7. GIS functionality

G.J. Meaden (FAO consultant, Canterbury, United Kingdom), J. Jenness (FAO consultant, Flagstaff, Arizona, United States of America), and S.Walker (NOAA Coastal Services Center, Charleston, United States of America)

7.1 INTRODUCTION

This chapter discusses the functional tools and techniques offered by typical GIS software packages. It is possible that most data processing discussed in earlier chapters could be performed without GIS software but, in most cases, the functionality that will be discussed here are specific to GIS software.[172] As mentioned in Chapter 2, for any user organization there are several choices when it comes to selecting a GIS software package. Each package has its own strengths and weaknesses and each varies in the range of operations that they perform. More expensive GIS packages may have over a thousand "tools" or other operations and even moderately priced packages will perform hundreds of operations. It is the goal of this chapter to introduce the user to the main functionality that may be useful in a wide range of applications. It is important to note that there is no standard terminology for tools (operations) among GIS packages, i.e. the name of tools discussed in this chapter will undoubtedly differ among the various packages. In addition, this chapter is presented using a number of broad "functional categories". Thus, while tools are discussed here under particular headings, this does not mean they could not be considered as applying to other functional groups. A further introductory point is that some of the functionality of the GIS operations might best be performed using either the raster or vector format. In most cases, attention is drawn to the preferred modelling format, but it is worth noting that many proprietary GISs are already programmed to use the most efficient format. Finally, it is assumed that all data to be used for GIS work has been validated and edited so that the user is confident that the data are as reliable as possible (see Section 5.3). For more information on GIS operations and functionality, see for example, Chrisman (2001), Chang (2004), Bolstad (2005), Heywood, Cornelius and Carver (2006), Delaney (2007) and Longley *et al.* (2011).

Before examining the types of tasks that GIS may perform, it is worth recalling that in order to perform any task the GIS software is dependent on its ability to be questioned or queried. The GIS analyst must have questions to which she/he requires answers, and these questions must be answerable using the tools available in the GIS. Querying is the act of searching attribute and location information stored in a spatial database and, as was shown in Section 5.5.3, the database management system (DBMS) within a GIS software package allows users to ask conditional questions of the data (tabular or spatial). Therefore, using so-called structured query language (SQL) and based on the information provided in Box 5.4, it is possible to ask questions such as "What is the proportion of fish caught in water having a flow velocity of 5 to 10 kilometres per hour?" Similarly, based on an extended version of Table 5.1, users could ask the software to "List all ports in the Federative Republic of Brazil where wooden vessels are registered", or more complex queries such as "What is the relationship between the length of fishing vessels in the Federative Republic of Brazil and the construction material used?" A spatial query might be "Show me the locations of vessels over 30 metres in length that are more than 12 years old." The reader can see that the GIS is

[172] It is certain that some of the pre-processing tasks discussed in Section 7.2 could be carried out in CAD (computer-aided design) or other remote sensing, drawing, photography and design packages.

making the best use of the attribute information contained in Table 5.1. As described in Section 5.5.3, SQL queries rely upon the fact that the DBMS of most GIS have the ability to perform numerous "operations" on data, such as Boolean logic, arithmetic operations, geometric operations, and various statistical or algebraic operations, all of which are designed to allow the user to manipulate the data in any way desirable.

In the rest of this chapter it will be important to realize that all of the functions that the GIS can perform will rely on this ability to ask suitable questions, and that these questions can only be answered if the appropriate data are accessible to the tools and the DBMS that are integrated in the GIS. Finally, it is the intention of this chapter only to describe what the typical functions are that GIS packages may perform; no attempt is made to tell readers how to perform these functions, i.e. as this will vary with the GIS software being used.

7.2 DATA PRE-PROCESSING AND TRANSFORMATIONS

This section covers some of the "preliminary functional" areas that most GIS packages should perform, i.e. those that are concerned with data pre-processing, data transformations and data manipulation. These are all necessary functions that can be applied to data sets, either in their raw or mapped form, which aim to change the data to suit the area, scale, theme, etc., of the specific GIS task being undertaken. This section discusses:

- data or map pre-processing
- generalization
- classification and reclassification
- buffering
- overlaying

While a great deal of time might be spent on the collection or creation of data for GIS projects and applications, it is very rare that data will be in a suitable form to answer the question being examined without first being manipulated. Pre-processing, therefore, allows the user to take the data collected and then adjust, update, refine or alter the data as necessary. Most GIS packages provide several tools to prepare data for use in analysis, though these tools will vary by software package. The ability to do the majority of pre-processing in the GIS provides users with the ability to experiment with the data in order to find out which type of output best suits individual project needs. Although this section describes mainly "pre-processing" functions that are often necessary to get the digital data into a useful state, in many cases these transformations might also represent the final output from the GIS project. For example, the "overlaying" function (see Section 7.3) can be effectively used to produce extremely informative new mapping information. Here, only the most important pre-processing and transformation functions available are discussed.

7.2.1 Data or map pre-processing

At a very basic level, digital maps and data sets may require a variety of obvious manipulations in order to best suit the task being undertaken. Most of these functions would also be available in other graphics or remote sensing packages. A range of the most basic pre-processing functions is given in Box 7.1.

GIS software often provides a range of more advanced transformation or pre-processing capabilities, and these are detailed in Box 7.2. These processes will be performed directly to the mapped output before the new output is stored in the database, and some of the processes may be applicable to remote sensing output as well as to GIS.

BOX 7.1

Examples of basic pre-processing functions available in most GIS or graphics programs

The following pre-processing functions may need to be applied to any spatial data set or digital map in order to make it available for a specific GIS project:

- **Delete** – Usually required in order to reduce a data set to suit a specified area, scale or theme. Individual entities can also be deleted.
- **Cut, clip or crop** – It is often possible to simply cut or crop mapping boundaries to suit a defined project area.
- **Move** – when changing scale it might be preferable to move entities to their more precise locations. Some entities may have been incorrectly located.
- **Recode** – If map layers are merged it might be necessary to establish new classes or coding for entities.
- **Dissolve** – It may be necessary to remove specific boundary lines, e.g. when two maps have been merged.
- **Update** – It will always be necessary to update maps or data sets.
- **Zoom** – It is very useful to be able to move in or out from mapped scenes, thus effectively changing scale. However, this might not be advantageous if data has only been captured at a small-scale.
- **Join or merge** – Allows neighbouring mapped areas to be merged into a single mapped area.
- **Rotate** – Allows a map to be conveniently viewed from any direction.
- **Labels** – Labels may need to be repositioned or changed in order to reflect different map designs or layout requirements.

BOX 7.2

Transformations that may be directly applied to digital maps

Unlike the pre-processing functions described in Box 7.1, these pre-processing transformations provide the capability to directly improve mapped images or to alter them in ways that might be necessary for specific GIS tasks. These transformations are unlikely to appear in more general graphics software.

- **Edge matching** – This function allows neighbouring maps in any series having the same scale (and usually theme) to be exactly matched up (registered) along their edges.
- **Projection change** – Because the world is spherical, there is no way to display the curved surface of the world on a flat map without distorting it. There are numerous ways to project the curved data to a flat surface, and, therefore, it must be ensured that the GIS recognizes the projection method used (see Section 3.3.4 and Figure 7.1). For excellent introductions to projections, please refer to Snyder (1993) and Iliffe (2008).
- **Datum (coordinate) transformation** – Similar to projection change, it is important that GIS can recognize a range of georeferencing systems (Section 3.3.4). Most modern GIS packages have "behind-the-scenes" functions to change between coordinate systems.
- **Rubber sheeting (warping)** – There are various reasons why maps might suffer from distortions, especially those derived from aerial photographs or scanned images, but it is important that they can be overlayed, if necessary, and they can be registered to appropriate geo-coordinates. Thus, rubber sheeting allows maps to be "contorted" as necessary so that the whole mapped area is adjusted to accurately locate control or registration points.
- **Orthorectification (geometric correction)** – This process corrects inevitable spatial distortions in satellite or aerial imagery caused by the topographic shape of the landscape, and is necessary if these images are to be used in mapping or GIS-based analyses.
- **Image enhancement** – This includes a range of techniques whose purposes are to improve satellite imagery so that images are clear and are well defined.
- **Structure conversion** – It is important that GIS packages have the ability to switch between working in vector or raster modes. Sometimes this is done automatically according to the GIS tool being used, but it can be user controlled. Figure 7.2 explains the procedure. It can be seen that there is some information loss if this procedure is carried out.

Figure 7.1 illustrates the importance of one of these transformations, i.e. that of projection change. Here, it can be seen that, for comparison purposes, the continent of Africa has been drawn using three different projections. All three versions of Africa in this image are perfectly accurate and yet, because of the projections used, they are clearly different from each other, with some countries appearing to occupy very different geographic locations.

FIGURE 7.1
Comparison of three different map projections of Africa

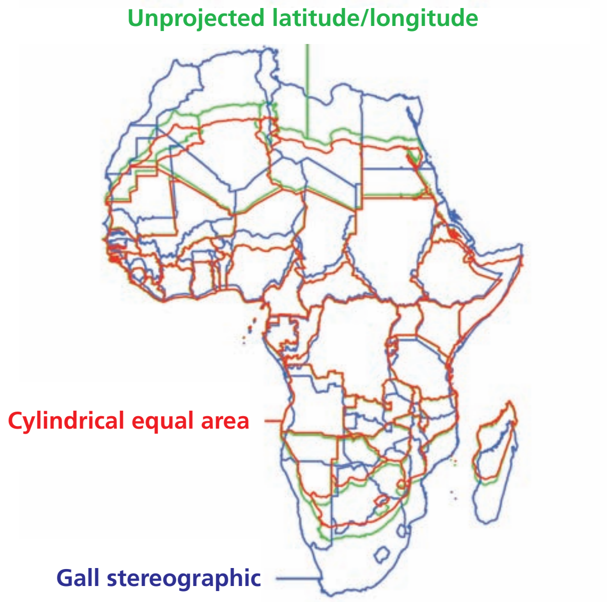
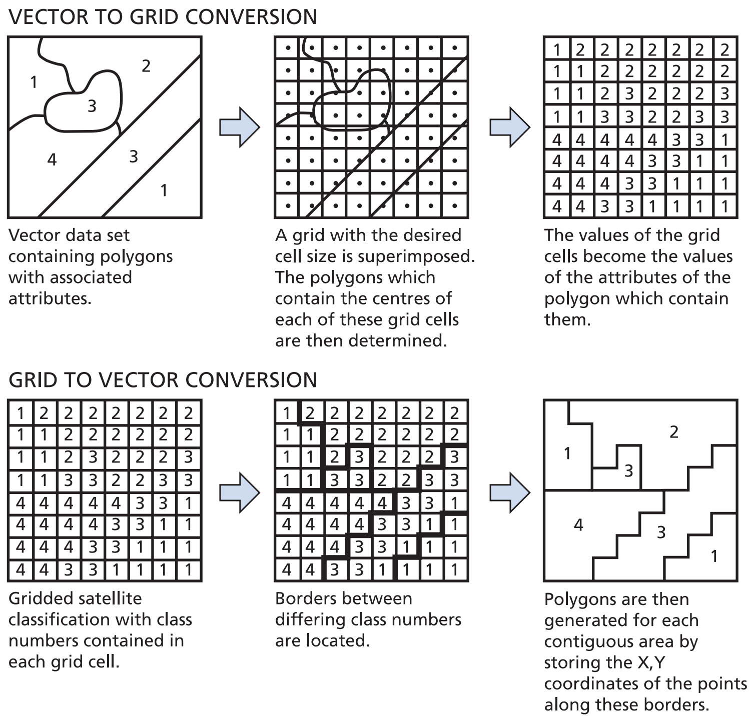


FIGURE 7.2
The vector to raster and raster to vector conversion process



Source: Meaden and Kapetsky (1991).

7.2.2 Generalization

Generalization is the abstraction, reduction and simplification of features so that they can be more easily used for particular mapping purposes. Thus, a major problem in cartography is that mapping at different scales requires different levels of detail to be captured at each scale. Box 7.3 shows the types of information that may reasonably be shown at three mapping scales. Details that are captured at a large-scale cannot possibly be shown on a map that is redrawn at a small-scale because individual mapped features would be too small. Likewise, the information captured at a small-scale often looks coarse and imprecise on a map redrawn at a large-scale because not enough detail will have been recorded.

BOX 7.3
Typical information shown on topographic maps at different scales

Large scale (1: 1 000 to 1: 10 000)
- plot boundaries and land parcels
- house or building outlines
- footpaths, grass verges, road widths
- small jetties

Medium scale (1:25 000 to 1: 100 000)
- all main roads and rail tracks
- all rivers over perhaps 3 metres in width
- most patches of woodland
- urban area shapes but perhaps not each residential road
- small villages and hamlets shown as dots
- contours at perhaps 20-metre intervals

Small scale (1:250 000 to 1:1 000 000)
- most towns as dots plus the location of larger villages only
- main rivers over perhaps 20 metres in width
- main roads and railways only
- contours at perhaps 100- to 250-metre intervals

Most GIS packages have several tools that fall under the heading of "generalization". Figure 7.3 shows brief examples of how a river, a group of buildings and patches of woodland might be generalized. Generalization typically includes the elimination of smaller features, the merging of features, the simplification of lines and a reduction of complexity, and is most appropriate when moving from large-scale towards a smaller scale. Figure 7.4 shows how "linear generalization" may be performed by a "line thinning" algorithm within the GIS software. Here, the original digitized outline of a river is shown in blue. In the first version of the river, the line thinning algorithm in the GIS will have reproduced the river's outline based on perhaps every tenth digitized location captured during the original digitization process, and thus is rather crude but appropriate for smaller-scale maps. The second river outline may be based on every fifth digitized location and is more accurate and suitable for larger-scale maps.

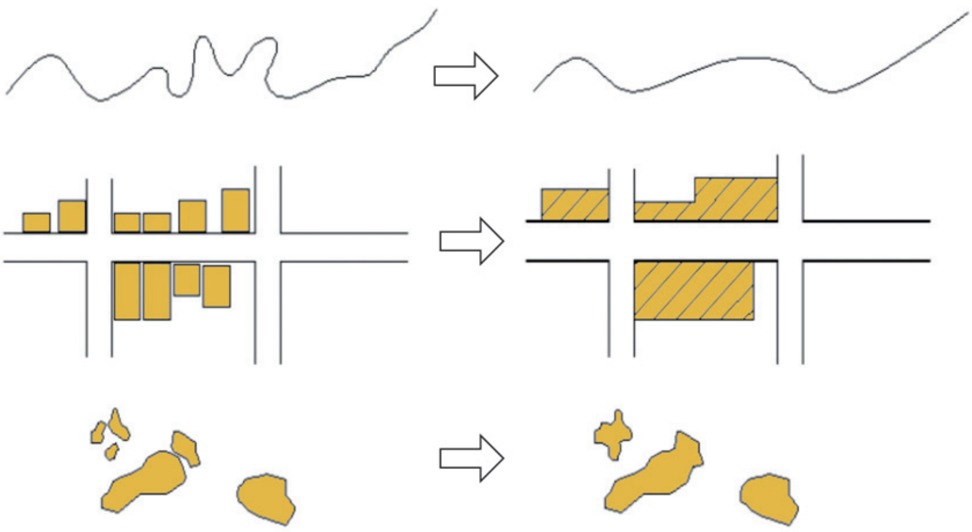
It should be noted that there are two classes of generalization:

(i) **Graphical generalization.** This is concerned with generalizing the graphical features on maps, i.e. those consisting of any points, lines and polygons.

