2003

Creating digital geological maps for geographic information systems

Eileen Patricia O'Halloran
San Jose State University

CREATING DIGITAL GEOLOGICAL MAPS FOR GEOGRAPHIC INFORMATION SYSTEMS

A Thesis

Presented to

The Faculty of the Department of Geography

San Jose State University

In Partial Fulfillment

of the Requirements for the Degree

Master of Arts

by

Eileen Patricia O'Halloran

August 2003

APPROVED FOR THE DEPARTMENT OF GEOGRAPHY

Dr. Richard Taketa

Dr. Gary Pereira

Michael J. Rymer, U.S. Geological Survey

APPROVED FOR THE UNIVERSITY

ABSTRACT

CREATING DIGITAL GEOLOGICAL MAPS FOR GEOGRAPHIC INFORMATION SYSTEMS

by Eileen Patricia O'Halloran

This thesis addresses the use of Geographic Information Systems to create accurate geological maps for a portion of the San Andreas fault zone near Parkfield, California. Such maps are needed in field studies currently underway at the San Andreas Fault Observatory at Depth (SAFOD). Geologic data recently collected for the area must be put into digital format for further scientific studies related to earthquakes.

Two software tools, Environmental Systems Research Institute's (ESRI) ArcGIS and Adobe Illustrator, are used to draw the fault linework. Then, the two sets of digital map data are displayed in ArcGIS as separate layers for analysis. Using the measurement tool in ArcGIS, the distances between fault lines drawn in each of the two software packages and fault lines drawn on the aerial photo in the field are determined. Testing and analysis concludes in acceptable software accuracy results. Detailed steps, findings, and comparison of the two methods ensues.

ACKNOWLEDGEMENTS

I would like to thank the helpful and kind people at the USGS in Menlo Park, California. This especially includes Michael J. Rymer, Luke Blair, Ben Sleeter, and Christian Raumann. I would also like to thank my thesis advisors at SJSU, Dr. Richard Taketa and Dr. Gary Pereira. And finally, I would like to thank my family and friends for their encouragement.

TABLE OF CONTENTS

LIST OF FIGURES

Introduction

The San Andreas fault is one of the most famous in the world. Horizontal slip that occurs along it causes a complex zone of faulting, significantly affecting the safety of structures and lives in the surrounding regions. Currently, the U.S. Geological Survey (USGS) and the National Science Foundation (NSF) is conducing a large, detailed, study underway near Parkfield, California, to gather and map geologic data and ultimately understand more about the mechanics and geophysical processes associated with earthquakes. Geologic data have been collected in the field around the study area recently and there is a great need to have these data in digital map format as quickly and efficiently as possible.

Many of the scientists at the USGS have access, training, and experience creating thematic maps in Adobe Illustrator. For them, drawing is much faster and easier in this program. Recently, however, the need for high level data manipulation and analysis functions to be performed require exporting the files from Adobe Illustrator into a GIS tool. Therefore, examining the efficiency and accuracy associated with the usage of each tool can lead to conclusions that may prove to be helpful.

In this study, fault linework for a small portion of the study area is drawn using two different sets of software tools: Adobe Illustrator and Environmental Systems Research Institute's (ESRI) ArcGIS. The two sets of linework are then brought into a Geographic Information System (GIS) and displayed as separate layers for analysis. Measurements are taken between sets of linework to determine the efficiency and accuracy of the digital data from each method as compared with the original field data.

The San Andreas Fault and SAFOD

The San Andreas fault makes up a portion of the boundary between the Pacific plate to the west and North American plate to the east. It extends for about 1300 kilometers through California from the south near the Salton Sea, to the north, near Cape Mendocino. The small and sparsely populated town of Parkfield is located about midway along this path, northeast of the towns of Paso Robles and San Luis Obispo.

Historically, moderate earthquakes (magnitude 6) have occurred along the Parkfield segment of the San Andreas fault every 20 to 30 years or so. However, this segment of the San Andreas Fault has not experienced a major earthquake in over 37 years. Analysis of historic earthquake records shows a small gap in seismic activity, suggesting that this fault segment is nearly ready for another moderate earthquake. This fault segment is also one of the most "consistently" seismically active regions along the San Andreas fault, producing small earthquakes (magnitude 2 or 3) just about on a weekly basis.

Geologic mapping along fault zones not only provides valuable information about the location of faults, rupture patterns, and recurrence intervals, but it also becomes a context for detailed studies. The area around the San Andreas fault near Parkfield is of great interest to scientists because it is one of the few segments along the that fault that has had recurring earthquakes of moderate magnitude since 1857. Therefore, various tests relating to the physics and mechanics of earthquakes in the Parkfield area are being conducted by earthquake geologists, seismologists, and geophysicists. The USGS and

National Science Foundation have teamed up on a project called "The San Andreas Fault Observatory at Depth" (SAFOD). Scientists working on this project are now carefully monitoring creep rates (slow fault slip without producing earthquakes) and comparing those rates to past ground displacement rates along the fault in an effort to learn more about the earthquake cycle. In segments where creep rates are higher than others, the accumulated stress on the rocks is believed to be less than in "locked" segments, or segments that have not experienced much movement. Scientists who study earthquakes believe that locked segments are subject to less frequent but more catastrophic movements.

A deep borehole will be drilled by SAFOD in an effort to study the seismically active fault segment. The hole will be approximately 4 kilometers deep and will bore through the fault so that instruments can be placed close to the hypocenter of recurring shallow earthquakes. Drilling will be performed at such an angle that the hole extends through the entire fault zone until relatively undisturbed rock is reached on the east side of the fault. Rock specimens and fluids will be retrieved from the fault zone so that laboratory analyses and scientific measurements can be made and analyzed. In the long-term, activities will include in-depth seismological observations of small to moderate earthquakes and ongoing measurements of rock deformation and other displacements during the earthquake cycle.

Data collected through scientific experiments conducted for SAFOD will provide valuable information surrounding the composition and mechanical properties of rocks in the fault zone, the nature of stresses responsible for earthquakes, the role of fluids in controlling faulting, earthquake recurrence patterns, and the physics of earthquake initiation and rupture. Through these experiments and data collection activities, SAFOD will make major advances in the pursuit of a rigorous scientific basis for assessing earthquake hazards and predicting earthquakes (USGS, 2002).

The geologic maps and aerial photos that are being used are in most cases decades old and lacking in detailed geologic information. This is due in part to the areas low population density but also to lack of previous interest in this type of information for this particular area. Michael J. Rymer, an earthquake geologist with the U.S. Geological Survey (USGS) in Menlo Park, California has gathered his own geological data in the field along the San Andreas fault zone in the Parkfield vicinity. This data is noted on aerial photos and will be used to create geologic maps for SAFOD.

Fault properties can be determined by examining geologic features on the ground around faults. As an example, evidence of past fault movement can be seen in the form of topographic features such as sag ponds, lines of springs, and fresh fault scarps. Some evidence can be seen by analyzing maps of the fault area over time. Subtle offset in layers of the sides of trenches can be mapped and the time intervals in between changes can help determine the ages of soil layers that have moved over time (Bolt, 2000). These geologic features along with earthquakes of every size, water table changes, and gas emissions are all being monitored and mapped in hopes of understanding the warning signs of forthcoming fault movements (Elam, 2002).

