College of Agricultural Technology
Theni

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

1 CONCEPTS AND FOUNDATIONS OF REMOTE SENSING

1 Introduction

Remote sensing is the science and art of obtaining information about an object, area, or phenomenon through the analysis of data acquired by a device that is not in contact with the object, area, or phenomenon under investigation.

In remote sensing, the sensors are not in direct contact with the objects or events being observed. The information needs a physical carrier to travel from the objects/events to the sensors through an intervening medium. The electromagnetic radiation is normally used as an information carrier in remote sensing. The output of a remote sensing system is usually an image representing the scene being observed. A further step of image analysis and interpretation is required in order to extract useful information from the image. The human visual system is an example of a remote sensing system in this general sense.

The generalized processes and elements involved in electromagnetic remote sensing are schematically illustrated in Figure 1.1. The two basic processes involved are data acquisition and data analysis.

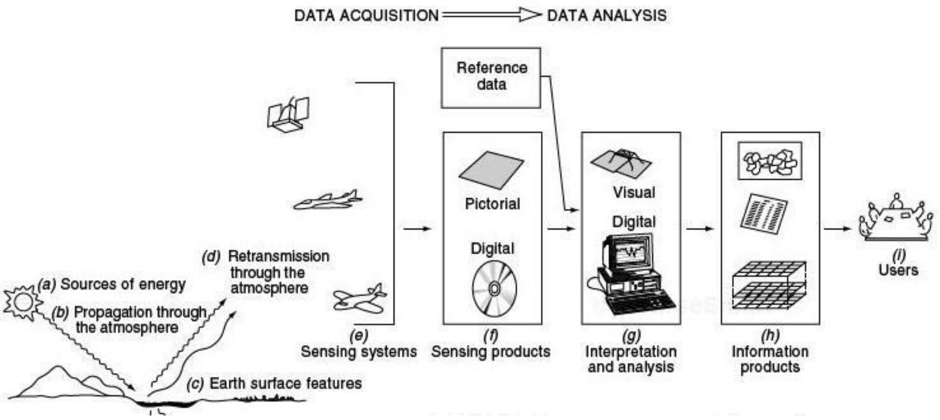


Figure 1.1 Remote Sensing Process

The elements of the data acquisition process are:

1. Energy sources
2. Propagation of energy through the atmosphere
3. Energy interactions with earth surface features
4. Retransmission of energy through the atmosphere
5. Airborne and / or spaceborne sensors
6. Generation of sensor data in pictorial and / or digital form

In short, we use sensors to record variations in the way earth surface features reflect and emit electromagnetic energy.

The data analysis process involves

7. **Examining the data using various viewing and interpretation devices to analyze pictorial data and / or a computer to analyze digital sensor data**

Reference data about the resources being studied (such as soil maps, crop statistics, or field-check data) are used when and where available to assist in the data analysis. With the aid of the reference data, the analyst extracts information about the type, extent, location, and condition of the various resources over which the sensor data were collected.

8. **Compilation of information in the form of hard copy maps and tables or as computer files that can be merged with other "layers" of information is a geographic information system (GIS)**

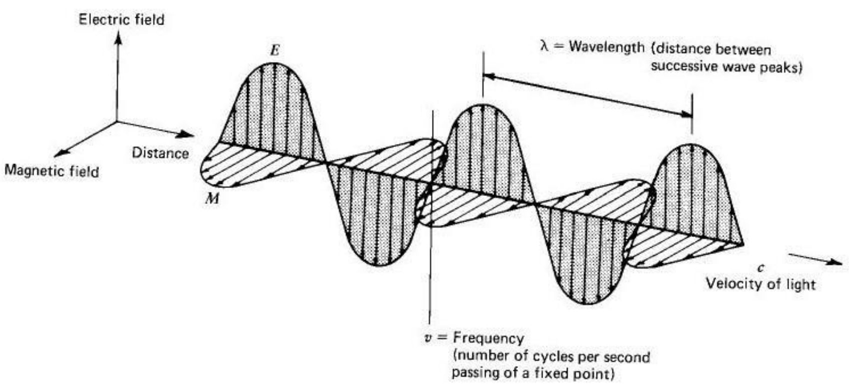
9. **Presented of information to users who apply it to their decision-making process**

2 Energy Sources and Radiation Principles

Visible light is only one of many forms of electromagnetic energy. Radio waves, heat, ultraviolet rays and X-rays are other familiar forms. All this energy is inherently similar and radiates in accordance with basic wave theory. This theory describes electromagnetic energy as traveling in a harmonic, sinusoidal fashion at the "velocity of light," c. The distance from one wave peak to the next is the wavelength λ, and the number of peaks passing a fixed point in space per unit time is the wave frequency v (Figure 1.2).

Waves obey the general equation

$$c = v \lambda \text{ ………………………………Equation 1.1}$$



(Adopted from Lilliesand et.al. 2004)

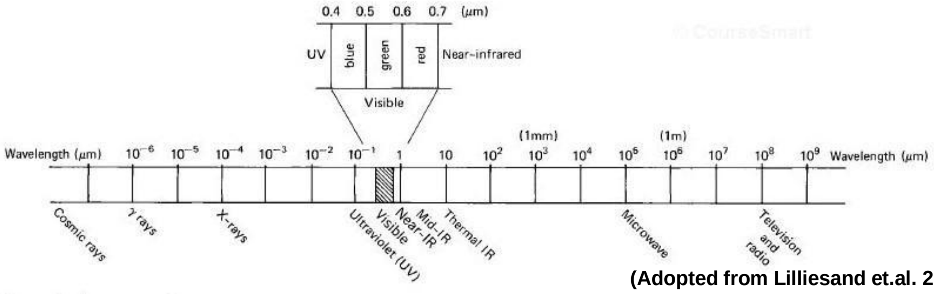
Figure 1.2 Electromagnetic wave

Components of a Electromagnetic wave include a sinusoidal electric wave (E) and a similar magnetic wave (M) at right angles, both being perpendicular to the direction of propagation

Since c is essentially a constant (here, 3 x 10^8 m/sec), frequency v and wavelength λ for any given wave are related inversely, and either term can be used to characterize a wave into a particular form.

In remote sensing, it is most common to categorize electromagnetic waves by their wavelength location within the electromagnetic spectrum (Figure 1.3). The most prevalent unit used to measure wavelength along the spectrum is the micrometer (1μm is 10^{-6}m and 1Angstrom (Å) is 10^{-10}m). Although names (such as "ultraviolet" and "microwave") are generally assigned to regions of the electromagnetic spectrum for convenience, there is no clear-cut dividing line between one nominal spectral region and the next. Divisions of the spectrum have grown out of the various methods for sensing each type of radiation more so than from inherent differences in the energy characteristics of various wavelengths. Also, it should be noted that the portions of the electromagnetic spectrum used in remote sensing lie along a continuum characterized by magnitude changes of many powers of 10. Hence, the use of logarithmic plots to depict the electromagnetic spectrum is quite common. The "visible" portion of such a plot is an extremely small one, since the spectral sensitivity of the human eye extends only from about 0.4 μm to approximately 0.7 μm. The color "blue" is ascribed to the approximate range of

0.4 to 0.5 μm, "green" to 0.5 to 0.6 μm, and "red" to 0.6 to 0.7 μm. Ultraviolet (UV) energy adjoins the blue end of the visible portion of the spectrum. Adjoining the red end of the visible region are three different categories of infrared (IR) waves: near-IR (from 0.7 to 1.3 μm), mid-IR (from 1.3 to 3 μm), and thermal-IR (beyond 3 μm). At much longer wavelengths (1 mm to 1 m) is the microwave portion of the spectrum.



(Adopted from Lilliesand et.al. 2004)

Figure 1.3 Electromagnetic spectrum

Most common sensing systems operate in one or several of the visible, IR, or microwave portions of the spectrum. Within the IR portion of the spectrum, it should be noted that only thermal IR energy is directly related to the sensation of heat; near- and mid-IR energy is not.

Although many characteristics of electromagnetic radiation are most easily described by wave theory, another theory offers useful insights into how electromagnetic energy interacts with matter. This theory - the particle theory - suggests that electromagnetic radiation is composed of many discrete units called photons or quanta. The energy of a quantum is given as

$$Q=hv$$ ………………………….Equation 1.2

Where,

Q = energy of a quantum, Joules (J)
h = Planck's constant, 6.626 x
10^{-34} J sec v = frequency

We can relate the wave and quantum models of electromagnetic radiation behavior by solving Eq. 1.1 for v and substituting into Eq. 1.2 to obtain

$$Q \quad \frac{hc}{} \text{..........................Equation 1.3}$$

Thus, we see that the energy of a quantum is inversely proportional to its wave length. The longer the wavelength involved, the lower its energy content.

The sun is the most obvious source of electromagnetic radiation for remote sensing. However, all matter at temperatures above absolute zero (0K, or -273ºC) continuously emits electromagnetic radiation. Thus, terrestrial objects are also sources of radiation, though it is of considerably different magnitude and spectral composition than that of the sun. How much energy any object (that behaves as a blackbody) radiates is, among other things, a function of the surface temperature of the object. This property is expressed by the Stefan Boltzmann Law, which states that

$$M = \sigma T^4 \text{..........................Equation 1.4}$$

Where,

M = total radiant exitance from the surface of a material, watts (W) m^{-2}
σ = Stefan-Boltzmann constant, 5.6697 x 10^{-8} W m^{-2} K^{-4}
T = absolute temperature (K) of the emitting material

It is important to note that the total energy emitted from an object varies as T^4 and therefore increases very rapidly with increases in temperature. Just as the total energy emitted by an object varies with temperature, the spectral distribution of the emitted energy also varies.