(ii) **Semantic (or conceptual) generalization.** This is concerned with the actual entities that are considered necessary for mapping at different scales, e.g. at a large scale there may be houses, apartments, shops and factories, but at a small-scale the map may only show "buildings".

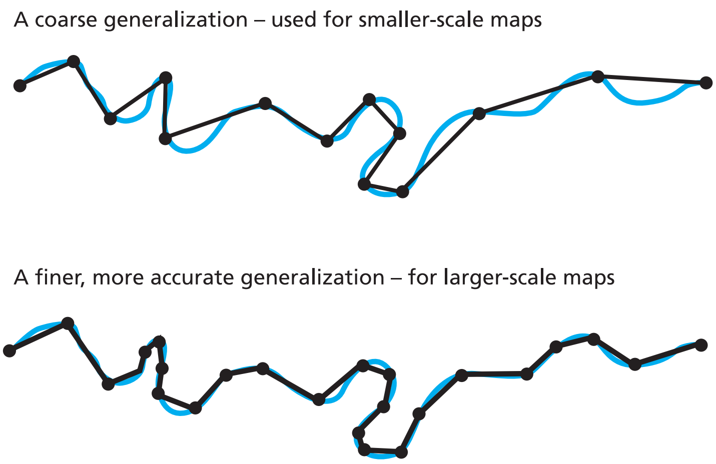
As well as generalization based on vectorized points, lines and polygons, it is possible to generalize rasters. Here, the GIS software may use an algorithm that calculates a mean value of, for instance, a central cell plus its eight surrounding neighbour cells, and the new raster will have a cell size equal to that original block of nine cells. This would create a more generalized map with the benefit of approximately an 88 percent reduction in data storage. As with most cartographic processes, there are no strict "right or wrong" methods regarding generalization and it will be up to the GIS operative to make sensible choices. Further details on generalization can be obtained from Laurini and Thompson (1992), Robinson *et al.* (1995), Li (2007) and Harvey (2008).

FIGURE 7.3
Examples of pre- and post-generalization in mapping



Source: ICIMOD (2011).

FIGURE 7.4
Line thinning at two levels of generalization



Source: Modified from Sinfogeo (2012).

7.2.3 Classification and reclassification

Classification and reclassification are subsets of generalization. The classification of data is necessary as a means of simplifying (generalizing) the complex real world and classification can be applied to each of the various types of data discussed in Box 3.5. Classification is the division of data into a specified number of classes using one of several classification methods, and Table 7.1 lists five common methods for classifying numeric data.

Figure 7.5 illustrates three of these classification methods in more detail. The histograms plot the number of counties in the United States of America on the vertical scale and the percentage of the population of each county who is more than 65 years old on the horizontal scale. Each example classifies the counties into ten classes, and these classes are then mapped such that the colour corresponds to the class. The top example divides the data into equal intervals, meaning that percentages shown on the horizontal scale are divided equally (perhaps 20 percent–30 percent; 30 percent–40 percent; 40 percent–50 percent, etc.). The middle example divides the data into what are termed "Jenks natural breaks", based on recognizable breaks in the data distribution (see Slocum *et al.*, 2005, for a full explanation of this method). The lower example divides the data into classes that each contain the same number of counties.

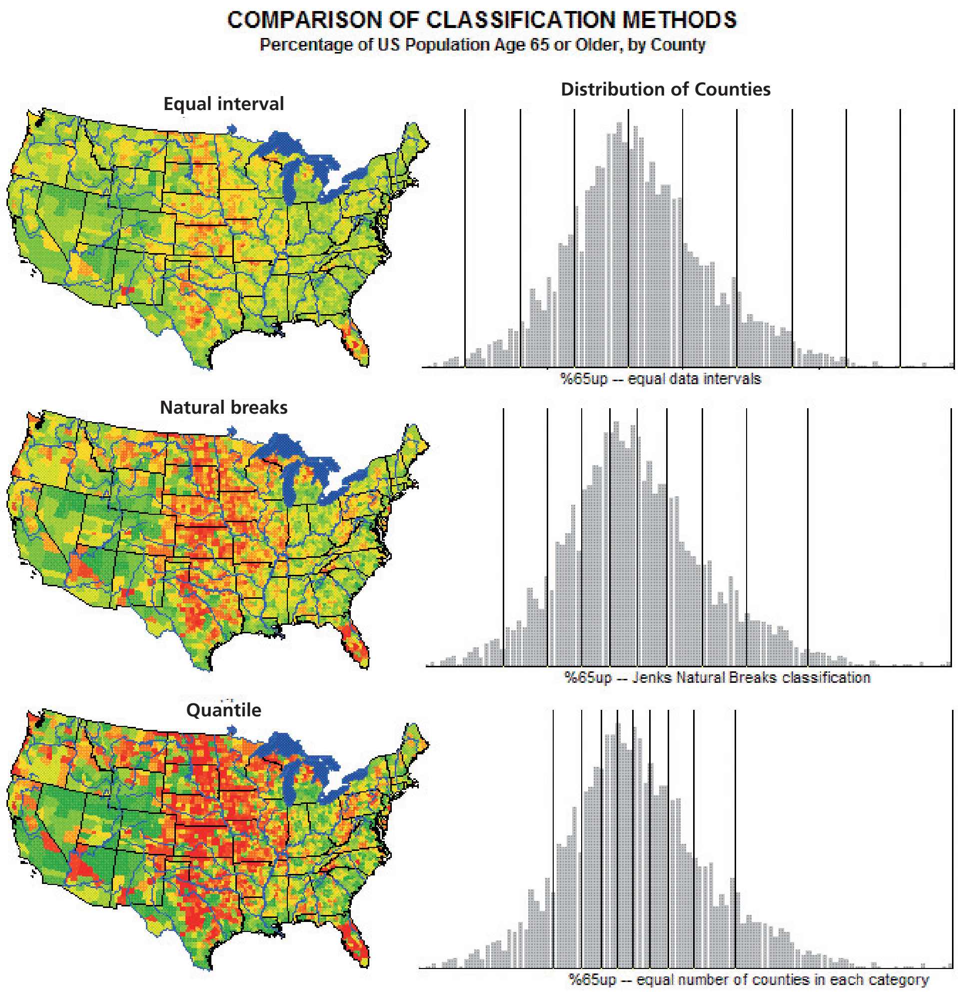
TABLE 7.1
Five methods to define class intervals for numerical data

Classification method	Description
Equal interval	Divides values into groups with an equal range of values.
Quantile	Creates classes that have the same number of features in each class being mapped.
Standard deviation	Finds the mean and then creates class breaks based on standard deviations above and below the mean until all values are contained.
Natural breaks (Jenks)	An algorithm defines classes based on natural break points in the data.
User defined	The user establishes class breaks at their own discretion.

Regardless of which classification method is used, the effect is to change the look of the data being displayed and often the spatial patterns highlighted. Data classification can also be used to create new data. For vector data, it is as simple as selecting the features that fall into the new class and then assigning a new value to those records that correspond to the new class. When working with raster data, it is possible to similarly create new data from an existing data set by classification, which is known as reclassification. Here, users can simply assign a new value to a cell or they can assign new values to a range of cell values. Table 7.2 shows the reclassification of data whereby the original population classification densities have been reassigned a new value (or coding). There are no specific rules regarding how many classes the data should be divided into, but recommendations usually say that five to seven classes are preferable for numerical data. If nominal data are being used, i.e. classes based on named categories such as soil or vegetation types, more classes can be used within the limits of easy colour and/or shading recognition. Finally, most proprietary GIS have menus or dialogue boxes that allow users to make easy classification selections, including choice of value ranges and colours or colour ramps.[173] For more information on classification, see Krygier and Wood (2005), Slocum *et al.* (2005) and Diaz (2006).

[173] Colour ramps are logical progressions of colour that depict ranges going from high to low values. These ramps may not logically apply to nominal values.

FIGURE 7.5
A comparison of three important methods of classifying data



Source: Mackenzie (2011).

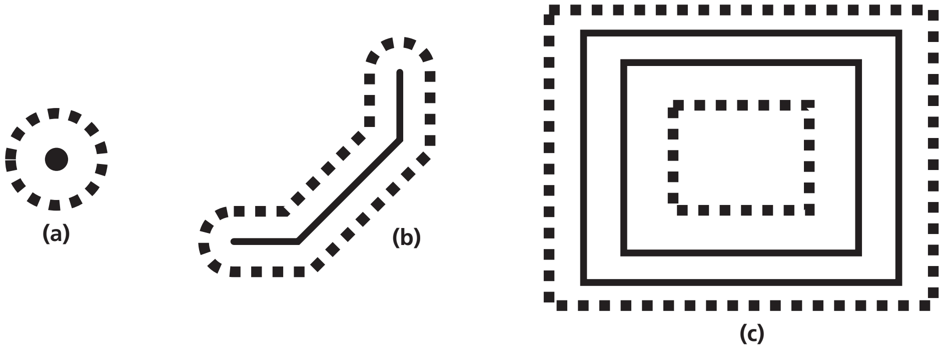
TABLE 7.2
Example of a reclassification table

Population density (inhabitants/km^2)	Reclassification coding	Possible interpretation
0–30	1	Low density
31–60	2	Moderate density
61–90	3	Fairly high density
91–120	4	High density

7.2.4 Buffering

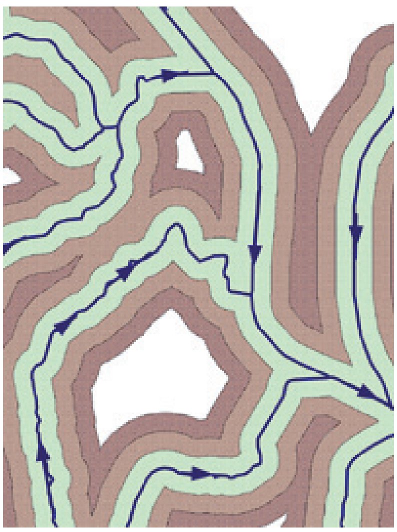
A buffer is a polygon created at some user-specified distance around a point, line or polygon feature. Figure 7.6 shows examples of buffers generated around the three main feature types. Notice that for polygons buffers may be drawn inside or outside any specified feature. Most GIS packages allow the user to create multiple buffers around a feature, as demonstrated by the three equal distance buffers created along sections of a stream network in Figure 7.7. Users are able to specify the number of buffer zones they wish to create as well as the distance increments at which the software creates the new buffers. The distances at which buffers are drawn, or the number of buffers used, are usually arbitrary decisions based on an "educated perception" of their need. For example, it would be possible to define mangrove areas as buffer zones (e.g. 50 m from the shrimp farms) to mitigate the potential impacts for wastewater discharges from shrimp farms. Clearly, some buffers can be more accurately defined than others and, indeed, some GIS allow for the drawing of variable buffers along portions of the same line or polygon.[174]

FIGURE 7.6
Buffers formed around (a) points, (b) lines, and (c) polygons



Buffers are the basis for much GIS analysis. For instance, if the land area shown in Figure 7.7 contained a number of mapped point sources of pollution, the GIS could provide an answer to the question "How many sources of pollution in the area are within 200 m of a stream?" The answer would simply be the number of points that intersect the 200-m wide buffer. More complex buffering algorithms exist that allow for differential movements or distances around a point source. For instance, if an oil spill occurs in an area of mixed geology, the oil may seep through some rocks at faster rates than others. If the seepage rates are known, then buffers can be calculated to show where (and how far) the oil will have reached after any specified time interval.

FIGURE 7.7
Multiple buffers surrounding part of a stream network



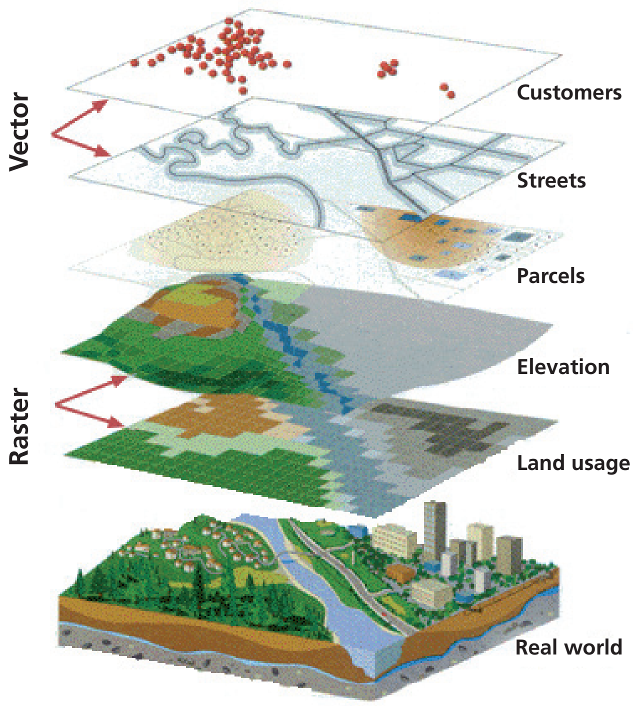
Source: United States Geological Survey (2007).

[174] This is particularly useful where lines represent features having variable widths, e.g. rivers.

7.3 OVERLAYING

Overlaying is basically the integration of mapped data from two or more different sources within the same area, and it is arguably the most powerful function of a GIS. Any mapped area of the earth's surface can be subdivided (classified) into any number of different thematic "layers". Thus, a typical map will conventionally show rail lines, the road network, forested areas, urban areas, the river network, the coastline, etc., and each of these themes can be mapped separately. Figure 7.8 illustrates that it is also possible to have different layers containing raster or vector maps, and perhaps topographic maps, historical maps, land use maps, remotely sensed images, etc. Any separate mapped layer can be overlaid with any other layer (or layers) as long as it is in the same projection and covers the same area. If, however, a raster layer is overlaid by a vector-based layer (or vice versa) with the intention of analyzing or integrating the data, then one or other layer would typically need to be converted (see Figure 7.2). Any new map produced will need to be named, filed and stored before it can contribute to further overlaying or other types of analysis.

