DRAFT

GIS Primer

Sonoran Desert Conservation Plan

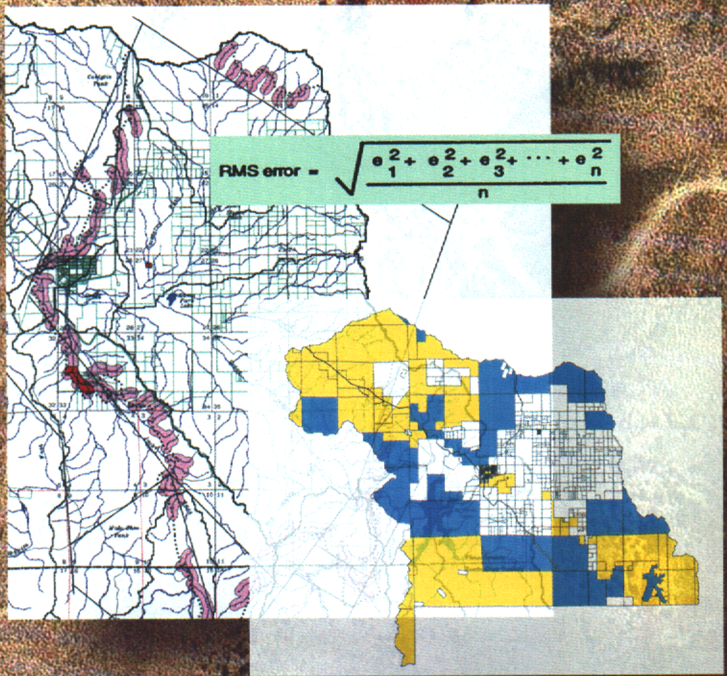
August 2000

Pima County, Arizona
Board of Supervisors

Mike Boyd, District 1
Dan Eckstrom, District 2
Sharon Bronson, Chair, District 3
Raymond J. Carroll, District 4
Raúl M. Grijalva, District 5

County Administrator

Chuck Huckelberry



$$\text{RMS error} = \sqrt{\frac{e_1^2 + e_2^2 + e_3^2 + \cdots + e_n^2}{n}}$$

MEMORANDUM

Date: August 10, 2000

To: The Honorable Chair and Members
Pima County Board of Supervisors

From: C.H. Huckelberry
County Administrator

Re: ***GIS Primer***

Pima County is one of the few local governments in the country that has a geographic information system (GIS) capable of doing the type of analysis we have come to take for granted in our day-to-day document writing. The attached report entitled *GIS Primer* defines some of the terms and provides explanations about our geographic information system. The *GIS Primer* goes further and places the technical and historical account of geographic information systems in the context of how Pima County's system has supported the development of the Sonoran Desert Conservation Plan. We have described in other documents how over 1000 data coverages have been transformed into a common coordinate system. Over 800 of these data layers are new additions to the system as a result of the Sonoran Desert Conservation Plan. Partnerships between Pima County and contributors from the community led to this region-wide compilation of data. An equally difficult job -- that of translating the variety of data sources into a single coordinate system -- has been carried out by a small group within the Technical Services Division of the County, led by Mr. John Regan, the author of the attached report, and including Mr. Felipe Morales and Mr. Mark Probstfeld.

Having employees who can work at peer level with the top geographic information system experts in the United States has led to tangible rewards. For example: On November 30, 1999, the Sonoran Desert Conservation Plan Decision Support System proposal was awarded $110,000 by the National Fish and Wildlife Foundation, and this amount was recently matched by the Packard Foundation. The Decision Support System project, described in detail in the November 1999 report entitled *Science and GIS Update,* is a collaborative effort between Pima County and University of California geographic information system experts, including Drs. Michael Gilpin and Peter Stine, who have a prestigious record of innovation in conservation mapping and modeling efforts. Last month, Pima County's biological consultant, Recon, received a national award for a geographic information system presentation at the Environmental Systems Research Institute Conference based on the *Land Cover and Data Assessment in Pima County* conducted for the Sonoran Desert Conservation Plan. County staff members working with technical teams and consultants to develop the information base of the Conservation Plan are peer level participants in these award winning efforts.

The most tangible contribution of the County's geographic information system is best described in Mr. Regan's report in this way: "GIS has played an important role in the SDCP process, producing maps by the hundreds -- literally several acres of maps, generating statistics and reports, and transferring data sets ... to the Graphic Design section where the data had the artist's touch added." The *GIS Primer* explains and helps make less mysterious some of the technical aspects of the County's geographic information system work, adding a written record to the outstanding data and map-based contributions by these County team members of the Sonoran Desert Conservation Plan initiative.

Sonoran Desert Conservation Plan
GIS Primer

by John Regan
SR GIS Analyst

Pima County Department of Transportation
Technical Services Division
Geographic Applications and Products Section
August 2000

Abstract

Maps have been used by humanity for thousands of years as portable abstractions of reality, to convey personal knowledge to others. Maps use concepts such as scale and coordinate systems to convey information to the reader in a variety of ways. In the last few decades, computer technology has evolved to the stage where it is able to handle geographic or spatial data in such a way as to link descriptive information to features located on the ground. Once this is accomplished, logical expressions can be formed to spatially join these electronic maps to produce entirely new information, then display it on a map. This innovation is called GIS - Geographic Information Systems and has transformed how we perceive the world. Topology is the concept describing relationships between spatial features mathematically and provides "intelligence" to vector spatial data used in a GIS. Questions about location, condition, trends, patterns and what is known as "what if..." modeling can be posed to GIS software and answered in the form of a map or tabular report. There are a number of analytical procedures that can be performed on GIS data sets: proximity analysis, boundary operations and spatial joins. The GIS software used by Pima County is ARC/INFO and programming is accomplished by using Arc Macro Language (AML) to run routine analytical functions and to produce maps. Pima County Department of Transportation's Technical Services Division maintains GIS hardware, software and databases with skilled personnel and has applied GIS technology to the Sonoran Desert Conservation Plan, adapting, creating or acquiring over 1000 data layers for use in the Plan. Metadata - data about data, has been developed for some of these layers. Expanded GIS functionality is described as the use of another data model - raster, and the development of a Decision Support System designed to balance conservation targets with socioeconomic objectives and constraints to meet the goals of the Sonoran Desert Conservation Plan.

Sonoran Desert Conservation Plan
GIS Primer

Introduction

Organized societies have recognized the importance of the spatial distribution of significant features on the earth's surface from the beginning of recorded history. Probably even before that, important data was collected by surveyors and geographers (or whatever they called themselves), and turned into maps. Maps were first used to describe places some distance away for a variety of purposes - military conquest or defense, taxation, colonization, agricultural exploitation. Land surveyors or *agrimensores* as they were known in Roman times, were important to governments and some of their work can still be seen on the European landscape to this day. Institutionalized mapping generally waned with the decline of the Roman Empire but by the 18th century governments again came to realize the importance of large scale systematic mapping of their domains. Branches of government and some private agencies were charged with the task of producing detailed maps of entire countries. As the scientific knowledge base grew, so did the amount of information requiring mapping. Natural resources - soils, geomorphology, geology and a new idea called *ecology* - the relationships between the environment and organisms , came to be recognized as important. Topographic maps had long been developed as a general purpose map since they don't usually have a specific purpose other than to display land relief. Now new maps with a single subject, with a more specific purpose or *thematic maps*, began to appear with themes such as soil series, rock types or land use.

The need for land evaluation increased along with human population growth and the required food production demands. Technology advanced and was applied to the making of maps. Up to this time, all maps had one thing in common: the spatial data was drawn on a sheet of paper or something similar, and the information was displayed as points, lines and polygons using a diverse set of symbols, colors and text. A legend explained the symbolization and additional narrative could discuss the purpose of the map or anything else the map maker desired. Essentially the map <u>was</u> the database. This meant the original data had to be simplified or classified in order to clearly represent it on a map. Many details were lost in the translation. Large areas had to be represented on a number of sheets and even today it has almost become an unwritten rule that the area one is interested in will always be at the intersection of four adjacent map sheets. If someone wanted to retrieve and reorder/recombine data for another purpose, it was a difficult task at best, not to mention expensive and time consuming.

Computers appeared on the scene in great proliferation in the last thirty years and it was a small step to transfer mapmaking tasks to computerized applications. There were a number of early map programs in the 1960s that came out of Harvard Graduate School of Design's Laboratory for Computer Graphics - SYMAP, GRID, IMGRID, GEOMAP. These programs used a raster approach to overprinting of lineprinter characters to produce grey scales, displaying the results of a geographic analysis. Cartographers had already adapted computer technology to the automation of map making, mostly automated drafting and the production of masters for printed maps, but rejected the crude products produced by innovative programs like SYMAP. By the late 1970s, computer-assisted cartography had progressed to the degree there were hundreds of mapping applications in government and the private sector in the U.S. alone. The introduction of the new technology did not always lead to immediate cost savings - systems were expensive, digital data was difficult to acquire and trained staff were hard to find. Software development was slow because the manufacturers viewed the computerized mapping market as so diverse that the huge investments required to create a software program to meet everyone's needs would not be profitable. This left purchasers of the technology no choice other than to hire their own programmers to adapt their system of choice to their own requirements. There was also uncertainty in organizations where the technology was being implemented. Were the new tools supposed to replace existing manual techniques to do work faster or were they supposed to provide a wider range of possibilities to the user? Sometimes computer-assisted mapping tools were added to an agency but work flow and organizational changes were not implemented to accommodate the new technology. Often the expectation was someone could handle this new technology during their lunch hour in addition to their existing job. After all, the computer did all the work! Hardware changed so quickly that often times the technology was not fully implemented and the hardware had become obsolete before any real benefit was realized to justify the initial expenditures.

The concept that a powerful tool for spatial analysis inherent in a mapped database was not realized at many sites where computerized mapping was installed. Poiker (1982) commented that *"...computer cartography is like a person with the body of an athlete in his prime time and the mind of a child"*. Computerized mapping consisted of two main trends: automation of existing tasks, emphasizing cartographic accuracy and visual quality, and the concern with spatial analysis at the expense of good graphics. It was extremely difficult to solve the computational and topological problems associated with the coding of spatial data.

A number of disciplines - military, engineering, planning, utilities - all needed some powerful tools to collect, store, retrieve, transform, analyze and display spatial data for specific purposes. In short, they all needed what became known as a Geographic Information System - GIS. No one seems to know who coined this moniker. Exactly who the first one to develop the link between spatial and descriptive data, apply topological models and package the whole system is also difficult to say. Some attribute it to the 1978 provincial Forest Act in Canada requiring all timber agencies to provide a detailed forest inventory at frequent intervals. Others attribute it to ESRI - Environmental Systems Research Institute, makers of ARC/INFO, sometime in the early 1980s. Since then there has been a proliferation of software packages that perform spatial analysis, from the complex to the simple, expensive to relatively cheap, from taking a full year to learn to a few hours to be somewhat functional, from a mainframe to the desktop PC. Taken as a whole, GIS technology has come a long way in the last two decades. The end result has been a change in how the world is viewed by a plethora of organizations. Many of them make decisions that affect us all and GIS has contributed to the overall picture by providing unique views of the world as we know it.

GIS technology is being applied to Pima County's Sonoran Desert Conservation Plan, both as a mapping and statistical tool as well as a spatial modeling device. The following pages describe some GIS basic concepts, how they are applied to the project and a brief overview of some of the techniques used in mapping procedures.