Figure 1.4 shows energy distribution curves for blackbodies at temperatures ranging from 200 to 6000 K. The units on the ordinate scale (W m^{-2} μm^{-1}) express the radiant power coming from a black body per 1µm spectral interval. Hence, the area under these curves equals the total radiant exitance, M, and the curves illustrate graphically what the Stefan-Boltzmann law expresses mathematically: the higher the temperature of the radiator, the greater the total amount of radiation it emits. The curves also show that there is a shift toward shorter wavelengths in the peak of a blackbody radiation distribution as temperature increases.

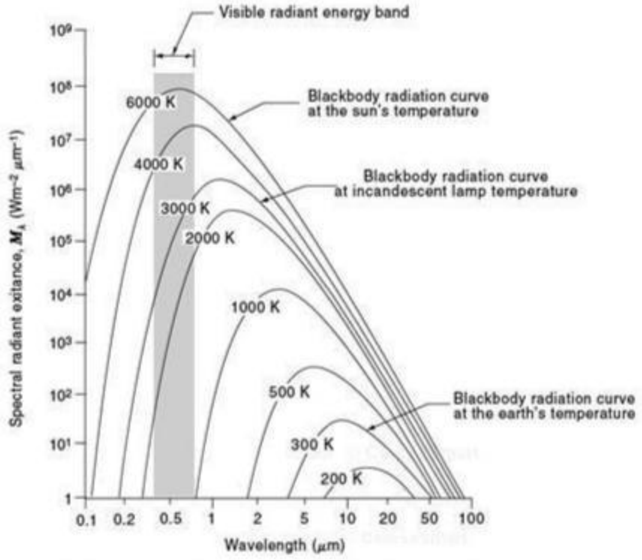


Figure 1.4 Spectral distribution of energy radiated from blackbodies of various temperatures

(Adopted from Lilliesand et.al. 2004)

The dominant wavelength, or wavelength at which a blackbody radiation curve reaches a maximum, is related to its temperature by Wien's displacement law,

m $\frac{A}{T}$ ………………………**Equation 1.5**

Where,

λ_m = wavelength of maximum spectral radiant exitance, µm A = 2898 µm K
T = temperature, K

Thus, for a blackbody, the wavelength at which the maximum spectral radiant exitance occurs varies inversely with the blackbody's absolute temperature. We observe this phenomenon when a metal body such as a piece of iron is heated. As the object becomes progressively hotter, it begins to glow and its color changes successively to shorter wavelengths - from dull red, to orange, to yellow, and eventually to white.

The sun emits in the same manner as a blackbody radiator whose temperature is about 6000 K (Figure 1.4). Many incandescent lamps emit radiation typified by a 3000 K blackbody radiation curve. Consequently, incandescent lamps have a relatively low output of blue energy, and they do not have the same spectral constituency as sunlight. The earth's ambient temperature (that is, the temperature of surface materials such as soil, water, and vegetation) is about 300 K (27º C). From Wein's displacement law, this means the

maximum spectral radiant exitance from earth features occurs at a wavelength of about 9.7 µm. Because this radiation correlates with terrestrial heat, it is termed "thermal infrared" energy. This energy can neither be seen nor photographed, but it can be sensed with such thermal devices as radiometers and scanners. By comparison, the sun has a much higher energy peak that occurs at about 0.5 µm, as indicated in Figure 1.4. Our eyes and photographic film are sensitive to energy of this magnitude and wavelength. Thus, when the sun is present, we can observe earth features by virtue of reflected solar energy. Once again, the longer wavelength energy emitted by ambient earth features can be observed only with a nonphotographic sensing system. The general dividing line between reflected and emitted IR wavelengths is approximately 3 µm . Below this wavelength, reflected energy predominates; above it, emitted energy prevails.

3 Energy Interactions in the Atmosphere

Irrespective of its source, all radiation detected by remote sensors passes through some distance, or path length, of atmosphere. The path length involved can vary widely. For example, space photography results from sunlight that passes through the full thickness of the earth's atmosphere twice on its journey from source to sensor. On the other hand, an airborne thermal sensor detects energy emitted directly from objects on the earth, so a single, relatively short atmospheric path length is involved. The net effect of the atmosphere varies with these differences in path length and also varies with the magnitude of the energy signal being sensed, the atmospheric conditions present, and the wavelengths involved.

The atmosphere can have a profound effect on, among other things, the intensity and spectral composition of radiation available to any sensing system. These effects are caused principally through the mechanisms of atmospheric scattering and absorption.

3.1 Scattering

Atmospheric scattering is unpredictable diffusion of radiation by particles in the atmosphere. *Rayleigh scatter* is common when radiation interacts with atmospheric molecules and other tiny particles that are much smaller in diameter than the wavelength of the interacting radiation. The effect of Rayleigh scatter is inversely proportional to the fourth power of wavelength. Hence, there is a much stronger tendency for short wavelengths to be scattered by this scattering mechanism than long wavelengths.

A "blue" sky is a manifestation of Rayleigh scatter. In the absence of scatter, the sky would appear black. But, as sunlight interacts with the earth's atmosphere, it scatters the shorter (blue) wavelengths more dominantly than the other visible wavelengths. Consequently, we see a blue sky. At sunrise and sunset, however, the sun's rays travel through a longer atmospheric path length than during midday. With the longer path, the scatter (and absorption) of short wavelengths is so complete that we see only the less scattered, longer wavelengths of orange and red.

Rayleigh scatter is one of the primary causes of "haze" in imagery. Visually, haze diminishes the "crispness", or "contrast", of an image. In color photography, it results in a bluish-gray cast to an image, particularly when taken from high altitude. Haze effect can often be eliminated, or at least minimized, in photography by introducing, in front of the camera lens, a filter that does not transmit short wavelengths.

Another type of scatter is *Mie scatter*, which exists when atmospheric particle diameters essentially equal the energy wavelengths being sensed. Water vapor and dust are major causes of Mie scatter. This type of scatter tends to influence longer wavelengths compared to Rayleigh scatter. Although Rayleigh scatter tends to dominate under most atmospheric conditions, Mie scatter is significant in slightly overcast ones.

A more bothersome phenomenon is *Non-selective scatter*, which comes about when the diameters of the particle causing scatter are much larger than the energy wavelengths being sensed. Water droplets, for example, cause such scatter. They commonly have a diameter in the range 5 to 100 µm and scatter all visible and near-to-mid-IR wavelengths about equally. Consequently, this scattering is "nonselective" with respect to wavelengths. In the visible wavelengths, equal quantities of blue, green, and red light are scattered, hence fog and clouds appear white.

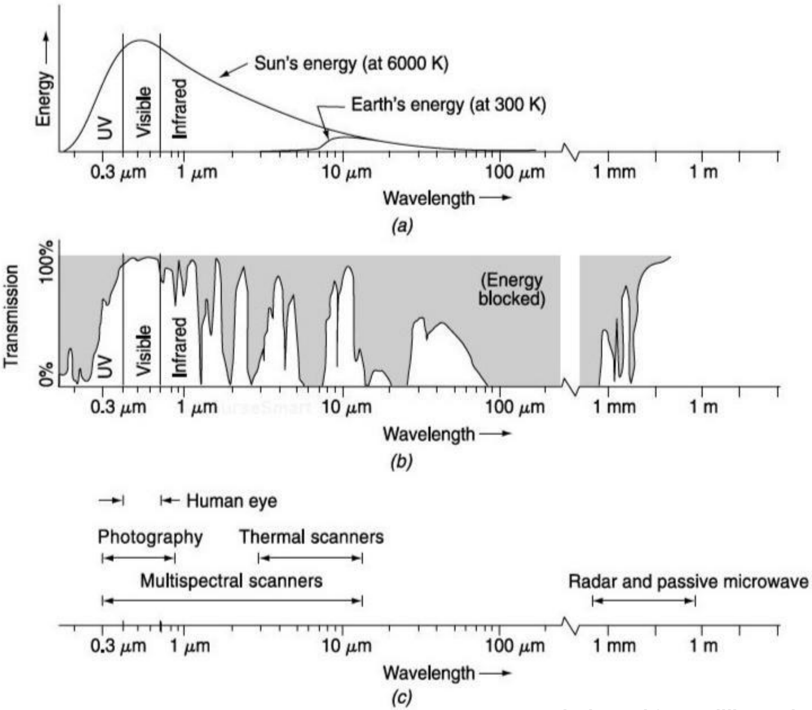
3.2 Absorption

In contrast to scatter, atmospheric absorption results in the effective loss of energy to atmospheric constituents. This normally involves absorption of energy at a given wavelength. The most efficient absorbers of solar radiation in this regard are water vapor, carbon dioxide, and ozone. Because these gases tend to absorb electromagnetic energy in specific wavelength bands, they strongly influence "where we look" spectrally with any given remote sensing system. The wavelength ranges in which the atmosphere is particularly transmissive of energy are referred to as *atmospheric windows*.

Figure 1.5 shows the interrelationship between energy sources and atmospheric absorption characteristics. Figure 1.5a shows the spectral distribution of the energy emitted by the sun and by earth features. These two curves represent the most common sources of energy used in remote sensing. In Figure 1.5b, spectral regions in which the atmosphere blocks energy as shaded. Remote sensing data acquisition is limited to the nonblocked spectral regions, called "atmospheric window". Note in Figure 1.5c that the spectral sensitivity range of the eye (the "visible" range) coincides both with an atmospheric window and the peak level of energy from the sun. Emitted "heat" energy from the earth, shown by the small curve in (a) is sensed through the windows at 3 to 5 µm and 8 to 14 µm using such devices as *thermal scanners. Multispectral scanners* sense simultaneously through multiple, narrow wavelength ranges that can be located at various points in the visible through the thermal spectral region. *Radar and passive microwave* systems operate through a window in the region 1 mm to 1 m.