FIGURE 7.8
Selected mapped layers for a specified area using points, lines, polygons, vectors and rasters



Source: Heard (2010).

Performing overlays using vector data can be computationally intensive because of the huge number of mathematical operations that will typically be performed in the course of the analysis. Furthermore, GIS operatives need to be cautioned that when doing overlays of vector data sets, there is often the problem of generating "false (or sliver) polygons". These are often due to the fact that the individual layers being overlaid were digitized separately and thus the points, lines or polygon boundaries on each map do not match up exactly (see Figure 5.3). Extensive editing may be needed in this case. Raster-based overlays are typically much simpler for the GIS to perform because individual cells are directly overlaid on other cells, although similar problems arise when the rasters have different cell sizes or origins. Raster overlay and the use of weightings are at the heart of map algebra (see later in this section).

A main purpose of overlaying is to see what relationship(s) might exist between different themes (layers). For instance, it is highly likely that there will be a relationship between soil type and land use, and GIS can tell the user facts about the closeness of this relationship. Other purposes for overlaying are described as follows:

(i) Creation of new features. It may often be useful to create a more complex map from existing data. For instance, data might be available for a specific area showing both "soil types" and "main vegetation classes". By overlaying and merging these two data sets a new map can be created showing "soil with vegetation classes".

(ii) Allocation of weightings. Sometimes layers are overlaid and merged in order to find the optimum location for some type of commercial activity. For instance, Figure 1.6 showed a map designed to seek optimum locations for shrimp production in western Democratic Socialist Republic of Sri Lanka. Fourteen layers were overlaid and merged to create this map and each was assessed as being of different importance to the success of shrimp production. Each layer was given a relative numerical weighting and this value was automatically assigned by the GIS to every pixel in that layer in the raster database. The 14 weighted layers were then merged in order to provide the final mapped output. The assignment of weightings can be rather arbitrary and, in fact, can be one of the most difficult parts of the analysis, but in many cases weightings can be agreed by a panel of experts.

(iii) Map algebra. Just as in mathematics where algebraic formulae have been designed to help with more complex computations, similar processes can be used in raster-based GIS. Basically this means that a formula can be developed, usually for modelling purposes, that allows the user to carry out a range of mathematical, e.g. addition, subtraction and multiplication, or algebraic manipulations to the data in order to achieve some goal. Figure 7.9 provides a very simple illustration of the basis upon which map algebra works. Most of the output achieved by the CHARM 2 team (see case study 10.4.1) utilized quite complex map algebra within the overlay process[175]. Map algebra functions are extremely powerful and allow the user to analyze and generate new data based on data they may already possess. A detailed description of map algebra functions is beyond the scope of this technical paper; for more information, see Arlinghaus and Griffith (1996), Batty and Longley (2003), O'Sullivan and Unwin (2003), de Smith, Goodchild and Longley (2007) and Bivand, Pebesma and Gomez-Rubio (2008).

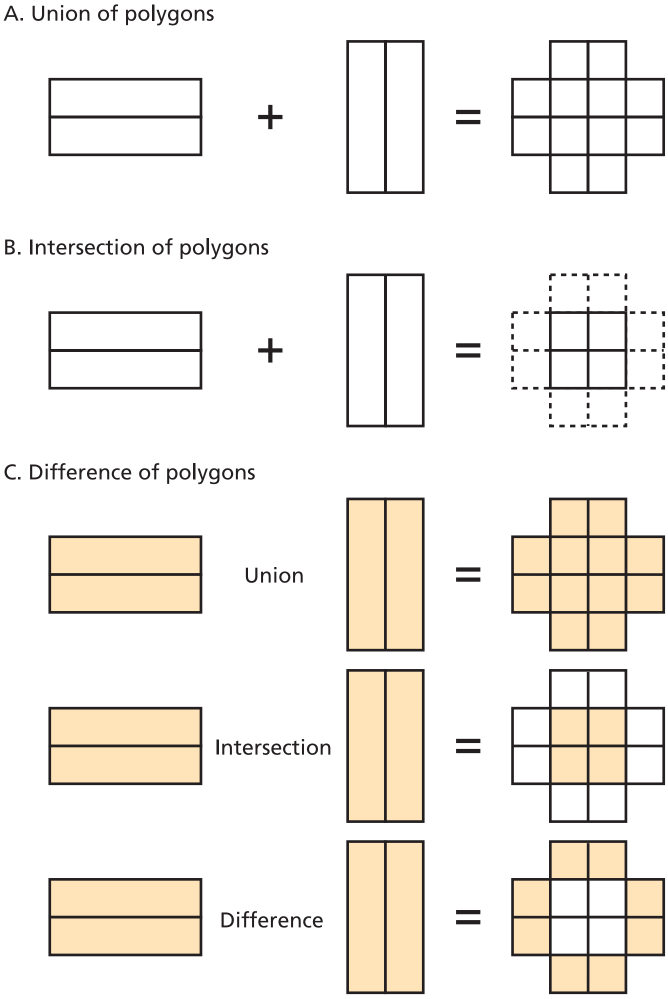
FIGURE 7.9
Illustration showing simple map algebra function

$$\begin{array}{|c|c|c|c|}\hline 6 & 4 & 2 & 6 \\ \hline 2 & 8 & 6 & 6 \\ \hline 6 & 4 & 8 & 8 \\ \hline 4 & 6 & 2 & 6 \\ \hline \end{array} + \begin{array}{|c|c|c|c|}\hline 3 & 2 & 1 & 3 \\ \hline 1 & 4 & 3 & 3 \\ \hline 3 & 2 & 4 & 4 \\ \hline 2 & 3 & 1 & 3 \\ \hline \end{array} = \begin{array}{|c|c|c|c|}\hline 9 & 6 & 3 & 9 \\ \hline 3 & 12 & 9 & 9 \\ \hline 9 & 6 & 12 & 12 \\ \hline 6 & 9 & 3 & 9 \\ \hline \end{array}$$

For most overlay operations, it is possible to perform a range of post-overlay procedures. This means that any resulting map can be edited in various ways so as to retain only desired information. For instance, Figure 7.10 (A) illustrates the overlaying (or union or merging) of two sets of polygons, which retains all the areas from both sets of polygons. Figure 7.10 (B) shows the intersection of two sets of polygons, which retains only the areas that are common to both data sets (see Lo and Yeung, 2002, for additional

[175] Map algebra is fundamental to the science of spatial statistics.

FIGURE 7.10
Overlay operation combines an input and overlay polygon



information). Figure 7.10 (C) shows the "difference" overlay operation, which subtracts the second set of polygons from the first set.

The simple overlaying of a polygon layer on top of a point layer can yield useful information about distributions. For example, if data are available to map survey results showing the distribution of a fish species in a particular area, then this point distribution can be overlaid by a vector map showing bottom sediment classes. The GIS can then calculate the numbers, the proportion and the density of the species within each sediment class.

Finally, it is worth considering the fairly complex range of analyses that may be performed by using a combination of buffering and overlay procedures. Again using Figure 1.6, which has many of its 14 layers based on buffering, the GIS-based map could be overlaid with a map of population density (by classes), and it would be possible to ask questions such as "How many people live within areas that are most suitable for shrimp farming?" or "What is the relationship between areas that are unsuitable for shrimp farming and the population distribution?

7.4 MEASUREMENT

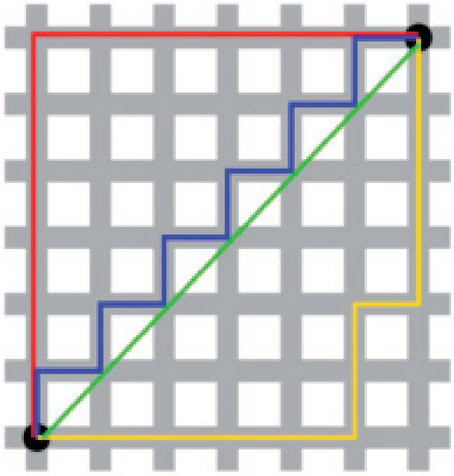
GIS software is capable of performing many types of measurement, including simple enumerations or counts, linear distances, areas, perimeters, volumes, directions and angles, plus a range of less frequently used measurements such as "cost" surfaces, weighted mean points, slope angles, directions, stream sinuosity, ratio of area to perimeter for polygons (edginess) and travel time distances. It is important to remember that various factors can affect the accuracy of measurements, including the projection of the data, whether the GIS accurately calculates areas and distances using different geographic coordinates, the precision and accuracy at which the data were collected, and the fact that the data typically do not exactly match the real world features (i.e. the real world may have curved edges, or fuzzy edges where one feature type gradually changes to another, while the GIS assumes clear boundaries made of straight lines). Because all forms of measurement are performed very differently in raster-based GIS compared with vector GIS, it is instructional to examine measurement under these two headings.

7.4.1 Raster-based measurement

If a raster is projected so that the units are linear (e.g. metres or feet), then line measurement using rasters is simple given that the cell (pixel) size is known. A measurement either horizontally or vertically along the raster columns or rows is simply the number of pixels multiplied by the width of a pixel. However, any other raster-based measurement will be more complex. Raster-based distances that diagonally cross the rasters (Figure 7.11) are 1.414 times as far as straight vertical or horizontal distances, so the green line is 8.5 distance units compared with a vertical or horizontal (red lines) measurement of 6 units. Figure 7.11 also illustrates that some raster-based measurements utilize so-called "Manhattan" distance (blue line) whereby distances are measured by traversing along an imaginary line going "across and up" the side of every pixel, thus increasing the real distance by approximately 30 percent from the straight diagonal distance, i.e. to 12 distance units. So, rasters are not a good basis for length measurements of linear distances. However, Figure 7.11 shows that rasters can easily be used to accurately calculate perimeter length (e.g. 6 × 6 = 36 distance units) and polygon area (6 × 6 = 36 aerial units). The same principle used to calculate area can be applied to measurements of spatial volume, i.e. width × length × depth.

Geographic (also known as unprojected) rasters are not defined in linear units (such as metres or feet) but rather as angular units (degrees), and thus are considerably more difficult to measure from and requires that the GIS knows how to handle geographic units. The basic problem is that the raster cells only look square when viewed in an X/Y plane but, in fact, the cells form trapezoids on the actual surface of the earth (Figure 7.12). Furthermore, the shape of the trapezoids depends on the latitude of the cell, with cells becoming narrower as they approach the poles.

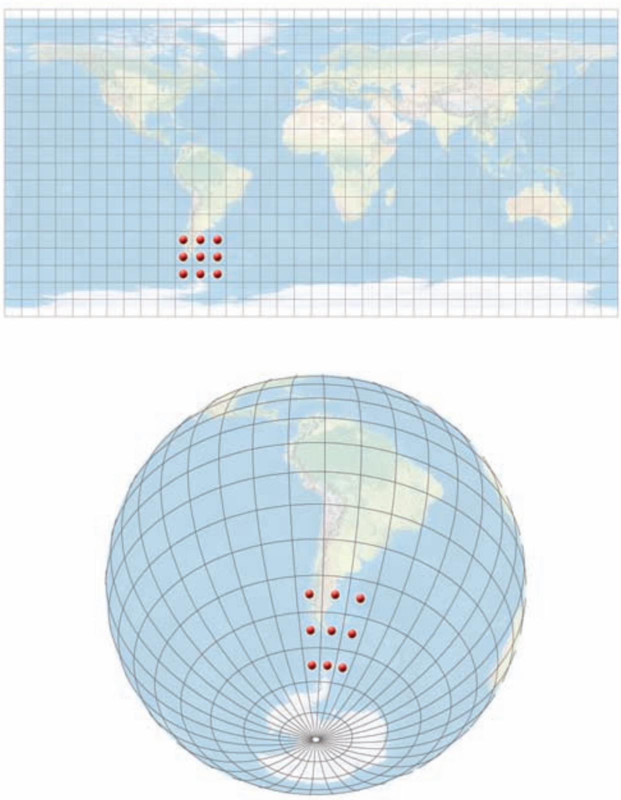
FIGURE 7.11
Raster GIS measurement showing diagonal distance (green); Manhattan distance (blue); and other random distances from start to finish (red or yellow)



Note: Red, blue and yellow lines are all the same length to show that whichever logical route is taken from one black dot to the other the travel will be the same distance.

Source: Ask.com (2012).

FIGURE 7.12
Trapezoidal raster cells in an unprojected raster



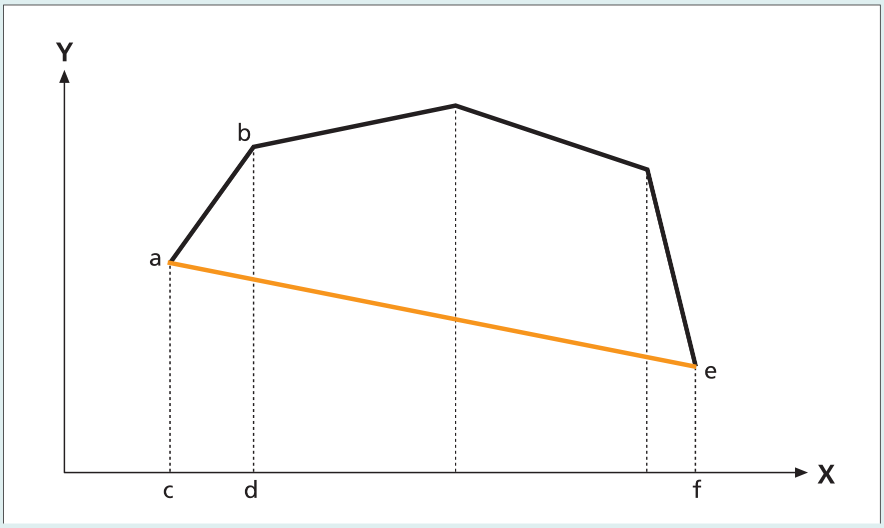
7.4.2 Vector-based measurement

Vector-based GIS provides a much more accurate basis for measurements of distance, area and centrality because the vertices, segments and edges are precisely defined. Measurements on projected data require simple Euclidean geometry, and the distance between any two points can easily be calculated using the Pythagorean theorem.[176] Euclidean geometry is also used to calculate perimeters or areas, with most GIS software using the so-called trapezoidal method (explained in Box 7.4) to calculate area.[177] More advanced GIS functions can incorporate the three-dimensional surface of the landscape to calculate the lengths and areas of vector features on hilly terrain.