Table of Contents

Table of Contents

List of Figures

Maps

Everyone knows what a map is: a map is a graphic representation of a region of the earth. It is also much more. Who has not looked at a map and felt the promise it offered of places to go and things to see as it silently communicated a sense of place to the viewer? The word "map" is often used as a metaphor to describe points of clarification, such as to "map out a strategy". In the ancient world maps were oriented to the east, presumably because that was where the sun rose and perhaps because it was believed to be the direction to paradise. Maps have been around for a long time, probably before written language, scratched out in the dirt before a camp fire, traced on a piece of hide with charcoal to show where the next water hole was or a map of the world etched on a clay tablet (figure 1). In essence a map is a mechanism of transferring personal knowledge to a portable graphic medium, putting things into perspective, so to speak.

Figure 1.
Early Mesopotamian Map
of the World on a Clay Tablet
-Wilford, 1981

Basic map concepts

Map scale

One of the most important parts of a map is distance, the relationship between two spatial features. How far this is from that is often the first question asked: how far is my house from the airport? How far from where I live to where I work? In order to extract this information from a map, the map must be converted to real world by using the ratio of map distance to distance on the ground. This ratio is called *map scale* and is expressed in three different ways.

1. Ratio or fraction: 1:63,360
-one inch on the map equals 63,360 inches on the ground

2. Verbal scale: one inch equals one mile
-one inch on the map equals one mile on the ground

3. Graphic or bar scale
-graphic representation of one inch on the map equals one mile on the ground

Maps are also referred to as large scale or small scale, often times resulting in confusion. The larger the scale, the larger the feature will appear on the map. This is expressed as a smaller number:

1:1,200 is a larger scale map than 1:1,000,000

Figure 2 displays differences in map scale. 1:1,200 shows more detail in a smaller area with a higher precision and resolution. 1:1,000,000 shows less detail over a larger area with less resolution.

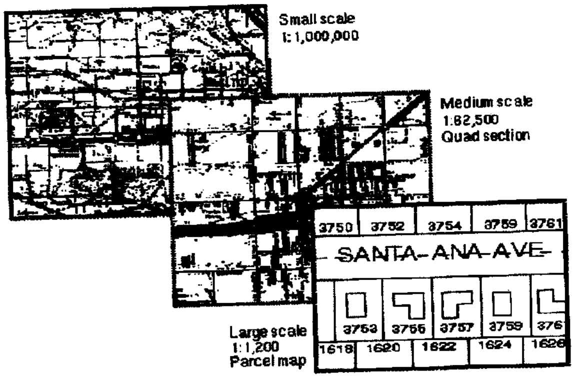


Figure 2. Map scales

Cartesian coordinate system

A coordinate system provides a design for defining the real world in terms of map locations. A Cartesian coordinate system uses two axes to represent direction: one for horizontal east-west (x), and one for vertical north-south (y) direction. The point at which they intersect is called the origin. Objects in geographic space are defined relative to the origin, using the x,y notation where x refers to the distance along the horizontal axis, and y refers to the distance along the vertical axis. The origin is defined as 0,0, as shown in figure 3.

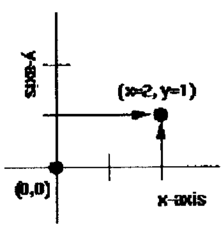


Figure 3.
Cartesian Coordinate System

This system assumes the surface to be flat or a plane similar to a paper map. Of course this is not the case, since the earth is round. The dilemma is similar to cutting a hollow ball in half then attempting to flatten it which cannot be done without altering its shape. Map projection is the process developed to deal with the problem of projecting a spherical surface to a flat surface. The results are always distortions in distance, area, shape or direction. This complex process has resulted in many projections, some minimizing distortion in one area at the expense of another. The familiar maps we grew up with in the classroom were projections designed for mid-latitude countries such as the U.S. that made it seem as if it was the center of the universe and all other parts of the world were diminished in size and by silent implication, importance.

The Development of GIS

Over the years map making has evolved into the production of more sophisticated products using standardized symbols and legends. The top of the map has been oriented to the north and complex approaches have been worked out to display the surface of a sphere on a flat surface, getting names like Albers Conic Equal-Area and Oblique Mercator. Many of the laborious tasks map making involves have been automated in recent years by electronic data processing involving the use of computers. Software known as *Geographic Information Systems (GIS)* has been developed to handle large complex spatial data sets.

Scarcely two decades ago, a technology focusing on the integration of map production and the management of spatial data was born and it was called GIS. GIS is defined as computer hardware, software, geographic data and most importantly, skilled operators - often termed "warmware", combined to capture, store, update, manipulate, analyze and display geographically referenced information. Geographically referenced refers to almost anything that occupies a place on the face of the Earth. The technology was destined to have a major impact on how spatial data was managed. The approach simply took some of the existing cartographic software packages - CAD or Computer-Aided Drafting, added a relational database management system (RDBMS) to link descriptive data to spatial features and a geometric concept known as *topology*. This combination allowed the user of the software to literally ask questions of the database and have the answers displayed graphically in the form of a map. The world has never been the same since.

One of the keys to this breakthrough was the application of the concept of topology to a *data model* or representation of spatial data used by a computer. Topology was nothing new. It has long been a concept used in geometry to describe spatial relationships in terms of area definition, connectivity and contiguity. What was new was the idea to apply the concept to lines and polygons stored digitally inside a computer as binary 0's and 1's.

Simply stated, topology is the mathematical representation of the relationships of spatial features to each other. This is something we do as humans without even thinking. The spatial world we live in is based on topological relationships but we don't even notice them. In our back yards, the door opens on to an area probably bounded by a fence defining our property - a polygon. The adjacent properties of our neighbors are also delimited by a fence. A driveway has a direction and connectivity to the street. There's a tree or two a specific distance and direction from our door - a point feature with its own unique attributes. Perhaps there's a tool shed in the northeast corner that could be considered an "island polygon" - an area with properties unique from the other properties that make up our back yard. There are probably water and gas lines that are buried under our yard and connect the utility meters to our homes - linear features crossing the polygon of our back yard, also with direction and connectivity. All these features are spatially related to each other, connected or adjacent, occupying unique places on the surface of the earth and having unique spatial relationships to each other. Our brains process this information, seemingly without conscious effort, but a computer has to have this information defined mathematically by topology. That's what ultimately gives a GIS "intelligence", allowing the user to perform a complex analysis with spatial information.

What is a GIS?

The following GIS discussion is based on a particular brand of software - ARC/INFO, developed by Environmental Systems Research Institute (ESRI). The concepts are common to other GIS software packages but some of the terminology is specific to ARC/INFO. An example is the *coverage* - the basic unit for vector data storage representing a single set of geographic features - parcels, roads, washes, wells, etc. A coverage supports the *georelational model*, a geographic data model that represents both spatial features and their related descriptive data.

The classic definition of a GIS is:

> *An organized collection of computer hardware, software, geographic data and personnel designed to efficiently capture, store, update, manipulate, analyze and display all forms of geographically referenced information. (ESRI 1992)*

Figure 4 shows a conceptual view of how GIS views the real-world as layers of information. Data is captured through digitizing or scanning of hard copy maps registered to a coordinate system, then attributed according to the needs of the user.

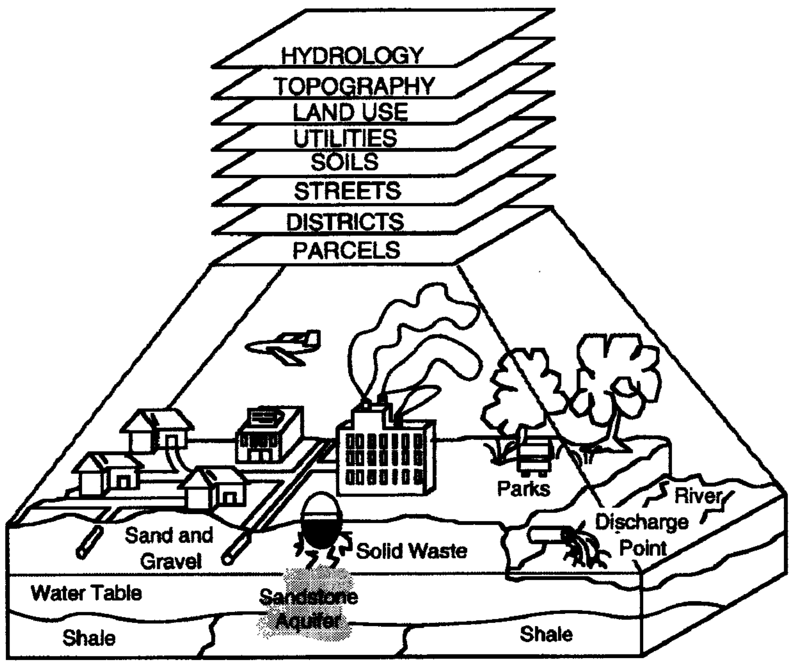


Figure 4.
GIS view of the real world

Components of a GIS include a skilled user, software tools, a database that is an abstraction or simplification of the real world, as shown in figure 5.

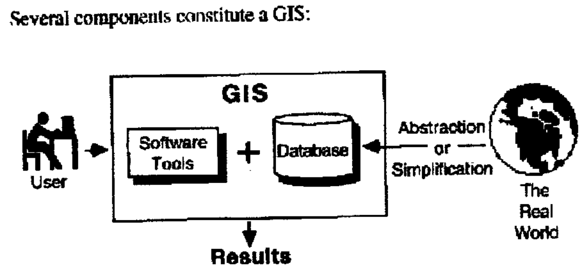


Figure 5.
GIS components

Representing descriptive information in a GIS

The power of a GIS comes from being able to link descriptive information or *attributes* to spatial features then perform a spatial analysis on this information before displaying it graphically. This is the essence of the georelational model For example, attributes are stored for each polygon in a parcel coverage that describe the parcel and include:

$RECNO	*TAXCODE*	*OWNER_NAME*	*OWNER_ADDRESS*	*FCV*
1	*21432265A*	*SMITH, JOHN*	*201 NORTH STONE AV*	*112,655*
2	*214366220*	*JONES, JIM*	*32 NORTH STONE AV*	*65,300*
3	*214422001B*	*MURAWSKI, JOE*	*66 EAST SPEEDWAY*	*125,200*

This set of descriptive information is known as an *attribute table*. Each row is a record and contains descriptive information about a single feature - a parcel of land in this case. The same columns or fields appear in each record and are referred to as *items*.

The georelational model links these attributes to spatial features and may access information stored in other tables outside the GIS. One important concept is how this information is linked to the spatial feature. The unique identifier associated with the feature is the same identifier attached to the attribute. This maintains a *one-to-one relate* between spatial features and the attribute records. When this connection is established, a map can be created based on attribute information, or display the attribute values in a tabular format. In the following example in figure 6, the polygon-arc topology table defines the polygon, while the polygon attribute table provides attributes of the polygon. The two tables are related to each other through the common attribute identifying the polygon's unique identifier, the polygon in this case, as shown in figure 6.

