

M.Sc. GEOGRAPHY LAB MANUAL

3rd Semester



Prepared By
Pure & Applied Science Dept.
Geography

MIDNAPORE CITY COLLEGE



MIDNAPORE CITY COLLEGE

MA/ MSc in Geography

Semester - III

(Study Material for Practical Courses)

Paper – GEO 395: GIS Application in Research

GEO 395.1: Applications of Remote Sensing & Geographic Information System

Course Outcome:

The course is designed for the general ideas of GIS and image-based information. Upon completion of this course, students get benefit from these baseline concepts to further increase their knowledge.

Full Marks- 25

Pattern of Setting Questions: Three compulsory questions bearing marks 7, 7 and 6 respectively, will be set covering the whole unit. 5 marks will be allotted for Laboratory Note Book and Viva-voce.

1. Geo-referencing of maps and satellite images.
 2. Preparation of different types of FCC. Image enhancement, Band rationing, Density slicing..
 3. Supervised and Unsupervised Classification Techniques, generation of signature statistics and signature reparability reports. Generation of NDVI.
 4. GIS: RS, GIS and GNSS.
 5. Map layers: Overlay analysis, buffering and map composition
-

1.1 Georeferencing of Maps and Satellite Images with Arc GIS

Georeferencing is the name given to the process of transforming a scanned map or aerial photograph so it appears “in place” in GIS. By associating features on the scanned image with real world x and y

coordinates, the software can progressively warp the image so it fits to other spatial datasets. This manual will explain how to georeference a raster image in ArcGIS, so it can then be used as an overlay or for digitizing purposes.

Georeferencing requires a spatially referenced dataset that will be used to provide locations on the scanned map with their associated coordinates. In this following example, we will match intersections represented on the scanned map with a shape file of city streets.

1. Click the **Add Data** button to bring up the *Add Data window* which allows you to browse to the folder and file location,
2. Click **ADD** to add the dataset to the *Table of Contents*.
3. Right-click on the name of the layer in the *Table of Contents window*, and select **Label Features**.

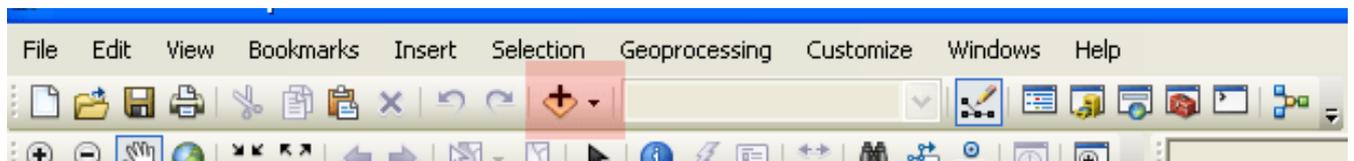


Figure 1

4. Before adding the raster file to be georeference make sure to open it in an image viewer or editor to ensure there is no excess areas which can often occur with scanned images.

For best results, the image should be rotated and cropped to so it is oriented in the right direction and areas outside the edge of the map sheet have been--removed. The image used in this manual (seen at left, below) is ready to be georeference.

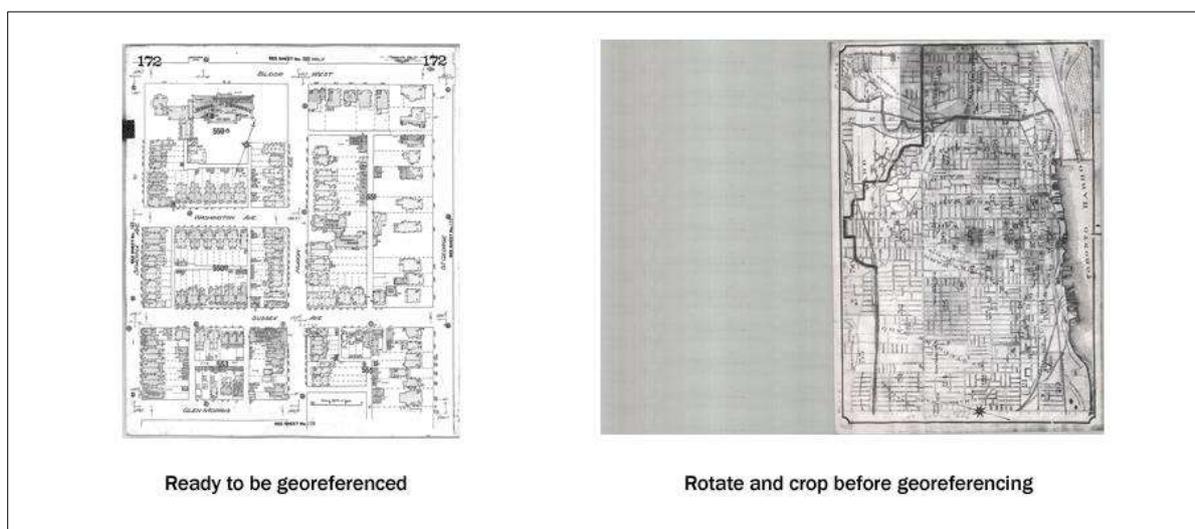


Figure 2

4. Use the **Add Data** button to add the scanned map image to the map document. If asked if you would like to build pyramids, click yes. Building pyramids improves the speed in which raster datasets load, and is optional.
5. You should receive the following warning message after adding the scanned image to Arc Map:

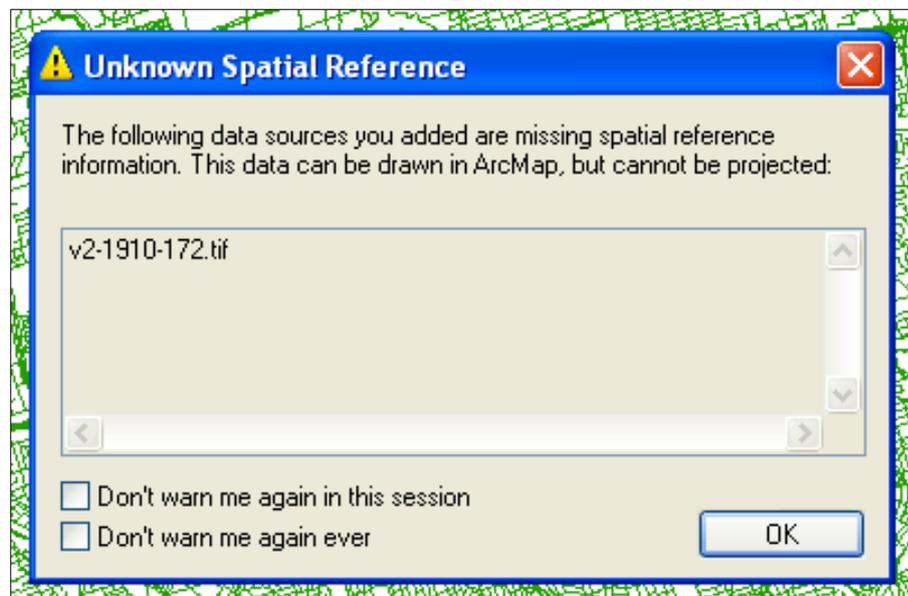


Figure 3

Note: The Georeferencing process in ArcGIS will solve the issue of missing spatial references. Click **OK**. Because Arc Map does not know where to locate this raster dataset, it will not display on the map, but will be visible in the Table of Contents.

6. Add the Georeferencing toolbar by going to Customize - Toolbars - Georeferencing.

The dropdown menu in the *Georeferencing toolbar* should display the filename of the raster file added to the map document.



Figure 4

Note: When multiple raster files are present in the Table of Contents, the dropdown menu

controls which raster the Georeferencing tools will be targeting, therefore it is important to make sure that you have the correct raster selected.

7. You will notice that while the raster file appears in the *Layers* portion of the *Table of Contents* it is not immediately visible on the screen. To make the image visible click the **Georeferencing** button to the left of the dropdown menu, and select **Fit to Display** from the menu that appears.

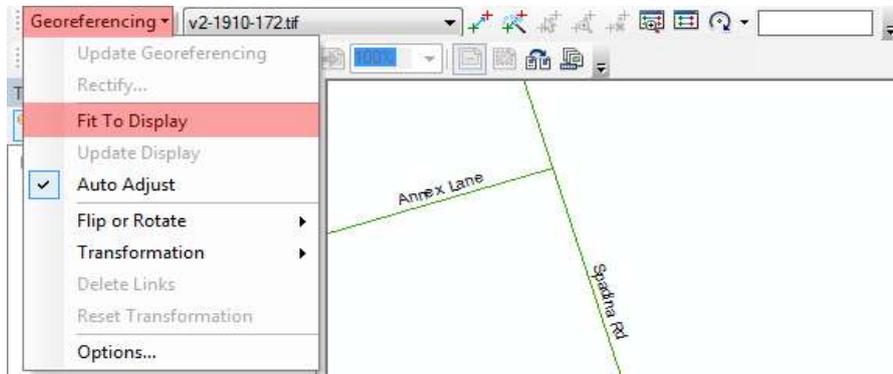


Figure 5

Note: The scanned map will appear with the data set as seen below, scaled to the extent of the current display – it should appear to be larger than the City of Toronto at the current zoom level.



Figure 6

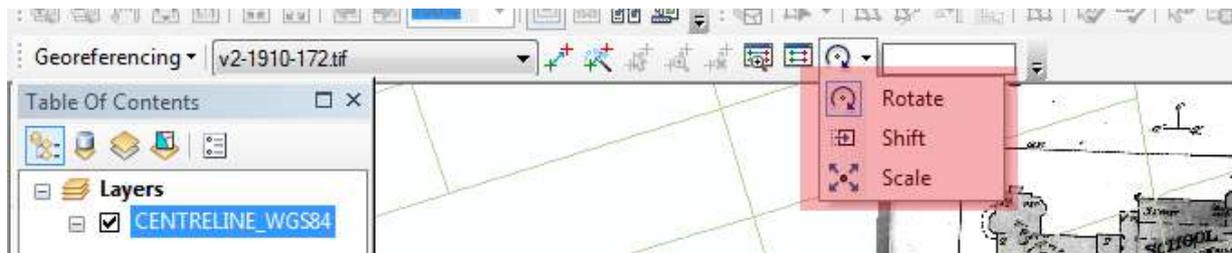


Figure 9

You can now begin to georeference the map.

11. In order to increase accuracy in selecting your control points it is best to turn on the **Vertex Snapping** option in the *Snapping* toolbar.

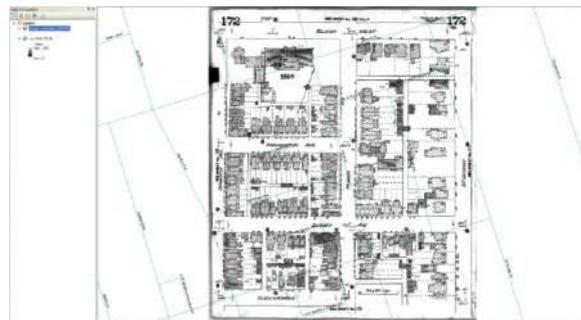


Figure 10

If this toolbar is not initially available it can be added to the main screen via *Customize > Toolbars > Snapping* and can then be dragged and dropped into position as below.

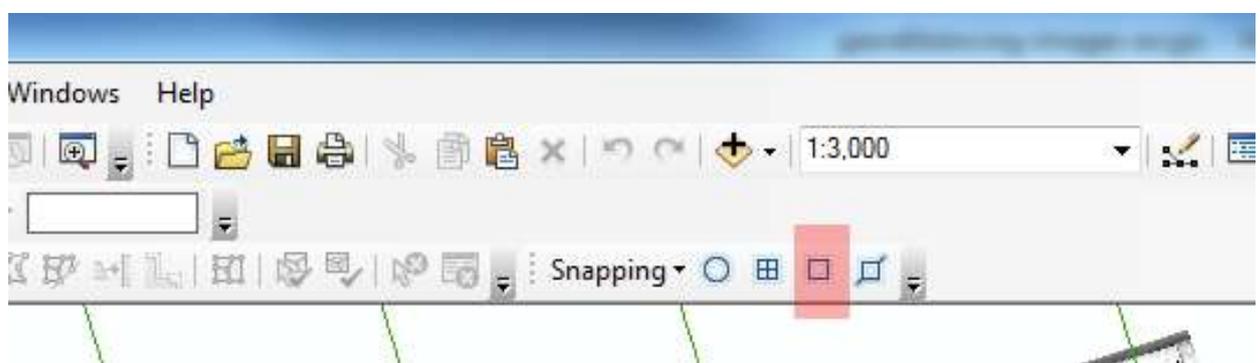


Figure 11

12. Another option which can be selected at this time is the **Auto-Adjust** option within the *Georeferencing toolbar* which automatically updates the display and adjusts the raster image with each new control point added. It is recommended that the Auto-Adjust option be selected for most GIS needs so that you can monitor the Georeferencing process in real time.

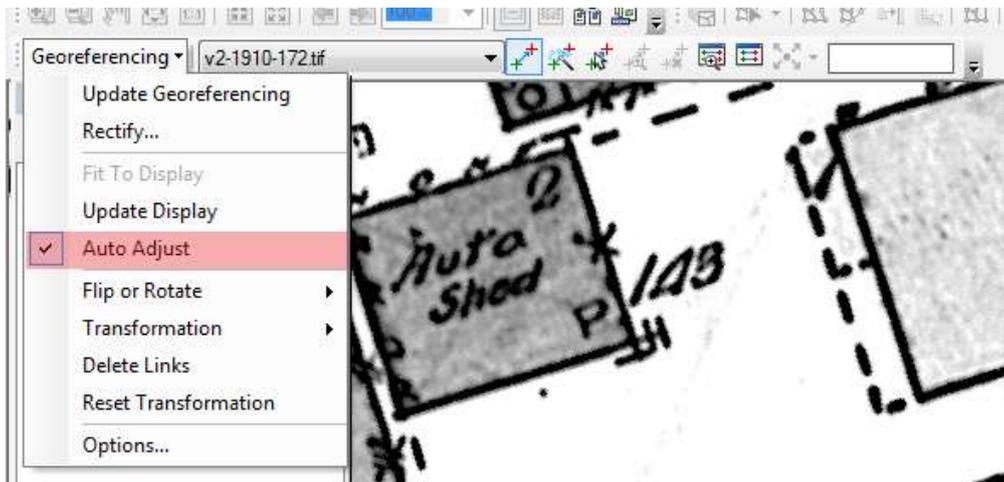
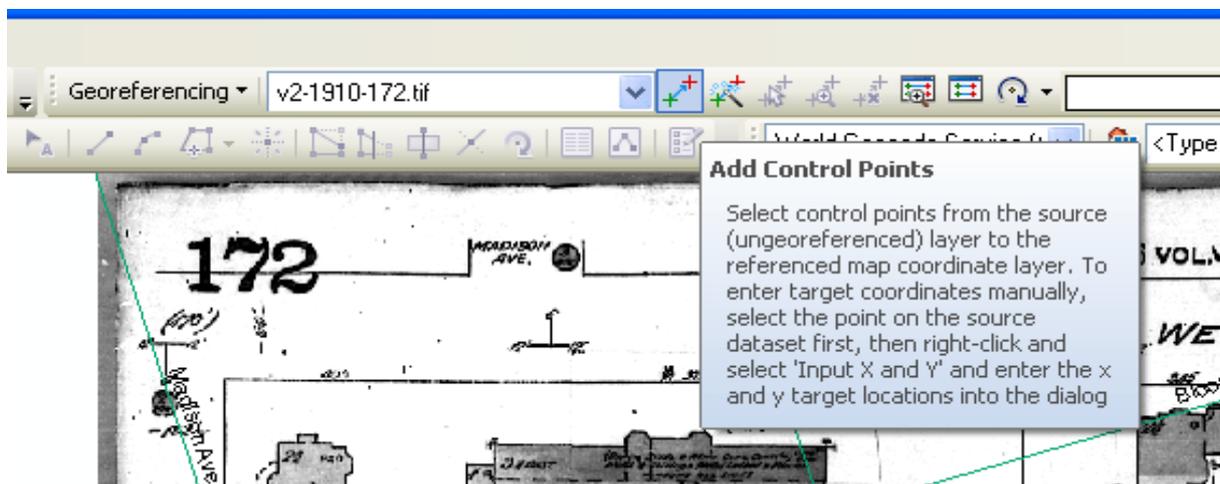


Figure 12

13. Click on the **Add Control Points** button in the Georeferencing toolbar. This tool allows you to associate points in a scanned map or aerial image with their real-world spatial coordinates within your GIS.



14. To add a **Control Point** to your map, first click on a location on the scanned map (a green 'plus' sign will appear as seen in the first image below), then click on the corresponding location in your street dataset to associate the two (a red 'plus' sign will appear as seen in the second image below). In this example, we will match the centre of the intersection of streets on the scanned map to the intersection of the same streets in our geospatial data, represented as line features in the spatial dataset. **Note that the lines in the middle of streets on the scanned map show the location of water pipes, not necessarily located at the centre of the intersection.**

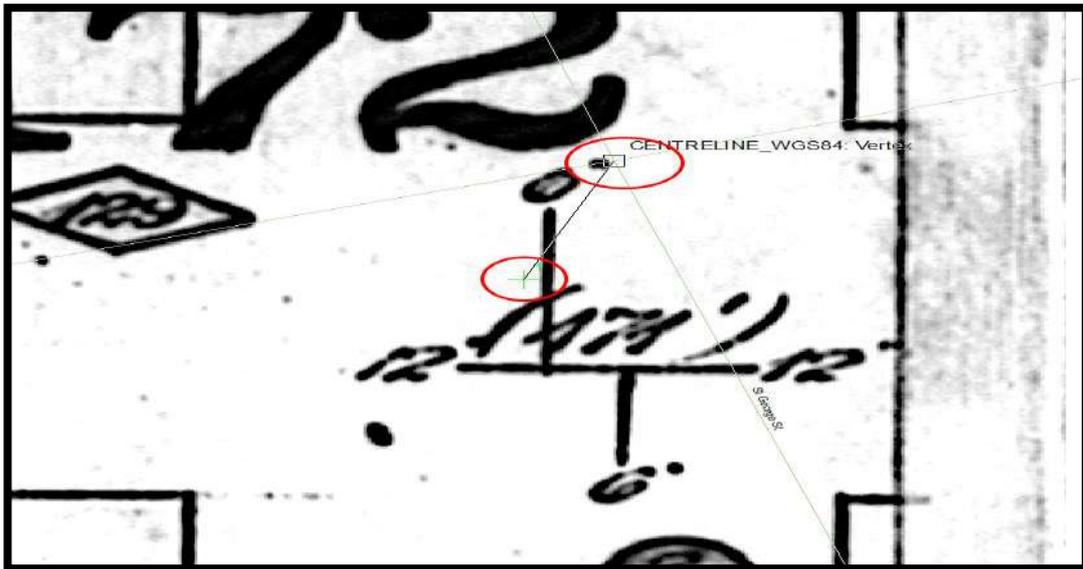


Figure 14

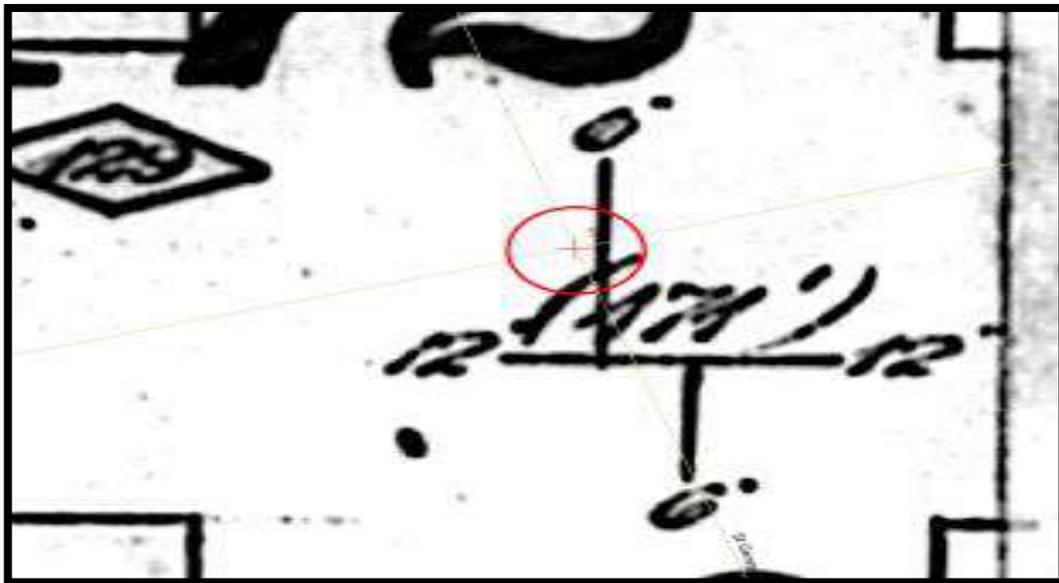


Figure 15

15. Continue adding control points until the raster image moves into place geographically, aligning itself with the digital data map. You will notice that after two control points were added the image automatically warped into position, in most cases several more control points should be added to increase the accuracy of the overlay. Select areas close to four corners of the map first, and then concentrate on your map area between the four corner points which should ensure that all points are georeference evenly.
16. If at any time you need to edit, save or delete any Control Points collected, this can be done by accessing the **View Link Table** button on the *Georeferencing taskbar* (highlighted in red below). This table displays the Control Point(s) information as well as the *Transformation* and *Auto-Adjust* options which control how the raster image responds to the addition of control points.

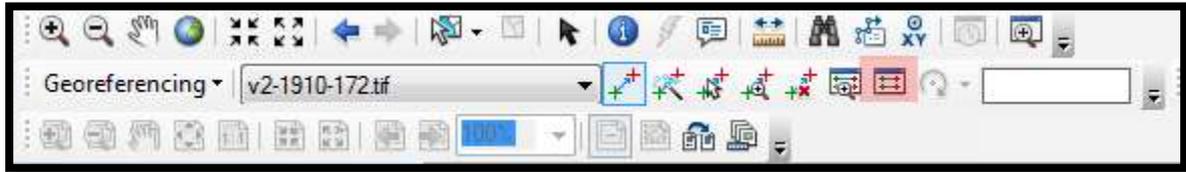


Figure 16

Each line represents a Control Point, a link between the raster image and the digital map data. Both the *Source* and *Map Points* can be edited by double-clicking on the values in the table. Clicking on the *Save* button allows the points table to be saved as a text file while the *Load* button allows for a set of control points stored in a text file to be loaded and edited. Saving a text file of your control points is strongly recommended

For images requiring many control points to georeference, periodically saving your control points to a text file allows you to come back to your work in case you are not able to finish it in one session.

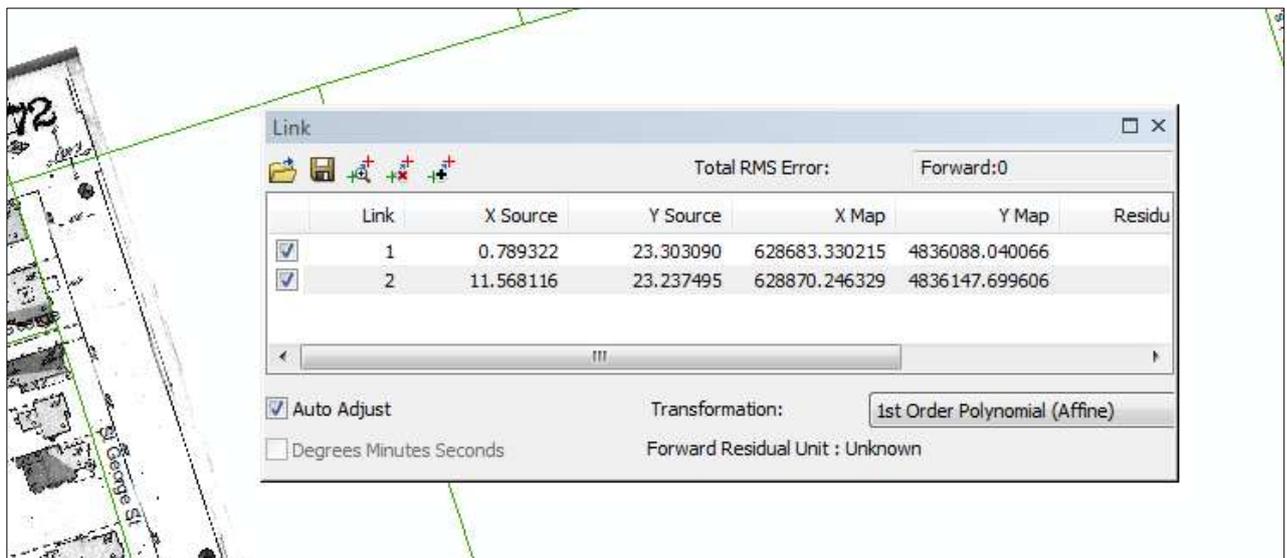


Figure 17

With the addition of control points at intersections and the auto-adjustment of the raster image to the digital dataset, your image should now look very similar to the image below:

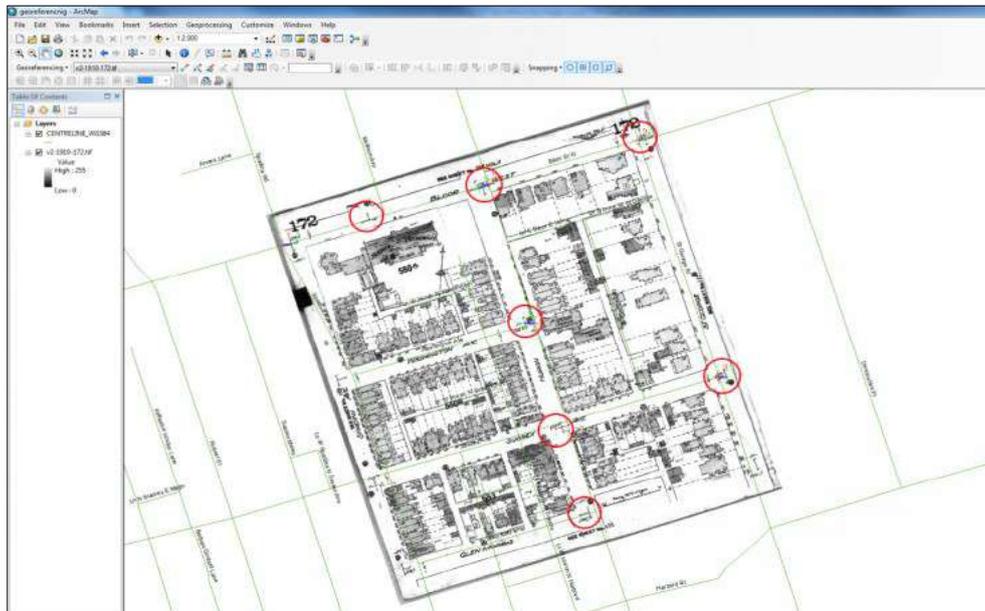


Figure 18

17. Open the Georeferencing menu in the toolbar.

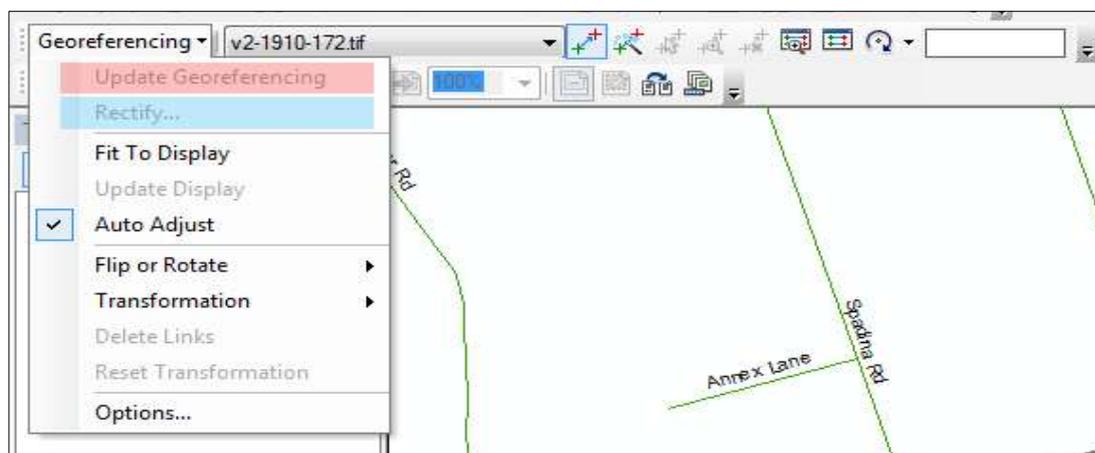


Figure 19

Using the **Update Georeferencing** tool will store the transformation information with the raster file, which can be internal or external depending on the type of raster dataset being used (i.e. for a the current raster dataset, that is a TIFF file, the transformation will be stored in what is known as a world file, with a .tfw or .tfwx extension).

18. Using the **Rectify** command will allow you to create a new raster dataset that is georeference using both map co-ordinates and spatial references which can be saved in a number of different formats (BIL, BIP, BMP, DAT, GIF, TIFF, ESRI GRID, IMG, JPEG, JPEG 2000, PNG). This

command can be useful if you wish to perform further analysis with the raster dataset or use it in an external program that does not recognize the external georeference information created in the ArcGIS world file (i.e. tfwx).

Note: Both functions can be accessed via the *Georeferencing toolbar*. For the **Rectify** option, another box will be displayed in which **Resample Type** and **Data Name, Location** and **Format** may be selected.

2. Preparation of different types of FCC. Image enhancement, Band rationing, Density slicing

2.1 False Colour Composites (FCC)

Colour assignment for any band of a multispectral image can be done in an/entirely arbitrary manner. The colour of a target in the displayed image does not have any resemblance to its actual colour. The resulting product is known as a false colour composite image. There are many possible schemes of producing false colour composite images. However, some scheme may be more suitable for detecting certain objects in the image.

A very common false colour composite scheme for displaying a SPOT/multispectral image is shown below:

R = XS3 (NIR band) G = XS2 (red band) B = XS1 (green band)

This false colour composite scheme allows vegetation to be detected readily in the image. In this type of false colour composite images, vegetation appears in different shades of red depending on the types and conditions of the vegetation, since it has a high reflectance in the NIR band as in the graph of spectral reflectance signature.



False colour composite multispectral SPOT image:
Red: XS3; Green: XS2; Blue: XS1

Figure 20

Another common false colour composite scheme for displaying an optical image with a short-wave infrared (SWIR) band is shown below:

R = SWIR band (SPOT4 band 4, Landsat TM band 5)

G = NIR band (SPOT4 band 3, Landsat TM band 4)

B = Red band (SPOT4 band 2, Landsat TM band 3)

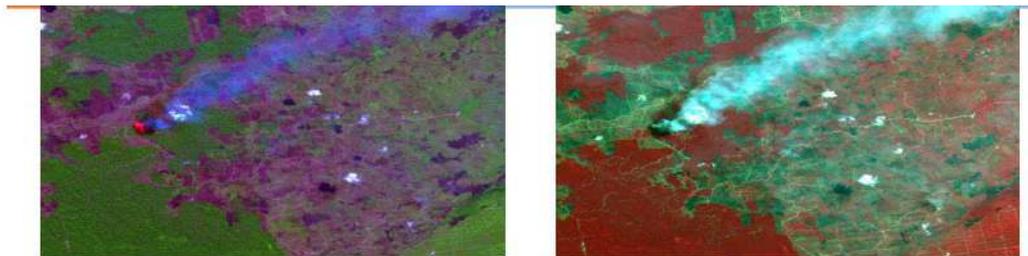


Figure 21

Left image: Including the SWIR band: Red: SWIR band; Green: NIR band; Blue: Red band. In this display scheme, vegetation appears in shades of green. Bare soils and clearcut areas appear purplish or magenta. The patch of bright red area on the left is the location of active fires. A smoke plume originating from the active fire site appears faint bluish in colour.

Right image: Excluding the SWIR band: Red: NIR band; Green: Red band; Blue: Green band. Vegetation appears in shades of red. The smoke plume appears bright bluish white.

2.2 Image Enhancement

Image enhancement techniques improve the quality of an image perceived by a human. These are applied separately to each band of a multispectral image in order to display or visually interpret the data more effectively. The objective is to create “new” images from original raw data in order to increase the amount of information that can be visually interpreted from the data. Image enhancement is attempted after the image is corrected for geometric and radiometric distortions.

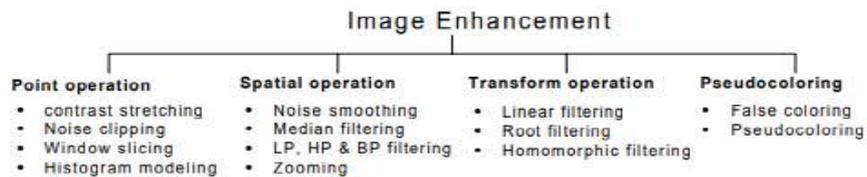


Figure 22

METHODS:

Image enhancement methods can be based on either spatial or frequency domain techniques.

Spatial domain enhancement methods:

- Spatial domain techniques are performed to the image plane itself and they are based on direct manipulation of pixels in an image.
- The operation can be formulated as $g(x,y) = T[f(x,y)]$, where g is the output, f is the input image and T is an operation on f defined over some neighbourhood of (x,y) .
- According to the operations on the image pixels, it can be further divided into 2 categories: Point operations and spatial operations (including linear and non-linear operations).

Frequency domain enhancement methods:

- These methods enhance an image $f(x,y)$ by convoluting the image with a linear, position invariant operator.
- The 2D convolution is performed in frequency domain with DFT.

Spatial domain: $g(x,y) = f(x,y) * h(x,y)$

Frequency domain: $G(w_1, w_2) = F(w_1, w_2) H(w_1, w_2)$

1. Enhancement by point processing

- These processing methods are based only on the intensity of single pixels.

Simple intensity transformation:

(a) Image negatives: Negatives of digital images are useful in numerous applications, such as displaying medical images and photographing a screen with monochrome positive film with the idea of using the resulting negatives as normal slides.

Transform function T: $g(x,y) = L - f(x,y)$, where L is the maximum Intensity.

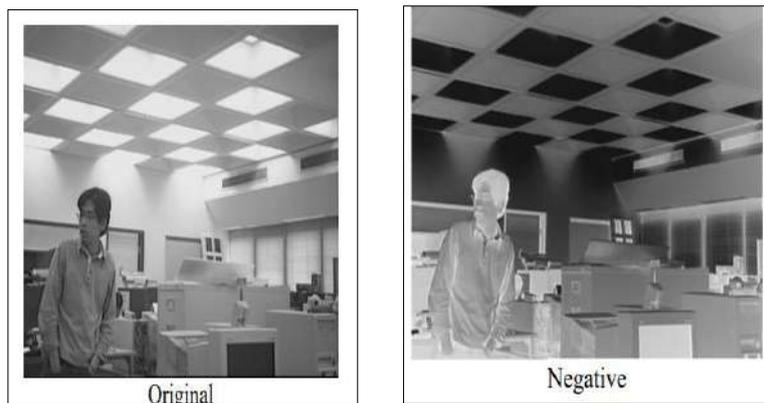
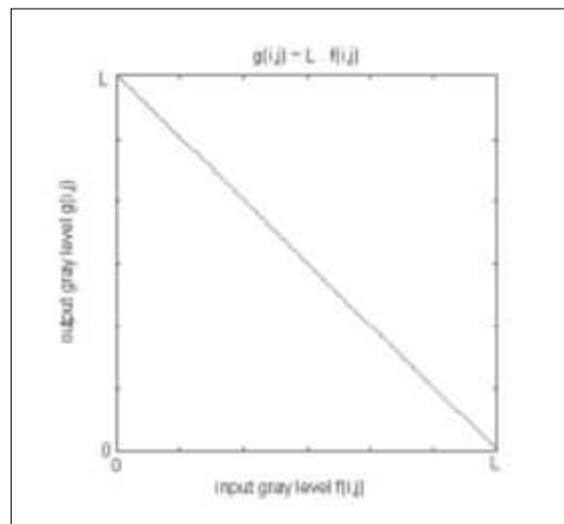


Figure 23

(b) Contrast stretching: Low-contrast images can result from poor illumination, lack of dynamic range in the image sensor, or even wrong setting of a lens aperture during image acquisition. The idea behind contrast stretching is to increase the dynamic range of the gray levels in the image being processed.

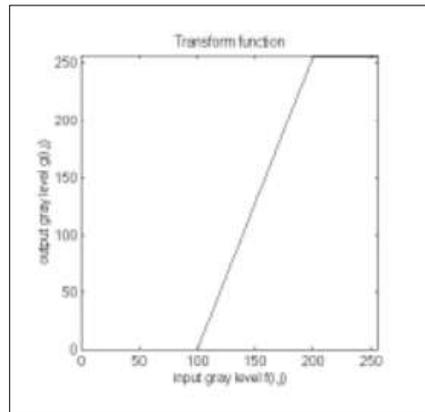


Figure 24

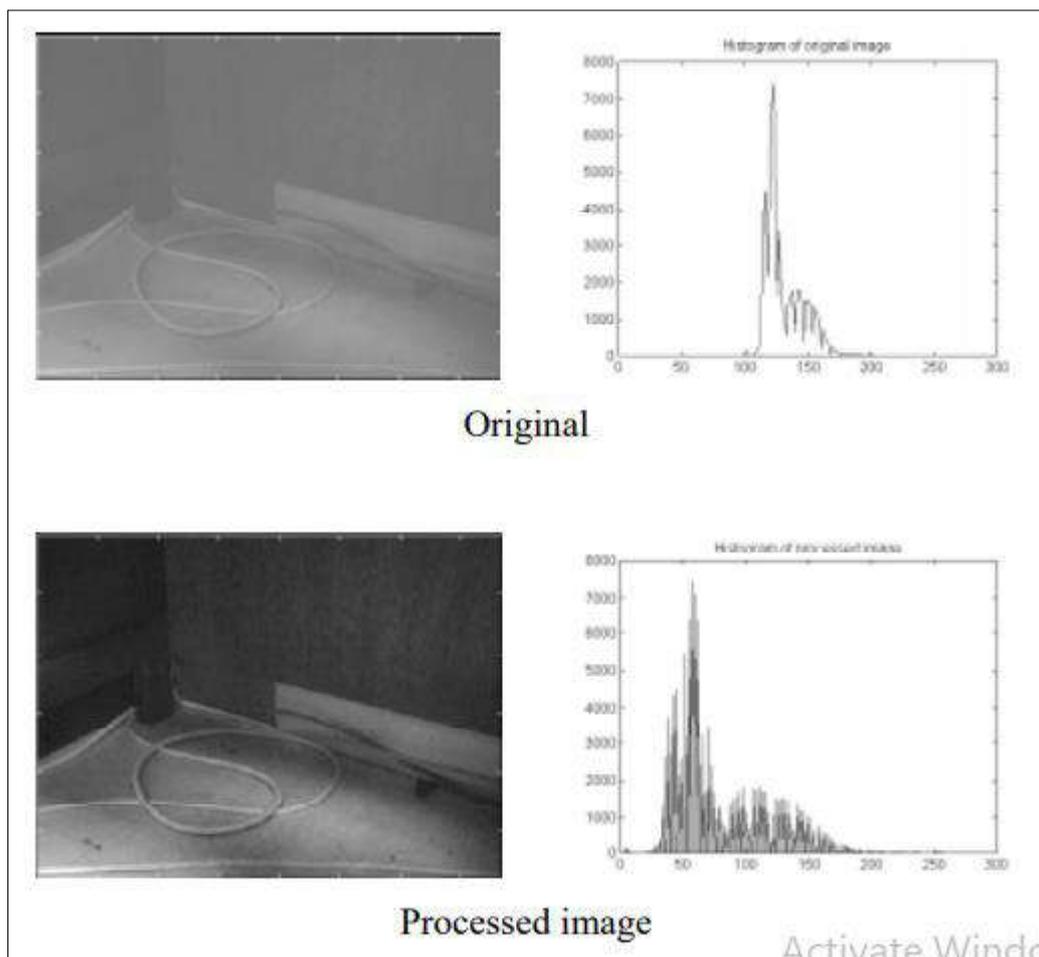


Figure 26

Special case: If $r_1=r_2=0$, $s_1=0$ and $s_2=L-1$, then it is actually a thresholding that creates a binary images.

(c) Compression of dynamic range: Sometimes the dynamic range of a processed image far exceeds the capability of the display device, in which case only the brightest parts of the images are visible on the display screen. An effective way to compress the dynamic range of pixel values is to perform the following intensity transformation function: $s = c \log (1+|r|)$ where c is a scaling constant and the logarithm function performs the desired compression.

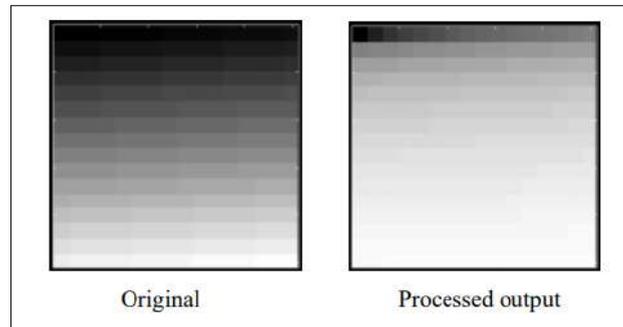


Figure 27

(d) Gray-level slicing: Highlighting a specific range of gray levels in an image often is desired. Applications include enhancing features such as masses of water in satellite imagery and enhancing flaws in x-ray images.

Histogram processing:

- The histogram of a digital image with gray levels in the range $[0, L-1]$ is a discrete function $p(r_k) = n_k/n$, where r_k is the k th gray level, n_k is the number of pixels in the image with that gray level, n is the total number of pixels in the image, and $k=0, 1, \dots, L-1$.
- $P(r_k)$ gives an estimate of the probability of occurrence of gray level r_k .
- The shape of the histogram of an image gives us useful information about the possibility for contrast enhancement.
- A histogram of a narrow shape indicates little dynamic range and thus corresponds to an image having low contrast.

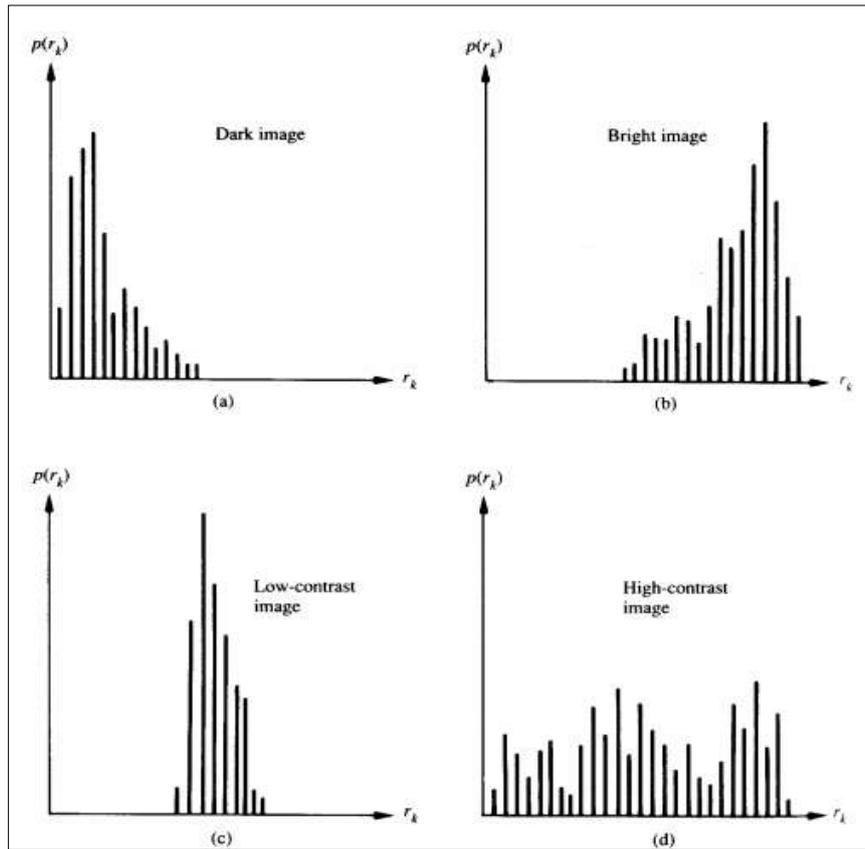


Figure 28

Spatial Filtering:

- The use of spatial masks for image processing is called spatial filtering.
- The masks used are called spatial filters.
- The basic approach is to sum products between the mask coefficients and the intensities of the pixels under the mask at a specific location in the image. $(2D \text{ convolution}) = \sum \sum w(i, j) f(x_i, y_j)$ where $(2d+1) \times (2d+1)$ is the mask size, $w(i, j)$'s are weights of the mask, $f(x, y)$ is input pixel at coordinates (x, y) , $R(x, y)$ is the output value at (x, y) . *
- If the center of the mask is at location (x, y) in the image, the gray level of the pixel located at (x, y) is replaced by R , the mask is then moved to the next location in the image and the process is repeated. This continues until all pixel locations have been covered.

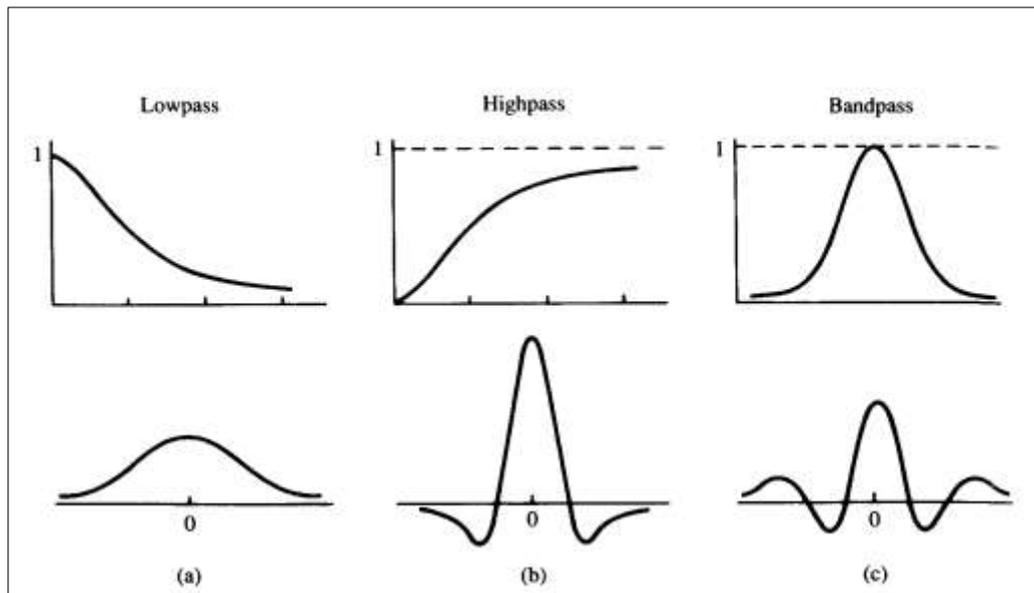


Fig: a) low pass b) band pass c) High pass filters

Figure 29

Smoothing filter:

- Smoothing filters are used for blurring and for noise reduction.
- Blurring is used in pre-processing steps, such as removal of small details from an image prior to object extraction, and bridging of small gaps in lines or curves
- Noise reduction can be accomplished by blurring with a linear filter and also by nonlinear filtering.

