1 Introduction to GIS

Introduction

GIS (Geographic Information System) has great value in our time as it is a comprehensive information system evolved and still developing parallel with the advancing technology. This era of the human kind is characterized as an information age, where the whole world is experiencing and interacting with a new revolution that change our traditional way to look at the things and do a business in completely different approach. The emphasis is on the technology and its use in every activity that range from agriculture, industry, business, social, research and education. The advancement in technology changed our world and our approaches to meet our need to rely completely on the technology and data. The value of information in our time becomes vital and important for development. GIS itself is an important module of the information system. The economy of all industrial countries and many other nations all over the globe, become more dependent on services. This means that the current economy rely more and more on computers, networking, accurate information and data. This shift required a mass of skilled labors that are capable to deal with the technology and data processing.

GIS technology is not an exception when it comes to its use in water resources, geology, and environmental related problems. It is a powerful tool for developing solutions for many applications ranging from creating a color coded geological map, interpolating the water quality of groundwater aquifer to managing water resources on a local or regional scale.

Water is the most precious and valuable resource and vital for socioeconomic growth and sustainability of the environment. In some arid countries, the water resources are limited, scarce, and mainly sourced from groundwater. In some Middle Eastern countries surface water is limited to few river systems and intermittent streams that are associated with rain during the winter time. Precipitation is vital and the primary source of recharge for various groundwater aquifers in these regions.

Groundwater in the region has been utilized through wells tapping various water-bearing formations to provide more fresh water to supply the increased demand water supply and irrigation. This practice negatively affected the whole hydrogeological setting of the basins. For example, total water withdrawal in the region (Israel, Jordan, and Palestinian territories) in 1994 was about 3050 million cubic meters. The estimated total renewable water supply that is practically available in the region is about 2400 million cubic meters per year. The water deficit is being pumped from the aquifers without being replenished. This practice caused the groundwater level to decline dramatically in some well fields, up to 20 m, which caused some major springs in 1990 to cease completely in Azraq basin, Jordan.

Therefore, management of water resources has become a major effort for governments in the region. Various ministries, water institution, and private companies worldwide are using the GIS as a tool to manage water resources in their countries. The GIS can be used to capture data and developing hydrologic dataset for all components of water resources. Understand the region's hydrology, map sources of contaminations, prepare water quality and water-rock interactions maps, delineate the watershed areas, and much more.

Electronic Supplementary Material: The online version of this chapter (https://doi.org/10.1007/978-3-319-61158-7_1) contains supplementary material, which is available to authorized users.

W. Bajjali, *ArcGIS for Environmental and Water Issues*, Springer Textbooks in Earth Sciences, Geography and Environment,
https://doi.org/10.1007/978-3-319-61158-7_1

What Is GIS?

GIS is an information technology system that stores all the digital data in one location, retrieving and analyzing the data quickly and efficiently in a color coded map format. The map creation depends on database and information. Users can use various types of databases to create different layers in one display. This will add an advantage to the data analysis by providing the spatial relationship between layers and will reveal any hidden relationship that the user can use for facilitating the analysis and strengthen the finding. For example, soil layers can be viewed with elevation layers and water table depth of an aquifer, in order to select the best location for building a landfill.

In previous years, all maps were created manually. A cartographer would use simple tools and paper to draw a map. This work was tedious and time consuming and if an error took place, the cartographers had to repeat all of their work and start over. This map was static in a way that any new development couldn't be added.

With the advancement in technology, a GIS map is more dynamic, can be modified in very little time, and can be stored, displayed, and printed out quickly and efficiently. GIS is a new methodology in science and applications, it is a new profession and a new business.

GIS refers to three integrated parts.

1. Geographic: The geographical location of the real world (coordinate system)
2. Information: The database
3. Systems: The hardware and software

GIS Description

A GIS is a computer-based tool that helps us visualize information with patterns and relationships that aren't otherwise apparent. The ability to ask complex questions about data and analyze many features at once, then instantly see the results on a map is what makes GIS a powerful tool for creating information. GIS can be used in many disciplines such as resource management, criminology, urban planning, marketing, transportation, etc. Primarily GIS is used for scientific analysis but is now being implemented in other disciplines.

