

Topic 1

Overview of the Case Study

1.1 Introduction

Geo-Business describes an emerging discipline that uses Geographic Information Systems (GIS) technology to visualize and analyze spatial patterns and relationships within and among mapped data. Targeted marketing, retail trade area analysis, site location and competition analysis are some examples of applications designed to improve decision making and ensure sound investment decisions.

While most of these application-specific procedures evolved independently, they share a common set of basic tools, techniques, approaches and concepts. In the early stage of development (1980s) efforts were narrowly-focused and incrementally evolved as the pieces of application puzzles fell into place. Geo-business innovators comprised a small group of researchers using complex hard/software environments and very little pre-packed digital data.

In just a couple of decades, the situation has dramatically changed with modern computers on every desk having ready access to comprehensive digital data bases—the technological conditions are in place. What remains, however, is an equivalent advance in the human-ware component producing a paradigm shift from non-spatial data analysis to spatial reasoning and problem solving. Understanding the new set of tools and how to apply them is complicated by the fact that most potential users have minimal exposure to quantitative analysis of spatial information.

Recognizing that “*maps are numbers first, pictures later*” is critical to the new mindset. Digital maps are radically different from paper maps as colorful patterns and symbols are replaced by organized sets of numbers that explicitly describe “where” as well as “what.” Similarly, procedures for analyzing digital mapped data often are significantly different from manual map processing and traditional statistical analysis. This book is designed to help readers appreciate the similarities and differences between the approaches, understand basic analytical procedures and identify important considerations in their use.

1.2 Mapping versus Analysis

Vector-based desktop mapping is rapidly becoming part of the modern business environment. The close link between these systems and traditional spreadsheet and database management programs has fueled the adoption. In many ways, “*a database is just picture waiting to happen*.” The direct link between attributes described in a database record and their spatial characterization is conceptually easy. Geo-query by clicking on a map to pop-up the attribute record or searching a database then plotting the selected records is an extremely useful extension of contemporary procedures. Increasing data availability and Internet access coupled with decreasing desktop mapping system costs and complexity make adoption of *spatial database technology* a practical reality.

Maps in their traditional form of point, lines and polygons identifying discrete spatial objects align with manual mapping concepts and experiences learned as early as girl and boy scouts. Grid-based maps, on the other hand, represent a different paradigm of geographic space. Whereas traditional vector maps emphasize “*precise placement* of physical features,” grid maps seek to “*statistically characterize* continuous space in both real and cognitive terms.” The tools for mapping of database attributes are extended to analysis of spatial relationships. This book focuses some of the basic concepts, considerations and procedures in grid-based data handling and analysis operations as they apply to geo-business applications. Three broad capabilities are discussed—1) *surface modeling*, 2) *spatial data mining* and 3) *map analysis*.

Surface modeling involves the translation of discrete point data into a continuous surface that represents the geographic distribution of that data. Traditional non-spatial statistics involves an analogous process when a numerical distribution (e.g., standard normal curve) is used to generalize the central tendency of a data set. The

derived mean (average) and standard deviation reflects the typical response and provides a measure of how typical it is. This characterization seeks to explain data variation in terms of the numerical distribution of measurements without any reference to the data's spatial distribution.

In fact, an underlying assumption in most traditional statistical analyses is that the data is randomly distributed in space. If the data exhibits spatial autocorrelation, many of the analysis techniques are less valid. Spatial statistics, on the other hand, utilizes geographic patterns in the data to further explain the variance. There are numerous techniques for characterizing the spatial distribution inherent in a data set but they can be characterized by three basic approaches:

- *Point Density* mapping that aggregates the number of points within a specified distance (number per acre),
- *Spatial Interpolation* that weight-averages measurements within a localized area (e.g., kriging), and
- *Map Generalization* that fits a functional form to the entire data set (e.g., polynomial surface fitting).