(a) Low pass filtering: The key requirement is that all coefficients are positive. Neighbourhood averaging is a special case of LPF where all coefficients are equal. It blurs edges and other sharp details in the image.

(b) Median filtering: If the objective is to achieve noise reduction instead of blurring, this method should be used. This method is particularly effective when the noise pattern consists of strong, spike-like components and the characteristic to be preserved is edge sharpness. It is a nonlinear operation. For each input pixel $f(x,y)$, we sort the values of the pixel and its neighbours to determine their median and assign its value to output pixel $g(x,y)$.

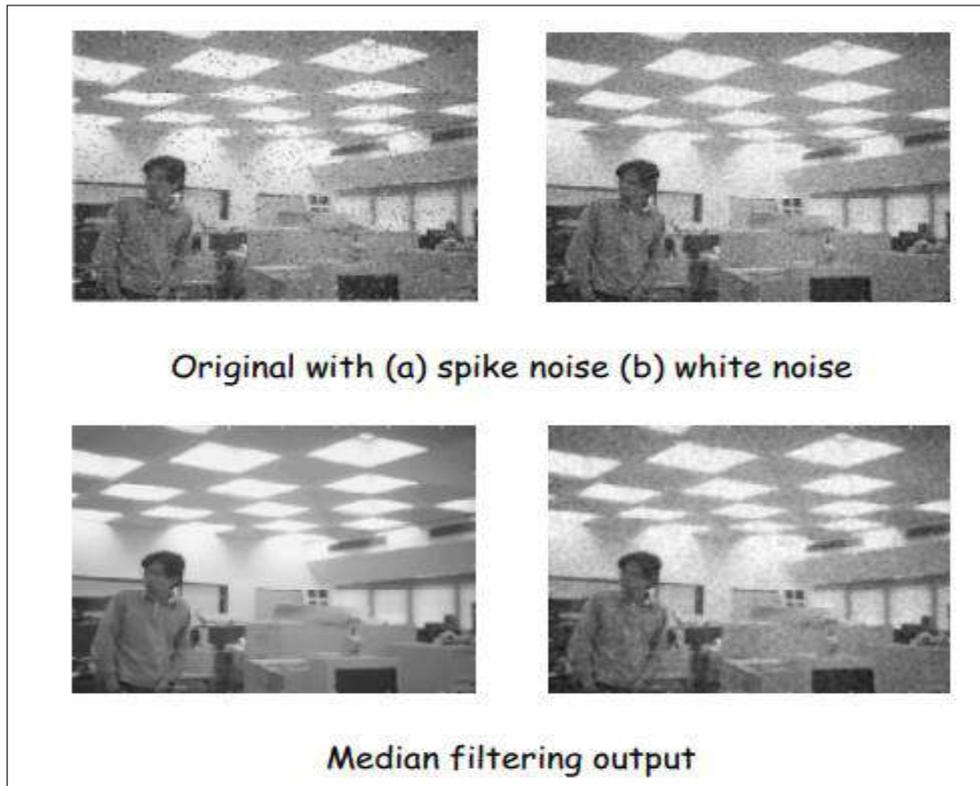


Figure 30

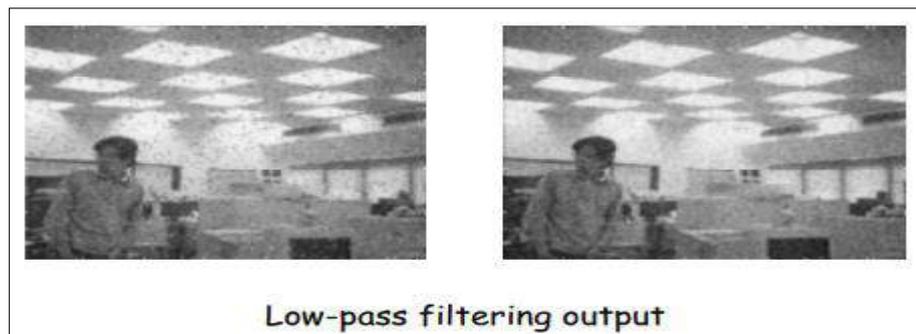


Figure 31

Sharpening Filters:

- To highlight fine detail in an image or to enhance detail that has been blurred, either in error or as a natural effect of a particular method of image acquisition.
- Uses of image sharpening vary and include applications ranging from electronic printing and medical imaging to industrial inspection and autonomous target detection in smart weapons.

(a) Basic highpass spatial filter: The shape of the impulse response needed to implement a highpass spatial filter indicates that the filter should have positive coefficients near its center, and negative coefficients in the outer periphery.

- Example: filter mask of a 3x3 sharpening filter

$$\frac{1}{9} \begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$$

Figure 32

- The filtering output pixels might be of a gray level exceeding $[0, L-1]$.
- The results of highpass filtering involve some form of scaling and/or clipping to make sure that the gray levels of the final results are within $[0, L-1]$.

(b) Derivative filters: Differentiation can be expected to have the opposite effect of averaging, which tends to blur detail in an image, and thus sharpen an image and be able to detect edges. The most common method of differentiation in image processing applications is the gradient. For a function $f(x,y)$, the gradient of f at coordinates (x',y') is defined as the vector.

$$\nabla f(x', y') = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} (x', y')$$

Figure 33

- Its magnitude can be approximated in a number of ways, which result in a number of operators such as Roberts, Prewitt and Sobel operators for computing its value.

Enhancement in the frequency domain:

We simply compute the Fourier transform of the image to be enhanced, multiply the result by a filter transfer function, and take the inverse transform to produce the enhanced image. Spatial domain: $g(x,y) = f(x,y)*h(x,y)$ Frequency domain: $G(w1,w2) = F(w1,w2)H(w1,w2)$ Lowpass filtering

- Edges and sharp transitions in the gray levels contribute to the high frequency content of its Fourier transform, so a lowpass filter smoothes an image.

- Formula of ideal LPF

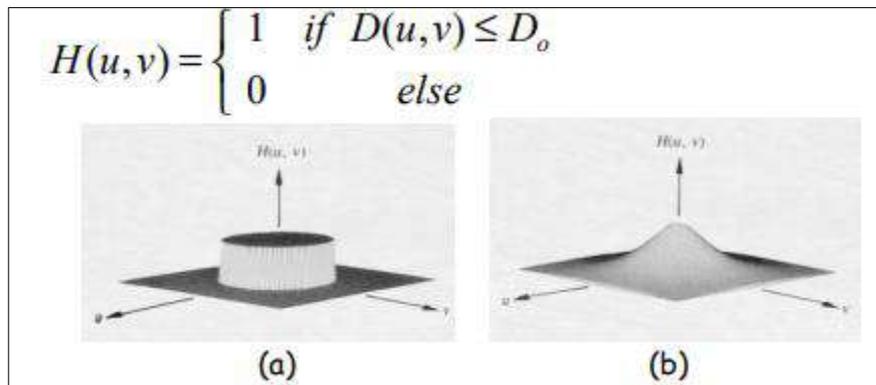


Figure 34

Highpass filtering:

- A highpass filter attenuates the low frequency components without disturbing the high frequency information in the Fourier transform domain can sharpen edges.

- Formula of ideal HPF function

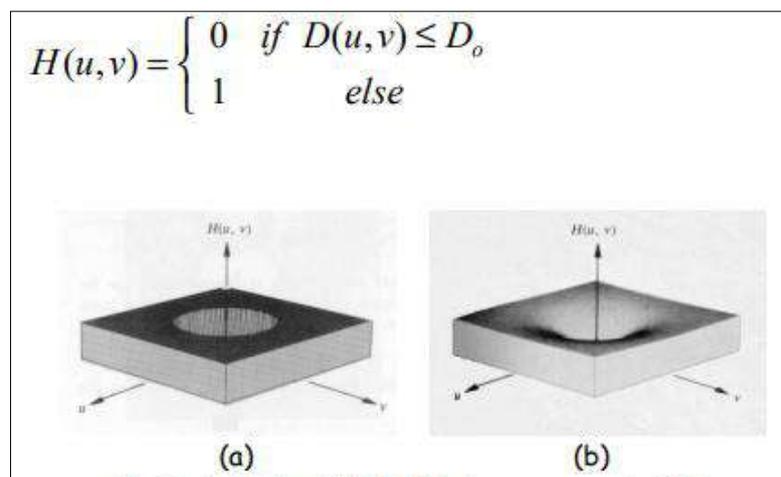


Figure 35

Pseudo colour image processing

- In automated image analysis, color is a powerful descriptor that often simplifies object identification and extraction from a scene.
- Human eye performs much better in discerning shades of colour than gray scale.
- A monochrome image can be enhanced by using colors to represent different gray levels or frequencies.

Gray level to colour transformation

- To perform 3 independent transformations on the gray level of any input pixel. The three results are then fed separately into the R, G, and B guns of a colour monitor.

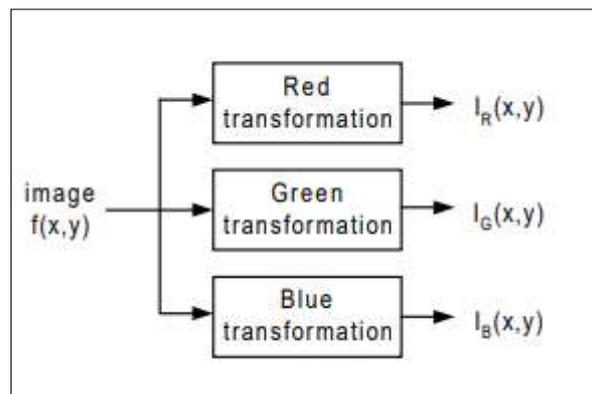


Fig. 36: Functional block diagram for pseudo-colour image processing

- This method produces a composite image whose colour content is modulated by the nature of the transformation functions.

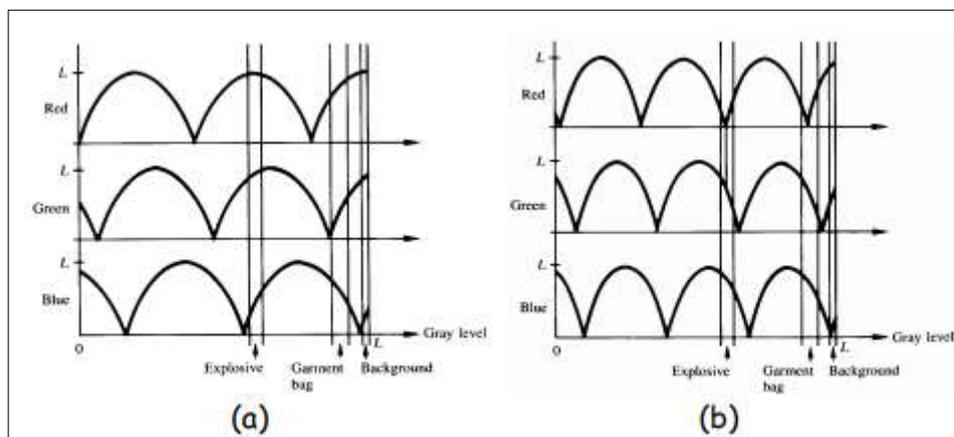
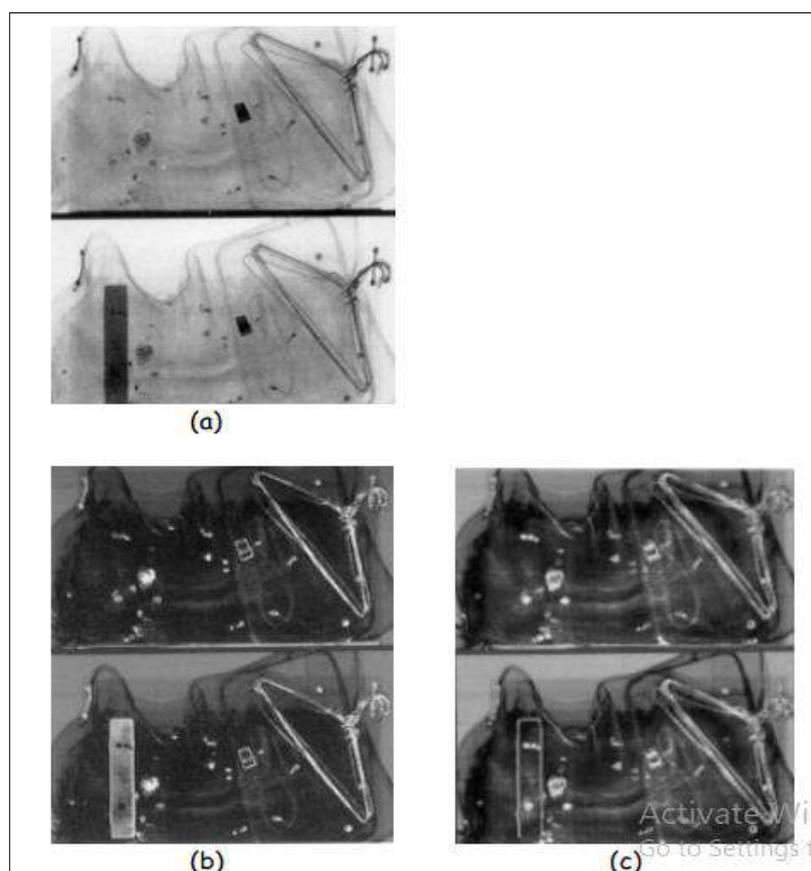


Fig. 37: Two examples of transformation functions

These sinusoidal functions contain regions of relatively constant value around the peaks as well as regions that change rapidly near the valleys.

- Changing the phase and frequency of each sinusoidal can emphasize ranges in the gray scale.
- A small change in the phase between the 3 transformations produces little change in pixels whose gray levels correspond to peaks in the sinusoidal (i.e. Case of $I(x, y) \approx I(x, y) \approx I(x, y)$ $R \approx G \approx B$).
- Pixels with gray level values in the steep section of the sinusoids are assigned a much stronger color content as a result of significant differences between the amplitudes of the 3 sinusoids caused by the phase displacement between them. (i.e. Case of $I(x, y) \approx I(x, y) \approx I(x, y)$ $R \neq G \neq B$)



Pseudo-color enhancement by using the gray-level to color transformations. (a) Original image, (b) and (c) results obtained with transformation functions (a) and (b) respectively.

Figure 38

2.3 Band Rationing

Sometimes differences in brightness values from identical surface materials are caused by various environmental factors such as topographic slope and aspect, shadows, or seasonal changes in sunlight illumination angle and intensity. These factors may limit the ability of an interpreter to correctly identify the surface materials or land use in a remotely sensed image. This ratio transformation can be applied to reduce the effects of such environmental conditions. Ratioing is the enhancement that provides unique information and differences between the spectral reflectance of surface features that are difficult to detect in a standard image. It is also useful to discriminate the soils from vegetation. A band ratio is a result of dividing each pixel of one band by the corresponding pixels of another band. Many researchers have developed many different band ratios over the years that exploit the specific reflectance properties of the surface materials. For example, vegetation has an extremely high reflectance in near infrared wavelength region, relative to red wavelengths. The number of possible ratio combinations for a multispectral sensor with K bands is $n=K(K-1)$. Thus for Thematic Mapper (for six reflectance bands), the number of ratio combinations is 30 out of which 15 are original and another 15 are reciprocal. Ratio function can be mathematically expressed as: $BV_{i,j,r} = BV_{i,j,k} / BV_{i,j,l}$ (1) Where $BV_{i,j,r}$ = output ratio value of the pixel at i th row and j th column $BV_{i,j,k}$ = brightness value at the same location for band K . $BV_{i,j,l}$ = brightness value at the same location for band l The output ratio value ranges from 0 to 255 (since the value of $BV_{i,j}$ can be zero). These new numbers can be stretched or expanded to produce images with considerable contrast variation in a black and white rendition. Certain features or materials can produce distinctive gray tones in certain ratios; TM band 3 (red) divided by 1 tends to emphasize on red or orange colored features or materials, such as natural hydrated iron oxide as light tones. Three band ratio images can be combined as color composites which highlight certain features in distinctive colors. Ratio images also reduce or eliminate the shadow effects.

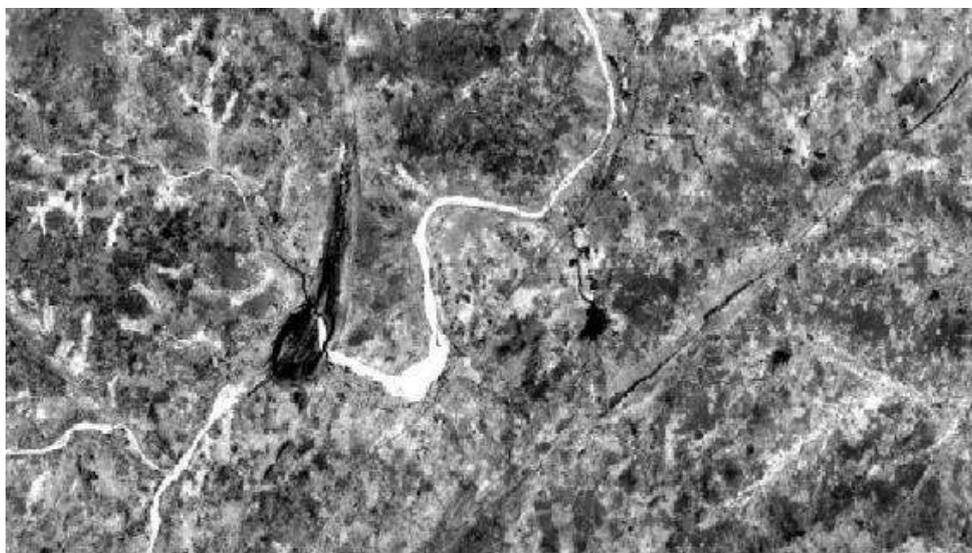


Fig. 39: Landsat TM image of Bisalpur Dam near Tonk, Rajasthan

Ratio images can be directly related to spectral properties of materials. Ratioing can be thought of as a method of enhancing minor differences between materials by defining the slope of the spectral curve between two bands. But we must understand that dissimilar materials having similar slopes but different Albedo values can be easily separable in the standard image but become inseparable in ratio images.

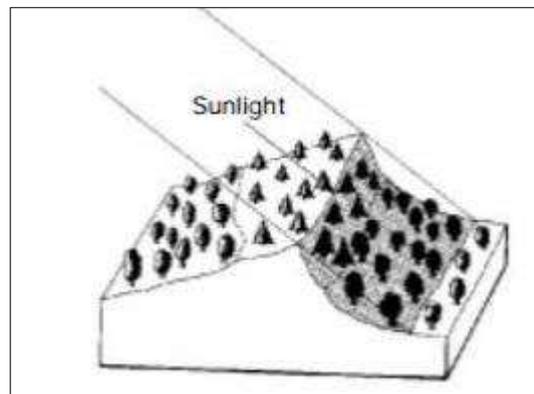


Fig. 40: Reduction of Scene Illumination effect through spectral ratioing (Source: Lillesand and Kiefer, 1993)

Selecting a proper band combination for an RGB image using Landsat Imagery:

Selecting the proper bands to use in the imagery has a huge impact to interpret the features in a particular image. The Below mentioned list describes some of the features of seven bands of Landsat Thematic Mapper Imagery for detecting different features.

Band 1 (0.45-0.52 μm): This band with short wavelength penetrates more than the other bands hence it is preferred more in the aquatic ecosystems. It is used to monitor sediments in the water, mapping coral reefs, and water depth. Since the blue light is scattered more than other bands, it has the maximum noise of all the Landsat bands.

Band 2 (0.52-0.60 μm): This has similar qualities to band1 but not extreme. It is sensitive to water turbidity differences. It matches the wavelength for the green features such as vegetation.

Band 3 (0.63-0.69 μm): It senses in the strong chlorophyll absorption region and strong reflectance region for most soils and thus it is sometimes called as chlorophyll absorption band. This band can be useful to discriminate soil and vegetation in monitoring vegetation health.

Band 4 (0.76-0.90 μm): It operates in the best spectral region in order to distinguish vegetation varieties and conditions. Since water absorbs nearly all light at this wavelength water bodies appear very dark.

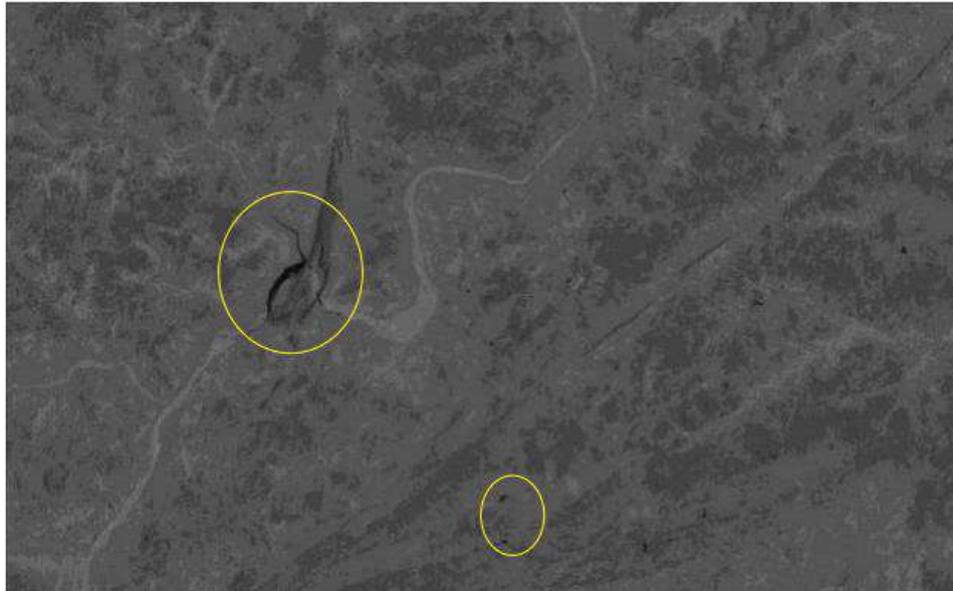


Fig. 41: Band 4 of Landsat TM (Bisalpur Dam near Tonk, Rajasthan)

Band 5 (1.55-1.75 μm): This band is very sensitive to moisture and hence is widely used to monitor vegetation and soil moisture. It has separated water body from barren and croplands (light gray tones). Since urban area and vegetation have the same spectral response it is difficult to distinguish between them.

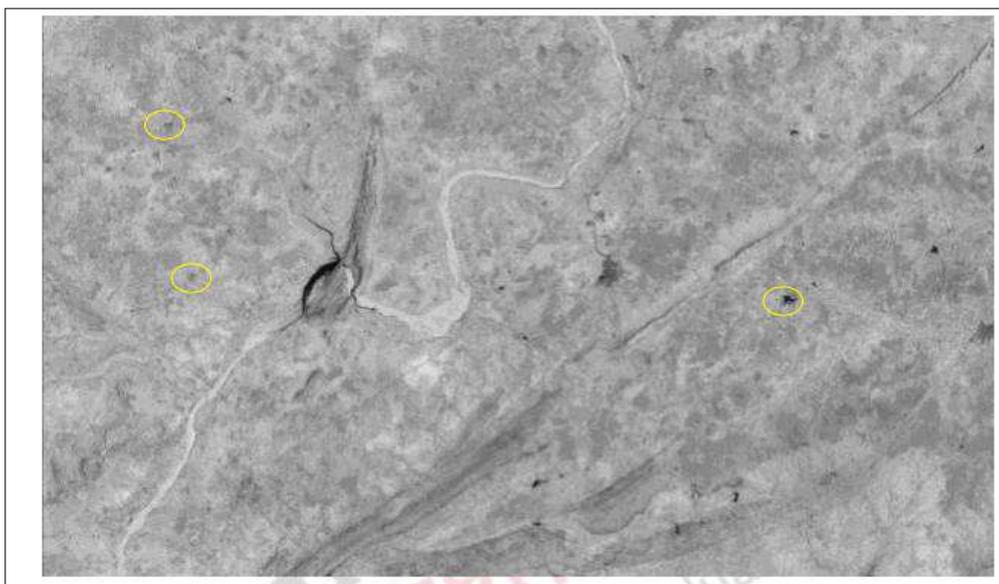


Fig. 42: Band 5 of Landsat TM not able to distinguish between settlement and vegetation since it has same gray tone for both the features.

Band 6 (10.4-12.5 μm thermal infrared): It is a thermal band and hence is used to measure surface temperature. It is primarily used for geologic applications but is sometimes used to measure plant heat

stress. It is used to distinguish clouds from soils since clouds tend to be very cold. It has half the resolution of other multispectral TM bands (i.e. 60m than 30m).

Band 7 (2.08-2.35 μm): It has the ability to sharply separate land and water features and hence is also used for vegetation moisture. Generally, band 5 is preferred for this as well as soil and geological mapping.

Band Ratios: Let us understand how different band ratio combinations are used for the interpretation of particular feature. Below is the sample image of Landsat TM of the area near Kota, Rajasthan and the natural drainage is the Chambal river flowing between Kota district.

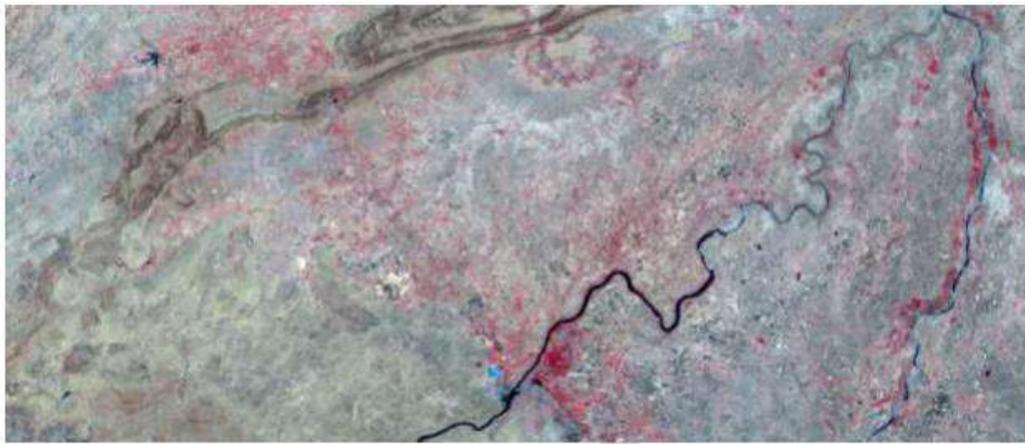


Fig. 43: Kota district, Rajasthan

Band 4/Band 3 (NIR/Red): It enhances the presence of vegetation, the brighter the tones the denser the vegetation. The below-given ratio image shows healthy vegetation by bright tone while water body is seen in dark tones. Along with these features line features such as roads are also clearly visible.

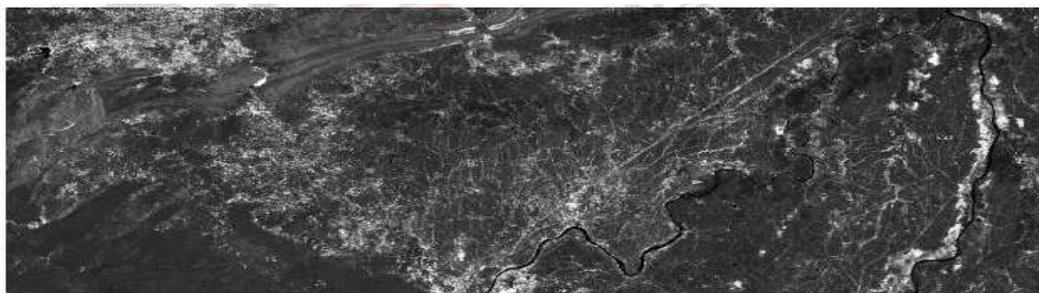


Fig. 44: Band4 / Band3 of Landsat TM showing the same area as in the previous figure.

Band 5/Band 2 (MIR/Green): This ratio can be used to detect the moisture level contained in the surface features. Darker the tone the greater the moisture content in the surface features. Hence water and soil moisture can be easily detected by this ratio.

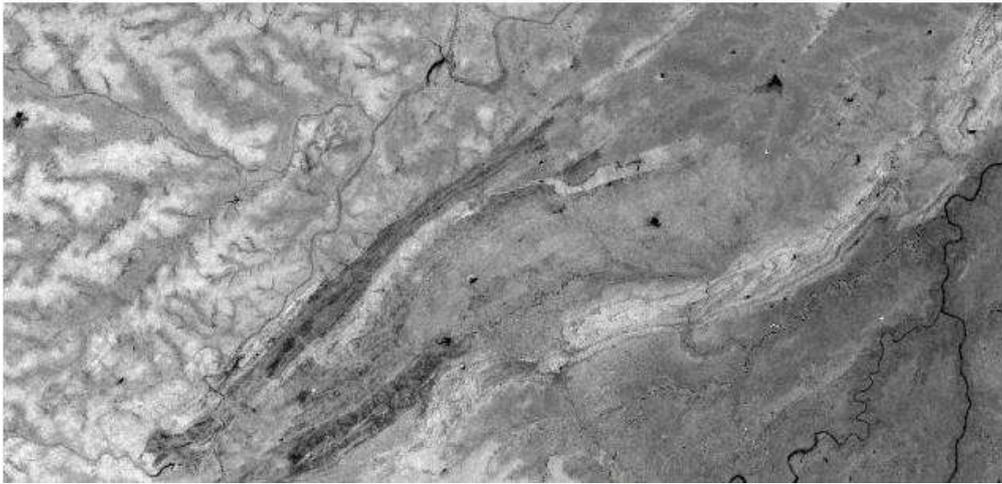


Fig. 45: Dark to bright tone represents high to low level of moisture content. Image shows the part of Bhilwara and Kota districts from left to right.

Band 3/Band 4 (Red/NIR): Urban, Water and line features like roads and drainage can be easily enhanced using this band ratio. Water bodies are seen in brighter tone while vegetation is visible in dark black tone. The level of water content is easily detectable in the imagery.

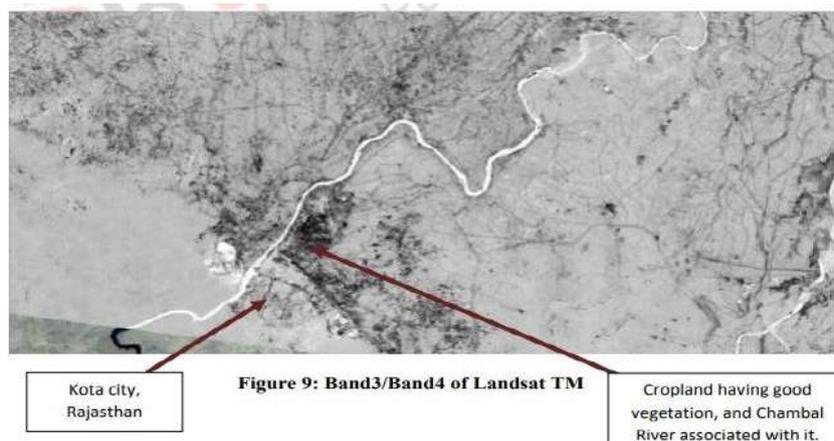


Figure 46

Apart from this, the ratio of band 2 to band 5 i.e. green by MIR is used to detect water bodies, ratio 3/1 to detect ferric iron rock, ratio 2/3 to sharply detect croplands & barren lands while ratio 4/5 enhances the water body, vegetation, and presence of moisture contents in the vegetation. In this combination, water body appears as dark tone and vegetation as light tone. Also, the ratio 7/2

separates forests from croplands but fails to discriminate between forests and water bodies as both appear as dark. It also enhances the highways, urban & built-up areas, and croplands and all of them appear in light tones. Aside from the above-listed ratios, many other combinations are also possible with the Thematic Mapper sensor (total of 30 ratios are possible). Landsat TM uses a different number of bands from those of Landsat MSS, Spot, and IRS (Indian Remote Sensing) sensors.

2.4 Density slicing

Density slicing is a digital data interpretation method used in analysis of remotely sensed imagery to enhance the information gathered from an individual brightness band. Density slicing is done by dividing the range of brightnesses in a single band into intervals, then assigning each interval to a colour. For example, in a black-and-white thermal image the temperature values in the image can be split into bands of 2°C, and each band represented by a colour of the spectrum. Therefore the temperature information in the image can be analyzed more easily because the differences between the colours are greater than the black and white and therefore it is easier to analyze the data contained in the image. Density slicing is sometimes referred to as "double threshold".

3. Supervised and Unsupervised Classification Techniques, generation of signature statistics and signature reparability reports, Generation of NDVI

Introduction:

The purpose of Image classification is to categorize all pixels in a digital image into different land use / land cover classes. Depending on the interaction between computer and interpreter during classification process, there are two types of classification. These two main categories used to achieve classified output are called Supervised and Unsupervised Classification techniques. Out of the two major methods of image classification, supervised classification is generally chosen when analysts have good knowledge of the area. In supervised classification, analyst select representative samples for each land cover class. The software then uses these "training sites" and applies them to the entire image. Supervised classification uses the spectral signature defined in the training set. There are many classification algorithms used in supervised classification such as maximum likelihood and minimum-distance classification. Supervised classification is based on the idea that a user can select sample pixels in an image that are representative of specific classes and then direct the image processing software to use these training sites as references for the classification of all other pixels in the image. Training sites are selected based on the knowledge of the user. The user also sets the bounds for how similar other pixels must be to group them together. These bounds are often set based on the spectral characteristics of the training area, plus or minus a certain increment. The user also designates the

number of classes that the image is classified into. Many analysts use a combination of supervised and unsupervised classification processes to develop final output analysis and classified maps. The schematics of the steps used in supervised classification are given below (Fig 1)

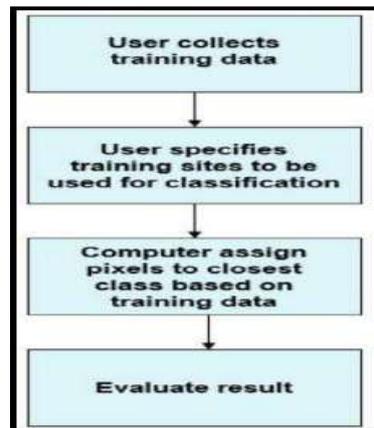


Fig.3.1: Steps of supervised classification

1. **Training Data:** Training fields are areas of known identity delineated on the digital image, usually by specifying the corner points of a rectangular or polygonal area using line and column numbers within the coordinate system of the digital image. The analyst must, of course, know the correct class for each area. Usually the analyst begins by assembling maps and aerial photographs of the area to be classified. Specific training areas are identified for each informational category following the key characteristics of training area. The objective is to identify a set of pixels that accurately represents spectral variation present within each information region.
2. **Selection of appropriate classification algorithm:** Various supervised classification algorithms may be used to assign an unknown pixel to one of a number of classes. The choice of a particular classifier or decision rule depends on the nature of the input data and the desired output. Parametric classification algorithm assumes that the observed measurement X_e for each class in each spectral band during the training phase of the supervised classification are Gaussian in nature i.e. they are normally distributed. Non parametric classification algorithm makes no such assumptions. Among the most frequently used classification algorithms are the parallelepiped, minimum distance, and maximum likelihood decision rules. An example of data samples is given below in Figure 2.

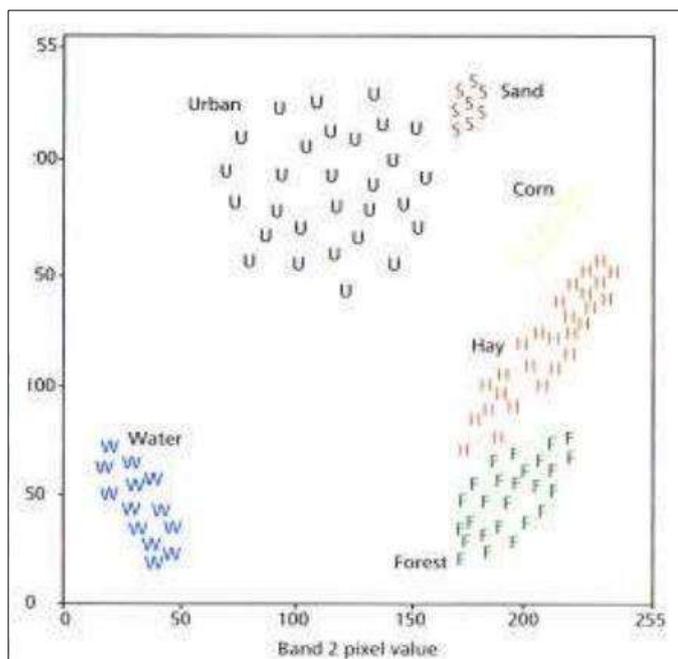


Fig. 3.2: algorithm

3. Parallelepiped Classification Algorithm: This is a widely used decision rule based on simple Boolean "and/or" logic. Training data in n spectral bands are used in performing the classification. Brightness values from each pixel of the multispectral imagery are used to produce an n-dimensional mean vector.

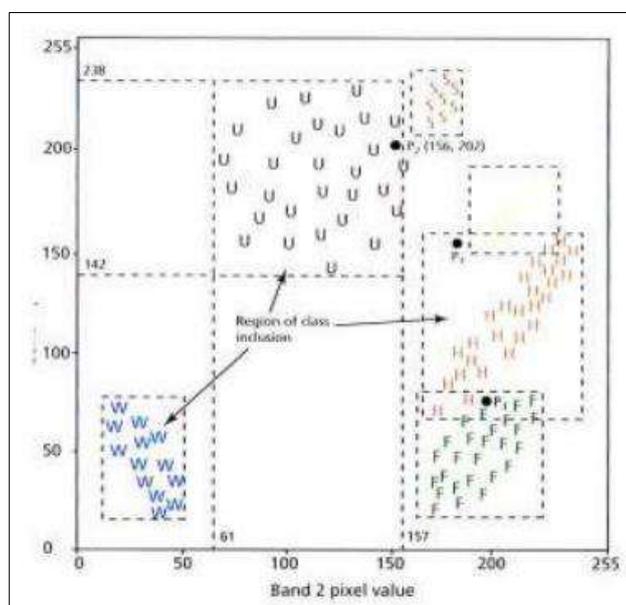


Fig. 3.3: Parallelepiped algorithm

4. Minimum Distance to Mean Classification Algorithm: This decision rule is computationally simple and commonly used. When used properly it can result in classification accuracy comparable to other more computationally intensive algorithms, such as the maximum likelihood algorithm. Like the parallelepiped algorithm, it requires that the user provide the mean vectors for each class in each band mck from the training data. To perform a minimum distance classification, a program must calculate the distance to each mean vector, mck from each unknown pixel (BVijk). It is possible to calculate this distance using Euclidean distance based on the Pythagorean Theorem (Figure 4).

The computation of the Euclidean distance from point to the mean of Class-1 measured in band relies on the equation:

$$\text{Dist} = \text{SQRT} (BV_{ijk} - m_{ck})^2 + (BV_{ijl} - m_{cl})^2$$

Where, mck and mcl represent the mean vectors for class c measured in bands k and l. Many minimum-distance algorithms analyst specify a distance or threshold from the class means beyond which a pixel will not be assigned to a category even though it is nearest to the mean of that category.

When more than two bands are evaluated in a classification, it is possible to extend the logic of computing the distance between just two points in n space using the equation:

$$n \text{ Dab} = \sum (a_i - b_i)^2 \text{ I}=1$$

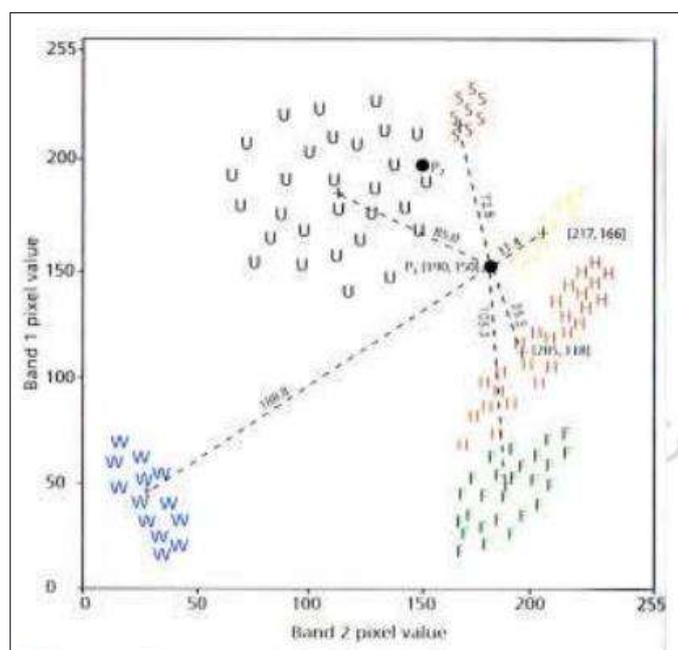


Fig. 3.4: Minimum distance to mean algorithm

5. Maximum Likelihood Classification Algorithm: The maximum likelihood decision rule assigns each pixel having pattern measurements or features X to the class c whose units are most probable or likely to have given rise to feature vector x . It assumes that the training data statistics for each class in each band are normally distributed, that is, Gaussian. In other words, training data with bi-or trimodal histograms in a single band are not ideal. In such cases, the individual modes probably represent individual classes that should be trained upon individually and labeled as separate classes. This would then produce unimodal, Gaussian training class statistics that would fulfill the normal distribution requirement. Maximum likelihood classification makes use of the statistics including the mean measurement vector, divergence, standard deviation and probability.

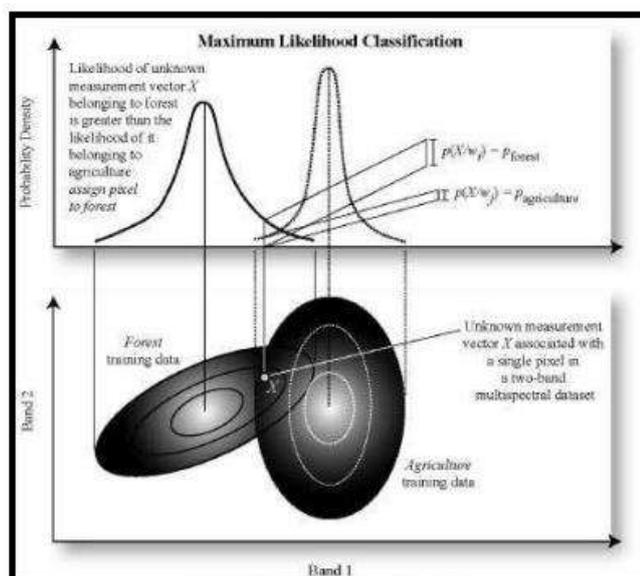


Fig. 3.5 Maximum Likelihood algorithm

6. Classification Accuracy Assessment: Quantitatively assessing classification accuracy requires the collection of some in situ data or a priori knowledge about some parts of the terrain, which can then be compared with the remote sensing derived classification map. Thus to assess classification accuracy it is necessary to compare two classification maps 1) the remote sensing derived map, and 2) assumed true map. The assumed true map may be derived from in situ investigation or quite often from the interpretation of remotely sensed data obtained at a larger scale or higher resolution.

UNSUPERVISED CLASSIFICATION

Introduction: Image classification refers to process of extraction of information from satellite image. The purpose of Image classification is to categorize all pixels in a digital image into one of several land use classes or themes. In order to make use of the multitude of digital data available from

satellite imagery, it must be processed in a manner that is suitable for the end user. For many projects this processing includes categorizing the land into its various use functions. Depending on the interaction between computer and interpreter during classification process, there are two types of classification.

These two main categories used to achieve classified output are called Supervised and Unsupervised Classification techniques unsupervised classification (commonly referred to as clustering) is an effective method of partitioning remote sensor image data in multispectral feature space and extracting land-cover information.

- Compared to supervised classification, unsupervised classification normally requires only a minimal amount of initial input from the analyst.
- This is because clustering does not normally require training data.
- The process where numerical operations are performed that search for natural groupings of the spectral properties of pixels, as examined in multispectral feature space.
- The clustering process results in a classification map consisting of m spectral classes. The analyst then attempts a posteriori to assign or transform the spectral classes into thematic information classes of interest (e.g., forest, agriculture).

In unsupervised classification, the software does most of the processing on its own generally resulting in more categories than the user is interested in. This is the point where the user has to make decisions on which categories can be grouped together into a single land use category. In either case additional image processing may be used to determine which method is better for a given situation. It must be kept in mind that maps are simple attempts to represent what actually exists in the world and are never completely accurate. The image used for demonstration is Landsat 5 Thematic Mapper (TM) scene. Ground resolution for this image is 30 meters. Landsat TM records data in seven different bandwidths. These bandwidths cover parts of the visible, infrared, and thermal infrared regions of the electromagnetic spectrum is shown in Table No. 1.

Table No. 1: bandwidths cover parts of the visible, infrared, and thermal infrared regions of the electromagnetic spectrum

Landsat 5 TM Band		
Descriptions(Jensen 2000)		
Band	Wavelength(μm)	Spectral region

1	0.45-0.52	Visible blue
2	0.52-0.60	Visible green
3	0.63-0.69	Visible red
4	0.76-0.90	reflective infrared
5	1.55-1.75	mid-infrared
6	10.40-12.50	thermal infrared
7	2.08-2.35	mid-infrared



Fig. 3.6: Perry Lake natural color

This composition is a window from LANDSAT image. This scene is centered on Perry Reservoir located in Jefferson county, north eastern Kansas and is shown in natural color.