What Can a GIS Do?

A GIS performs six fundamental operations that make it a useful tool for finding solutions to real-world problems. Throughout this course, you will gain experience with the ArcGIS tools used for these operations.

1. **Capture data:** You can add data from many sources to a GIS, and you can also create your own data from scratch. You will learn about getting data into a GIS in Chap. 4.
2. **Store data:** You can store and manage information about the real world in ways that makes sense for your application. You will learn about organizing data in Chap. 3, 4, and 7.
3. **Query data:** You can ask complex questions about features based on their attributes or their location and get quick results. You will gain experience with querying in Chap. 7, 8, 12, and 16.
4. **Analyze data:** You can integrate multiple datasets to find features that meet specific criteria and create information useful for problem solving. You will perform analysis in Chap. 2, 4, 5, 6, 8, 9, 11, 12, 13, 14, 15, 16, 17 and 18.
5. **Display data:** You can display features based on their attributes, a powerful feature you'll come to appreciate. You will learn how to symbolize features in different ways in Chap. 2 and 3.
6. **Present data:** You can create and distribute high-quality maps, graphs, and reports to present your analysis results in a compelling way to your audience. You will learn how to design an effective map in Chap. 3.

GIS is a computerized system that deals with spatial data in terms of the following:

1. ***Storage***: Digital and database storage.
2. ***Management of Data***: Integration of the database into the GIS system.

3. ***Retrieval***: The capacity to view the various database data formats.
4. ***Conversion***: Convert different sets of data from one form to another.
5. ***Analysis***: Manipulating data to produce new information.
6. ***Modeling***: Simplifying the data and its process.
7. ***Display:*** Presenting the output works.

Organization

GIS is a complete system that consists of sophisticated hardware and software. It performs many integrated functions:

1. GIS accepts data from multiple sources, which can be in a variety of formats. For example, if you are dealing with ArcGIS you can work not only with Shapefile, but also with Coverage (ARC/INFO format), Geodatabase, DXF (AutoCAD format), DBS (database system), and other types of database and digital formats.
2. Data types include the following:
   - Maps (Tiff, Jpeg, etc.)
   - Images from aircraft and satellite
   - Global Positioning System (GPS): (Coordinates, elevation)
   - Text data (report and text)
   - Tabular data (excel file)

GIS Infrastructure

1. Hardware: The machine where the GIS can be run (computer, digitizer, plotter, printer).
2. Software: The program needed to run the GIS (ArcGIS and its extensions)
3. Data: The digital and database (information)
4. Organization and People: This is the most important part of the GIS structure. The GIS is too important and so costly that it cannot be considered just equipment. It requires organization and staff to utilize this technology. Unfortunately many organizations treat the GIS as equipment rather than an important analysis tool.

GIS Principles

1. The computer is an unavoidable technology in our time. We are living in the digital age, which has become an important element in nearly all professions.
2. Computer training in most scientific disciplines is essential. Without this technology all professionals will be handicapped.
3. The GIS is an inevitable technology that will be used in all scientific fields. The GIS has become the accepted and standard means of using spatial data.
4. GIS is more ***Accurate Flexible***, ***Object Efficient***, and ***Rapid Fun*** comparing with the traditional method of spatial data inventory.
5. GIS is replacing traditional cartography. Much of traditional "pen and ink" cartography done by skilled draftsperson and artist is being replaced by GIS.
6. GIS is opening new horizons. New mode of analysis and applications are constantly discovered.

Spatial Data Representation

Spatial data is a fundamental component in any GIS environment. The data is based on the perception of the world as being occupied by features. Each feature is an entity which can be described by its attribute or property, and its location on earth can be mapped using a spatial reference. The most common representation of spatial data that measures the landscape is using discrete data (vector model) and continuous data (raster model). The data models are a set of rules used to describe and represent real world features in a GIS software.

Vector Data Model

The vector data model is a representation of the world of distinct features that have definite boundaries, identities, and has a specific shape using point, lines, and polygons. Vector data is structured with two specific elements (node, vertex) and coordinates. This model is useful for storing data that has discrete boundaries, such as groundwater wells, streams, and lakes. Each entity has a dimension, boundary and location. For example, a well has a specific measurement and its location can be described using a coordinate system such as latitude-longitude. The following represents the three fundamental vector types that exist in GIS.