The Pythagorean theorem and the trapezoidal methods, however, only work with projected data because these methods assume the vector features lie on a flat plane. Unprojected vector data (i.e. in latitude/longitude coordinates) lie on the curved surface of the planet and, therefore, require more complex methods (drawn from spherical and spheroidal geometry) to calculate distance, area and centrality. Most GIS systems use Vincenty's method to calculate the great circle distance[178] and direction between any two points (Vincenty, 1975).

BOX 7.4

Method used by GIS to calculate the area of vector-based polygons



The polygon consists of four heavy grey lines plus a brown line forming the fifth side. Trapezia are dropped from each edge to the x-axis and their areas are calculated, i.e. for trapezium a,b,d,c the area is the length c,d times the average of height a,c and height b,d. The areas for the four trapezia that meet with the x-axis are summed. The area of the trapezium e,a,c,f is then calculated and the result is subtracted from the sum of the four trapezia.

In general, if the two polygon vertices used in the trapezium move to the right (as in trapezium a,b,d,c), then the area of the trapezium is added to the total. If the two polygon vertices move to the left (as in trapezium e,a,c,f), then the area of the trapezium is subtracted from the total.

Source: Adapted from Longley *et al.* (2005a).

176 See Purplemath.com (www.purplemath.com/modules/distform.htm), which provides detail on the Pythagorean measurement of length.

177 Information on this can be found at Geocomputation (www.geovista.psu.edu/sites/geocomp99/Gc99/076/gc_076.htm).

178 The great circle distance is the shortest distance between any two points on the surface of a sphere measured along a path on the surface of the sphere (as opposed to going through the sphere's interior).

Although vector-based measurement methods are generally more accurate than raster-based methods, they are more computationally demanding. Often vector objects can be composed of thousands of segments and can, therefore, easily require tens or hundreds of thousands of mathematical operations to calculate areas, lengths or centroids. This is especially the case with unprojected data or when the vector object approximates a curved line (like a circle). Curves are generally approximated using many short straight-line segments and, therefore, typically have large numbers of vertices and segments. Longley *et al.* (2005a), Dale (2005) and DeMers (2008) provide good detail on various facets and means of measurement in GIS, and the first of these publications makes the important point that most GIS distance measurements tend to result in underestimations of the true distances.

7.5 SPATIAL RELATIONSHIPS

Most modern GISs offer a wide range of spatial relationship functions. In this section, the various ways in which objects or areas are associated in space are described. The types of questions for which answers are sought may be as follows:

- Are there any "spatial patterns" across the landscape?
- Are distributions of features or objects random or regular?
- Are there density variations of features and what might this be related to?
- What is the degree of association between any features or objects?
- Have there been temporal or spatial changes in distributions across space?

It is easy to see how answers to these questions could have a profound influence on fisheries and/or aquaculture management, notably in terms of the appropriate scale or level of detail in which to carry out a project, or in deciding on the best sampling or data collection strategies. To derive answers to these types of questions, the broad range of GIS functions will each be examined under those headings that are conventionally applied, though different authors or software packages may have alternative preferred names. It should be mentioned that, according to the density and distribution of sampling information held, it might be necessary to utilize interpolation methods in order to gain additional data. These methods are described in Section 7.5.3.

7.5.1 Measures of centrality

Most readers will be aware that in statistics or mathematics there are various concepts of "centrality" such as the mode, the mean and the median. As these measures represents centrality in the single dimension, so there are various concepts of centrality in two dimensional space. Here, two useful methods of establishing centrality using GIS are briefly described, i.e. establishing central points and constructing Thiessen polygons.

Establishing central points

Knowledge of spatial centrality can be useful in several ways. For instance, because it minimizes the average distance that customers may have to travel, there are many reasons why a central location can give economic advantages for businesses. Thus, retail outlets prefer central locations and processing plants may wish to locate in a location that is central to their raw material sources. This latter point may be particularly important to both fisheries and aquaculture where fish processing is now a large-scale commercial activity.

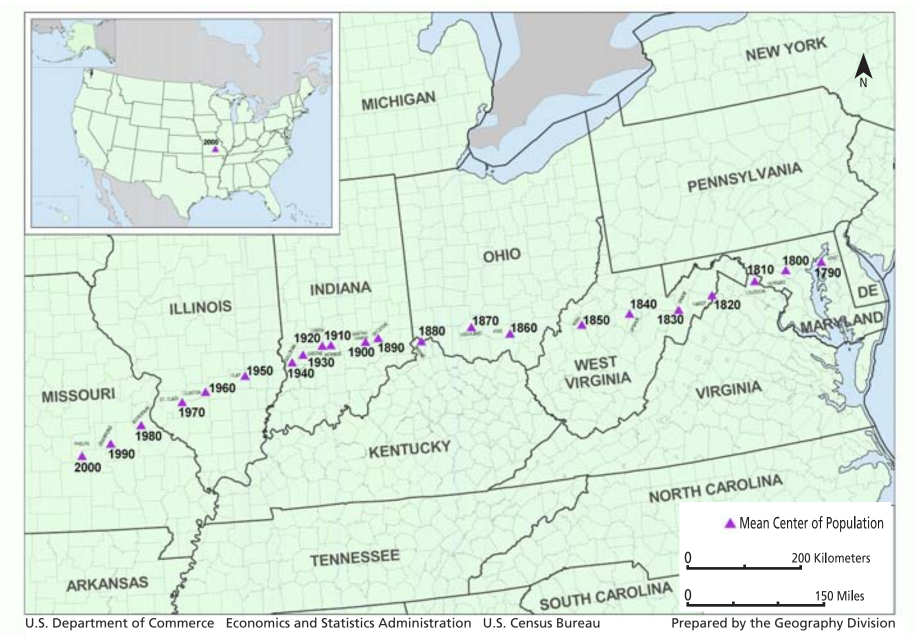
Most GIS packages have the functionality to locate various forms of centrality. One measure of centrality is that of locating the central point of a polygon – the so-called "centroid", or mean centre position[179]. This point can be envisaged as the point at which an imaginary cardboard cut-out of the polygon would be able to balance on the point of a pin. In vector analysis, centroids are sometimes used as "handles" that facilitate

[179] It is also the point that minimizes the sum of squared distances.

the default position for placing map labels, and they may be used in several analysis operations, for example, when performing an operation like distance calculations, the software will find the centroid of the polygon and use that point from which to do any measurements. Centroids are also used in raster analysis, for example, when determining a water flow path, centroids are used to generate lines representing this path.

Centrality can also be thought of as the spatial centre of gravity. This means that for many distributions of points within a given area the distribution will show some level of clustering. The centre of gravity will be the point within the given area around which the distribution is best centred. For instance, in an unpublished dissertation, Meaden (1978) calculated the centre of gravity of catfish farms in the lower Gulf States of the United States of America for successive years from 1970 to 1976. During this period, the author found that the centre of gravity migrated about 70 miles south-eastwards from near Indianola to near Benton in Mississippi state. On a longer time scale, Figure 7.13 shows how the centre of gravity for the population living in the United States of America has moved westwards from a position in Maryland in 1790 to a position in Missouri in 2000. A variation of the straight forward centre of gravity is the "weighted mean centre". This means that the items forming the distribution may not all be of equal importance, and therefore more "important" items may be allocated a weighting before the centre of gravity is calculated. Almost all measures of centrality are most efficiently calculated using vector-based GIS.

FIGURE 7.13
The mean population centre of the United States of America (1790 to 2000)



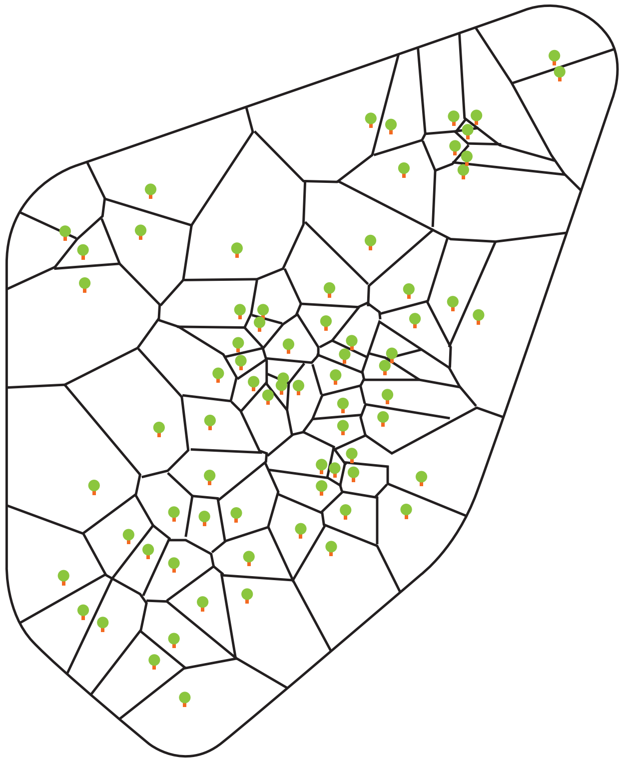
Source: Jacobs (2009).

Thiessen polygons (or Voronoi diagrams)

Another measure of centrality that can be performed by GIS is the creation of proximal polygons, better known as Thiessen polygons[180] or Voronoi diagrams. Using this method, users can input point data and then use the GIS software to generate polygons that surround each of the points. Figure 7.14 shows an example of the polygons created

[180] These polygons are named after Alfred Thiessen, who in 1911 devised a method for best assigning single data values (in this case to rain gauge readings) to areas on a map where data collection points are highly irregularly placed.

FIGURE 7.14
An illustration of Thiessen polygons



Source: Chou (1997).

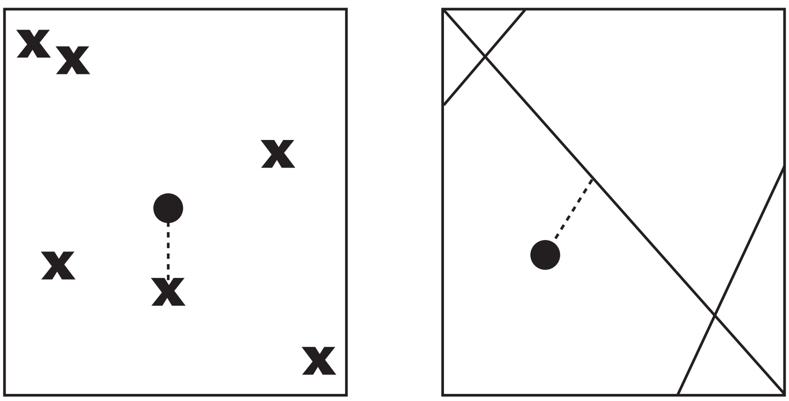
by a Thiessen polygon operation. This type of operation can be useful for delineating territories or regions of influence. Note that the measure of centrality being established here is an answer to the question "From where I am situated, which is my closest central point?" So, the central point is likely to represent some kind of service centre, or the points of the data set could represent seafood distributors, and the Thiessen polygons might show their potential market areas based on distance from the distributor.

Thiessen polygons are not intended to have a regular shape or to form a tesselation of identical shapes. They are intended to divide the landscape into polygons around a set of points, with each polygon constructed around a single point in such a way that all the area within that polyon is closer to that point than to any other point. Figure 7.14 shows that there is no place within any of the polygons that is closer to another point than it is to the polygon point. The only time Thiessen and/or Voronoi polygons would have a constant size and shape would be if the points were distributed in a regular array.

7.5.2. Measures of proximity

Proximity analysis is concerned with the distances between different features, and the numbers of a feature *a* that may be near to a feature *b*. GIS analysts may often want to know the nearest feature to some other feature (such as the nearest market to a fish farm, or the nearest fishery to an oil spill). Figure 7.15 illustrates the basic "nearest feature" proximity operation, and it is easy to see how this function can be expanded to identify all the features within a specified distance, or perhaps

FIGURE 7.15
Examples of the point-to-point and point-to-line proximity operations



the nearest x feature to a specified point. Typically, users may perform a variety of proximity analyses as a step in their overall analysis process, meaning that the results of this type of analysis may generate data, often displayed in the form of tabular or graphical output, for use in other analytical operations or GIS output.

Quadrat counts or points and/or lines in polygons

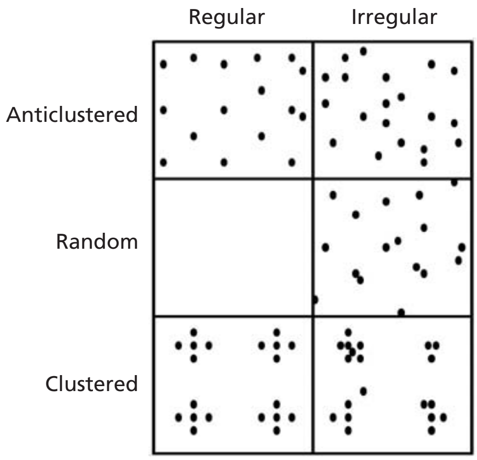
The measurement of proximity differs between feature types. Proximity between polygons can be assessed by calculating the shortest distance between polygon perimeters or by computing the distance between the centroids of the polygons (see Section 7.4) (Chou, 1997). The distance between points is established by simple measurement between points.

One of the issues with using any "count within polygon" method is determining the size and placement of the polygon being used.[181] Count operations make use of buffering techniques, so if the user wanted to know "How many fish farming operations are within x kilometres of a fish processing or fish meal plant", then a buffer would be drawn at the required distance around the plant so that the count could be made by the GIS.

Point pattern (or nearest neighbour) analysis

Another type of spatial proximity analysis is examination of the point pattern of the data, with a view to gaining insights into underlying causes of these patterns. These points could be fish hatcheries, cities, etc., and they can represent features of varying scale. As shown in Figure 7.16, point patterns can be classified as being regular or irregular (organized or disorganized) and these could be subdivided into anticlustered (or uniform), random or clustered.[182] In practice, the points representing the distribution of any single object can vary from being absolutely regularly spaced through to being absolutely clustered. However, these patterns are very much determined by the scale or resolution of the study, i.e. what appears to be a random distribution at a very large scale (small area) may in fact be a clustered distribution at a small scale (large area).

FIGURE 7.16
Possible spatial distribution patterns in point data



Source: Cole and King (1968).