(a)

(b)



(Adopted from Lilliesand et.al. 2004)

Figure 1.5 Spectral characteristic of (a) energy sources, (b) atmospheric transmittance, and (c) common remote sensing system (note that wavelength scale is logarithmic)

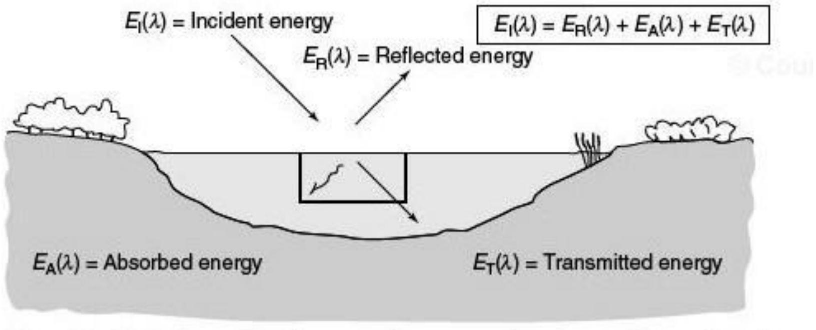
The important point to note from Figure 1.5 is the interaction and the interdependence between the primary sources of electromagnetic energy, the atmospheric windows through which source energy may be transmitted to and from earth surface features, and the spectral sensitivity of the sensors available to detect and record the energy. One cannot select the sensor to be used in any given remote sensing task arbitrarily; one must instead consider (1) the spectral sensitivity of the sensors available, (2) the presence or absence of atmospheric windows in the spectral range(s) in which one wishes to sense, and (3) the source, magnitude, and spectral composition of the energy available in these ranges. Ultimately, however, the choice of spectral range of the sensor must be based on the manner in which the energy interacts with the features under investigation.

4 Energy Interactions with Earth Surface Features

When electromagnetic energy is incident on any given earth surface feature, three fundamental energy interactions with the feature are possible. This is illustrated in Figure 1.6 for an element of the volume of a water body. Various fractions of the energy incident on the element are reflected, absorbed, and/or transmitted. Applying the principle of conservation of energy, we can state the interrelationship between these three energy interactions as

$$E_I(\lambda) = E_R(\lambda) + E_A(\lambda) + E_T(\lambda)$$ **………………………....Equation 1.6**

Where E_I denotes the incident energy, E_R denotes the reflected energy, E_A denotes the absorbed energy, and E_T denotes the transmitted energy, with all energy components being a function of wavelength λ. Equation 1.6 is an energy balance equation expressing the interrelationship among the mechanisms of reflection, absorption, and transmission.

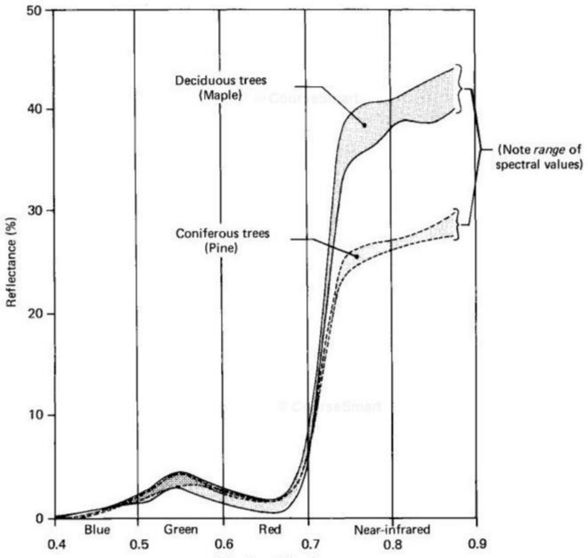


(Adopted from Lilliesand et.al. 2004)

Figure 1.6 Basic interactions between electromagnetic energy and an earth surface feature

Two points concerning this relationship should be noted. First, the proportions of energy reflected, absorbed, and transmitted will vary for different earth features, depending on their material type and condition. These differences permit us to distinguish different features on an image. Second, the wavelength dependency means that, even within a given feature type, the proportion of reflected, absorbed, and transmitted energy will vary at different wavelengths. Thus, two features may be indistinguishable in one spectral range and be very different in another wavelength band. Within the visible portion of the spectrum, these spectral variations result in the visual effect called color. For example, we call objects "blue" when they reflect highly in the blue portion of the spectrum, "green" when they reflect highly in the green spectral region, and so on. Thus, the eye utilizes spectral variations in the magnitude of reflected energy to discriminate between various objects.

The geometric manner in which an object reflects energy is also an important consideration. This factor is primarily a function of the surface roughness of the object. *Specular reflectors* are flat surfaces that manifest mirror like reflections, where the angle of reflection equals the angle of incidence. *Diffuse (or Lambertian) reflectors* are rough surfaces that reflect uniformly in all directions. Most earth surfaces are neither perfectly specular nor diffuse reflector. Their characteristics are somewhat between the two extremes.



(Adopted from Lilliesand et.al. 2004)

Figure 1.7 Generalized spectral reflectance envelopes for deciduous (broad-leaved) and coniferous (needle-bearing trees)

Figure 1.7 illustrates the geometric character of specular, near- specular, near-diffuse, and diffuse reflectors. The category that characterizes any given surface is dictated by the surface's roughness in comparison to the wavelength of the energy incident upon it. For example, in the relatively long wavelength radio range, rocky terrain can appear smooth to incident energy. In comparison, in the visible portion of the spectrum, even a material such as fine sand appears rough. In short, when the wavelength of incident energy is much smaller than the surface height variations or the particle sizes that make up a surface, the surface is diffuse.

Diffuse reflections contain spectral information on the "color" of the reflecting surface, whereas specular reflectance does not. Hence, in remote sensing, we are most often interested in measuring the diffuse reflectance properties of terrain features.

The reflectance characteristics of earth surface features may be quantified by measuring the portion of incident energy that is reflected. This is measured as a function of wavelength and is called *spectral reflectance*, ρ_λ. It is mathematically defined as

$$\frac{E_R(\quad)}{E_I(\ )} \text{ ....Equation 1.7}$$

$$= \frac{\text{energy of wavelength reflected from the object}}{\text{energy of wavelength incident upon the object}} \times 100$$

where ρ_λ is expressed as a percentage.

5 Spectral reflectance curve

A graph of the spectral reflectance of an object as a function of wavelength is termed a *spectral reflectance curve*. The configuration of spectral reflectance curves gives us insight into the spectral characteristics of an object and has a strong influence on the choice of wavelength region(s) in which remote sensing data are acquired for a particular application.

6 Spectral Reflectance of Vegetation, Soil, and Water

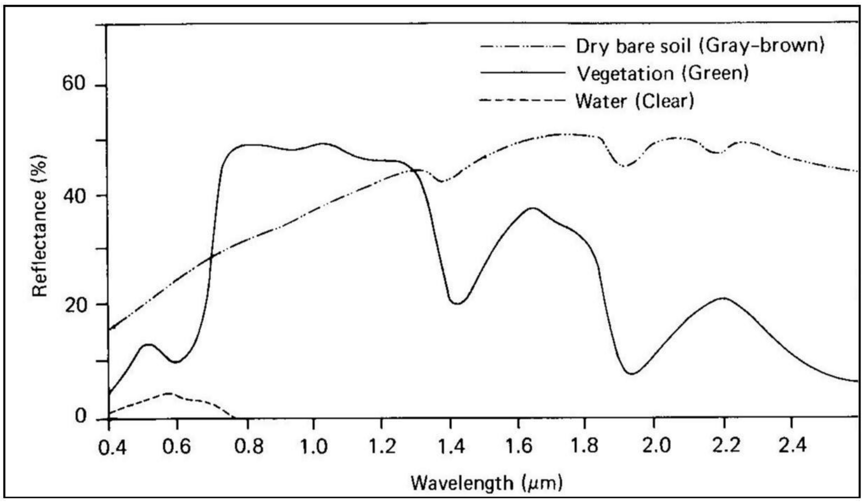


Figure 1.8 Typical spectral reflectance curve of dominant earth features (adopted from Lillesand et al., 2004)

Figure 1.8 shows typical spectral reflectance curves for three basic types of earth features viz., healthy green vegetation, dry bare soil (gray-brown loam), and clear lake water. The lines in this figure represent average reflectance curves compiled by measuring a large sample of features. Note how distinctive the curves are for each feature. Spectral reflectance curves for healthy green vegetation almost always manifest the "peak-and-valley" configuration. The valleys in the visible portion of the spectrum are dictated by the pigments in plant leaves. Chlorophyll, for example, strongly absorbs energy in the wavelength bands centered at about 0.45 and 0.67 μm. Hence, our eyes perceive healthy vegetation as green in color because of the very high absorption of blue and red energy by plant leaves and the very high reflection of green energy. If a plant is subject to some form of stress that interrupts its normal growth and productivity, it may decrease or cease chlorophyll production. The result is less chlorophyll absorption in the blue and red bands. Often the red reflectance increases to the point that we see the plant turn yellow (combination of green and red).

As we go from the visible to the near-infrared portion of the spectrum at about 0.7 μm, the reflectance of healthy vegetation increases dramatically. In the range from about 0.7 to 1.3 μm, a plant leaf typically reflects 40 to 50 percent of the energy incident upon it. Most of the remaining energy is transmitted, since absorption in this spectral region is less

than 5 percent. Plant reflectance in the range 0.7 to 1.3 µm results primarily from the internal structure of plant leaves. Because this structure is highly variable between plants species, reflectance measurements in this range often permits us to discriminate between species, even if they look the same in visible wavelength. Figure 1.9 explain the dominant factor controlling the spectral reflectance.