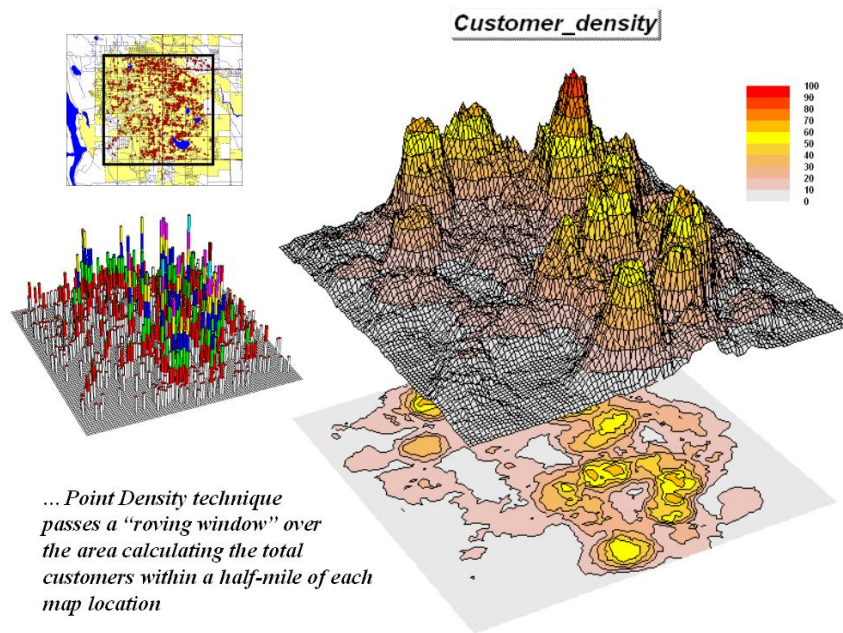


Figure 1-1. Point density map aggregating the number of customers within a specified distance.

For example, consider figure 1-1 showing a point density map derived from customer addresses. The project area is divided into an analysis frame of 250-foot grid cells ($100c \times 100r = 10,000$ cells). The number of customers for each grid space is determined by street addresses in a desktop mapping system ("spikes" in the 3D map on the left).

A neighborhood summary operation is used to pass a "roving window" over the project area calculating the total customers within a half-mile of each map location. The result is a continuous map surface indicating the relative density of customers—peaks where there is a lot of nearby customers and valleys where there aren't many. In essence, the map surface quantifies what your eye sees in the spiked map—some areas with lots of customers and others with very few.

Spatial data mining seeks to uncover relationships within and among mapped data. Some of the techniques include coincidence summary, proximal alignment, statistical tests, percent difference, surface configuration, level-slicing, map similarity, and clustering that are used in comparing maps and assessing similarities in data patterns.

Another group of spatial data mining techniques focuses on developing predictive models. For example, an early use of predictive modeling was in extending a test market project for a phone company (figure 1-2). Customers' address were used to "geo-code" map coordinates for sales of a new product enabling distinctly different rings to be assigned to a single phone—one for the kids and one for the parents. Like pushpins on a map, the pattern of sales

throughout test market emerged with some areas doing very well, while in other areas sales were few and far between.

The demographic data for the city was analyzed to calculate a prediction equation between product sales and census block data. The prediction equation derived from the test market sales was applied to another city by evaluating existing demographics to “solve the equation” for a predicted sales map. In turn the predicted map was combined with a wire-exchange map to identify switching facilities that required upgrading before release of the product in the new city.