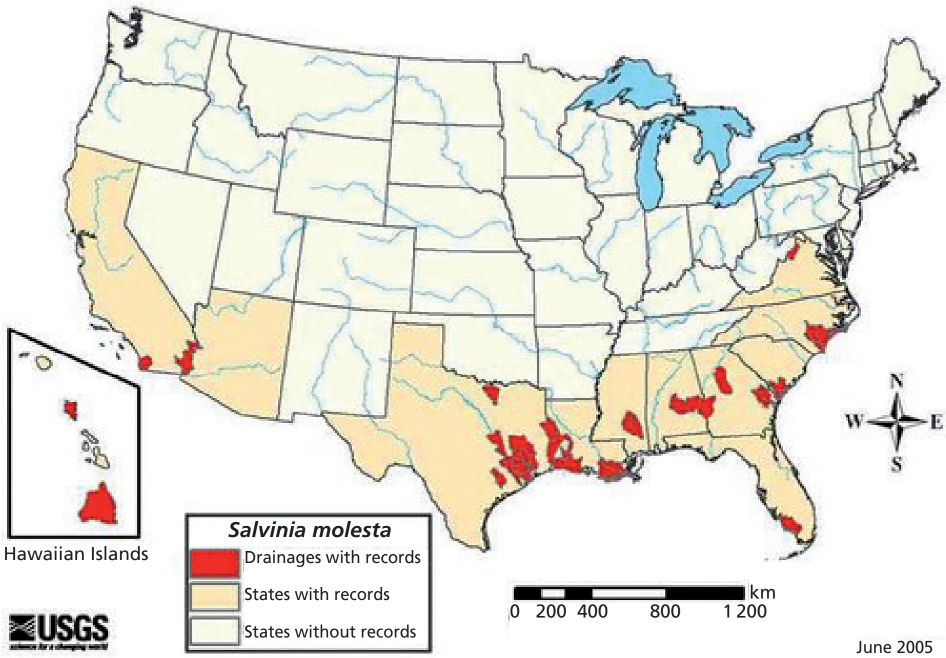
[181] This is part of the family of "modifiable areal unit" problems highlighted by Openshaw and Alvanides (1999) whereby GIS (and other software) output results can be severely compromised according to factors such as the spatial scale of study, the size of aereal units, the number and position of class boundaries, the location (placement) of polygons or cells on the map, etc. See also Reynolds (1998) available at www.badpets.net/Thesis/index.html.

[182] The box "Random/Regular" in Figure 7.16 is blank because it would be impossible to have a "Regular/Random" point distribution pattern.

Figure 7.17 shows that the occurrence of the invasive pond weed *Salvinia molesta* in the United States of America appears to be clustered mainly in the south-eastern portion of the country, but, obviously, an examination at a local level would suggest that the aquatic weed is more randomly distributed.

Most GIS are able to perform a statistical test on point distributions called the "nearest neighbour analysis", the result of which gives a numeric value giving a relative indication of the degree of regularity, randomness or clustering of the points within a specified area, plus a probability value indicating whether the pattern is statistically different from a perfectly random distribution. This analysis can be performed using either vector- or raster-based GIS functionality. A look at the distribution of fish ponds in many areas shows a fairly high degree of clustering, and this is likely to be around areas that have good soils for water retention, or where fish food sources are available or close to fish processing facilities. If marine species are found to show clustering in an area, it can be invaluable to try to ascertain what independent spatial factor is the cause of the clustering.

FIGURE 7.17
The distribution of the invasive weed *Salvinia molesta* in fish ponds in the United States of America

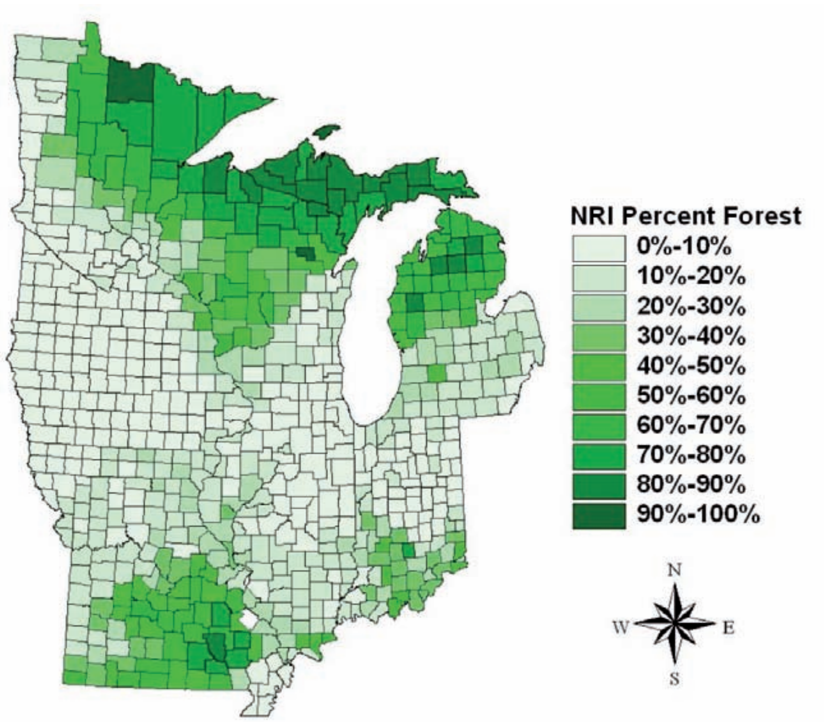


Source: Toledo-Bend Guide (2005).

Contiguity and spatial autocorrelation

In a similar way to point distributions, polygons may also be distributed in various patterns that range from uniform through random to clustered. However, with polygon distributions there is more interest in attempts to measure the dispersion of polygons in terms of their contiguity or spatial autocorrelation. Contiguity refers to whether polygons showing the same type (or classification) are more or less likely to be adjacent to one another. Consider the forest cover shown in Figure 7.18. Here, it is clear that counties having similar proportions of forest cover are frequently contiguous (next to) counties having the same proportion of forest cover. The other extreme would be the pattern seen on a chessboard where none of the black or white squares share boundaries with a similar colour. So a greater degree of contiguity means that a much more regular polygon dispersal pattern could be mapped. DeMers (2008) describes how a useful measure of contiguity is the "joint count statistic", which indicates either how many similar polygons are co-joined or the relative number of polygons that are contiguous within a defined area. Contiguity is important in the natural world because it can indicate how easy it is for plants or animals to move or migrate within favourable environments.

FIGURE 7.18
Distribution of forest cover by county in the north central United States of America



Source: Rutherford (2012).

Similar to a measure of contiguity, most GIS software can provide a measure of the similarity between neighbouring cells or polygons, i.e. by use of the Moran or Geary statistic. This measurement is called spatial autocorrelation. For example, elevation data tends to have high spatial autocorrelation, e.g. the elevations at two locations close to each other tend to be similar.

Using Figure 7.18, the question could be posed "What proportions of cells having 40–50 percent forest cover are adjacent to polygons having the same density of forest cover?" Clearly, the spatial distribution of forest density shows a high positive degree of spatial autocorrelation, i.e. polygons that are close to each other tend to have similar forest densities and, therefore, the landscape has a locally homogeneous forest distribution. When such a relationship is noticed, then it is often useful to look for the underlying causes. In the case of forest cover, it is likely that neighbouring counties tend to share similar climatic and soil conditions, which directly affect forest density, though it might also be related to the fact that the land may be unsuitable for alternative uses. As indicated earlier, extreme negative spatial autocorrelation would be exhibited by the colour pattern of squares on a chessboard. Like the nearest neighbour analysis, measurements of contiguity or spatial autocorrelation are scale dependent.

7.5.3 Measures of statistical surfaces

In this section, the concern is with measuring any of a range of so-called "statistical surfaces". A statistical surface can be thought of as any section of the earth's surface (including the seabed) that has an interval or ratio numerical value attached to it indicating a value on the z-axis (vertical). Surfaces can be mapped and come in two general forms – those relating to: (i) the human environment; and (ii) the physical environment. Measurements of the z-axis of the human environment might include average incomes per unit area, number of people per household, number of fish markets per districts and rates of fish consumption per country. Measurements of the physical environment might include height above sea level, mean air temperatures, depth to the water table and ocean depth. All of these measurements (and countless others) can be measured and mapped.

All of these examples can also be described as being either "spatially continuous" or "spatially discrete". Spatially continuous data are those where a value can be allocated to any place, and values might continuously vary over space. These data sets are usually best represented by rasters. For example, it is possible to measure the ocean temperature at any location, so an infinite number of continuous measurements could be made. Discrete data, by contrast, only occur where the object (factor) being measured is located, e.g. the postcode, the household and plant types. These types of data can be represented by either raster or vector formats, but the vector format is used more often because it precisely defines the area in question.

These statistical surfaces showing the distribution of any feature can be relatively smooth or very rugged. For instance, a surface showing air temperatures or ocean water salinity is likely to be very smooth because these factors only change very gradually through any spatial area. By contrast, a mapped surface showing the distribution of a single fish species might be very rugged, with surface values varying significantly over short distances. Although this statistical surface data can be used in a range of GIS analyses, two main kinds of analysis are specifically examined: interpolation and trend surface analysis.

Interpolation

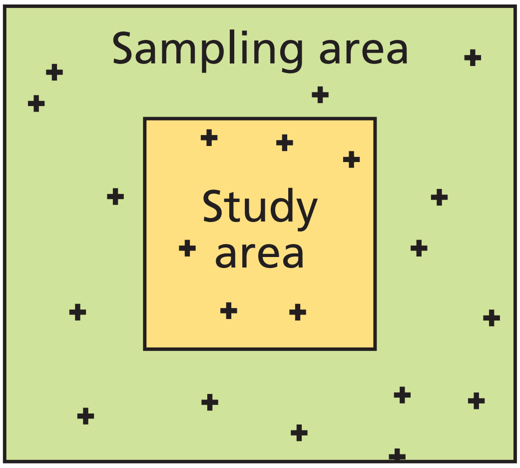
Interpolation is the process of estimating values of properties at unsampled locations based on known values in surrounding areas (existing sampled observations). Because of the usefulness of estimating missing data values, it is important that GIS has the tools to estimate these.

Interpolation is based on the fact that it is possible to predict or estimate missing values inside a set of numbers by using numbers on either side of the gap. Interpolation can be applied to spatially continuous surfaces and it generally relies on spatial autocorrelation, i.e. the idea that things closer together are more similar than things further apart. For example, to make a temperature map it is not likely that there will be sample points (observations from weather stations) at regularly spaced intervals, but it is likely that nearby temperature readings will be very similar to each other. A good sample set for interpolating a surface should have enough points to represent the important details of the surface. The sample set should also be able to represent changes in the surface as well as the dips and troughs of the surface. There are several different interpolation methods, or, to be more precise, interpolation algorithms that the software implements to create these estimations.

Interpolation may be described as being either "global" or "local". Global interpolation methods use a single mathematical function for every available sampled point in a region to create the estimation of any cell or point value it is looking for. This type of interpolation is typically used if values are fairly consistent over the area being studied. However, when creating, for instance, an elevation surface (e.g. a contour map) using point sample data, there may be features in the real world that could skew the estimation of the surface. Thus, the presence of a large canyon a short distance away could unduly influence any sample values generated if it is taken into account. To avoid this problem, the "local" interpolation method uses only a local sample of the available known points to complete the estimation. Therefore, the use of a local method that excludes distant irrelevant values produces an interpolation result that more closely resembles the real world.

It is the responsibility of the user to understand the nature of the data being collected to know whether or not a global or local method would produce the most accurate results. It is also important to note that the samples used to interpolate any surface should extend well beyond the area to be estimated. Having the sample data extending outside of the study area decreases what is known as the "edge effect", i.e. without sample data outside the study area the values estimated along the edge of the study area are likely to be skewed. Figure 7.19 shows an example of properly collected samples in relation to the study area to be interpolated.

FIGURE 7.19
Correct sample point strategy when using interpolation methods within a defined study area



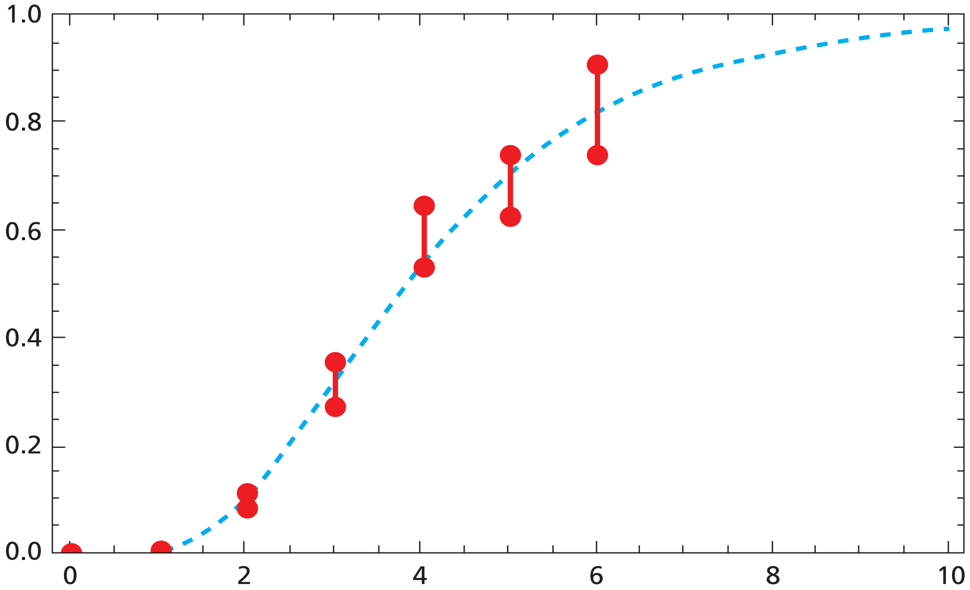
Another way of classifying interpolation methods is by whether they use linear or non-linear techniques. Linear methods of interpolation assume that values across a surface change in a regular arithmetic progression, e.g. 10, 20, 30, ?, 50, 60…etc., and it is easy to interpolate the missing value here as "40". Values across a surface rarely change in an exact form, but linear interpolation methods might be the safest method to use when there is any degree of uncertainty.

Other distributions may change in a logarithmic or exponential form (Figure 7.20). From this figure, distributions can be interpolated from the best-fit line that for a value of 4.5 on the "x" axis, the value on the "y" axis would be 0.6. A final interpolation method illustrated here is that known as inverse distance weighting. This method assumes that those known sampling points that are located nearest to a missing data point will probably have more influence in determining any missing value than known points located at a distance. So, in calculating a missing value, the near points are allocated a weighting that is higher than those that are further away. An example of the use of inverse distance weighting interpolation is shown in Figure 7.21, which shows estimates of commercial values of all groundfish species for all locations in the New York Bight, United States of America. Here, the interpolation was based on the known values of groundfish caught at each haul site (blue dots) and from these values weighted interpolated estimates could be obtained for the whole exploited groundfish marine area.