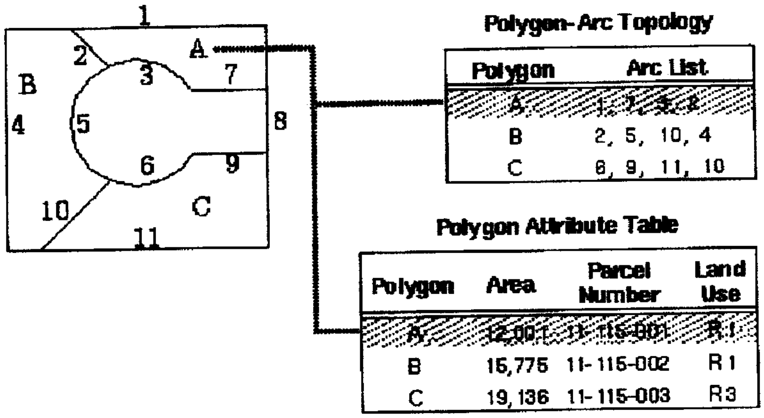


Figure 6.
One-to-one relate

A *many-to-one relate* can also be performed by using a *relate* as long as there is a common attribute in both tables. In this case, the common item is the parcel number and is referred to as a *key*. Parcel C has multiple owners as shown in figure 7.

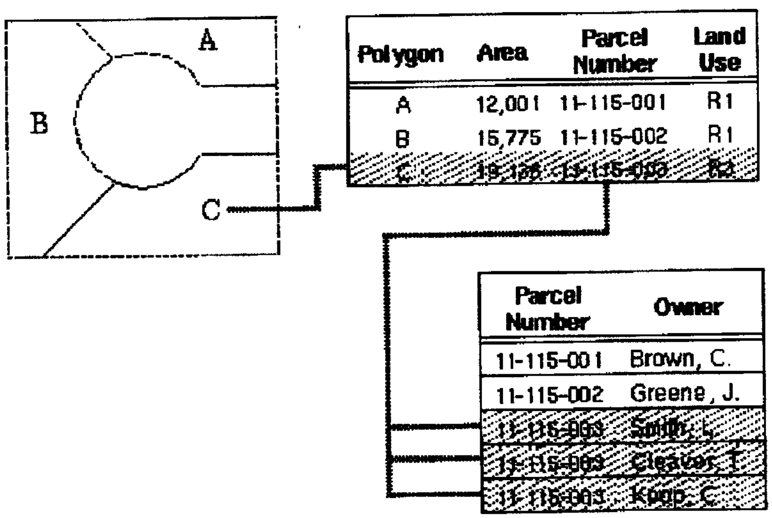


Figure 7.
Many-to-one relate

Why use GIS?

1. Provide better information to decision makers! A succinct presentation of complex sets of information can make all the difference in the outcome of an issue. Often times decisions have to be made without adequate information; GIS can help avoid these moments. For example, Pima County has over 330,000 individual parcels. Sorting out specific characteristics for an area of choice could result in a complete meltdown of staff productivity if they have to perform calculations by hand, viewing and handling literally thousands of records. When computerized, it takes a matter of minutes. Many features of the landscape cannot be seen any other way than through the perception of a GIS, using the power of modern desktop computers.

2. Share data with other departments and agencies. It is estimated that 80% of the data local government uses is spatially related, that is, it can be drawn on a map and used in a GIS. Once the data is captured in a digital format, it can be used by many groups, avoiding costly duplication of effort and saving enormous amounts of time.

3. Make maps of combinations of data sets as well as the results of a specific analysis or query. Maps can be centered on any location, at almost any scale, symbolized to the client's specifications. Pattern recognition comes into play once rows and columns of numbers are turned into a map. The way data is viewed can have a profound effect on the conclusions that are developed.

What are some of the questions GIS can answer?

1. Location. What is at...?

What exists at a particular place? Example: what town or towns are in zipcode 85601?

2. Condition. Where is it?

Condition requires spatial analysis, going beyond merely identifying what exists at a location. The object here is to identify where specific conditions occur. Example: where are residential properties with dwellings constructed before 1950?

3. Trends. What has changed since...?

This may involve both location and condition but seeks to find the differences within a specific area over time. Example: how many wells have been added since 1975, where are they and what is their legal status?

4. Patterns. What spatial patterns exist?

Here we're seeking apparent patterns contributing to cause/effect relationships and anomalies that don't fit. Example: why is there an abnormally low property tax rate in a certain area? Why is that different than other areas?

5. Modeling. What if...?

"What if..." questions are posed to determine what might happen if some action was taken or if no action was taken on a particular issue. Example: what would happen to the water table if an additional 2000 wells were drilled in a small watershed over the next ten years? This is the highest level of analysis and requires extensive spatial and descriptive information to correctly model the outcome of events.

GIS Data Models

A data model is simply an approach to representing spatial or geographically-referenced data in a GIS. There are two data models used in Pima County GIS: vector and raster. The first portion of the discussion will focus on vector data in some detail; raster will be addressed later in the modeling section. All figures were developed by ESRI, the developer of the Arc/Info software, and adapted to this report.

We're all familiar with the way maps represent geographic data - as points, lines and polygons colored to represent the specific information the map is intended to convey to the reader. This is essentially the vector data model at work on the paper map being held in your hands. A Cartesian coordinate system of x,y pairs refer to real-world locations where each feature is recorded as a single x,y coordinate. Points are features too small to be represented as lines or polygons and show up as a single point. Lines are series of ordered x,y coordinates while polygons represent areas by a series of coordinate pairs forming a line that closes on itself. Figure 8 shows a typical representation of how georaphic features are stored in ARC/INFO.

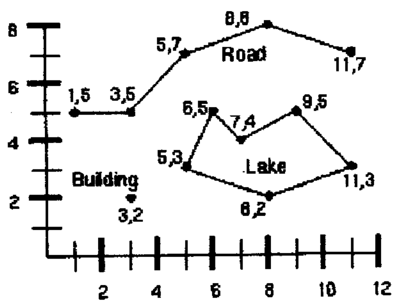
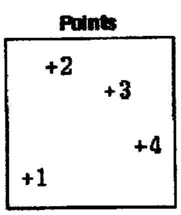


Figure 8.
Geographic
Feature Storage

The software assigns a unique identifier to each feature to keep track of them. The list of coordinates is then associated with the tag or identifier, point numbers in figure 9.



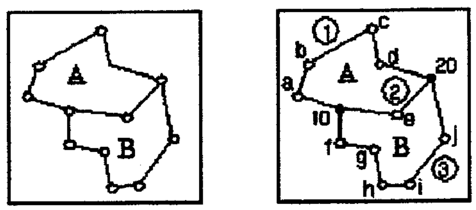
Point Number	x,y Coordinates
1	2,2
2	3,6
3	5,5
4	6,3

Figure 9.
Unique identifiers

Arc-node data structure

Boundaries of two adjacent polygons share a common boundary rather than having two lines or *arcs* duplicate the function of separating the two features. Storing coordinates for points shared by a number of lines is inefficient as is storing each polygon as a closed string of coordinates, so an *arc-node data structure* is used. This approach uses nodes to construct arcs and arcs to construct polygons. The two endpoints of an arc are called *nodes* and may connect two or more arcs. *Arcs* are made up of two nodes and an ordered series of points or *vertices* defining its shape. Vertices and nodes are stored as x,y coordinates.

In figure 10, polygons A and B are represented by a series of connected coordinate pairs. Nodes are created where lines intersect, arcs are created between the nodes. Vertices provide shape and the two polygons, A and B, are formed by the arcs.



Arc Number	Start Node	Vertices	End Node
1	20	d, c, b, a	10
2	10	e	20
3	10	f, g, h, i, j	20

Polygon	Arc List
A	1, 2
B	2, 3

Figure 10.
Arc-node data structure

Topology

Topology, as previously described, is the relationship between spatial features. In the case of polygons, these relationships are expressed as lists of arcs that make up the boundary. There are a number of advantages to creating and storing topological relationships in a computer. First, data is stored efficiently, resulting in faster processing time and less disk space utilized. When topology is present, analytical functions such as combining adjacent polygons with similar characteristics, overlay analysis, identifying adjacent features and modeling flow through connecting lines are possible.

Three major topological concepts are supported with the arc-node data structure:

1. Area definition: connected arcs surrounding an area define a polygon
2. Connectivity: arcs connect to each other at nodes
3. Contiguity: arcs have left and right sides as well as direction

Area definition

Parcels of land, administrative districts, soil associations and other geographic features are represented in this data model by one or more boundaries defining a polygon. If, for example, a lake had an island in the middle, it would actually be represented by two boundaries, one for the outer edge and one for the inner edge or the island's outline. Topology is used to define these relationships by the arc-node data structure as shown in figure 11.

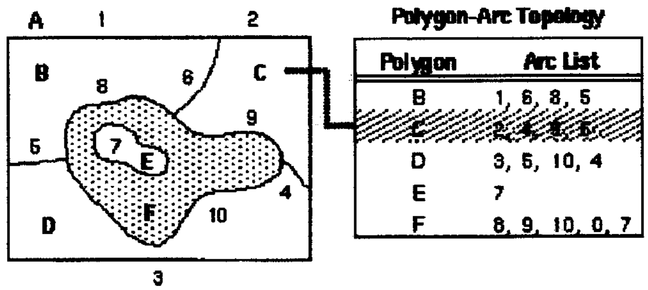


Figure 11.
Polygon-arc
topology

A polygon is represented as an ordered list of arcs rather than a closed loop of x,y coordinates. Polygon F is made up of arcs 8,9,10 and 7. The 0 before the 7 indicates this arc creates an island in the polygon. Each arc appears in the list in two polygons, arc 5 is in polygon B and polygon D. Since the polygon is defined by the list of arcs that make up its boundary, arc coordinates are stored only once. Polygon-arc topology reduces the amount of data stored and ensures that adjacent polygon boundaries do not overlap.

Connectivity

Connectivity identifies a route to a specific location, or the connections between stream segments. An arc is defined by its two endpoints - the from-node indicates where the arc begins while the to-node shows where it ends. This is called *arc-node topology* and is supported through the arc-node list displayed below in figure 12. The list shows the from-node and to-node of each arc. Connected arcs are discerned by a search of the list for common node numbers. Below, arcs 1,2 and 3 all intersect because they share node 11. The computer software determines it is possible to travel along arc 1, turn on to arc 3 because they share common node 11, but it is not possible to turn directly from arc 1 on to arc 5 - they don't have a node in common.

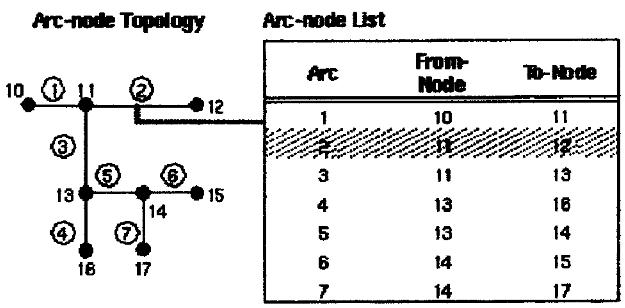
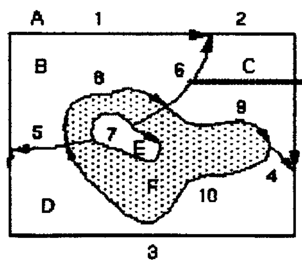


Figure 12.
Arc-node
topology

Contiguity

Two geographic features sharing a common boundary are called *adjacent*. *Contiguity* is the topological concept where adjacency is determined in the vector data model. An arc is defined by the from-node and the to-node, indicating the direction of the arc. This information determines which polygons are on the left and right side of the arc in question. In figure 13, polygon B is on the left side of arc 6; polygon C is on the right, leading to the conclusion polygons B and C are adjacent.