UNSUPERVISED CLASSIFICATION: IDEA

An attempt made to classify the various land uses in Idrisi was done using unsupervised classification techniques. Unsupervised classification techniques do not require the user to specify any information about the features contained in the images. This example was conducted using the ISOCLUST module in Idrisi. With ISOCLUST, the user simply identifies which bands Idrisi should use to create the classifications, and how many classes to categorize the land cover features into. Again Landsat TM bands 1-5 and 7 were used. The resulting classified image is seen below in Fig no. 2.

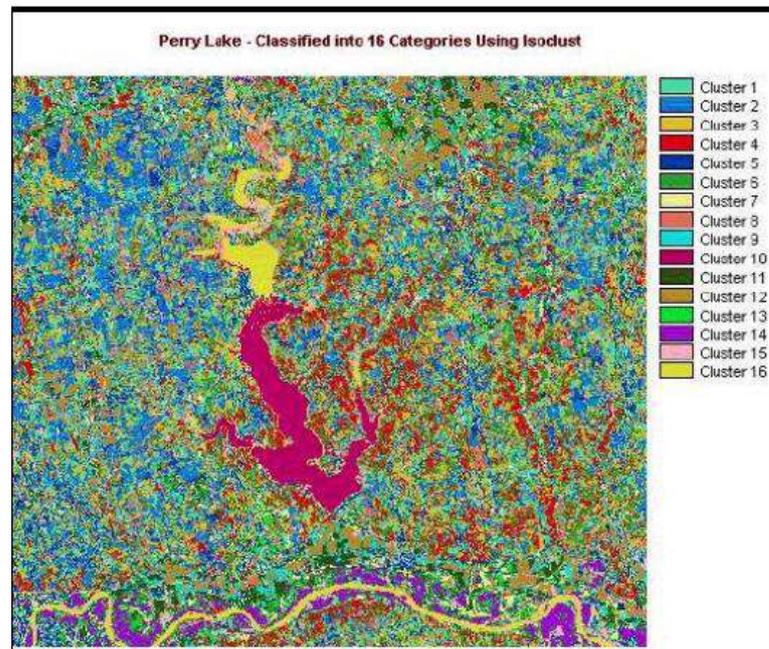


Fig. 3.7: ISOCLUST of Perry Lake

(Source: <http://academic.emporia.edu/aberjame/student/banman5/perry3.html>)

At this point, the image is difficult to interpret. Looking to the image and classified output, decisions need to be made to convert spectral groups into feature or thematic classes. To make these decisions, other materials and knowledge of the area are useful. Ground truthing what is seen in the digital image with what was actually present at the time the image was recorded makes this task more efficient and more accurate. If this knowledge is not available, scientific reasoning may be used to group the various categories together into land use categories. In the demonstration example six land cover types were identified from the original 16 categories as shown in the image.

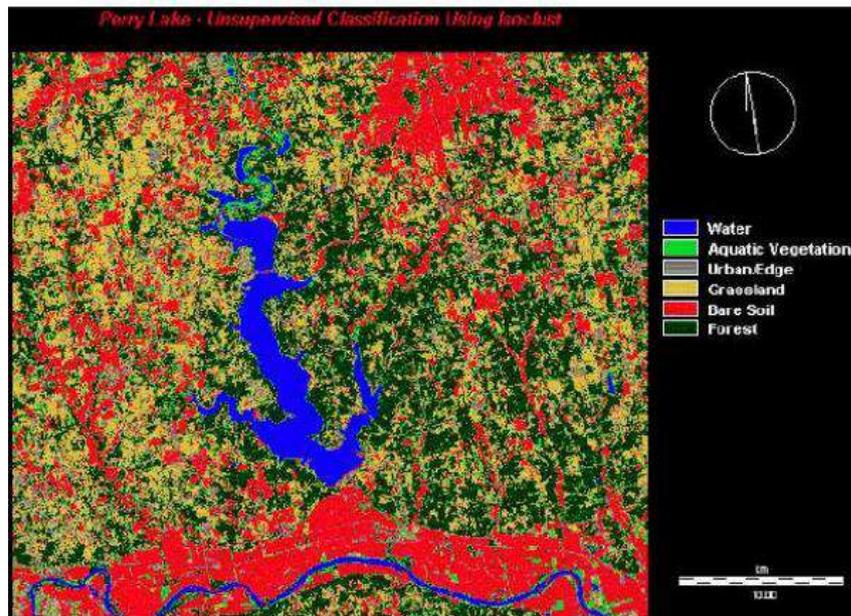


Fig. 3.8: Classified image with 6 categories

There are two major methods of unsupervised classification.

Clustering

1. K-means clustering
2. Isodata clustering

1. K-means: K-means is one of the simplest unsupervised learning algorithms that solve the well known clustering problem. The procedure follows a simple and easy way to classify a given data set through a certain number of clusters (assume k clusters) fixed a priori. The main idea is to define k centroids, one for each cluster. These centroids should be placed in a cunning way because of different location causes different result. So, the better choice is to place them as much as possible far away from each other. The next step is to take each point belonging to a given data set and associate it to the nearest centroid. When no point is pending, the first step is completed and an early grouping is done. At this point we need to re-calculate k new centroids as barycenters of the clusters resulting from the previous step. After we have these k new centroids, a new binding has to be done between the same data set points and the nearest new centroid. A loop has been generated. As a result of this loop we may notice that the k centroids change their location step by step until no more changes are done. In other words centroids do not move any more. Finally, this algorithm aims at minimizing an objective function, in this case a squared error function as in this equation.

The algorithm is composed of the following steps:

1. Place K points into the space represented by the objects that are being clustered. These points represent initial group centroids
2. Assign each object to the group that has the closest centroid.
3. When all objects have been assigned, recalculate the positions of the K centroids.
4. Repeat Steps 2 and 3 until the centroids no longer move. This produces a separation of the objects into groups from which the metric to be minimized can be calculated. Although it can be proved that the procedure will always terminate, the k -means algorithm does not necessarily find the most optimal configuration, corresponding to the global objective function minimum. The algorithm is also significantly sensitive to the initial randomly selected cluster centers. The k -means algorithm can be run multiple times to reduce this effect. K -means is a simple algorithm that has been adapted to many problem domains.

There are two schemes to use k -means to classify data

Scheme-1: One method used is to separate the data according to class labels and apply k -means to every class separately. If we have two classes, we would perform k -means twice, once for each group of data. At the end, we acquire a set of prototypes for each class. When we have a new data point, we put all of the prototypes together and find which one is closest to the new data point. This prototype is associated with a class because the prototypes are created by clustering each class of data individually. The class of this prototype is taken as the class of the new data point. Scheme 1 of using k -means clustering to the training data in each class separately, using R prototypes per class.

♣ Assign a class label to each of the $K \times R$ prototypes.

♣ Classify a new feature vector x to the class of the closest prototype. Below is a result from the textbook using this scheme (Figure No. 6). There are three classes green, red, and blue. K -means is applied using 5 prototypes for each class. We can see below that for each class, the 5 prototypes chosen are shown by filled circles.

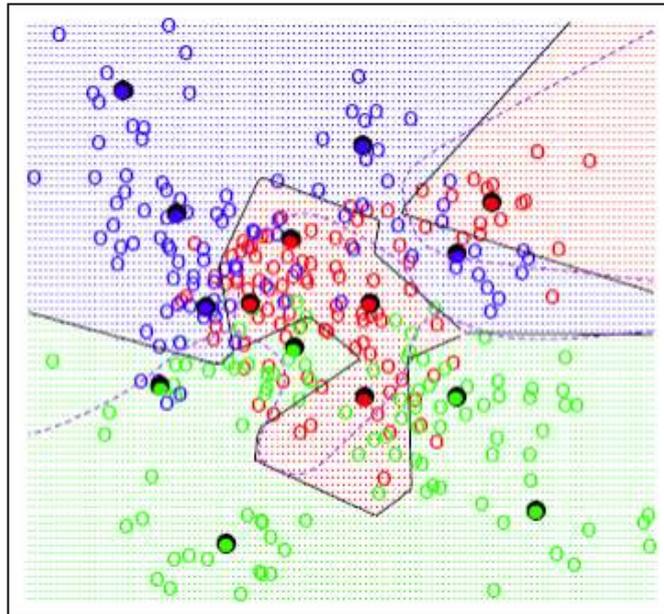


Fig. 3.9: K-means clustering

(Source: <https://onlinecourses.science.psu.edu/stat857/node/113>)

According to the classification scheme, for any new point, among these 15 prototypes, we would find the one closest to this new point. Then, depending on the color code of that prototype, the corresponding class will be assigned to the new point. The black lines are classification boundaries generated by the k-means algorithm. Specifically, these are the classification boundaries induced by the set of prototypes based on the nearest neighbor. The decision boundary between any two prototypes based on the nearest neighbor rule is linear. Every prototype occupies some region in the space. The region around each prototype is sometimes called the voronoi region and is bounded by hyperplanes. Because we have more than one prototype for each class, the classification boundary between the two classes is connected segments of the straight lines, which gives a zigzag look

Scheme- 2

The second scheme of classification by k-means is to put all the data points together and perform k-means once. There's no guarantee that points in the same group are of the same class because we conduct k-means on the class blended data. To associate a prototype with a class, we count the number of data points in each class that are assigned to this prototype. The dominant class with the most data points is associated with the prototype. During the classification of a new data point, the procedure then goes in the same way as Scheme 1.

Steps of Scheme 2:

- ♣ Apply k-means clustering to the entire training data, using M prototypes.
- ♣ For each prototype, count the number of samples from each class that are assigned to this prototype. Associate the prototype with the class that has the highest count.
- ♣ Classify a new feature x to the class of the closest prototype.

2. ISODATA Clustering: The Iterative Self-Organizing Data Analysis Technique (ISODATA) represents a comprehensive set of heuristic procedures that have been incorporated into an iterative classification algorithm. Many of the steps incorporated into the algorithm are a result of experience gained through experimentation. The ISODATA algorithm is a modification of the k-means clustering algorithm, which includes a) merging clusters if their separation distance in multispectral feature space is below a userspecified threshold, and b) rules for splitting a single cluster into two clusters.

- ISODATA is iterative because it makes a large number of passes through the remote sensing dataset until specified results are obtained, instead of just two passes.

- ISODATA does not allocate its initial mean vectors based on the analysis of pixels in the first line of data the way the two-pass chain algorithm does. Rather, an initial arbitrary assignment of all C_{max} clusters takes place along an n -dimensional vector that runs between very specific points in feature space. The region in feature space is defined using the mean and standard deviation of each band in the analysis. This method of automatically seeding the original C_{max} vectors makes sure that the first few lines of data do not bias the creation of clusters. ISODATA is self-organizing because it requires relatively little human input. Typical ISODATA algorithms normally require the analyst to specify the following criteria:

- C_{max} : the maximum number of clusters to be identified by the algorithm (e.g., 20 clusters). However, it is not uncommon for fewer to be found in the final classification map after splitting and merging take place.

- T : the maximum percentage of pixels whose class values are allowed to be unchanged between iterations. When this number is reached, the ISODATA algorithm terminates. Some datasets may never reach the desired percentage unchanged. If this happens, it is necessary to interrupt processing and edit the parameter. M : the maximum number of times ISODATA is to classify pixels and recalculate cluster mean vectors. The ISODATA algorithm terminates when this number is reached. Minimum members in a cluster (%): If a cluster contains less than the minimum percentage of members, it is deleted and the members are assigned to an alternative cluster. This also affects whether a class is going to be split (see maximum standard deviation). The default minimum percentage of members is often set to 0.01. Maximum standard deviation (σ_{max}): When the standard

deviation for a cluster exceeds the specified maximum standard deviation and the number of members in the class is greater than twice the specified minimum members in a class, the cluster is split into two clusters. The mean vectors for the two new clusters are the old class centers ± 1 of standard deviation (σ). Maximum standard deviation values between 4.5 and 7 are typical. Split separation value: If this value is changed from 0.0, it takes the place of the standard deviation in determining the locations of the new mean vectors plus and minus the split separation value. Minimum distance between clusters means (C): Clusters with a weighted distance less than this value are merged. A default of 3.0 is often used. The below figure (7) shows the schematic of iso-data clustering:

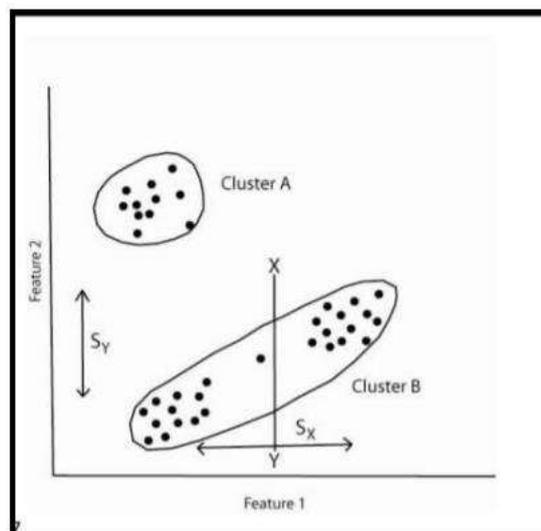


Fig. 3.10: Iso-data clustering

(Source: https://thesai.org/Downloads/Volume7No4/Paper_25

[Hyperspectral Image Classification Using Unsupervised Algorithms.pdf](#))

Preparation of NDVI with ArcMap 10.3

NDVI computes the normalized difference vegetation index (NDVI) from a multispectral image or two selected image layers. The NDVI is a standardized index allowing you to generate an image displaying greenness (relative biomass). This index takes advantage of the contrast of the characteristics of two bands from a multispectral raster dataset—the chlorophyll pigment absorptions in the red band and the high reflectivity of plant materials in the near-infrared (NIR) band.

To view the **Image Analysis** window, click **Windows** on the main menu and click **Image Analysis**.

If you are selecting a multiband dataset, you can specify which band will be used for the red and infrared inputs by doing the following:

Click the **Options** button on the **Image Analysis** window.

Click the **NDVI** tab.

Change the inputs for the **Red Band** and **Infrared Band**.

By default, the Red Band is 3 and the Infrared Band is 4. You may need to look at the metadata for your imagery to determine which bands to use. If, for example, you have a three-band color infrared image, the Infrared Band may be 1 and the Red Band may be 2.

Optionally, check **Use Wavelength**.

This option will attempt to identify the correct bands to use if the wavelength information is in the dataset. If not, then the band numbers will be used.

Optionally, check **Scientific Output**.

The Band Arithmetic function will be used instead of the NDVI function, since the Band Arithmetic function will output values between -1.0 and 1.0. Whereas, the NDVI function will scale the values to a range of 0–200, which can easily be rendered with a specific color ramp or color map.

Click **OK**.

Click a multiband raster, mosaic dataset, or image service layer in the **Image Analysis** window.

Optionally, select two single-band raster, mosaic dataset, or image service layers in the table of contents.

Click the **NDVI** button .

The output is added as a temporary layer to the table of contents. By default, a color map renderer is applied using a color map where green represents vegetation, unless the Scientific Output option is checked.

By default, the **NDVI** operation adds at least two functions to the temporary layer: the **NDVI function** and the **Colormap function**.

4. GIS: RS, GIS and GNSS.

4.1. REMOTE SENSING

Introduction

Remote sensing is an art and science of obtaining information about an object or feature without physically coming in contact with that object or feature. Humans apply remote sensing in their day-to-day business, through vision, hearing and sense of smell. The data collected can be of many forms: variations in acoustic wave distributions (e.g., sonar), variations in force distributions (e.g., gravity meter), variations in electromagnetic energy distributions (e.g., eye) etc. These remotely collected data through various sensors may be analyzed to obtain information about the objects or features under investigation. In this course we will deal with remote sensing through electromagnetic energy sensors only.

Thus, remote sensing is the process of inferring surface parameters from measurements of the electromagnetic radiation (EMR) from the Earth's surface. This EMR can either be reflected or emitted from the Earth's surface. In other words, remote sensing is detecting and measuring electromagnetic (EM) energy emanating or reflected from distant objects made of various materials, so that we can identify and categorize these objects by class or type, substance and spatial distribution [American Society of Photogrammetry, 1975].

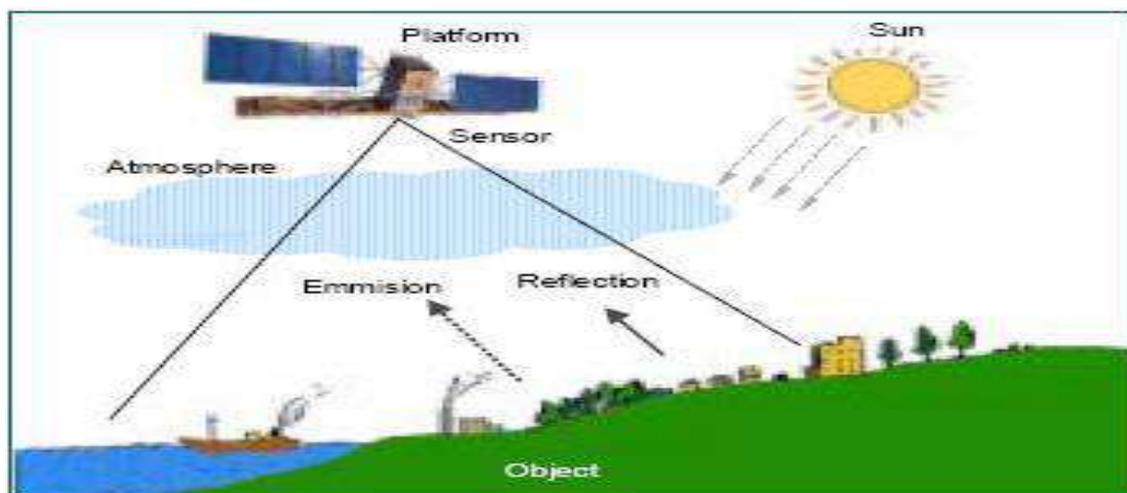


Fig. 4.1: Schematic representation of remote sensing technique

Main stages in remote sensing are the following.

A. Emission of electromagnetic radiation

- The Sun or an EMR source located on the platform
- B. Transmission of energy from the source to the object
- Absorption and scattering of the EMR while transmission
- C. Interaction of EMR with the object and subsequent reflection and emission
- D. Transmission of energy from the object to the sensor
- E. Recording of energy by the sensor
- Photographic or non-photographic sensors
- F. Transmission of the recorded information to the ground station
- G. Processing of the data into digital or hard copy image
- H. Analysis of data

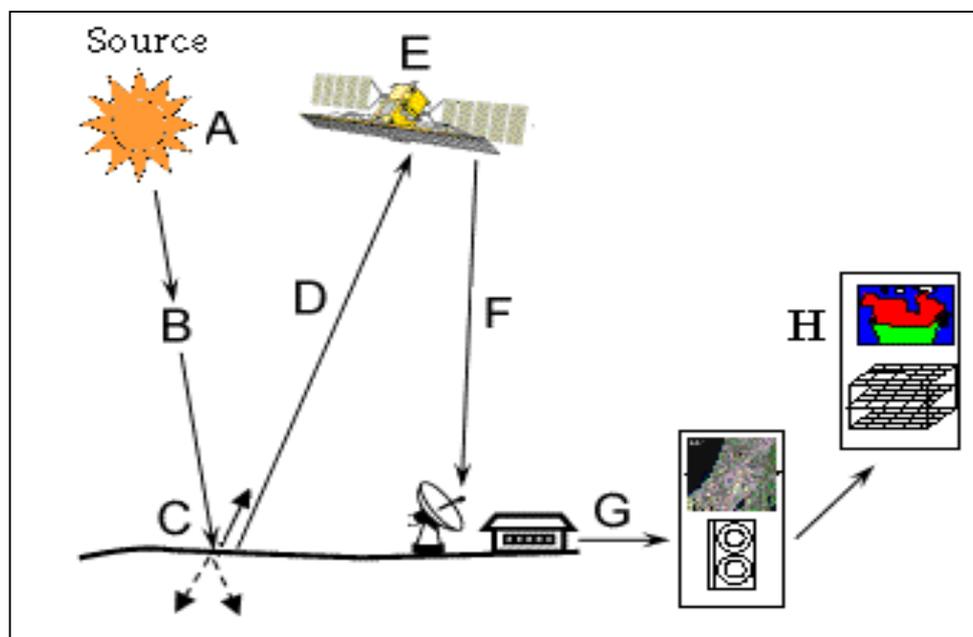


Fig. 4.2: Important stages in remote sensing

Types of Remote Sensing

Depending on the source of electromagnetic energy, remote sensing can be classified as passive or active remote sensing.

1. Passive Remote Sensing: In the case of passive remote sensing, source of energy is that naturally available such as the Sun. Most of the remote sensing systems work in passive mode using solar energy as the source of EMR. Solar energy reflected by the targets at specific wavelength bands are recorded using sensors onboard air-borne or space borne platforms. In order to ensure ample signal strength received at the sensor, wavelength / energy bands capable of traversing through the atmosphere, without significant loss through atmospheric interactions, are generally used in remote sensing.

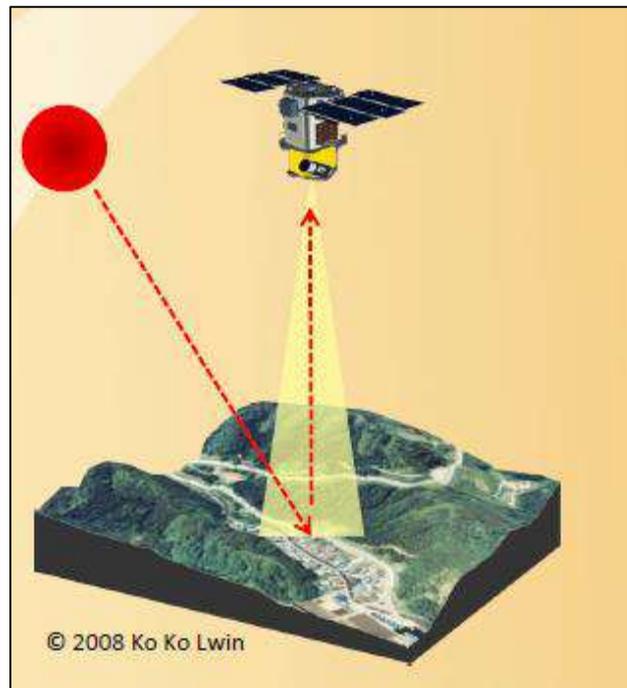


Fig. 4.3: Passive remote sensing

Any object which is at a temperature above 0o K (Kelvin) emits some radiation, which is approximately proportional to the fourth power of the temperature of the object. Thus the Earth also emits some radiation since its ambient temperature is about 300° K. Passive sensors can also be used to measure the Earth's radiance but they are not very popular as the energy content is very low.

2. Active Remote Sensing: In the case of active remote sensing, energy is generated and sent from the remote sensing platform towards the targets. The energy reflected back from the targets are recorded using sensors onboard the remote sensing platform. Most of the microwave remote sensing is done through active remote sensing.

As a simple analogy, passive remote sensing is similar to taking a picture with an ordinary camera whereas active remote sensing is analogous to taking a picture with camera having built-in flash.

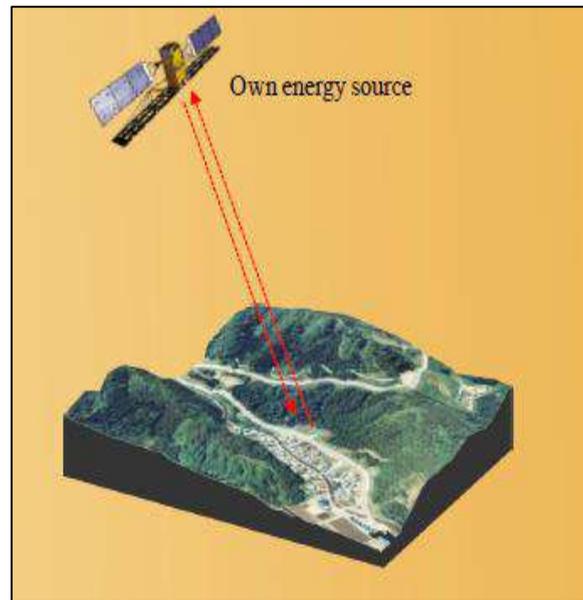


Fig. 4.4: Active remote sensing

Remote Sensing Platforms

Remote sensing platforms can be classified as follows, based on the elevation from the Earth's surface at which these platforms are placed.

- **Ground level remote sensing:** Ground level remote sensors are very close to the ground. They are basically used to develop and calibrate sensors for different features on the Earth's surface.
- **Aerial remote sensing:** Low altitude aerial remote sensing. High altitude aerial remote sensing
- **Space borne remote sensing:** Space shuttles, Polar orbiting satellites, Geo-stationary satellites.

From each of these platforms, remote sensing can be done either in passive or active mode.

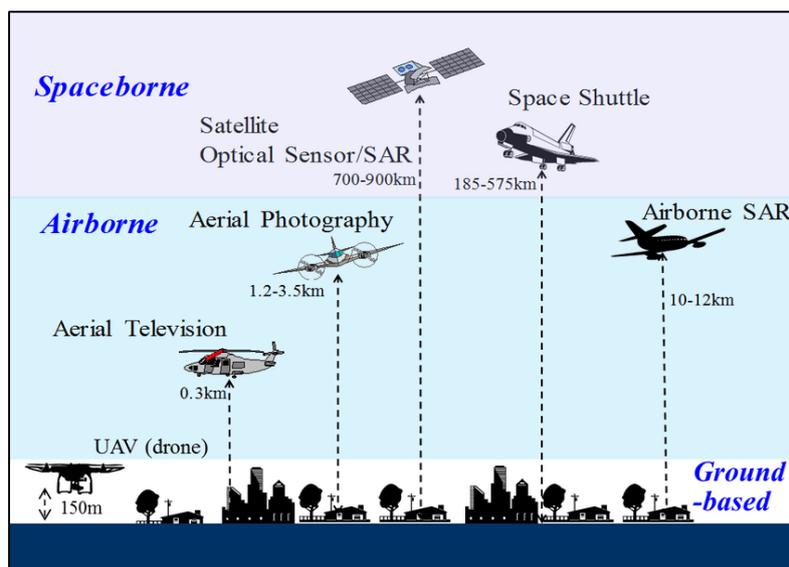


Fig. 4.5: Different types of remote sensing platforms

Types and Uses of Satellite:

Types of satellites can be classified by their orbit characteristics.

- **Type 1:** Low Earth Orbits/Satellites: Normally used in spy satellite (Military purposes)
- **Type 2:** Sun-synchronous Orbits/Satellites: a polar orbit where the satellite always crosses the Equator at the same local solar time. Most of the earth resources satellites are sun-synchronous orbit.
- **Type 3:** Geostationary Orbits/Satellites: Satellites at very high altitudes, which view the same portion of the Earth's surface at all times. Especially used in metrological applications.

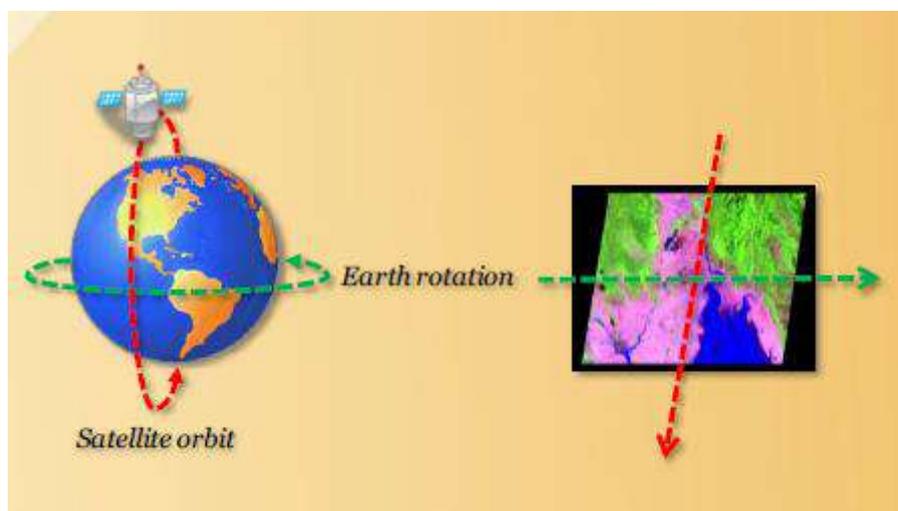


Fig. 4.6: Sun-synchronous satellite

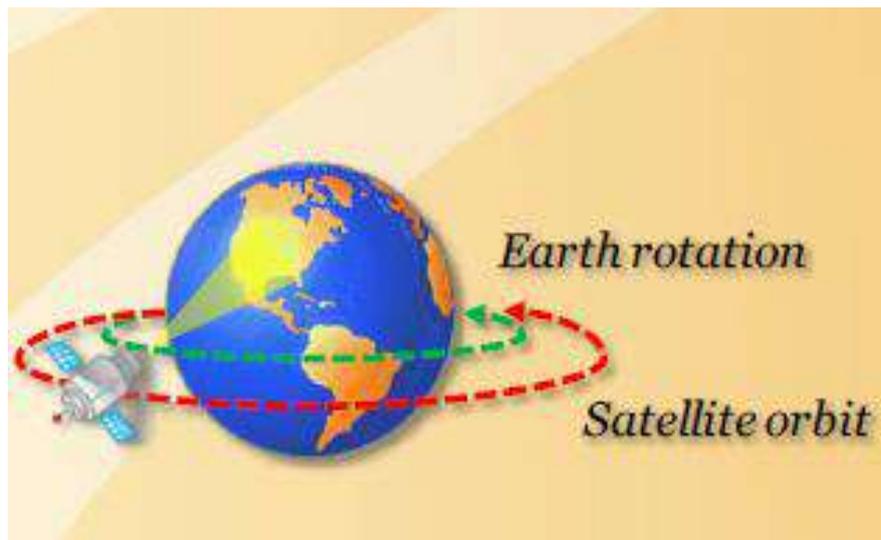


Fig. 4.7: Geostationary satellite

Remote Sensing Data Acquisition:

Electromagnetic energy or electromagnetic radiation (EMR) is the energy propagated in the form of an advancing interaction between electric and magnetic fields (Sabbins, 1978). It travels with the velocity of light. Visible light, ultraviolet rays, infrared rays, heat, radio waves, X-rays all are different forms of electro-magnetic energy.

Electro-magnetic energy (E) can be expressed either in terms of frequency (f) or wave length (λ) of radiation as follows:

$$E = h c f \text{ or } h c / \lambda \text{ (1)}$$

where h is Planck's constant (6.626×10^{-34} Joules-sec), c is a constant that expresses the celerity or speed of light (3×10^8 m/sec), f is frequency expressed in Hertz and λ is the wavelength expressed in micro meters ($1\mu\text{m} = 10^{-6}$ m).

As can be observed from equation (1), shorter wavelengths have higher energy content and longer wavelengths have lower energy content.

Distribution of the continuum of energy can be plotted as a function of wavelength (or frequency) and is known as the EMR spectrum.

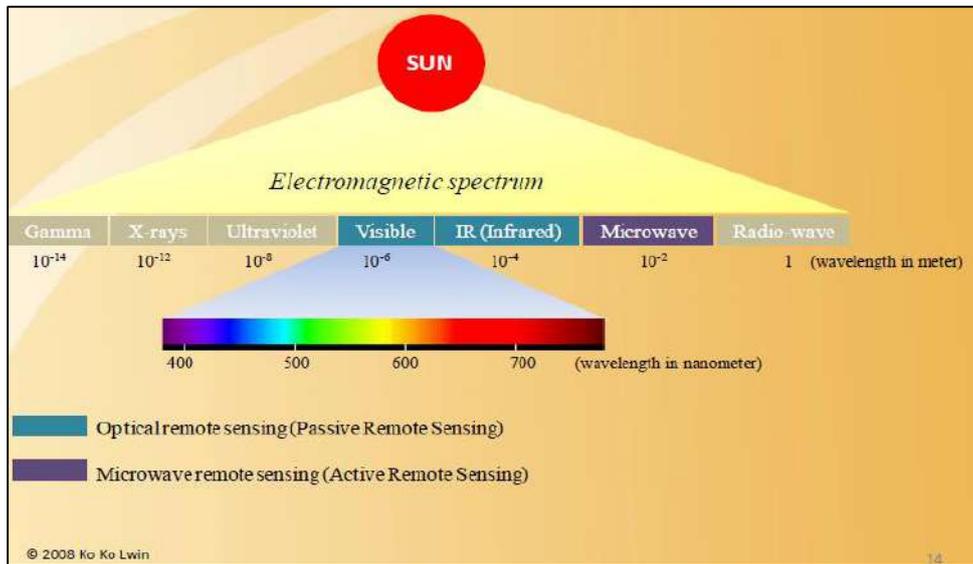


Fig. 4.8: Electromagnetic spectrum

As the energy travels from the source to the target it will come in contact with and interact with the atmosphere it passes through. This interaction may take place a second time as the energy travels from the target to the sensor.

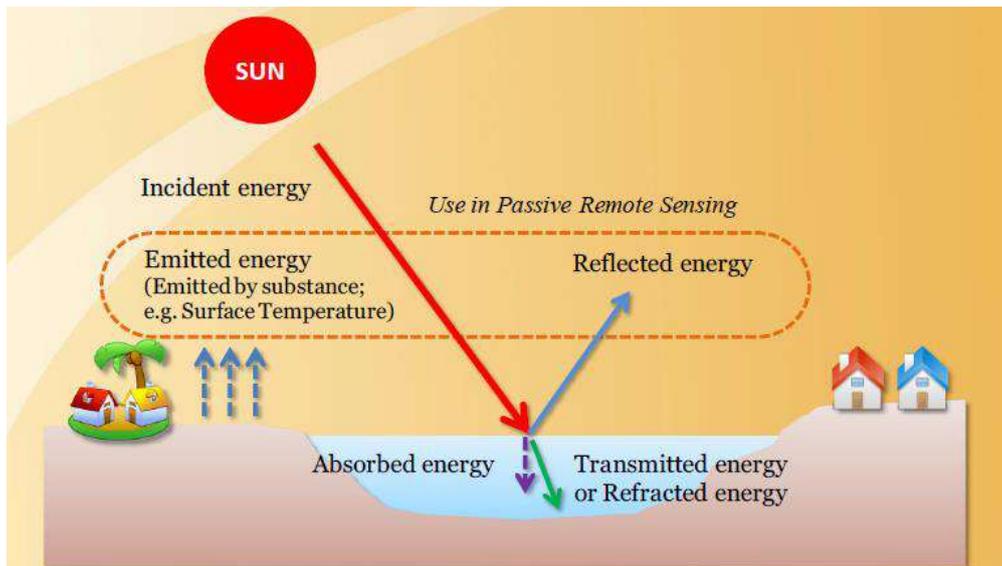


Fig. 4.9: Types of energy

Spectral Reflectance:

Different surface features reflect and absorb the sun's electromagnetic radiation in different ways. The reflectance properties of an object depend on the material and its physical and chemical state, the surface roughness as well as the angle of the sunlight. The reflectance of a material also varies with the wavelength of the electromagnetic energy. The amount of reflectance from a surface can be

measured as a function of wavelength; this is referred to as Spectral Reflectance. Spectral Reflectance is a measure of how much energy (as a percent) a surface reflects at a specific wavelength. Many surfaces reflect different amount of energy in different portions of the spectrum. These differences in reflectance make it possible to identify different earth surface features or materials by analyzing their spectral reflectance signatures. Spectral reflectance curves graph the reflectance (in percent) of objects as a function of wavelengths.

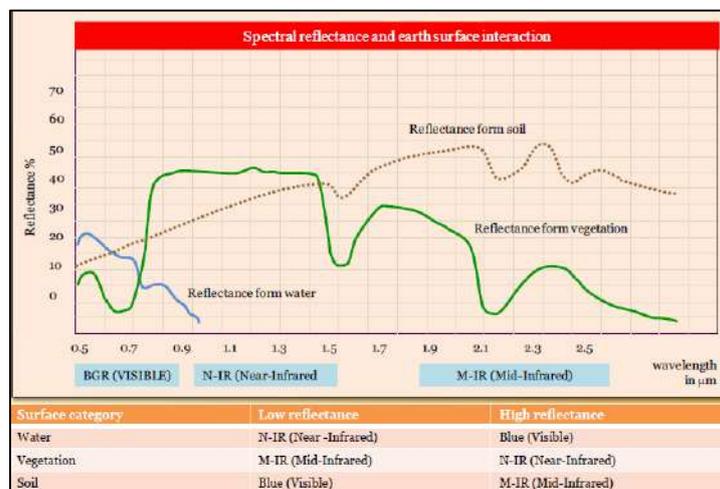


Fig. 4.10: Spectral reflectance curves

Measuring Spectral Reflectance

There are many different ways that the spectral reflectance pattern of an object or surface can be recorded. Reflectance measurements can be made in the laboratory, in the field using a field spectrometer or captured by other remote sensors including those mounted on aircraft and satellite.



Fig. 4.11: Collection of field spectral reflectance measurements for the USGS Spectral Library. Image

Credit: USGS

The level of detail of the reflectance curve (number of data points collected) depends on the spectral resolution of the sensor. Laboratory and field spectrometers typically collect hundreds of data points, measuring the percent reflectance of a material at hundreds of wavelengths. The Landsat 8 OLI (Operational Land Imager) only measures reflectance at nine different bands (or specific wavelengths) between 400 to 2500 nanometres, while advanced hyperspectral sensors like Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) measure 224 bands in the same wavelength range.

Table 4.1: Spectral Properties and Its Application

Band	Wavelength (μm)	Principal applications
B-1	0.45 - 0.52 (Blue)	This band is useful for mapping coastal water areas, differentiating between soil and vegetation, forest type mapping, and detecting cultural features.
B-2	0.52 - 0.60 (Green)	This band corresponds to the green reflectance of healthy vegetation. Also useful for cultural feature identification.
B-3	0.63 - 0.69 (Red)	This band is useful for discriminating between many plant species. It is also useful for determining soil boundary and geological boundary delineations as well as cultural features.
B-4	0.76 - 0.90 (Near-Infrared)	This band is especially responsive to the amount of vegetation biomass present in a scene. It is useful for crop identification and emphasizes soil/crop and land/water contrasts.
B-5	1.55 - 1.75 (Mid-Infrared)	This band is sensitive to the amount of water in plants, which is useful in crop drought studies and in plant health analyses. This is also one of the few bands that can be used to discriminate between clouds, snow, and ice.
B-6	10.4 - 12.5 (Thermal Infrared)	This band is useful for vegetation and crop stress detection, heat intensity, insecticide applications, and for locating thermal pollution. It can also be used to locate geothermal activity.
B-7	2.08 - 2.35 (Mid-Infrared)	This band is important for the discrimination of geologic rock type and soil boundaries, as well as soil and vegetation moisture content.

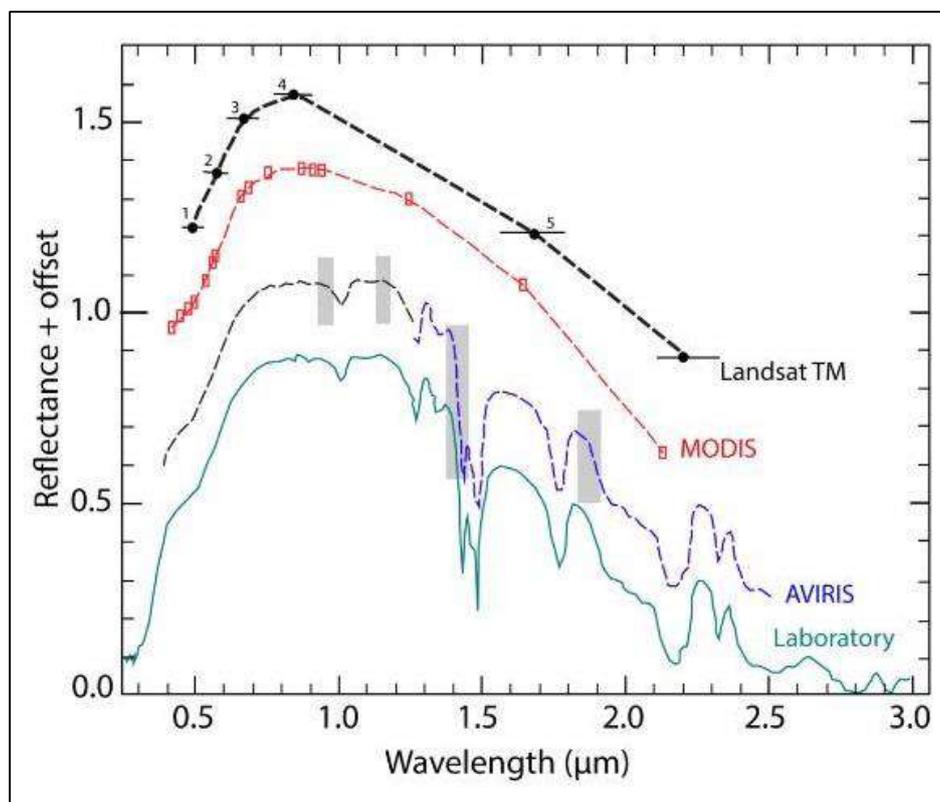


Fig. 4.12: Comparison of spectral reflectance curves for alunite from four sensors (Landsat TM, MODIS, AVIRIS and laboratory sensor) with different spectral resolutions.

Image Credit: USGS

In the above image, the Landsat Thematic Mapper (TM) has only six bands in the visible to shortwave-infrared spectrum which isn't sufficient resolution to detect any absorptions typical in minerals. AVIRIS, however, has sufficient spectral range and resolution to resolve many common absorption bands found in a wide variety of minerals and other compounds.

Spectral Reflectance to DN (Digital Number)

In fact, remote sensing data is converting of spectral reflectance value to digital number (DN) known as a pixel. Each spectral wavelength represents as a single layer in remote sensing data called "Band" or "Channel". The more bands or channels present, the more spectral properties in remote sensing data.

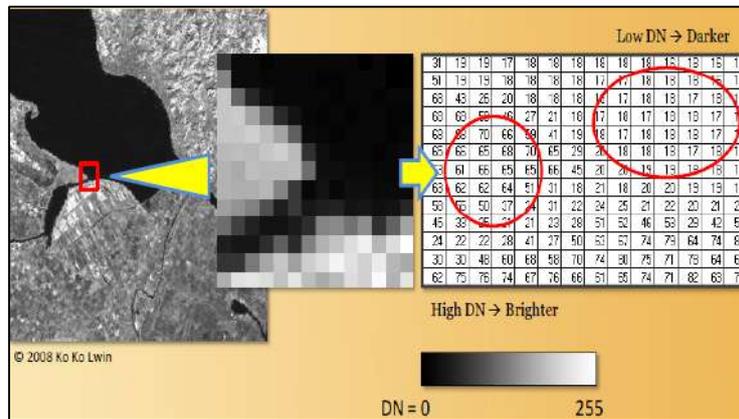


Fig. 4.13: Spectral Reflectance to DN (Digital Number)

Resolutions in Remote Sensing:

The resolution of remote sensed raster data can be characterized in several different ways. There are four primary types of "resolution" for rasters:

- Spatial
- Spectral
- Radiometric
- Temporal

Spatial Resolution: Spatial resolution is the type of resolution most people are familiar with. Spatial resolution is usually reported as the length of one side of a single pixel. For example, Landsat 8 has 30 meter spatial resolution. In other words, an image with 30 meter spatial resolution means that a single pixel represents an area on the ground that is 30 meters across.

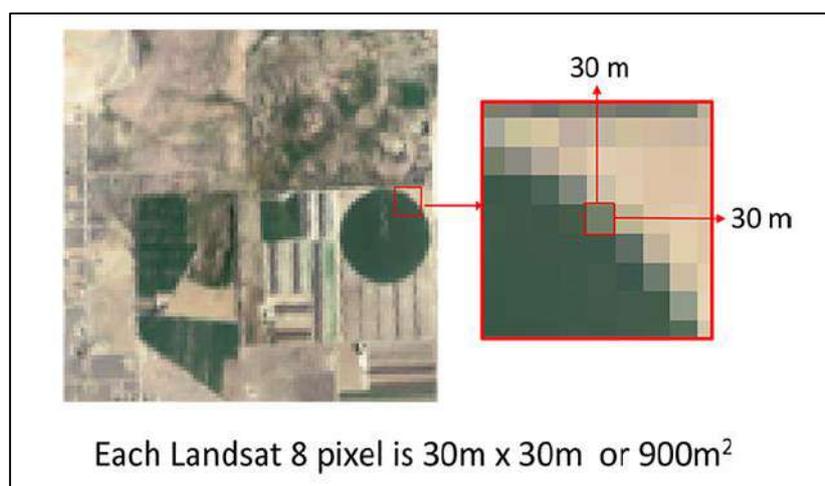


Fig. 4.14: Spatial resolution

In analog imagery (film), the dimension (or width) of the smallest object on the ground that can be distinguished in the imagery defines the spatial resolution. The spatial resolution of a raster is determined by sensor characteristics for digital imagery and film characteristics including field of view, altitude for film photography.



Sensor	Spatial Resolution
MODIS	250-1000 m
Landsat MSS	60 m
Landsat TM, ETM+, OLI	30 m
Sentinel-2	10-60 m
ASTER	15 m
Digital Globe	0.3 m (30 cm)

Fig. 4.15: Satellite Data Spatial Resolution

Spectral Resolution: Spectral resolution refers to how many spectral “bands” an instrument records. Spectral resolution is also defined by how “wide” each band is or the range of wavelengths covered by a single band. Black and white photos contain only 1 band that covers the visible wavelengths, color (RGB) images contain 3 bands and Landsat 8 has a total of 11 bands. For example, MODIS (Moderate Resolution Imaging Spectroradiometer) has greater spectral resolution than Landsat 8

because it has 36 relatively narrow bands that cover wavelengths from 0.4 to 14 micrometers. Landsat 8, on the other hand, has a total of 11 bands that cover less wavelengths and each band is wider in terms of wavelength.