There are a number of other more specialized interpolation methods that rely on complex mathematical procedures that go beyond the scope of this technical paper, for example, kriging and the drawing of splines. Some of these are illustrated and discussed at Spatial Analysis Online (www.spatialanalysisonline.com/output). Most

FIGURE 7.20

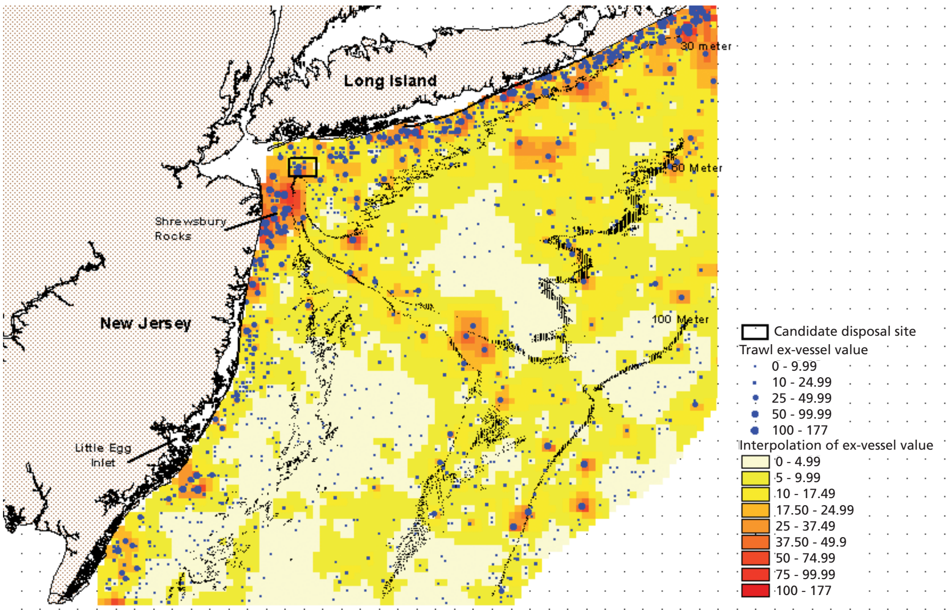
Interpolation based on a sample appearing to show an exponential distribution



Source: Stack Overflow (2012).

FIGURE 7.21

An inverse distance-weighted interpolation of the commercial values for groundfish in the New York Bight, United States of America



Source: Nuckols III (2012).

GIS packages will have several interpolation methods available for use with point, line or area data. For more information on interpolation, see Dubois (1998), Longley *et al.* (2005a), de Smith, Goodchild and Longley (2007) and DeMers (2008). For more advanced texts on different types of kriging, see Chilès and Delfiner (1999) and Cressie (1993).

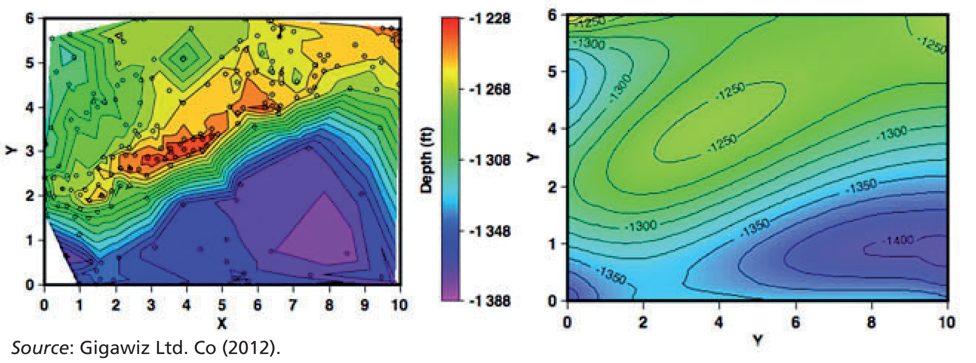
Trend surface analysis

Trend surface analysis is a frequently used global form of interpolation. The term "global" in this case does not mean that it covers the entire globe or is used all over the world, but rather that the entire landscape being analysed is described by a single equation or polynomial model.[183] This is in contrast to "local" interpolation techniques (such as inverse distance weighted, splines and kriging, mentioned above), which interpolate the surface at any location based on data points in the nearby neighbourhood. Local methods will use a different set of data to interpolate any point in the study area, while global methods use all the data to generate a single equation that describes the entire study area.

The trend surface method creates maps that may be used to identify general patterns of regional trends without being skewed by localized anomalies that often have no overall pattern. A way to envisage the trend surface is to think of it as the three dimensional equivalent of the conventional "best-fit" line, which may be drawn through a two dimensional set of data points. However, instead of a single line being drawn, a whole "surface" is drawn based on a polynomial model.

For example, suppose that detailed quantitative data are sampled allowing for the mapping of a fish species throughout a regional sea. For mobile species such as fish, each individual data sample is likely to vary, perhaps quite significantly, between neighbouring samples. When mapped, the data might look highly irregular and the map might mask underlying trends. However, through the use of trend surface analysis, it is possible to obtain a generalized map for the regional sea showing a broader trend of the species distribution. For most species, this would show a high fish distribution in areas where habitat conditions were favourable and the map might show declining fish quantities with distance from the highest count area(s). It is frequently useful to compare the trend surface map with the original data sampling map so as to, for instance, compare the differences in recorded fish density at any x, y location on the maps, or to identify areas where fish densities are seen to be particularly high, i.e. identifying anomalous fish numbers. Because fish are constantly on the move, it is quite likely that a trend surface map of fish distributions is more reliable than any actual sampling made on a particular day. Figure 7.22 shows an application of trend surface analysis to sea-water depths. The left-hand map illustrates the seabed depths calculated from the numerous sampling points shown. The right-hand map shows the overall trend of sea depths in the same area. Trend surface mapping can easily be accomplished by most GIS. Details of how it is undertaken are not discussed, but further information can be found in Lo and Yeung (2002) and Heywood, Cornelius and Carver (2006).

FIGURE 7.22
A trend surface to show the general trends of sampled water depths



Source: Gigawiz Ltd. Co (2012).

---

183 For details on polynomial model, including examples of alternatives to polynomial trend analysis, please refer to Chapter 13, Section 2 of Legendre and Legendre (1998).

7.6 NETWORK AND SPATIAL INTERACTION ANALYSES

In this section, attention turns away from an examination of phenomena that are distributed over continuous surfaces, concentrating instead on potential GIS analyses that focus on linear patterns in the landscape. These analyses are frequently described as network analyses, though this section also includes other considerations of spatial allocation within networks, such as gravity modelling, which is not network analysis per se but which is usually dependent on networks. Section 5.8.4 explained in some detail how networks are modelled for GIS purposes, and it also pointed out that networks could represent a wide array of routes or pathways. Figure 5.18 illustrated a typical network structure consisting of links and nodes[184], information on these being retained in the database management system of a GIS through the use of topological tables. In this section, spatial interaction via gravity models, routes and pathways via forms of connectivity analyses, spatial allocation and service areas, plus some analyses of river and stream networks are described. Users of network analysis should be aware that not all GIS packages have the tools necessary to work with networks. For example, Environmental Systems Research Institute's (ESRI) ArcGIS has a geometric network data model, but it is primarily for use by utility companies, and ESRI's ArcHydro data model is designed for hydrologic modelling (mainly river networks).

7.6.1 Gravity models

In a topological network of links and nodes, it might be assumed that each node is of equal value or importance. While for some networks this might be true, in the majority of cases this is far from the real world situation. If the case of towns or cities is considered, then it is clear that larger urban centres are likely to provide more shops, jobs, markets and other facilities. Similarly, larger lakes are likely to attract more wildfowl or fishers. It is easy to envisage that a hierarchy can be produced that assigns a value to town nodes or to lake nodes according to their size or to their "attraction capacity". The value may be in terms of population, or number of shops or number of boat-mooring places, etc., and the higher the value assigned, the greater will be the "gravitational pull" of any node.[185] Effectively, this means that people are usually prepared to travel further to get to a larger "node" because that node is able to offer a greater range of services or other attractions.

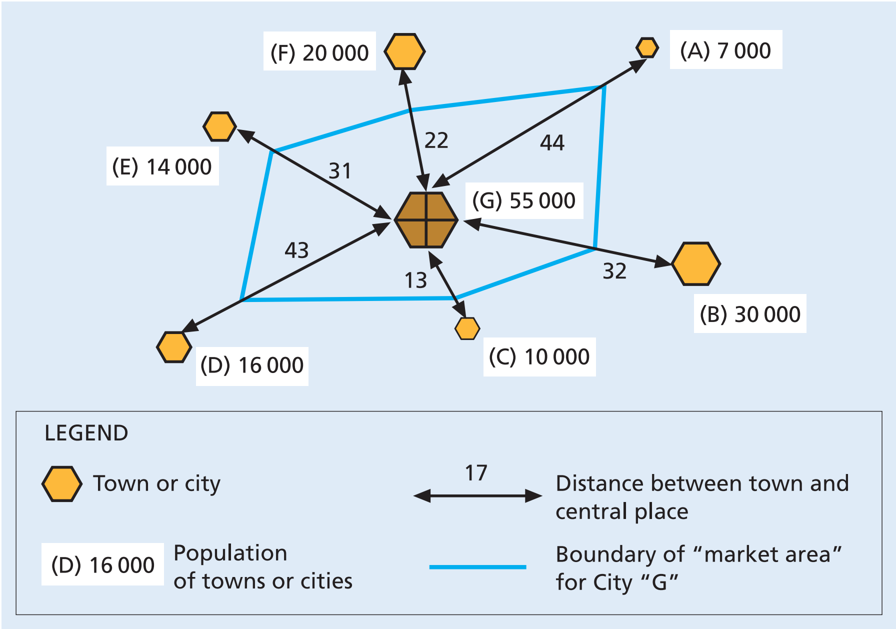
Figure 7.23 shows a large central place (G) with a population of 55 000. It is surrounded by various towns (A to F) that have populations varying from 7 000 to 30 000 and that vary in distance from the central place between 13 and 44 kilometres. Most GIS contain algorithms that can calculate the likely number of people who will utilize the facilities in the larger central place and the likely break-even distance threshold between each town and the central place. By linking the six distance thresholds, it is possible to derive a market or service area that surrounds the central place. Calculations to establish the market area can be simply based on distance, but there are many other calculation methods. For instance, it could be based on travel times,travel costs, the number of shops in each urban area or on the specific services offered. From a fisheries or aquaculture perspective, this form of analysis might be useful in detecting the optimum location for fish processing plants, i.e. relative to the disposition of aquaculture production, or maps such as Figure 7.21 could compare values of groundfish at different ports relative to the distance necessary to travel to fishing grounds. Clearly, what has been described here is a simplification of the real world, because in reality there might be many other factors that influence the interactions between centres.

---

[184] Links and nodes are described in section 5.8.4.

[185] Gravity models utilize Newton's law of gravity, which states that the force of attraction of two bodies is proportional to the product of their masses but inversely proportional to the squared distance separating them.

FIGURE 7.23
Likely market area for city "G" based on calculations using a typical gravity model within a GIS



7.6.2 Optimum routes and pathways

There are a number of different types of network analyses that can be performed relating to optimizing routes along networks or pathways, and these are sometimes described under the general heading of "connectivity analyses". The main types of network analyses are described as follows.

Shortest path analysis

Recall from Section 5.8.4 that routes or pathways can comprise not only transport routes such as roads, rail, air and tram, but also pipelines, cables, sewers, electricity grids, etc. Shortest path analysis aims to find the quickest or shortest route between two nodes (or which have the least impedance, as explained in Section 5.8.4). This is the type of analysis used when consulting Web-based travel planners or when getting directions from an in-vehicle navigation system. The analysis, which is based on a network such as Figure 5.18, uses an algorithm that finds the best route along a network from an origin to a destination (Fu and Rilett, 1998). The algorithm looks at the topological table associated with the network that details any impedance that might be found between two nodes, and the software looks at all possible routes between nodes to determine which has the lowest aggregate impedance values. For instance, using Figure 5.18, it is possible to easily calculate that to get from node "B" to "F" there are three apparently logical route choices, but each would have total impedance values as follows:

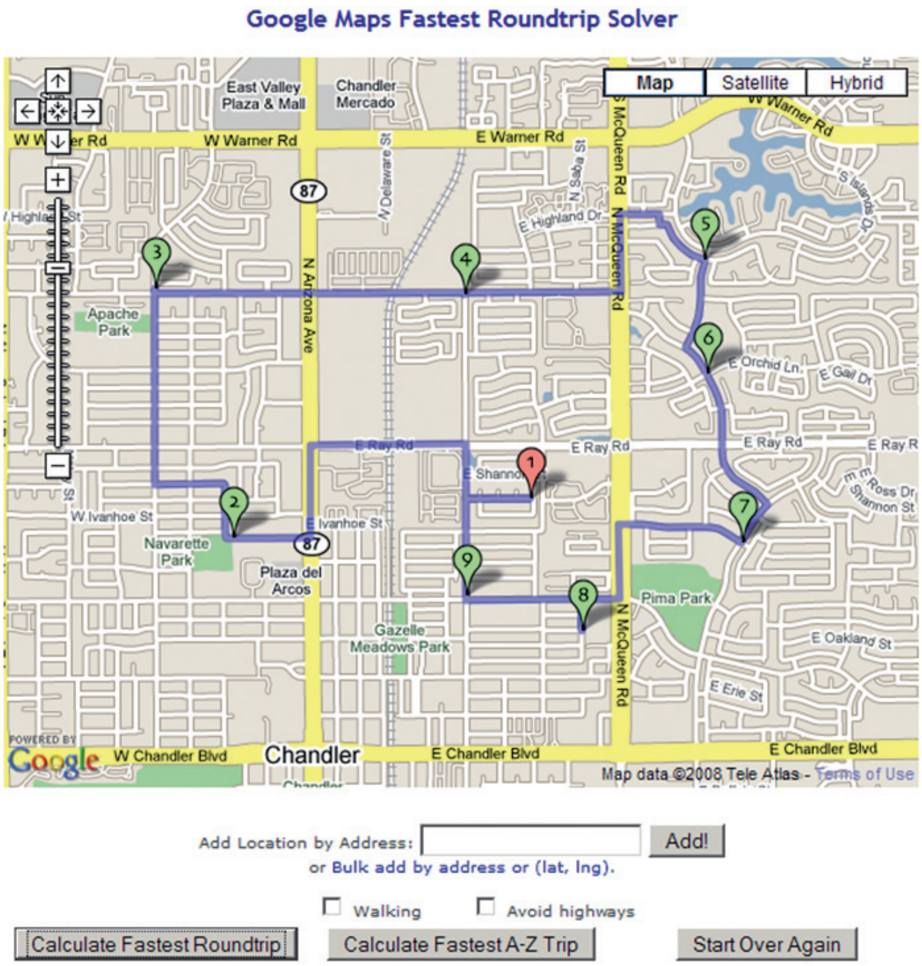
(i) Route B, A, C, G, F = impedance values 4 + 5 + 3 + 4 = 16.
(ii) Route B, D, C, G, F = impedance values 3 + 2 + 3 + 4 = 12.
(iii) Route B, D, E, F = impedance values 3 + 7 + 5 = 15.