The label for polygon A is outside the boundary of the defined area. This is a peculiarity of ARC/INFO known as the *universal polygon*, representing the world outside the study area. This ensures that all arcs in a coverage always have a left and right side defined.

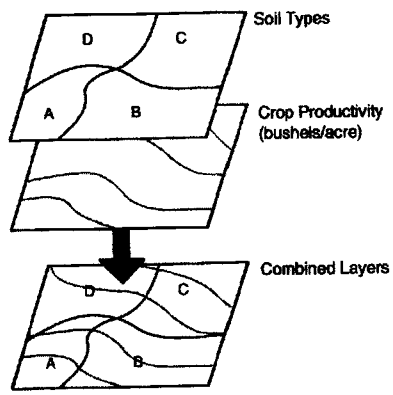


Left-Right Topology

Arc	Left Polygon	Right Polygon
1	A	B
2	A	C
3	A	D
4	C	D
5	D	B
6	B	C
7	F	E
8	B	F
9	C	F
10	D	F

Figure 13.
Left-right topology

<u>Application of topology</u>
This mathematical representation of spatial relationships opens a variety of analytical doors to the GIS user. Once topology has been added to our data model, some interesting recombinations of spatial data can be performed to come up with entirely new information. Most commonly used is the overlay analysis (figure 14), where two attributed data layers are combined to create a third layer with the attributes of both original layers. An example of this approach is a simple query where we want to know the acreage of crops by soil type. The operation results in a new series of polygons where a frequency adds up the acres of crops by soil type..



When data boundaries between layers don't match, the layers can be joined, creating a new layer containing the characteristics of both.

Figure 14.
Overlay analysis

GIS Analysis in the Vector Data Model

A geographic analysis is generally performed by combining one or more data layers in a number of ways. It is not unusual to also run an analysis on a single coverage such as a parcel base. There are a number of tools available to the ARC/INFO user that take advantage of the topological data structure in the GIS database. These are described in general conceptual terms rather than attempting to provide a technical guide to actually perform the operations. They include proximity analysis, boundary operations, logical operations and spatial joins.

<u>Proximity Analysis</u>

This set of commands creates information about the distances between, among or around coverage features. Their function is to determine which features are closest to other features, how far apart features are and how much area is within a specific distance.

1. BUFFER generates a buffer zone around coverage features - polygons, lines or points. The user specifies the distance around the feature and the command creates a polygon coverage that can be used in subsequent operations. Example: a property owner applies for a rezoning and the county must notify by mail all property owners within 500 feet of the parcel in question. A 500 foot buffer around the parcel is used to clip neighboring parcels from the full parcel base. Names and addresses of affected property owners are selected so notification can be sent out.

2. NEAR compares the distance from specific points to features in a second coverage. Example: determine the distance from a set of wells to the nearest stream for regulatory purposes.

3. POINTDISTANCE is used to compare the distance between points in one coverage to points in another coverage. A search radius may be specified. Example: several wells have reported contamination from localized pollutants. The task is to determine the distance to other wells in the vicinity in an attempt to monitor progress of the contamination plume.

4. THIESSEN converts a set of points to Thiessen or proximal polygons. Each point is apportioned to a region so that every location in a region is nearer to that region's point than to any other point. Example: rain gauge locations are converted to provide an approximation of how complete the network of monitoring station represents an area.

Figure 15 on the next page displays the proximity commands.

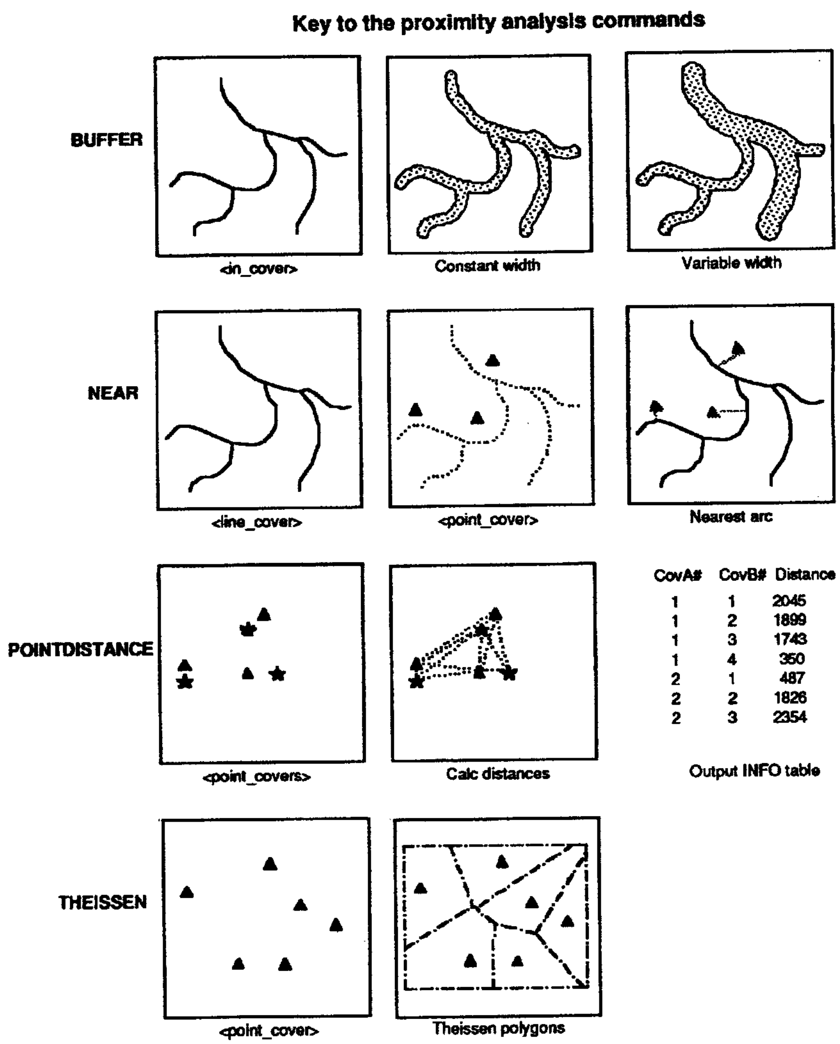


Figure 15. Proximity analysis commands

Boundary Operations

Boundary operations allow the user to process features within a geographic boundary. These operations include coverage automation, updating and isolating specific areas. A geographic boundary must be defined, and all commands create new coverages. Figure 16 displays boundary operations.

1. CLIP uses a polygon coverage to reduce another coverage to the boundaries of the study area defined by the clip coverage. If features fall within the boundary of the clip coverage they are retained in the new coverage. CLIP can also be used in the automation process to clip digitized coverages to standard boundaries, ensuring edgematching and correct topology. Example: a watershed boundary is used as the "clip" coverage to determine acreage of natural vegetation series just in that study area.

2. ERASE uses a polygon coverage to "erase" the features within the boundaries of the study area in question. If features fall within the boundary of the erase coverage they are removed from the new coverage. Example: a proposed subdivision has been added to the parcel coverage and the study calls for the display of only existing plats in a specific township. An erase coverage of the proposed subdivisions is used to clear the linework of proposed lot splits, leaving only the platted lots for display.
3. UPDATE takes a coverage of new or updated features and replaces the existing features of another coverage just inside the boundary of the new coverage. Example: a vacant parcel of land is subdivided and platted. A coverage of that area is developed independently and added to the master coverage using the UPDATE command.
4. SPLIT breaks a single coverage into many coverages.
5. MAPJOIN allows the user to join two or more adjacent coverages into one coverage. The command also allows the use of a CLIP coverage to reduce the size of the final coverage. Example: there are a dozen subdivision coverages covering more than a single watershed. By using the watershed boundary as the CLIP coverage, we can MAPJOIN the dozen coverages into one and have the excess areas clipped to the watershed boundary.
6. APPEND is similar but does not provide an option for use of a CLIP coverage.

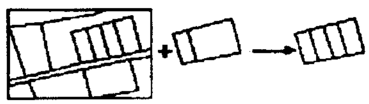
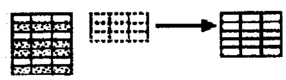
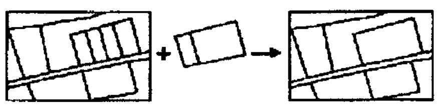
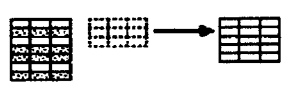
Boundary operations	Spatial results	Tabular results
CLIP		
POLY LINE POINT NET LINK	Subset of existing features ■ retains features inside the external polygon boundary ■ clipping uses topological overlay	Subset of records ■ same attributes and User-IDs as <in_cover> but fewer records, areas are adjusted
ERASE		
POLY LINE POINT NET LINK	Subset of existing features ■ erases features inside the external polygon boundary ■ erasing uses topological overlay	Subset of records ■ same attributes and User-IDs as <in_cover> but fewer records, areas are adjusted

Figure 16.
Boundary operations

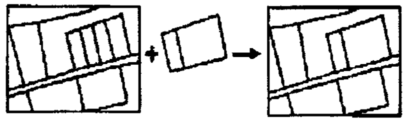
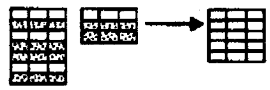
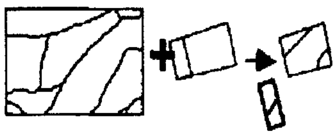
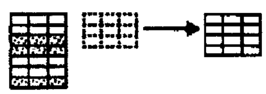
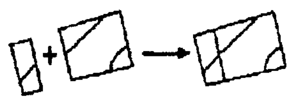
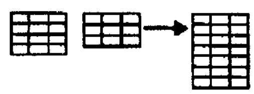
Boundary operations	Spatial results	Tabular results
UPDATE		
POLY NET	Replaces existing polygons ■ the external polygon boundary is used to 'cut-and-paste' new features ■ updating uses topological overlay	Replaces old records ■ attributes in <update_cover> must match <in_cover> ■ <update_cover> User-IDs are used on updated areas, creating possibility of duplicate User-IDs
SPLIT		
POLY LINE POINT NET LINK	Subset of existing features ■ retains features inside internal polygon boundaries ■ splitting uses topological overlay	Subset of records ■ same attributes and User-IDs as <in_cover> but fewer records, areas are adjusted
MAPJOIN		
POLY NET	Merge adjacent polygons ■ polygons from many coverages are combined, including their previous borders ■ mapjoining uses topological 'clean'	Records are combined ■ attributes in all coverages being combined must match ■ User-IDs can be offset or remain the same (duplicate values possible)

Figure 16.
Boundary operations

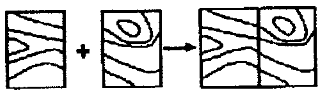
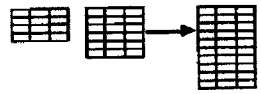
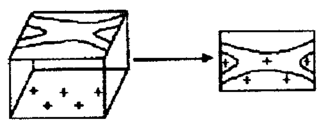
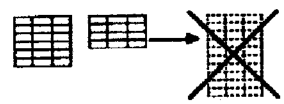
Boundary operations	Spatial results	Tabular results
APPEND POLY LINE POINT NET LINK	 Merge similar features – adjacent or coincident ■ features from several coverages are combined ■ topology must be reestablished with CLEAN or BUILD	 Records are combined ■ attributes in all coverages being combined must match ■ User-IDs can be offset or remain the same (duplicate values possible)
APPEND (all types) NOTEST option	 Merge different features – adjacent or coincident ■ features from several coverages are combined ■ topology must be reestablished with CLEAN or BUILD	 Feature attributes are lost ■ attributes don't matter because they aren't appended

Figure 16.
Boundary operations

Performing Logical Operations

Using logical expressions to select features by their attributes can create new coverages. Features can be merged or extracted by the use of three basic commands. Figure 17 displays the logical operator commands.