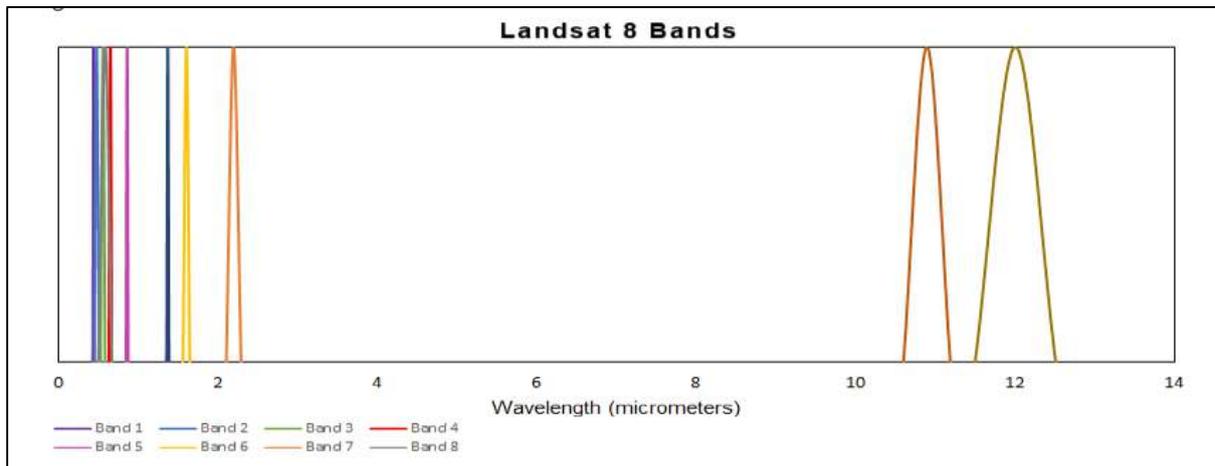


Fig. 4.16: Spectral resolution of LANDSAT

Radiometric Resolution: Radiometric resolution is how finely a satellite or sensor divides up the radiance it receives in each band. The greater the radiometric resolution the greater the range of intensities of radiation the sensor is able to distinguish and record. Radiometric resolution is typically expressed as the number of bits for each band. Traditionally 8-bit data was common in remote sensed data; newer sensors (like Landsat 8) have 16-bit data products. 8 bits = $2^8 = 256$ levels (usually 0 to 255) 16 bits = $2^{16} = 65,536$ levels (0 to 65,535).

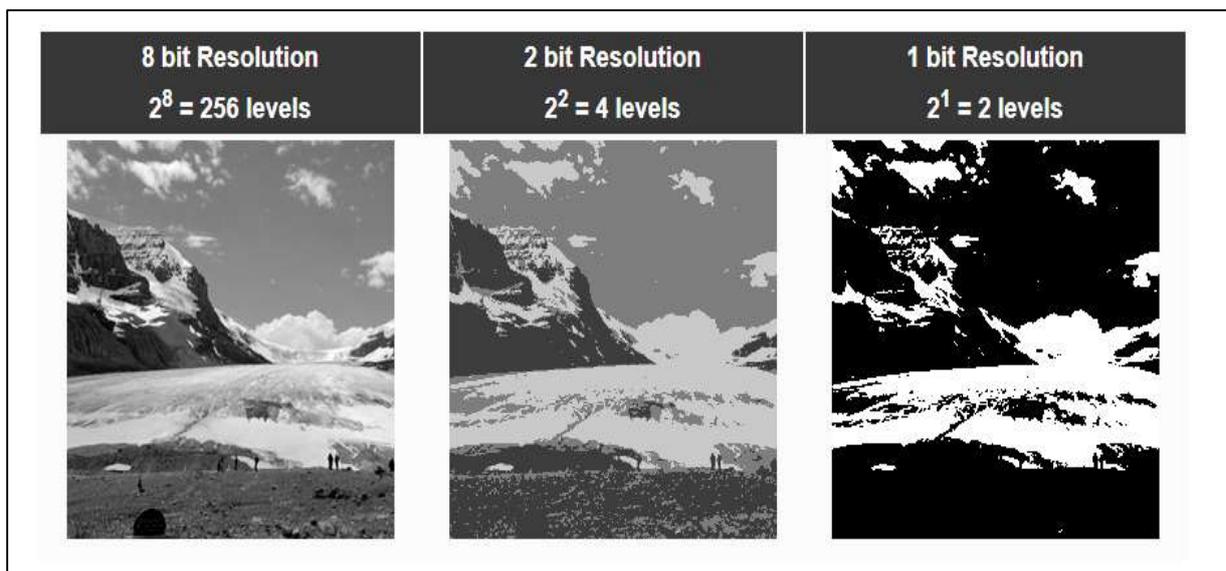


Fig. 4.17: Radiometric resolution of LANDSAT 8

Temporal Resolution: Remote sensed data represents a snap shot in time. Temporal resolution is the time between two subsequent data acquisitions for an area. This is also known as the “return time” or "revisit time". The temporal resolution depends primarily on the platform, for example, satellites usually have set return times and while sensors mounted on aircraft or unmanned aircraft systems (UAS), have variable return times. For satellites, the return time depends on the orbital characteristics (low vs. high orbit), the swath width and whether or not there is an ability to point the sensor. Landsat has a return time of approximately 16 days, while other sensors like MODIS have nearly daily return times.

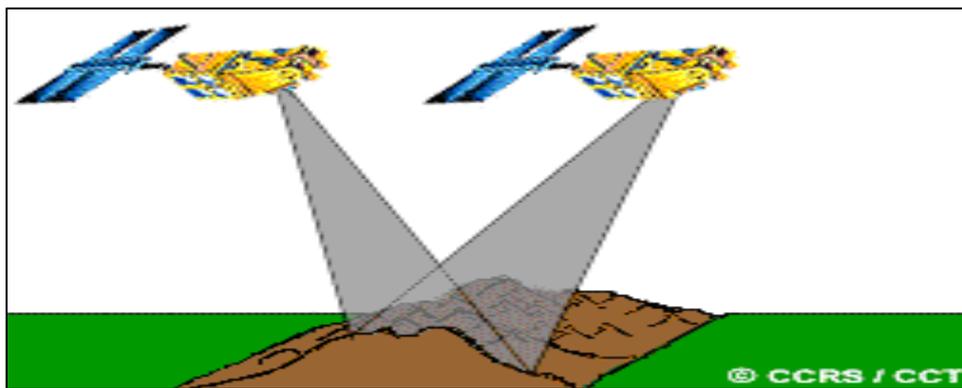


Fig. 4.18: Temporal resolution

Remote Sensing Data Processing and Analysis

Digital Image Processing is largely concerned with four basic operations: image restoration, image enhancement, image classification, image transformation. Image restoration is concerned with the correction and calibration of images in order to achieve as faithful a representation of the earth surface as possible—a fundamental consideration for all applications. Image enhancement is predominantly concerned with the modification of images to optimize their appearance to the visual system. Visual analysis is a key element, even in digital image processing, and the effects of these techniques can be dramatic. Image classification refers to the computer-assisted interpretation of images—an operation that is vital to GIS. Finally, image transformation refers to the derivation of new imagery as a result of some mathematical treatment of the raw image bands.

Image Restoration

Remotely sensed images of the environment are typically taken at a great distance from the earth's surface. As a result, there is a substantial atmospheric path that electromagnetic energy must pass through before it reaches the sensor. Depending upon the wavelengths involved and atmospheric

conditions (such as particulate matter, moisture content and turbulence), the incoming energy may be substantially modified. The sensor itself may then modify the character of that data since it may combine a variety of mechanical, optical and electrical components that serve to modify or mask the measured radiant energy. In addition, during the time the image is being scanned, the satellite is following a path that is subject to minor variations at the same time that the earth is moving underneath. The geometry of the image is thus in constant flux. Finally, the signal needs to be telemetered back to earth, and subsequently received and processed to yield the final data we receive. Consequently, a variety of systematic and apparently random disturbances can combine to degrade the quality of the image we finally receive. Image restoration seeks to remove these degradation effects. Broadly, image restoration can be broken down into the two sub-areas of radiometric restoration and geometric restoration.

Radiometric Restoration: Radiometric restoration refers to the removal or diminishment of distortions in the degree of electromagnetic energy registered by each detector. A variety of agents can cause distortion in the values recorded for image cells. Some of the most common distortions for which correction procedures exist include: uniformly elevated values, due to atmospheric haze, which preferentially scatters short wavelength bands (particularly the blue wavelengths); striping, due to detectors going out of calibration; random noise, due to unpredictable and unsystematic performance of the sensor or transmission of the data; and scan line drop out, due to signal loss from specific detectors. It is also appropriate to include here procedures that are used to convert the raw, unitless relative reflectance values (known as digital numbers, or DN) of the original bands into true measures of reflective power (radiance). See the chapter on Image Restoration in the IDRISI Guide to GIS and Image Processing Volume 2 for a more detailed discussion of radiometric restoration and how it can be implemented in IDRISI.

Geometric Restoration: For mapping purposes, it is essential that any form of remotely sensed imagery be accurately registered to the proposed map base. With satellite imagery, the very high altitude of the sensing platform results in minimal image displacements due to relief. As a result, registration can usually be achieved through the use of a systematic rubber sheet transformation process¹⁰ that gently warps an image (through the use of polynomial equations) based on the known positions of a set of widely dispersed control points. This capability is provided in IDRISI through the module RESAMPLE. With aerial photographs, however, the process is more complex. Not only are there systematic distortions related to tilt and varying altitude, but variable topographic relief leads to very irregular distortions (differential parallax) that cannot be removed through a rubber sheet transformation procedure. In these instances, it is necessary to use photogrammetric rectification to remove these distortions and provide accurate map measurements¹¹. Failing this, the central portions of high altitude photographs can be resampled with some success. RESAMPLE is a module of major

importance, and it is essential that one learn to use it effectively. Doing so also requires a thorough understanding of reference systems and their associated parameters such as datums and projections. The chapter on Georeferencing later in this volume provides an in-depth discussion of these issues.

Image Enhancement

Image enhancement is concerned with the modification of images to make them more suited to the capabilities of human vision. Regardless of the extent of digital intervention, visual analysis invariably plays a very strong role in all aspects of remote sensing. While the range of image enhancement techniques is broad, the following fundamental issues form the backbone of this area:

Digital Filtering

Content stretch

Image Classification

Image classification refers to the computer-assisted interpretation of remotely sensed images. The procedures involved are treated in detail in the IDRISI Guide to GIS and Image Processing Volume 2 chapter Classification of Remotely Sensed Imagery. This section provides a brief overview. Although some procedures are able to incorporate information about such image characteristics as texture and context, the majority of image classification is based solely on the detection of the spectral signatures (i.e., spectral response patterns) of land cover classes. The success with which this can be done will depend on two things: 1) the presence of distinctive signatures for the land cover classes of interest in the band set being used; and 2) the ability to reliably distinguish these signatures from other spectral response patterns that may be present. There are two general approaches to image classification: supervised and unsupervised. They differ in how the classification is performed. In the case of supervised classification, the software system delineates specific landcover types based on statistical characterization data drawn from known examples in the image (known as training sites). With unsupervised classification, however, clustering software is used to uncover the commonly occurring landcover types, with the analyst providing interpretations of those cover types at a later stage.

Supervised Classification

The first step in supervised classification is to identify examples of the information classes (i.e., land cover types) of interest in the image. These are called training sites. The software system is then used to develop a statistical characterization of the reflectances for each information class. This stage is often called signature analysis and may involve developing a characterization as simple as the mean or the range of reflectance on each band, or as complex as detailed analyses of the mean, variances

and covariance over all bands. Once a statistical characterization has been achieved for each information class, the image is then classified by examining the reflectances for each pixel and making a decision about which of the signatures it resembles most. There are several techniques for making these decisions, called classifiers. Most Image Processing software will offer several, based on varying decision rules. IDRISI offers a wide range of options falling into three groups depending upon the nature of the output desired and the nature of the input bands.

Unsupervised Classification

In contrast to supervised classification, where we tell the system about the character (i.e., signature) of the information classes we are looking for, unsupervised classification requires no advance information about the classes of interest. Rather, it examines the data and breaks it into the most prevalent natural spectral groupings, or clusters, present in the data. The analyst then identifies these clusters as landcover classes through a combination of familiarity with the region and ground truth visits. The logic by which unsupervised classification works is known as cluster analysis, and is provided in IDRISI primarily by the CLUSTER module. CLUSTER performs classification of composite images (created with COMPOSITE) that combine the most useful information bands. It is important to recognize, however, that the clusters unsupervised classification produces are not information classes, but spectral classes (i.e., they group together features (pixels) with similar reflectance patterns). It is thus usually the case that the analyst needs to reclassify spectral classes into information classes. For example, the system might identify classes for asphalt and cement which the analyst might later group together, creating an information class called pavement. While attractive conceptually, unsupervised classification has traditionally been hampered by very slow algorithms. However, the clustering procedure provided in IDRISI is extraordinarily fast (unquestionably the fastest on the market) and can thus be used iteratively in conjunction with ground truth data to arrive at a very strong classification. With suitable ground truth and accuracy assessment procedures, this tool can provide a remarkably rapid means of producing quality land cover data on a continuing basis. In addition to the above mentioned techniques, two modules bridge both supervised and unsupervised classifications. ISOCLUST uses a procedure known as Self-Organizing Cluster Analysis to classify up to 7 raw bands with the user specifying the number of clusters to process. The procedure uses the CLUSTER module to initiate a set of clusters that seed an iterative application of the MAXLIKE procedure, each stage using the results of the previous stage as the training sites for this supervised procedure. The result is an unsupervised classification that converges on a final set of stable members using a supervised approach (hence the notion of "self-organizing"). MAXSET is also, at its core, a supervised procedure. However, while the procedure starts with training sites that characterize individual classes, it results in a classification that includes not only these specific classes, but also

significant (but unknown) mixtures that might exist. Thus the end result has much the character of that of an unsupervised approach.

4.2 GEOGRAPHICAL INFORMATION SYSTEM (GIS)

GIS

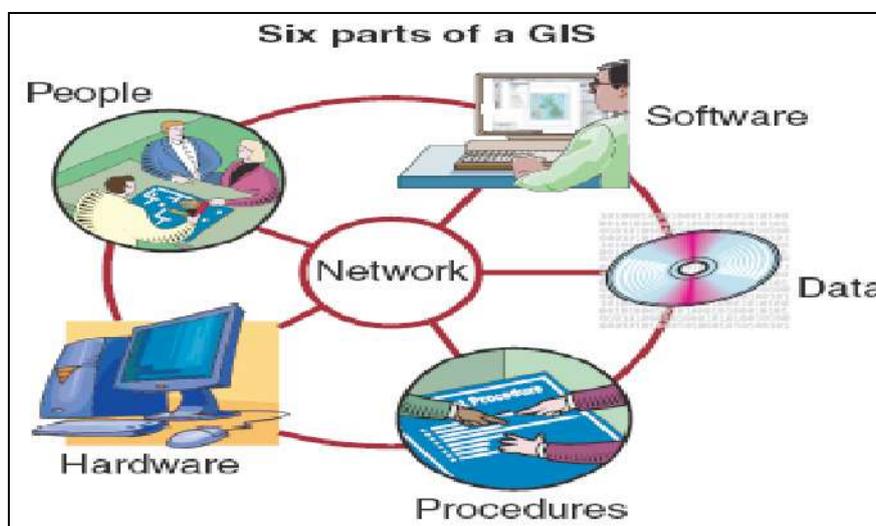
A Geographical Information System (GIS) is a system for capturing, storing, analyzing and managing data and associated attributes, which are spatially referenced to the Earth. The geographical information system is also called as a geographic information system or geospatial information system. It is an information system capable of integrating, storing, editing, analyzing, sharing, and displaying geographically referenced information. In a more generic sense, GIS is a software tool that allows users to create interactive queries, analyze the spatial information, edit data, maps, and present the results of all these operations. GIS technology is becoming essential tool to combine various maps and remote sensing information to generate various models, which are used in real time environment. Geographical information system is the science utilizing the geographic concepts, applications and systems.

Evolution of GIS:

Stage of Development	The Formative Years	Maturing Technology	GI Infrastructure
Time Frame	1960-1980	1980-Mid-1990	Mid-1990s-present
Technical environment	Mainframes and minicomputers. Proprietary software. Proprietary data structure. Mainly raster based.	Mainframes and minicomputer. Geo-relational data structures. Graphical users interface. New data acquisition technologies like GPS, Remote sensing	Workstations and PCs. Network/Internet. Open system design. Multimedia. Data Integration. Enterprise Computing. Object-Relational data model.
Major Users	Government, Universities, Military	Government, Universities, Military, Utilities, Business	Government, Universities and schools, Military, Utilities, Business, General Public
Major Application Areas	Land and Resource Management, Census, Surveying and Mapping	Land and Resource Management, Census, Surveying and Mapping, Facilities Management, Market Analysis	Land and Resource Management, Census, Surveying and Mapping, Facilities Management, Market Analysis, Utilities, Geographic Data Browsing

Components of GIS

GIS enables the user to input, manage, manipulate, analyze, and display geographically referenced data using a computerized system. To perform various operations with GIS, the components of GIS such as software, hardware, data, people and methods are essential.



A. Software

GIS software provides the functions and tools needed to store, analyze, and display geographic information. Key software components are (a) a database management system (DBMS) (b) tools for the input and manipulation of geographic information (c) tools that support geographic query, analysis, and visualization (d) a graphical user interface (GUI) for easy access to tools. GIS software is either commercial software or software developed on Open Source domain, which are available for free. However, the commercial software is copyright protected, can be expensive and is available in terms number of licensees.

Currently available commercial GIS software includes Arc/Info, Intergraph, MapInfo, Gram++ etc. Out of these Arc/Info is the most popular software package. And, the open source software is AMS/MARS etc.

B. Hardware

Hardware is the computer on which a GIS operates. Today, GIS runs on a wide range of hardware types, from centralized computer servers to desktop computers used in stand-alone or networked configurations.

C. Data

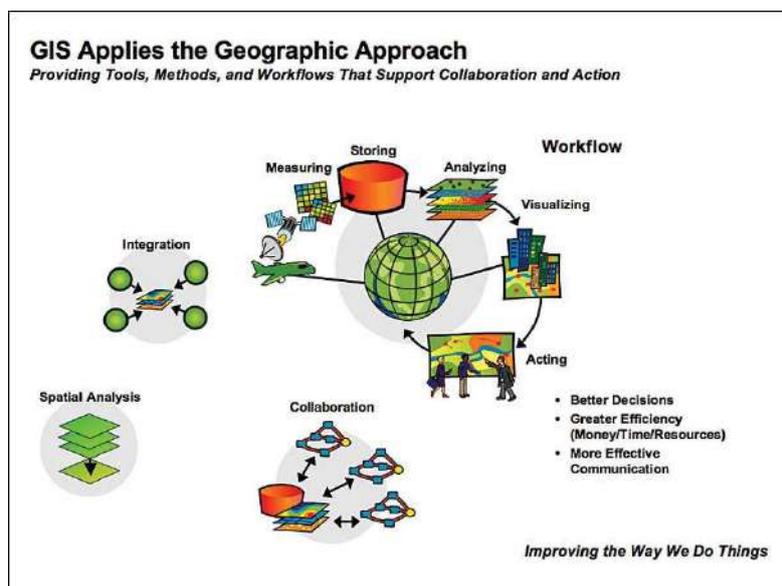
The most important component of a GIS is the data. Geographic data or spatial data and related tabular data can be collected in-house or bought from a commercial data provider. Spatial data can be in the form of a map/remotely-sensed data such as satellite imagery and aerial photography. These data forms must be properly geo referenced (latitude/longitude). Tabular data can be in the form attribute data that is in some way related to spatial data. Most GIS software comes with inbuilt Database Management Systems (DBMS) to create and maintain a database to help organize and manage data.

D. Users

GIS technology is of limited value without the users who manage the system and to develop plans for applying it. GIS users range from technical specialists who design and maintain the system to those who use it to help them do their everyday work.

Functions of GIS

General-purpose GIS software performs six major tasks such as input, manipulation, management, query and analysis, Visualization.



1. Input

The important input data for any GIS is digitized maps, images, spatial data and tabular data. The tabular data is generally typed on a computer using relational database management system software. Before geographic data can be used in a GIS it must be converted into a suitable digital format. The DBMS system can generate various objects such as index generation on data items, to speed up the information retrieval by a query. Maps can be digitized using a vector format in which the actual map

points, lines, and polygons are stored as coordinates. Data can also be input in a raster format in which data elements are stored as cells in a grid structure (the technology details are covered in following section). The process of converting data from paper maps into computer files is called digitizing. Modern GIS technology has the capability to automate this process fully for large projects; smaller jobs may require some manual digitizing. The digitizing process is labor intensive and time-consuming, so it is better to use the data that already exist. Today many types of geographic data already exist in GIS-compatible formats. These data can be obtained from data suppliers and loaded directly into a GIS.

2. Manipulation

GIS can store, maintain, distribute and update spatial data associated text data. The spatial data must be referenced to geographic coordinate systems (latitude/longitude). The tabular data associated with spatial data can be manipulated with help of data base management software. It is likely that data types required for a particular GIS project will need to be transformed or manipulated in some way to make them compatible with the system. For example, geographic information is available at different scales (scale of 1:100,000; 1:10,000; and 1:50,000). Before these can be overlaid and integrated they must be transformed to the same scale. This could be a temporary transformation for display purposes or a permanent one required for analysis. And, there are many other types of data manipulation that are routinely performed in GIS. These include projection changes, data aggregation, generalization and weeding out unnecessary data.

3. Management

For small GIS projects it may be sufficient to store geographic information as computer files. However, when data volumes become large and the number of users of the data becomes more than a few, it is advised to use a database management system (DBMS) to help store, organize, and manage data. A DBMS is a database management software package to manage the integrated collection of database objects such as tables, indexes, query, and other procedures in a database. There are many different models of DBMS, but for GIS use, the relational model database management systems will be highly helpful. In the relational model, data are stored conceptually as a collection of tables and each table will have the data attributes related to a common entity. Common fields in different tables are used to link them together with relations. Because of its simple architecture, the relational DBMS software has been used so widely. These are flexible in nature and have been very wide deployed in applications both within and without GIS.

4. Query

The stored information either spatial data or associated tabular data can be retrieved with the help of Structured Query Language (SQL). Depending on the type of user interface, data can be queried using the SQL or a menu driven system can be used to retrieve map data.

For example, you can begin to ask questions such as:

- Where all the soils are suitable for sunflower crop?
- What is the dominant soil type for Paddy?
- What is the groundwater available position in a village/block/district? Both simple and sophisticated queries utilizing more than one data layer can provide timely information to officers, analysts to have overall knowledge about situation and can take a more informed decision.

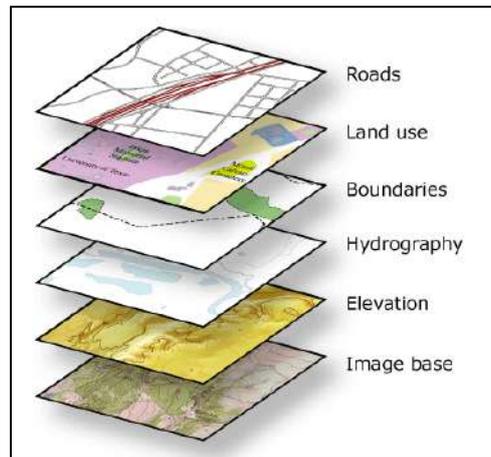
5. Analysis

GIS systems really come into their own when they are used to analyze geographic data. The processes of geographic analysis often called spatial analysis or geo-processing uses the geographic properties of features to look for patterns and trends, and to undertake "what if" scenarios. Modern GIS have many powerful analytical tools to analyze the data.

The following are some of the analysis which is generally performed on geographic data.

a. Overlay Analysis

The integration of different data layers involves a process called overlay. At its simplest, this could be a visual operation, but analytical operations require one or more data layers to be joined physically. This overlay, or spatial join, can integrate data on soils, slope, and vegetation, or land ownership. For example, data layers for soil and land use can be combined resulting in a new map which contains both soil and land use information. This will be helpful to understand the different behavior of the situation on different parameters.

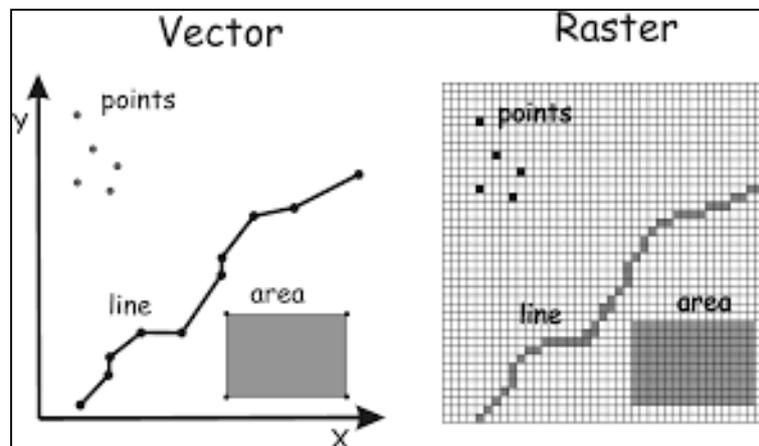


b. Proximity Analysis

GIS software can also support buffer generation that involves the creation of new polygons from points, lines, and polygon features stored in the database. For example, to know answer to questions like; How much area covered within 1 km of water canal? What is area covered under different crops? And, for watershed projects, where is the boundary or delineation of watershed, slope, water channels, different types water harvesting structures are required, etc. Visualization GIS can provide hardcopy maps, statistical summaries, modeling solutions and graphical display of maps for both spatial and tabular data. For many types of geographic operation the end result is best visualized as a map or graph. Maps are very efficient at storing and communicating geographic information. GIS provides new and exciting tools to extend the art of visualization of output information to the users.

c. Data representation

GIS data represents real world objects such as roads, land use, elevation with digital data. Real world objects can be divided into two abstractions: discrete objects (a house) and continuous fields (rain fall amount or elevation). There are two broad methods used to store data in a GIS for both abstractions: Raster and Vector.



A. Raster

A raster data type is, in essence, any type of digital image. Anyone who is familiar with digital photography will recognize the pixel as the smallest individual unit of an image. A combination of these pixels will create an image, distinct from the commonly used scalable vector graphics, which are the basis of the vector model. While a digital image is concerned with the output as representation of reality, in a photograph or art transferred to computer, the raster data type will reflect an abstraction of reality. Aerial photos are one commonly used form of raster data, with only one purpose, to display a detailed image on a map or for the purposes of digitization. Other raster data sets will contain information regarding elevation, a DEM (digital Elevation Model), or reflectance of a particular wavelength of light.

Digital elevation model, map, and vector data, Raster data type consists of rows and columns of cells each storing a single value. Raster data can be images (raster images) with each pixel containing a color value. Additional values recorded for each cell may be a discrete value, such as land use, a continuous value, such as temperature, or a null value if no data is available. While a raster cell stores a single value, it can be extended by using raster bands to represent RGB (red, green, blue) colors, color maps (a mapping between a thematic code and RGB value), or an extended attribute table with one row for each unique cell value. The resolution of the raster data set is its cell width in ground units.

Raster data is stored in various formats; from a standard file-based structure of TIF, JPEG formats to binary large object (BLOB) data stored directly in a relational database management system (RDBMS) similar to other vector-based feature classes. Database storage, when properly indexed, typically allows for quicker retrieval of the raster data but can require storage of millions of significantly sized records.

B. Vector

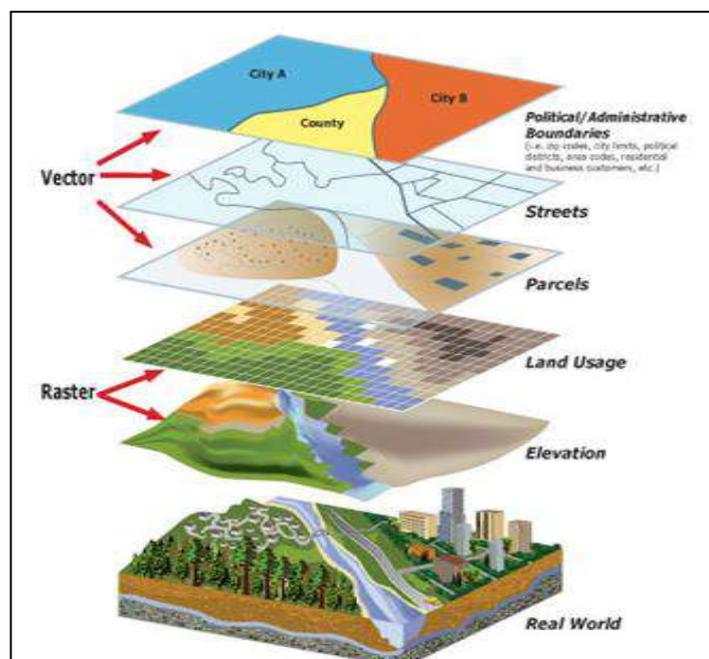
A simple vector map, using each of the vector elements: points for wells, lines for rivers, and a polygon for the lake. In a GIS, geographical features are often expressed as vectors, by considering those features as geometrical shapes. In the popular ESRI Arc series of programs, these are explicitly called shape files. Different geographical features are best expressed by different types of geometry:

Points: Zero-dimensional points are used for geographical features that can best be expressed by a single grid reference; in other words, simple location. For example, the location of wells, peak elevations, features of interest or trailheads. Points convey the least amount of information of these file types.

Lines or polylines: One-dimensional lines or polylines are used for linear features such as rivers, roads, railroads, trails, and topographic lines.

Polygons: Two-dimensional polygons are used for geographical features that cover a particular area of the earth's surface. Such features may include lakes, park boundaries, buildings, city boundaries, or land uses.

Polygons convey the most amount of information of the file types. Each of these geometries is linked to a row in a database that describes their attributes. For example, a database that describes lakes may contain a lake's depth, water quality, pollution level. This information can be used to make a map to describe a particular attribute of the dataset. For example, lakes could be colored depending on level of pollution. Different geometries can also be compared. For example, the GIS could be used to identify all wells (point geometry) that are within 1-mile (1.6 km) of a lake (polygon geometry) that has a high level of pollution.



Use of Raster and Vector map layers in GIS.

Applications of Geographic Information System (GIS)

At a broad range of scales, map has become increasingly important as legal document that convey land ownership and jurisdictional boundaries, as tools to support decision making (For example, in urban planning). The GIS provides the ability to completely model utility networks, such as those supplying water, power and telecommunications to a large number of consumers. Such a system may operate at variety of scales, modeling service Connections to consumers, service districts as well as detailed facilities inventories and layouts, such as transformers, valves conduits, and schematic diagrams. Applications of the GIS are listed below:

1. ❖ **Navigation (Routing and Scheduling):** Web-based navigation maps encourage safe navigation in waterway. Ferry paths and shipping routes are identified for the better routing. ArcGIS supports safe navigation system and provides accurate topographic and hydrographic data. Recently DNR, s Coastal Resources Division began the task of locating, documenting, and cataloging these no historic wrecks with GIS. This division is providing public information that makes citizens awareness of these vessel locations through web map. The web map will be regularly updated to keep the boating public informed of these coastal hazards to minimize risk of collision and injury.
2. ❖ **Surveying:** Surveying is the measurement of location of objects on the earth's surfaces. Land survey is measuring the distance and angles between different points on the earth surface. An increasing number of national and governments and regional organizations are using GNSS measurements. GNSS is used for topographic surveys where a centimeter level accuracy is provided. These data can be

incorporated in the GIS system. GIS tools can be used to estimate area and also, digital maps can prepare.

3. ❖ **GIS Applications in Geology:** Geologists use GIS in a various applications. The GIS is used to study geologic features, analyze soils and strata, assess seismic information, and or create three dimensional (3D) displays of geographic features. GIS can be also used to analyze rock information characteristics and identifying the best dam site location.
4. ❖ **GIS for Planning and Community Development:** GIS helps us to better understand our world so we can meet global challenges. Today GIS technology is advancing rapidly, providing many new capabilities and innovations in planning. By applying known part of science and GIS to solve unknown part, that helps to enhance the quality of life and achieve a better future. Creating and applying GIS tools and knowledge allow us integrating geographic intelligence into how we think and behave.
5. ❖ **Worldwide Earthquake Information System:** One of the most frightening and destructive phenomena of nature is the occurrence of an earthquake. There is a need to have knowledge regarding the trends in earthquake occurrence worldwide. A GIS based user interface system for querying on earthquake catalogue will be of great help to the earthquake engineers and seismologists in understanding the behavior pattern of earthquake in spatial and temporal domain.
6. ❖ **Energy Use Tracking and Planning:** GIS is a valuable tool that helps in the planning organizing and subsequent growth in the energy and utilities industries. The effective management of energy systems is a complex challenge. GIS has enormous potential for planning, design and maintenance of facility. Also it provide improved services and that too cost effectively.
7. ❖ **GIS for Fisheries and Ocean Industries:** GIS tools add value and the capability to ocean data. ArcGis is used to determine the spatial data for a fisheries assessment and management system. It is extensively used in the ocean industry area and we get accurate information regarding various commercial activities. To enhance minimizing cost for the fishing industry. Also it can determine the location of illegal fishing operations.
8. ❖ **Traffic Density Studies:** GIS can effectively use for the management of traffic problems. Today's population along with the road traffic is increasing exponentially. The advantage of GIS makes it an attractive option to be used to face the emerging traffic problems. By creating an extensive database that has all the traffic information such as speed data, road geometry, traffic flow and other spatial data and processing this information will provide us the graphical bigger picture for the traffic management.

9. ❖ **Space Utilization:** GIS helps managers to organize and spatially visualize space and how it can best be used. Operational costs can be decreased by more efficiently using space including managing the moves of personal and assets as well as the storage materials. The 3D visualization in GIS platforms helps planers to create a feeling of experience like virtual walk inside the building and rooms before construction.
10. ❖ **Development of Public Infrastructure Facilities:** GIS has many uses and advantages in the field of facility management. GIS can be used by facility managers for space management, visualization and planning, emergency and disaster planning and response. It can be used throughout the life cycle of a facility from deciding where to build to space planning. Also it provides facilitate better planning and analysis.
11. ❖ **Location Identification:** This technique is used to find a location for a new retail outlet. It helps to find out what exists at a particular location. A location can be described in many ways, using, for instance, name of place, post code, or geographic reference such as longitude or latitude or X/Y.
12. ❖ **River Crossing Site Selection for Bridges:** The important geotechnical consideration is the stability of slope leading down to and up from the water crossing. It is advisable to collect historical data on erosion and sedimentation. On the basis of these information asses the amount of river channel contraction, degree of curvature of river bend, nature of bed and bank materials including the flood flow and the flow depth, all these can be done in GIS within estimated time and accurately. This information has been often used for river crossing site selection for bridges.
13. ❖ **Regional Planning:** Every day, planners use Geographic Information System (GIS) technology to research, develop, implement, and monitor the progress of their plans. GIS provides planners, surveyors, and engineers with the tools they need to design and map their neighborhoods and cities. Planners have the technical expertise, political savvy, and fiscal understanding to transform a vision of tomorrow into a strategic action plan for today, and they use GIS to facilitate the decision-making process. (ESRI, GIS Solutions for Urban and Regional Planning).

4.3 GLOBAL NAVIGATION SATELLITE SYSTEM (GNSS)

Concept of GNSS

Global Navigation Satellite System (GNSS) refers to a constellation of satellites providing signals from space that transmit positioning and timing data to GNSS receivers. The receivers then use this data to determine location. By definition, GNSS provides global coverage. Examples of GNSS include Europe's Galileo, the USA's NAVSTAR Global Positioning System (GPS), Russia's

Global'naya Navigatsionnaya Sputnikovaya Sistema (GLONASS) and China's BeiDou Navigation Satellite System. GNSS (Global Navigation Satellite System) is a satellite system that is used to pinpoint the geographic location of a user's receiver anywhere in the world. Two GNSS systems are currently in operation: the United States' Global Positioning System (GPS) and the Russian Federation's Global Orbiting Navigation Satellite System (GLONASS).

Component of GNSS

The GNSS consist of three main satellite technologies: consists mainly of three segments: (a) space segment, (b) control segment and (c) user segment. These segments are almost similar in the three satellite technologies, which are all together make up the GNSS.

Space Segment

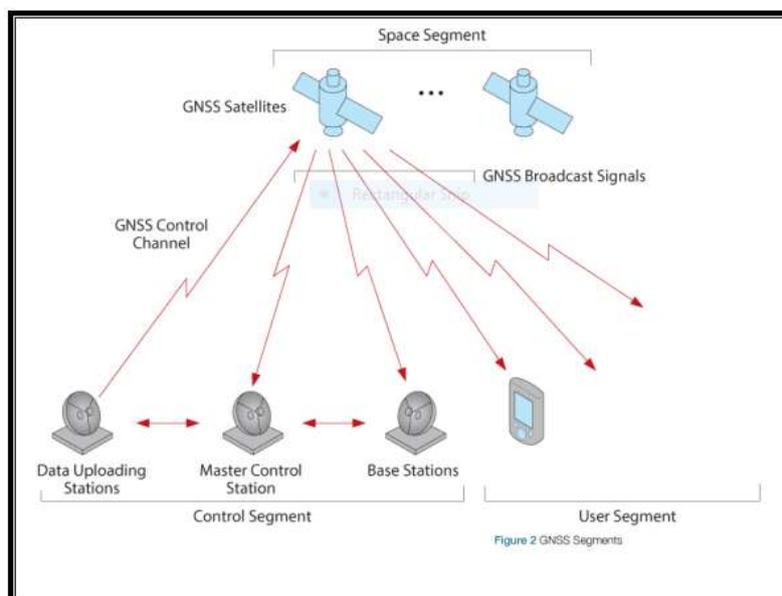
The space segment consists of GNSS satellites, orbiting about 20,000 km above the earth. Each GNSS has its own "constellation" of satellites, arranged in orbits to provide the desired coverage. Each satellite in a GNSS constellation broadcasts a signal that identifies it and provides its time, orbit and status. To illustrate, consider the following. You are downtown. You call a friend. Your friend is not at home, so you leave a message.

Control Segment

The control segment comprises a ground-based network of master control stations, data uploading stations and monitor stations; in the case of GPS, two master control stations (one primary and one backup), four data uploading stations and 16 monitor stations, located throughout the world.

In each GNSS system, the master control station adjusts the satellites' orbit parameters and onboard high-precision clocks when necessary to maintain accuracy.

Monitor stations, usually installed over a broad geographic area, monitor the satellites' signals and status, and relay this information to the master control station. The master control station analyses the signals then transmits orbit and time corrections to the satellites through data uploading stations.



GNSS Segments.

User Segment

The user segment consists of equipment that processes the received signals from the GNSS satellites and uses them to derive and apply location and time information. The equipment ranges from smart phones and handheld receivers used by hikers, to sophisticated, specialized receivers used for high end survey and mapping applications.

Types of GNSS:

The GNSS consist of three main satellite technologies: GPS, GLONASS and Galileo. Each of them consists mainly of three segments: (a) space segment, (b) control segment and (c) user segment. These segments are almost similar in the three satellite technologies, which are all together make up the GNSS. These are as follows.

A. G.P.S.

As of today, the complete satellite technology is the GPS technology and most of the existing worldwide applications related to the GPS technology. The GNSS technology will become clearer after the operation of Galileo and the reconstruction of Glonass in the next fe years. The United States Department of Defense (DoD) has developed the Navstar GPS, which is an all-weather, space based navigation system to meet the needs of the USA military forces and accurately determine their position, velocity, and time in a common reference system, any where on or near the Earth on a continuous basis.

GPS comprises three main components:

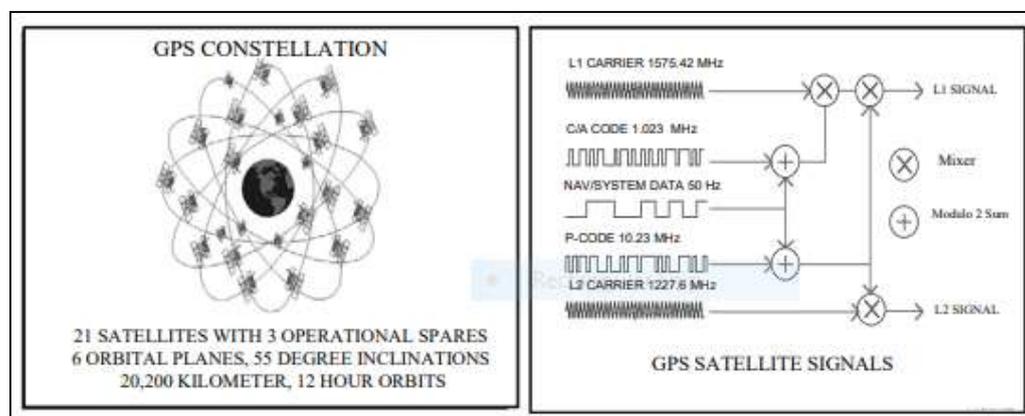
Space segment: The Space Segment of the system consists of the GPS satellites. These space vehicles (SVs) send radio signals from space as shown in Figure.

1. Control segment:

The Control Segment consists of a system of tracking stations located around the world. The Master Control facility is located at Schriever Air Force Base (formerly Falcon AFB) in the State of Colorado, USA.

2. User segment:

The GPS User Segment consists of the GPS receivers and the user community. GPS receivers convert space vehicle (SV) signals into position, velocity, and time estimates.



GPS Constellation and GPS Satellite Signals.

3. GALILEO:

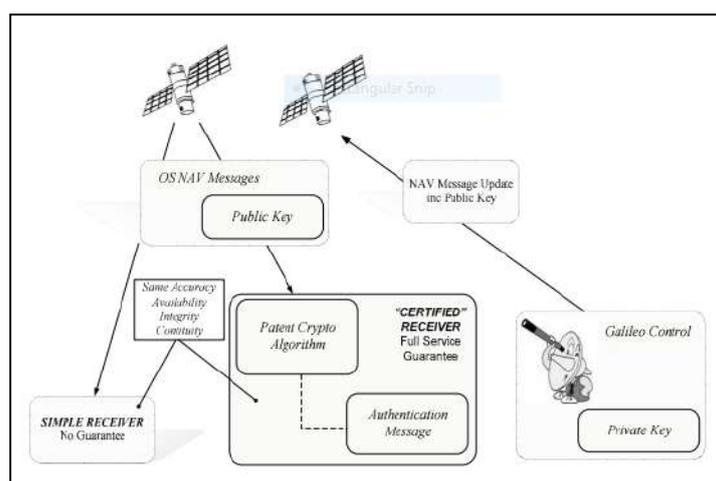
Galileo segments are almost similar to GPS, but with some modification. The main extension of Galileo compared to GPS is the implementation of a global/ regional segment for integrity monitoring.

4. Space Segment

The space segment or the constellation features consists of 30 Medium Earth Orbiting (MEO) satellites (27 and 3 active spare satellite), distributed evenly and regularly over three orbit planes. The projected altitude is slightly larger than for GPS 23,616 km and the inclination is 56°.

5. Ground Segment

The Galileo ground segment is responsible for managing the constellation of navigation satellites, controlling core functions of the navigation mission such as orbit determination of satellites, and clock synchronization, and determining and disseminating (via the MEO satellites) the integrity information, such as the warning alerts within time-to-alarm requirements, at global level. The Global ground segment will also provide interfaces with service centers. The Ground Control Segment will consist of about 12-15 reference stations, 5 up-link stations and two control centers. The ground segment also will include 16-20 monitor stations, three up-link stations for integrity data and two central stations for integrity computations.



Galileo Segments.

User Segment:

The user segment consists of different types of user receivers, with different capabilities related to the different GALILEO signals in order to fulfill the various GALILEO services.

Observation Techniques

The basic concept of GNSS is to measure the signal travelling time between artificial satellite and receiver. By multiplying this time by the light velocity (c), we get the range between the satellite and the receiver. The time or phase measurement performed by the receiver is based on the comparison between the received signal at the antenna of the receiver and the generated reference signal by the receiver. The two signals are affected by the clocks errors. Therefore, the range measured is not true and it is called pseudo range. Since the signal travels through the atmospheric layers, further noise should be modelled in order to compute the precise range.

- I. **Code Pseudo range Measurements:** Code correlation technique is used to measure the time difference between the received and generated replica code.
- II. **Phase Pseudo range Measurements:** Phase pseudo range is based on the measurements of phase difference between the received and generated signal $S \Delta \phi R$ at the receiver. The received carrier is Doppler shifted due to the motion of satellite.
- III. **GNSS observable errors:** The code and phase measurements are affected by noise and errors due to the propagation of signals through atmospheric layers and due to the noise measurements.

5. Map layer

A map layer is a GIS database containing groups of point, line, or area (polygon) features representing a particular class or type of real-world entities such as customers, streets, or postal codes. A layer contains both the visual representation of each feature and a link from the feature to its database attributes. Maps in a Geographic Information System are made by combining multiple layers.

5.1 Overlay analysis

Overlay analysis is one of the spatial GIS operations. Overlay analysis integrates spatial data with attribute data. (Attributes are information about each map feature.) Overlay analysis does this by combining information from one GIS layer with another GIS layer to derive or infer an attribute for one of the layers.

Overlay Analysis

(Using arc map 10.3)

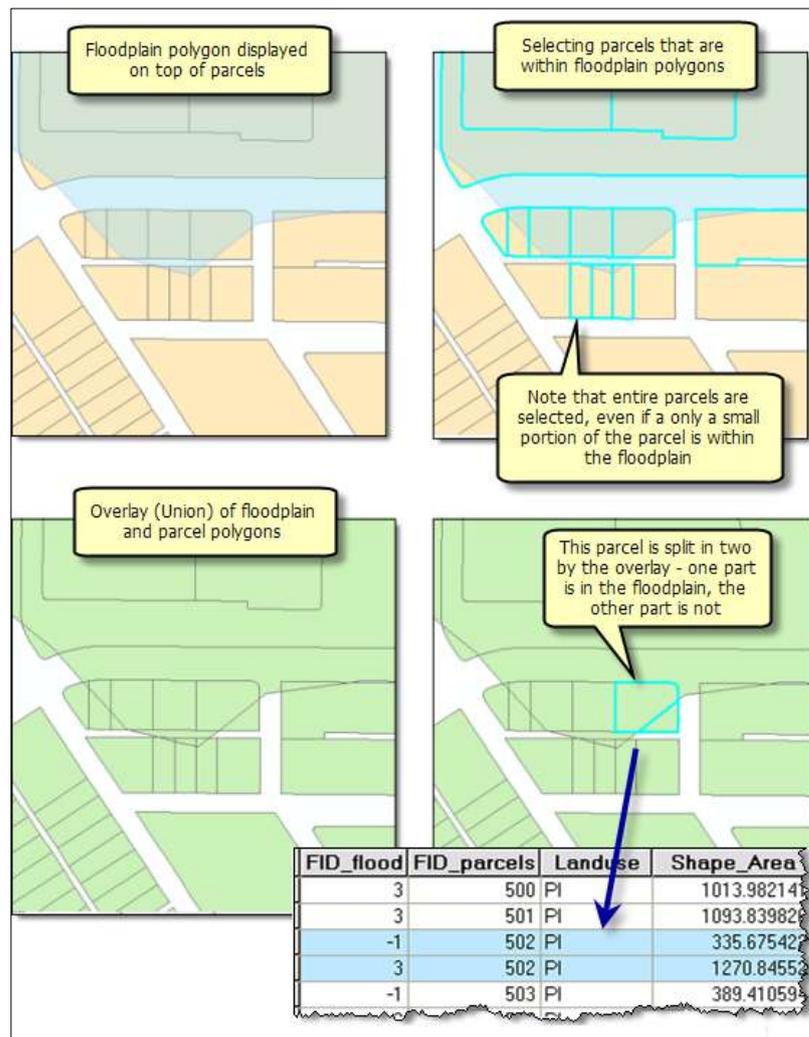
One of the most basic questions asked of a GIS is "What's on top of what?" For example:

- What land use is on top of what soil type?
- What parcels are within the 100-year floodplain? ("Within" is just another way of saying "on top of.")
- What roads are within what counties?
- What wells are within abandoned military bases?