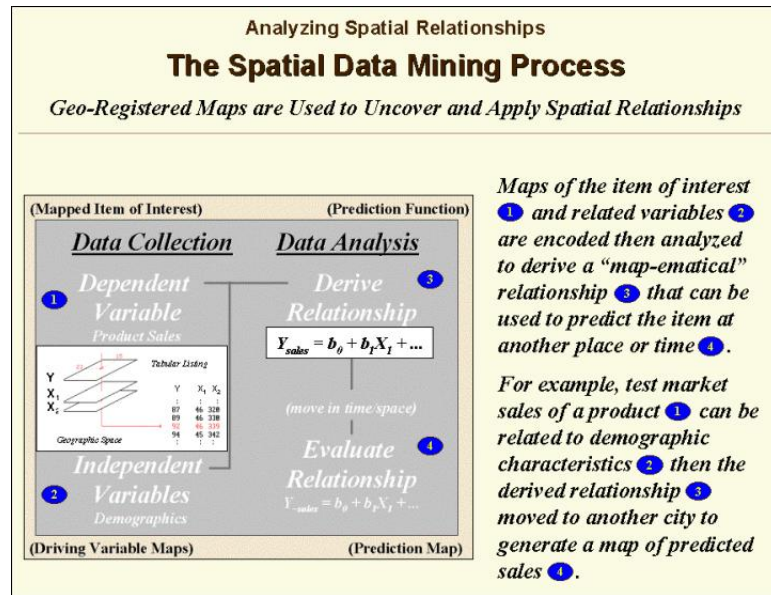


Figure 1-2. Spatial data mining can be used to derive predictive models of the relationships among mapped data.

Whereas spatial data mining responds to “numerical” relationships in mapped data, **map analysis** investigates the “contextual” relationships. Tools such as slope/aspect, buffers, effective proximity, optimal path, visual exposure and shape analysis, fall into this class of spatial operators. Rather than statistical analysis of mapped data, these techniques examine geographic patterns, vicinity characteristics and connectivity among features.

The example shown in figures 1-3 and 1-4 builds on two specific map analysis capabilities—*effective proximity* and *accumulation surface analysis*. The following discussion focuses on the application of these tools to competition analysis between two stores.

The top-left side of figure 1-3 shows the travel-time surface from Kent’s Emporium. It is calculated by starting at the store then moving out along the road network like waves propagating through a canal system. As the wave front moves, it adds the time to cross each successive road segment to the accumulated time up to that point.

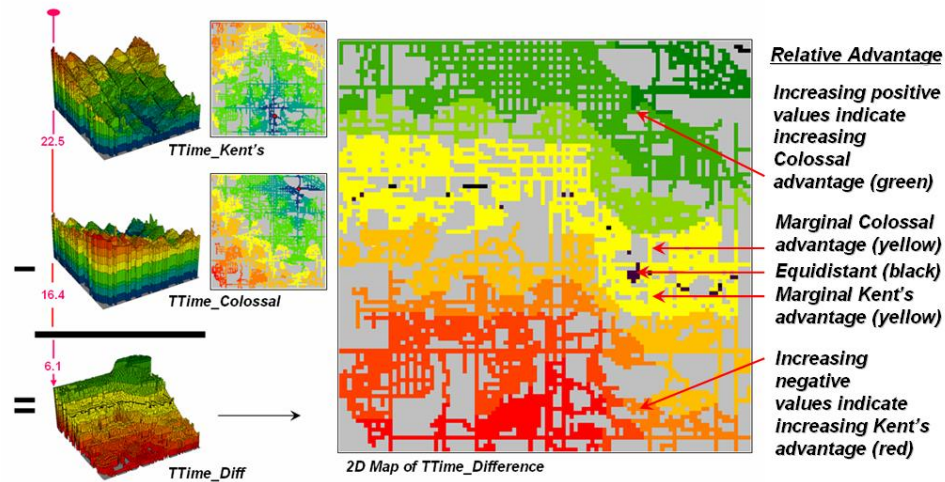


Figure 1-3. Two travel-time surfaces can be combined to identify the relative advantage of each store.

The result is the estimated travel-time from Kent's to every location in the city. The surface starts at 0.0 and extends to 24.4 minutes away (red peaks). Note that it is shaped like a bowl with the bottom at the store's location (south central portion of the city). In the 2D display, travel-time appears as a series of rings that form increasing distance zones. The critical points to conceptualize are 1) that the surface is analogous to a football stadium (continually increasing) and 2) that every road location is assigned a distance value (minutes away).