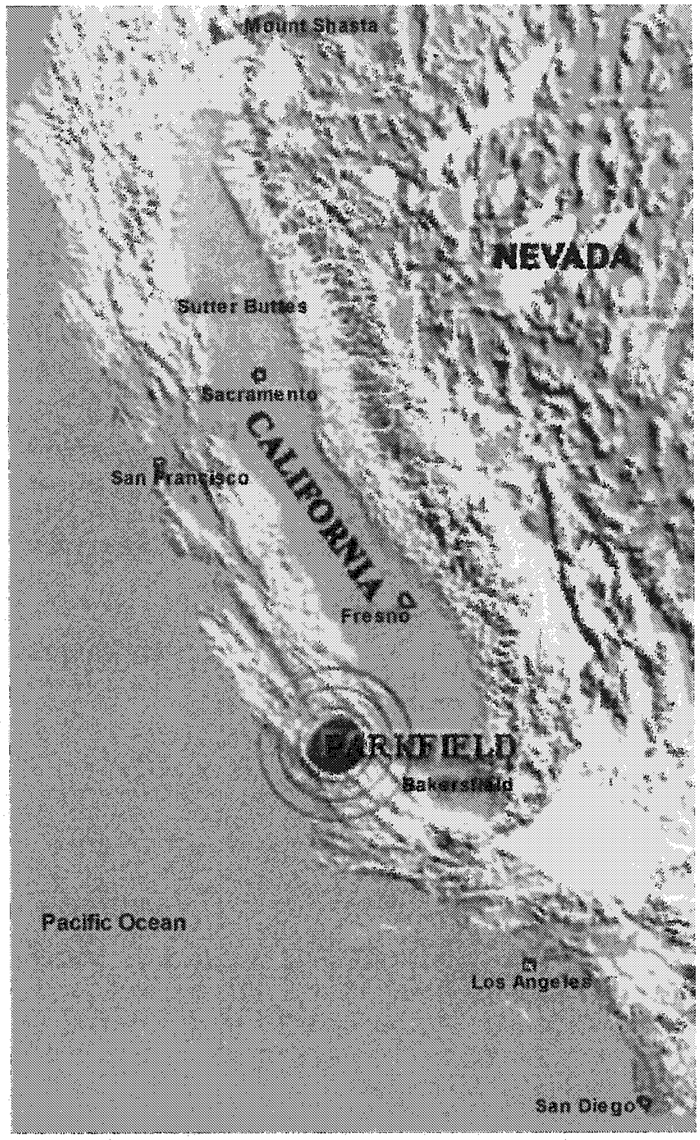

Figure 1. Location map of the town of Parkfield, CA in Central California. (Source: U.S. Geological Survey, 2002 http://quake.usgs.gov/research/parkfield/index.html)

Figure 2. Photograph of Parkfield, CA. The town lies a few hundred meters east of main trace of the San Andreas fault. (Source: U.S. Geological Survey, 2002 http//quake.usgs.gov/research/parkfield/geology.html)

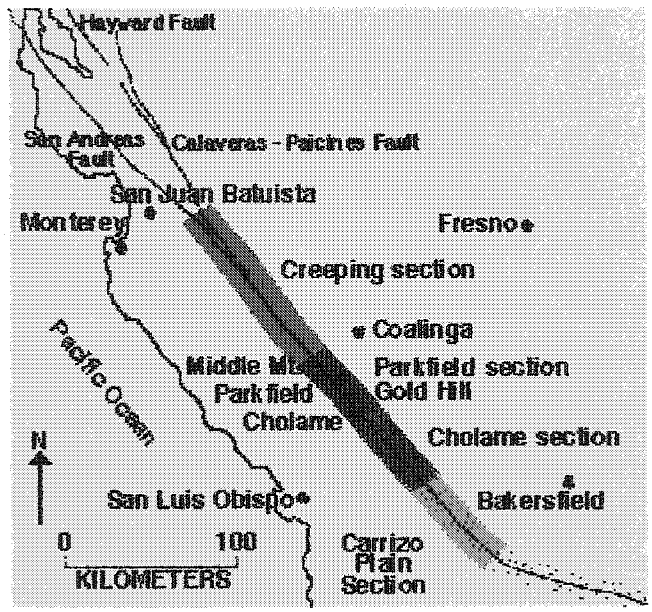

Figure 3. Map of the various sections of the San Andreas fault. The creeping section (green) separates locked stretches. The Parkfield section (red) is a transition zone between the creeping and locked sections of the fault. (Source: U.S. Geological Survey, 2003 http://quake.usgs.gov/research/parkfield/geology.html)

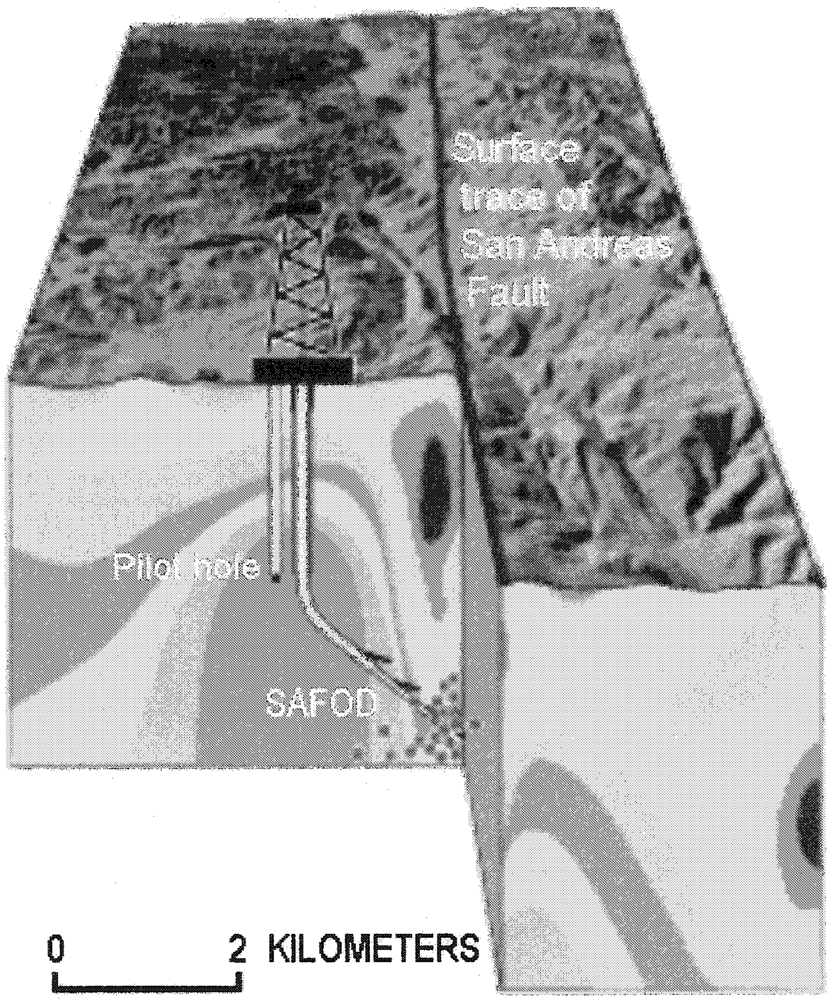

Figure 4. Schematic cross section of the San Andreas Fault Zone at Parkfield. The planned drill hole for the San Andreas Fault Observatory at Depth (SAFOD) and the pilot hole drilled in 2002 are viewed here spanning the various layers of rock. (Source: U.S. Geological Survey, 2002 http://quake.usgs.gov/research/parkfield/safod_pbo.html)