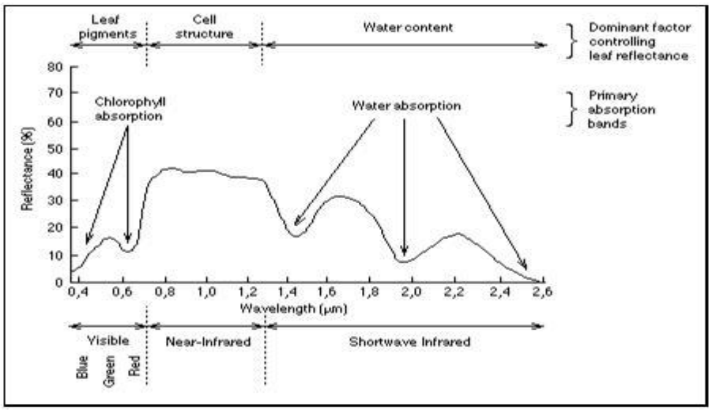


Figure 1.9 Spectral reflectance curve of typical vegetation showing chlorophyll and water absorption bands

Beyond 1.3 µm, energy incident upon vegetation is essentially absorbed or reflected, with little to no transmittance of energy. Dips in reflectance occur at 1.4, 1.9, and 2.7 µm because water in the leaf absorbs strongly at these wavelengths. Accordingly, wavelengths in these spectral regions are referred to as water absorption bands. Reflectance peaks occur at about 1.6 and 2.2 µm, between the absorption bands. Throughout the range beyond 1.3 µm, leaf reflectance is approximately inversely related to the total water present in a leaf. This total is function of both the moisture content and the thickness of a leaf.

The soil curve in Figure 1.10 shows considerably less peak-and-valley variation in reflectance. That is, the factors that influence soil reflectance act over less specific spectral bands. Some of the factors affecting soil reflectance are moisture content, soil texture, surface roughness, presence of iron oxide, and organic matter content. These factors are complex, variable, and interrelated. For example, the presence of moisture in soil will decrease its reflectance. As with vegetation, this effect is greatest in the water absorption bands at about 1.4, 1.9, and 2.7 µm (clay soils also have hydroxyl absorption bands at about 1.4 and 2.2 µm). Soil moisture content is strongly related to the soil texture: coarse, sandy soils are usually well drained, resulting in low moisture content and relatively high reflectance; poorly drained fine textured soils will generally have lower reflectance (Figure 1.10).

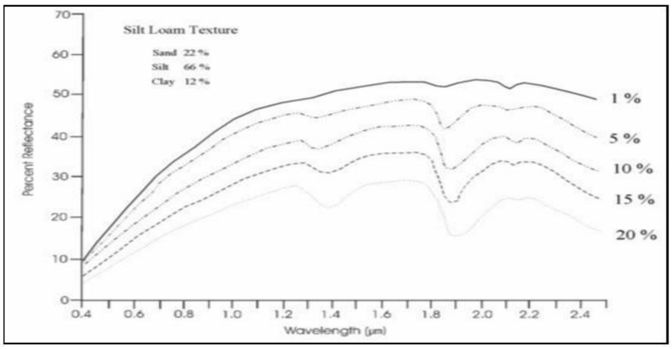


Figure 1.10 Influence of moisture content on spectral reflectance of soil

In the absence of water, however, the soil itself will exhibit the reverse tendency: coarse textured soils will appear darker than fine textured soils. Thus, the reflectance properties of a soil are consistent only within particular ranges of conditions. Two other factors that reduce soil reflectance are surface roughness and content of organic matter. The presence of iron oxide in a soil will also significantly decrease reflectance, at least in the visible wavelengths. In any case, it is essential that the analyst be familiar with the conditions at hand.

Considering the spectral reflectance of water, probably the most distinctive characteristics is the energy absorption at near-infrared wavelengths. In short, water absorbs energy in these wavelengths whether we are talking about water features per se (such as lakes and streams) or water contained in vegetation or soil. Locating and delineating water bodies with remote sensing data are done most easily in near-infrared wavelengths because of this absorption property. However, various conditions of water bodies with remote sensing data are done most easily in near-infrared wavelengths because of this absorption property. However, various conditions of water bodies manifest themselves primarily in visible wavelengths. The energy / matter interactions at these wavelengths are very complex and depend on a number of interrelated factors. For example, the reflectance from a water body can stem from an interaction with the water's surface (specular reflection), with material suspended in the water, or with the bottom of the water body. Even with deep water where bottom effects are negligible, the reflectance properties of a water body are a function of not only the water per se but also the material in the water.

Clear water absorbs relatively little energy having wavelengths less than about 0.6 μm. High transmittance typifies these wavelengths with a maximum in the blue-green portion of the spectrum. However, as the turbidity of water changes (because of the presence of organic or inorganic materials), transmittance - and therefore reflectance - changes dramatically. For example, waters containing large quantities of suspended sediments resulting from soil erosion normally have much higher visible reflectance than other "clear" water in the same geographic area (Figure 1.11).

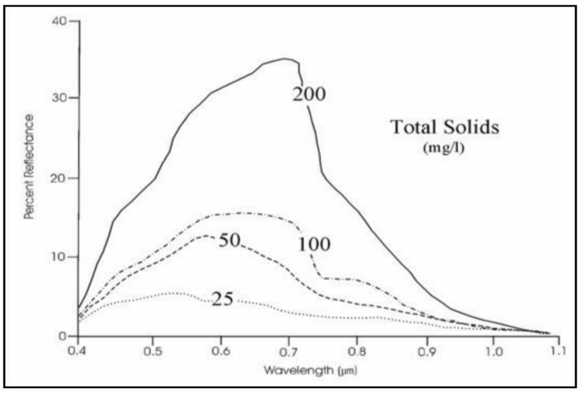


Figure 1.11 Spectral response of water with varying suspended solids

Likewise, the reflectance of water changes with the chlorophyll concentration involved. Increases in chlorophyll concentration tend to decrease water reflectance in blue wavelengths and increase it in green wavelengths. These changes have been used to monitor the presence and estimate the concentration of algae via remote sensing data. Reflectance data have also been used to determine the presence or absence of tannin dyes from bog vegetation in lowland areas and to detect a number of pollutants, such as oil and certain industrial wastes.

Many important water characteristics, such as dissolved oxygen concentration, pH, and salt concentration, cannot be observed directly through changes in water reflectance. However, such parameters sometimes correlate with observed reflectance. In short, there are many complex interrelationships between the spectral reflectance of water and particular characteristics. One must use appropriate reference data to correctly interpret reflectance measurements made over water.

7 Spectral reflectance pattern

The spectral reflectance characteristics of vegetation, soil, and water are normally spectrally separable. However, the degree of separation between types is a function of “where we look spectrally”. For example, water and vegetation might reflect nearly equal in visible wavelengths, yet these features are almost always separable in near-infrared wavelengths.

Because spectral response measured by remote sensors over various features often permit an assessment of type and/or condition of the features, these response have often been referred to as spectral signatures.

Although it is true that many earth surface features manifests very distinctive spectral reflectance and/or emittance characteristics, these characteristics result in spectral response pattern rather than spectral signatures. The reason for this is that the term signature tends to imply a pattern that is absolute and unique. This is not the case with the spectral patterns observed in the natural world. Spectral response pattern measured

by remote sensors may be quantitative but they are not absolute. They may be distinctive but they are not necessarily unique.

8 Advantages of Remote Sensing

Provides data of large areas

Provides data of very remote and inaccessible regions

Able to obtain imagery of any area over a continuous period of time through which the any anthropogenic or natural changes in the landscape can be analyzed

Relatively inexpensive when compared to employing a team of surveyors Easy and rapid collection of data

Rapid production of maps for interpretation

9 Disadvantages of Remote Sensing

The interpretation of imagery requires a certain skill level Needs cross verification with ground (field) survey data Data from multiple sources may create confusion

Objects can be misclassified or confused

Distortions may occur in an image due to the relative motion of sensor and source

10 References

Campbell, J.B. 1996. Introduction to Remote Sensing. Taylor & Francis, London. Colwell, R.N. (Ed.) 1983. Manual of Remote Sensing. Second Edition. Vol I: Theory, Instruments and Techniques. American Society of Photogrammetry and Remote Sensing

John R. Jensen, 2007. Remote Sensing of Environment – An Earth resource perspective. Pearson education published as Prentice Hall. Pp 583

Lillesand, T. M., R.W. Kiefer. and J.W. Chipman 2004. Remote Sensing and Image Interpretation, 5th edition, John Wiley and Sons, ASIA Pte Ltd. Singapore, 763 p.

Klaus Tempfli, Norman Kerle, Gerrit C. Huurneman and Lucas L.F. Janssen. 2009. Principles of remote sensing – An introductory textbook, ITC Educational Textbook Series, ITC

Sabins, Floyd F. 1996. Remote Sensing: Principles and Interpretation. 3rd ed. New York: Freeman.

Thomas M. Lillesand, Ralph W. Keifer, Jonathan W. Chipman. 2004. Remote Sensing and Image Interpretation, Fifth edition, John Wiley and Sons, Inc. pp 1-16.

2 PLATFORMS AND SENSORS: TYPES AND CHARACTERISTICS

1 Sensors

A Sensor is a device that gathers energy (EMR or other), converts it into a signal and presents it in a form suitable for obtaining information about the target under investigation.

2 Classification of Sensors

2.1 Sensors based on Source of Energy

Sensors can be divided into two groups: Passive sensors which depend on an external source of energy, usually the Sun, and sometimes the Earth itself. Current operational passive sensors cover the electromagnetic spectrum in the wavelength range from less than 1 picometer (gamma rays) to larger than 1 meter (microwaves). The oldest and most common type of passive sensor is the photographic camera.

Active sensors have their own source of energy. Measurements by active sensors are more controlled because they do not depend upon varying illumination conditions. Active sensing methods include radar (radio detection and ranging), lidar (light detection and ranging) and sonar (sound navigation ranging), all of which may be used for altimetry as well as imaging.