To answer such questions before the days of GIS, cartographers would create maps on clear plastic sheets and overlay these sheets on a light table to create a new map of the overlaid data. Because overlay yields such valuable information, it was paramount to the development of GIS.

An overlay operation is much more than a simple merging of line work; all the attributes of the features taking part in the overlay are carried through, as shown in the example below, where parcels (polygons) and flood zones (polygons) are overlaid (using the Union tool) to create a new polygon dataset. The parcels are split where they are crossed by the flood zone boundary, and new polygons created. The FID_flood value indicates whether polygons are outside (-1) or inside the flood zone, and all polygons retain their original land-use values.

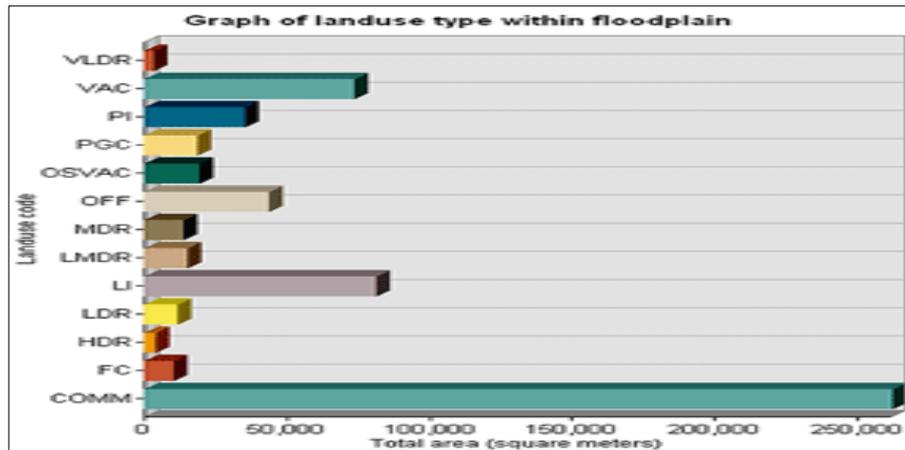
original land-use values.



their category

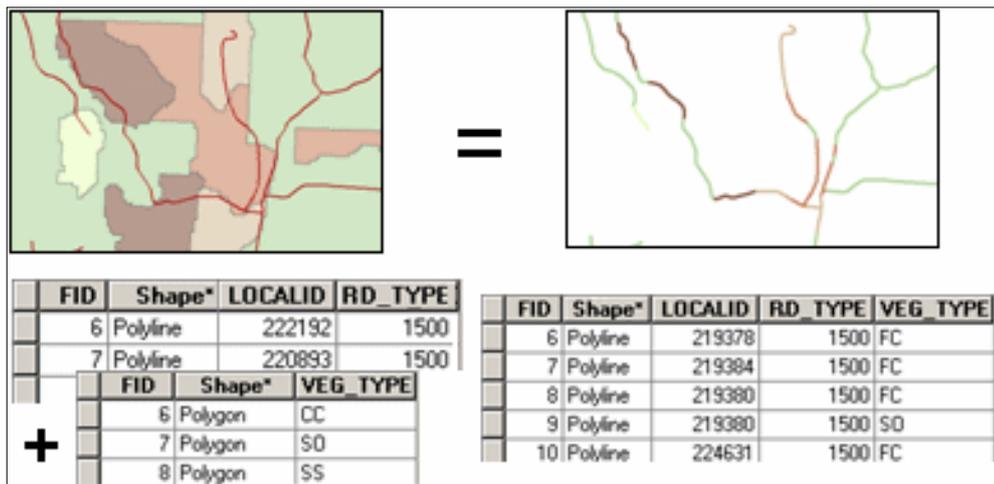
The total area of each land-use type in the flood zone can be calculated by selecting all polygons within the flood zone (using the Select Layer by Attribute tool, for example) and summarizing the

area by land-use type (using the Frequency tool). Following is a chart illustrating the result of this calculation.



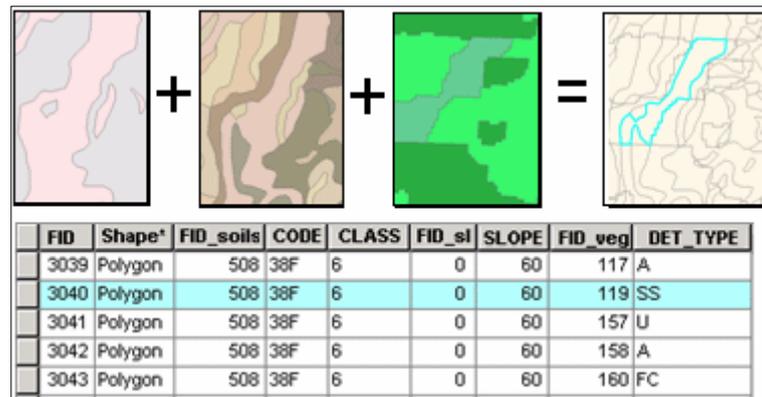
Similarly, you'd overlay watershed boundaries with a vegetation layer to calculate the amount of each vegetation type in each watershed.

In the illustration below, logging roads (lines) and vegetation types (polygons) are overlaid to create a new line feature class. The lines have been split where they were intersected by polygons, and each line feature has been assigned the attributes of both original layers. The lines are shown symbolized by the vegetation type associated with each.



You can use overlay analysis to combine the characteristics of several datasets into one. You can then find specific locations or areas that have a certain set of attribute values—that is, match the criteria you specify. This approach is often used to find locations that are suitable for a particular use or are susceptible to some risk. For example, you'd overlay layers of vegetation type, slope, aspect, soil moisture, and so on, to find areas susceptible to wildfire.

Below is an example of an overlay of steep slopes, soils, and vegetation. New polygons are created by the intersection of the input polygon boundaries. The resulting polygons have all the attributes of the original polygons.



Overlay analysis is often used in conjunction with other types of analysis. For example, you might include datasets derived from proximity analysis (such as the Buffer tool) or surface analysis (the Slope or Aspect tool). Similarly, you'll likely perform additional analysis on the results of the overlay, such as extraction to select a subset of features, or generalization (to dissolve polygons, for example). Often, overlay is one step in an analysis process or model and may occur at various points in the process.

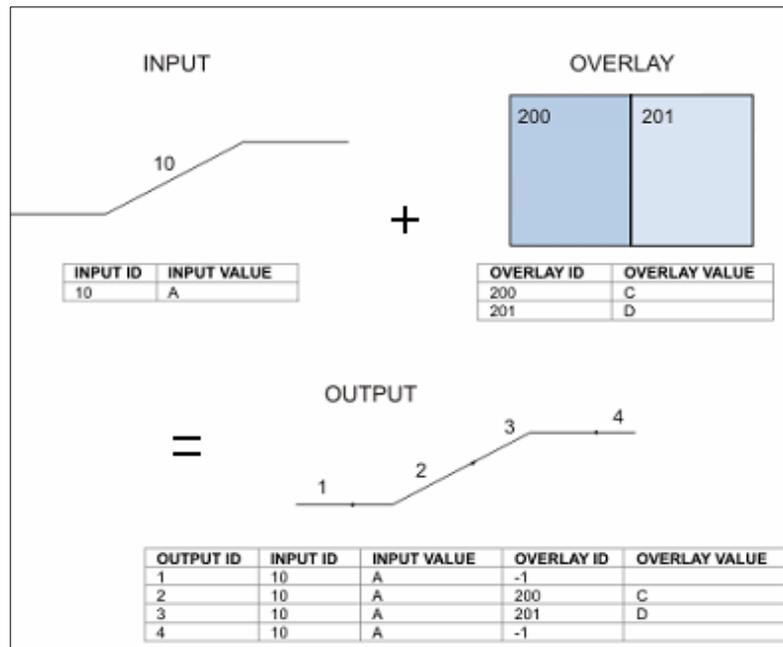
Overlay methods

In general, there are two methods for performing overlay analysis—feature overlay (overlapping points, lines, or polygons) and raster overlay. Some types of overlay analysis lend themselves to one or the other of these methods. Overlay analysis to find locations meeting certain criteria is often best done using raster overlay (although you can do it with feature data). Of course, this also depends on whether your data is already stored as features or rasters. It may be worthwhile to convert the data from one format to the other to perform the analysis.

Feature overlay

The key elements in feature overlay are the input layer, the overlay layer, and the output layer. The overlay function splits features in the input layer where they are overlapped by features in the overlay layer. New areas are created where polygons intersect. If the input layer contains lines, the lines are split where polygons cross them. These new features are stored in the output layer—the original input layer is not modified. The attributes of features in the overlay layer are assigned to the appropriate new features in the output layer, along with the original attributes from the input layer.

Below is an example of line-on-polygon overlay. The line is split at the polygon boundaries, and each of the resulting line features has the original line attributes plus the attributes of the polygon it fell within.



Raster overlay

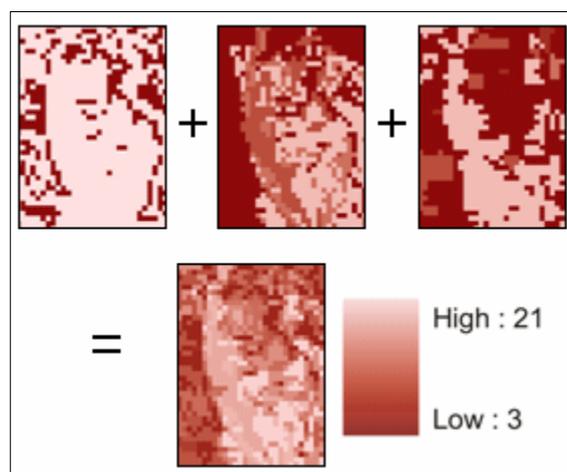
In raster overlay, each cell of each layer references the same geographic location. That makes it well suited to combining characteristics for numerous layers into a single layer. Usually, numeric values are assigned to each characteristic, allowing you to mathematically combine the layers and assign a new value to each cell in the output layer.

Below is an example of raster overlay by addition. Two input rasters are added together to create an output raster with the values for each cell summed.

INPUT 1				INPUT 2		
3	3	1		11	12	10
4	2	2	+	12	12	10
3	1	1		14	12	11
OUTPUT						
14	15	11				
16	14	12				
17	13	12				

This approach is often used to rank attribute values by suitability or risk, then add them to produce an overall rank for each cell. The various layers can also be assigned a relative importance to create a weighted ranking (the ranks in each layer are multiplied by that layer's weight value before being summed with the other layers).

Below is an example of raster overlay by addition for suitability modelling. Three raster layers (steep slopes, soils, and vegetation) are ranked for development suitability on a scale of 1 to 7. When the layers are added (bottom), each cell is ranked on a scale of 3 to 21.



Alternatively, you can assign a value to each cell in the output layer based on unique combinations of values from several input layers.

Overlay tools

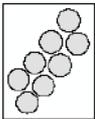
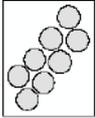
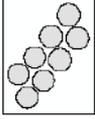
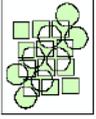
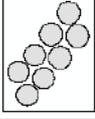
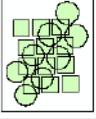
Vector overlay tools

Feature overlay tools are located in the Analysis toolbox in the Overlay toolset. Conceptually, the tools are similar—they differ by the feature types they allow you to overlay, by whether you can overlay multiple layers at one time, and by which input and overlay features are maintained in the output layer.

Tool	Binary or multiple overlay	Input data type	Overlay data type	Output
Identity	Binary	Any	Polygon or same as input	Input features, split by overlay features
Intersect	Multiple	Any	N/A	Only features common to all input layers
Symmetrical difference	Binary	Any	Same as input	Features common to either input layer or overlay layer but not both
Union	Multiple	Polygon	N/A	All input features
Update	Binary	Any	Polygon	Input feature geometry replaced by update layer

Overlay operations summary table

The table below shows the results of overlaying an input dataset and an overlay dataset using each tool.-

Input features	Overlay features	Operation	Result
		Identity	
		Intersect	
		Symmetrical difference	
		Union	

Overlay results visualization

Raster overlay tools

Raster overlay tools are located in several toolsets in the Spatial Analyst toolbox. Spatial Analyst is an ArcGIS extension that is licensed separately. If your site has a Spatial Analyst license and the Spatial Analyst extension has been installed, you will have access to the Spatial Analyst toolbox in ArcToolbox.

Tool	Location	What it does
Zonal Statistics	Zonal toolset	Summarizes values in a raster layer by zones (categories) in another layer—for example, calculate the mean elevation for each vegetation category.
Combine	Local toolset	Assigns a value to each cell in the output layer based on unique combinations of values from several input layers.
Weighted Overlay	Overlay toolset	Automates the raster overlay process and lets you assign weights to each layer before adding (you can also specify equal influence to create an unweighted overlay).
Weighted Sum	Overlay toolset	Overlays several rasters, multiplying each by their given weight and summing them together.

Raster overlay operations summary table

5.2 Buffers

A buffer is a reclassification based on distance: classification of within/without a given proximity. Buffering involves measuring distance outward in directions from an object. Buffering can be done on all three types of vector data: point, line, area. The resulting buffer is a polygon file.

Most often buffers are measured in uniform distance. For example, creating a 50' buffer around all rivers. A buffer based on different distances is called a variable buffer. For example, the noise level surrounding surround a street network may be based on the traffic load. Therefore a variable buffer may be used to illustrate the noise level by using a larger distance for high traffic roads and a shorter distance for quieter roads.

Creating a buffer around a feature

You can create a buffer around selected points, lines, or area features by using the Buffer command. Buffered features are created as line or polygon features using the settings of the chosen feature template.

For instance, you might use buffers to show an ecological zone around a waterway, distances from schools or public buildings that certain retail stores are located or the area around a contaminated well. You can buffer more than one feature at once, but a separate buffer will be created around each feature.



The buffer distance is given in map units by default. You can also give the value in other units by specifying a distance unit's abbreviation with the value that you enter.

1. Click the **Edit** tool  on the **Editor** Toolbar.
2. Click the feature around which you want to create a buffer.
3. Click the **Editor** menu and click **Buffer**.
4. Type the distance in map units for the buffer area around the feature.
5. Choose the target in which the new feature will be created.
 - If you have feature templates for the layers in your map, click the **Template** button and click the template to use to create the new feature. You can also double-click the preview of the template to choose a different template.
 - If you do not have feature templates, click the layer in which to create the feature.
6. Click **OK**.

Map Composition

Browse Geospatial Data

1.1. Load data

To launch Arc Map, please click: *Start > All Programs > ArcGIS > Arc Map 10*.

Click **OK** to proceed. You can ignore all other options. They are redundant.

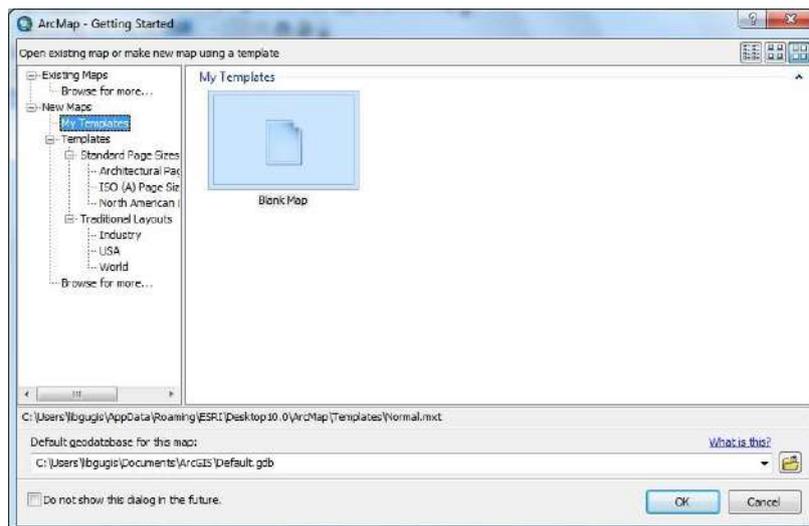


Figure 1. The start-up splash window of Arc Map

There are two ways to add data: Click the Add *Data*  button (Figure 2), which opens a window similar to Figure.

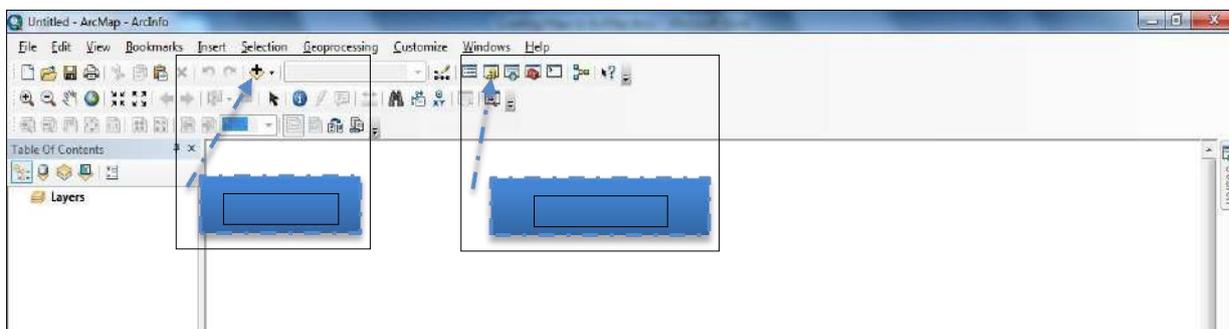


Figure 2. Main Windows with Marks on Add Data and ArcCatalogue

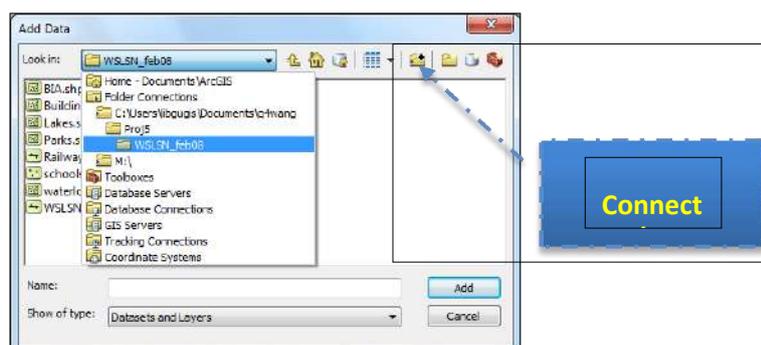


Figure 3. The Pop-up Window of "Add Data"

To browse to the file you want to play with, you can link to the folder containing all these files by clicking



The **ConnecttoFolder** button in Figure 3, then browse to add the folder that contains your files and click **OK**. After that, you can locate the files listed below and add them in at one time by holding the **Ctrl** or **Shift** button when you are clicking mouse button to make selections. Please click the **Add** button just like any other file explorer dialogs in Windows to close the dialog.

- **schools_dec06.shp**: All schools in the City of Waterloo;
- **waterloo_city.shp**: The administration area of the City of Waterloo;
- **WLSLN_feb08.shp**: Roads in the City of Waterloo;
- **AirPhoto.jpg**: A resolution-decreased air photo image of the City of Waterloo.

When data is loaded, all files will be listed in the left pane (table of contents) in Figure 4. The geographic features (contents) are displayed in the right pane. And most controls can be found on the top pane (or by right-clicking on the object you are working at and checking the pop-up window).

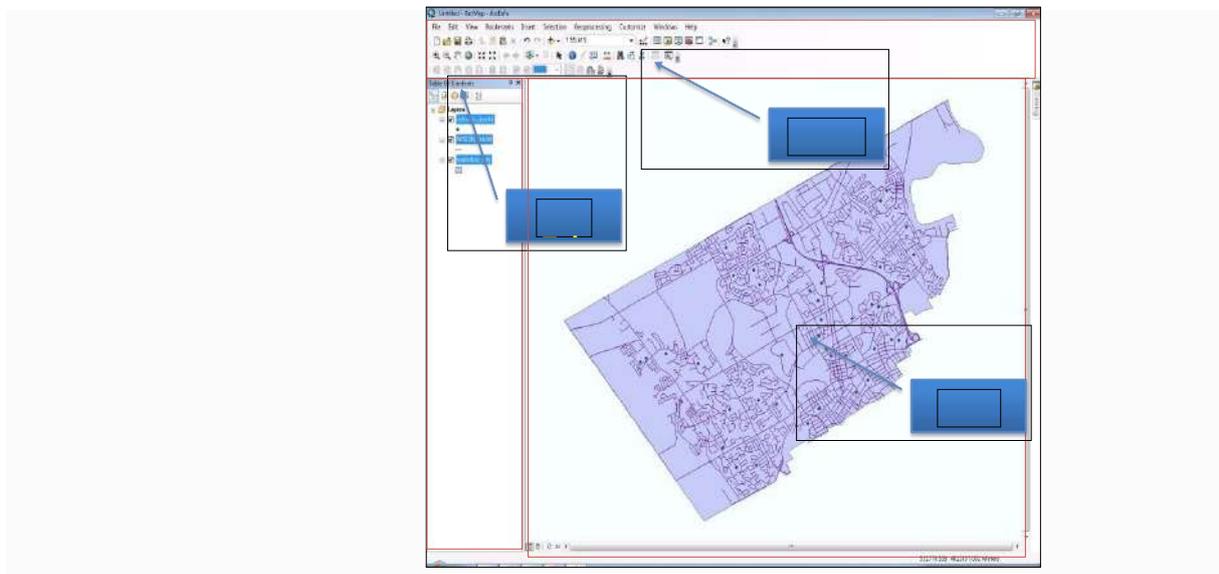


Figure 4: Structure of the Main Window

Browse Geographic Features

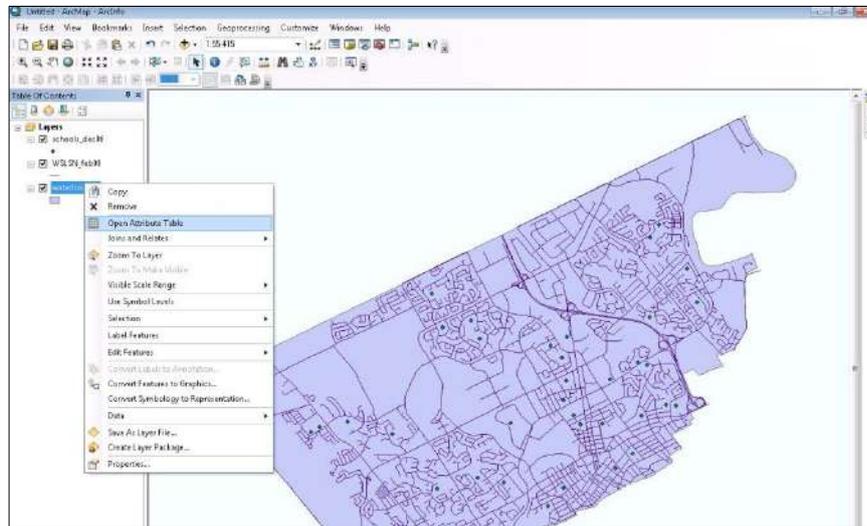
In ArcMap 10, ESRI starts to put most controls to browse data into one toolbar (Figure 5). If you cannot find this toolbar, please go to **Customize -> Toolbars** and check the **Tools** on.



Figure 5. The Toolbar with Data Browsing Controls

Most icons are intuitive and self-explained. If you are not sure what function it has, please put your mouse on that icon. A pop-up text will show and further explanation. In Figure 5, an example is shown when the mouse is hovering over the . The pop-up window explains its functionality: “Select Features by Rectangle”, which means that you can draw a rectangle over some area, and all features within (or have parts within) this rectangle will be selected for further operations.

In addition to the toolbar, some functions on attribute data are put in the pop-up window when you



right click on one layer (a geospatial data file or a feature class) in the table of contents (Figure 6). When you click on “Open Attribute Table”, a table with all information on features will be shown (Figure 7).

Figure 6. Pop-up Window with Options on a Layer.

FID	Shape	NAME	TYPE
0	Point	Bluevale Collegiate	PUBLIC
1	Point	Brighton Public	PUBLIC
2	Point	Cedarbrae Public	
3	Point	Centennial Public	
4	Point	Elizabeth Ziegler Public	
5	Point	Empire Public	
6	Point	Harold Wagner Public	PUBLIC
7	Point	Holy Rosary Separate	SEPARAT

Figure 7. Attribute Data Window

	Geographic Operations	Attribute Operations
--	------------------------------	-----------------------------

These controls can be classified and summarized into the following table:

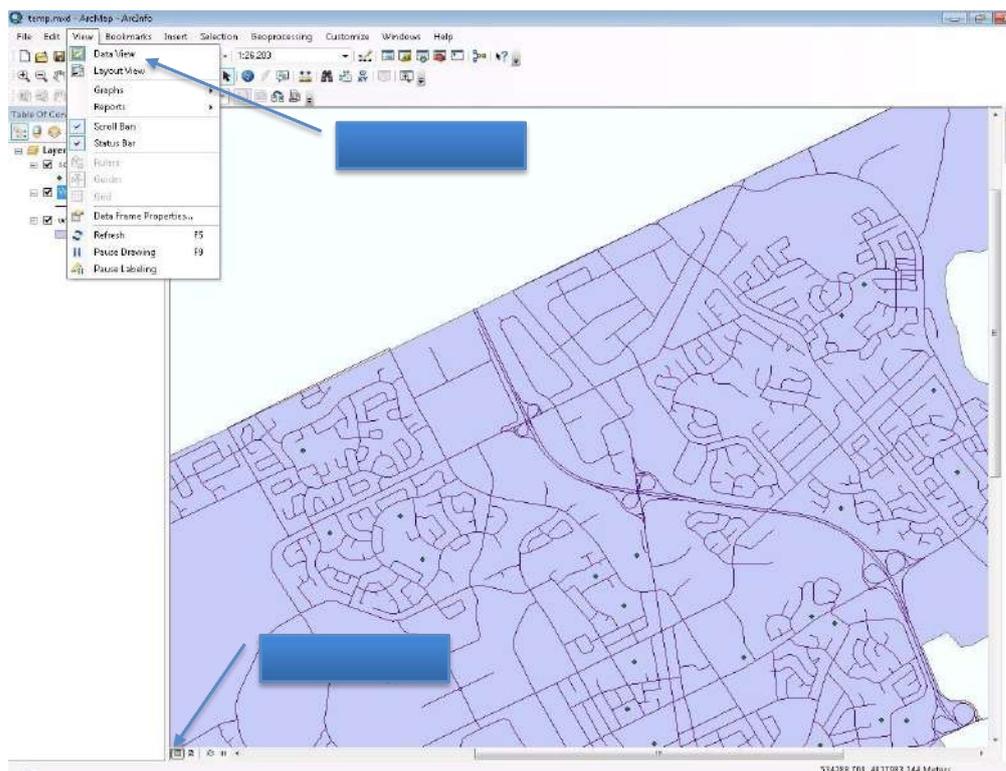
Mapping

Distinction between Geospatial Data and Map Components

This section is more concept- oriented. To simplify mapping process, ArcGIS divides mapping into two categories: data view, which focuses on data manipulation and representation (symbology), and layout view, which provides functionality of adding legend, scale bar, north arrow, etc.

All data representations should be set in data view. Key operations include the change of layer order, symbology customization, layer transparency, labeling, and annotations. Most of data representations can be saved in a layer file (*.lyr), because they are more data associated.

Layout view shows the virtual map you will get. Layout view works in a What You See Is What You Got (WYSIWYG) fashion. The output will be the same as what you see in the layout view, where you can add map elements and change the paper and output settings. Map setting will be save in a *.mxdfile.



Browse		Right-click on the layer (Figure 6 and 7)
Search/Identify		
Select		(Figure 7)

Figure 8: Switch between Data View and Layout View

Key Options of Geospatial Data Representations

Main options of changing geospatial data representations are layer order, layer transparency, symbology, label, and annotation. Apart from the first one (layer order) and last one (annotations), all the rest locate in the pop-up window (Figure 9) when you right-click the layer you want to modify and select property (Figure 10). They are under either the symbology or labels tab.

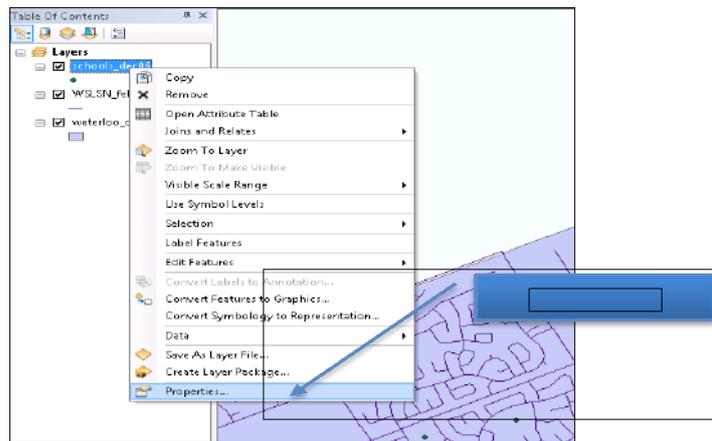


Figure 9. Pop-up Window of a Layer's Property

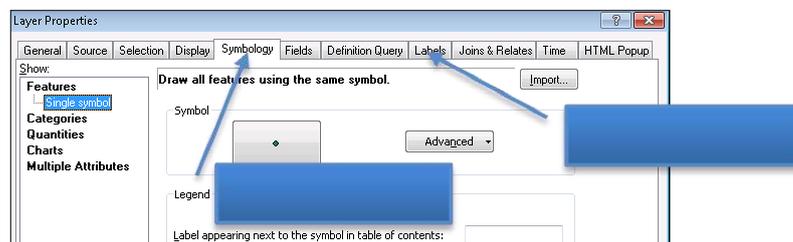


Figure 10. The Property Window

All options will be introduced as subsections in the following.

Layer Order and Transparency

When users want to overlay air photos (airphoto.jpg) and parcels (waterloo_city.shp) together to produce a map like Google Satellite View the parcel layer (waterloo_city.shp) is better be made as “halo” polygons, which will be introduced after the concept of layer order.

Layer Order

ArcMap displays geospatial data according to the order in the table of contents: the bottom layer will be drawn on the screen first and covered by upper layer. Hence, the layer on the top in the table of content will be displayed as the top layer in the map. By default, ArcMap shows points, lines, polygons, and raster data (images, DEMs, etc) from the top to the bottom respectively so that the visibility is maximized. If two layers belong to the same feature class, i.e., point features, the newly added one will be on top of the older one.

When there are multiple raster files, layer visibility and order are important. You can change layer visibility by switching the checkboxes left to the layer name in table of contents (Figure 11). And the display order can be changed by simply dragging the layer toward or away from the top.

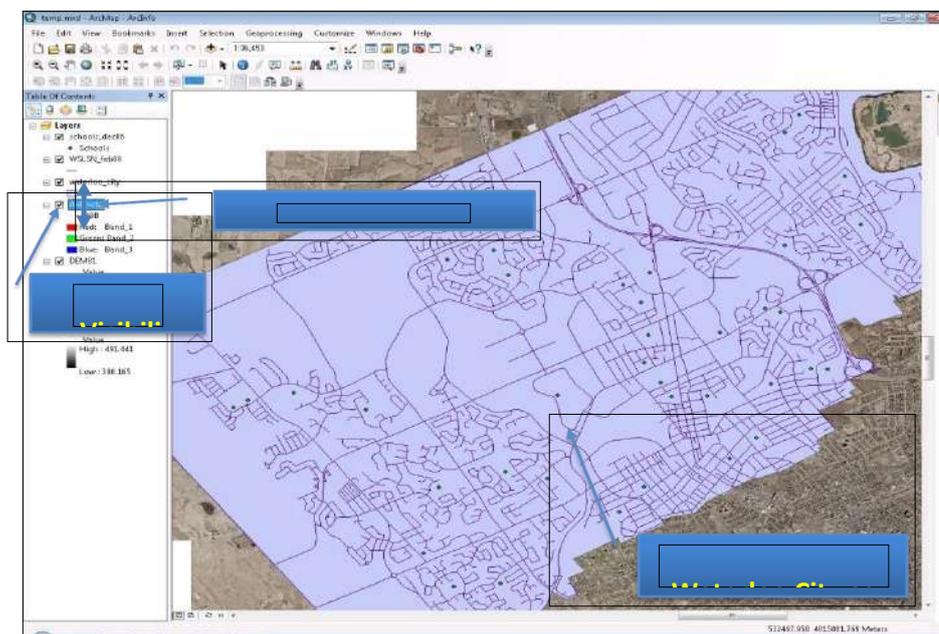


Figure 11. Layer Visibility and Order Control

Layer Visibility

When a polygon layer, such as waterloo city, is added over an image (air photo), the content on image under the building is invisible (Figure 11). But users often like to only show the parcel’s boundary and remain the image displayed. In this case, layer transparency and hollow symbology can help.

To do so, please right-click on the polygon layer (waterloo city) and select Properties first (Figure 9). Then please switch to the tab Symbology and click on the symbol itself (Figure 12).

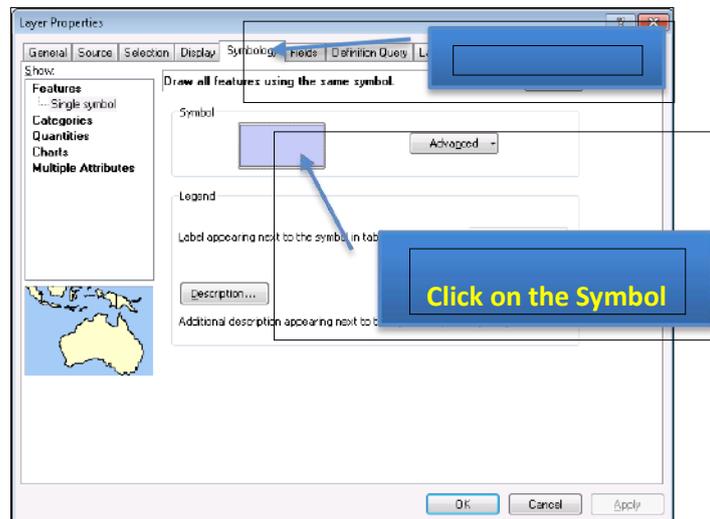


Figure 12. Properties of Symbology

When the pop-up window shows (Figure 13), you can change the **Fill Color** to **No Color** and increase the **Outline Width** (0.4 point is somewhat narrow).

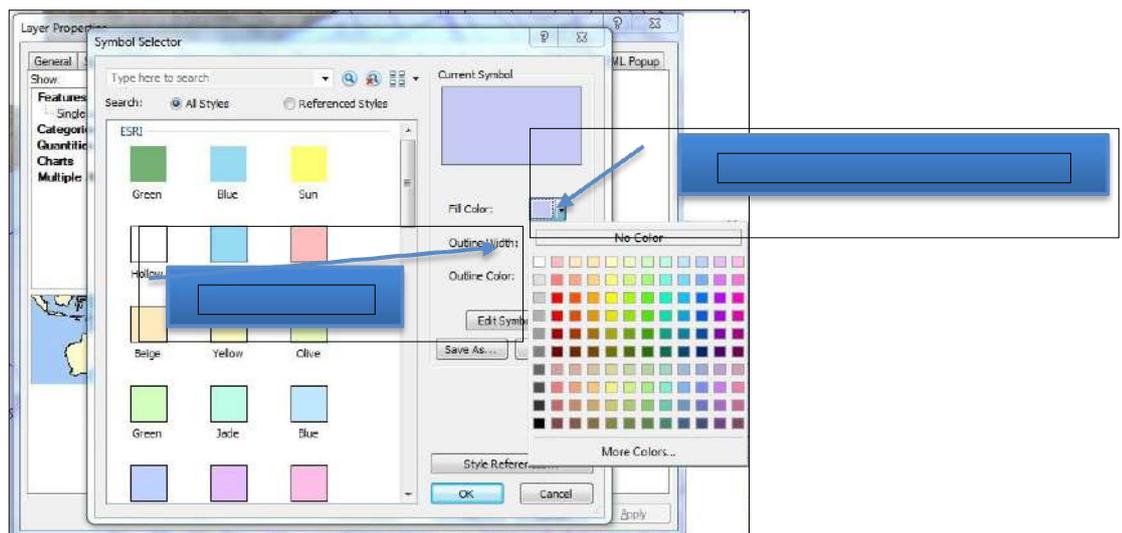


Figure 13. Make Hallow Polygons

Symbology and Label

Symbology

Symbology is critical in making maps, which are classified into four categories in ArcMap. Due to its complexity, these four categories and their normal usages will be briefly introduced in the note region with further reading. We will only focus on symbol customization and classification symbols, which are mostly used.

Please right-click the **schools_dec06** layer and select **Properties**. Then switch to the **Symbology** tab and click on the **Symbol** (Figure 12). The pop-up window (Figure 14) contains three main parts. In most cases, you will simply choose a symbol from the symbol library (other than making hollow polygons mentioned in 2.2.1.). Depending on the type of the feature (point, line, or polygon), the options change accordingly.

You can type in **school** and click the search icon (magnifier) to search all symbols related to school representations.

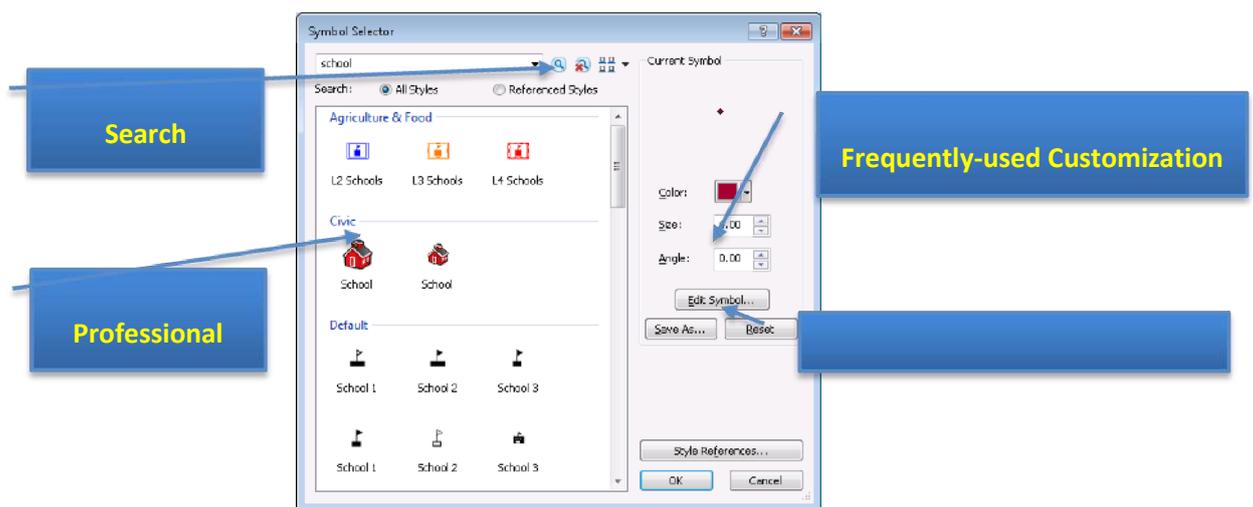


Figure 14. Symbology Customization Window with Explanations

There are cases when you want to differentiate features in a layer based on some attribute. For instance, I want to know which schools are public, separate, or for higher education in the schools layer, which is the top layer in Figure 11.

Different types will be marked in colors. To do so, we should **right-click** on the school layer, select **properties**, and go to the **symbology** tab, which is the same as shown in Figure 12. But, instead of by-default **Features** option, we should click and select the **categories** option (Figure 15).

The steps to follow are also shown in Figure 15:

1. Click on the **Categories** first and make sure the “**Unique Value**” is chosen;
2. Draw down the drop box of **Value Field**, and select TYPE as the classification attribute;
3. Click on the **Add All Values** button to add all distinct values under the TYPE

- attribute into consideration;
- (Optional) Click on the **Color Ramp** drop box and select your favourite color schema.

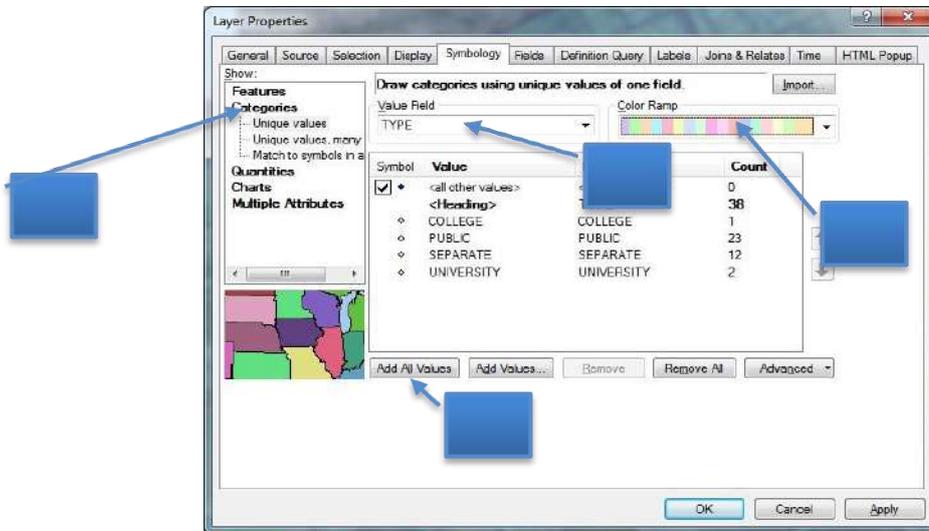


Figure 15. Classifying Features using Colors

The result is shown in Figure 16. Please note that you can still customize individual symbol for each category (size, shape, etc) by double clicking corresponding class on either the left pane or the line in symbology window (Figure 16). The pop-up window is very similar to the window in Figure 14. Please,

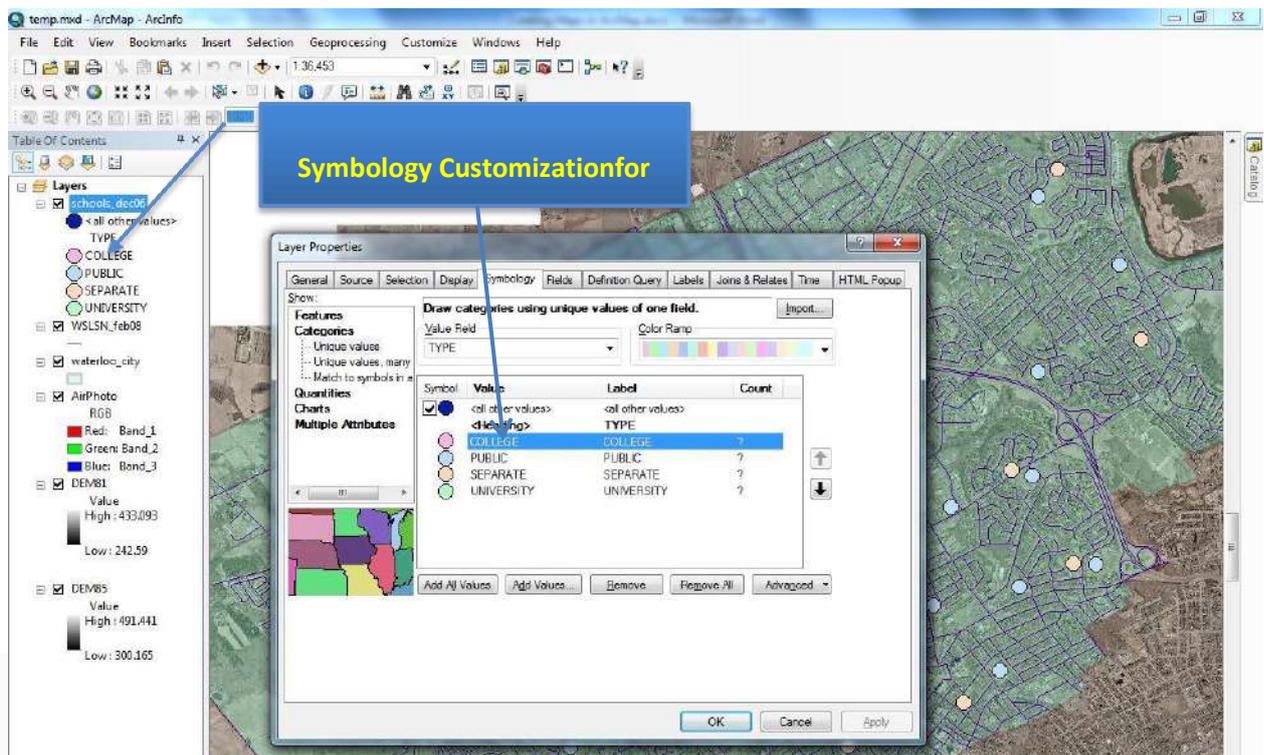


Figure 16. Classified Symbols and Further Customization.

Label

The labels are an important feature of a map. By marking some property up on the map, for stance, the name of the marked location, labels can make your map more useful, informative, and visually appealing.

To add labels to your map, please **right-click** on the layer that you want to add marks on (**school_dec06**) and select **Properties** (Figure 12). Then follow the steps below:

1. Select the **Labels** tab and check on (off by default) “Label features in this layer”.
2. Choose the attribute you want to display (**NAME**) on the map in the dropdown of **Label Field** (Figure 17).
3. Change Font (**10 pt size and Bold**), so that it is clearly shown on the map.
4. You can click on the **Apply** button to see the effect until it is satisfactory. Then click **OK** to save your setting and exit.