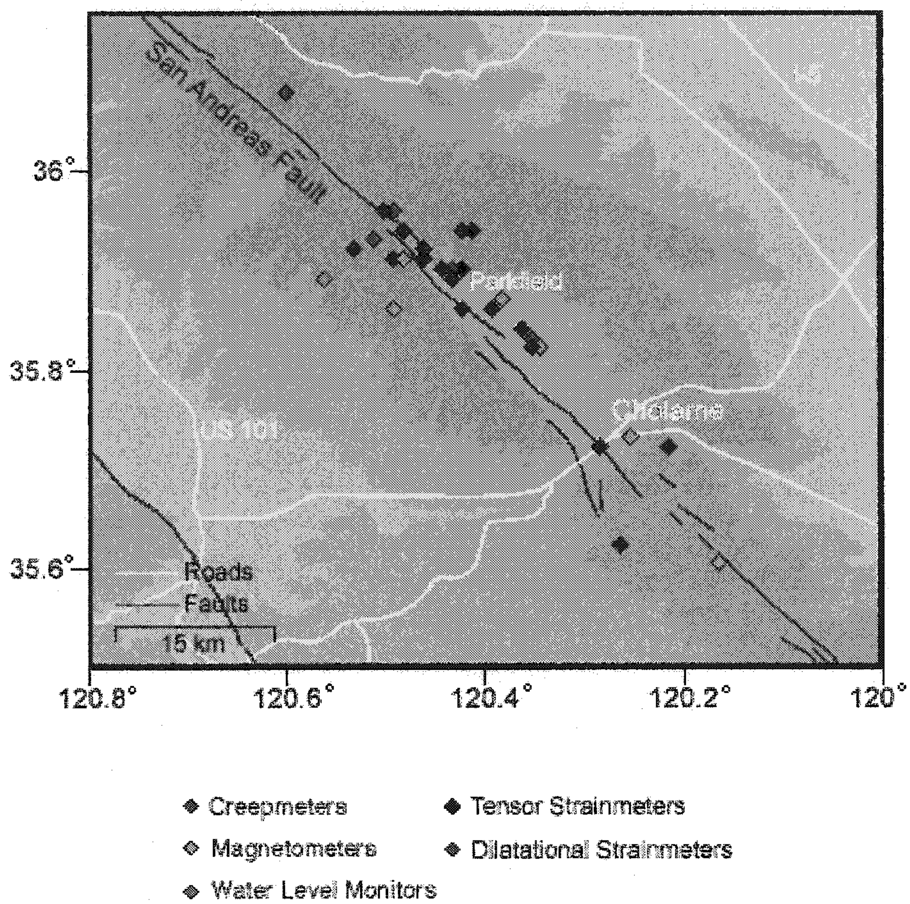

Figure 5. Parkfield area map showing some of the geophysical instrument networks installed in the study area. The town of Parkfield is located near the center of the instrument cluster.
(Source: U.S. Geological Survey, 2002
http//quake.usgs.gov/research/deformation/monitoring/pklocations.html)

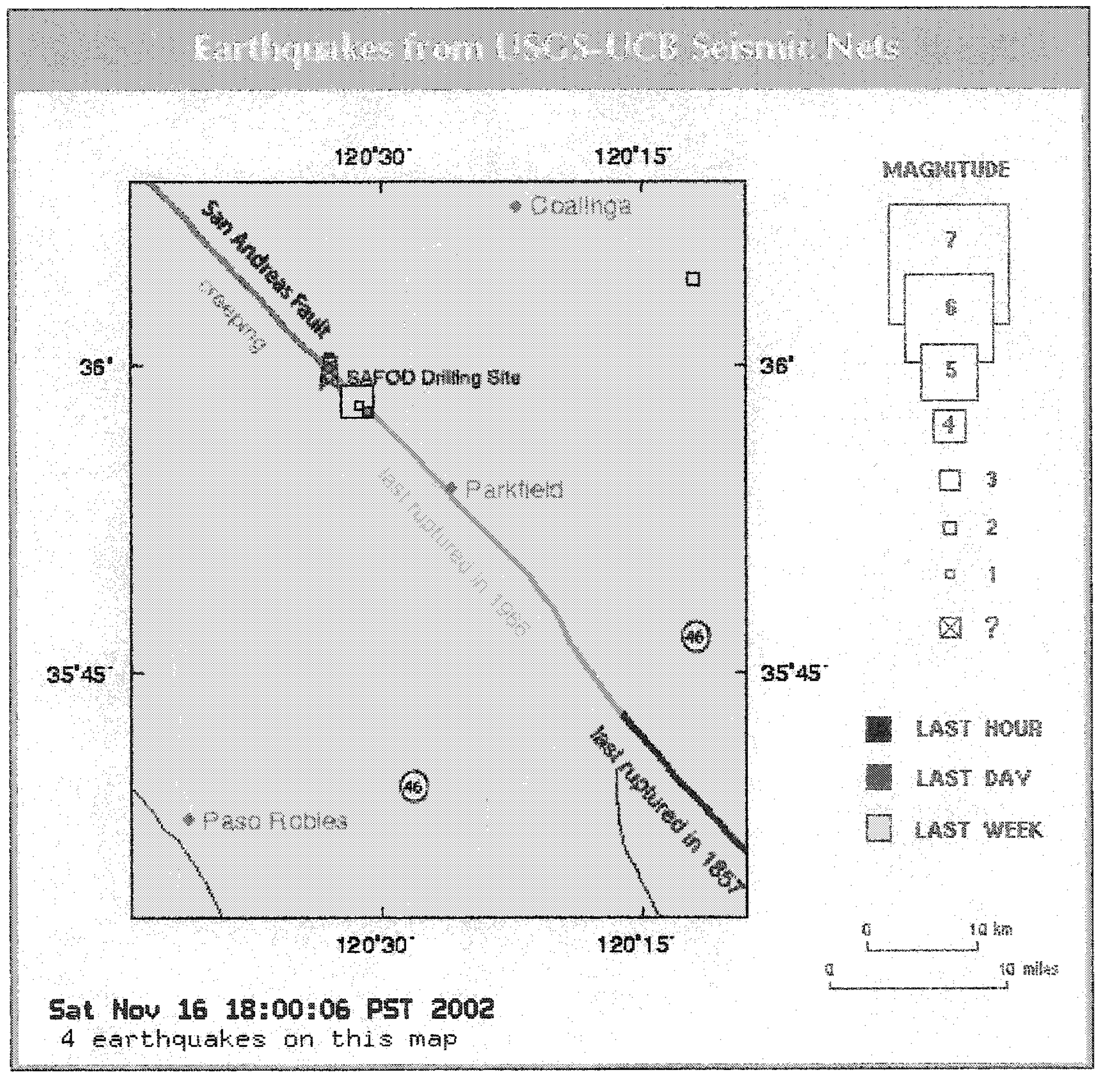

Figure 6. Seismicity map of Parkfield illustrating the location and size of recent seismic events. This map is updated online by the USGS every hour or within five minutes of an earthquake. (Source: U.S. Geological Survey, 2002 http://quake.usgs.gov/recenteqs/Maps/Parkfield.html)

Geologic Mapping Efforts - Past and Present

William Smith was an English surveyor and engineer of canals in the late 18th century. Through his work in the canals, he began to notice patterns in the Earth layers that the canals were cut through. He soon observed that the fossils found in a section of sedimentary rock were always in a certain order from the bottom to the top and the various rock layers contained distinctly different fossils. Eventually, he recognized that rock sections could be identified by the fossils within them, regardless of where the section was found in the region or even throughout the country. Smith then assembled a fossil-based map of England and Wales that showed the composition below the surface. This geologic map and those that followed armed the industrialists and capitalists at the time with good information about where to seek the various mineral and coal deposits, which in part helped fuel the industrial revolution (Winchester, 2001).