Point: A point entity is simply a location that can be described using the coordinate system (longitude, latitude or X, Y). The point has no actual spatial dimension and has no actual length and width but has a specific location in space (single coordinate pair). Point can be represented by different symbols. Points generally specify features that are too small to show properly at a given scale. For example, buildings, schools, or a small farm at a scale of 1:25,000 can be represented as a point.

Figure 1.1 shows the location of five groundwater wells with each well representing a point feature. Table 1.1 shows the coordinate locations of the wells in (X, Y). The coordinate system allows users to integrate the wells into GIS and make them subject to mapping. The well feature is associated with an attribute table. The attribute of each well has information related to the depth and the yield of each well.

Line: A spatial feature that is given a precise location that can be described by a series of coordinate pairs. Each line is stored by the sequence of the first and last point together with the associated table attribute of this line. Line is one dimensional feature and has length but no width. Lines are a linear feature such as rivers, pipelines, and fences. The more points used to create the line, the greater the detail. The recent requirement that the line features include topology, which means that the system stores one end of the line as the starting point and the other as the end point, giving the line “direction”.

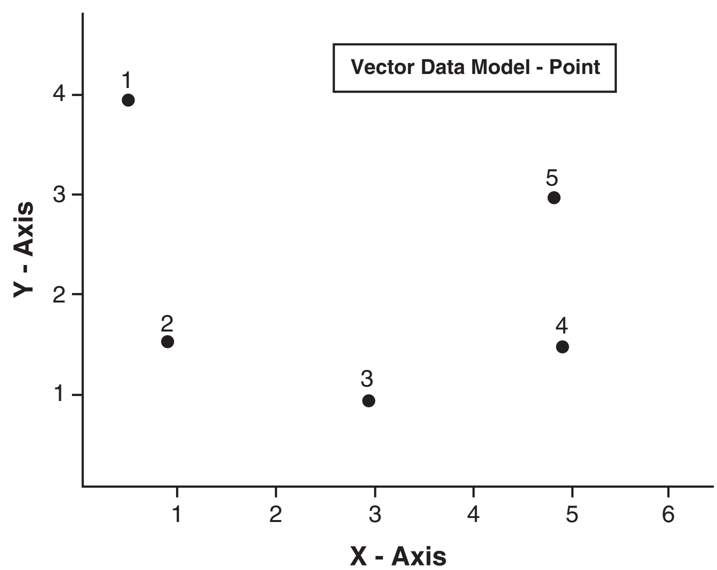


Fig. 1.1 Point feature representation

Table 1.1 Attribute table of point feature

No.	Well	X	Y	Well	Depth (m)	Yield (m^3/h)
1	WAJ-1	0.5	4.0	WAJ-1	78	90
2	WAJ-2	1.0	1.5	WAJ-2	48	68
3	WAJ-3	2.95	0.9	WAJ-3	35	54
4	WAJ-4	4.95	1.5	WAJ-4	58	75
5	WAJ-5	4.90	3.0	WAJ-5	55	75