Feature Extraction

1. RESELECT allows the selection of a subset of geographic features using a logical expression on feature attributes. It creates a new coverage containing only the features selected by the logical expression. Example: reselect from the parcel base only properties owned by the State of Arizona. The logical expression would read RESELECT PARCELS POLY MAIL1 = 'STATE OF ARIZONA'

Feature Merging

2. DISSOLVE selects an item used to evaluate existing polygon boundaries and merge adjacent polygons with the same values. Example: dissolve a soils coverage into the primary soil classifications based on attribute values, generalizing the data.

3. ELIMINATE acts on what is known as sliver polygons that often occur when overlapping polygons from one or more cover are joined together. ELIMINATE performs a merging by deleting polygons identified, usually by a specified area threshold size.

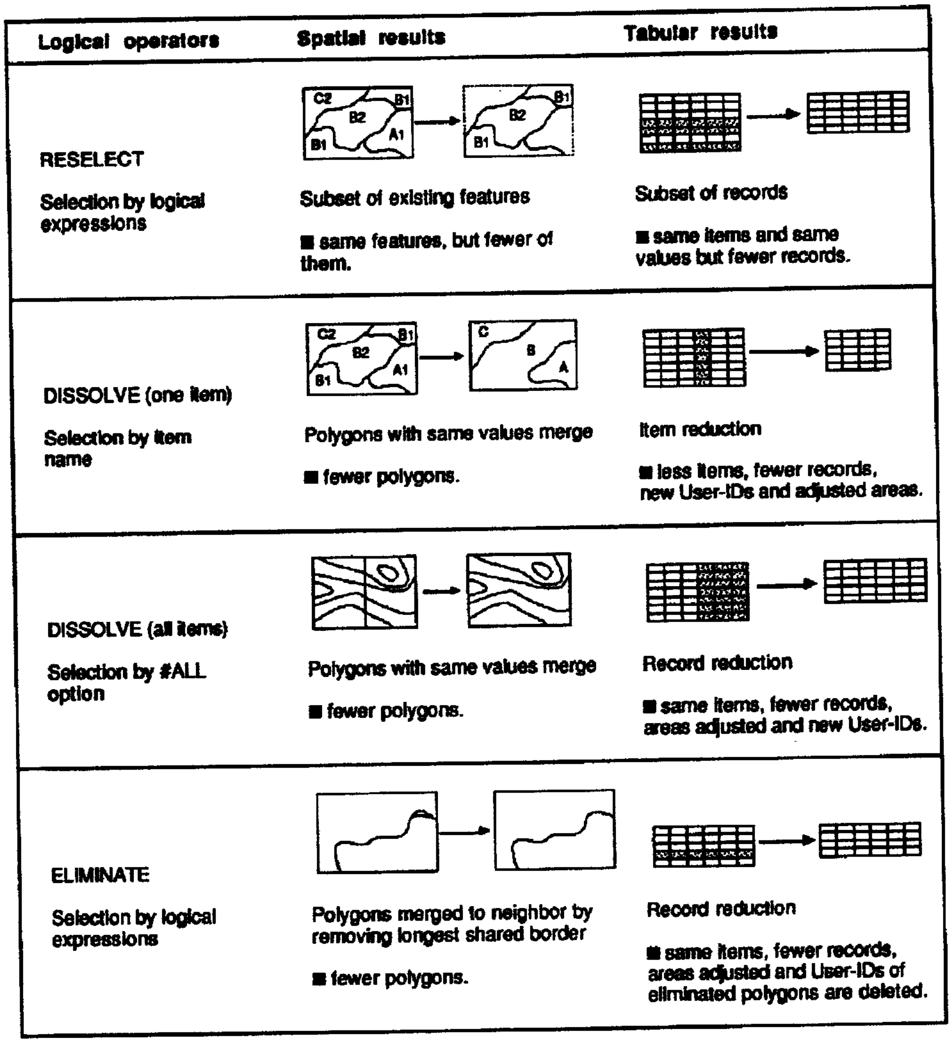


Figure 17. Logical operations

Spatial Joins

When features of two coverages are combined by some form of overlay analysis, the attributes and features are joined to create a new coverage. There are three types of operations: point-in-polygon, line-in-polygon and polygon-in-polygon. Polygon-on-polygon has three commands used to produce different results: UNION, INTERSECT and IDENTITY.

1. POINT-IN-POLYGON OVERLAY essentially combines the attributes of a point coverage with those of a polygon coverage. Example: what wells are located in specific property holdings? Figure 18 shows the point-in-poly operation.

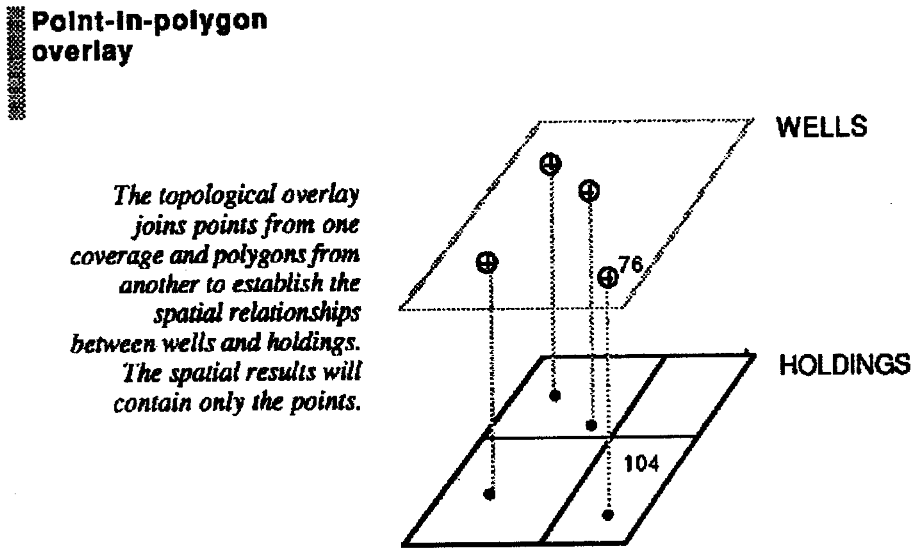


Figure 18.
Point-in-poly overlay

2. LINE-IN-POLYGON OVERLAY provides a mechanism to join lines to polygons to determine which lines fall within what polygons. As before, attributes of both coverages are joined in this operation. Example: what roads fall in unincorporated Pima County? Figure 19 illustrates this concept.

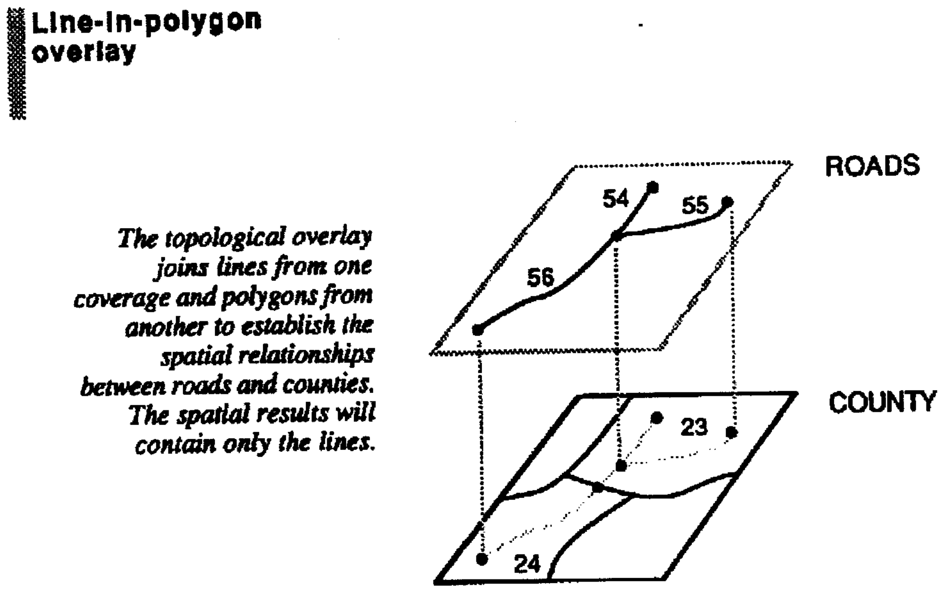


Figure 19.
Line-in-polygon overlay

3. POLYGON-ON-POLYGON OVERLAY combines two coverages of polygons into one new coverage with the attributes of both original coverages. Example: what types of soils are found in specific parcels? Figure 20 shows the polygon-on-polygon overlay.

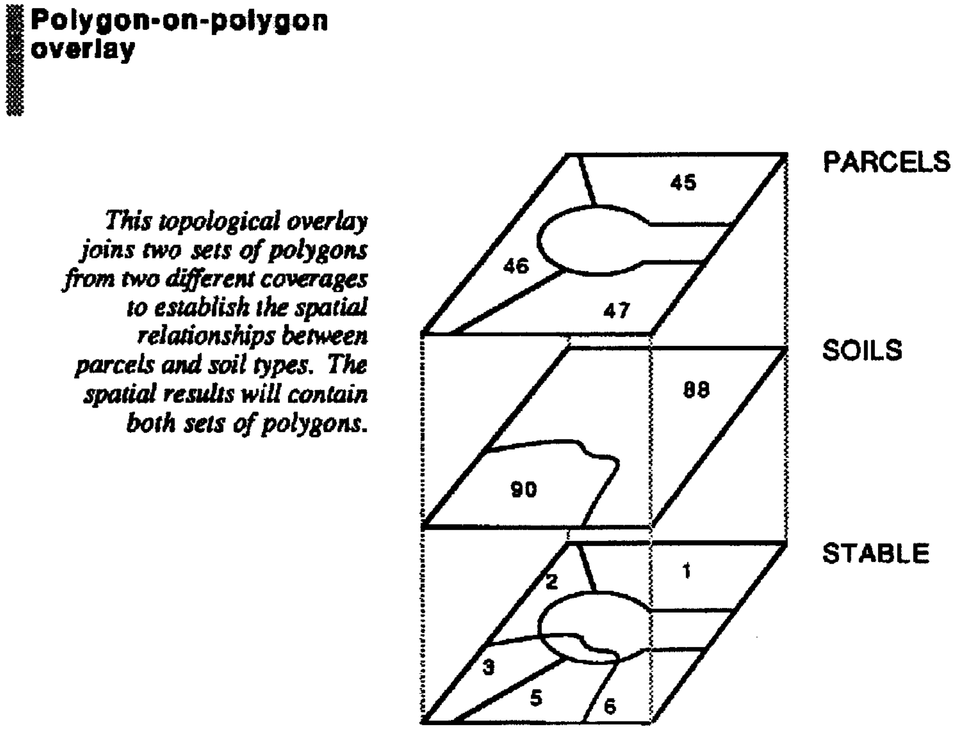


Figure 20. Polygon-on-polygon overlay

There are several ways to approach the polygon-on-polygon overlay, using Venn diagrams to predict the results. The concept of set theory is used in geometry and is used to illustrate the different possible combinations and the Venn diagrams in figure 21 illustrate these differences when the UNION, INTERSECT or IDENTITY commands are used for the polygon-in-polygon operation is performed.