In 1908, the Lawson report on the 1906 earthquake included some of the first fault maps of California. This was the first time that people understood the length of the San Andreas Fault. In 1923, a fault map of California was prepared by Bailey Willis and H.O. Wood and published by the Seismological Society of America. It was at a scale of 1:500,000 and was entitled "Fault map of California." In the words of Willis, the map was "designed to show the lines on which earthquakes may occur and which, therefore, should be avoided by structures liable to damage by earthquakes." There was some dispute as to what was considered an "active" fault, but this map was still to be the largest scale fault map of the State at that time, until the California Division of Mines and

Geology published a geologic map that included faults of California in 1938. The coordinator of this mapping project was Olaf P. Jenkins, then Chief Geologist of the Division of Mines. The USGS also collaborated on the preparation. The scale of this map was larger than any previous geologic map of the State up to that time, and therefore careful attention was paid to detail. Geologic boundaries and source data were examined closely to maintain accuracy. This was the first map of California to include faults depicted together with the spatial distribution of rocks. The USGS collaborated on the preparation of this map. A great deal of coordination had to be done to correlate and adjust the map units and legends of the various sources. The map was not published at its original scale but instead was incorporated in the geologic map of the 1932 map of the United States. Soon after, the map was entirely recompiled (Jennings, 1966). This compilation of geologic data from various sources and all cartographic work including photographic reduction, transcribing, and drafting of the maps required an extremely large number of resources and time.

Throughout the decades after these first geologic maps were created, map compilation methods were improved, but were still very manual and time-consuming. GIS usage did exist to some extent on the earlier, large mainframe computers of the 1960's. However, the prevalence of personal computers in the following decades was what really enabled digital mapping programs such as GIS to be developed and enhanced, making them a viable option for widespread mapping needs.

In the most basic of definitions, Geographic Information Systems provide a set of tools that enable management of information according to location. The many benefits to using them and the types of applications are almost limitless. In a GIS, data can be managed and displayed, then sophisticated analysis of data can be performed. Special features of the GIS can aid activities such as comparing the distribution of two types of data within a study area, searching for locations that provide a selected set of characteristics, or searching for the nearest neighbor of a special feature. The information generated from this analysis can then be presented to the user in tables and graphs as a visual aid (Campbell, 1991).

Since their inception, GIS tools have proven to be extremely useful for creating, combining, and displaying various types of map information for analysis by interested parties such as scientists, community developers and planners, structural engineers, and many others. Additionally, they have proven to be great assets to the field of cartography not only by making the process of creating maps more efficient, but also by making updating, printing, and analysis easier and more sophisticated. Once the map is in a GIS, the data can be easily rearranged or manipulated and displayed in various ways to help analyze the study area or problem without needing to be totally re-created, which provides a highly economical benefit. The map and its corresponding data can also be easily printed and taken anywhere, including the field. Some GIS tools even allow for portability through devices such as a Global Position System, or GPS, that uses signals sent by satellites to precisely determine a location (Breslin, 1999).

Decades ago, cartographers at the USGS started to realize that the benefits of creating maps digitally were starting to outweigh the benefits of creating them by hand. Increased quality and decreased production costs were discovered to be two of the main benefits. Data incorporated into digital databases (including coding of attributes and spatial relationships) could later be used for other projects. Additionally, a higher level of map detail could now be created which leads to more precise maps. Maps of a higher complexity could now be produced and cartographers could have more control over such features as colors and appearance (Soller, 1990). Today, the USGS produces geologic quadrangle maps using GIS technology. They are composed of standard geologic maps on topographic bases in 7 ½ -minute or 1 degree quadrangle formats (scales mainly 1:24,000 or 1:100,000) displaying bedrock, surface, or engineering geology in an array of colors. The USGS produces DOQs mostly in black and white in standard quadrangle format and of 1 m resolution (Campbell, 1991).

A GIS tool that was created specifically to meet the challenges of digitizing paper maps is an ARC/INFO (a software package from Environmental Systems Research Institute or ESRI) AML (Arc Macro Language) program called GRIDVECTOR. This AML program reduces or replaces the need to do manual tracing of lines on mylar sheets and scanning in the mylar, as this traditional method of tracing and scanning is less accurate and more time-consuming. With GRIDVECTOR, the scanned image is automatically converted to lines using specific computer commands, and the process of adding new features is simplified (Persits, 1998).

GIS tools will be used for quickly compiling and delivering the detailed geologic information for the SAFOD project at Parkfield. Therefore, studying previous data conversion projects is a useful exercise in that it can inform on past findings and may reduce the number of inefficiencies in future projects.

One large map data conversion project took place in the spring of 1990 as a joint effort by the USGS, the Tennessee Department of Environment and Conservation, the U.S. Army Corp of Engineers, the Tennessee Valley Authority and the U.S. Soil Conservation Service. In this project, 368 geologic maps of 7 ½ – minute quadrangles in Tennessee were converted to GIS coverages. To carry out the conversion, film copies of the scribe coats were made, then scanned, vectorized, and written into a computer program format file and finally loaded into ARC/INFO. Polygons were tagged with symbols identifying geologic units, and coverage lines were tagged with line types showing stratigraphic contacts. Faults were tagged as FAULT or A-Fault designating the location and/or approximate location, respectively. For quality assurance, the completed geologic coverage was plotted and cross-checked with the original geologic map in paper format.

The project encountered the following problems:

a) Not all geologic maps could be converted because only some had a scannable scribe coat;

b) The lake polygons were erroneously tagged as alluvium due to the fact that the polygon outlines were not delineated;

c) Human judgments were made regarding the locations of some of the lines in order to close polygons, which led to inaccurate map data;

d) Fault locations were only grouped into observed or approximate

> categories, not by type of fault;
>
> e) In some cases, “stratigraphic contacts did not exactly overlay the corresponding contacts on the published map. This may be the result of different types of projections used;
>
> f) Geologic maps were constructed over a period of 30 years by different geologists; consequently, stratigraphic delineation may be more detailed on some geologic quadrangles than on adjacent quads and geologic interpretation may differ between adjoining quadrangles.” (Connell, 1995)

When creating new geologic maps, a pre-defined set of standards should be followed in order to be consistent. The USGS follows a set of national map standards which ensures that all products prepared by the USGS National Mapping Division (NMD) under the National Mapping Program (NMP) reflect current mapping and data policies. All maps created following these standards must be accurate and consistent in style and content. Standards are also necessary for efficient sharing of products and to provide information about geospatial data (USGS, 2003).

The United States National Map Accuracy Standards are the benchmark for horizontal and vertical map accuracy and many other rules surrounding map accuracy. These rules include: the methods used to test accuracy, representation of accuracy on published maps, representation when a map is an enlargement of another map, and basic information for map construction as to latitude and longitude boundaries. For horizontal accuracy, the standard states:

> “For maps on publication scales larger than 1:20,000, not more than 10 percent of the points tested shall be in error by more than 1/30 inch, measured on the publication scale; for maps on publication scales of

> 1:20,000 or smaller, 1/50 inch. These limits of accuracy shall apply in all cases to positions of well-defined points only. Well-defined points are those that are easily visible or recoverable on the ground, such as the following: monuments or markers, such as bench marks, property boundary monuments; intersections of roads, railroads, etc.; corners of large buildings or structures (or center points of small buildings); etc. In general what is well defined will be determined by what is plottable on the scale of the map within 1/100 inch. Thus, while the intersection of two road or property lines meeting at right angles would come within a sensible interpretation, identification of the intersection of such lines meeting at an acute angle would obviously not be practicable within 1/100 inch. Similarly, features not identifiable upon the ground within close limits are not to be considered as test points within the limits quoted, even though their positions may be scaled closely upon the map. In this class would come timber lines, soil boundaries, etc." U.S. Bureau of the Budget (Issued June 10, 1941 Revised April 26, 1943 Revised June 17, 1947) (USGS, 2003)

Another set of standardized features of geologic maps are symbols that represent data features. Geologic maps contain a tremendous amount of information, therefore, symbolization is a key feature of each map. The Federal Geographic Data Committee creates the digital cartographic standard for geologic map symbolization.