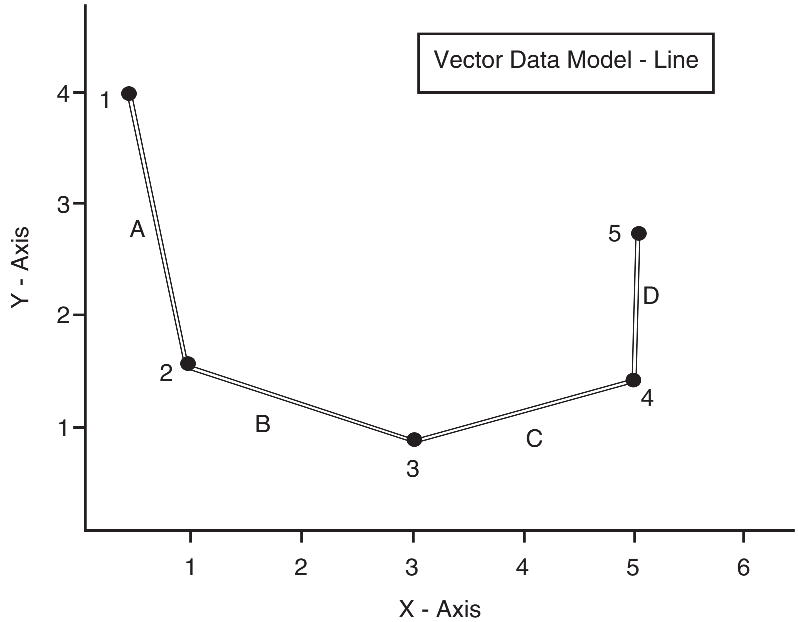


Fig. 1.2 Line feature representation

Table 1.2 Attribute table of line feature

Node No	X	Y
1	0.5	4.0
2	1.0	1.5
3	2.95	0.9
4	5.0	1.5
5	5.1	2.8
Line	1st Node	Last Node
A	1	2
B	2	3
C	3	4
D	4	5
Line	Length	Discharge
A	25.5	5
B	20.6	4
C	20.6	4
D	15.0	3

Figure 1.2 shows four pipelines (A, B, C, and D). Each pipeline is represented by a line that has its first and last node to distinguish its location. Each line has attributes of length and discharge. Notice that each node has coordinates (X, Y) stored in another table (Table 1.2).

Polygon: The polygon is an area fully encompassed by a series of connected lines. The first point in the polygon is equal to the last point. Polygon is a 2-D feature with at least three sides and because lines have direction, the area that falls within the lines compromise the polygon and the perimeter can be calculated. All of the data points that form the perimeter of the polygon must connect to form an unbroken line. Polygons are often an irregular shape such as parcels, lakes, and political boundaries.

Figure 1.3 shows polygon A, which represents an agricultural field. The polygon has its first and last node in node number 1 to settle its location. Node 1, vertexes 2, 3, 4, 5, and 6 have coordinate (X, Y) that is stored in another table (Table 1.3). Aside from location attributes, the polygon has associated attributes of area and crop.

Features on maps have spatial relationships which shows how those features are related to each other in space. The most important spatial relationships are:

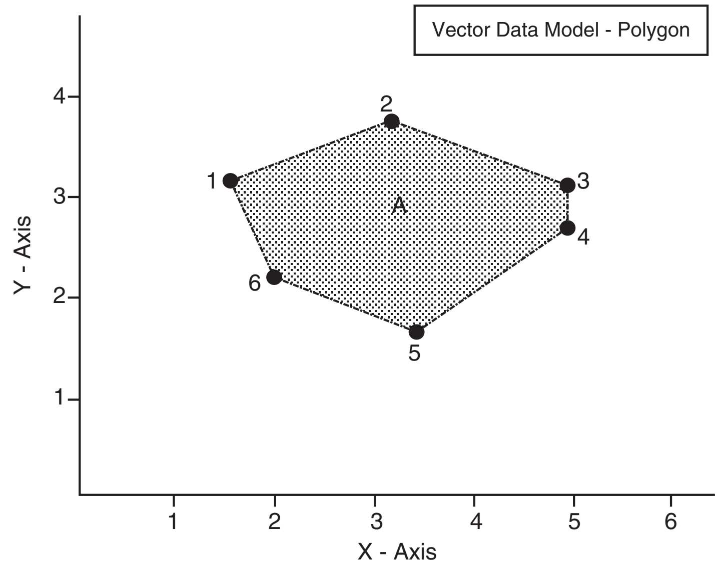


Fig. 1.3 Polygon feature representation

Table 1.3 Attribute table of polygon feature

Node No 1 and vertexes	X	Y
1	1.6	3.1
2	3.2	3.8
3	5.0	3.1
4	5.0	2.6
5	3.4	1.8
6	2.0	2.2
Polygon	Node	
A	1, 2, 3, 4, 5, 6	
Polygon	Area	Crop
A	520	Tomato

1. **Distance**: This measures the distance from one feature to another in the GIS map. The distance concept is an important relationship as the distance between features can be measured in any unit regardless of the map's coordinate system.
2. **Distribution**: This is the collective location of features where relationships can show the feature among themselves or their spatial relationships with other features in the map.
3. **Density**: This is the number of features per unit area or simply how close features are to each other.
4. **Pattern**: This is the consistent arrangement of a feature.