From these calculations, it can be seen that the second choice above would seem to offer the best option because it has the lowest total impedance value, in this case aggregate travel time. However, impedance can take many forms, so while option (ii) might be preferable in terms of the lowest cumulative impedance, if the impedance were changed to reflect public transport costs or actual distance travelled, then this route might not be the cheapest, shortest or the quickest.

The "travelling salesman problem"

Closely linked to shortest path analysis is the so-called "travelling salesman problem". For example, numerous delivery companies are required to make x number of deliveries per day, delivering to centres (nodes) that may be widely scattered. Most GIS network analysis programs have the capability of working out the optimum delivery route taking into account any desired set of impedance values. From a fisheries perspective, it might be extremely useful to establish the optimum daily route for the collection of landed fish catches. There is now an interactive Web site (http://gebweb.net/optimap) where users can utilize Optimap, which is linked to Google Maps, in order to calculate any optimum round-trip problem involving visits to a maximum of 24 locations. Figure 7.24 gives an example of the output from Optimap where, from a start location at "1", nine sites have been visited using the shortest possible routeway.

FIGURE 7.24
Illustration of the travelling salesman problem solved by the use of Google Maps in Optimap



Source: Free Geography Tools (2009).

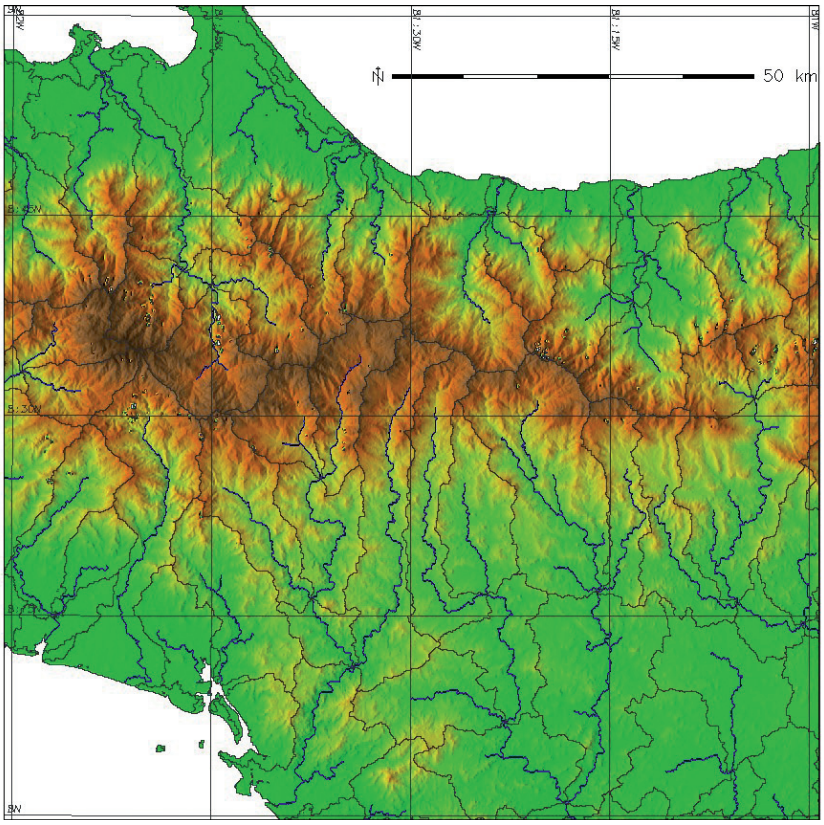
Drive time analysis or service area definition

It is frequently useful to map the region that is within a certain drive time of some central location. This type of analysis is commonly done to determine the most efficient placement of such locations as commercial businesses or emergency response services. Hospitals can use this information to determine the likely number of patients residing within their service areas. If fish farmers make direct deliveries to customers, they could calculate their potential customer base by identifying the region (service area) that is within a 60-minute travel time from their farms. From any given point on the network, a GIS can calculate the likely travel distance achievable in a given unit of time, or alternatively, it can spatially distribute a network of service centres in an efficient manner.

7.6.3 River and stream networks

Figure 7.25 provides a vector cartographic illustration of an oriented network, in this case the stream networks in a portion of the central part of the Republic of Panama. Here the landscape is dominated by a dense network of streams that generally radiate outwards from a core highland area. Given this stream network, and working in either raster or vector formats, a wide range of different modelling investigations can be accomplished, as is shown in Box 7.5. All the modelling described can be directly relevant to both inland fisheries and to aquaculture site location considerations. For instance, with respect to detecting flood storage areas, it can be envisaged that a similar modelling process can be used to determine optimum sites for the building of dams or reservoirs. An increasing proportion of the latter are likely to be utilized for fisheries or aquaculture, especially if water quality can be maintained.

FIGURE 7.25
The stream network in a portion of the central part of Panama



Source: Mitasova (2006).

BOX 7.5
Examples of directed (oriented) network modelling relating to rivers or streams

The following network modelling examples can be mostly accomplished using either vector or raster data structuring.

- **Calculating flow direction and water accumulation –** Rasters containing elevation values (i.e. digital elevation models, or DEMs) enable the GIS to calculate water flow direction. By using known rates for a particular rainfall event and geographic area, then it is possible to calculate approximate water concentrations along stream stretches or at particular points on the network.

- **Deriving stream catchments –** Raster DEMs allow river water catchments to be delineated, i.e. as connected high points between basins.

- **Detecting flood storage areas –** Using both triangulated irregular networks (TINs) and stream network information, areas with the potential for holding excess flood waters can be delineated. With increasing urbanization along rivers, this is more frequently a problem that needs addressing

- **Probability of flooding –** Again, with increasing population pressures and changed rainfall patterns, there is the need to identify which areas along a river are likely to be flooded. This is known as inundation modelling.

- **Rate of pollution dispersal –** If known quantities of point-source pollutants enter a stream, their trajectory and likely dispersal rates can be calculated. Similar modelling can be done relating to sediment dispersal.

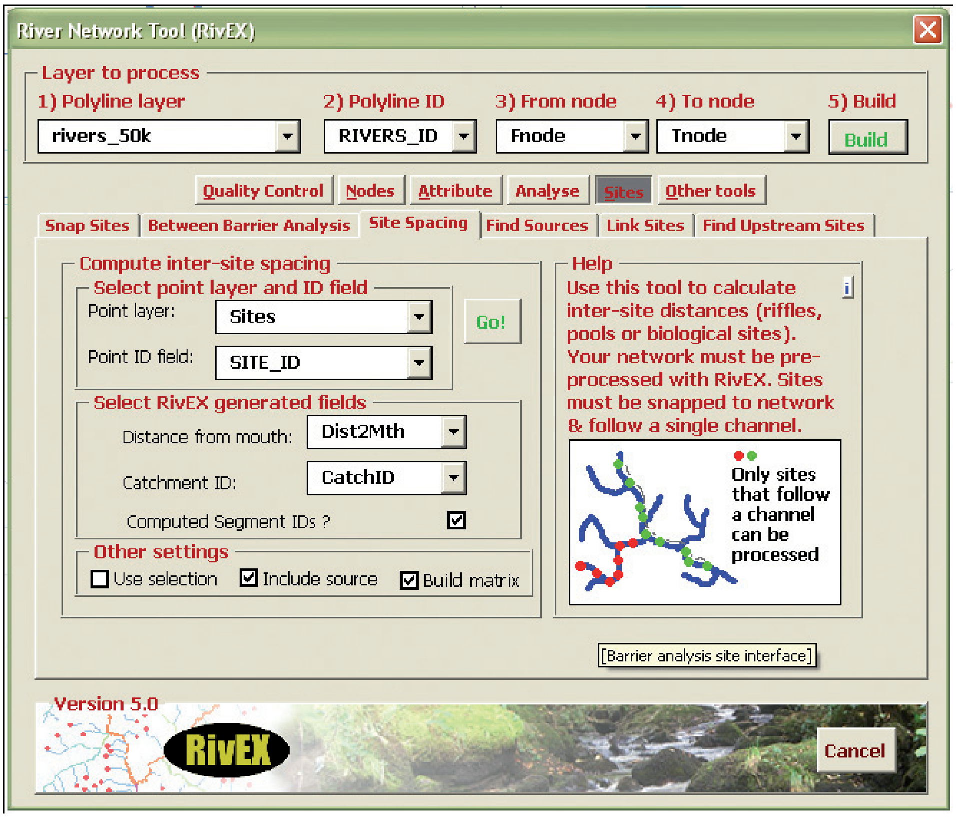
- **Managing water extraction –** Knowledge of river networks, water flow rates and water demand allows for the GIS-based management of water extraction.

- **Creating stream networks –** If digital data on stream networks are not available, then TINs or DEMs can be used to create the likely stream network.

- **Peak flow prediction –** Given a storm event over a part of a river catchment, then the peak flow in a stream can be modelled in terms of time and location.

Most proprietary GIS packages available today can perform a range of network analyses, but if network analyses are going to be important for future anticipated GIS work, the user should check the exact functionality before investing in the software. For example, ESRI's ArcGIS includes specialized tools designed specifically for hydrologic networking. There is also a wide range of separate network analysis software available online that can be linked to GIS. See Spatial Analysis Online (www.spatialanalysisonline.com) and Spatial Hydrology (www.spatialhydrology.com/software_hydrostat.html) for two gateways to hydrologic tools listing approximately 100 software packages for modelling hydrology. Examples of low cost or free GIS-based river analysis software capability are: www.rivertools.com, http://grass.itc.it and www.rivex.co.uk. Figure 7.26 provides a screen shot from RivEX Version 5.0 showing one of the many capabilities that the software provides. For more information on network modelling, see Lo and Yeung (2002), Vieux (2004) and de Smith, Goodchild and Longley (2007).

FIGURE 7.26
Screen shot of the RivEX GIS software for modelling stream networks



Source: RivEX (2011).

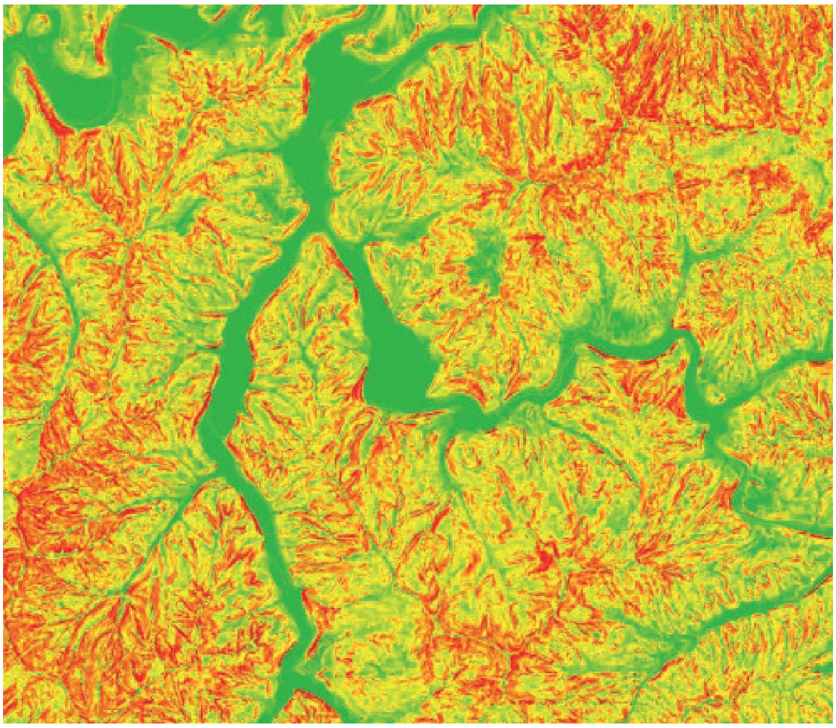
7.7 TOPOGRAPHIC SURFACE ANALYSES

Topographic surface analysis is concerned with describing and accounting for various attributes associated with the shape of the earth's surface. To perform topographic analyses, various types of data must be collected and stored, and these typically include point elevations, angle of slope, aspect of slope, curvature of the landscape and hillshades, through the use of triangulated irregular networks (TINs) or digital elevation modelling (DEMs). For more information on the TINs and DEMs that form the basis of topographic surface analyses, see Section 5.8, which also gives the range of analyses that can be performed by GIS. Here, the range of analyses are classified under: (i) gradient (slope) and aspect; (ii) visibility (viewshed) analyses; and (iii) watershed and river flow analyses. This section does not consider contour derivation (interpolation) as this has been reviewed in Section 7.5.3.

7.7.1 Gradient (slope) and aspect analyses

Gradient (or slope) is a measure of the rate of change in elevation between points in the terrain (O'Sullivan and Unwin, 2003). This gives the user a percent of elevation change between the points, or the actual angle of the slope, or the ratio of the amount of vertical rise compared with the amount of the horizontal distance between the two points, e.g. a slope could be 1 unit in 10 units. It is also possible for the GIS to "calculate" the slope in terms of its length. Different GIS packages use different algorithms to make their calculations of slope. This type of analysis could be of significant use in seeking areas suitable for aquaculture pond locations, and this is well illustrated in Figure 7.27 where only areas of green shading are likely to be of use for pond locations.

FIGURE 7.27
Illustration of GIS output using slope analysis in an area near Harlan, Kentucky, United States of America

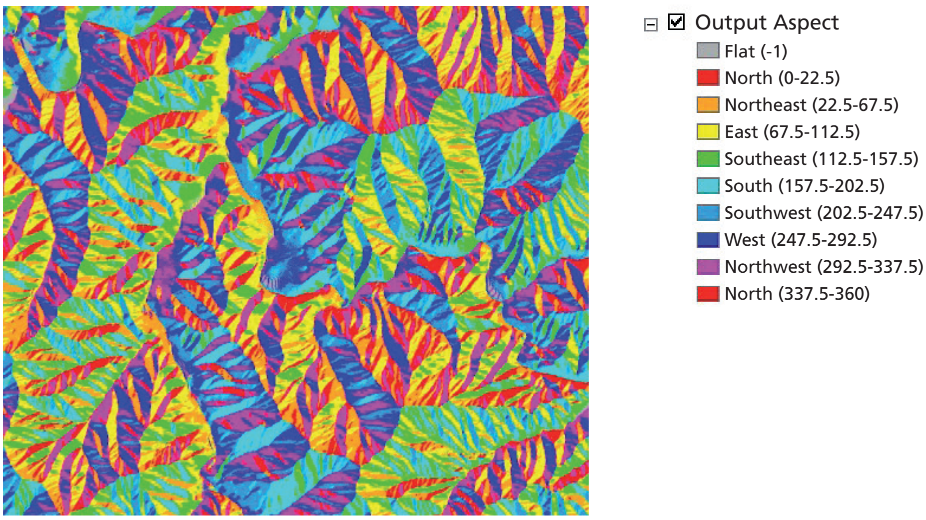


Note: Green areas are flat, yellow are intermediate and red are steep.