One example of a symbol type is the representation of "strike-and-dip". This symbol can represent a great amount of information. It illustrates the third dimension of the strata—the direction they extend into the ground. Geologists measure the orientation of strata using a compass. In sedimentary rocks, they look for the the layers of sediment. In metamorphic rocks, the direction of foliation (or layers of minerals) is measured. The strike of bedding or foliation is the direction of the level line across its surface (the direction you would walk without going uphill or downhill). Dip is how steeply (in degrees) the bed or foliation slopes from a horizontal plane. Those two measurements are all that is needed to characterize the orientation of the rock. Each symbol usually represents the average of many measurements. Direction of lineation can be symbolized with an extra arrow. Lineation might be a set of folds, a slickenside, a group of stretched-out mineral grains, or some other similar feature. The number represents the plunge, or the dip angle in that direction (About, Inc. 2002).

The many types of faults, such as thrust faults, are symbolized with a set of solid, dashed or dotted lines and various directional arrows and notations (Figure 7). This project will focus only on the fault linework (symbolization will not be digitized).

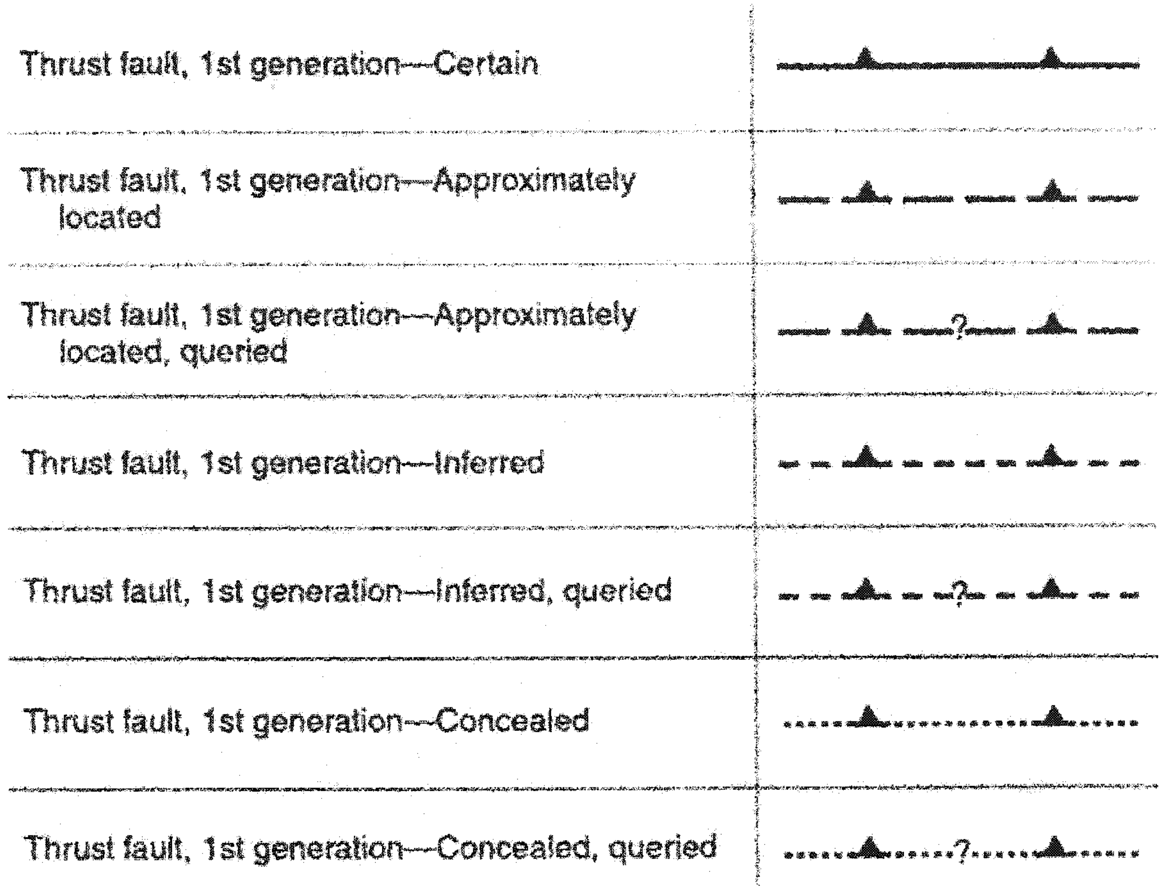

Figure 7. Sample fault symbolization for geologic maps. (Source: Federal Geographic Data Committee, 1999)

Figure 8. Aerial photo of a portion of the SAFOD study area with field etchings and notes. (Source: Michael J. Rymer, U.S. Geological Survey, 2002)

Presentation of the Method

To begin the process of creating maps with GIS for SAFOD, the digital orthophotoquads (DOQs) of the Stockdale Mountain, Smith Mountain, and Parkfield quadrangles were obtained from the USGS. Only portions of the DOQs were needed in order to cover the study area. The process of creating a photomosaic of the DOQ was performed in Adobe PhotoShop, and saved in .img format. The aerial photographs (non-orthorectified images) of the study area were then obtained from the USGS and scanned into digital format using PhotoShop on a Macintosh computer.

The DOQs and the aerial photos were imported to ERDAS Imagine in order to georeference the aerial photographs to the DOQ. The projection on the DOQ was set to UTM Zone10N NAD 1983. Georeferencing was accomplished by selecting an object on the aerial photo (such as a tree or road intersection) and locating and clicking on the same object on the DOQ. More control points were added (a minimum of three) until the match was satisfactory. Once a good match existed, the photo was then rectified and the steps were repeated in succession for adding additional aerial photographs to the background DOQ, thereby creating a photomosaic of the SAFOD study area to be mapped. Finally, the file was exported from ERDAS and imported into ArcGIS (ArcMap).

The first method of digitizing fault data was through the use of ArcGIS. The first task was to create new data "coverages." This involved some set up work including defining the projection as well as adding data and setting coverage type to polygon. Next, the editor was used to create new features. Fault information such as location, length, and geometry from the aerial photograph sketched on in the field was geocoded in detail. Multiple points were added to create the curvature of the lines.

The second method involved creating the fault linework in Adobe Illustrator. The first step was to import the photomasaic of the DOQ and aerial photo with the field etchings on it into Illustrator. Next, the pen tool was used to draw or trace in the precise geometry and location of the fault lines as they appeared on the aerial photo. The pen tool has a feature called "Bezier curves" which enables drawing and modifying lines by pulling out and moving the control handler to adjust sections of the line or curve.

The next step was to export the file from Adobe Illustrator. In order to decrease the file size for exporting efficiency, paths were simplified using built in tools. The map was then saved as an AutoCAD Interchange file (DXF format). Next, the file was imported to ArcGIS. This file was then converted to shapefiles and added as a new ArcInfo coverage. Using ArcToolbox, the coordinate system was set to the same parameters (Transverse Mercator, False Easting, False Northing, Central Meridian and Scale Factor) as the existing layers from Method One. Line type, color and feature names were then assigned to the fault lines.

According to National Mapping Standards, "the accuracy of any map may be tested by comparing the positions of points whose locations or elevations are shown upon it with corresponding positions as determined by surveys of a higher accuracy" (USGS, 2003). Therefore, to begin the process of testing the accuracy of the linework, the layers of fault linework were turned on to display on top of the DOQs in ArcGIS. The Measure tool was used to measure the differences in location by selecting a location on the fault linework and then selecting the corresponding location on the DOQ. Thirty locations (illustrated in Figure 10) were measured for each set of linework to determine if less than 10% were in error by more than 1/50". Since measurement was based upon the scale of the 1:24,000 orthophotoquad, the maximum displacement allowable is 40 ft. All of this was done in accordance with the national mapping standards.

Findings

Out of the 30 locations tested for each of the two sets of fault lines, results concluded in all measurements of fault lines drawn in ArcGIS falling between 0 and 10 ft. in offset distance. Results from testing the linework drawn in Illustrator showed 27 of the 30 measurements falling between 0 and 10 ft. in offset distance, with one very large line displacement measuring 286 ft. Therefore, testing resulted in an overall accuracy of 90%.