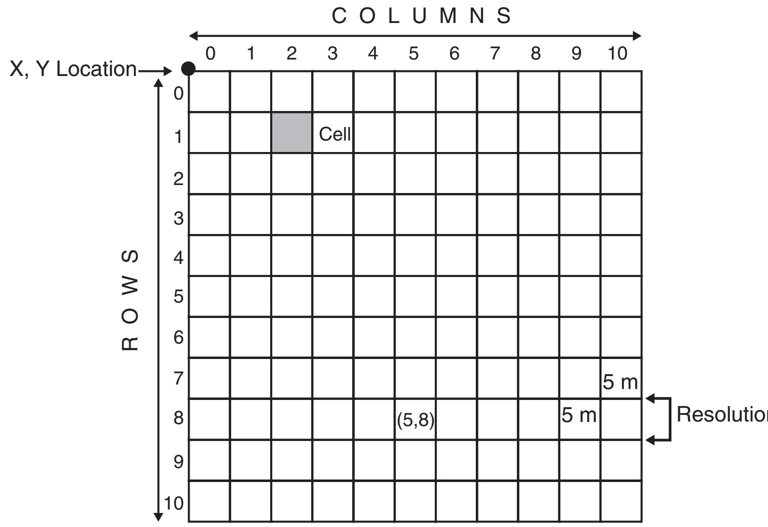
Raster Data Model

A representation of an area or region as a surface divided into a grid of cells (Fig. 1.4). It is useful for storing data that varies continuously, as in an aerial photograph, a satellite image, a surface of humidity, or digital elevation model (DEM) (Fig. 1.5).

Figure 1.4 depicts the basics of the raster data model. The cell is the minimum mapping unit and the smallest size at which any landscape feature can be represented.

These cells in the raster dataset are used as building blocks for creating points, lines, and polygons. In the raster data model, points, lines, polygons are represented by grid cells. The location of each cell in the grid is determined by two things: (1) the origin of the grid (the upper left-corner, which is (0, 0) and the resolution (size of the cell). The resolution is determined by measuring one side of the square cell. For example, a raster model with cell representing 5 m by 5 m (25 m^2) in the real

Fig. 1.4 Characteristics of raster data structure



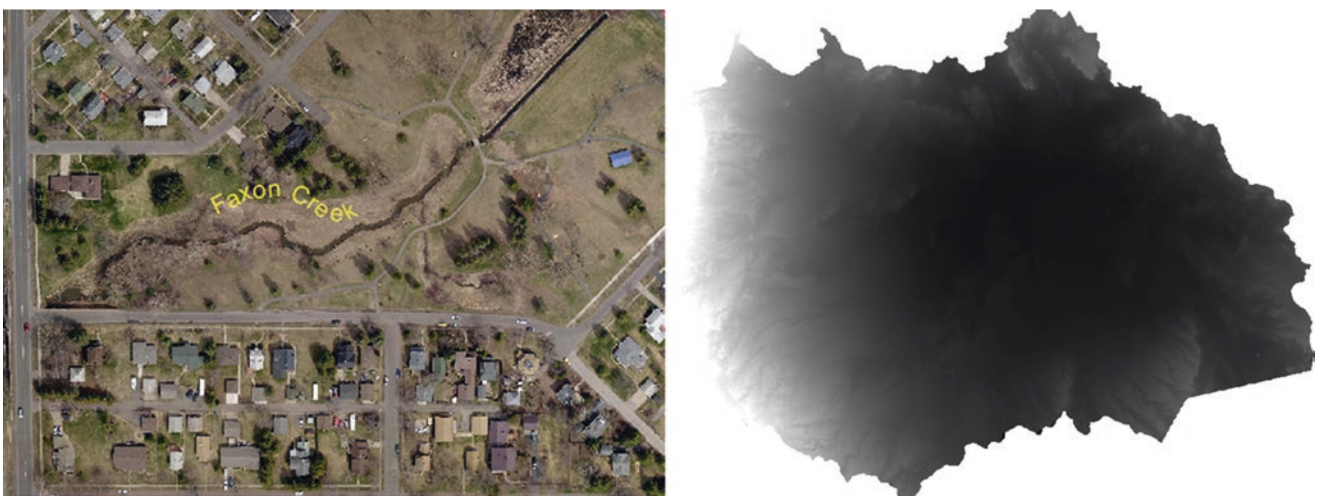


Fig. 1.5 Aerial photograph of Faxon Creek, Superior, WI (*left*) and DEM Jafr, Jordan (*right*)