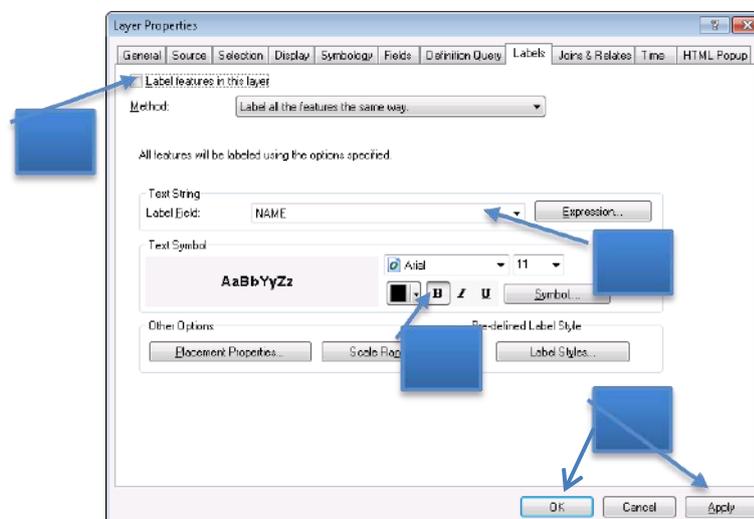


Figure 17. Label Setting Window

To ensure a clear label, the option of mask can be used **Symbol**, which locates to the right of Font options as a **Symbol** button (Figure 17). Please click it and then click the **Edit Symbol** button, since label is a kind of symbol (text symbol) too. Navigate to the **Mask** tab shown in the figure below. Change the option of **Style** to **Halo** (Figure 18).

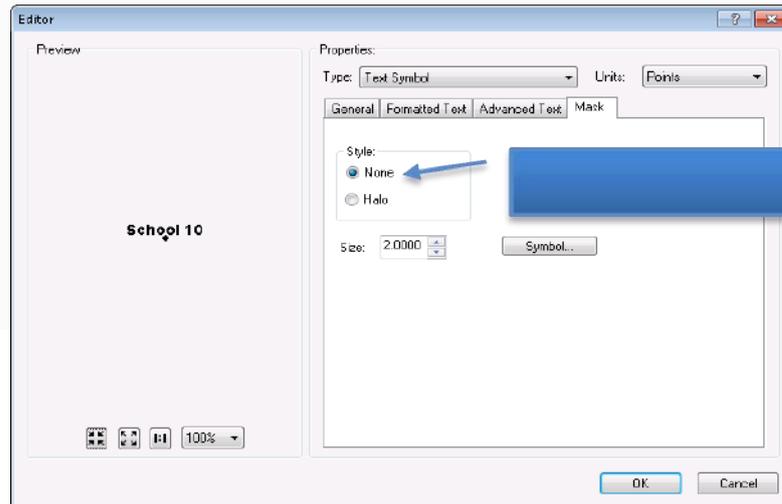


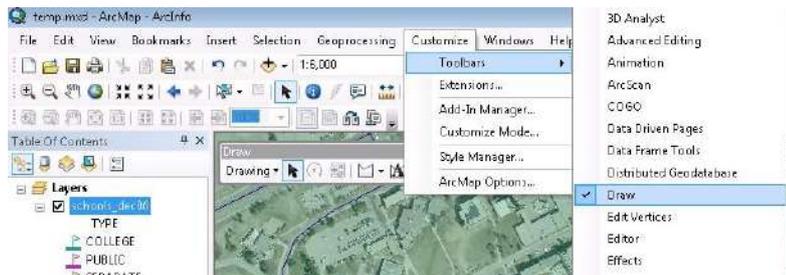
Figure 18. Change Mask Option for Clearer Labels.

Annotations

Annotations look like labels. The key difference is that annotation can be any text you want to add on the map, regardless to whether the information has been included in the geospatial data.

For instance, if you want to add a point at the entrance of UW, which is not in the geospatial data, please follow the steps below (*Please note that annotations can only be edited under the data view*):

1. Turn on the **Draw** toolbar. To do so, please click on the menu of **Customize->Toolbars->Draw** (Figure 19).



2. Add Marker. Click on  and select **Marker**. Then browse and point to the location you want to add the marker (South Campus Hall in this case) and click. A marker will be added to the location.



Figure 20. The Draw Toolbar

3. Add Text. Click  and choose **Call out** (or **New Text**). Click on the map, somewhere close to the marker you just made and type in accordingly (i.e. "UW Entrance").

Result is shown in Figure 21.

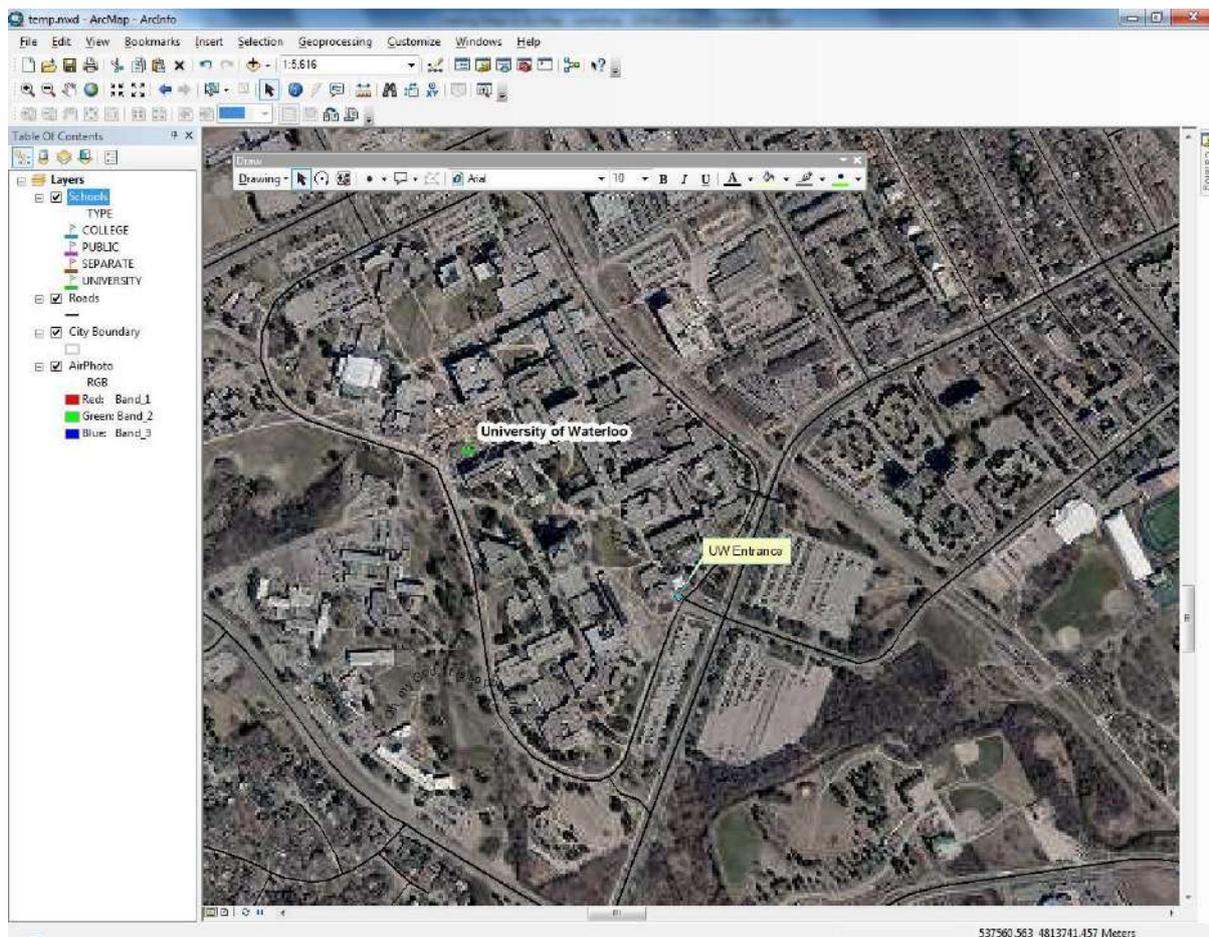


Figure 21. Map after Adding Labels (with Mask) and Annotations

Map Components

When you select **Layout View** to start adding map components, all components are gathered in the menu of **Insert** (Figure 22). The key ones are:

- The title of the map(Purpose/Content);
- The scale of the map(Measurement);
- The North Arrow (Orientation);
- The Legend (how to interpret the map);
- (Optional) The reference (where you obtain the data).

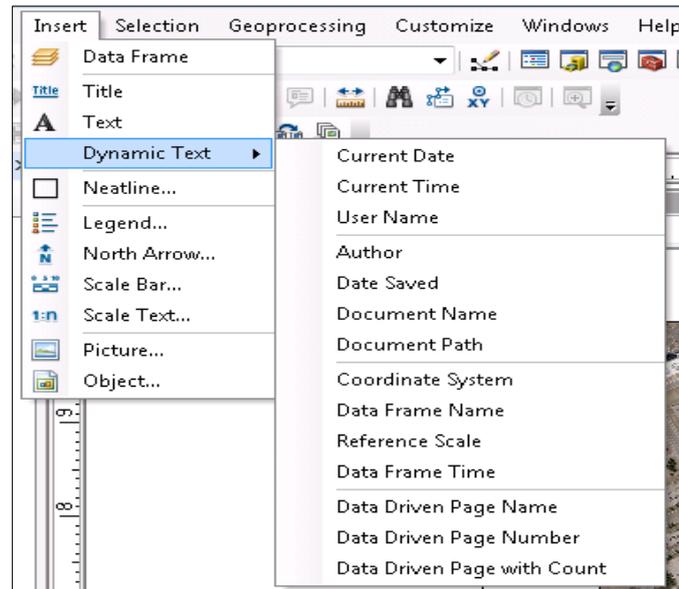


Figure 22. All Map Components Available

Map titles are just a specific kind of text box with some pre-defined fonts. Hence, the usage of title and text are identical and intuitive: type in the content and customize the text properties.

Add scale bar or scale text is also intuitive. The only thing that should be noted as the display unit, such as meters, degrees, miles, etc. You can customize it by clicking the **Properties** button in the **Scale Bar (or Text) Selector**, which is the pop-up window when you choose **Insert ->Scale Bar (or Scale Text)**. In the **Scale and Units** tab, you can choose an appropriate unit in the **Division Units** drop box (Figure 23). In addition, you can change the font style under the **Format** tab as well.

In this example, please use Scale Text with meters as unites shown like-



Add North Arrow is the easiest task, which is very similar to add a scale bar. Just choose a pre-defined north arrow from the library shown in the left pane is ok. If your North Arrow is difficult to be distinguished from the background, you can add some background color, which has similar effect of “mask” for labels. This option normally locates in the **Frame** tab (Figure 24). *Legend scales, and texts all have similar options.*

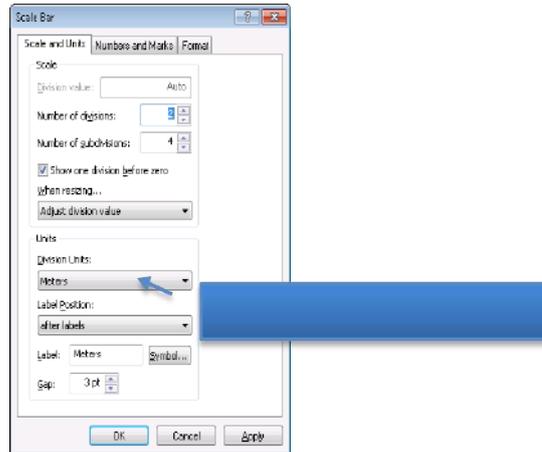


Figure 23. The Scale Bar (or Text) Property Window



Figure 24. Change Background Color in Frame Tab

Adding Legend is somewhat complex and deserves some explanation. When **Insert ->Legend** is selected, ArcMap will use a wizard guiding you through various options of making a legend.

At the first step, you will be asked about which layers need legend and the order of adding legend (Figure 25). The left pane **Map Layers** displays all layers (geospatial data) available for the map. The right pane **Legend Items** shows the layers that will be added to the legend in order.

- Add/Remove Layers: You can decide the layers to be added to the legend by clicking  to add  to remove. Please **remove the air photo layer**, since the band information for raster images is not useful (Figure 24).
- Change the Display Orders: The legend items will be added strictly in the order shown in the **Legend Items** pane. You can change the order by clicking a layer first (displayed in the blue bar) and move it up  or move it down .

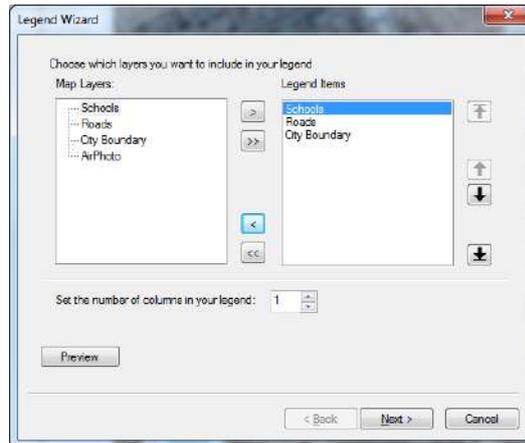


Figure 25. The First Step of Adding Legend

In the following three steps (four steps in total), ArcMap will ask you about some options to display the legend. None of them is critical and requires customization (especially for normal users). You can just accept all default configurations and click **next** until **Finish** it. You will get the legend that looks like in Figure 26.

It is easy to find that the name displayed in the legend is difficult to understand. For instance, users won't know what "WLSLN_feb08" means. In this case, you can modify the name in the **table of content** pane (Figure 4). To do so, you can **right-click** on the "WLSLN_feb08" layer and select **Properties** (Figure 12). At this time, please switch to the **General** tab. Then you can type the more meaningful name in the **Layer Name**: Roads (Figure 27).



Figure 26. An Example of Generated Legend

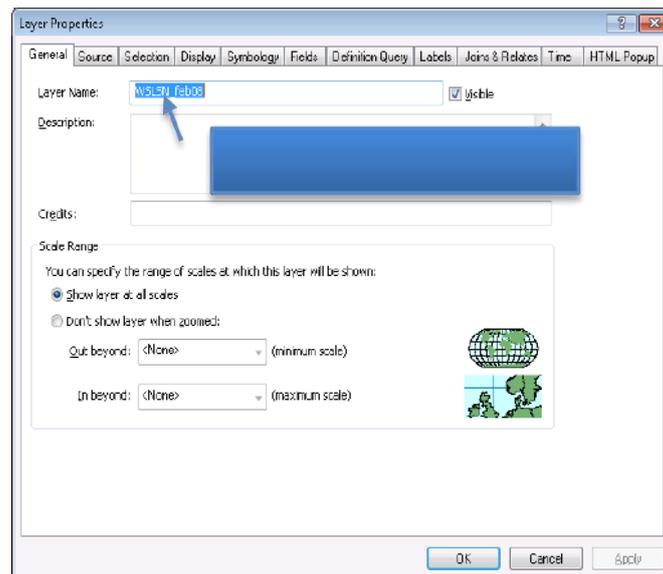


Figure 27. The General Tab to Change Layer Name

UML would like to emphasize the importance of adding references to your map. The reference clarifies the data source, making others easier to track, and also is good for academic integrity.

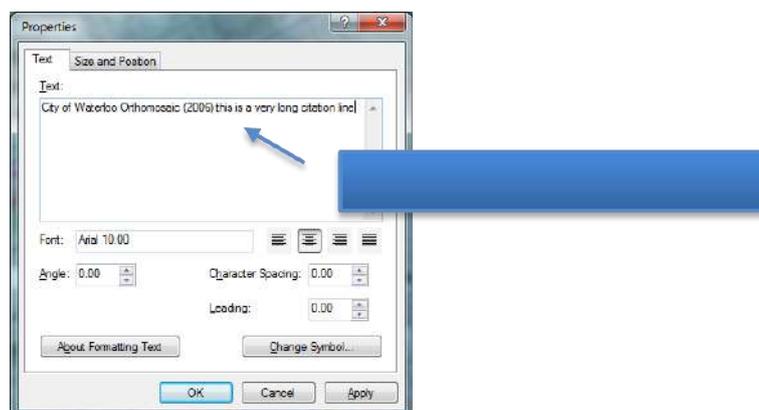


Figure 28. Property Window of Text in Layout View

Export Maps

When you are satisfied with everything in the map and are ready to deliver, you can export your map by the menu of **File ->Export Map**, which will pop-up a dialog (Figure 29).

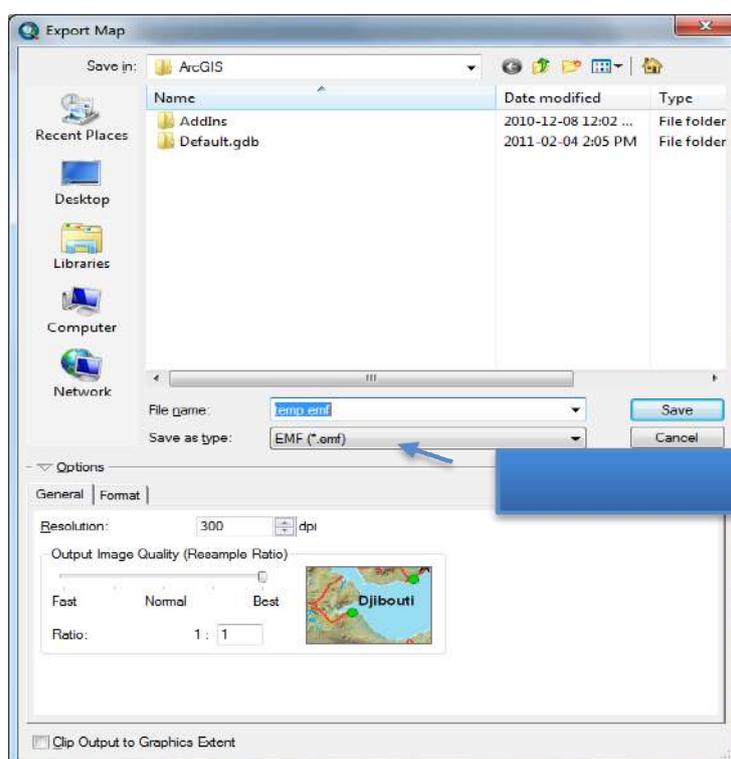


Figure 29. Export Map Dialog

You can specify the directory and file name of the map in this dialog. But, most importantly, you need to choose an appropriate file format by click on the **Save as type** drop box. The most frequent-used formats are JPEG and PDF. JPEG is a popular graphic file format, which is easy to be inserted into word as a graph, while PDF is best used for sharing. The printed copy of PDF will be identical regardless of your computer (and printer) environment.

Please note that if the purpose of your map is to be shown on screen, **96 dpi** resolutions are good enough. It is especially true when you want to publish your map online, given that lower resolution maps are much smaller, and therefore, easier to transfer via Internet. If you need hard copies of your map, please make sure that the resolution is set to **300 dpi** or **600 dpi**. Otherwise the output map might be obscure.

Other than export map to other file formats, you can directly print out hardcopies of your map by **File ->Print**. You can preview the result via **File ->Print Preview** to ensure satisfactory result. And, of course, you can customize page and printer setup via **File ->Page and Print Setup**, such as page orientation, page size, printer configuration (color or grey), and so on. All these operations are intuitive and very similar to other applications, such as Microsoft Word.

GEO 395.2: RESEARCH METHODOLOGY

Course Outcome

The learners will get the initial training on various steps involved in geographical research. They will develop the idea on fundamentals of research methodology including data collection, methodology and report writing. This course aims to develop fundamental research aptitude among all the students.

Pattern of Setting Questions

Full Marks – 25: Three compulsory questions bearing marks 7, 7 and 6 respectively, will be set covering the whole unit. 5 marks will be allotted for Laboratory Note Book and Viva-voce.

Detailed Syllabus

1. Research ethics and paradigm shift of research methodology in Geography.
 2. Need for research, basic research types.
 3. Identification of research problems, development of theoretical background - literature review, research gap and research question and specification of the objectives of study; hypothesis building, Framework of research writing.
 4. Methods of data collection- primary and secondary; Preparation of questionnaire and survey schedule and their differences, research ethics.
 5. Methods of writing notes, style of referencing, bibliography and appendices, abstract and synopsis writing.
-

Geography is a tremendously wide ranging field to study the earth, its natural and physical environments, human activities and social changes, the interrelationships and interactions of these and their effects, from local to global scales. Conducting research in geography requires clarity of concepts and theoretical knowledge of scientific behaviour studies. For much of the twentieth century, as geographers' concerns ranged over contemporary physical and human space and into their past arrangements in so far as they could be documented, the methods they used to explain, model, and predict different aspects of the human and physical worlds became progressively more quantitative. But the new technologies and theoretical perspectives that emerged in the latter decades of the twentieth century helped to redefine the objects of geographers' inquiry and extend the methods in use for collecting and analyzing data and evaluating research.

The Concept of Research

Research is a careful and exhaustive investigation of a phenomenon with an objective of advancing knowledge. According to Theodorson and Theodorson (1969), “it is a systematic and objective attempt to study a problem for the purpose of deriving general principles”. Robert Burns (2000) describes it as a systematic investigation to find solution to a problem. The investigation is guided by previously collected information. Kerlinger (1964) has defined scientific research as “a systematic, controlled, empirical, and critical investigation of hypothetical propositions about the presumed relations among phenomena”. Henry Johnson (Black and Champion, 1960) has stated following four characteristics of Scientific Research:

- **It is empirical**, i.e. it is based on observation and reasoning and not on speculation.
- **It is theoretical**, i.e. it summarises data precisely giving logical relationship between propositions which explain casual relationship.
- **It is cumulative**, i.e. generalisations/theories are corrected, rejected and newly developed theories are built upon one another.
- **It is non-ethical**, i.e. scientists do not say whether particular things/ events/ phenomena/ institutions/ systems/ structures are good or bad. They only explain them.

Research Ethics

Research ethics are the set of ethics that govern how scientific and other research is performed at research institutions such as universities, and how it is disseminated. Research ethics are important for a number of reasons as described by Resnick (2015).

- They promote the aims of research, such as expanding knowledge.
- They support the values required for collaborative work, such as mutual respect and fairness. This is essential because scientific research depends on collaboration between researchers and groups.
- They mean that researchers can be held accountable for their actions. Many researchers are supported by public money, and regulations on conflicts of interest, misconduct, and research involving humans or animals are necessary to ensure that money is spent appropriately.
- They ensure that the public can trust research. For people to support and fund research, they have to be confident in it.

- They support important social and moral values, such as the principle of doing no harm to others.

Codes of Ethics

The following is a general summary of some ethical codes postulated by US National Institutes of Health:

1. **Honesty:** Honestly report data, results, methods and procedures, and publication status. Do not fabricate, falsify, or misrepresent data.
2. **Objectivity:** Strive to avoid bias in experimental design, data analysis, data interpretation, peer review, personnel decisions, grant writing, expert testimony, and other aspects of research.
3. **Integrity:** Keep your promises and agreements; act with sincerity; strive for consistency of thought and action.
4. **Carefulness:** Avoid careless errors and negligence; carefully and critically examine your own work and the work of your peers. Keep good records of research activities.
5. **Openness:** Share data, results, ideas, tools, resources. Be open to criticism and new ideas.
6. **Respect for Intellectual Property:** Honour patents, copyrights, and other forms of intellectual property. Do not use unpublished data, methods, or results without permission. Give credit where credit is due. Never plagiarize.
7. **Confidentiality:** Protect confidential communications, such as papers or grants submitted for publication, personnel records, trade or military secrets, and patient records.
8. **Responsible Publication:** Publish in order to advance research and scholarship, not to advance just your own career. Avoid wasteful and duplicative publication.
9. **Responsible Mentoring:** Help to educate, mentor, and advise students. Promote their welfare and allow them to make their own decisions.
10. **Respect for Colleagues:** Respect your colleagues and treat them fairly.
11. **Social Responsibility:** Strive to promote social good and prevent or mitigate social harms through research, public education, and advocacy.
12. **Non-Discrimination:** Avoid discrimination against colleagues or students on the basis of sex, race, ethnicity, or other factors that are not related to their scientific competence and integrity.

- 13. Competence:** Maintain and improve your own professional competence and expertise through lifelong education and learning; take steps to promote competence in science as a whole.
- 14. Legality:** Know and obey relevant laws and institutional and governmental policies.
- 15. Animal Care:** Show proper respect and care for animals when using them in research. Do not conduct unnecessary or poorly designed animal experiments.
- 16. Human Subjects Protection:** When conducting research on human subjects, minimize harms and risks and maximize benefits; respect human dignity, privacy, and autonomy.

Research Misconducts

1. Fabrication - making up data or results and recording or reporting them.
2. Falsification - manipulating research materials, or changing or omitting data or results such that the research is not accurately represented in the research record.
3. Plagiarism - the appropriation of another person's ideas, processes, results, or words without giving appropriate credit.
4. Research misconduct does not include honest error or differences of opinion.

Paradigm Shift of Research Methodology in Geography

The shift in Geography Methodology has never been smooth. Geography has over the years been at crossroads with other disciplines and having ups and downs; its definition, scope, techniques and methodology of geography have changed, refined and some even discarded.

Early geographies: Exploration and human-environment relations

Early explorers might have caught the travel bug independently, but they have to find sponsors for their excursions. They observed, asked questions, they tested, smell, heard, and felt different places; they took notes and wrote diaries and drew maps and brought 'specimens' back to the folks at home. In sum, they produced qualitative and quantitative data about what they experienced synthesized these data with their existing knowledge and produced representations that were heavily and unabashedly skewed by their own perspectives.

Early geographies: Regional approach

The concept of ‘the region’ has a long standing position in emerging geographies over centuries, but in the late nineteenth and early twentieth centuries, the task of regional descriptions became the distinct purview of geography of various players focusing on different strategies. For example: (i) some scholar worked to identify the core commonalities that would allow a region to be classified as a region (such as H. J. Mackinder); (ii) some took more systematic approach to define categories of regions across the globe (such as A. J. Herbatson’s attempt to use climate categories to define the earth’s regions).

The primary methodology of ‘old’ regional geography could perhaps be summed up by the term ‘cataloguing’. Counting and quantifying the instances of phenomena, whether they were characteristics of slope and elevation or agricultural practices, and linking them to the locations.

The early twenty-first century cataloguing and description seemed to employ a mixed bag of what we would now call qualitative and quantitative methods, using field observations, historical climate records, population census, and production records and talking with residences directly. While in many ways ‘old regional geography’ and its tendency toward uncritical, theoretical description are derisively held up as a dark chapter in our past.

Before the Second World War

Geographers, for more than two hundred years, had been confronted with the problems of generalization and theory building (Cone, 1986). According to ĀtaśaBhāratīya (2008), Geography as an academic discipline encountered a lot of challenges on its methodological approaches leading to crisis, divisions (dualisms) and even closing of many geography departments and courses in universities, some of these challenges include:

- i. The closing of many geography departments and courses in universities took place, e.g., the abolition of the geography program at Harvard University (a highly prestigious institution) in 1948.
- ii. There was continuing division between human and physical geography – general talk of human geography becoming an autonomous subject.
- iii. Geography was regarded as overly descriptive and unscientific –it was claimed that there was no explanation of why processes or phenomena occurred.
- iv. Geography was seen as exclusively educational –there were few if any applications of contemporary geography.

- v. Continuing debates regarding what geography is –science, art, humanity or social science–took place.
- vi. After World War II, technology became increasingly important in society, and as a result, nomothetic-based sciences gained popularity and prominence.

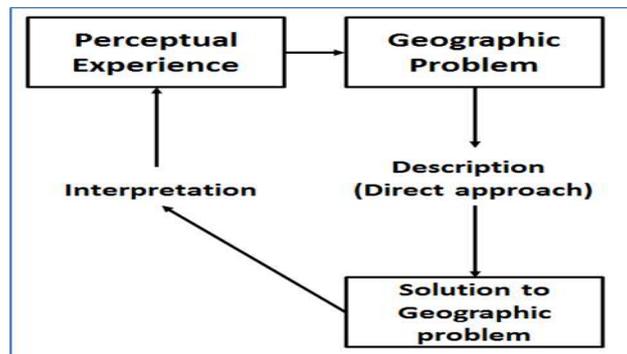


Figure 1: How geographers use their experiences to address geographical problems before the Quantitative Revolution (ĀtaśaBhāratīya 2008)

After the Second World War (i.e. between the 1950s and 1960s)

The Quantitative Revolution began in the universities of Europe with the support of geographers and statisticians in both Europe and the United States. The empirical descriptive geography was discarded and more stress was laid on the formulation of abstract models. The diffusion of statistical techniques in geography, to make the subject and its theories precise, is known as the Quantitative Revolution in geography (Nel, 2010).

Some techniques that epitomize the Quantitative Revolution include the following:

- i. Descriptive statistics
- ii. Inferential statistics
- iii. Basic mathematical equations and models, such as gravity model of social physics, or the Coulomb equation.
- iv. Stochastic models using concepts of probability, such as spatial diffusion processes
- v. Deterministic models, e.g. Von Thünen's and Weber's location models.

The earliest use of quantitative techniques started in climatic studies such as (i) Koeppen's climatic classification, (ii) R.R. Croze's *The Analysis of Rainfall Probability*, (iii) M.G. Kendall's "The Geographical Distribution and Crop Productivity in England", (iv) H.A. Matthew's "A New View of

Some Familiar Indian Rainfall”. Christaller made a major contribution to location theory, by applying quantitative techniques enormously in his study Central Places in Southern Germany.

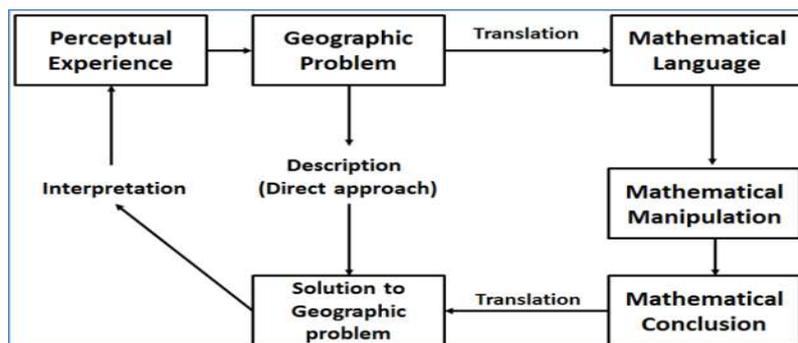


Figure 2: How geographers use their experiences to address geographical problems after the Quantitative Revolution (ĀtaśaBhāratīya 2008)

Post-revolution Geography

The greater use of computers in geography also led to many new developments in geomatics, such as the creation and application of GIS and remote sensing. These new developments allowed geographers for the first time to assess complex models on a full-scale model and over space and time. Further advances also led to a greater role of spatial statistics and modelling within geography.

Methodological Developments in Social Geography

Social geography is a subfield of human geography. It gained importance during the 1960s and 1970s. The focus in social geography has shifted from the study of ‘social area analysis’, ‘poverty’, and ‘segregation’ using a quantitative approach to unravelling the understanding of ‘social inequalities’ and the study of ‘geography of differences’ and ‘social justice’ using the humanistic and Marxist approaches (Harvey, 1973). Jones (1984) posited that quantitative methodology in social geography cannot explain social processes since it does not take cultural values into consideration. To study social issues a variety of approaches have been utilized: quantitative, qualitative, GIS, feminist and mixed method approaches.

According to Thakur (2012) two methodologies were important in social geographic research – quantitative and qualitative. The quantitative methodology became popular in social geography since it was based on scientific principles and methods. This approach, identified with positivism (the philosophy of science), was applied in natural and engineering sciences to understand scientific phenomenon. The methodology was subsequently applied to study social geographic problems such as urban social areas, poverty, crime, housing, food and nutrition, culture, health and diseases.

The second methodology adopted by social geographers was the qualitative approach. These tools provide access to the motives, aspirations and relationships that account for how places, people and events are composed and represented. This methodology includes various techniques such as biographic study, phenomenological studies, grounded theory studies, ethnographic studies, and case studies.

A review of selected areas or themes in social geographic research demonstrates the shifts in the application of any one methodology towards a hybrid approach to understand social geographic phenomenon.

- In social area analysis, the factorial ecology models have been important to unravel the urban socio-spatial differentiation in developing and developed nations.
- A correlation and regression analysis has been applied to identify the determinants of the spatial dimensions of poverty.
- Housing analysis has utilized indices to measure the degree of separation between various ethnic or racial groups.
- A quantitative and geographical systems approach has been utilized to study the spatial patterns of healthcare and the relationships between environment and patterns of ill-health of the population.
- Location-allocation modelling approach provided an evaluation of the location and accessibility to health facilities in developing countries.
- The study of crime in social geography has focused on the study of offenders, offences and the location of crime in geographical space.
- A quantitative and GIS approach has been utilized to identify the determinants of crime in various cities and localities.
- The social dimension of urban and regional planning has been studied by the application of spatial analysis and geographic technologies to provide insights into the development process in both the developed and developing world.

Therefore it can be said that, no single methodology is appropriate or sufficient to analyze social geographic issues or problems and there is a need for a hybrid approach to analyze social geographic problems.

Table 1: Understanding paradigms in geographical research along with methodology

Geographical paradigm	Theoretical assumptions	Data and analytic approaches	Researcher – object relations
<p>Spatial Science: mapping spatial relations</p>	<p>Space is viewed as a two - dimensional container within which places exist as discrete phenomena with distinct attributes.</p> <p>Variations across and interactions between places are measurable, usually quantifiable.</p> <p>Each object of analysis has a series of distinct characteristics, including its location within this two dimensional container.</p> <p>A descriptive mapping of the spatial locations and characteristics of objects is the first step in assessing causal relations.</p> <p>If theory suggests a causal relation and characteristics are found to spatially vary in systematic ways, then the theory has empirical support.</p>	<p>Data: Census (socio-demographic, economic); field methods (tree ring width, snow - pack depth); remote sensing (vegetation cover; slope); social surveys (consumer behaviour, political opinion).</p> <p>Analytic approaches: Descriptive and multivariate statistical analysis; mathematical analysis; simulation; GIS overlay; survey research methods; behavioural modelling; carbon dating.</p>	<p>Subjectivity controlled in favour of objective forms of research analysis and presentation.</p> <p>Researcher maintains a distance between her/himself and the object under study.</p>
<p>Humanism: interpreting places of meaning</p>	<p>Space as a two - dimensional container is but one way of looking at the world, commonly used by scientists but also governments.</p> <p>In the everyday sense, place is a more useful concept because it describes the attachment people have to particular parts of the world such as their home, workplace, or car, but also their neighbourhood and country.</p> <p>A “sense of place” differs for each individual and changes through time.</p> <p>Specific objects are associated with a particular</p>	<p>Data: Archival material (diaries, local histories); oral and other folk traditions; landscape reconstruction; place histories; individual life stories and experiences.</p> <p>Analytic Approaches: Ethnographic immersion; participant observation; textual analysis; structured and unstructured interviews; culturally – specific interpretation; “thick description”.</p>	<p>Researcher rejects the notion of objectivity and focuses on developing an empathic understanding of individual’s experiences as they relate to space.</p> <p>Focus can also be on what meanings imply “to me”, the researcher who is charged with the task of interpretation.</p>

	<p>“sense of place”, or produced as a means of expressing this sense.</p> <p>Specific objects are placed within the landscape in order to represent particular emotions and feelings.</p>		
<p>Critical Realism: uncovering the structures and mechanisms behind events</p>	<p>Events in the world are embedded within wider structures (e.g., capitalism and patriarchy) and their casual mechanisms (e.g., worker exploitation, gender division of labour).</p> <p>For many physical processes, explanation is of a “necessary” type: the causal powers operated the same from one place to another.</p> <p>For social relations, the mechanisms for causing change are usually contingent, or context specific.</p> <p>Paying attention to contingent contexts, including the different spaces in which things happen, requires in – depth case study.</p> <p>Social structures and mechanisms are differentiated across local and global contexts.</p>	<p>Data: Participatory observation and ethnography; interviews; field study of context - specific causal operations.</p> <p>Analytic Approaches: Reflection on and interpretation of the nature of events and the wider mechanisms that embed them, as well as the causal forces (structures) from which they emerge; in the physical sciences, researcher remains open to the possibility that causal mechanisms are complexly, contingently, and non linearly related.</p>	<p>In social analysis, researchers assume that individuals have a situated capacity to reflect upon structures and mechanisms, even though they might not have a well - developed theoretical language for articulating them.</p> <p>Even though the world exists objectively, our capacity for knowing it in an objective way is limited by our own contingently situated knowledge.</p>
<p>Post - structuralism: interrogating the production and import of spatial discourses</p>	<p>Discourses – the socially constructed ways we have to describe the world – name or “fix” the character of phenomena as well as the boundaries between the insane and rational, the possible and the impossible.</p> <p>Discourses have material effects; it “matters” how people and objects are characterized and brought into relation with one another.</p> <p>Though they are the vehicles through which</p>	<p>Data: Texts and visual data; ethnographic immersion and participant observation; descriptive accounts of micro - scale actions (practices) and events.</p> <p>Analytic Approaches: Textual and visual interpretation; genealogy of the emergence and deployment of discourses; deconstruction (analysis of the work of binaries).</p>	<p>Researcher and researched, human and non - human, are thoroughly embedded within wider fields of relation and meaning.</p> <p>Researcher eschews the notion of total knowledge, understanding that his or her views are from “somewhere” and thus are just as partial as their subjects.</p> <p>Reflexivity includes a monitoring of the researcher’s positions with respect to the settings</p>

	<p>power – laden relations are formed and operate; discourses are variable and therefore open to transformation.</p> <p>Space is socially constructed from the unequal relations of power that seek to fix the nature of people and things. Not just discourses, but matter too is a force in the world.</p>		<p>of which they are a part. Researcher is her/himself an object of analysis.</p>
--	--	--	---

Note: Adapted from Del Casino et al. (2000) and Del Casino (2009)

Need for Research

The aims of scientific research coincide with the type of research, i.e., whether it is exploratory research or explanatory research or descriptive research. In other words, it depends upon the general goals (understanding for its own sake), the academic goals, the theoretical goals and the pragmatic goals of research. Broadly speaking, the important aims of scientific research are:

Need for scientific research

The goal of scientific research is to discover laws and postulate theories that can explain natural or social phenomena, or in other words, build scientific knowledge. Given that theories and observations are the two pillars of science, scientific research operates at two levels: a theoretical level and an empirical level.

- **The theoretical level** is concerned with developing abstract concepts about a natural or social phenomenon and relationships between those concepts (i.e., build “theories”).
- **The empirical level** is concerned with testing the theoretical concepts and relationships to see how well they reflect our observations of reality, with the goal of ultimately building better theories.

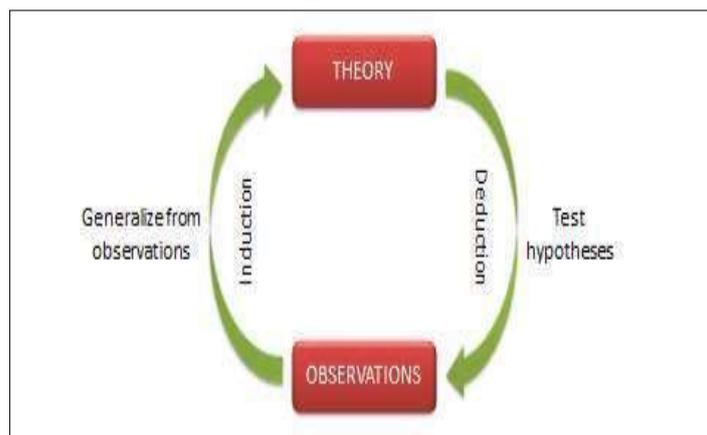


Fig. 3: The Cycle of Scientific Research.

The need of scientific research can be summarised as:

- it improves decision-making;
- it reduce uncertainty;
- it enables adopting new strategies;

- it helps in planning for future; and
- it helps in ascertaining trends.

The scientific inquiry should not be conducted when availability of adequate data is doubtful, there is a time constraint.

Need for social research

- To understand the functioning of society.
- To study individual behaviour and social action.
- To evaluate social problems, their effects on society, and to find out possible solutions.
- To explore social reality and explain social life.
- To develop theories.

Becker (1989) and **Sarantakos** (1998) have referred to the following goals of social research:

- **General goals:** Understanding for its own sake.
- **Theoretical goals:** Verification, falsification, modification or discovery of a theory.
- **Pragmatic goals:** Solution of social problems.
- **Political goals:** Development of social policy, evaluation of programmes, planning of reconstruction, empowerment and liberation.

Mahr (1995) has outlined the following motives of social research:

- **Educational:** to educate and inform the public.
- **Personal:** to promote the academic status of the researcher.
- **Institutional:** to enhance the research quantum of the institutions for which the researcher works.
- **Political:** to provide support to political plans and programmes.
- **Tactical:** to delay decision or action for as long as the investigation is under way.

According to **Kothari** (2004), the significance of research can also be understood keeping in view the following points:

- To those students who are to write a master's or Ph.D. thesis, research may mean a careerism or a way to attain a high position in the social structure;
- To professionals in research methodology, research may mean a source of livelihood;
- To philosophers and thinkers, research may mean the outlet for new ideas and insights;
- To literary men and women, research may mean the development of new styles and creative work;
- To analysts and intellectuals, research may mean the generalisations of new theories.

Thus, research is the fountain of knowledge for the sake of knowledge and an important source of providing guidelines for solving different business, governmental and social problems. It is a sort of formal training which enables one to understand the new developments in one's field in a better way

Basic Research Types

The main purpose of research area: to explore, to describe and to explain. On this basis, we may illustrate three types of research:

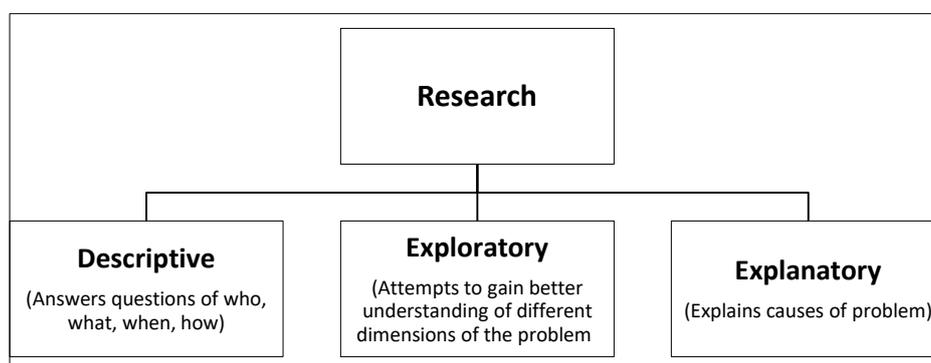


Fig. 4: Major Types of Research

The other types of research area: (i) pure and applied, (ii) experimental and evaluation, (iii) qualitative and quantitative and (iv) longitudinal and comparative.

Exploratory research

This research studies subject about which either no information or little information is available. Generally, this type of research is qualitative which becomes useful in formulating hypotheses or testing hypotheses and theories. In this research, the assumption is that the researcher has little or no knowledge about the problem or situation under study, or he is unfamiliar with the structure of the group he is studying (say, prison, industry, university, village, and so on). As an example, in an exploratory study of a prison, the researcher points out how a prison is divided in barracks and wards; what type of work is assigned to different types of prison officers; what recreational, medical, educational facilities are provided to prisoners; what rules they have to follow in interacting with other inmates or with officials; how are they to maintain contacts with the outside world; and so on.

Descriptive research

This type of research describes social-physical situations, events, systems, structures, etc. The researcher observes/studies and then describes what did he find? Take, for example, the research on drug abuse. The Ministry of Social Welfare, Government of India assigned this study in 1976, 1986 and 1996 to teams of scholars (doctors, sociologists, criminologists) to study the extent of drug abuse among college students, nature of drugs taken, causes of taking drugs, sources of getting drugs, effects of taking drugs, and so on.

Pure research

This research, also called basic research, is concerned with quest for knowledge and knowing more about the phenomenon without concern for its practical use and also with developing and testing hypotheses and theories. For example, developing a theory pertaining to the functioning of group mind (collective behaviour) or group dynamics. This type of research is also used to reject or support the existing theories about social phenomena.

Applied research

This type of research is concerned with search for ways of using scientific knowledge to analyse and solve practical problems. This research is generally conducted on large scale basis, it is expensive. This type of research is conducted on interdisciplinary basis by the financial support of World Bank, UNICEF, UGC, ICSSR etc.

Quantitative research

This research employs quantitative measurement and the use of statistical analysis. for example, percentage of medical, engineering, law, arts, science and commerce students take drugs or use alcohol. How many man days have been lost due to strikes and lockouts in industries in India in last two decades? This type of research is based on the methodological principles of positivism.

Qualitative research

This research presents non-quantitative type of analysis. It describes reality as experienced by the groups, communities, individuals etc. For example, how does the structure and organisation of wall-less prisons (or minimum security jails) differ from that of the central or district jails (or maximum security jails) and contribute to the reformation and re-socialisation of criminals?

Comparative research

In this research, the similarities and differences between different socio-physical units or cultural or social groups are studied. For example, comparing marriage system of Hindus with Muslims, art and culture of tribals with non-tribals, customs and social practices of rural people with urban people, the difference between Indian family and American family, nature and causes of crimes committed by females in India with those of females in America, England, Canada, etc.

Longitudinal research

A longitudinal study (or longitudinal survey, or panel study) is a research design that involves repeated observations of the same variables (e.g., people) over short or long periods of time (i.e., uses longitudinal data). Example: prevalence of HIV-AIDS among males and females in India in 1979, 1989 and 1999.

Cross sectional research

A cross-sectional study (also known as a cross-sectional analysis, transverse study, prevalence study) is a type of observational study that analyzes data from a population, or a representative subset, at a specific point in time—that is, cross-sectional data.

3. Identification of research problems, development of theoretical background - literature review, research gap and research question and specification of the objectives of study; hypothesis building, Framework of research writing.**Steps in scientific research**

Manheim (1980) has suggested nine steps in scientific research which includes:

- Step 1: Casual observation of event/phenomenon
- Step 2: Wonder/curiosity about how, why, what etc.