The middle-left side of figure 1-3 shows the travel-time surface for another store (Colossal Mart) with its origin in the northeast portion of the city. While Kent's travel-time surface appears to "grow" away from you, Colossal's surface seems to grow toward you.

Simply subtracting the two surfaces derives the relative travel-time advantage for the stores (right-side figure 1-3). Keep in mind that the surfaces actually contain geo-registered values and a new value (difference) is computed for each map location. The inset on the left side of the figure shows a computed Colossal Mart advantage of 6.1 minutes ($22.5 - 16.4 = 6.1$) for the indicated location in the extreme northeast corner of the city.

Locations that are the same travel distance from both stores result in zero difference and are displayed as black. The green tones on the difference map identify positive values where Kent's travel-time is larger than its competitor's—advantage to Colossal Mart. Negative values (red tones) indicate the opposite—advantage to Kent's Emporium. The yellow tone indicates the "combat zone" where potential customers are about the same distance from either store—advantage to no one.

Figure 1-4 displays the same information as a 3D surface. The combat zone is shown as a yellow valley dividing the city into two marketing regions—peaks of strong travel-time advantage. Targeted marketing efforts, such as leaflets, advertising inserts and telemarketing might best be focused on the combat zone.

At a minimum the travel-time advantage map enables retailers to visualize the lay of the competitive landscape. However the information is in quantitative form and can be readily integrated with other customer data. Knowing the relative travel-time advantage (or disadvantage) of every street address in a city can be a valuable piece of the marketing puzzle. Like age, gender, education, and income, relative travel-time advantage is part of the soup that determines where one shops.

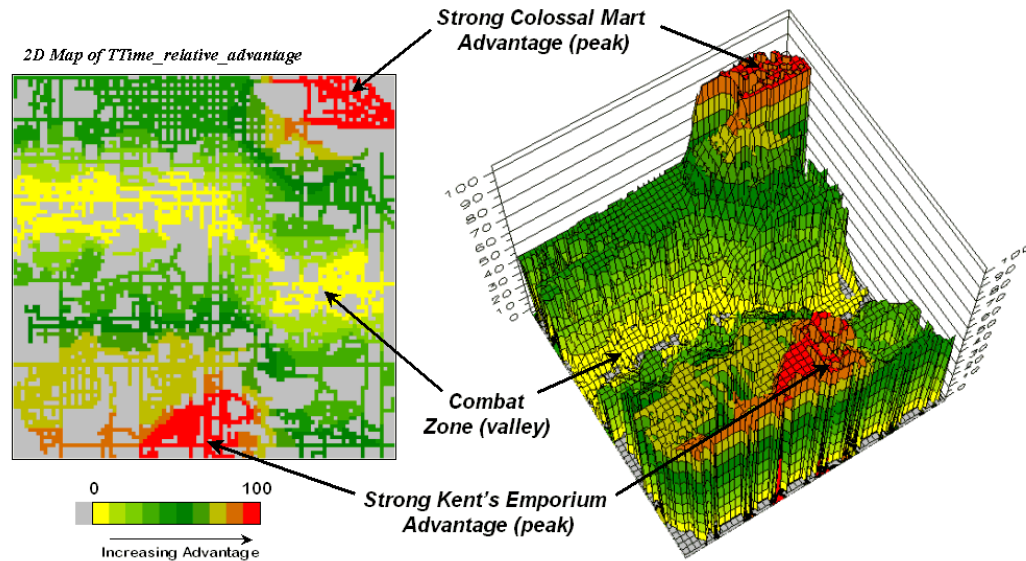


Figure 1-4. A transformed display of the difference map shows travel-time advantage as peaks (red) and locations with minimal advantage as an intervening valley (yellow).

There are numerous other map analysis operations in the grid-based “toolbox.” The examples of customer density surface, sales prediction map and travel-time/competition analysis were used to illustrate a few of geo-business applications capitalizing on the new tools—focus of the remaining topics in this book.