world would be said to have a spatial resolution of 5 m. Each cell in the raster carries a single value, which represents the characteristic of the spatial phenomenon at a location denoted by its row and column. The precision of raster data is ruled by the resolution of the grid data set. The data type for that cell value can be either integer or floating-point.

The raster model will average all values within a given cell to yield a single value. Various techniques are used to assign cell code such as presence-absence, cell center, dominant area, and percent coverage. The more area covered per cell, the less accurate the associated data values. The area covered by each cell determines the spatial resolution of the raster model from which it is derived. Raster coding produces spatial inaccuracies as the shape of features is forced into an artificial grid cell format. Therefore, there is no way to know where any small feature occurs within the cell as the location according to the raster format is simply the entire cell. If the raster cell representing 100 m by 100 m, and the cell represents a well that has 0.5 m diameter. The cell in this case represents the well and this makes the raster format imprecise.

Advantages and Disadvantages of the Raster and Vector Model

There are several advantages and disadvantages for using either the raster or vector data model for storing and displaying spatial data.

Raster Model

Advantages

1. Simple data structure
2. Efficient for remotely sensed or scanned data
3. Simple spatial analysis procedures

Disadvantages

1. Requires greater storage space on a computer
2. Depending on pixel size, graphical output may be less pleasing
3. Projection transformations are more difficult
4. Difficult to represent topological relationships

Vector Model

Advantages

1. Data can be represented in its original resolution without generalization
2. Requires less disk storage space
3. Topological relationships are readily maintained
4. Graphical output closely resembles hand-drawn maps

Disadvantages

1. More complex data structure
2. Inefficient for remotely sensed data
3. Some spatial analysis procedures are complex and process intensive
4. Overlaying multiple vector maps is often time consuming

GIS Project

To carry a GIS project, users need to integrate spatial data into the GIS software where the data can have a vector or raster dataset. GIS data comes from many resources

1. Hard copy maps
2. Digital files
3. Imagery
4. GPS
5. Excel, text delimited, and dbf files
6. Reports

The GIS analysis is based on the database, which is powerful and important in GIS.

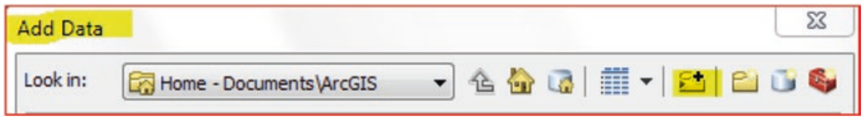
Lesson 1: Explore the Vector Data

Layers in vector data format in GIS can be a point, line and polygon, either in shapefiles, coverage, or geodatabase feature classes. These layers can be integrated in ArcMap and classified and symbolized by different symbols. In this lesson you are going to display Newton Creek that starts from Murphy Oil Inc. (Superior, WI) and discharges into Hog Island including the 5-sampling sites.

