used to understand how the land use changed from agriculture to industrial and the simultaneous population change in these tracts.



**Figure 11: Hazardous Sites in Phoenix**

## **Results and Conclusions**

Some of the results that were eventually analyzed from the maps are:

- The calculations and the maps show that there is a strong relation between CHDI, income and race of residents.
- From the land use maps, it is understood that there is an increase of commercial land use zoning in tract 1142 since 1970 growing from 44% to 56%. Economic decline in this tract has not affected the level of environmental hazards. (Similar such analysis was made for other important tracts in Phoenix area.)
- The tract 1139 indicates industrial/transportation encroachment wherein the historical residential character of the tract was sacrificed in the interest of highway construction and airport expansion.

More results and conclusions were made but for this present case study these were the most relevant points. This case study exemplifies how spatial, logistic, and technical aspects modified and implemented in a GIS act a step towards developing landscape ecological models for decision making, analysis and reasoning.

## CHAPTER 5

### CONCLUSIONS AND FUTURE WORK

#### **Conclusions**

Having worked as an intern it has given me experience in not just one but several fields like IT, Math, Biology and even Archeology. It was an excellent opportunity for me to combine my knowledge and skills that I gained as an Industrial Engineer with a Masters degree in CBS. My objective to provide a solution to the most IT-lacking field of agriculture and public health care was now being realized by this program.

An important issue in agricultural environmental modeling is that all the basic units (water, soil and chemicals) have a spatial distribution, and since this distribution affects the processes and dynamics of their interaction considerably, geographic information system (GIS) is emerging as an important tool in modeling. Traditional decision support systems based on crop simulation models are normally site-specific. In order to address the effect of spatial variability of soil conditions, topography and weather variables on crop production, a Geographic Information Systems (GIS) can be linked with crop simulation models. Some of the most important decisions in agricultural production, such as what crops to grow and on how much land to allocate depends on the existing knowledge base of current and future

physical conditions like soil and climate, yields and prices. Modeling of the various processes in the system helps us to understand its flow and intricacies. Its emphasis on agriculture has been increasing especially in countries like India, Africa where there exists the most diversified agro-ecosystem.

Another important field where the GIS technology is being used is the public health care. The National Center of Health Statistics (NCHS) of the US Health Department of Health and Services has been implementing GIS to study geographic relationships that affect health outcomes, public health risks, disease transmission, access to health care, and other public health concerns [10]. Geographic information systems (GIS) provide ideal platforms for the convergence of disease-specific information and their analyses in relation to population settlements, surrounding social and health services and the natural environment. They are highly suitable for analyzing epidemiological data, revealing trends and interrelationships that would be more difficult to discover in tabular format. Moreover GIS allows policy makers to easily visualize problems in relation to existing health and social services and the natural environment and so more effectively target resources.

### **Future Work**

My future work will be concentrated in the area of Health Care where Geographic Information Systems can be used for the study and control of *infectious diseases*. For long, epidemiologists have traditionally used maps when analyzing associations between location, environment, and disease. GIS is now being used in the surveillance and monitoring of vector-borne diseases, water-borne diseases, in environmental health, analysis of disease policy and planning, health situation in an area, generation and analysis of research hypotheses, identification of high-risk health groups, planning and programming of activities, and monitoring and evaluation of interventions. GIS would

enable researchers to locate high prevalence areas and populations at risk, identify areas in need of resources, and make decisions on resource allocation.

*Line Patterns* a special spatial data analytic technique in GIS can aid in the analysis of disease diffusion of several epidemics, such as the international spread of AIDS and patient-to-health care facilities flow. For example, in studying the surveillance of poliomyelitis in India, it is important to find out which type of polio is occurring in which parts of the country, as this has important implications for the disease eradication strategy employed. GIS can help in generating thematic maps - ranged color maps or proportional symbol maps to denote the intensity of a disease or a vector. In comparison with tables and charts, maps developed using GIS can be extremely effective means for communicating messages clearly even to those who are not familiar with technology. GIS keeps track of the geographical locations of service providers, customers, resources, and health plans and programmers. The most well known case of health mapping first started by John Snow's cholera map in mid-nineteenth century is yet another excellent example of how the GIS technology works.

GIS has a very vital role to play in the future. The possibilities that can be explored are limitless, depending on the skill and imaginative use of the researchers and the willingness of health sector management to resource its implementation. Future work in public health care is highly necessary to explore the tremendous potential of GIS.

## LIST OF REFERENCES

- [1] Central Arizona – Phoenix, Long-Term Ecological Research (CAP LTER), Fifth Annual Poster Symposium February 19, 2003. Toxic tracts: A historical geography of environmental inequality in Phoenix, Arizona. Bob Bolin, Scott Smith, Ed Hackett, Sara Grineski, Tim Collins, Deepa Vuppaladadium, and Jennie Kronenfeld.
- [2] Prehistoric Desert People, Retrieved from World Wide Web:  
[http://www.desertusa.com/ind1/du\\_peo\\_hoh.html](http://www.desertusa.com/ind1/du_peo_hoh.html)
- [3] Creating a Slope Grid in ArcView Spatial Analyst, Retrieved from the World Wide Web: <http://www.ar.utexas.edu/Courses/parmenter/gis/tips/slope.html>
- [4] Satya Priya and Ryosuke Shibasaki (1999). GIS-Based Regional Spatial Crop Yield Modeling.
- [5] Satya Priya, Shibasaki, Ryosuke and Shiro Ochi (1998) Modeling Spatial Crop Production: A GIS approach, Proceedings of the 19th Asian Conference on Remote Sensing, 16-20 Nov., 1998 held at Manila pp A-9-1 to A-9-6.
- [6] ESRI Support Center, Retrieved from the World Wide Web:  
<http://arcscripts.esri.com/>
- [7] NGDC, 1997 GTOPO30, Global Land One-Km Base Elevation, (Average 30-Second Elevations Grids). National Geophysical Data Center 325 Broadway, Boulder, Colorado.

- [8] Norango, S.E., and A.J. Sawyer. 1989. A Simulation Model of Northern Corn Rootworm, *Diabrotica Barberi*, Population Dynamics, Oviposition: Significance of Host-Plant Phenology. *Can. Ent.* 121:169-191.
- [9] GIS Development Tutorials. Retrieved from the World Wide Web:  
<http://www.gisdevelopment.net/tutorials/>
- [10] National Center of Health Statistics, GIS and Public Health. Retrieved from World Wide Web: <http://www.cdc.gov/nchs/gis.htm>
- [11] Knowles-Yáñez, Kim, Cherie Moritz, Jana Fry, Charles L. Redman, Matt Bucchin, and Peter H. McCartney. August 1999. Historic Land Use: Phase I Report on Generalized Land Use. Central Arizona - Phoenix Long-Term Ecological Research Contribution No. 1, Center for Environmental Studies, Arizona State University, Tempe.
- [12] Introduction to Data Analysis Using Geographic Information Systems, Retrieved from World Wide Web:  
<http://www.extension.umn.edu/distribution/naturalresources/DD5740.html#toc>