2.2 Sensors based on Method of Scanning

Across Track Scanning (Whisk broom)

Across-track scanner scans the Earth in a series of lines. The lines are oriented perpendicular to the direction of motion of the sensor platform. Each line is scanned from one side of the sensor to the other, using a rotating mirror. As the platform moves forward over the Earth, successive scans build up a two-dimensional image of the Earth's surface.

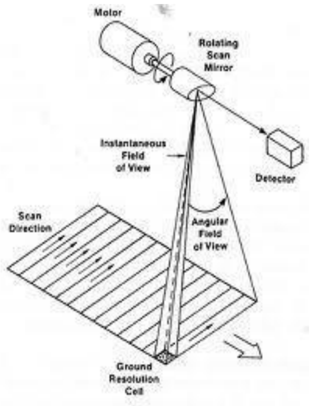


Figure 2.1 Across Track Scanning

Along track scanner (push broom)

These types of scanners use a linear array of detectors located at the focal plane of the image (B) formed by lens systems, which are "pushed" along in the flight track direction (i.e. along track). Along-track scanners also use the forward motion of the platform to record successive scan lines and build up a two-dimensional image. The array of detectors combined with the push broom motion allows each detector to "see" and measure the energy from each ground resolution cell for a longer period of time (dwell time). This allows more energy to be detected and improves the radiometric resolution. Because detectors are usually solid-state microelectronic devices, they are generally smaller, lighter, require less power, and are more reliable and last longer because they have no moving parts

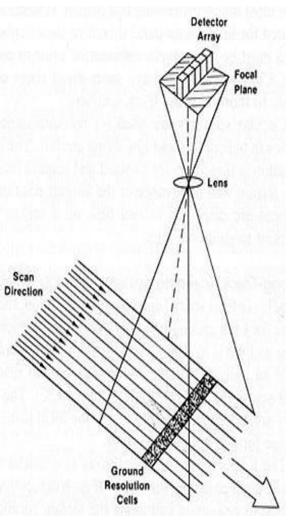


Figure 2.2 Along Track Scanning

3 Resolution

Resolution is defined as the ability of the system to render the information at the smallest discretely separable quantity in terms of distance (spatial), wavelength band of EMR (spectral), time (temporal) and/or radiation quantity (radiometric).

Types of resolution

- **Spatial Resolution**
- **Spectral Resolution**
- **Radiometric Resolution**
- **Temporal Resolution**

3.1 Spatial resolution

Spatial resolution is the capability of sensor to discriminate the smallest object on ground. This is a measure of the area or size of the smallest dimension on the earth's surface over which an independent measurement can be made by the sensor.

The spatial resolution specifies the pixel size of satellite images covering the earth surface.

High spatial resolution: 0.6 - 4 m
Medium spatial resolution: 4 - 30 m
Low spatial resolution: 30 - > 1000 m

3.2 Spectral resolution

Spectral resolution describes the ability of a sensor to define fine wavelength intervals. The finer the spectral resolution, the narrower the wavelengths range for a particular channel or band. Hyperspectral sensors detect hundreds of very narrow spectral bands throughout the visible, NIR and MIR of the electromagnetic spectrum. High spectral resolution helps in fine discrimination between different targets based on their spectral response in each of the narrow bands.

3.3 Radiometric resolution

The ability to distinguish the finer variation of the reflected or emitted radiation from different objects is characterized by the radiometric resolution. Imagery data are represented by positive digital numbers which vary from 0 to (one less than) a selected power of 2. This range corresponds to the number of bits used for coding numbers in binary format. Each bit records an exponent of power 2 (e.g. 1 bit = 21 = 2). The maximum number of brightness levels available depends on the number of bits used in representing the energy recorded. Thus, if a sensor used 8 bits to record the data, there would be 2_8 = 256 digital values available, ranging from 0 to 255. If only 4 bits were used, then only 2_4 = 16 values ranging from 0 to 15 would be available.

3.4 Temporal resolution

Temporal resolution is the capability of the satellite to image the exact same area at the same viewing angle at different points of time. It is the revisit period of the satellite which refers to the length of time taken for a satellite to complete one entire orbit cycle. It is also the frequency of obtaining data over a given area. Actual temporal resolution depends on the sensor capabilities, swath, overlap and latitude.

It is an important aspect in remote sensing when

- Persistent cloud offers limited clear views of the earth's surface
- Short lived phenomenon need to be imaged (flood, oil slicks etc.)
- Multi temporal comparisons are required (agriculture application)
- Changing appearance of a feature over time can be used to distinguish it from near similar features (wheat/maize)

4 Platforms

Platform is a stage to mount the camera or sensor to acquire the information about a target under investigation. Based on its altitude above earth surface, platforms may be classified as (1) Ground borne, (2) Air borne and (3) Space borne

4.1 Ground-based platforms

The ground based remote sensing system for earth resources studies are mainly used for collecting the ground truth or for laboratory simulation studies.

Ground- based sensors are often used to record detailed information about the surface which is compared with information collected from aircraft or satellite sensors. In some cases, this can be used to better characterize the target which is being imaged by these other sensors, making it possible to better understand the information in the imagery.

4.2 Air-borne platforms

Aircraft's are generally used to acquire aerial photographs for photo-interpretation and photogrammetric purposes. Scanners are tested against their utility and performance from these platforms before these are flown onboard satellite missions.

Sensors may be placed on a ladder, scaffolding, tall building, cherry-picker, crane, etc. Aerial platforms are primarily stable wing aircraft, although helicopters are occasionally used. Aircraft are often used to collect very detailed images and facilitate the collection of data over virtually any portion of the Earth's surface at any time.

4.3 Space-borne platforms

Platforms in space are not affected by the earth's atmosphere. These platforms are freely moving in their orbits around the earth, and entire earth or any part of the earth can be covered at specified intervals. The coverage mainly depends on the orbit of the satellite.

It is through these space-borne platforms, we get the enormous amount of remote sensing data and as such the remote sensing has gained international popularity.

For remote sensing purpose the following orbital characteristics are relevant.

Altitude: It is the distance (in Km) from the satellite to the mean surface level of the earth. The satellite altitude influences the spatial resolution to a large extent.

Inclination angle: The angle (in degrees) between the orbit and the equator. The inclination angle of the orbit determines the field of view of the sensor and which latitudes can be observed. If the inclination angle is 60° then the satellite flies over the earth between the latitudes 60° South and 60° North, it cannot observe parts of the earth above 60° latitude.

Period: It is the time (in minutes) required to complete one full orbit. A polar satellite orbiting at an altitude of 800 km has a period of 90 mins.

Repeat Cycle: It is the time (in days) between two successive identical orbits.

Swath: As a satellite revolves around the Earth, the sensor sees a certain portion of the Earth's surface. The area is known as swath. The swath for satellite images is very large between tens and hundreds of kilometers wide.

Depending on their altitude and orbit these platforms may be divided in two categories.

(a)Geostationary and (b) Polar orbiting or Sun-synchronous a. Geostationary satellites

An equatorial west to east satellite orbiting the earth at an altitude of 35000 km., the altitude at which it makes one revolution in 24 hours, synchronous with the earth's rotation. These platforms are covering the same place and give continuous near hemispheric coverage over the same area day and night. These satellites are put in equatorial plane orbiting from west to east Its coverage is limited to 70°N to 70°S latitudes and one satellite can view one-third globe. These are mainly used for communication and meteorological applications viz. GOES METEOSAT, INTELSAT, and INSAT satellites.

b. Sun-synchronous satellites

An earth satellite orbit in which the orbital plane is near polar and the altitude is such that the satellite passes over all places on earth having the same latitude twice in each orbit at the same local sun-time. This ensures similar illumination conditions when acquiring images over a particular area over a series of days.

5 References

Klaus Tempfli, Norman Kerle, Gerrit C. Huurneman and Lucas L.F. Janssen. 2009. Principles of remote sensing – An introductory textbook, ITC Educational Textbook Series, ITC

3 ELEMENTS OF VISUAL PHOTO / IMAGE INTERPRETATION

Introduction

So far we have been focusing primarily on the concepts of remote sensing and on how images of Earth's surface is represented based on the reflection of electromagnetic energy by the surface features. It is understood that, such images bring on minute and comprehensive details which cannot be derived from other sources. Yet the information is not available to us directly. The information we need is encoded in various tones or textures as we see on the image.

To translate these images into valuable information, we must apply specialized knowledge and skill, which forms the basis for image interpretation. What makes interpretation of imagery more difficult than the everyday visual interpretation of our surroundings. The first reason being, we lose our sense of depth when viewing a two-dimensional image. Secondly, viewing objects from directly above the head also provides a new and different perspective that we are not familiar with. So, combining an unfamiliar perspective with a different scale and lack of recognizable detail can make the most familiar object unrecognizable in an image. Finally, we are used to seeing only the visible wavelengths and the imaging outside the visible wavelength is more difficult for use to understand.

An image interpreter has to take clue from lot of aspects in an image and study the image carefully before arriving at a conclusion. Recognizing targets is the key to interpretation and information extraction. Observing the differences between targets and their backgrounds involves comparing different targets based on any, or all of the visual elements. *By tradition, image interpreters use some combination of eight elements of image interpretation viz., tone, size, shape, texture, pattern, shadow, site and association.* In the coming pages, we will discuss the elements of visual image interpretation in detail. The following figure (Figure 3.1) shows the elements of interpretation in decreasing order of complexity

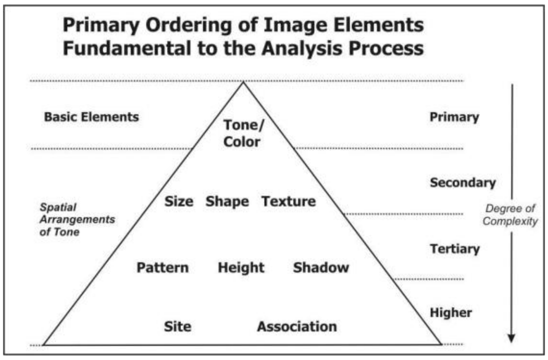


Figure 3.1 Primary ordering of Elements of Image Interpretation

1 Tone

Tone refers to the relative brightness or colour of objects in an image. Tone is the fundamental aspect for discerning different features. For black and white images, tone may be differentiated as "light", "dark grey", "dark" and so on, whereas for colour images, tone refers to terms such as "dark green", "light blue", etc. Wet sand appears in a dark tone in an image, whereas a dry sand appears in a light tone. Variations in tone also allows the elements of shape, texture, and pattern of objects to be distinguished.