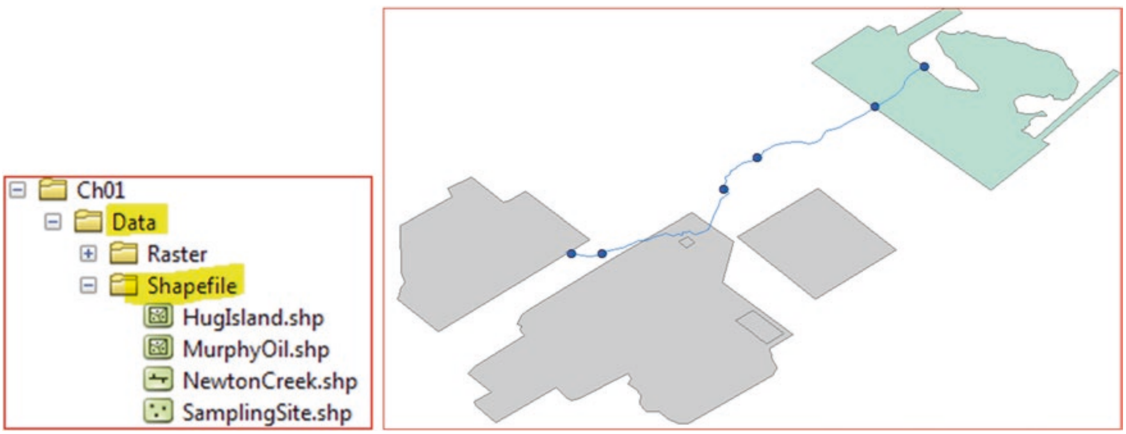
1. Launch ArcMap[1]

2. Click Add Data  in the Standard Toolbar

3. In the Add Data Dialog box click the Connect To Folder  button



4. Browse to Ch01/click OK
5. D-click Data folder/open Shapefile folder and highlight **HugIsland.shp**, **MurphyOil.shp**, **NewtonCreek.shp**, and **SamplingSite.shp**
6. Click Add



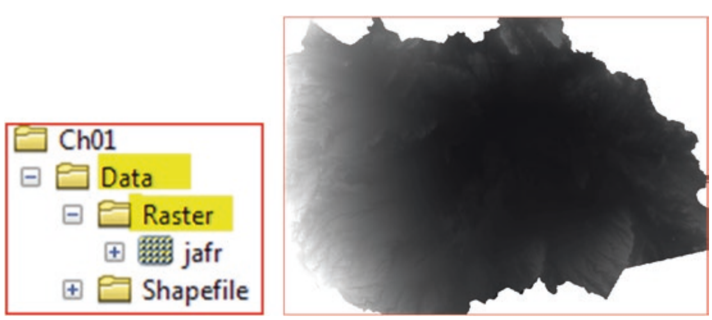
Result: The 4-layers display in ArcMap

Lesson 2: Explore the Raster Data

A raster map is full of continuous data like a photograph and uses different formats. In this lesson you are going to use a Digital Elevation Model (DEM) of Jafr basin (Jordan). It is a depression basin located in the east of the country and has groundwater resources used for irrigation.

7. Insert menu/click Data Frame

8. Click Add Data  in the Standard Toolbar
9. Navigate to \\Ch01\Data\Raster and integrate **jafr** (DEM)



[1] From the window taskbar/Start/All Programs/ArcGIS/ArcMap.

10. In the Table of Content (TOC)/R-click Jafr image/Properties/click Source tab

    Note: the Raster property includes four sections that provide detailed information about the Jafr DEM

- **Raster Information**: This shows that the DEM has 15,689 columns and 11,633 rows. It shows that the image consists of 1 band and the cell size (resolution) is 10 × 10 m. It also indicates that the pixel type is a signed integer, meaning that the raster has an attribute table and can be opened in ArcGIS. The Pixel depth of the raster is 16 Bit.
- **Extent**: this shows the coordinate extent in meter.
- **Spatial Reference**: shows that the coordinate of the Jafr DEM is registered in Palestine_1923_Palestine_Belt (customized UTM projection) and the datum is D_Palestine_1923. It also provides information about the parameters of the projection such as the false easting and false northing.
- **Statistics**: this shows the minimum and maximum elevation of the area.

Raster Information	
Columns and Rows	15689, 11633
Number of Bands	1
Cell Size (X, Y)	10, 10
Uncompressed Size	696.22 MB
Format	GRID
Source Type	Generic
Pixel Type	signed integer
Pixel Depth	16 Bit

Extent	
Top	1013205.07347
Left	191250.731268
Right	348140.731268
Bottom	896875.073475

XY Coordinate System	Palestine_1923_Palestine_Belt
Linear Unit	Meter (1.000000)
Angular Unit	Degree (0.0174532925199433)
False_Easting	170251.555
False_Northing	1126867.909
Central_Meridian	35.21208055555556
Scale_Factor	1
Latitude_Of_Origin	31.73409694444445
Datum	D_Palestine_1923