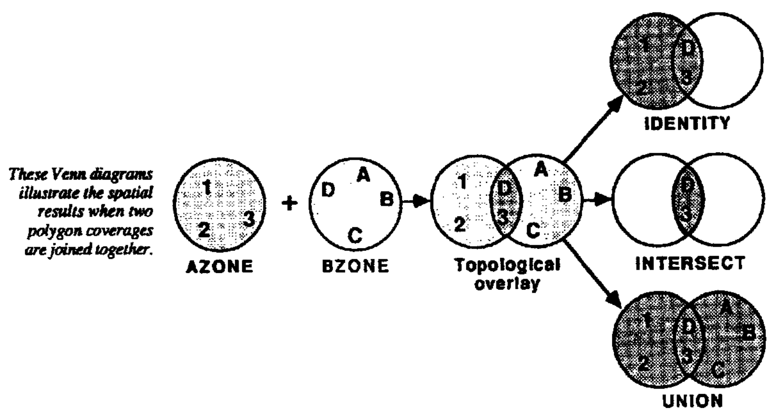


Figure 21. Venn diagrams

How to do a GIS analysis

There are five basic steps to performing a GIS analysis.

1. Frame the question.
Begin by figuring out what information you need, framed in the form of a question. Be as specific as possible

- -Where are the parcels zoned RH?
- -Who are the owners of those properties?
- -What are the full cash values of the land and improvements?

2. Select your data.
The type of data available can determine the methodology used in the analysis. This is the hard part - getting data for what you need to do. Ideally, there is already an adequate database available and unless this is the case, performing an analysis could turn into a painful and frustrating experience. Databases are not easily created - they take time, money and intensive labor to develop. Accuracy is another important issue.

Assuming the data is available to answer the questions we have asked, we would select a parcel base coverage in this case.

3. Choose an analysis method.
Select a method of analysis based on your original questions, taking into consideration how the results of the analysis will be used. We need to decide if we want to have a coverage of only the parcels meeting our selection criteria or only a tabular list of the results of our query.

4. Process the data.
Execute the appropriate commands to answer the questions we have posed. Generate a map of the results if that is the approach selected or capture the data in a file in the correct format for your subsequent application such as a spreadsheet or word processing package.

5. Look at the results.
This is a very important step. NEVER assume the software is smarter than you! The results must be checked to assure you framed the questions appropriately and the information that came back is what you were expecting. For example, if you got back 300,000 parcels out of a possible 330,000, something probably happened and the wrong criteria was used. Check for typos in the query, review your assumptions and logic to see if anything is wrong with the approach. Adjust where needed and rerun the analysis as many times as necessary.

Programming in ARC/INFO

Programming the ARC/INFO software to perform required tasks can be done at the command line - typing in each command to be executed, then waiting for it to finish before the next command is issued. Sometimes this is the preferred method, especially if the results of the analysis are in question and require scrutiny before being introduced into the subsequent analysis.

For more routine tasks like map making, an AML program is usually used to generate the plot file. *AML* stands for *Arc Macro Language*, an interface between the user and the ARC environment. ESRI describes AML:

> *AML provides full programming capabilities and a set of tools to tailor the user interface of your application. AML can be used throughout the ARC system to perform a variety of tasks. AML allows you to automate frequently performed actions, create your own commands, provide startup utilities to help new or inexperienced users perform operations that require specific command settings, and develop menu-driven user interfaces to meet the needs of end users.* *-AML User's Guide, ESRI, 1991*

Pima County Technical Services use AMLs extensively to generate map products. A template has been developed that results in all the maps having the same look even when the subject is different. The AML file is created on a system text editor and is merely a text file of commands, variables and directives typed in a logical sequence. The AML is run in the ARC/INFO environment and the software executes each command in sequence, writing the results to a 1040 graphics file that is drawn on the computer screen after the program is complete. This allows the user to check for errors, text size and placement, color combinations and dozens of other details on the map before it is plotted. If corrections are warranted, the AML is reopened in the text editor, changes made to the code and rerun. The end product is a map.

See Appendix A for an example of a typical AML used to generate a map. Appendix B displays the map generated by this AML.

GIS Infrastructure in Pima County DOT Technical Services

This brief description of the GIS capabilities of Pima County Department of Transportation's Technical Services touches on the basic components necessary to make GIS technology viable.

<u>Network & Hardware</u>

Network:

All Technical Services computers are networked via a 10/100 megabit Ethernet system, allowing data to be easily transferred between workstations and servers.

Hardware:

Technical Services operates on dual platforms: UNIX and NT. Both systems perform a variety of GIS and related functions but utilize different hardware and operating systems .

Servers:

A SUN Sparc 630MP serves as an archiving unit while two SUN Sparc 2's serve the ARC/INFO software to the UNIX workstations in the Division. Technical Services is also connected to the City of Tucson's server, Mars, a SUN Enterprise 5000. NT servers are Pentium 2 processor-based machines, serving software and data to the individual desktop machines throughout the Division.

UNIX:

A variety of SUN workstations are used in the division, mainly for processing GIS data for display and analysis. This platform is where all the heavy lifting is done - processing the parcel base, digital orthophotos, corporate limits, street centerlines, taxmaster, etc.

NT:

Most of the database development and maintenance is performed on NT workstations, which provide an inexpensive desktop solution. Experience has shown this platform will not handle the high level processing accomplished by the UNIX workstations, and required by large GIS data sets.

Software

Technical Services operates primary with two software packages: AutoCAD and ARC/INFO. AutoCAD is used mostly as a database development and maintenance tool while ARC/INFO performs GIS analysis and mapping tasks. The software programs operate on both hardware platforms.

Two auxiliary products from each of these software manufacturers are also put to extensive use in the Division: ArcView, a simplified desktop version of ARC/INFO, and MapGuide, a browser plug-in from AutoDesk used to view spatial data sets over the Internet.

Personnel

Contrary to popular belief, computers do not run themselves. The most important part of the GIS operation is people. Highly skilled in the use of intricate software, they build and maintain the database that is used to find answers to the spatially-related questions users ask. Network Operations, another competent group of people, maintain the network, making sure everything is in proper working condition.

Database

Technical Services GIS database has over 1000 coverages developed or adapted to the Sonoran Desert Conservation Plan, a variety of digital orthophoto quadrangles (DOQ's), other imagery, and a number of related tables such as the Assessor's Taxmaster, Treasurer's records, and Development Services' Permit files. All these data are accessible through ARC/INFO while the tables can be queried through a variety of software programs. There is a full set of maintenance procedures in place to keep the database current. Software routines process data nightly and perform updates to the parcel base, street network and administrative boundaries coverages along with related database tables.

The Sonoran Desert Conservation Plan and GIS

GIS is being used to support the Sonoran Desert Conservation Plan (SDCP) in Pima County. Designed to protect and enhance the natural environment with long-range planning, most of the data is geographic and ideally suited to GIS applications. Starting out with about 200 coverages already in the Technical Services GIS database, an additional 800 coverages were developed, borrowed or converted from a variety of sources and transformed into a common coordinate system. GIS has played an important role in the SDCP process, producing maps by the hundreds - literally several acres of maps, generating statistics and reports, transferring data sets to Adobe Illustrator software in the Graphics Design section where the data had the artist's touch added to produce eye-pleasing illustrations.

The GIS database is at the heart of this effort and requires a great deal of care and feeding, acting like a living creature at times, hungry and ever-changing, restless and demanding. Rather than providing a laundry list of pages of coverage names, the main data sets will be listed by seven SDCP plan elements, catalogued by this structure into a natural data grouping. There is a great deal of overlap - many layers will apply to more than one element - roads, for example.

Sonoran Desert Conservation Plan Data by Plan Element

I. RIPARIAN RESTORATION

- bank protection
- CAP
- climate
- CIP bridge, drainage, detention projects
- dams
- depth to water
- detention basins
- FEMA floodplain
- floodway
- hazardous plumes
- hydromeso areas
- lakes
- perennial streams
- precipitation min/max annual averages
- riparian habitat
- sewer treatment plants
- springs
- Tucson Active Management Area
- washes
- water companies
- watershed delineation
- well sites

II. HABITAT and CORRIDORS

- bedrock
- biological corridors
- exotic species
- faults
- GAP vegetation
- geology
- habitat
- Heritage Data / AZGF TE species
- min & max elevations
- Organ Pipe NM vegetation
- PAG vegetation - 1977
- riparian connections
- Saguaro NP vegetation
- soils
- topography

III. MOUNTAIN PARKS
- bike routes
- CIP parks
- city parks
- county parks
- federal parks
- proposed parks

MOUNTAIN PARKS
- peaks and ridges (restricted)
- open space acquisitions
- trails
- wilderness boundaries

IV. CULTURAL/HISTORIC PRESERVATION
- archaeological sites
- archaeological surveys
- historic sites

V. RANCH CONSERVATION
- grazing allotments
- potential National Conservation Areas
- ranch conservation areas
- stock tanks

VI. LAND USE, LEGAL and FISCAL
- cemeteries
- golf courses
- jurisdictions
- landfills; wildcat dumps
- land ownership (federal, state, county)
- land parcels
- land use (existing)
- mines
- railroads
- school districts; schools
- scenic streets and routes
- streets
- subdivisions
- towns
- zoning

Metadata

One of the demands databases place on their owners is metadata. Simply stated, metadata is data about data. This important feature can be handled in a number of ways - as an INFO file, text file, an html file on a web site. It is often the first thing to be put aside when project demands exceed available resources, and that is the case with the SDCP metadata. Portions of the database in existence before the advent of the SDCP are completely documented and available on-line via the Internet. The plan is to accomplish the same level of documentation for the full set of SDCP coverages, but until then, the fallback position is the old style of metadata maintenance - the origin and other details remain in the minds of project participants.

An example of existing metadata for the parcel base is shown below.

Descriptive name	*Parcels - Regional*
File name	*paregion*
Category	*MM-NonTrans*
Spatial domain	*Pima County*
Abstract	*paregion displays all the existing parcels in Pima Co.*
Feature type	*polygon*
Feature count	*338011*

Scale	*1:3,000*
Control tics	*ADOT*
Rectification	*parcel*
Known errors/qualifications	*paregion reflects current parcel count when coverages were appended.*
Source organization	*PC Assessor*
Source contact	-
Source document or file name	-
Source date	*1997*
Source scale	*1:200*
Source format	*Unknown*
Date of last update	*199909*
On maintenance?	*Yes*
Maintenance organization	*PC DOT Technical Services*
Library write access	*Technical Services*
Maintenance frequency	*Daily*
Maintenance format	*Coverage*
Maintenance description	-
Rectification process	*N/A*
Original conversion by	*PC DOT Technical Services*
Projection	*State Plane Feet*
Path	*az/counties/pima/covers*
Technical Services contact	*Steve Whitney*

For a full review of the existing metadata, see the Technical Services web page at:

http://www.dot.co.pima.az.us/gis/data/

Appendix C provides the field definitions used in this metadata file.

Expanded Functionality of GIS

There are still a number of vector functions that have not been discussed - surface modeling using TIN - triangulated irregular networks, network modeling - routing, resource allocation and a number of other approaches to using vector data to solve real-world problems. These topics are outside the scope of this report.

Departing from the traditional vector data model, GIS functionality expands by opening the door to new modeling possibilities using the raster model and a Decision Support System - the topic of another report. These topics will be briefly introduced and discussed in more detail in later reports.