2 Size

Size of objects in an image is a function of scale. It is important to assess the size of a target relative to other objects in an image, as well as the absolute size to aid in the interpretation. A quick approximation of size of target feature can aid an interpreter to appropriate result quicker. For example, if an interpreter had to distinguish zones of land use and has identified an area with a number of buildings in it, large buildings such as factories, warehouses would suggest commercial property, whereas small buildings would indicate residential use. Another example is though roads may reflect same tone on the image, the size (or width) of highways and narrow streets differ considerably thereby helping in differentiating them.

3 Shape

Shape indicates the form of an object. For example, individual structures and vehicles have characteristic shapes that, if visible in sufficient detail, provide basis for their identification. Features in nature often have such distinctive shapes that shape alone might be sufficient to provide clear identification. For example, ponds, lakes, rivers occurs in specific shapes unlike others found in nature. For example, the crown of a conifer tree looks like a circle, while that of a deciduous tree has an irregular shape. Airports, harbors, factories and so on, can also be identified by their shape.

4 Texture

Texture is the frequency of tonal variations in an image. If the tone changes abruptly across an image, it is rough textured. If there is very little tonal variation, then the image will be smooth textured. Golf courses, Grasslands appear smooth textured, whereas forest canopy results in a rough textured appearance.

5 Pattern

Pattern is the spatial arrangement of individual objects into distinctive recurring forms that help in their recognition. Pattern on an image usually follows from a functional relationship between the individual features that compose the pattern. The best example is the distinctive spacing of orchard trees leading to a pattern and helps it distinguishable from a forest, although orchard and forest may have same tone.

6 Shadow

Shadow helps the interpreter in understanding the elevation profile and relative height of targets. The following figure (Figure 3.2) depicts an open field in which scattered shrubs

and bushes are separated by areas of open land. At the edges between the trees in the hedgerows and the adjacent open land, trees cast shadows that form a dark strip that enhances the boundary between the two zones as seen on the imagery.

Figure 3.2 Significance of shadow in image interpretation

However, shadows can also reduce or eliminate interpretation in their area of influence, since targets within shadows are much less discernible from their surroundings.

7 Site

Site refers to topographic location. For example, tea estates in higher elevations, sewage treatment facilities are positioned at low topographic sites.

8 Association

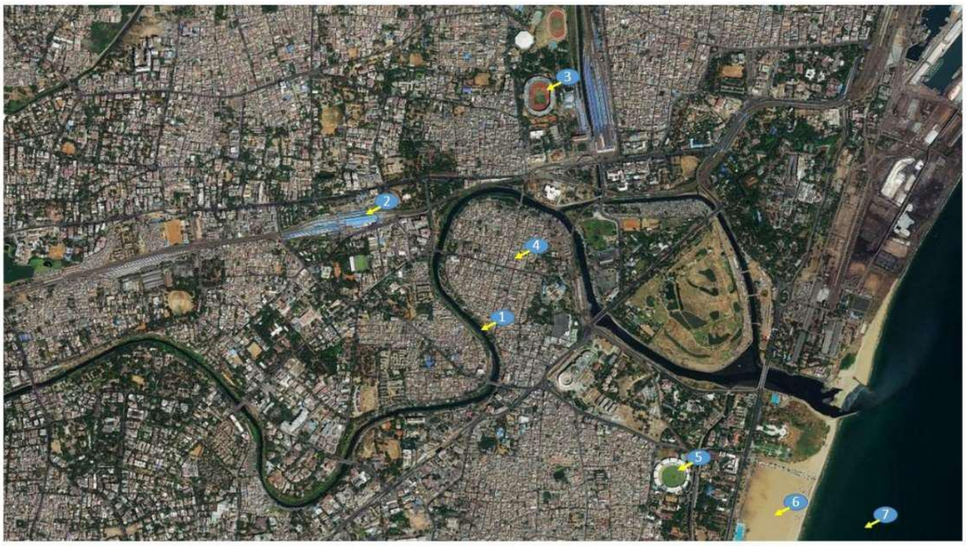
Association takes into account the relationship between other recognizable objects or features in proximity to the target of interest. Association specifies the occurrence of certain objects or features, usually without the strict spatial arrangement implied by pattern. For example schools or colleges are always associated with playground, which can be easily identified on an image (Figure 3.3).

Figure 3.3 Example for Association

9 Ancillary information

In addition to all these elements of image interpretation, an interpreter may or will use non-image information known as ancillary information to assist in the process of interpretation. Ancillary information can consist of information from books, maps, statistical tables, field observations or other disparate sources. All image interpreters use ancillary information in the form of implicit, often intuitive, knowledge that every interpreter brings to an interpretation in the form of everyday science.

Identify various features / objects (1 to 7) on the given image using elements of image interpretation.



4 DIGITAL IMAGE PROCESSING

1 Introduction

The goal of digital image processing is to produce a processed image that is suitable for a given application. For example, we might require an image that is easily inspected by a human observer or an image that can be analyzed and interpreted by a computer. There are two distinct strategies to achieve this goal. First, the image can be displayed appropriately so that the conveyed information is maximized. Hopefully, this will help a human (or computer) extract the desired information. Second, the image can be processed so that the informative part of the data is retained and the rest discarded. This requires a definition of the informative part, and it makes an enhancement technique application specific. Nevertheless, these techniques often utilize a similar framework.

2 Digital data

Digital images are arrays of numbers, i.e. an image is represented logically as a matrix of rows and columns. These image data arrays are included in the general class of 'raster data', which means that the individual data value is not explicitly associated with a particular location on the ground. The location of each data value (or picture element, corrupted into 'pixel') is implied by its position in the array (Figure 4.1). The values of the numbers stored in the array elements lie in a specified range, commonly 0–255, which corresponds to the brightness range of the colour associated with that image array. The value 0 indicates lack of the associated colour (red, green or blue), and the value 255 is the brightest level at which that colour is displayed.

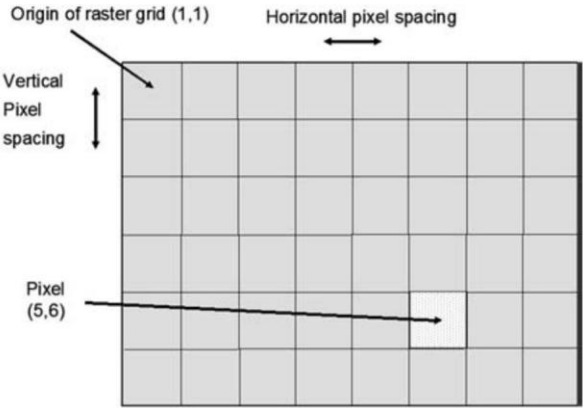


Figure 4.1 View of raster data

A colour image is produced by using three raster arrays, which hold pixel values that represent the levels of the three primary colours of light. Levels 0 to 255 represent the range of each primary colour from 0 (black) to 255 (maximum intensity of red, green, or blue, hereafter RGB). Different combinations of RG and B produce the colours of the spectrum, as demonstrated by Sir Isaac Newton's famous prism experiment. The primary colours of light are 'additive' – for example, red+green = yellow.

3 Radiometric Characteristics of Image Data

Not all remotely-sensed images have pixel values that lie in the range 0–255. For example, AVHRR data use a 0–1023 range. IKONOS pixels lie in the range 0–2047, and the thermal bands of ASTER images are measured on a 0–4095 scale. Specific use is made of the lowest and highest counts ASTER data, for example '0' and '4095'; these are used to indicate 'bad data' and 'saturated pixel', respectively. The values stored in the cells making up a digital image (the 'pixel values' or 'pixel intensities') are represented electronically by a series of binary (base two) digits that can be thought of as 'on/off' switches, or dots and dashes in Morse code. In base two form the decimal numbers 0, 1, 2, 3, are written as 0, 1, 10, 11 . . . with each column to the left representing a successively higher power of two, rather than ten as in the everyday decimal system. If eight binary digits are used to record the value stored in each pixel, then 0 and 255 are written as 00000000 and 11111111. Thus, a total of eight binary digits (bits) are needed to represent the 256 numbers in the range 0–255. The range of pixel intensities is termed the dynamic range of the image.

4 Geometric Characteristics of Image Data

Remote sensing data are data digitized by a process of sampling and quantization of the electro-magnetic energy which is detected by a sensor. IFOV (Instantaneous Field Of View) is defined as the angle which corresponds to the sampling unit as shown in Figure 4.2. Information within an IFOV is represented by a pixel in the image plane. The maximum angle of view which a sensor can effectively detect the electromagnetic energy, is called the FOV (Field Of View). The width on the ground corresponding to the FOV is called the swath width.