Source: Nelstead (2009).

Aspect measures the direction that the slope is facing – also called the slope's orientation. Direction is represented in degrees starting from 0, which represents north, and proceeding clockwise around to 360 degrees. In Figure 7.28, which illustrates aspect, the legend displays both a colour showing the direction which the slope is facing, as well as the compass degrees from north.

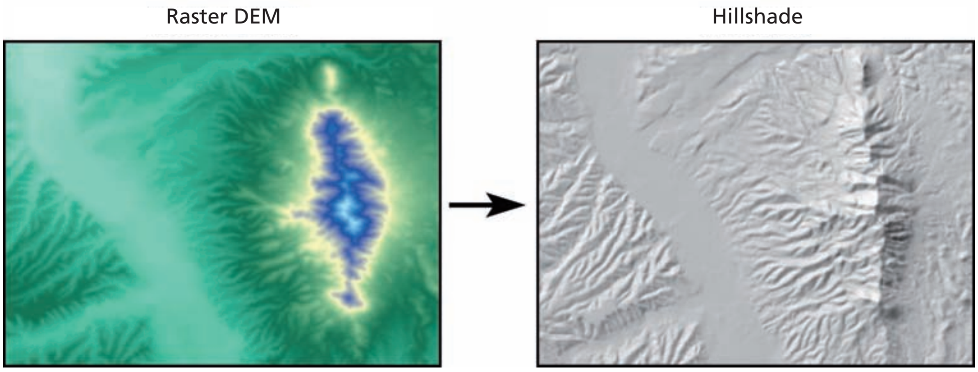
FIGURE 7.28
Illustrating aspect for the same area shown in Figure 7.27



Source: Nelstead (2009).

A variant of aspect representation is that of hillshading (Figure 7.29). Here, the GIS assumes that the sun is in a designated position in the sky (in terms of direction and angle of inclination), and GIS output is obtained showing the areas of the land that would then be in shade or in sunlight. It is clear from the "Hillshade" output map that the sun is shining from a north-west direction – hill slopes facing south-east are the most heavily shaded. Aspect can be extremely important in terms of the amount of solar radiation, wind or rainfall received and the consequent effects on microclimates, habitats, length of the growing season, and so on.

FIGURE 7.29
Hillshade obtained from a raster digital elevation modelling (DEM)



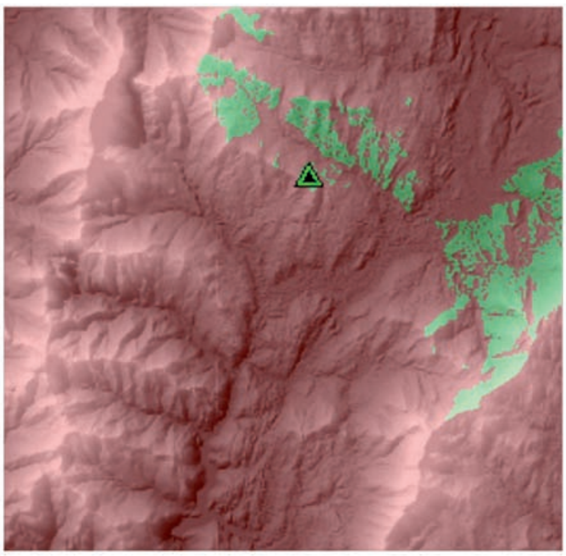
Source: Jenness (2011).

7.7.2 Visibility (viewshed) analysis

Visibility analysis allows the user to determine what is visible from certain locations based on surrounding elevations, such as hills, valleys and mountains. The software does a so-called "line-of-sight" analysis from an observation point to every cell in a raster layer in order to calculate what cells would be visible or not visible from that point. This binary data set makes it very easy to visualize the results of the analysis. Thus, Figure 7.30 shows in green all the areas of a landscape that would be visible from the location indicated by the triangle. Depending on the size of the data set and the number of observation points, a visibility analysis can be time and computer intensive.

Visibility analysis may be very useful for aquaculture site selection. When siting a new aquaculture facility, government agencies may require that the visual impact is kept to a minimum. To do this, the farmer can create a viewshed analysis from various observation points near the proposed site to visualize the potential impact the facility might have on the surrounding area.

FIGURE 7.30
Visibility analysis showing areas (in green) visible from the located triangle



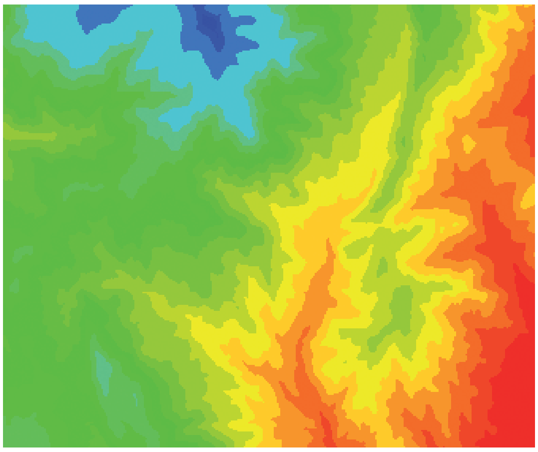
Source: Neteler and Mitasova (2007).

7.7.3 Watershed and river flow analysis

Some modelling relative to water flow has already been described in Box 7.5, but here the basis on which the GIS performs its analyses is described. GIS may be used to examine the characteristics and likely stream flow in river catchments. For workers in inland fisheries or aquaculture, this could be an important area for spatial analysis. The main types of GIS work that can be accomplished include:

- defining water catchment areas (watersheds or river basins);
- determining catchment boundaries;
- determining subcatchments and so-called "pour-points"[186];
- establishing stream networks (or stream ordering);
- calculating likely water flow directions;
- defining likely water flow quantities (or water accumulation);
- calculating stream lengths.

FIGURE 7.31
Raster elevation data in a hypothetical area

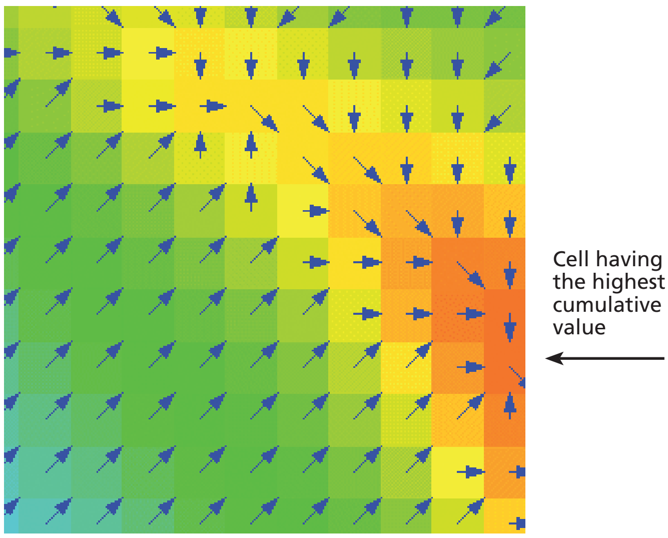


Source: Nelson, Jones and Smemoe (1997).

Examining watersheds mostly involves the use of elevation raster data to determine where water would flow and accumulate given the topography. The broad principles upon which water flow and analysis works are described using Figures 7.31 to 7.33. Figure 7.31 shows a raster DEM with low areas coloured red and higher elevations coloured progressively with yellows, greens and finally to blues at the highest areas.

Given that every pixel will have its own value, and that water will always flow downhill using the steepest gradient, the GIS is able to examine the pixel value for each pixel plus its eight surrounding pixels to determine where rainfall water (or snowmelt) is likely to flow, i.e. the surrounding pixel having the lowest elevation value. Figure 7.32 shows predicted water flow directions for each pixel in an enlarged sub-area of Figure 7.31. Thus, it is clear that water is flowing generally from blue areas towards the red colour pixels. The pixel values allocated to any cell in Figures 7.32 may now relate to the cumulative number of cells through which water has flowed to get to that cell. So, the cell indicated as having the highest value has accumulated (received) water from 113 cells is shown on this map extract.

FIGURE 7.32
Predicted water flow direction across a section of the hypothetical area shown in Figure 7.29

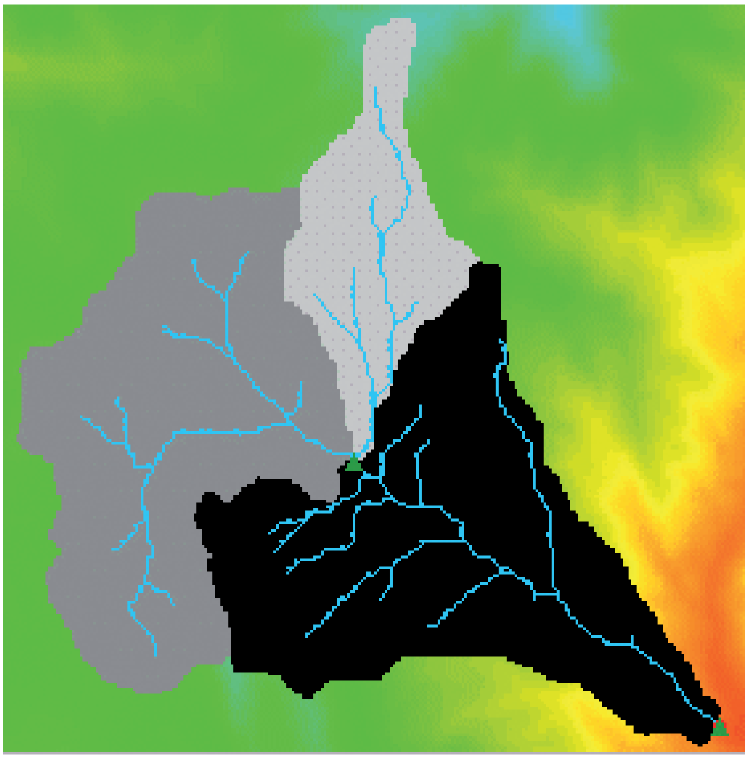


Source: Nelson, Jones and Smemoe (1997).

[186] A "pour-point" is the point where one subcatchment meets the next subcatchment. This will be at the lowest point in the subcatchment.

Where relevant, each pixel will accumulate water from neighbouring higher value cells. Once cells have accumulated water from more than x number of cells (a threshold value), then the GIS can be programmed to start drawing in the likely path of the river.[187] Figure 7.33 shows (in light blue) the likely river network in the main part of the same area as covered by Figure 7.31. It also shows that this particular river catchment can be divided into three main subcatchments, with "pour-points" occurring at the sites of the olive green triangles. At the "pour-point" in the lower right-hand corner, the river depicted clearly merges with a larger river that is not shown but which would be flowing down the valley (red area) sited near the lower right-hand side of the map. The main catchment boundary is also clearly shown. It can be seen that the GIS has produced information that potentially allows any of the bullets listed at the start of this Section (7.7.3) to be evaluated and discussed. The reader may appreciate that this account simply outlines the basics of water flow and river catchment analysis. For further information, see Nelson, Jones and Smemoe (1997), Maidment (2002), Chang (2004) and DeMers (2008). For examples of watersheds analysis as applied to inland fisheries and aquaculture, see the African Water Resource Database (AWRD) by Jenness *et al.* (2007a,b).[188]

FIGURE 7.33
The main river catchment and subcatchments identified by the GIS for the area shown in Figure 7.29



Source: Nelson, Jones and Smemoe (1997).

[187] The value of x will depend upon what size river the GIS operator wishes to consider.

[188] The African Water Resource Database includes some 5.5 GB of data plus a set of data and custom-designed GIS-based tools covering many aspects of inland fisheries and aquaculture. Concepts, application case studies and a technical manual and workbook are also provided.

7.8 CUSTOMIZATION AND SCRIPTING

Even the best GIS will not have every tool and function that the user could want. Most users have questions that cannot be answered simply with the tools at hand and may require a long series of functions strung together. Some questions require complex statistical analysis that simply cannot be answered at all with the available tools.

A good GIS will provide a means for users to create their own functions and tools. This may be a simple macro function. Macros are collections of frequently used commands that are grouped together in a file. Instead of entering commands individually, users can call the macro and the commands stored therein will be executed automatically, and then the user is able to run several tools in sequence without having to set up each tool individually each time. Macro functions are especially useful for repetitive functions, or functions that have to be done on a regular basis. For example, a common function in both terrestrial and benthic habitat analysis is to categorize the landscape by slope position (i.e. find the ridgetops, midslope regions, valley bottoms and flat areas). This process uses a statistic called the Bathymetric Position Index in marine environments and Topographic Position Index in terrestrial environments (Weiss, 2001), and it involves several steps including calculating slopes from a DEM or bathymetric data set, a neighbourhood analysis, map algebra and multiple reclassifications based on multiple data sets. The steps are relatively intuitive and straightforward, but repetitive and tedious and prone to error if done manually. It is much easier to run this type of analysis in a macro.

The most advanced GIS software will allow the user to write sophisticated tools using the most current software programming languages, such as C++, C#, Java, Python or VB.NET. Such tools give users the ability to do any type of analysis on the data, limited only by their imagination and ability to write the algorithms. Many high-end GIS packages have enthusiastic user groups of people who write and share custom tools, and often provide them free online. With experience in GIS it seems that most functions are possible.

An example of a script for inland fisheries is one developed by de Graaf *et al.* (2003) to conduct a linear regression analysis between the number of fishers and the number of gillnets in villages around a lake. Another example is the African Water Resource Database (AWRD) (Jenness *et al.*, 2007a,b). The AWRD includes an extensive collection of custom-made tools[189] to analyse watersheds, surface waterbodies and aquatic species, and includes many additional tools to assist with general statistical analysis, cartography and creating and editing metadata.

[189] Scripts and macros generally just run an analysis. The AWRD tools include many other internal aspects, including graphical user interface (GUI) windows, buttons, graphics and user interface parameters, that make them a lot more sophisticated and user-friendly than simple scripts. The technical name for the functions provided in the AWRD are "tools", "command buttons" and "dialogs", and these call a number of background scripts to gather analysis parameters and to run the analysis.