<u>Raster Data Model</u>
The raster data model views the world by representing features as cells in a matrix organized in rows and columns. Each row has groups of cells with numeric values that represent geographic features. Figure 22 portrays this concept. Spatial relationships are implicit because the data model is a regular grid and storing explicit spatial relationships as in the vector data model is not necessary.

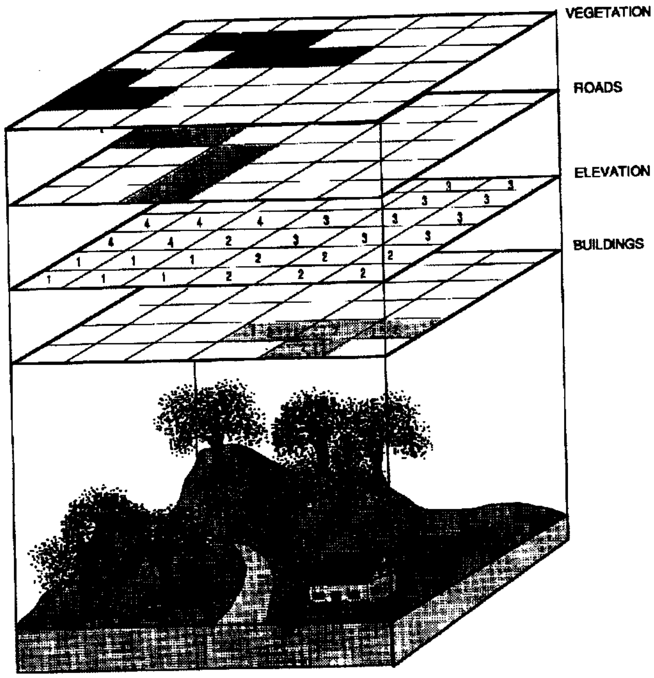


Figure 22.
Raster view of
the real world

Accuracy depends on cell size. The smaller the cell size, the greater the resolution of the data. ARC/INFO provides utilities where vector coverages can be easily converted to a raster format known as a *grid*. The user specifies the cell size and the resulting grid is georeferenced to the same coordinate system as the original coverage. Points are represented by a value in a single cell, shown in figure 23.

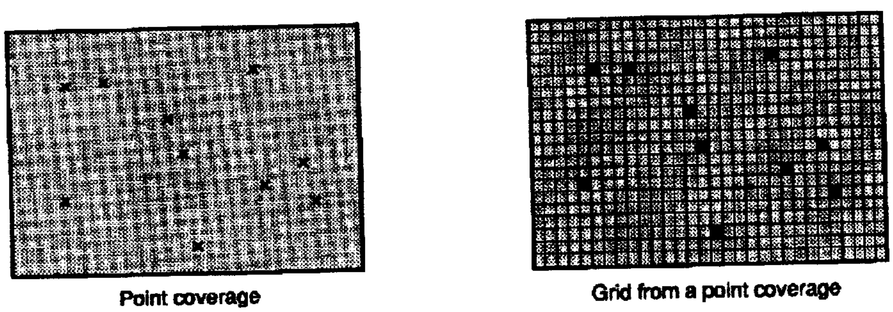
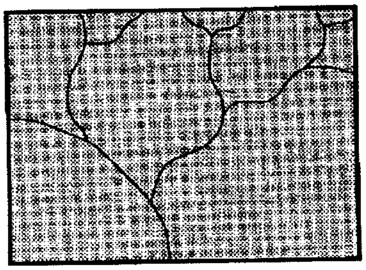
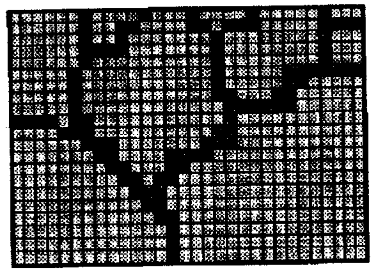


Figure 23.
Point coverage
converted to a grid

Lines are interpreted as a series of connected cells with values as seen in figure 24.



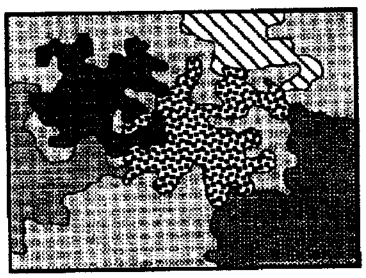
Line coverage



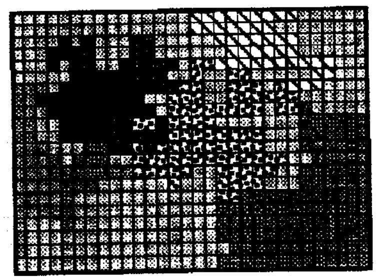
Grid from the line coverage

Figure 24.
Line coverage
converted to a grid

Polygons are displayed as a group of connected cells with values as shown in figure 25.



Polygon coverage



Grid from a polygon coverage

Figure 25.
Polygon coverage
converted to a grid

Individual cells are identified by their position in the grid - row and column number, 3,2 in this case. The origin is usually the upper left and cell identification begins with that point. Figure 26 illustrates this concept.

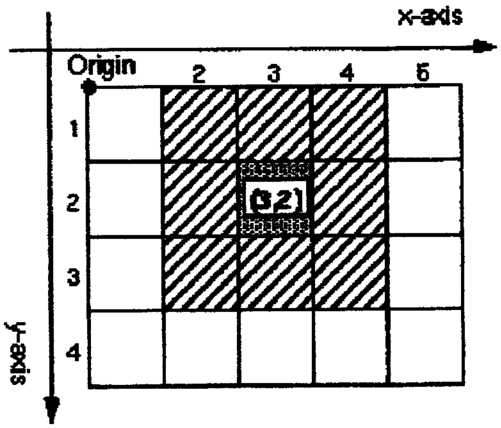


Figure 26.
Origin of
a grid

Decision Support System (DSS)

The Sonoran Desert Conservation Plan is a complex multi-faceted approach to achieving an array of conservation objectives. Different alternatives will be considered throughout the process - a series of answers to "what if ... statements" that need to be visualized and subjected to analysis by a variety of experts in fields concerned with conservation planning.

The DSS will use GIS derived habitat suitability surfaces as the basis for analysis under the assumption that SDCP objectives can be captured with a set of biological, cultural and civic objectives referred to as *conservation targets*.

The fundamental components of the DSS model are:

1. *Conservation targets or biological objectives spatially defined using GIS technology*
2. *Socioeconomic objectives spatially defined using GIS technology*
3. *Biological and socioeconomic constraints established using GIS technology*

Conservation targets are converted from selected vector coverages such as elevation, slope, aspect, vegetation, soils, geology and a number of other factors into raster grids, manipulated and represented by a cumulative score on a cell by cell basis. This is what is being called a *habitat suitability surface.*

These will be compared to the *socioeconomic "cost" surfaces* generated by GIS analysis, depicting the cost of conserving any particular cell of land based on full cash value, current and planned land uses and other factors.

Constraints are site-specific locations included as necessary elements to any solution or conversely, a site that cannot be included in an y solution. Examples of necessary elements may be a single location where a rare endangered species is found while a site that may not be included is slated for another use such as a freeway.

Conclusion

GIS will continue to play a strong role in the Sonoran Desert Conservation Plan as certain key data layers such as soils, geomorphology and vegetation are improved to allow for further development of the Decision Support System. The extensive GIS database will also expand and be updated wherever appropriate information exists. Metadata development will continue and eventually it will be available along with the SDCP data for viewing on a web site.

Remotely sensed data - satellite and aerial photography, will occupy an important niche as well, becoming integrated into the existing database, providing timely information to assist with change detection and a view of current conditions.

GIS will be used in other projects in Pima County in the future as well. The Comprehensive Plan will benefit from GIS technology. A project similar to the SDCP - currently being referred to as "The Sonoran People's Protection Plan" will use GIS to evaluate human ecology issues such as housing, neighborhoods, health, economic development, job training, education and crime.

The next report in this series will further explore the raster model and the development of the Decision Support System.

Appendix A - ARC/INFO plot AML program

```
/******************************************************************
/* AML: soils.aml
/* PURPOSE: Creates E. Pima Soils Map
/* CONTACT: Julia Fonseca
/* PROJECT #: 3PSDCP
/* BY: Felipe Morales
/* DATE: 8/30/99
/* PLATFORM: NT 4.0,  ARC/INFO 7.2.1
/******************************************************************
 &echo &on

/* SET UP TO RUN ON UNIX OR NT
 &select [extract 1 [show &os]]
  &when SunOS
   &do
    &s path1 = /net/mars/mars1/az/counties/pima/covers/
    &s path2 = /net/mars/mars1/az/legend/
    &s path3 = /net/simba/sonora$/soils/covers/
    &s path4 = /net/simba/sonora$/soils/files/
   &end

  &when Windows_NT
   &do
    &s path1 = //mars/mars1/az/counties/pima/covers/
    &s path2 = //mars/mars1/az/legend/
    &s path3 = //simba/sonora/soils/covers/
    &s path4 = //simba/sonora/soils/files/
   &end
  &otherwise
   &type ERROR: Unknown Operating System.
 &end /* End Select

/* SET COVER VARIABLES
 &s adm  = %path1%limjuris
 &s soil  = %path1%soilsgen
 &s soil2 = %path3%surv668_ft
 &s tr    = %path1%tr

/* SET FILE VARIABLES
 &s lin   = %path4%c1.lin
 &s shd   = %path4%soils.shd

/* SET LEGEND VARIABLES
 &s pima  = %path1%pima_all
/* &s bnd   = %path3%c1-bnd
 &s north = %path2%north_02
 &s seal  = %path2%s_pima
 &s cot   = %path2%s_cot
```

```
&s logo  = %path2%l_tecsrv
&s add   = %path2%address.txt
&s disc  = %path2%disclaim.txt

/* <<<<<<<<<<<< BEGIN MAIN PROGRAM >>>>>>>>>>>>

arcplot
display 1040
soils.gra

shadeset cal999.shd
lineset plotter8.lin

/* CREATE NEATLINES
mapunits inches
pagesize 48 36
linesymbol 1
line 0 0 .000001 .0
linesymbol 73
box .04 .375 47 35.2
line 36 .396 36 35.179
line 36.02 31 46.98 31
line 36.02 12.5 46.98 12.5
line 36.02 7.5 46.98 7.5

/* DEFINE MAP PARAMETERS
mapunits feet
mape 508468 133970 983277 597350
mapscale 160000
/* mapscale automatic
/* MAPINFO
/* &pause
mappos ll .4 1

/* <<<<<<<<<< DRAW MAP FEATURES >>>>>>>>>

linesymbol 1

/* DRAW SOILS
polygonshades %soil% psym
polygonshades %soil2% 725
linesymbol 1
arcs %soil%
arcs %soil2%

/* DRAW TR
linesymbol 20
arcs %tr%
annotext %tr% trest

/*<><><><><><><><><><><><><><><><><><><><><><>
```

```
/* DRAW INDEX MAP
 mape %pima%
 mapscale 1500000
 linesymbol 1
 mappos ll 37.5 8
 arcs %pima%
 polygonshades %soil% 7
 polygonshades %soil2% 725
/* linesymbol 2
/* arcs %soil%
 linesymbol 1
 arcs %adm%
 annotext %adm%

/*<><><><><><><><><><><><><><><><><><><><><><><><>

/* PLACE TITLE
 textquality proportional
 textset font
 textsymbol 24
 textsize 1
 move 41.5 34.5
 text 'Eastern' cc
 move 41.5 33.27
 text 'Pima County' cc
 move 41.5 32
 text 'Soils Map' cc

/* PLACE INDEX MAP TEXT
 textsymbol 22
 textsize .4
 move 41.5 12.1
 text 'Pima County Index Map' cc
 textsize .2
 move 41.5 7.8
 text 'Index Map Scale 1:1,500,000' cc

/* PLACE AML PATH AND DATE
 textdelete all
 textset plotter
 linesymbol 1
 textsize .15
 move 36.2 .5
 text '/simba/sonora/soils/amls/soils.aml'
 move 44.3 .5
 text 'Plotted: 8/31/99'

/* PLACE SCALE TEXT
 textsize .16
 move 45.3 5
 text 'Scale 1:37,000' cc
```