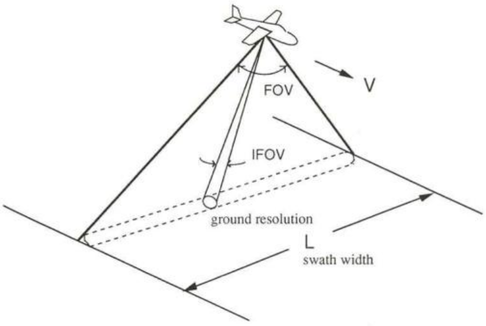


Figure 4.2 FOV and IFOV

The minimum detectable area, or distance on the ground is called the ground resolution. Sometimes the projected area on the ground corresponding to a pixel or IFOV is also called the ground resolution. In remote sensing, the data from a multiple number of channels or bands which divide the electromagnetic radiation range from Ultra Violet to Radio Waves are called multi-channel data, multi-band data or multi- spectral data. In general, multi-channel data are obtained by different detectors as shown in Figure 4.2. Because the detectors are located at slightly different positions, and the light path of different wavelengths is a little different from each other, the images of multi-channel data are not identical in geometric position. To correct such geometric errors between channels

is called registration. The term registration is also used for registration of multi-temporal (or multi-date) images.

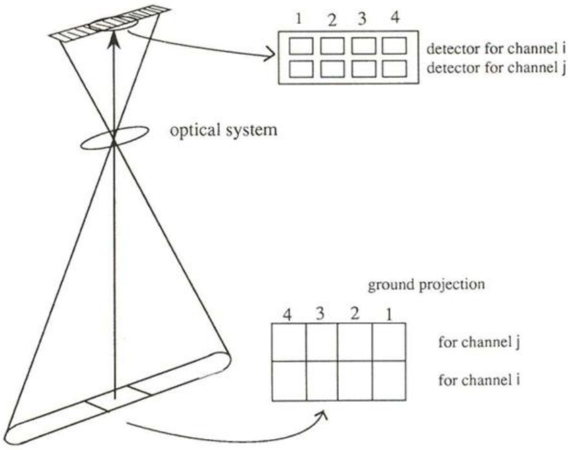


Figure 4.3 Relationship between a detector and its ground projection

5 Digital Image Processing in Remote Sensing

In this cyber era, most remote sensing data are recorded in digital format only and hence virtually all image interpretation and analysis involve digital processing. Digital image processing may involve numerous procedures including formatting and correcting of the data, digital enhancement to facilitate better visual interpretation, or even automated classification of features by the computer. In order to process remote sensing imagery digitally, firstly, the data must be recorded and available in digital form. Secondly, appropriate image analysis system, encompassing both hardware and software should be available.

Most of the common image processing functions available in a image analysis system can be categorized as follows:

1. **Preprocessing**
2. **Image Enhancement**
3. **Image Transformation**
4. **Image Classification and analysis**

5.1 Preprocessing

Preprocessing involves those operations that are normally required prior to the main data analysis and extraction of information. They are generally grouped as radiometric and geometric corrections. Radiometric corrections include correcting the data for sensor irregularities and unwanted sensor or atmospheric noise, and converting the data so that they accurately represent the reflected energy measured by the sensor. Geometric corrections include correcting for geometric distortions due to sensor-Earth geometry

variations and conversion of the data to real world coordinates (e.g., latitude and longitude) on the Earth’s surface.

5.2 Image enhancement

Image enhancement is the process by which the appearance of the imagery is improved so as to assist in a better way for visual interpretation and analysis. Examples of image enhancement functions include Contrast stretching to increase the tonal distinction between various features in an image, Spatial filtering to enhance or suppress specific spatial patterns in an image.

5.3 Image transformation

Image transformations are operations similar to that of image enhancement. But, unlike image enhancement functions which are applied only to a single band of data at a time, image transformation operations usually involve multiple spectral bands. Arithmetic operations (subtraction, addition, multiplication, division) are performed to combine and transform the original bands into “new” images which better display or highlight certain features in an image. Examples are spectral ratioing, principal components analysis etc.

5.4 Image classification and analysis

Image classification and analysis operations are used to digitally identify and classify pixels in the data. Classification is usually performed with multi-band data sets and this process assigns each pixel in an image to a particular class of theme based on statistical characteristics of the pixel brightness values. The two methods of digital image classification are supervised and unsupervised classification

The term classifier refers loosely to a computer program that implements a specific procedure for image classification. Over the years scientists have devised many classification strategies. From these alternatives the analyst must select the classifier that will best accomplish a specific task. At present it is not possible to state that a given classifier is "best" for all situations because characteristics of each image and the circumstances for each study vary so greatly. Therefore, it is essential that analyst understand the alternative strategies for image classification. There are two kinds of image classification, viz., (a) supervised classification and (b) unsupervised classification.

6 Supervised classification

In supervised classification, we identify examples of the Information classes (i.e., land cover type) of interest in the image. These are called "training sites". The image processing software system is then used to develop a statistical characterization of the reflectance 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 bands, 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 reflectance for each pixel and making a decision about which of the signatures it resembles most based on suitable classifier algorithm

7 Unsupervised classification

The goal of unsupervised classification is to automatically segregate pixels of a satellite image into groups of similar spectral character. Classification is done using one of several statistical routines generally called "clustering" where classes of pixels are created based on their shared spectral signatures. Clusters are split and /or merged until further clustering doesn't improve the explanation of the variation in the scene.

Unsupervised Classification is the identification of natural groups, or structures, within multi- spectral data by the algorithms programmed into the software. The following characteristics apply to an unsupervised classification:

There is no extensive prior knowledge of the region that is required for unsupervised classification unlike supervised classification that requires detailed knowledge of the area.

The opportunity for human error is minimized with unsupervised classification because the operator may specify only the number of categories desired and sometimes constraints governing the distinctness and uniformity of groups. Many of the detailed decisions required for supervised classification are not required for unsupervised classification creating less opportunity for the operator to make errors.

Unsupervised classification allows unique classes to be recognized as distinct - units.

8 References

C.H. Chen. 2008. Image processing of remotely sensed images. CRC press, Taylor and Francis group., pp 380

Jay Gao, 2009. Digtial analysis of remotely sensed imagery. The McGraw-Hill Companies, Inc., pp 645.

Paul M. Mather, 2004. Computer Processing of Remotely-Sensed Images - An Introduction, John Wiley & Sons Inc.,

Thomas M. Lillesand, Ralph W. Kiefer, Jonathan W. Chipman. 2004.Remote Sensing and Image Interpretation. Fifth edition, John Wiley and Sons., pp 763

Remote Sensing Notes edited by Japan Association of Remote Sensing © JARS. 1999. Produced by National Space Development Agency of Japan (NASDA) / Remote Sensing Technology Center of Japan (RESTEC). Prepared by Asian Center for Research on Remote Sensing (ACRoRS) in Asian Institute of Technology (AIT).

5 GIS : DEFINITION, COMPONENTS AND FUNCTIONS

1 What is GIS?

Geographic Information System (GIS) is defined as an information system that is used to input, store, retrieve, manipulate, analyze and output geographically referenced data or geospatial data, in order to support decision making for planning and management of land use, natural resources, environment, transportation, urban facilities, and other administrative records.

2 GIS application potential

The strength of GIS depends upon how good is the geospatial database. It can be used for natural resource application (i.e. forestry, agriculture and water resources etc.) in combination with remote sensing and earth observation. In addition it is used for infrastructure development (i.e. highways, railways etc.); utility services like water supply distribution network, telephone network management, gas supply distribution etc.; business application such as real estate, establishment of new retailer shops; heath services; investigation services like crime incidences and their distribution etc. In addition GIS can be used for research and scientific investigations, particularly for water budgeting, atmospheric modelling, climatic studies and global warming.

3 Component of

GIS 3.1 Hardware

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

3.2 Software

GIS software provides the functions and tools needed to store, analyze, and display geographic information. Key software components are:

Tools for the input and manipulation of geographic

information A database management system

Tools that support geographic query, analysis and visualization A

graphical user interface (GUI) for easy access to tools

3.3 Data

Possibly the most important component of a GIS is the data. Geographic data and related tabular data can be collected in-house or purchased from a commercial data provider. A GIS will integrate spatial data with other data resources and can even use a DBMS, used by most organizations to organize and maintain their data, to manage spatial data.

3.4 People

GIS technology is of limited value without the people who manage the system and develop plans for applying it to real-world problems. GIS users range from technical specialists who design and maintain the system to those who use it to help them perform

3.5 Methods

A successful GIS operates according to a well-designed plan and business rules, which are the models and operating practices unique to each organization.

4 GIS Software: Commercial versus Open/free

GIS software is one of the bottlenecks in GIS industry as the major junk money (~50 % or so) is invested towards its procurement and maintenance annually. Because of it many users have apprehensions to change from conventional methods to GIS. In the recent past there is a paradigm shift in usage of GIS software. There are many new and open/free software are launched into the market. The free software where it is freely available and mostly through WWW but the user do not have access to program coding, so not possible to modify or update it. In case of open source, it is free as well as available with full access to program coding so user can modify/update it according to his requirements. Table 18.1 below provides list of some of commercial, open and free GIS software:

Table 5.1 List of GIS software available commercially / as a open source/ freely to the user community

S.No.	Software	Functionality /Remarks
Commercial Software / Proprietary software		
1	ArcGIS	Core modules, Market leader but high cost, many more to be bought for other applications
2	Geomedia	Core modules of GIS, supports education and research institutions
3	MapInfo	Moderate cost
4	Autocad Map	Better input and database creation facility
5	JTMaps (India)	Quite economical and works in vector model
Open Source		
6	GRASS GIS	Satellite Data Analysis & GIS (http://grass.itc.it/)
7	Quantum GIS	Desktop GIS, supports all OS (http://qgis.org/)
8	ILWIS	Satellite Data Analysis & GIS (www.itc.nl)
9	JUMP	Read shp and gml format, display facility and support for wms and wfs, limitations of working with large data files (http://jump-project.org/)

10	PostGIS	With Spatial extensions for the open source. PostgreSQL database, allowing geospatial queries. http://postgis.refractions.net/
11	Mapserver	Web server GIS Software (http://mapserver.gis.umn.edu/)

5 Advantages of GIS

- – Exploring both geographical and thematic components of data in a holistic way
- – Stresses geographical aspects of a research question
- – Large volumes of data
- – Integration of data from widely disparate sources
- – Allows a wide variety of forms of visualisation

6 Disadvantages of GIS

- – Data are expensive
- – Learning curve on GIS software can be long
- – Shows spatial relationships but does not provide absolute solutions
- – Origins in the Earth sciences and computer science. Solutions may not be appropriate for humanities research

7 Need for GIS

Any organization, government or private is in some way or another strongly linked to the geography in which it operates. A GIS that has been designed in a proper manner has the capability of providing quick and easy access to large volumes of data of these geographical features. The user can access & select information by area or by theme to merge one data set with another, to analyze spatial characteristics of data, to search for particular features, to update quickly and cheaply and asses alternatives.