1.3 The Case Study

The case study in this book begins with a series of techniques used to visualize and summarize mapped data and ends with several techniques that develop maps that can be used to guide business decisions. The focus is on methodology (the geo-business “toolkit”) and not on the scientific results or generalizations about the complex economic, demographic and psychological factors driving business conditions in a particular area. These intricate relationships are left to readers to scrutinize, interpret and put into context with their business environment.

The objective of the series of exercises is to demonstrate the analysis procedures on a consistent data set. The intent is to build a case study for a single project area that illustrates the wealth of emerging analytical capabilities that are available.

One might argue that “the technical cart is in front of the scientific horse.” In many instances that appears true. We can collect data and view the numbers as maps or summary statistics, but struggle to make sense of it all. In part, the disconnection isn’t technology versus science but the predicament of “the chicken or the egg” coming first.

We have a radically different set of analysis tools than those available just a decade ago. Concurrently, we have a growing amount of geo-registered data, such as demographics, housing prices, zoning and product sales. The computer industry is turning out software and machines that are more powerful, yet costing less and less and easier to use. Agriculture equipment companies are developing instruments and implements that respond to “on the go” movements throughout a field. In short, the pieces are coming together.

The next decade in precision agriculture will focus on making sense of the site-specific relationships we measure. The technology is here (an analytical toolbox) to bridge the gap between maps and decisions. Its widespread application will be determined by how effectively map analysis is applied.

Table 1-1 lists the topics that will be covered in the case study series. Many of the topics are discussed in the “Beyond Mapping” columns appearing in

Table 1-1. Workbook Topics
Overview of the Case Study
Desktop Mapping
Grid-Based Mapping
Spatial Analysis
Analyzing Location
Spatial Interpolation
Spatial Data Mining
Predictive Modeling

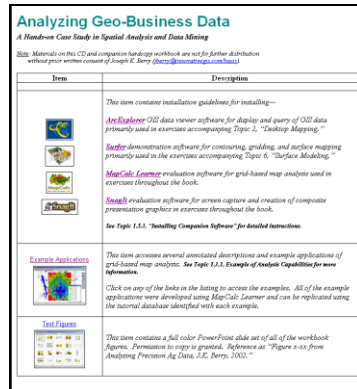
GeoWorld (Adams Business Media, 1996-present) and are compiled online at—
<http://www.innovativegis.com/basis/MapAnalysis/Default.html>.

The book is organized into eight topics and two appendices that lead the reader from an understanding of the fundamental nature of mapped data through a series of basic procedures used in deriving, analyzing and applying spatial information. A case study approach is used with each topic area describing the application of a set of related analysis techniques. The discussion is followed by a series of hands-on exercises providing practical experience in applying the techniques. The exercises include step-by-step instructions that are thoroughly annotated.

A consistent geo-business data set is used to demonstrate several of the emerging data analysis capabilities. Hopefully, the examples will stimulate innovative thinking and new techniques as much as it outlines how existing map analysis procedures can be applied. All of the processing is done within a standard personal computer environment using software provided on the companion CD.

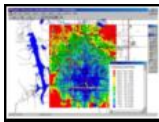
1.4 Exercises

The case study companion CD contains the MapCalc Learner software used in the hands-on exercises. The following exercises provide instructions for installing the software and introduce some of software's basic capabilities. To begin, insert the companion CD into your computer and the main menu will automatically appear.



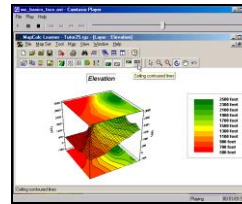
If this main menu doesn't appear select "Start" → "Run" → "Startup.bat" from the Windows Task Bar.

1.4.1 Example of Analysis Capabilities



From the main menu select **Example Applications** → **Short Video Demos** → **MapCalc Basics** → and choose **Open** to launch a short video on basic use of the software including interactive 2D/3D plotting, display

types, display navigation, accessing maps for display and graphical overlay.