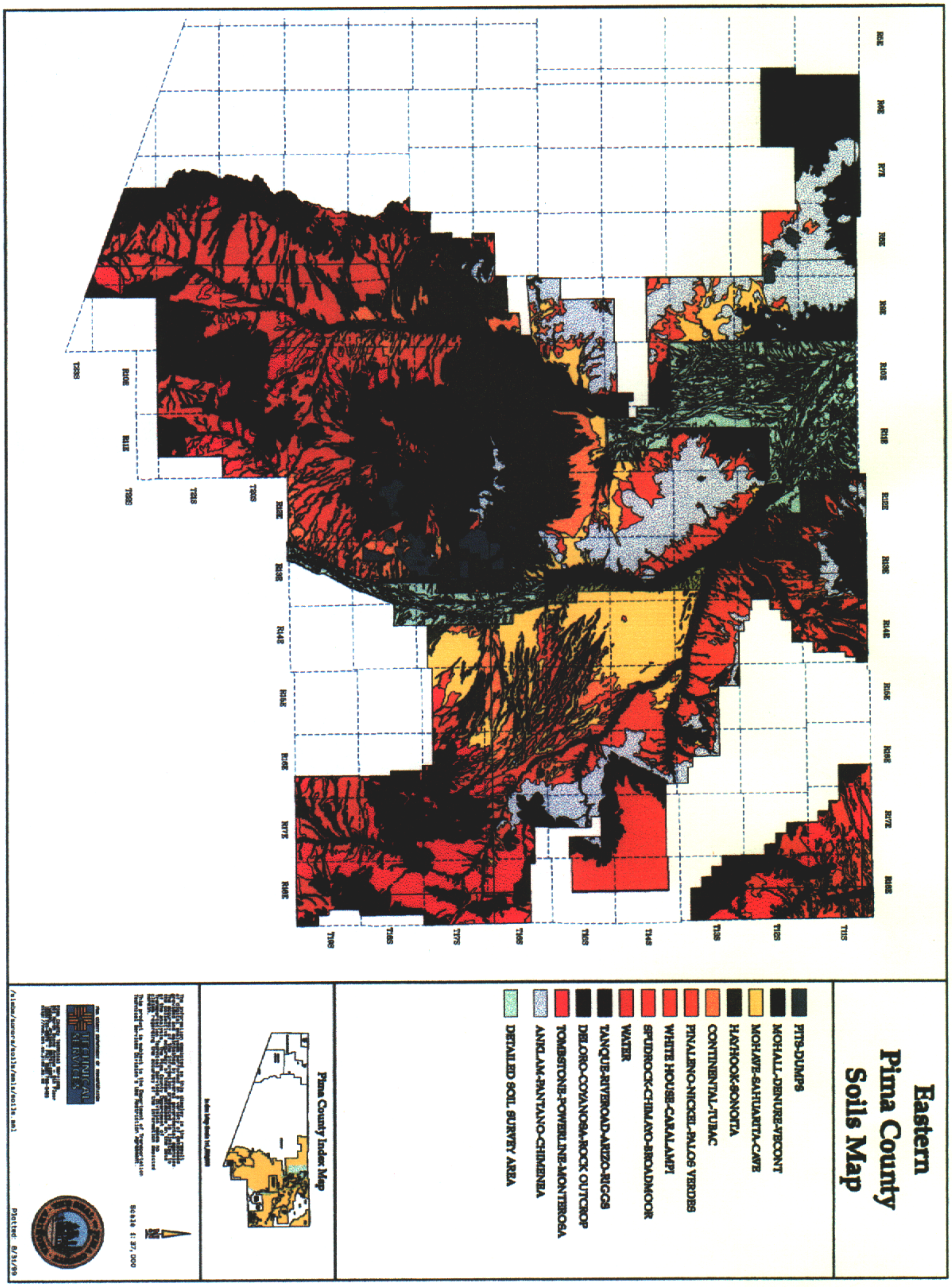
```
/* DEFINE KEYS
linesymbol 1
textquality constant
textset font
textsymbol 22
textsize .4
keybox 1 .5
keyseparation .3 .3
keyposition 36.3 30

/* PLACE KEY BOXES
lineset plotter8
/* keyline %lin%
keyshade %shd%

textquality proportional
textdelete all
textset plotter

/* ADD ADDRESS
textsize .1
move 36.8 2.2
textfile %add%

/* ADD DISCLAIMER
move 36.4 6.8
textsize .11
textfile %disc%

/* DRAW NORTH ARROW
mapunits feet
mape %north%
mapscale 35
mappos ll 45 5.4
polygonshades %north% psym
linesymbol 1
arcs %north%

/* DRAW TECHNICAL SERVICES LOGO
mapscale 15
mappos ll 36.6 2.4
mape %logo%
arcs %logo%
polygonshades %logo% psym999
clipmapextent off
textscale 1.3
annotext %logo%
textscale 1
clipmapextent on
```

```
/* DRAW COUNTY SEAL
 mapscale 30
 mappos ll 43.7 1.1
 mape %seal%
 shadedelete all
 shadeset seal.shd
 polygonshades %seal% psymseal
 arcs %seal%

 shadedelete all
 shadeset cal999.shd

/* <<< END PROGRAM >>>

 quit
 draw soils.gra 9999 3

 &echo &off
```

Appendix C - Metadata field definitions

Metadata Field	Field Description	Acceptable Values
Descriptive name	Full name of dataset	Up to 75 characters
File name	Coverage and shapefile name	Up to 8 characters
Category	Broad data classification	See the six category definitions.
Spatial domain	Area included in the data	Pima County, Eastern Pima County, downtown, etc. - up to 20 characters
Abstract	Short, descriptive paragraph and explanation	Up to 255 characters
Feature type	Point, line, polygon, raster, or tics	
Feature count	Number of features in the dataset	An integer number
Scale	Appropriate viewing scale	24000, etc.
Control tics	Source of control points used in georegistering the dataset quads	ADOT, HARN/HPGN, N/A, or unknown
Rectification	Other dataset used to rectify or align this dataset	COT basemaps, orthophoto, parcel, quads, source map, TIGER, FIRM panels, N/A, or unknown
Known errors or qualifications	Assessment of known errors and problems	Up to 255 characters
Source organization	Name of organization from which data originated.	Up to 35 characters
Source contact	Name of contact in organization from which data originated, possibly including phone number. Optional.	Up to 35 characters
Source document Or file name	Name of document or file name from which data originated. Optional.	Up to 50 characters
Source date	Date of source map or source data	Up to 8 characters in form 19990728 (ymmdd)
Source scale	Scale of source map. Optional	24000, etc.
Source format	Format of source data.	Unknown, Coverage, Shape, DWG, TIFF, or Paper
Second source organization	Name of organization for additional data used in the layer. Optional.	Up to 35 characters
Second source contact	Name of contact in second organization from which data originated, possibly including phone number. Optional	Up to 35 characters
Second source	Name of document or file name from	Up to 50 characters

Appendix B - Map produced from SOILS.AML



Metadata Field	Field Description	Acceptable Values
Document or file name	which data originated. Optional	
Second source date	Date of second source map or source data creation. Optional.	Up to 8 characters in form 19990728 (yyyymmdd)
Second source scale	Scale of second source map.	24000, etc
Second source format	Format of second source data.	Unknown, Coverage, Shape, DWG, TIFF, or Paper
Date of last update	Date that cover was last updated with new or significantly altered features	Up to 8 characters in form 19990728 (yyyymmdd)
On maintenance?	Is the dataset being maintained/ updated?	Yes, No, or Not required
Maintenance organization	Name of organization that maintains the data in the specified maintenance format. It is not the organization that determines content. That would be the source organization	Up to 35 characters
Library write access	Describes the maintenance organization's access to the GIS data library. "External organization direct" means they have direct write access to the GIS library, for example PC Development Services or COT DOT. "External organization via import" means the maintaining organization gives the data to the Technical Services contact to write the data to the GIS library, for example PAG, or COT Planning.	
Maintenance frequency	If on maintenance, then this field describes how often it is updated	Daily, Weekly, Monthly Quarterly, Annually, Varies Unknown, N/A.
Maintenance description	A brief description of who does what and when to maintain the data. Includes how we get the updated info or data, who does the update, and when it is updated - particularly for those layers where the update frequency is "Varies". May also include the formats being converted and the organizations involved, etc.	Optional.

Metadata Field	Field Description	Acceptable Values
Maintenance format	If on maintenance, then this field describes the format that is used to maintain the data.	Unknown, Coverage, Shape, DWG, SDE, Geodatabase, TIFF
Rectification process	Process(es) to be applied after from/to points are determined from orthophotos.	N/A, DEL (Delete), ADJ (Adjust/move features), PRJ (Project to NAD83 only), RS (Rubbersheet), RS-ADJ (Rubbersheet and adjust), RS-EXT (Rubbersheet data sourced in external organization)
Original conversion by	Point of contact name or organization that did the original conversion to electronic form.	Up to 25 characters
Projection	Projection, coordinate system, units	State plane feet
Path	Path to the data. Optional.	Up to 50 characters
Technical Services contact	Name of point of contact person in Technical Services	Up to 25 characters Technical Services contact initials, Point of contact initials Up to 9 characters

References

Arnoff, Arnold. (1989). *Geographic Information Systems: A Management Perspective.* WDL Publications, Ottawa. Canada.

Burroughs, P.A. (1986). *Principles of Geographical Information Systems for Land Resources Assessment.* Clarendon Press, Oxford.

Environmental Systems Research Institute, Inc. (1991). *AML Users Guide*. Redlands, CA.

Environmental Systems Research Institute, Inc. (1991). *ARC Commands References.* Redlands, CA.

Environmental Systems Research Institute, Inc. (1991). *Cell-based Modeling with GRID.* Redlands, CA.

Environmental Systems Research Institute, Inc. (1990). *Introduction to Spatial Analysis.* Redlands, CA.

Environmental Systems Research Institute, Inc. (1990). *Understanding GIS: The ARC/INFO Method.* Redlands, CA.

Poiker, T.K. (1982). *Looking at Computer Cartography.* GeoJournal 6, 241-9.

Snyder, John P. (1987). *Map Projections - A Working Manual.* U.S. Geological Survey Professional Paper 1395. United States Government Printing Office, Washington D.C.

Stine, Peter. (2000). *Conservation Targets.* Personal Communication, email July 27, 2000.

Technical Services, Department of Transportation, Pima County. (2000) *Metadata.* Tucson.

Wilford, John Noble. (1981). *The Mapmakers.* Alfred A. Knopf, New York.