To investigate the reason for the extremely large offset drawn in Illustrator, the original linework was re-checked for accuracy back in the original software before it was exported. At first look, shape and geometry were still symmetric, however after further exploration, fault locations in relation to the original source were inaccurate. The displacement only occurred along that one fault line. When the aerial photo layer was turned back on in Illustrator, it became apparent that the linework for that fault was positioned over the original sketchwork, when it should have been positioned on the DOQ. Therefore, it can be deduced that human error was responsible for the displacement of that fault linework.

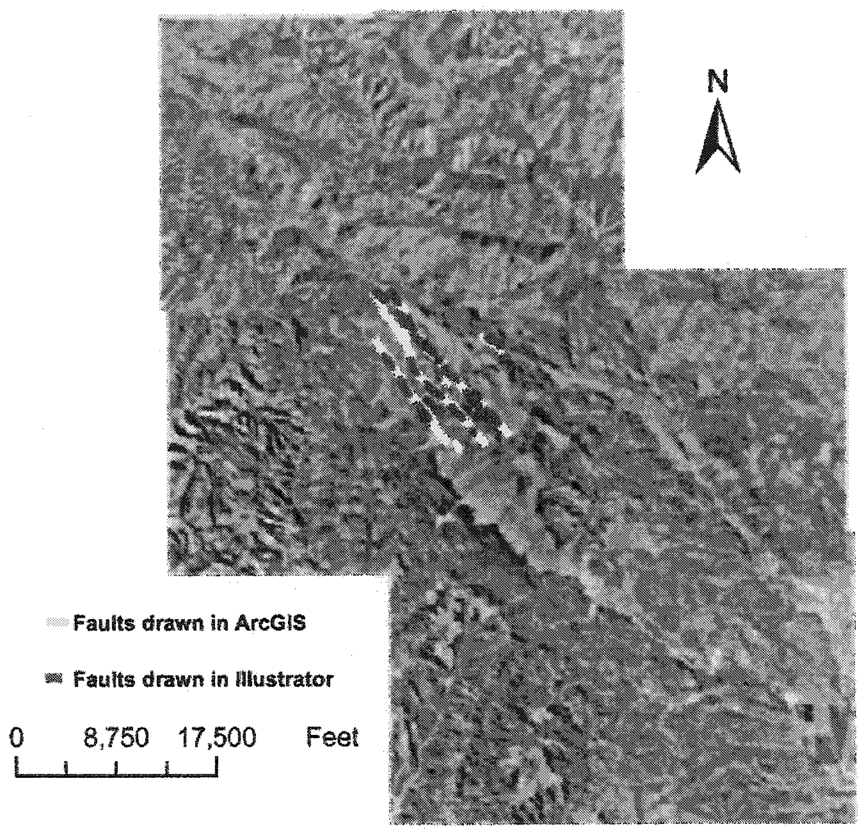

Figure 9. Aerial photo location and corresponding linework displayed on the photomosaiced DOQ.

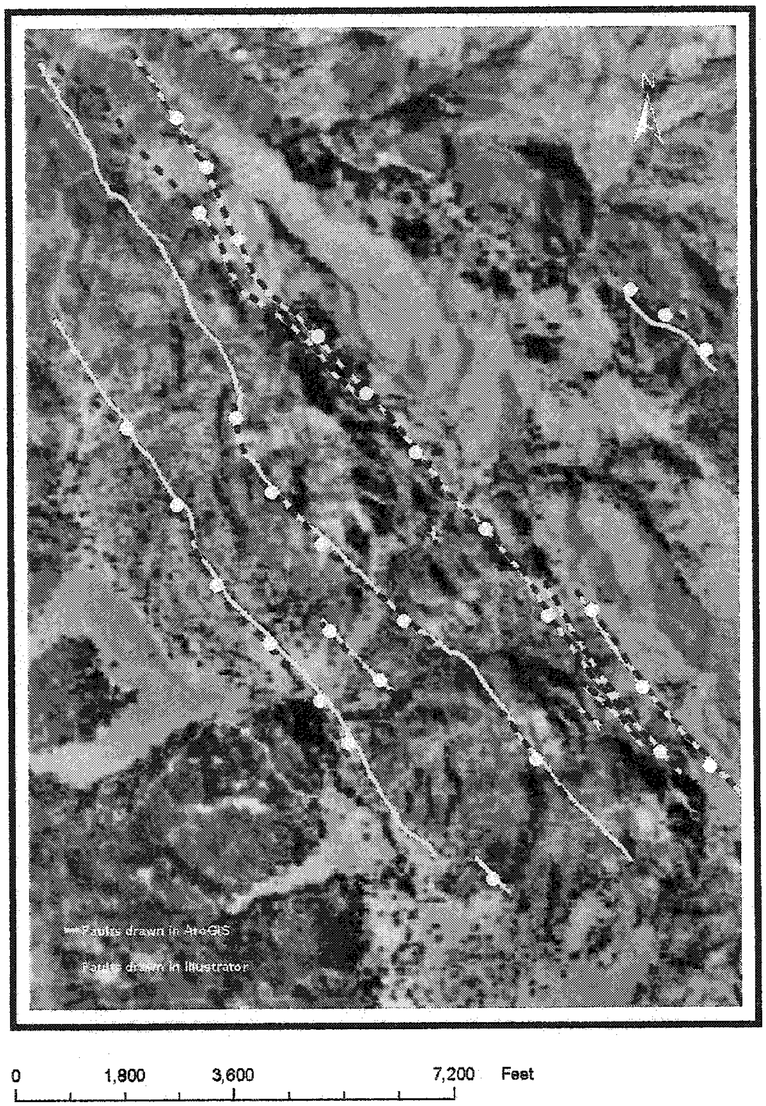

Figure 10. Fault linework drawn using both ArcGIS and Adobe Illustrator displayed as layers over the DOQ. Thirty measurements were taken for each set of linework. Points showing location of measurements appear as white dots.

Figure 11. Image depicting the location discrepancy ranges from 0% to 10%. In most test areas, the faults drawn in Illustrator and ArcMap met National Map Accuracy standards of less than 40 ft. of offset.