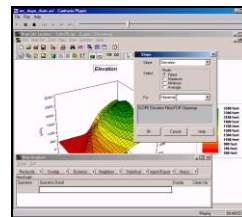


instructional video

"MapCalc Basics"

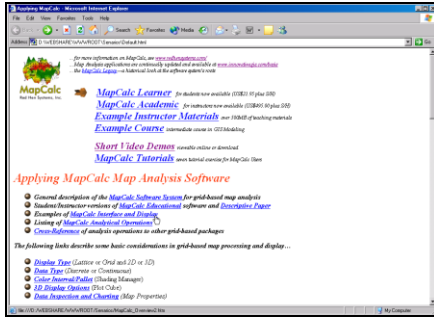
You can replay the video by clicking the "right arrow" button (*Play*) in the upper left portion of the window. Close the video by clicking the "X" button in the upper right portion of the window.

From the "Short Video Demos" menu, select **Calculating Slope and Surface Flow** → **Open** to launch another short video demonstrating use of some basic map analysis operations. The application describes calculating slope and flow maps using an elevation surface.



"Calculating Slope and Surface Flow" instructional video

From the Video Demo menu, press the **Back** button in the top-left portion of your browser window to return to the *Applying MapCalc* menu (see below).

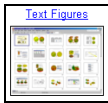


Main Menu for annotated example applications

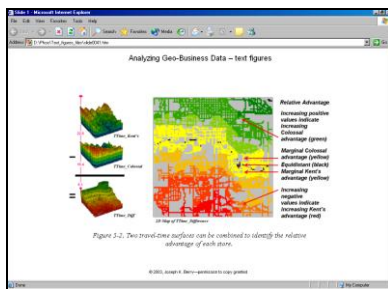
There are numerous annotated examples of MapCalc operation and applications. For example the *Travel-Time and Customer Access* example creates a travel-time map from a store then transfers this information to a standard desktop mapping system.

As time and interest permits, click on other links of interest to activate additional examples. Examples using the *Smallville.rgs* database demonstrate geo-business applications.

1.4.2 Viewing Full Color Figures



Select the **Text Figures** link to access a full color PowerPoint slide set of all of the workbook figures. Choose **Open** to view the figures provided PowerPoint is installed on your system. Choose **Save** to save the slide set to your hard drive.



Color display of maps is critical to visual analysis. You are encouraged to refer to the corresponding full color slides as you read the various topics in this book.

Permission to copy and use the slides in presentations and papers is granted. Reference slides as “Figure x-xx from *Analyzing Geo-Business Data*, J.K. Berry, 2003 (BASIS Press).”

1.4.3 Installing Companion Software

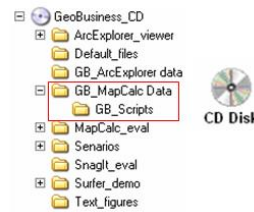
Re-access the main menu for the *Analyzing Geo-Business Data* CD and close all other open programs.

Step 1: install programs



In turn, install the four companion software programs by clicking on their respective links. Click on the **Open** button then follow the on-screen instructions using the default specifications (recommended).

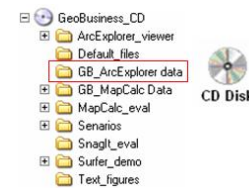
Step 2: copy special data



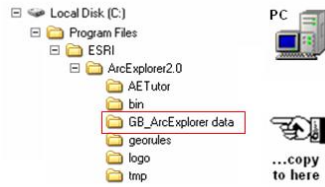
Install the data and support materials for *MapCalc* by copying the **\GB_MapCalc Data** folder on the CD to the—



...**\MapCalc Data** folder where the *MapCalc* program was installed.