- Step 3: Hypothesis (about relationship between two variables)
- Step 4: Preparing design of research
- Step 5: Data collection, processing, analysis and interpretation
- Step 6: Determining whether hypothesis is true or false
- Step 7: Description/inference
- Step 8: Prediction (using inductive process)
- Step 9: Practical application

Horton and Hunt (1984) pointed out eight steps in scientific research:

- Step 1: Define the problem
- Step 2: Review of literature
- Step 3: Formulation of hypothesis
- Step 4: Plan the research design
- Step 5: Collected the data
- Step 6: Analysis the data
- Step 7: Draw conclusion
- Step 8: Replicate the study

All the steps of scientific research have been discussed one by one in the following sections:

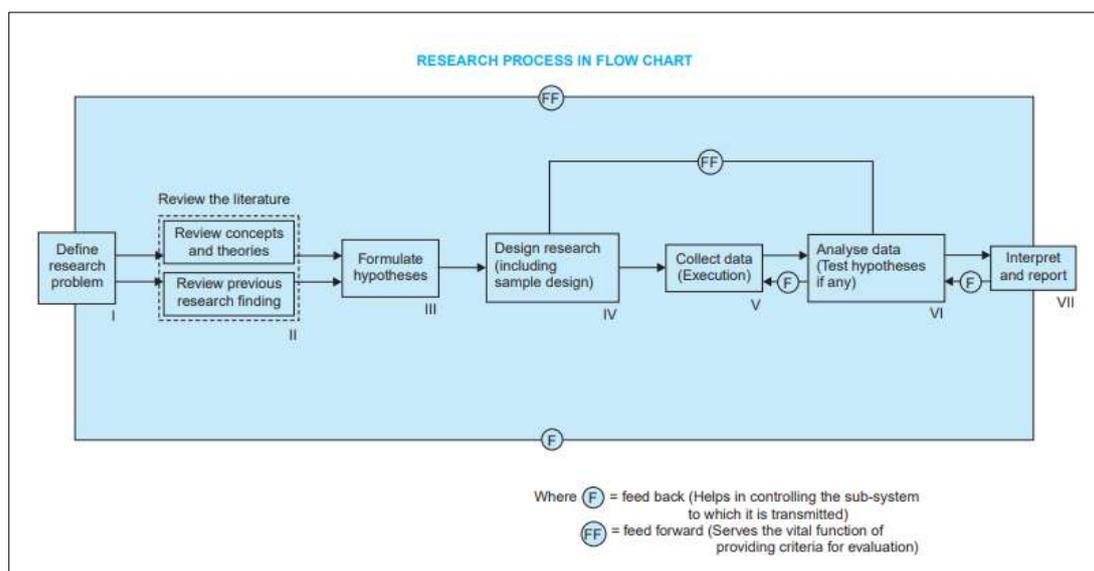


Fig. 5: Research Process in Flow Chart, **Source:** Kothari (2004)

Identification of Research Problem/Define Research Problem

A Research problem is a question that a researcher wants to answer or a problem that a researcher wants to solve. A research problem is an issues or a concern that an investigator / researcher presents and justifies in a research study. Identification and formulation of a research problem is the first step of the research process. It is the most challenging and difficult phase of the research process. Selection of a research problem depends on several factors such as researcher’s knowledge, skills, interest, expertise, motivation & creativity with respect to the subject chosen. A good research studies needs a lots of time for selection of research problem. Initially every researcher faces a predicament of identifying, selecting and formulating a good research problem.

According to **Kerlinger**, ‘A problem is an interrogative sentence or statement that asks what relation exists between two or more variable. The answer to question will provide what is having sought in the research.

R.S. Woodworth defines problem as ‘a situation for which we have no ready & successful response by instinct or by previous acquired habit. We must find out what to do’, i.e. the solution can be found out only after an investigation.

Selecting the problem

According to **Kothari** (2004), the following points may be observed by a researcher in selecting a research problem or a subject for research:

1. Subject which is overdone should not be normally chosen, for it will be a difficult task to throw any new light in such a case.
2. Controversial subject should not become the choice of an average researcher.
3. Too narrow or too vague problems should be avoided.
4. The subject selected for research should be familiar and feasible so that the related research material or sources of research are within one's reach.
5. The importance of the subject, the qualifications and the training of a researcher, the costs involved, and the time factor are few other criteria that must also be considered in selecting a problem.
6. The selection of a problem must be preceded by a preliminary study.

Technique involved in defining a problem

The technique for the purpose involves the undertaking of the following steps generally one after the other:

1. Statement of the problem in a general way: First of all the problem should be stated in a broad general way, keeping in view either some practical concern or some scientific or intellectual interest. The problem stated in a broad general way may contain various ambiguities which must be resolved by cool thinking and rethinking over the problem. At the same time the feasibility of a particular solution has to be considered and the same should be kept in view while stating the problem.

2. Understanding the nature of the problem: The next step in defining the problem is to understand its origin and nature clearly. The best way of understanding the problem is to discuss it with those who first raised it in order to find out how the problem originally came about and with what objectives in view. The researcher should also keep in view the environment within which the problem is to be studied and understood.

3. Survey of available literature: All available literature related to research problem must be analysed and examined. Studies on related problems are useful for indicating the type of difficulties that may be encountered in the present study as also the possible analytical shortcomings. At times such studies may also suggest useful and even new lines of approach to the present problem.

4. Developing ideas through discussions: Various new ideas can be developed through such an exercise. Hence, a researcher must discuss his problem with his colleagues and others who have

enough experience in the same area or in working on similar problems. This is quite often known as an experience survey.

5. Rephrasing research problem: Finally, the researcher must sit to rephrase the research problem into a working proposition. Through rephrasing, the researcher puts the research problem in as specific terms as possible so that it may become operationally viable and may help in the development of working hypotheses.

In addition to what has been stated above, the following points must also be observed while defining a research problem:

- Technical terms and words or phrases, with special meanings used in the statement of the problem, should be clearly defined.
- Basic assumptions or postulates (if any) relating to the research problem should be clearly stated.
- A straight forward statement of the value of the investigation (i.e., the criteria for the selection of the problem) should be provided.
- The suitability of the time-period and the sources of data available must also be considered by the researcher in defining the problem.
- The scope of the investigation or the limits within which the problem is to be studied must be mentioned explicitly in defining a research problem.

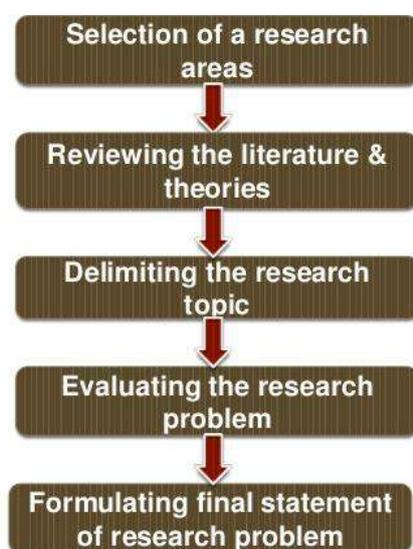


Fig. 6: Technique involved in defining a research problem

Some sources of identifying the research problem

Personal experience: Day to day experience of the researcher serves as a good source of ideas to formulate research problem

Practical experience: Clinical experience provides nurses ample opportunities for problem identification.

Critical appraisal of literature: When one looks at books, articles, summaries of clinical issues relating to the subject of our interest, pertinent questions may arise.

Previous research: A body of knowledge is developed on a sound research finding. Usually at the end of the research problems are suggested based on the shortcoming of the previous research, these could be investigated.

Existing theories: Research is a process of theory development and theory testing. If an existing theory is used in developing a researchable problem, a specific statement from theory must be isolated. Generally a part or parts of a theory are subjected to testing in a clinical situation. The testing of a theory is needed therefore they serve as a good source of research problem.

Social issues: Sometimes issues of global contemporary or political issues of relevance are suggested. An idea for a research may stem from a familiarity with social concerns or controversial social issues.

Brainstorming: Brainstorming sessions are good techniques to find new research questions. Brainstorming refers to intensified discussions among interested people of the profession in order to find more ideas to formulate a good research problem.

Intuition: It is believed that reflective mind is a good source of ideas, which may be used to find out new and good research problems.

Exposure to field situations: During field exposure researchers get a variety of experiences which may provide plenty of ideas to formulate research problems.

Consultations with experts: Experts are believed to have sound experience of their respective field, which may suggest a significant problem to be studied.

Development of theoretical background

Literature review

A literature review or narrative review is a type of review article. A literature review is a scholarly paper that presents the current knowledge including substantive findings as well as theoretical and methodological contributions to a particular topic.

Writing a literature review involves finding relevant publications (such as books and journal articles), critically analyzing them, and explaining what you found. There are five key steps:

1. Search for relevant literature
2. Evaluate sources
3. Identify themes, debates and gaps
4. Outline the structure
5. Write your literature review

A good literature review doesn't just summarize sources – it analyzes, synthesizes, and critically evaluates to give a clear picture of the state of knowledge on the subject.

What are the possible sources of information?

1. Individuals, groups, and organisations
2. Published information (books, articles, indexes, abstract journals)
3. Unpublished information (other research proposals in related fields, reports, records, computer data bases)

Types of literature review

The main types of literature reviews are: evaluative, exploratory, and instrumental.

Systematic review: It focused on a research question, trying to identify, appraise, select and synthesize all high-quality research evidence and arguments relevant to that question. A meta-analysis (statistical analysis) is typically a systematic review using statistical methods to effectively combine the data used on all selected studies to produce a more reliable result.

Integrative literature: The purpose of an integrative literature review is to generate new knowledge on a topic through the process of review, critique, and then synthesis of the literature under investigation.

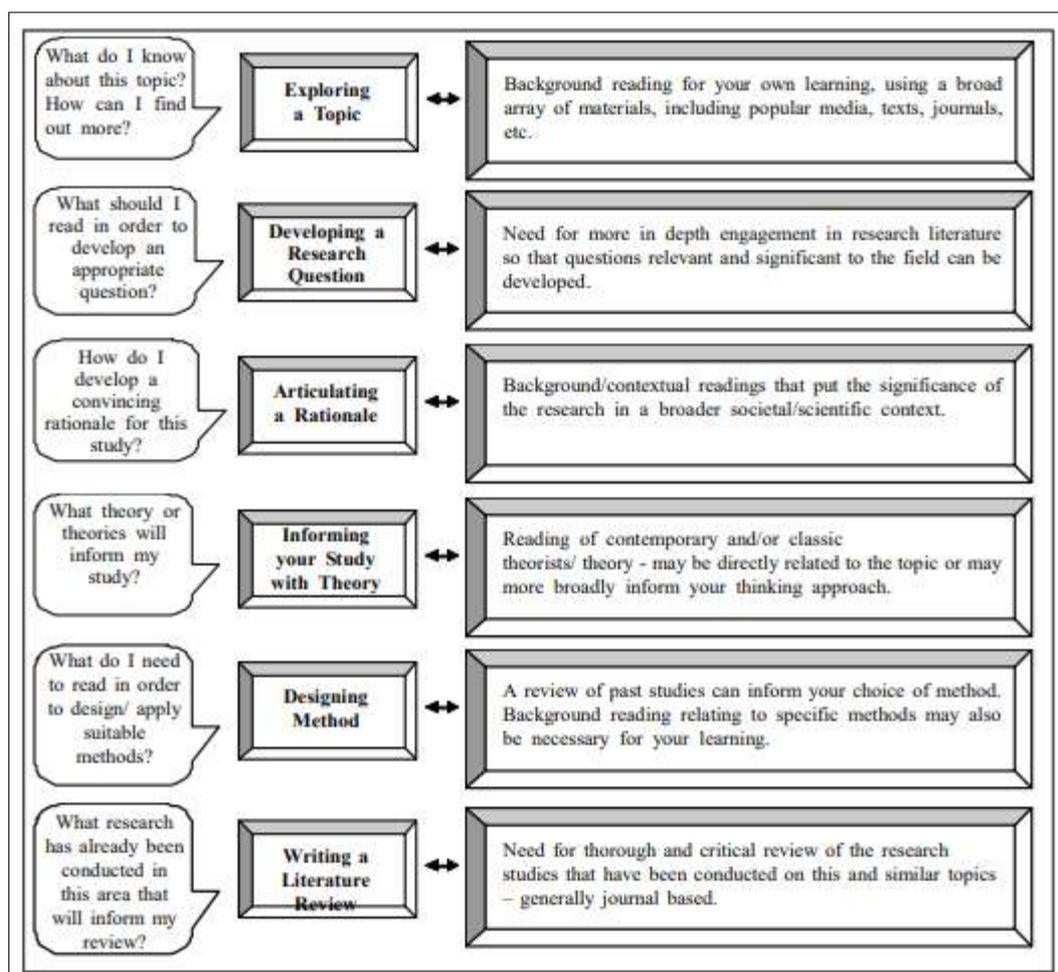


Fig. 7: Methods of literature review (Source: O'Leary, 2004)

Research Gap

The gap, also considered the missing piece or pieces in the research literature, is the area that has not yet been explored or is under-explored. This could be a population or sample (size, type, location, etc.), research method, data collection and/or analysis, or other research variables or conditions. Research gaps are particularly useful for the advance of science in general.

How to identify gaps in literature?

- Look for inspiration in published literature
- Seek help from your research advisor
- Use digital tools to seek out popular topics or most cited research papers
- Check the websites of influential journals

- Make a note of your queries
- Research each question

Objective of Research

Having identified a real world problem and framed it in terms of its relevance to a policy maker, the researcher must, then, pose an answerable research question or hypothesis. This essentially defines the objective of the project. The objective of a certain research summarizes what is to be achieved by the study. It should clearly define the question for which a solution is being sought. You should consider the following points on stating research objectives in order to be clearly related to the statement of the problem cover the different aspects of the problem be clearly phrased in operational terms specifying exactly what you are going to do, where and for what purpose.

- Be realistic considering local conditions and available resources.
- Use specific action verbs i.e. to determine, to identify, to verify, to describe, to calculate.
- Avoid vague non-action verbs i.e. to study, to appreciate, to understand.

Types of research objectives

1. General Objectives: These will define what is expected to be achieved by the study in general terms. For example, if the research problem chosen is “**Atrocities against tribals in the Area X**”, the general objective could be “to identify the reasons for atrocities against tribals in the Area X, so as to suggest remedial measures”.

2. Specific Objectives: Here, the general objective is broken down into smaller and logically connected units to address the various aspects of the problem and the key factors that are assumed to cause or influence the problem. The specific objectives of a research work should focus on the following points:

- 1) To identify the distribution and patterns of a problem.
- 2) To examine the possible factors that may influence the problem.
- 3) To indicate how the research results will be used. Returning to our example on atrocities against tribals, the specific objectives for a research work in this problem could be set in the following manner:
 - i. To determine the prevalence of atrocities against tribals in the Area X.

- ii. To identify the socio-economic characteristics of the exploited tribals and the exploiters.
- iii. To examine the mode of exploitation and oppression of the tribals in the Area X.
- iv. To identify the socioeconomic and cultural factors that may influence atrocities against tribals in the Area X.

Hypothesis Building

According to **Theodorson** and **Theodorson** (1996), “a hypothesis is a tentative statement asserting a relationship between certain facts”. **Kerlinger** (1973) describe it as “a conjectural statement of the relationship between two or more variables”.

A hypothesis can be defined as, a ‘Logical conjecture (hunch or educated guess) about the nature of relationships between two or more variables expressed in the form of a testable statement’. A hypothesis is a suggested or assumed explanation for the research problem that can be tested. It is a prediction for a relationship between one or more factor and the problem under study.

In our example (Example: **Atrocities against tribals in the Area X**) about atrocities against tribals we can formulate and test the following hypotheses:

- (1) The higher caste, landowning, non-tribals attack and exploit the tribals from poor socioeconomic backgrounds.
- (2) Women, children, and elderly tribals are more vulnerable to such violence.
- (3) Police inaction/leniency is the root cause of persistence of atrocities against tribals.

Importance of hypothesis

1. Hypothesis enables the researcher to objectively investigate new areas of discovery, thus helps in theory building.
2. Hypothesis provides objectivity to the research activity.
3. Hypothesis provides directions to conduct research.
4. Hypothesis provides clear and specific goals to the researchers. These goals provide the investigator with a basis for selecting sample and research procedures to meet the set goals.

5. Hypothesis provides a link between theories and actual practical research. It serves as a bridge between theory and reality.
6. A hypothesis suggests the methodology to be chosen to conduct the study.
7. Hypothesis is a tentative statement of the anticipated results, it guides the researcher towards the direction in which the research should proceed.
8. It stimulates the thinking process of the researcher.
9. It provides directions about the methodology to be chosen & techniques of data analysis.
10. Hypothesis provides an understanding of the researchers about what to expect from the results of the study. It serves as a framework for drawing conclusions of a research study.
11. Without a hypothesis a research would be an aimless wandering.

Characteristics of a good hypothesis

Conceptual clarity: Clearly defined and understandable concepts. The meaning and implications of which cannot be doubted.

Empirical referents: A research must have an ultimate empirical referent and must have empirical basis from the area of enquiry.

Objectivity: Hypothesis must be objective (objectivity in data collection).

Specificity: Hypothesis must be specific, not general and should explain the expected relations between variables. (E.g. Exercise and stress)

Relevant: The hypothesis should be relevant to the problem and objectives under enquiry.

Testability: Hypothesis should be testable and should not be a moral judgment. It should be directly or indirectly measurable.

Causal link: A good hypothesis states the causal link between independent and dependent variables, which is later evaluated by using inferential statistical tests.

Consistency: It should correspond with existing knowledge.

Availability of technique: The researcher must ensure that scientific methods are available for testing their proposed thesis.

Purposiveness: It refers to the relevance of hypothesis to the research problem and its objectives.

Verifiability: A good hypothesis should be verifiable in practical terms.

Economical: The expenditure of resources can be controlled if the hypothesis underlying the research that which is undertaken is good.

Sources of hypothesis

Theoretical framework: Theoretical framework or conceptual frameworks are the most important sources of hypothesis.

Deductive reasoning: Through a deductive approach these hypothesis are drawn for testing them. E.g., Pavlov's theory – a hypothesis can be drawn about modification of behaviour.

Previous research: Findings or previous research may be used for framing the hypothesis for a new study that have an area of relevance.

Real-life experiences: Real life experiences may contribute in the formulation of hypothesis.

Academic literature: Academic literature (formal theories, empirical evidences, experiences, observations and conceptualizations of academicians) may serve as good source for formulating hypothesis for research studies.

Types of hypothesis

A. Simple/Complex hypothesis

1. **Simple hypothesis:** It is a statement which reflects the relationship between two variables. (E.g., low levels of haemoglobin & risk of infection)
2. **Complex hypothesis:** It is a statement which reflects the relationship between more than two variables. (E.g., satisfaction among patients who are older & dwelling in rural areas than those who are younger & living in urban areas)

B. Associative/Causal hypothesis

1. **Associative hypothesis:** Reflects a relationship between variables that occurs or exists in natural settings without manipulation. This hypothesis is use in correlational research studies.
2. **Causal hypothesis:** Predicts the cause-and-effect relationship between two or more dependent and independent variables in experimental or interventional setting, where

independent variable is manipulated by researcher to examine the effect on dependent variable.

C. Directional/Non-directional hypothesis

1. **Directional hypothesis:** Specifies not only the existence, but also the expected direction of the relationship between variables. E.g., there is a positive experience between years of teaching experience and job satisfaction among teachers.
2. **Non-directional hypothesis:** Reflects the relationship between two or more variables, but it does not specify the anticipated direction and nature of relationship such as positive or negative. E.g., there is a relationship between years of teaching experience and job satisfaction among teachers.

D. Null hypothesis/Research hypothesis

1. **Null hypothesis (H₀):** It is also known as *Statistical Hypothesis* and is used for statistical testing and interpretation of statistical outcomes. E.g., there is no relationship between smoking and incidence of lung cancer.
2. **Research hypothesis (H₁):** It states the existence of relationship between two or more variables. E.g., there is relationship between smoking and incidence of lung cancer.

The Concept of Theory

A theory is a well-established principle that has been developed to explain some aspect of the natural world. Theories arise from repeated observation and testing and incorporate facts, laws, predictions, and tested assumptions that are widely accepted [e.g., rational choice theory; grounded theory; critical race theory].

The key distinctions between theory and hypothesis are:

1. A theory predicts events in a broad, general context; a hypothesis makes a specific prediction about a specified set of circumstances.
2. A theory has been extensively tested and is generally accepted among scholars; a hypothesis is a speculative guess that has yet to be tested.

Source:

Cherry, Kendra. Introduction to Research Methods: Theory and Hypothesis. About.com Psychology; Gezae, Michael et al. Welcome Presentation on Hypothesis. Slideshare presentation.

Framework of Research Writing

A research framework clearly illustrates the structure of the research plan and helps the researcher formulate relevant research questions. A research framework provides an underlying structure or model to support our collective research efforts.

1. Theoretical framework

Imenda (2014) defined a theoretical framework “as the application of a theory, or a set of concepts drawn from one and the same theory, to offer an explanation of an event, or shed some light on a particular phenomenon or research problem”.

A theoretical framework consists of concepts and, together with their definitions and reference to relevant scholarly literature, existing theory that is used for your particular study. The theoretical framework must demonstrate an understanding of theories and concepts that are relevant to the topic of your research paper and that relate to the broader areas of knowledge being considered. Following are some strategies to develop of an effective theoretical framework:

- Examine your thesis title and research problem.
- Brainstorm about what you consider to be the key variables in your research.
- Review related literature to find how scholars have addressed your research problem.
- List the constructs and variables that might be relevant to your study.
- Review key social science theories that are introduced to you in your course readings
- Discuss the assumptions or propositions of this theory and point out their relevance to your research.

Structure and writing style of theoretical framework

The theoretical framework may be rooted in a specific theory, in which case, you are expected to test the validity of an existing theory in relation to specific events, issues, or phenomena.

However, you may not always be asked by your professor to test a specific theory in your paper, but to develop your own framework from which your analysis of the research problem is derived. Given this, it is perhaps easiest to understand the nature and function of a theoretical framework if it is viewed as the answer to two basic questions:

- What is the research problem/question? [e.g., "How should the individual and the state relate during periods of conflict?"]

- Why is your approach a feasible solution? [I could choose to test Instrumentalist or Circumstantialists models developed among Ethnic Conflict Theorists that rely upon socio-economic-political factors to explain individual-state relations and to apply this theoretical model to periods of war between nations].

The answers to these questions come from a thorough review of the literature and your course readings and the gaps in the research that emerge from the review process. With this in mind, a complete theoretical framework will likely not emerge until after you have completed a thorough review of the literature.

In writing this part of your research paper, keep in mind the following:

- Clearly describe the framework, concepts, models, or specific theories that underpin your study.
- Position your theoretical framework within a broader context of related frameworks, concepts, models, or theories.
- The present tense is used when writing about theory.
- You should make your theoretical assumptions as explicit as possible. Later, your discussion of methodology should be linked back to this theoretical framework.
- Don't just take what the theory says as a given.

2. Conceptual framework

Conceptual framework is the researcher's own model illustrating variables that specify the problem and gives direction to the study. It may be adaptation of a model in an early theory, with modifications to suit the theory. There are various types of conceptual framework in research writing, e.g. (1) **Working hypothesis** – exploration or exploratory research; (2) **Pillar questions** - exploration or exploratory research; (3) **Descriptive categories** – description or descriptive research; (4) **Practical idea type** – analysis; (5) **Models of operation research** – decision making; (6) **Formal hypothesis** – explanation and prediction etc.

Steps on how to make the conceptual framework

Before you prepare your conceptual framework, you need to do the following things:

- **Choose your topic:** Decide on what will be your research topic. The topic should be within your field of specialization.

- **Do a literature review:** Review relevant and updated research on the theme that you decide to work on after scrutiny of the issue at hand. Preferably use peer reviewed, and well-known scientific journals as these are reliable sources of information.
- **Isolate the important variables:** Identify the specific variables described in the literature and figure out how these are related. Some abstracts contain the variables, and the salient findings thus may serve the purpose. If these are not available, find the research paper’s summary. If the variables are not explicit in summary, get back to the methodology or the results and discussion section and quickly identify the study variables and the significant findings.
- **Generate the conceptual framework:** Build your conceptual framework using your mix of the variables from the scientific articles you have read. Your problem statement serves as a reference for constructing it. In effect, your study will attempt to answer the question that other researchers have not explained yet. Your research should address a knowledge gap.

Example 1: Conceptual framework of malnutrition

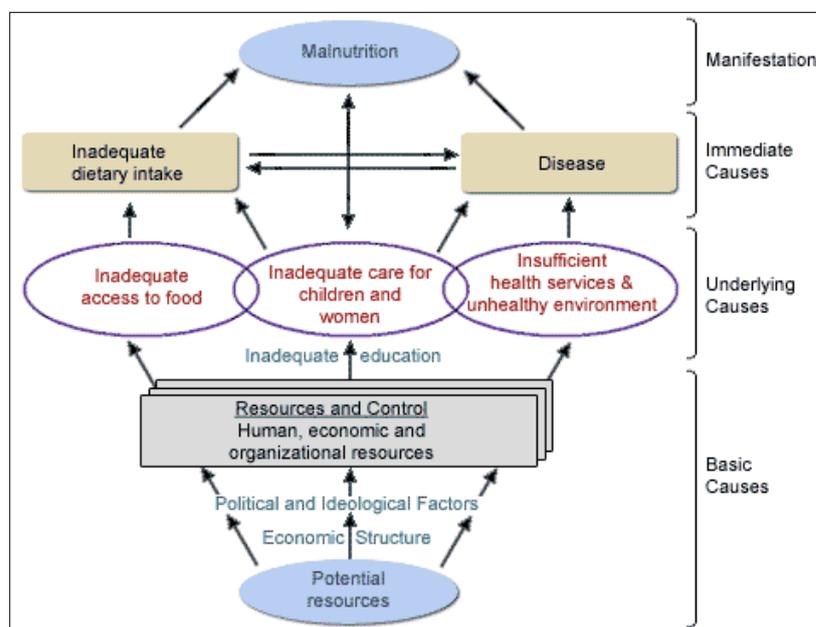


Fig. 8: Conceptual framework of malnutrition

Example 2: Conceptual framework on the relationship between malnutrition, infections and poverty.

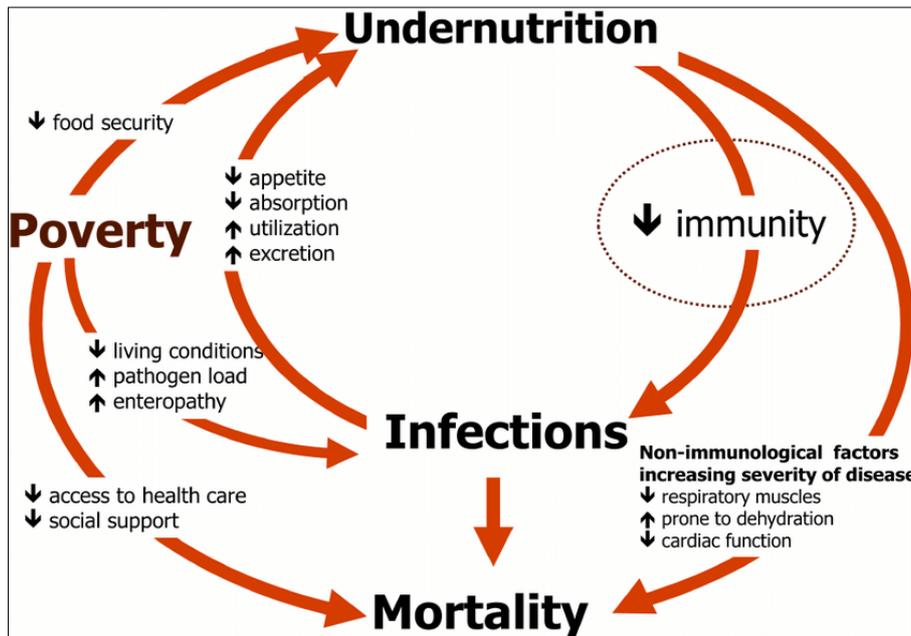


Fig. 9: Conceptual framework on the relationship between malnutrition, infections and poverty.

Example 3: Broad society: Patriarchal relationships with high gender inequality especially in sexual matters.

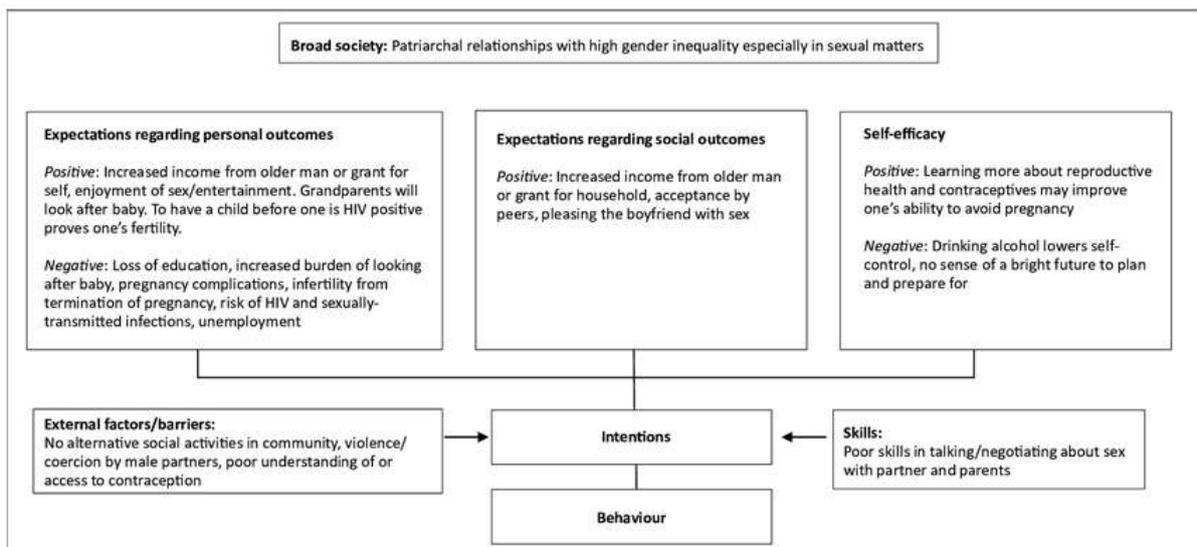


Fig. 10: Conceptual framework on Broad society: Patriarchal relationships with high gender inequality especially in sexual matters.

Table 2: Differences between conceptual and theoretical framework in research writing

Variable	Conceptual Framework	Theoretical framework
Genesis	a) Created by the researcher from a variety of conceptual or theoretical	Evolves or ‘takes shape’ from reviewed literature and/or the data

	perspectives	collected. Adopted / adapted from a pre-existing theory or theoretical perspective.
Purpose	<p>a) Helps the researcher see clearly the main variables and concepts in a given study;</p> <p>b) Provides the researcher with a general approach (methodology – research design, target population and research sample, data collection and analysis);</p> <p>c) Guides the researcher in the collection, interpretation and explanation of the data, where no dominant theoretical perspective exists.</p> <p>d) Guides future research – specifically where the conceptual framework integrates literature review and field data.</p>	<p>a) Helps the researcher see clearly the main variable and concepts in a given study.</p> <p>b) Provides the researcher with a general approach (methodology – research design, a target population and research sample, data collection & analysis); and</p> <p>c) Guides the researcher in the collection, interpretation and explanation of the data.</p>
Conceptual Meaning	Synthesis of relevant concepts.	Application of a theory as a whole or in a part.
Process Underlying Review of Literature	<p>a) Mainly inductive, as in social sciences where research problems cannot ordinarily be explained by one theoretical perspective;</p> <p>b) Some social science research also gets driven by theories, but theories in the social 'power' as those in the natural sciences.</p>	Mainly deductive, as in the natural sciences where hypothesis testing takes place to verify the 'power' of a theory.
Methodological Approach	<p>a) May be located in both quantitative and qualitative research paradigms; increasingly, mixed-methods approaches are recommended;</p> <p>b) Data mostly collected through both empirical and descriptive survey instruments, interviews and direct observations – hence, a preponderance of qualitative data;</p> <p>c) Strong on consideration of context.</p>	<p>a) Located mainly in the quantitative research paradigm;</p> <p>b) Data collected mainly through experimental designs, empirical survey and tests;</p> <p>c) Efforts made to standardize context, or else ignore it.</p>
Scope of Application	Limited to specific research problem and or context.	Wider application beyond the current research problem and context

Source: Imenda (2014)

Methods of Data Collection

1. Primary data collection

Primary data is a type of data that is collected by researchers directly from main sources through interviews, surveys, experiments, etc. Primary data are usually collected from the source—where the

data originally originates from and are regarded as the best kind of data in research. Primary data collection may involve the collection of both qualitative and quantitative primary information.

- **Quantitative information** is collected through measurement and may be processed using statistical and other techniques. Example: age, cost, length, height, area, volume, weight, speed, time, and temperature etc.
- **Qualitative information** is collected through observation or subjective judgment and does not involve measurement. Qualitative information may be processed or quantified where appropriate or it may be presented through images or as text (students are advised to remember the word limit when presenting qualitative information as text only). The nature of qualitative data should provide sufficient information for analysis and conclusion. Example: sex (male or female), name, state of origin, citizenship, etc.

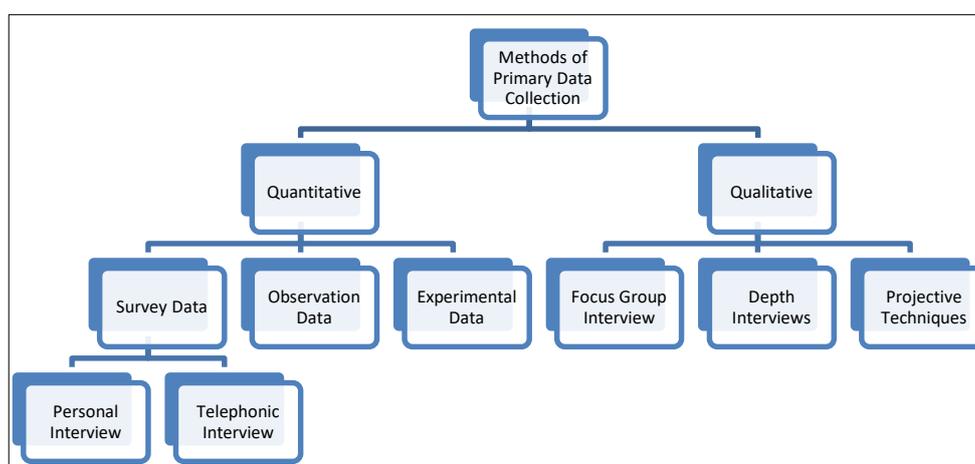


Fig. 11: Different Methods of Primary Data Collection

There are several methods of collecting primary data, particularly in surveys and descriptive researches. Important ones are:

i. Interviews

Interview is the verbal conversation between two people with the objective of collecting relevant information for the purpose of research. There are various types of interview which are as follows:

- **Personal interview:** Personal Interview communication is between a face to face two way interviewer and the respondents.
- **Telephone Interview:** Telephone interview the information is collected from the respondent by asking him questions on the phone is called as telephone interview.

- **Focus Group Interview:** Focus group interview is an unstructured interview which involves a moderator leading a discussion between small groups of respondents on a specific topic.
- **Depth Interview:** Depth interview is nondirective in nature where the respondent is given freedom to answer within the boundaries of the topic of interest.
- **Projective Techniques:** Projective Techniques involve the presentation of an ambiguous, unstructured object, activity or person that a respondent is asked to interpret and explain.

Some advantages of Interview Method are as follows:

- Opportunity for Feedback
- Probing Complex Answers
- Respondent will answer all questions asked
- Interviewers have the opportunity of showing respondents items such as sample products, graphs and sketches, which can aid in their answers.
- Interviewing respondents personally can increase the likelihood of their participation, many people participate directly.
- Interviewing selected subjects may provide more information, economically, than waiting for the spontaneous occurrence of the situation.
- The interviewer can collect supplementary information about the respondent's personal characteristics and environment which is often of great value in interpreting results.

Some disadvantages of Interview Method are as follows:

- Usually more expensive than mail, telephone and internet surveys
- There remains the possibility of the bias of interviewer
- This method is relatively more-time-consuming
- Interviewing at times may also introduce systematic errors.
- The interviewers approach and individual questioning demean or may influence the respondents' answers.

- It has been shown that the demographic characteristics of the interviewer can influence the answers of the respondents.
- Cultural aspects may influence peoples' willingness to participate in an interview (e.g. Brutal Middle Eastern cultures discourage females from being questioned by male interviewers)

ii. Observations

Observation is a technique that involves systematically selecting, watching, listening, reading, touching, and recording behaviour and characteristics of living beings, objects, or phenomena. The researchers, adopting this method, attempt to understand behaviour and societies by getting to know the persons involved and their values, rituals, symbols, beliefs, and emotions.

Observation is Helpful when:

- need direct information
- trying to understand ongoing behaviour
- there is physical evidence, products, or outputs than can be observed
- need to provide alternative when other data collection is infeasible or inappropriate

The major steps in an observational study as enumerated by Bailey (1987) are:

1. Decide upon the goals of the study;
2. Decide upon the group for subjects to be observed;
3. Gain entry to the group;
4. Gain rapport with the subjects being studied;
5. Conduct the study by observing and recording field notes over weeks, months or even years;
6. Deal with crises that occur, such as confrontations with subjects who think you are some sort of spy;
7. Exit from the observational study area;
8. Analyze the data;
9. Write a report presenting the findings.

Various types of observation (Classification)

A. Classification-1

- (1) **Structured Observation:** When the observation is characterized by a careful definition of the units to be observed, the style of recording the observed information, standardized conditions of observation and the selection of related data of observation.
- (2) **Unstructured Observation:** When it takes place without the above characteristics.

B. Classification-2

- (1) **Participant Observation:** When the observer is member of the group which he is observing then it is Participant Observation.

Examples:

- **Community meetings:** One of the most common participatory methods and it must be well organized. Must establish ground rules such as agree on purpose, who will speak, time allotted for speakers, format for questions and answers etc.
 - **Mapping:** Mapping of natural resources, social, health, individual or civic assets, wealth, land use and demographics through drawing or using existing maps.
 - **Transect walks:** Evaluator walks around community observing people, surroundings, and resources. Walk a transect line through a map of a community — line should go through all zones of the community. Need good observation skills.
- (2) **Non-Participant Observation:** When observer is observing people without giving any information to them then it is Non-Participant Observation.

C. Classification-3

- (1) **Uncontrolled Observation:** When the observation takes place in natural condition i.e., uncontrolled observation. It is done to get spontaneous picture of life and persons.
- (2) **Controlled Observation:** When observation takes place according to pre arranged plans, with experimental procedure then it is controlled observation generally done in laboratory under controlled condition.

Some advantages and disadvantages of observation method

Advantages

- Produces large quantities of data
- All data obtained from observations are usable
- The observation technique can be stopped or begun at any time
- Relative Inexpensive

Disadvantages

- It is expensive method
- Limited information
- Extensive training is needed

iii. Questionnaires (discussed in next sections)**iv. Survey schedule** (discussed in next sections)**v. Focus groups**

It is a qualitative research where small homogenous groups of people are brought together to informally discuss specific topics under the guidance of a moderator. Major purpose of this method is to identify issues and themes, not just interesting information, and not “counts”

Focus Groups are inappropriate when:

- Language barriers are insurmountable
- Evaluator has little control over the situation
- Trust cannot be established
- Free expression cannot be ensured
- Confidentiality cannot be assured

Table 3: Focus Group Process

Phase	Action
Opening	Ice-breaker; explain purpose; ground rules; introductions
Warm-up	Relate experience; stimulate group interaction; start with least threatening and simplest questions
Main body	Move to more threatening or sensitive and complex questions; elicit deep responses; connect emergent data to complex, broad participation
Closure	End with closure-type questions; summarize and refine; present theories, etc; invite final comments or insights; thank participants

Some advantages and challenges of Focus Groups

Advantages

- Can be conducted relatively quickly and easily;
- May take less staff time than in-depth,
- In-person interviews;
- Allow flexibility to make changes in process and questions;
- Can explore different perspectives;
- Can be fun

Challenges

- Analysis is time consuming;
- Participants not be representative of population,
- Possibly biasing the data;
- Group may be influenced by moderator or dominant group members

vi. Other methods: (a) warranty cards; (b) distributor audits; (c) pantry audits; (d) consumer panels; (e) using mechanical devices; (f) through projective techniques; (g) depth interviews, and (h) content analysis.

2. Secondary data collection

Secondary data refers to data that is collected by someone other than the primary user. Common sources of secondary data for social science include censuses, information collected by government departments, organizational records and data that was originally collected for other research purposes. This research involves gathering information from sources that have already been compiled in written, statistical or mapped forms. Secondary information may supplement primary information but must only play a small part in the investigation.

Before using secondary data, the researcher must see that they possess following characteristics:

1. Reliability of data: The reliability can be tested by finding out such things about the said data:

- Who collected the data?
- What were the sources of data?
- Were they collected by using proper methods?
- At what time were they collected?
- Was there any bias of the compiler?
- What level of accuracy was desired?
- Was it achieved?

2. Suitability of data: The data that are suitable for one enquiry may not necessarily be found suitable in another enquiry. Hence, if the available data are found to be unsuitable, they should not be used by the researcher. In this context, the researcher must very carefully scrutinise the definition of various terms and units of collection used at the time of collecting the data from the primary source originally.

3. Adequacy of data: If the level of accuracy achieved in data is found inadequate for the purpose of the present enquiry, they will be considered as inadequate and should not be used by the researcher. The data will also be considered inadequate, if they are related to an area which may be either narrower or wider than the area of the present enquiry.

From all this we can say that it is very risky to use the already available data. The already available data should be used by the researcher only when he finds them reliable, suitable and adequate. At

times, there may be wealth of usable information in the already available data which must be used by an intelligent researcher but with due precaution.

Published data are available in various resources including:

- Government publications
- Public records
- Historical and statistical documents
- Business documents
- Technical and trade journals
- Unpublished data includes
- Diaries
- Letters
- Unpublished biographies etc.

Advantages of Secondary Data:

- Ease of access
- Low cost or free
- Time-saving
- Allow you to generate new insights from previous analysis
- Longitudinal analysis
- Anyone can collect the data
- A huge amount of secondary data with a wide variety of sources
- Disadvantages:
- Might be not specific to your needs
- You have no control over data quality

- Biasness (As the secondary data is collected by someone else than you, typically the data is biased in favour of the person who gathered it.)
- Not timely (Secondary data is collected in the past which means it might be out-of-date.)
- You are not the owner of the information

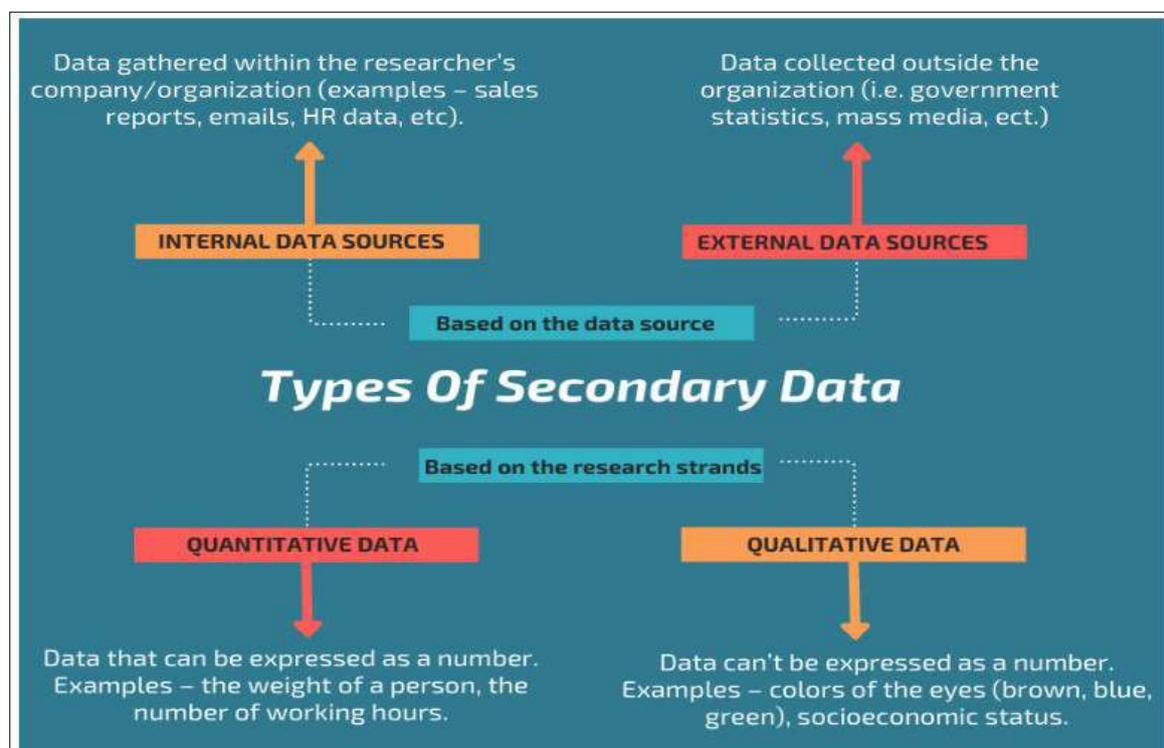


Fig 12: Types of secondary data (Source: intellspot.com)

Collection of Data through Questionnaire

Questionnaire is a commonly used tool for collecting a variety of data. A questionnaire may include a series of questions pertaining to psychological, social, educational, or any such issues which are sent to an individual or a group, with the aim of obtaining relevant data on the topic of research.

- A questionnaire is a popular means of collecting different kinds of data in research.
- A questionnaire is administered personally either individually or to a group of individuals or is mailed to them to save a great deal of time and money in travel.
- Questionnaires are used both to initiate a formal inquiry and also to supplement and check data previously accumulated.

Types of questionnaires

1. Structured questionnaires

Structured questionnaires are those which pose definite and concrete questions. They are prepared well in advance and not on the spot. The form of questions may require responses which are either closed or open.

- Closed-form of questionnaires is used when categorised data are required. They include a set of questions to which a respondent can reply in a limited number of ways—‘yes’, ‘no’, ‘no-opinion’, or an answer from a short list of possible responses.
- The open-ended responses on the other hand are free and spontaneous expressions by the respondent to the questions posed to him/her. The open-ended responses are used mainly for intensive study of a limited number of cases or preliminary exploration of new problems and situations.