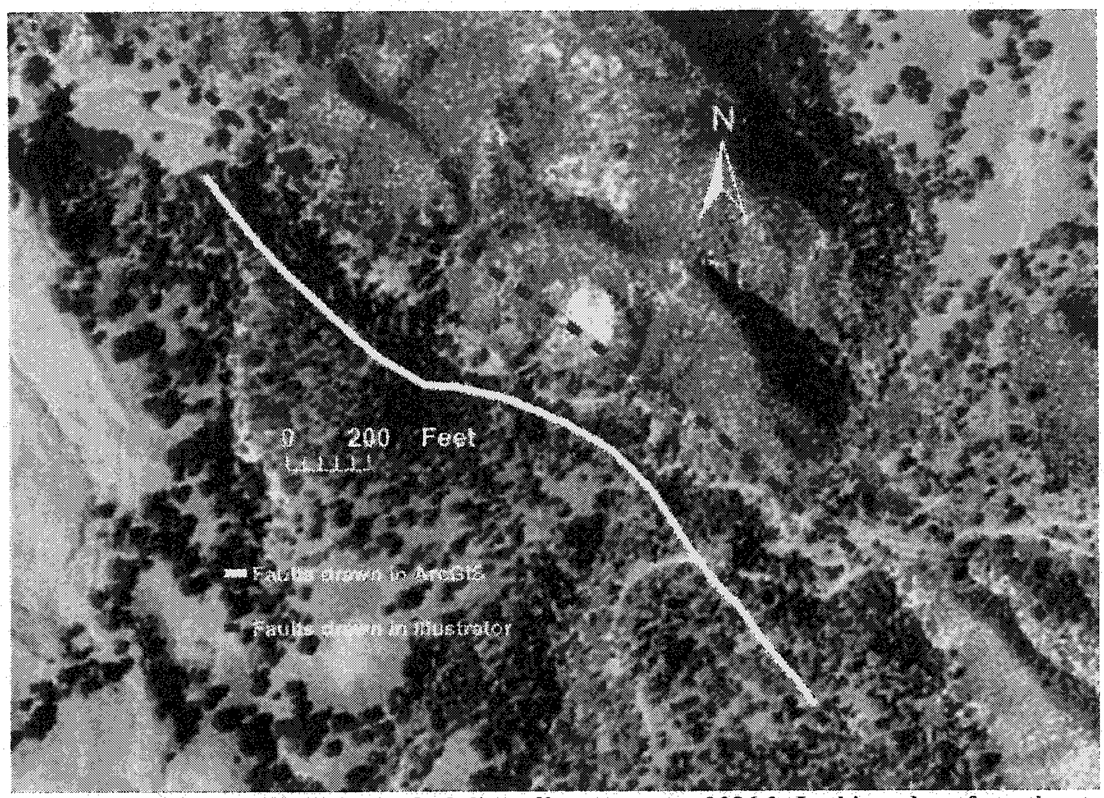

Figure 12. Image depicting maximum fault line offset measure of 286 ft. In this and a a few other test areas, faults drawn in Illustrator did not match up within close proximity to the original sketch work or to the faults drawn in ArcMap.

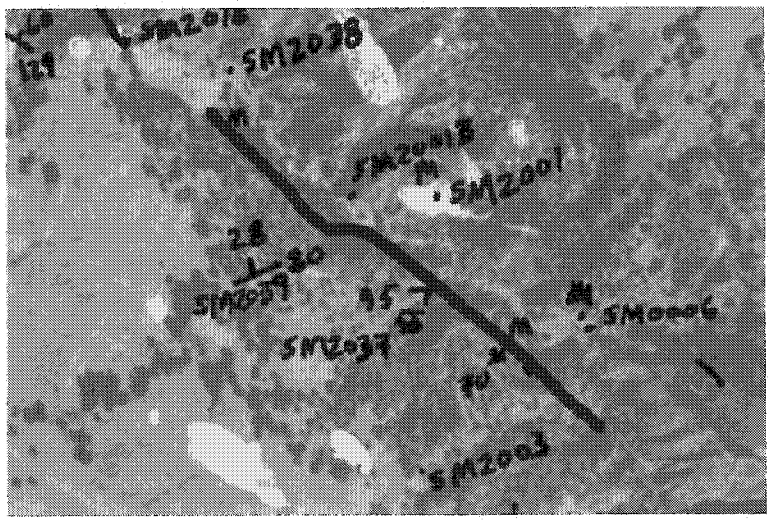

Figure 13. Image depicting the displaced linework drawn in Illustrator. The linework was positioned on top of original fault line sketched on aerial photo in the field.

Discussion

As with any study, human errors can be made quite easily, especially when working with a high volume of extremely detailed data as well as a number of layers. The error found in the linework drawn in Illustrator is a good example of this. When examining the cause of the inaccurate map data, the decision was made that if the original linework was correct, then it would be the case that the data was slightly skewed in the export or import process. However, re-checking the linework in Illustrator resulted in the discovery that these errors in location displacement were not attributable to data lost in the export/import process but to errors in digitizing the linework. As with any projects, human judgments can make for erroneous data. Inferences made in drawing the linework can lead to high offset or location errors and therefore should not be made without checking validity of the data. When striving for a high level of accuracy, a great deal of time and detail must be spent on even the smallest of lines. For example, when drawing the fault lines in this study, straying from the field etched lines even slightly could lead to the fault on the final product appearing several hundred feet away from the actual location on the ground. This could result in a large number of errors in data gathered by the scientists on the SAFOD project who are performing precise measurements.

Another lesson learned was that it is best to draft data directly onto an uncropped DOQ. Having the geologist take the DOQ into the field to gather and note data would help avoid the step of georeferencing. DOQs do not have distortion, are already georeferenced and can be viewed in both Illustrator or a GIS program.

Georeferencing was an important step in this process because it is necessary to create a relationship between the data displayed in GIS software and its actual location on the Earth's surface (ESRI, 2002). In the case of aerial photos, georeferencing must be done to correct the distorted view of the ground that is creating simply due to the camera on the airplane being closer to the objects directly below it, as compared with the ground further out which appear smaller and more distant. However, DOQs by definition are aerial photos that are corrected for lens distortion, air plane pitch, roll, and topographic relief. Without this correction, air photos can not create an accurate basemap, as experienced with the set of distorted lines drawn in Illustrator. As the georeferencing process involves matching a point on the DOQ with the same point on the aerial photo, it proved to be a challenge and often tedious to perform without having a familiarity with the sparsely developed area. In this study, trees and dirt roads were usually the only objects available for matching. More unique appearing features such as buildings or major roads with intersections would have been easier to match up. The USGS often uses pre-installed objects called benchmarks, which allow for easier and more precise location determination. Drafting tic marks or selecting points that have known coordinates would also have been extremely helpful to this end. However, as mentioned, georeferencing an air photo to a DOQ is not going to be completely accurate because distortion will always exist throughout the image. Also, the process of cropping DOQs in a program such as PhotoShop can strip the coverage of its georeferencing and therefore requires the additional steps of performing this task once again.

ERDAS Imagine was used for georeferencing the aerial photo to the DOQ because this process was more user-friendly than ArcGIS. However, a large amount of time was spent in preparing the photo for each new software it was going into for digitizing. The extra time spent on this task outweighed the time-saving benefit of using that tool in the first place. Georeferencing in ArcGIS would have been less time-consuming.

ArcGIS (ArcMap and ArcCatalog) by ESRI is an industry leading software tool that has a sole purpose of GIS functionality. The software was initially chosen as the GIS application for this study, mainly on the basis of three criteria:

1. Ability to import and combine layers of map data
2. User-friendly graphical user interface (GUI)
3. Accessibility to the product at the USGS in Menlo Park

ArcGIS was used as a tool to display both sets of linework because it allows for more sophisticated data manipulation and analysis.

Drawing and editing the fault linework in ArcGIS was a tedious process that involved adding a large number of points in order to achieve precise curvature. To edit the curvature of a line often meant moving all the individual points that made up the curve one by one. Besides being tedious to edit points, the other disadvantage is that by having more points a larger file size is created, and to manipulate the number of points or nodes for reduced file size would lead to corruption of line detail.

Illustrator is a vector-graphics tool mainly used for general artwork and web graphics creation. Therefore, many of the features of this product made it an ideal tool for drawing geologic linework and corresponding labels. Illustrator allows for importing of images and also for multiple layers to be displayed over an image. The layers can be edited and saved separately, making it a useful tool for working with geological data. Additionally, the pen tool allows for sophisticated click-and-drag line creation which helps in drawing long, curvaceous faults. In this study, creating the fault linework in Illustrator was advantageous because of two user-friendly tools within the product: Bezier curves and batch increasing or decreasing of the number of points. Bezier curves are a feature of the pen tool which involves the clicking and dragging of the line points along with a handle which can be adjusted to make the curvature of the line more precise without adding more control points. Then, the number of control points can be easily changed by selecting all of them and then either adding control points or simplifying the linework with built-in commands. The advantage of being able to change the number of control points is that file size can be decreased for ease of exporting and importing.