Install the data and support materials for *ArcExplorer* by copying the **\GB_ArcExplorer Data** folder on the CD to the—

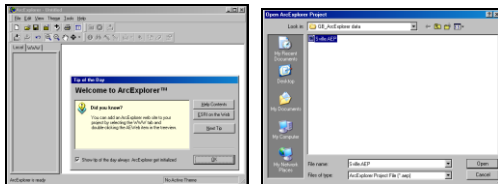


...\GB_ArcExplorer Data folder where the ArcExplorer program is installed.

Step 3: test installation

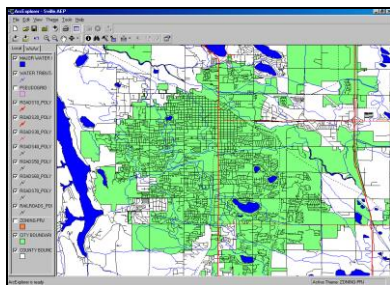
Once the programs and data are successfully installed, test the systems as described below.

- To test the **ArcExplorer** installation, click on **Start→ Programs→ ESRI→ ArcExplorer 2.0→ ArcExplorer**.



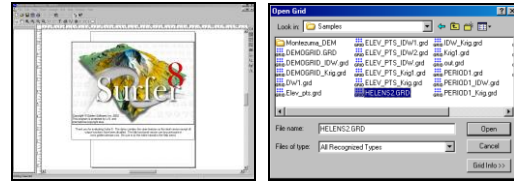
From ArcExplorer's main menu, select **File→ Open** and specify ...\\Sville.AEP as the project to open. Press the **Open** button to plot the data.

If ArcExplorer has been properly installed plot of the Smallville case study area will appear (see below; make take several moments).



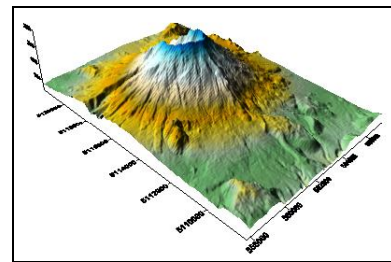
You can exit the program by selecting **File → Exit** or by clicking on the "X" in the upper-right corner of the ArcExplorer program window. Do not save the project.

- To test the **Surfer** installation, click on **Start→ Programs→ Golden Software Surfer 8→ Surfer 8**.



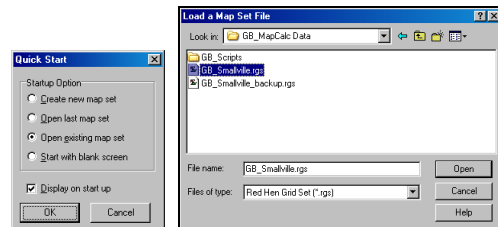
From Surfer's main menu, select **Map→ Surface** and specify ...\\Samples\\Helens2.grd as the grid layer to open. Press the **Open** button to plot the data.

If Surfer has been properly installed a three-dimensional plot of Mt. St. Helens will appear (see below).



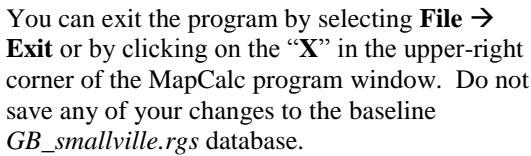
You can exit the program by selecting **File → Exit** or by clicking on the "X" in the upper-right corner of the Surfer program window. Do not save the plot file.

- To test the **MapCalc** installation, click on **Start→ Programs→ MapCalc Learner→ MapCalc Learner**.



Select **Open existing map set** and browse to the ...\\GB_MapCalc Data folder containing the workbook data and support materials. Specify **GB_Smallville.rgs** to access the data base used in this case study.

If MapCalc has been properly installed a three-dimensional plot of total customers will appear (see below).



-

If this banner page
has been successfully installed,
press the keys **Ctrl/Shift/P** to
open the screen capture utility.
Select the portion of the screen then
press the keys to capture and press the
keys to save.



Analyzing Geo-Business Data
9