2. Unstructured questionnaires

Unstructured questionnaires are frequently referred to as interview guides. They also aim at precision and contain definite issues that are covered while conducting an interview. Flexibility is the chief advantage of the unstructured questionnaire.

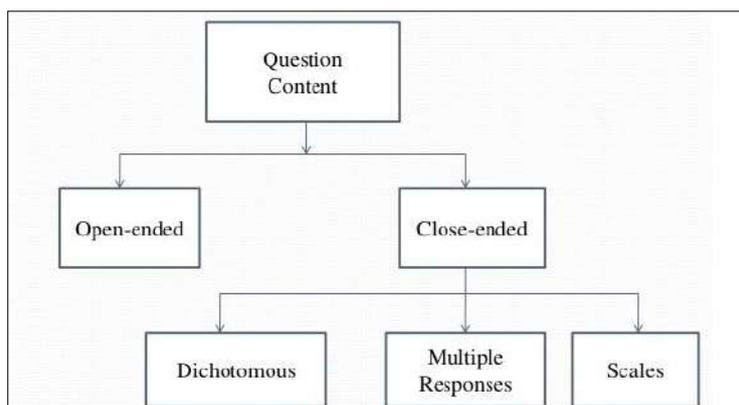


Fig 13: Types of questions

Characteristics of a good questionnaire

1. **Purpose:** A good questionnaire must serve two purposes. First, it must translate the objectives of an investigation into specific questions. Secondly, the questionnaire must motivate the respondents to communicate the required information.
2. **Language:** The language of a good questionnaire should be concise and directed towards producing uniformity of understanding among the respondents.

3. **Frame of reference:** Questions on controversial issues should be broken down into components, so that the researcher can determine the respondent's feeling about various aspects of the problem.
4. **Arrangement of questions:** The questions placed first in the questionnaire should be the easiest to answer. A proper sequence of questions proceeds from the general to specific, from simple to complex one.
5. **Length of the questionnaire:** A questionnaire should not be longer than necessary. The total number of questions must not be too large to tire or bore the respondents.
6. **Respondents must be competent to answer**
7. **Respondents must be willing to answer**
8. **Clearly typed, i.e. not be hard to read**
9. **Adequate space for answers**
10. **Avoid abbreviations**
11. **Proper instructions**
12. **Questions should be relevant**
13. **Biased terms should be avoided**

Limitations of questionnaires

- 1) A questionnaire cannot be used with children and illiterates.
- 2) The return of the mailed questionnaires is often as low as 40 percent to 50 percent. As a result of this poor response, the data obtained are sometimes of limited validity. The respondents who return the questionnaires may not be representative of the entire group. It will make the sample a biased one and thus vitiate the findings.
- 3) Sometimes respondents may not like to respond in writing to questions of intimate and confidential nature or to questions involving controversial issues. It is sometimes difficult to formulate the phrase questions on certain complex and delicate problems.
- 4) There is no check on the respondent who misinterprets a question or gives incomplete or indefinite response.

- 5) Sometimes the respondent may modify his/her earlier/original responses to the questions when he/she finds that his/her responses to latter questions are contradicting the previous ones.

Steps in questionnaire construction

Questionnaires are constructed in a systematic manner. The process goes through a number of interrelated steps. The most commonly steps are (Sarantakos, 1998):

1. **Preparation:** The researcher thinks of various items to be covered in the questionnaire, arrangement of these items in relation to one another, and taking into consideration questions prepared and used in other similar studies.
2. **Constructing the first draft:** The researcher formulates a number of questions including direct/indirect, closed/open-ended and primary/secondary/tertiary questions.
3. **Self evaluation:** The researcher thinks about relevance, symmetry, clarity in language, etc.
4. **External evaluation:** The first draft is given to one or two experts/colleagues for scrutiny and suggestions for changes.
5. **Revision:** After receiving suggestions, some questions are eliminated, some changed and some new questions added.
6. **Pre-test or pilot study:** A pre-test or a pilot study is undertaken to check the suitability of the questionnaire as a whole.
7. **Revision:** The minor and major changes may be made on the basis of experience gained in pre-testing.
8. **Second pre-testing:** The revised questionnaire is then subjected to a second test and amended, if necessary.
9. **Preparing final draft:** After editing, checking spelling, space for response, pre-coding, the final draft is prepared.

Pre-testing of questionnaire

Howsoever carefully the questionnaire might have been prepared; there could be some ambiguities, and some confusing, missing, inappropriate, redundant, inadequate or unanswerable questions. There could also be insufficient space for answering open-ended questions. It is, therefore, necessary to pre-test the questionnaire and remove such questions. For, this the questionnaire may be administered to a

few persons who are similar to those who are to be studied ultimately. Pretesting should not be done on the actual respondents. Many “don’t know” answers may indicate a poorly-worded question which needs to be deleted. The pre-test should be conducted in the same manner as the final study. If it is a mailed questionnaire, the pre-test should also be mailed. If it is interview schedule, the pre-test should be through interview.

After the pre-test, the researcher has first to deal with unanswered questions, followed by those questions which get similar responses from all respondents and may therefore be deleted. He should then take up suggestions, comments and opinions given by the respondents for adding some words or deleting some offending questions. However, the researcher need not accept all suggestions of the respondents.

The cover letter

The cover letter aims at explaining the research topic introducing the objectives of the study, describing how the respondent was selected, giving some important instructions for answering the questions, motivating respondents to participate in the study and assuring anonymity and confidentiality and removing their doubts or mistrust. Here is an example:

We are surveying students and teachers in selected colleges and university departments for a UGC sponsored programme to find out to what extent the teaching is considered satisfactory, and to ascertain the deficiencies in the existing educational system. Your name was selected at random from the list of students/teachers supplied by the colleges/departments. Our questionnaire will take no more than 20 minutes. Kindly answer all questions.

- The cover letter is thus to contain the following main points: Identifying the researcher and the research sponsor.
- Explaining the social significance of the study.
- Pointing out the main objectives of the study.
- Specifying requirements for completion of questionnaires through brief instructions.
- Giving reasons for respondent’s cooperation.
- Assuring anonymity and confidentiality.
- Giving approximate time the questionnaire is likely to take in filling up.

Collection of data through schedules

This method of data collection is very much like the collection of data through questionnaire, with little difference which lies in the fact that schedules (proforma containing a set of questions) are being filled in by the enumerators who are specially appointed for the purpose. These enumerators along with schedules go to respondents, put to them the questions from the proforma in the order the questions are listed and record the replies in the space meant for the same in the proforma. In certain situations, schedules may be handed over to respondents and enumerators may help them in recording their answers to various questions in the said schedules. Enumerators explain the aims and objects of the investigation and also remove the difficulties which any respondent may feel in understanding the implications of a particular question or the definition or concept of difficult terms.

- This method requires the selection of enumerators for filling up schedules or assisting respondents to fill up schedules and as such enumerators should be very carefully selected.
- The enumerators should be trained to perform their job well and the nature and scope of the investigation should be explained to them thoroughly so that they may well understand the implications of different questions put in the schedule.
- Enumerators should be intelligent and must possess the capacity of cross examination in order to find out the truth.
- Above all, they should be honest, sincere, hard working and should have patience and perseverance.

This method of data collection is very useful in extensive enquiries and can lead to fairly reliable results. It is, however, very expensive and is usually adopted in investigations conducted by governmental agencies or by some big organisations. Population census all over the world is conducted through this method.

The questions in the questionnaire or schedule seek three type of information:

1. Demographic information which identifies the interviewee,
2. Substantive information which is focused on the subject under study, and
3. Additional information which may support the substantive information.

Table 4: Differences between Questionnaire and Survey Schedule

Questionnaire	Schedule
The questionnaire is one of the methods used for data collection. The questionnaire will have many questions, with each question having multiple choices.	The schedule is also one of the methods of data collection. It will have a set of statements, questions and space given to note down the answers.
Questionnaire method of data collection is preferred when the respondents are willing to cooperate. In addition, to deploy this method the respondents need to be literate.	The Schedule method of data collection can be utilised irrespective of the respondent's literacy. It can be used when the respondents are literate and can be used even when the respondents are illiterate.
The type of technique used in the Questionnaire method is Quantitative.	The type of technique used in the Schedule method is Qualitative.
In the Questionnaire method, the grouping is made on the basis of different categories like location, age, gender etc.	In the schedule method of data collection, the grouping may exist or may not exist.
Informants receive questionnaires through emails, posts and the answers will be given as per instructions given in the cover letter.	Answers in the Schedule method of data collection are filled by research workers/enumerators.
In the Questionnaire method, there is no scope for direct personal contact with the respondents.	In the Schedule method, there is direct personal contact of the respondents with the enumerators.
The cost incurred in the questionnaire method of data collection is economical in comparison with the schedule. The cost is less even if the sample size used is very large. Predominantly the money is spent on preparing questionnaires only.	The cost incurred in the Schedule method of data collection is very expensive since there is the cost involved in preparing the schedule, cost incurred on enumerators in addition to the training imparted to them.
The coverage of Questionnaire method is large as the questionnaires can even be sent to respondents who are not easily accessible.	The coverage of this method is relatively small as there are constraints in sending enumerators to larger areas.
In the questionnaire, there is a higher possibility of collecting wrong or incomplete information when respondents are unable to have a clear understanding of the given question.	The possibility of receiving inaccurate answers or incomplete answers due to difficulties in understanding the question can be ruled out in this method of data collection as the enumerators will be present and they can resolve any doubts and queries of respondents.
In the Questionnaire method, respondents will get sufficient time to think before answering questions.	The time available for respondents while answering questions is limited in the Schedule method when compared to the Questionnaire method.
In Questionnaires, responses are filled by the respondents.	In Schedule, method responses are filled by the enumerators themselves.
In the Questionnaire method, there is no scope for bias or the answers getting influenced by the interviewers thought process as the answers are filled by the respondents themselves.	In the Schedule method of data collection, there is scope for bias or the answers getting influenced by the enumerator as the answers to the questions are filled by enumerators although the answers are given by the respondents.

The response rate of the Questionnaire method is low compared to the Schedule method.	The response rate in the Schedule method of data collection is high.
In the Questionnaire method, the identity of the respondent is not known.	In the Schedule method, the identity of the respondent is known.
The Questionnaire quality determines the success of the questionnaire method of data collection.	The success of the Schedule method of data collection is dependent on the efficiency, integrity and honesty of the Schedule method of data collection.

Methods of Writing Notes

Note taking is an indispensable part of writing a documented essay or research paper. Note taking allows students/researchers to gather information from lectures, books, or any other situation that they will later have to memorize or use in order to successfully complete their academic program. Major role of note taking are:

- Helps you to organize material
- Easier to identify sources
- Information at your fingertips
- Good study tool
- Helps you to avoid plagiarism

To take effective notes, the researcher must do the following:

1. **Understand the information:** Reading a source, making decisions about what is useful for your essay, and writing the notes on index cards or in a notebook
2. **Select the information:** when you take notes, you must sort out the material you need from other information surrounding it in the text.
3. **Record the information:** you should choose a style of recording that best suits you and that lists all necessary information-- title, author, publisher, etc., which you will need later for your bibliography.

Note Taking Tips

- Keep your topic, controlling purpose, and audience in mind at all times.
- Do not record material unrelated to your topic.

- Make sure that summaries or paraphrases correctly reflect the meaning of the original.
- Be accurate. Direct quotations are picked up word-for-word. Use quotation marks at the beginning and end of quotation.
- Always double-check page references. It's easy to copy these incorrectly.
- Double-check statistics and facts.
- Nonessential parts of a quotation can be cut if the overall meaning of the quotation is not changed. Indicate omissions of nonessential material from a quotation by using ellipsis points, a series of 3 or 4 spaced dots. Use 3 dots (. . .) when cutting material within a single sentence; use 4 dots (. . . .) when cutting a full sentence, a paragraph, or more than a paragraph from a quotation.

Forms of note taking

Summarize: Capture the main ideas of the source succinctly by restating them in your own words.

Paraphrase: Restate the author's ideas in your own words.

Quote: Copy the quotation exactly as it appears in the original source. Put quotation marks around the text and note the name of the person you are quoting.

Tips for Taking Notes By Hand

- Use index cards to keep notes and track sources used in your paper.
- Create Work Cited cards for each source.
 - Include the citation (i.e., author, title, publisher, date, page numbers, etc.) in MLA format. It will be easier to organize the sources alphabetically when creating the Work Cited page.
 - Number the source cards.
- On each note card:
 - Use only one side to record a single idea, fact or quote from one source. It will be easier to rearrange them later when it comes time to organize your paper.
 - Include a heading or key words at the top of the card.
 - Include the Work Cited source card number.
 - Include the page number where you found the information.
- Taking notes:
 - Use abbreviations, acronyms, or incomplete sentences to record information to speed up the notetaking process.
 - Write down only the information that answers your research questions.
 - Use symbols, diagrams, charts or drawings to simplify and visualize ideas.

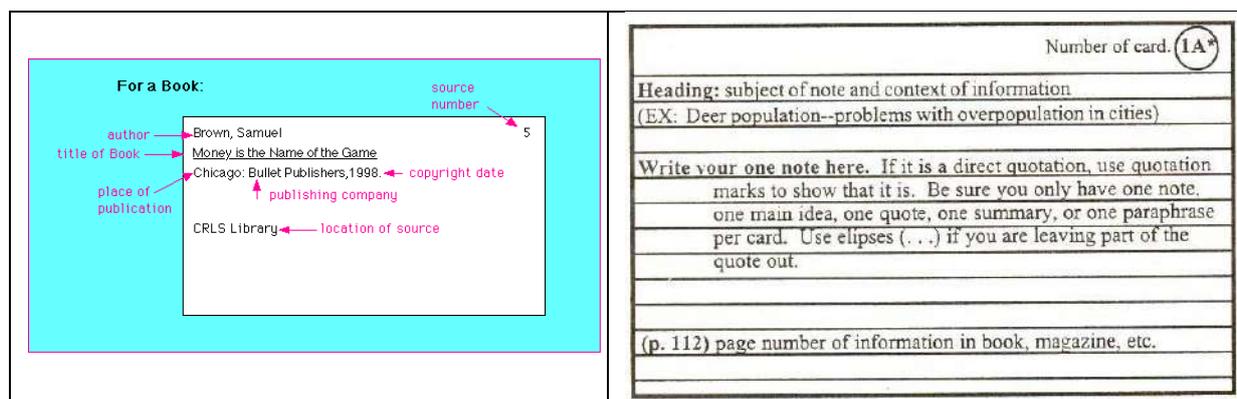


Fig. 14: Example of a Note Card

Style of Referencing

A referencing style is a set of rules on how to acknowledge the thoughts, ideas and works of others in a particular way. Referencing is a crucial part of successful academic writing, avoiding plagiarism¹ and maintaining academic integrity in your assignments and research.

There are (3) major citation styles used in academic writing:

- APA (American Psychological Association) is used by Education, Psychology, and Sciences.
- MLA (Modern Language Association) style is used by the Humanities.
- Chicago/Turabian style is generally used by Business, History, and the Fine Arts.

1. APA referencing style

APA referencing style is an author-date referencing system published by the American Psychological Association. There are two components in the APA referencing style: in-text citations and their corresponding reference list entries.

Example

¹ Plagiarism is the representation of another author's language, thoughts, ideas, or expressions as one's own original work. In educational contexts, there are differing definitions of plagiarism depending on the institution.

Material Type	In-text Citation	Bibliography
A book	(Sapolsky, 2017)	Sapolsky, R. M. (2017). <i>Behave: The biology of humans at our best and worst</i> . Penguin Books.
Chapter in an edited book (If the chapter is from an authored book, use the book citation)	(Dillard, 2020)	Dillard, J. P. (2020). Currents in the study of persuasion. In M. B. Oliver, A. A. Raney, & J. Bryant (Eds.), <i>Media effects: Advances in theory and research</i> (4th ed., pp. 115–129). Routledge.
An article in a print journal	(Weinstein, 2009)	Weinstein, J. (2009). "The market in Plato's Republic." <i>Classical Philology</i> , 104(4), 459–456.
An article in an electronic journal	(Grady et al., 2019)	Grady, J. S., Her, M., Moreno, G., Perez, C., & Yelinek, J. (2019). Emotions in storybooks: A comparison of storybooks that represent ethnic and racial groups in the United States. <i>Psychology of Popular Media Culture</i> , 8(3), 207–217. https://doi.org/10.1037/ppm0000185
A website	(Bologna, 2019)	Bologna, C. (2019, October 31). <i>Why some people with anxiety love watching horror movies</i> . HuffPost. https://www.huffpost.com/entry/anxiety-love-watching-horror-movies_15d277587e4b02a5a5d57b59e

2. MLA referencing style

With MLA style, you must include a Works Cited page at the end of your paper. A Works Cited page is an alphabetical listing of the resources cited in your paper.

Example

Material Type	Works Cited
Book in print	Card, Claudia. <i>The Atrocity Paradigm: A Theory of Evil</i> . Oxford UP, 2005.
eBook	Gaither, Milton. <i>Homeschool: An American History</i> . Palgrave Macmillan, 2017. <i>SpringerLink</i> . doi-org.pitt.idm.oclc.org/10.1057/978-1-349-95056-0.
An article in a print journal	Doggart, Julia. "Minding the Gap: Realizing Our Ideal Community Writing Assistance Program." <i>The Community Literacy Journal</i> , vol. 2, no. 1, 2007, pp. 71-80.
An article in an electronic journal	Sherrard-Johnson, Cherene. "'A Plea for Color': Nella Larsen's Iconography of the Mulatta." <i>American Literature</i> , vol. 76, no. 4, 2004, pp. 833-869, doi: 10.1215/00029831-76-4-833.
A encyclopedia entry	"Patanjali." <i>Benét's Reader's Encyclopedia</i> , edited by Bruce Murphy, 4th ed., HarperCollins Publishers, 1996, p. 782.
A government publication	United States, Federal Maritime Commission. <i>Hawaiian Trade Study: An Economic Analysis</i> . Government Printing Office, 1978.

3. Chicago referencing style

Chicago is a documentation style that has been published by the Chicago University Press since 1906. This citation style incorporates rules of grammar and punctuation common in American English. Typically, Chicago style presents two basic documentation systems: (1) notes and bibliography and (2) author-date.

Material Type	Notes/Bibliography Style
A book in print	<p><u>Note Style:</u> 1. Michael Pollan, <i>The Omnivore's Dilemma: A Natural History of Four Meals</i> (New York: Penguin, 2006), 99–100.</p> <p><u>Duplicate Note:</u> 2. Pollan, <i>Omnivore's Dilemma</i>, 3.</p> <p><u>Bibliography:</u> Pollan, Michael. <i>The Omnivore's Dilemma: A Natural History of Four Meals</i>. New York: Penguin, 2006.</p>
An article in a print journal	<p><u>Note Style:</u> 1. Joshua I. Weinstein. "The Market in Plato's Republic," <i>Classical Philology</i> 104 (2009): 440.</p> <p><u>Duplicate Note:</u> 2. Weinstein, "Plato's Republic," 452–53.</p> <p><u>Bibliography:</u> Weinstein, Joshua I. "The Market in Plato's Republic." <i>Classical Philology</i> 104 (2009): 439–58.</p>

4. Harvard (Author-Date) referencing style

The Harvard referencing style is another popular style using the author-date system for in-text citations. In-text citation: It consists mainly of the authors' last name and the year of publication (and page numbers if it is directly quoted) in round brackets placed within the text.

Example

Banerjee, A. and Watson, T.F. (2011) Pickard's manual of operative dentistry. 9th edn. Oxford: Oxford University Press.

Davidson, A. (2013) 'The Saudi Marathon Man', *The New Yorker*, 16 April. Available at: <http://www.newyorker.com/news/daily-comment/the-saudi-marathon-man> (Accessed: 22 June 2015).

Guy, J. (2001) *The view across the river: Harriette Colenso and the Zulu struggle against imperialism*. Charlottesville, Virginia: University Press of Virginia.

Bibliography

A reference list is the detailed list of references that are cited in your work. A bibliography is a detailed list of references cited in your work, plus the background readings or other material that you may have read, but not actually cited. Different courses may require just a reference list, just a bibliography, or even both.

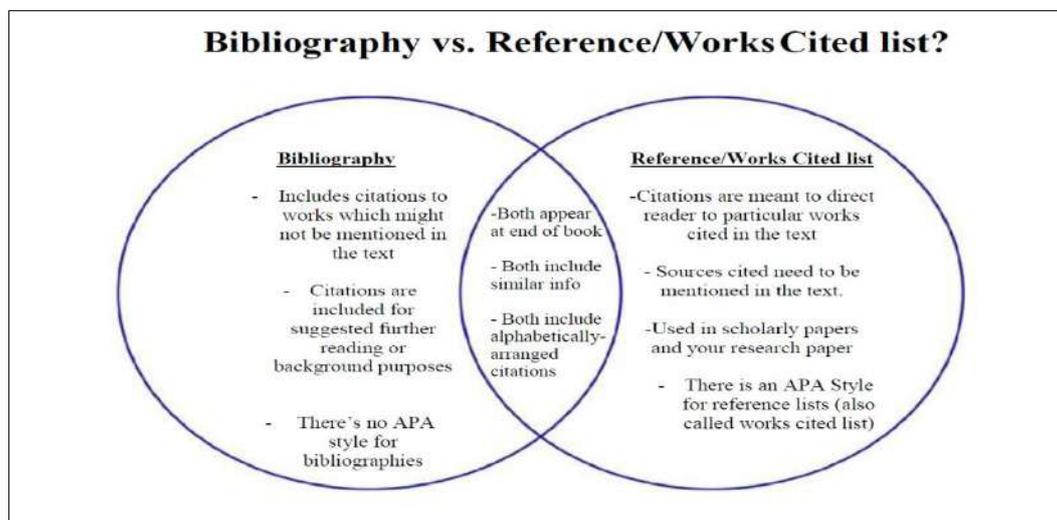
Example of bibliography

Centres for Disease Control and Prevention (1994). *Epi Info (Version 6) [Computer program]*. Available at <http://www.cdc.com/download.html> (Accessed: 23 June 2015).

Gregory, S. (1970). *English military intervention in the Dutch revolt*. B.A. Thesis. University of Birmingham. Available at: <http://findit.bham.ac.uk/> (Accessed: 18 June 2015).

Jones, B., (1997) Methods in tumour research. National Agency for Tumour Research, volume. 7.

Peart, N. (1976) Something for Nothing. Toronto: Toronto Sound Studios.



Appendices

Appendices are used when the incorporation of material in the body of the work would make it poorly structured or too long and detailed.

1. Appendices may be used for helpful, supporting or essential material that would otherwise clutter, break up or be distracting to the text.
2. Other people's work in the appendix will be referred to (e.g. see Appendix A), not quoted (e.g. using short or long quotes) from the appendix.
3. Appendices must be referred to in the body of the text; for example, 'details of the questionnaire are given in Appendix B.

Examples of appendix items Appendices may include some of the following:

- supporting evidence (e.g. raw data that is referred to in the text)
- contributory facts, specialised data (raw data appear in the appendix, but summarised data appear in the body of the text)
- sample calculations (referred to in the text)
- technical figures, graphs, tables, statistics (referred to in the text)

- detailed description of research instruments (referred to in the text)
- maps, charts, photographs, drawings (referred to in the text)
- letters, copies of emails (referred to in the text)
- questionnaires/surveys (questionnaire/survey results appear in the body of the text)
- transcripts of interviews (summarised in the text)
- specification or data sheets (summarised in the text)

Abstract

An abstract is a self-contained, short, and powerful statement that describes a larger work. Components vary according to discipline. An abstract of a social science or scientific work may contain the scope, purpose, results, and contents of the work. An abstract of a humanities work may contain the thesis, background, and conclusion of the larger work. An abstract is not a review, nor does it evaluate the work being abstracted. While it contains key words found in the larger work, the abstract is an original document rather than an excerpted passage.

Types of abstract

1. Descriptive abstracts

A descriptive abstract indicates the type of information found in the work. It makes no judgments about the work, nor does it provide results or conclusions of the research. It does incorporate key words found in the text and may include the purpose, methods, and scope of the research. Essentially, the descriptive abstract describes the work being abstracted. Some people consider it an outline of the work, rather than a summary. Descriptive abstracts are usually very short—100 words or less.

2. Informative abstracts

The majority of abstracts are informative. While they still do not critique or evaluate a work, they do more than describe it. A good informative abstract acts as a surrogate for the work itself. That is, the writer presents and explains all the main arguments and the important results and evidence in the complete article/paper/book. An informative abstract includes the information that can be found in a descriptive abstract (purpose, methods, and scope) but also includes the results and conclusions of the research and the recommendations of the author. The length varies according to discipline, but an informative abstract is rarely more than 10% of the length of the entire work. In the case of a longer work, it may be much less.

Abstracts are used for:

1. library services
2. scientific publications
3. speaker proposals at conferences
4. grant or scholarship applications in foreign institutions
5. present-day Internet - related occupations such as bloggers, copywriters

An abstract should reflect:

1. Problem statement
2. Actuality
3. Ways of problem solution
4. Used techniques and methods
5. Conclusion (for whom it may be interesting)

General requirements for writing abstracts:

1. Consider the abstract purpose
2. Follow stylistic peculiarities
3. Write your statements briefly and clearly
4. Avoid repetitions, including the title
5. Observe terminological consistency
6. Use impersonal constructions such as: is/are considered..., is/are analyzed..., is/are reported
7. Write an abstract within 50 – 400 words depending on the material and publisher's requirements

Synopsis Writing

Before starting to work on Dissertation/Article, the Researcher/ trainee have to prepared a Synopsis and get it approved. The synopsis is a brief out line (about four A-4 size pages or 1000 words is the maximum limit) of your future work. A synopsis must have the following headings:

1. Title: Should reflect the objectives of the study. It must be written after the whole synopsis has been written so that it is a true representative of the plan (i.e. the synopsis).

2. Introduction: Should contain brief background of the selected topic. It must identify the importance of study, its relevance and applicability of results. It must clearly state the purpose of the study.

3. Objectives: An objective is intent of what the researcher wants to do state in clear measurable terms.

4. Operational definition: Is the definition of the exposure and outcome variables of interest in context to objective in a particular study and their means of measurement/determination. Examples: Landform, Coastal tourism, Livelihood etc.

5. Hypothesis: A hypothesis is a statement showing expected relation between two variables. A hypothesis is needed in the following study designs: i. All interventional studies, ii. Cohort, and iii. Case control.

6. Material and methods:

A. **Study design:** Mention the name of the appropriate study design.

B. **Setting:** Name and place where the research work is to be conducted.

C. **Duration of study:** How long will the study take with dates?

D. **Sample size:** How many patients will be included. If there are groups how many per group?

E. **Sampling technique:** Type of sampling technique employed, i.e. simple random, stratified, systematic etc.

F. **Sample selection:**

Inclusion criteria: On what bases will patients be inducted in the study?

Exclusion criteria: On what bases will patients be excluded from the study?

G. **Data collection procedure:** A detailed account of how the researcher will perform research; how s/he will measure the variable. It includes: Identification of the study variables Methods for collection of data collection tools (proforma/questionnaire)

H. **Data analysis procedure:** Relevant details naming software to be used, which descriptive statistics and which test of significance if and when required, specifying variables where it will be applied.

7. **References:** In various styles (APA, MLA, Harvard).

8. **Data collection instrument:** The researcher must attach, as an annex, the proforma or questionnaire with the help of which he/she intends to collect data. The proforma/ questionnaire must match the objectives and must not contain irrelevant sections like inclusion and exclusion criteria etc.

Some Important Terms used in Research Methodology

Case Study: The collection and presentation of detailed information about a particular participant or small group, frequently including the accounts of subjects themselves.

Citation: The act of acknowledging or documenting a reference source used in preparing an assignment, report or project. It is also described as documentation. A full citation lists accurate information about author, title, publication date and related facts. There are a number of different citation styles.

Cohort: A group of people sharing a common demographic experience who are observed through time. For example, all the people born in the same year constitute a birth cohort.

Cohort Analysis: Group by group analytic treatment of individuals having a statistical factor in common to each group. Group members share a particular characteristic [e.g., born in a given year] or a common experience [e.g., entering a college at a given time].

Construct: A concept. A theoretical creation that cannot be directly observed.

Cross-Sectional Data: Data collected about individuals at only one point in time. This is contrasted with longitudinal data, which is collected from the same individuals at more than one point in time.

Empirical Research: The process of developing systematized knowledge gained from observations that are formulated to support insights and generalizations about the phenomena being researched.

Ethnography: A research methodology associated with anthropology and sociology that systematically

describes the culture of a group of people. The goal of ethnographic research is to understand the natives' / insiders' view of their own world (an emic view of the world).

Exploratory Study: A study that aims to identify relationships between variables when there are no predetermined expectations as to the nature of those relations. Many variables are often taken into account and compared, using a variety of techniques in the search for patterns.

Longitudinal study: A longitudinal study is a research design that involves repeated observations of the same variables over short or long periods of time. It is often a type of observational study, although they can also be structured as longitudinal randomized experiments.

Methodology (or Research Methods): A particular procedure or set of procedures. These may include the methods, techniques and instruments used in a research experiment.

Methods: A systematic approaches to the conduct of an operation or process. It includes steps of procedure, application of techniques, systems of reasoning or analysis, and the modes of inquiry employed by a discipline.

Participatory action research: Action research intended not only to address a local problem but also to empower individuals and to bring about social change.

Pilot study: A small-scale study administered before conducting an actual study _ its purpose is to reveal defects in the research plan.

Research proposal: A detailed description of a proposed study designed to investigate a given problem.

Research Question: A clear statement in the form of a question of the specific issue that a researcher wishes to answer in order to address a research problem. A research problem is an issue that lends itself to systematic investigation through research.

Researcher bias: A situation in which the researcher's hopes or expectations concerning the outcomes of the study actually contribute to producing various outcomes, thereby creating a threat to internal validity.

References

- ĀtaśaBhāratīya, LalitaRāṇā. (2008). Geographical Thought. New Delhi: Concept Pub. Co.
- Black, J. A. and Champion, D. J. (1976). Methods and Issues in Social Research. John Wiley and Sons, New York.
- Burns, R. B. (2000). Introduction to Research Methods. Sage Publications, London, UK.
- Cone, J. D. (1986). Idiographic, nomothetic, and related perspectives in behavioral assessment. In: R. O. Nelson & S. C. Hayes (eds.): Conceptual foundations of behavioural assessment (pp. 111–128). New York: Guilford.
- Del Casino, V. J. (2009). Social Geography: A Critical Introduction. Chichester: Wiley- Blackwell.

Del Casino , V. J. , Hanna , S. , Grimes , A. , and Jones III , J. P. (2000). Methodological frameworks for the geography of organizations. *Geoforum* 31: 523 – 38.

Imenda, S. (2014). Is There a Conceptual Difference between Theoretical and Conceptual Frameworks. *Journal of Social Science*, 38(2): 185-195

Kerlinger, F. N. (1964). *Foundations of Behavioural Research*, Holt Rinehart and Winston, Inc. New York, USA.

O’Leary, Z. (2004). *The Essential Guide to Doing Research*. Vistaar Publications, New Delhi.

Kothari, C.R. (2004). *Research Methodology: Methods and Techniques*. New Age International (P) Limited, Publishers, ISBN (13): 978-81-224-2488-1

Resnick, D. B. (2015). What is Ethics in Research and Why is it Important?

Thakur, S.K. (2012). Methodological developments in social geography. DOI: 10.1017/UPO9788175969360.007

The Office of Research Integrity, U.S. Department of Health & Human Services

Theodorson, G. A. and Theodorson, A. G. (1969). *A Modern Dictionary of Sociology*. Thomas Y Crowell Co., New York.

Paper – GEO 396: Special Paper Based Field Work

396A: Advanced Geomorphology

GEO 396A.1: FIELD WORK

Field works in Advanced Geomorphology (Generation of report (within about 100 A4 size pages including 30-40 maps/diagrams/field photographs) on the basis of field works carried out under Unit-47.

396B: Coastal Management

GEO 396B.1: FIELD WORK

Field works in Deltaic/ Non-deltaic or Rocky Coastal Environments

1. Studies on the shore profile forms and channel cross sections using echo-sounder and current meter. Littoral environment observation with special reference hydrological parameters: Waves, tides, winds and currents.
2. Coastal sediment budget analysis.
3. Analysis of population pressure on the coastal zones.

4. Analysis of coastal ecosystem diversity using remote sensing.
5. Hazard mapping and community vulnerability mapping Mapping CRZ violation areas using GPS handset, Nearshore geomorphological mapping with conventional surveying equipments or with Total Station. Mapping of the forms of coastline changes

396C: Urban Geography and Regional Planning

GEO 396C.1: FIELD WORKS

Field works on focused urban issue(s) in an urban area e.g. Small town/ Big city/ Few wards of a Big City.

1. Ward-wise land use/land cover survey using high resolution Remote Sensing data. Wardwise distribution of population and its change over time.
2. Survey of spatial distribution of urban facilities (Bank, School, College, Offices, Hospitals etc.), urban transport: Nodes, Network and flow: Collection of database
4. (Primary and secondary), Water supply system: waste collection and disposal system.
5. Study of Urban sprawling, renewal, economics, environmental problems etc.: Collection of database (Primary and Secondary).
6. Income/Social grouping, Ethnic groups Education; Occupation; Age-sex Composition etc.
7. Enquiries on future prospects and planning strategies

396D: Remote Sensing and Geographic Information System

GEO 396D.1: Field Works

Field works for ground truth verification in a selected field area

1. Remote Sensing in Earth Sciences: Geomorphic Mapping; Visual interpretation of landforms; Basic Concepts, Recognition elements, interpretation of drainage pattern, erosion and deposition landforms.
2. Remote Sensing in Agricultural Applications: Soil Mapping, Crop Mapping/ Crop stress determination. Forest Management-Forest density mapping, Forest type mapping.

3. Remote Sensing in Land and Water Management: Landuse/ Land cover Planning, Land resource management; Water Resources: Surface water-ground water, water deciphering, quality inventory and monitoring, quality assessment; Watershed Management: Morphometric Analysis, Hydro-morphogeologic interpretation techniques for targeting ground water potential zones in alluvial, sedimentary and hard rock areas; flood assessment and watershed management.
 4. Remote Sensing in Urban and Rural Development: Mapping of human habitation and type.
 5. Remote Sensing in Coastal Management: Coastal Landuse, Spatial and temporal changes, SST, Phytoplankton assessment, Sediment assessment.
-

FIELD REPORT

Generation of report (within about 100 A4 size pages including 30-40 maps/diagrams/field photographs) on the basis of field works carried out under Special Papers (Advanced Geomorphology, Coastal Management, Urban Geography and Regional Planning, Remote Sensing and Geographic Information System).

The official format of the field report has been given in the following section.

Title of Your Thesis (Capitalized Each Word except Preposition)

*Thesis Submitted to Midnapore City College
for the Partial Fulfilment of the Degree of
Master of Science (Your Subject)*

Submitted by

Your Name

Guided by

Guide name



Department of Biological Sciences

MIDNAPORE CITY COLLEGE

Kuturiya, P.O. Bhadutala,

Paschim Medinipur, Pin-721129

West Bengal, India

Year

Certificate (Should be printed in College letter head)

This is to certify that the project report entitled <Title of the project> submitted by <name of the candidate> <Roll No.> to the Midnapore City College, Midnapore, West Bengal, India during the year of 2019 in partial fulfilment for the award of the degree of M.Sc. in <Stream> is a bona fide record of project work carried out by him/her under my/our supervision. The contents of this report, in full or in parts, have not been submitted to any other Institution or University for the award of any degree

Guide name and Designation

Principal

MIDNAPORE CITY COLLEGE

Director

MIDNAPORE CITY COLLEGE

Date:

Place: Midnapore City College, Paschim Medinipur

Declaration

I do hereby declare that the present Master thesis entitled '*title of your work*' embodies the original research work carried out by me in the Department of Biological Sciences, Midnapore City College, Paschim Medinipur, West Bengal, India under the supervision of Your Guide Name, Designation, Department, College/ Institution name and address. No part thereof has been submitted for any degree or diploma in any University.

Date:

Place: Midnapore City College, Paschim Medinipur

(Your Name)

Approval Sheet

This project report entitled (Title) by (Author Name) is approved for the degree of _____ (Degree details).

(Signature of Examiners)

(Name :.....)

(Signature of Guide)

(Name :.....)

(Signature of Principal)

(Name: Dr. Sudipta Chakrabarti)

(Dr. Pradip Ghosh)

Director

Midnapore City College

Date: _____

Place: _____



(Optional)

Acknowledgment

I would first like to acknowledge Dr. Pradip Ghosh, Hon'ble Founder Director, Midnapore City College, Paschim Medinipur for providing me the opportunity to study and complete my thesis work in this college. I am gratefully indebted to him for his very valuable comments on this thesis.

I would like to thank my thesis adviser [Dr./Mr.] [Name Surname] of the [Department name] at [College name]. The door to Prof. [Last name] office was always open whenever I ran into a trouble spot or had a question about my research or writing. He/She consistently allowed this paper to be my own work, but steered me in the right the direction whenever he/she thought I needed it.

I would also like to thank the other Faculties (Title, Name and Surname of all of the Other Faculties in the Department) and other non-teaching staffs for their support to carry out this research project. Without their passionate participation and input, the validation survey could not have been successfully conducted.

Finally, I must express my very profound gratitude to my parents (or any other person) for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them. Thank you.

Author

[Name Surname]

Abstract

Brief summary of your thesis. Please be specific at your point.

List of Tables

Table 1.

List of Figures

Fig. 1.

Table of Contents

PAGE No.	CONTENTS
.....
.....
.....
.....
.....

Chapter 1: Introduction

1. Introduction

A. Broad introduction to thesis topic and method. One or two page.

B. Research problem. State broadly. Explain carefully. In one sense, usually the problem is to expand the body of knowledge examined in the literature review.

Chapter 2: Literature Review

2. Literature Review

- A. Brief Overview. Theoretical foundations.
- B. Literature. Group articles by ideas. For a given idea, first discuss common strands in the literature.
- C. Model. Of a process, usually. Based on the literature reviewed.
- D. Point out the gap
- E. Reveal the research question and sometimes Hypotheses (in broad sense of the term; also called Propositions). For each, give brief restatement of justification tied to earlier sections; explain derivation and implications. Include assumptions. Explicitly state plausible rival hypotheses (explanations of process) of a substantive nature.

Chapter 3: Aims and Objective

3. Aims and Objective

From the gap found in the literature review, set a particular aim(s). To fulfill that aim, set the objectives.

Chapter 4: Materials and Methods

4. Materials and Methods

1.1. Materials: All of the chemicals were procured from the company name, country name.

1.2. Sample collection: Briefly describe the sample collection and storage method.

1.3. Other methods: Broadly describe the methods used for the thesis work.

Or

4.1. Design. Experiment, quasi-experiment, survey, and so forth. Detailed description.

4.2. Sample. Universe, population, element, sample design, tolerance, probability.

4.3. Measurement. Operational definitions. Include, as applicable, detailed discussion of indexes/ scales. Specify methods used to assess validity and reliability.

4.4. Analysis. Techniques to be used; justification. Nature of relationships expected (e.g., asymmetrical, symmetrical, reciprocal; linear, monotonic, other curvilinear; necessary, sufficient, necessary and sufficient). Include dummy tables and worked examples of statistics.

4.5. Validity. Design: Internal and external, with relevant subtypes.

4.6. Methodological assumptions. Discuss limitations they impose.

Please use past tense while writing the methods. If you want to use any other tense, please contact with your guide.

Chapter 5: Results

5. Results

5.1. Brief overview.

5.2. Results of application of method; any unusual situations encountered. Nature of sample.

5.3. Descriptive analysis. One-way frequency distributions on central variables.

5.4. Validity/reliability analysis.

5.5. Tests of hypotheses. ANOVAs, cross tabulations, correlations, and such, depending on techniques used; give in same order as hypotheses.

Chapter 6: Discussion

6. Discussion

Prepare a short introduction about the broad problems (2-4 sentences)

While discussing implications, deal with both the theoretical and the practical. Present only interpretations of the findings, not opinion.

You may combine different results and make a solid outcome from those combined results. Discuss the other reported findings with your results and make a conclusive line to justify your findings.

Discussion of results of application of method. Implications.

Chapter 7: Conclusions

7. Conclusions

May include writer's opinion.

A. Summary of entire findings.

B. Conclusions.

C. Implications. Speculate about broadest possible consequences, both theoretical and practical. Label speculation clearly.

D. Limitations. Theory, method.

Chapter 8: Future Scope

8. Future Scope

Brief write up about the future scope of the present research.

Appendices. Bibliographic essay. Questionnaire and coding manual, if any. Raw data.

References. Include all relevant sources examined, whether cited or not.

Reference style Text (Only within the text): All citations in the text should refer to:

1. Single author: the author's name (without initials, unless there is ambiguity) and the year of publication;
2. Two authors: both authors' names and the year of publication;
3. Three or more authors: first author's name followed by 'et al.' and the year of publication. Citations may be made directly (or parenthetically).

Examples: 'as demonstrated (Allan, 2000a, 2000b, 1999; Allan and Jones, 1999). Kramer et al. (2010) have recently shown'

List: References should be arranged first alphabetically.

More than one reference from the same author(s) in the same year must be identified by the letters 'a', 'b', 'c', etc., placed after the year of publication.

Examples: Reference to a journal publication: Van der Geer, J., Hanraads, J.A.J., Lupton, R.A., 2010. The art of writing a scientific article. *Journal of Physics Communications*, 163, 51–59.

Reference to a book: Strunk Jr., W., White, E.B., 2000. *The Elements of Style*, fourth ed. Longman, New York. Reference to a chapter in an edited book: Mettam, G.R., Adams, L.B., 2009.

How to prepare an electronic version of your article, in: Jones, B.S., Smith, R.Z. (Eds.), *Introduction to the Electronic Age*. E-Publishing Inc., New York, pp. 281–304. [dataset] Oguro, M., Imahiro, S., Saito, S., Nakashizuka, T., 2015. Mortality data for Japanese oak wilt disease and surrounding forest compositions. *Mendeley Data*, v1. <http://dx.doi.org/10.17632/xwj98nb39r.1>.

References in the list should be placed first alphabetically.

Mettam, G.R., Adams, L.B., 2009. How to prepare an electronic version of your article, in: Jones, B.S., Smith, R.Z. (Eds.), Introduction to the Electronic Age. E-Publishing Inc., New York, pp. 281-304.

PREPARATION OF THE FINAL PROJECT THESIS

TYPING

The style may be Courier or Times New Roman, or another equivalent word processor style. The font size should be 12. Standard black font is used throughout the text. Color may be used in tables, figures, charts and diagrams. The style must be consistent throughout the thesis.

CORRECTIONS

All corrections (including typographical errors) must be made prior to submission of the final version.

MARGINS

The Printed Sheets shall have the following written area and margins:

Top Margin 15 mm

Head Height 3 mm

Head Separation 12 mm

Bottom Margin 22 mm

Footer 3 mm

Foot Separation 10 mm

Text Height 245 mm

Text Width 160 mm

When header is not used the top margin shall be 30 mm.

Left and Right Margins

The candidates shall have the options of single or double sided printing

Single sided/odd number page (in double sided printing)

Left Margin 30mm

Right Margin 20 mm

Double sided/ even numbered page

Left Margin 20mm

Right Margin 30mm

HEADER

When the header style is chosen, the header can have the Chapter number and Chapter title on the page header.

SPACE AND INDENTATION

The text of the thesis must be one and half spaced. All paragraphs begin indented five spaces from the left margin. The spacing of long quotations, footnotes, tables, bibliographies, captions, etc. should conform to the specifications of this document or conform to the style manual specified for a student's discipline.

PAGINATION

Preliminary pages: The pages preceding the first chapter (title page, approval page, abstract, table of contents, etc.) are counted, however, the first page with the number printed on the page is the ABSTRACT. Preliminary pages are numbered consecutively using lower case Roman numerals (iii, iv, v...) centered ½ inch from the bottom of the page. The title, copyright, approval, dedication and acknowledgment pages are counted, but not numbered. The page numbering begins appearing on the Abstract page. DO NOT use punctuation marks with the Roman numerals.

Text and reference pages: Starting with the first page of the text, all pages are numbered; including the bibliography, appendices, and index. Pages are numbered consecutively in Arabic numbers (1, 2, 3...). The Arabic numbers used for text must be centered and ½ inch from the bottom of the page. Do not use punctuation marks before or after page numbers, do not use the word “page” followed by the number and do not use letter suffixes such as 10a, 10b, etc.

Plagiarisms: No plagiarism of any form will be allowed and thesis will be returned if found any.

Important Notes: Students are directed to check & verify the project report by their concern Guide/HOD before taking printout.

1. STUDENTS HAVE TO SUBMIT FOUR COPIES OF PROJECT REPORT. (One for project guide, one for department, one for college Authority, one for him/her. For example, suppose there are 4 members in a group, then they have to submit one copy for their guide, one copy for their department, one copy for college Authority and four copies for four students).

2. STUDENTS HAVE TO SUBMIT ONE CD. (contains: soft copy of project report, Implemented project & ppt presentation. CD must be leveled with students name & Roll numbers)

3.1 Binding. Final Project Report should be in Hard Book Binding (Rexine Binding) with good quality white executive bond paper of A4 size. The final hard bound copies to be submitted after the viva voce examination will be accepted during the submission of project report.

3.2 Front Covers.

- Full title of project report in 6 mm 22 point's size font properly centered and positioned at the top.
- Full name of the candidate in 4.5 mm 15 point's size font properly centered at the middle of the page.
- A 40 mm dia replica of the Institute emblem followed by the name of department, name of the Institute and the year of submission, each in a separate line and properly centered and located at the bottom of page.

3.2.1 Lettering

All lettering shall be embossed in gold.

3.2.2 Bound back

The degree, the name of the candidate and the year of submission shall also be embossed on the bound (side) in gold.

3.3 Blank Sheets

In addition to the white sheets (binding requirement) two white sheets shall be put at the beginning and the end of the project report.

3.4 Title Sheet

This shall be the first printed page of the project report and shall contain the submission statement: the Project Report submitted in partial fulfilment of the requirements of the Degree M.Sc., the name and Roll No. of the candidate, name of the Supervisor, Department, Institute and year of submission. Sample copy of the 'Title Sheet' is appended in the first page of this proforma.