In both tools, separate data layers are possible, therefore faults can be drawn in one layer, and other linework can be created and stored in other layers so that data manipulation can be completed easily and effectively and analysis can be done by turning on and off the separate layers. However, there are many advantages to putting the final product into ArcGIS, so that sophisticated data storage and analysis could be performed. For example, a database of attributes could be linked with linework or other map features. Illustrator does not have this or other GIS-related functionalities.

Ultimately, the software of choice for creating specialized digital geologic maps depends mainly on the user's ability to quickly and efficiently create the necessary linework and then manipulate and analyze the data. However, the final product should reside in a software tool that is readily available for use by interested parties.

Beyond choosing an appropriate software tool, learning new software programs is always a challenge. While doing research for this study, it became apparent that the challenge of the digital cartographer is not just the need for proficiency and experience with the software programs, but also in keeping up with and following all the necessary cartographic standards. These standards can be complex and extreme care should be taken in complying with them.

Costs for this study were kept to a minimum, but those incurred were mainly paid for by the USGS. Generally, the costs involved were for the aerial photos (some were flown specifically for this project), DOQs, software (ArcGIS, ERDAS Imagine, Adobe Illustrator, scanner), paper, and the geologists and cartographer's time.

Summary

Storing geologic information and creating maps in GIS software can be very useful for scientists in a study area such as Parkfield where features such as fault displacement and sag ponds may be changing often and require frequent map updates. ArcGIS and Adobe Illustrator are both useful tools for drawing geologic linework and displaying the map in layers. Although testing and analysis resulted in an overall map accuracy of 100% for ArcGIS and 90% for Illustrator, the 10% inaccuracy can mainly be attributed to a human error.

Anyone creating digital maps with Adobe Illustrator or any other software program should be cognizant of the need to check the data for accuracy once it has been imported to another software tool. However, working within one software tool could prove to be highly beneficial by decreasing the chance for introducing errors in the data conversion process as well as reducing the number of time-consuming steps.

Appendix

Geocoding Steps

ArcMap (ArcInfo) Georeferencing steps:

Add Data
File open (open the aerial photo .tif format)
File open (open the DOQ .tif format)
Choose: Georeferencing: "fit to display"
Choose: Layer (turn layer .tif off and zoom in)
Add Control Points ("++" icon)
First select image you want to geoference
Choose: save point on other (on corners is best)
Create at least five control points
Choose: Georeferenced – Rectify
Save as .mxd file

ERDAS Imagine 8.5 Georeferencing Steps:

Open DOQ, move window to left side of screen
Right click and choose: Fit Image to Window
Open aerial photo (.tif), move window to right side of screen
Select: Fit image to window
Tile the viewers
Choose: Raster (aerial photo)
Choose: Geometric correction
Choose: Polynomial (1^{st} order)
Projection: map units in meters
UTM GRS 1980 NAD83N
UTM Zone 10
Set projection from GCP tool
Image Layer (brings up GCP tool)
Move ">" symbol down
Resample – output file
File Save – Input As – Save Ref. As

ArcGIS (ArcCatalog) Steps for creating coverages:

New – coverage
Name: Faults
Define (projection):UTM, Meters, 10, X: 0, Y: 0
Add Data: polygon
Editor: Start Editing – create new features
Snap: vertex, edge, end
Editor: share feature editing

ArcMap Steps for Geocoding:

(Once data is added and new map is open)
Editor: Start Editing (using the Edit tool)
(Optional: Editor – Snapping to turn on snapping environment)
Task: Create new features
Target: faults
Draw lines and features using the Sketch tool

Add labels (in separate label layer):

Tools – Editor Toolbar – Start Editing
Choose: Add label
Type names and select text sizes
Save as .mxd file

Adobe Illustrator Geocoding Steps:

New
Place - .tif
Lock
Select/open pentool
Choose color: 100% magenta and red
Choose stroke: .2pt
Draw faults by clicking and dragging over underlying linework
Repeat sequences in separate layers for other geology
Save as .ai file
(Before exporting:
Select all
Object

Path
Add Control Points (Repeat several times)

Select all
Object
Path
Simplify (checkmark: straight lines)

Once in ArcGIS:
open as a new coverage (add data)
convert coverage to .shp file
Edit- Start Editing
Complete Spatial Adjustment with Georeferencing tool

BIBLIOGRAPHY

About, Inc., Symbolizing Geology on Paper: Reading Geologic Maps. http://geology.about.com/library/weekly/aa100800b.htm December 11, 2002.

Bolt, Bruce A. Earthquakes. New York: W.H. Freeman and Company, 2000.

Breslin, Pat., Nick Frunzi, Eileen Napoleon and Tim Ormsby. Getting to Know ArcView GIS. Redlands, CA. ESRI, 1999.

Campbell, John. Introductory Cartography (Second Edition). Dubuque, IA: Wm. C. Brown Publishers, 1991.

Connell, Joseph F., William R. Barron, Jr., and Reavis L. Mitchell, III. Conversion of Geologic Quadrangle Maps to Geologic Coverages. U.S. Geological Survey Open File Report 94-359, 1995.

Elam, Tim and The Buena Vista Museum of Natural History, The San Andreas Fault. http://sharktoothhill.com/fault.html October 10, 2002.

Environmental Systems Research Institute, Inc. Introduction to ArcGIS I. Redlands, CA: ESRI, 2002.

Jennings, Charles W. State Geologic Maps of California – A Brief History. San Francisco: California Division of Mines and Geology, 1966.

Persits, Feliks. An ARC/INFO AML Program to Extract Linear Features from a Gray-Scale Image of a Paper Geologic Map. U.S. Geological Survey Open File Report 97-713, 1998.

Rymer, Michael J., U.S. Geological Survey. Aerial Photo of a Portion of the Stockdale Mountain Topographic Quadrangle Map. October, 2002

Soller, David R., W.R. Stettner, K.J. Lanfear, and D.S. Aitken. A User's Manual for a Method of Map Scaning and Digital Editing for Thematic Map Production and Data-Base Construction. U.S. Geological Survey Circular 1054, 1990.

Federal Geographic Data Committee. Digital Cartographic Standard for Geologic Map Symbolization. U.S. Geological Survey Open File Report 99- 430, 1999.

U.S. Geological Survey. Map of Creeping Sections of the San Andreas Fault in Central California. http://quake.usgs.gov/research/parkfield/geology.html December 7, 2002.

U.S. Geological Survey. Location map of the town of Parkfield, CA in Central California. http://quake.usgs.gov/research/parkfield/index.html December 7, 2002.

U.S. Geological Survey. National Mapping Program Standards. http://mapping.usgs.gov/standards/ March 30, 2003.

U.S. Geological Survey. National Mapping Program Standards – United States National Map Accuracy Standards. http://rockyweb.cr.usgs.gov/nmpstds/nmas647.html March 30, 2003.

U.S. Geological Survey. Parkfield Area Crustal Deformation Measurements. http://quake.usgs.gov/research/deformation/monitoring/pk.html December 1, 2002.

U.S. Geological Survey. Parkfield Area Map Showing Some of the Geophysical Instrument Networks. http://quake.usgs.gov/research/deformation/monitoring/pklocations.html December 7, 2002.

U.S. Geological Survey. Parkfield: Plans for a San Andreas Fault Observatory at Depth. http://quake.usgs.gov/research/parkfield/safod_pbo.html December 1, 2002.

U.S. Geological Survey. Photograph of the Town of Parkfield, CA. http://quake.usgs.gov/research/parkfield/geology.html April 11, 2003.

U.S. Geological Survey. Schematic cross section of the San Andreas Fault Zone at Parkfield. http://quake.usgs.gov/research/parkfield/safod_pbo.html December 7, 2002.

U.S. Geological Survey. Seismicity Map of Parkfield. http://quake.usgs.gov/recenteqs/Maps/Parkfield.html November 16, 2002.

Winchester, Simon. The Map That Changed the World. New York: HarperCollins Publishers, 2001.