In simpler terms, GIS allows the user to understand geographic information in an easy manner without having to go through large volumes of confusing data that is in tabular form. Visualizing the geography of a particular location is no doubt easier that trying to analyze raw data. The potential and substantial benefits of using GIS makes it a very important tool making the work of any organization easier and more productive. Some of the potential benefits of GIS are:

Opportunity to reduce sets of manual maps held and associated storage costs. Faster and more extensive access to geographic information.

Improved analysis e.g. areas, distances, patterns, etc.

Better communication of information to public officers, members. Improved quality of services.

Better targeting and coordination of services.

Improved productivity in providing public information.

Improved efficiency in updating maps.

The ability to track and monitor growth and development over time. Improved ability to aggregate data for specific sub areas.

Thus GIS's have become indispensable tools for governance, commerce, and environmental and social science.

8 Functions of GIS

The Functions of GIS describe the steps that have to be taken to implement a GIS. These steps have to be followed in order to obtain a systematic and efficient system. The steps involved are:

I) DATA CAPTURE:

Data used in GIS often come from many sources. Data sources are mainly obtained from Manual Digitization and Scanning of aerial photographs, paper maps, and existing digital data sets. Remote-sensing satellite imagery and GPS are promising data input sources for GIS. In this stage Digitization (A conversion process which converts paper maps into numerical digits that can be stored in the computer. Digitizing simplifies map data into sets of points, lines, or cells that can be stored in the GIS computer) is carried out. There are two basic methods of Digitization : Manual Digitizing & Scanning.

II) DATA COMPILATION:

Following the digitization of map features, the user completes the compilation phase by relating all spatial features to their respective attributes, and by cleaning up and correcting errors introduced as a result of the data conversion process. The end results of compilation is a set of digital files, each accurately representing all of the spatial and attribute data of interest contained on the original map manuscripts. These digital files contain geographic coordinates for spatial objects (points, lines, polygons, and cells) that represent mapped features.

III) DATA STORAGE (GIS DATA MODELS):

Once the data have been digitally compiled, digital map files in the GIS are stored on magnetic or other digital media. Data storage is based on a Generic Data Model that is used to convert map data into a digital form. The two most common types of data models are Raster and Vector. Both types are used to simplify the data shown on a map into a more basic form that can be easily and efficiently stored in the computer.

9 References

Wright, D. J. October 28, 1993a. Department of Geography, UC-Santa Barbara. Re: Value of Peer Review [Discussion]. Geographic Information Systems Discussion List [Online].

Petican, D. J. October 29, 1993. University of Waterloo, Canada. Re: GIS as a Science [Discussion]. Geographic Information SystemsDiscussion List [Online].

Feldman, M. November 1, 1993a. Community Planning, University of Rhode Island. Re: GIS as a Science & Value of Peer Review [Discussion]. Geographic Information Systems Discussion List

6 DATA MODELS

1 Data Model

In order to represent the spatial information and their attributes, a data model – a set of logical definitions or rules for characterizing the geographical data is adopted. The data model represents the linkages between the real world domain of geographical data and the computer and GIS representation of these features. As a result, the data model, not only helps in organizing the real-world geographical features into a systematic storage/retrieval mechanism, but also helps in capturing the user's perception of these features. The model:

> **Structures the data to be amenable to computer storage/retrievaland manipulation. The data structure is the core of the model and it is based upon this that features of real world are represented. The ability of the data structure to totally represent the real world determines the success of the model.**
>
> **Abstracts the real world into properties, which is perceived by a specific application. For example, a Landuse map is perceived to be made up of different classes with symbols and legends. The district information is perceived to be made up of district maps and different attribute tables.**
>
> **Helps organize a systematic file structure, which is the internal organization of real world data in a computer.**

2 Entity Definition

An entity is the element in reality. It is a phenomenon of interest in reality that is not further subdivided into phenomena of the same kind. For example, a city can be considered an entity. A similar phenomena stored in a database are identified as entity types or objects. All geographical phenomena can be represented in two dimensions by three main entity types: points, lines, and areas. Figure 6.1 shows how a spatial data model could be constructed using points, lines, and areas. Figure 6.1 also introduces two additional spatial entities: networks and surfaces. These are an extension of the area and line concepts.

A surface entity is used to represent continuous features or phenomena. For these features there is a measurement or value at every location, as in the case of elevation, temperature and population density. This makes representation by a surface entity appropriately. The continuous nature of surface entities distinguishes them from other entity types (points, lines, areas, and networks) which are discrete, that is , either present or absent at a particular location.

A network is a series of interconnecting lines along which there is a flow of data, objects or materials, for example, the road network, along which there is a flow of traffic to and from the areas. Another example is that of a river, along which there is a flow of water. Others not visible on the land surfaces, include the sewerage and telephone systems considered network type of entities.

The dynamic nature of the world poses two problems for the entity-definition phase of a GIS project. The first is how to select the entity type that provides the most appropriate representation for the features being modelled. Is it best to represent a forest as a collection of points (representing the location of individual trees}, or as an area (the boundary of which defines the territory covered by the forest)? The second problem is how to represent changes over time. A forest, originally represented as an area, may decline until it is only a dispersed group of trees that are better represented by using points.

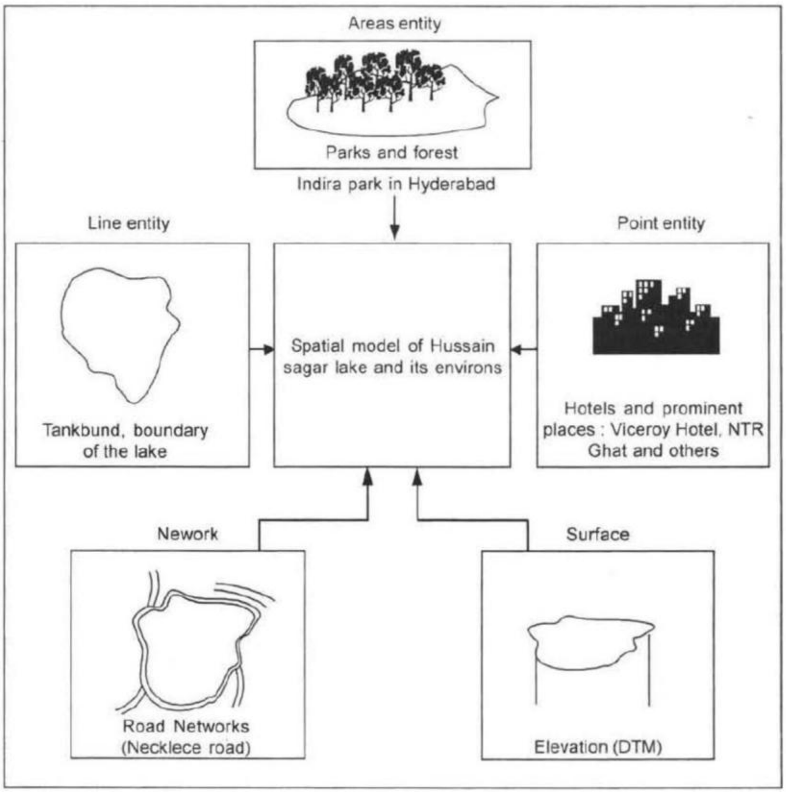


Figure 6.1 Spatial entity data model

The definition of entity types for real- world features is also hampered by the fact that many real-world features simply do not fit into the categories of entities available. An area of natural woodland does not have a clear boundary as there is normally a transition where trees are interspersed with vegetation from a neighbouring habitat type. In this case, if we wish to represent the woodland by an area entity, where do we place the boundary? The question is avoided if the data are captured from a paper map where a boundary is clearly marked, as if someone has already made a decision about the location of the woodland boundary. But is this the true boundary? Vegetation to an ecologist may be a continuous feature (which could be represented by a surface), whereas vegetation to a forest is better represented as series of discrete area entities.

Features with 'fuzzy' boundaries, such as the woodland, can create problems for the GIS designer and the definition of entities, and may have an impact on later analysis. Deciding which entity type should be used to model a real-world feature is not always straightforward. The way in which individuals represent a spatial feature in two dimensions will have a lot to do with how they conceptualise the feature, In turn this will be related to their own experience and how they wish to use the entity they produce. An appreciation of this issue is central to the design and development of all GIS applications.

3 Spatial Data Models

Burrough (1986) recognizes that the human eye is highly efficient at recognizing shapes and forms, but the computer needs to be instructed exactly how spatial patterns should be handled and displayed. Computers require unambiguous instructions on how to turn data about spatial entities into graphical representations. This process is the second stage in designing and implementing a data model. At present there are two main ways in which computers can handle and display spatial entities. These are the raster and vector approaches.

3.1 Raster Spatial Data Model

The raster spatial data model is one of a family of spatial data models described as tessellations. In the raster world individual cells are used as the building blocks for creating images of point, line, area, network and surface entities. Figure 18.2 shows how a range of different features from Happy Valley, represented by the five different entity types, can be modelled using the raster approach. Hotels are modeled by single, discrete cells; the ski lifts are modeled by linking cells into lines; the forest by grouping cells into blocks; and the road network by linking cells into networks. The relief of the area has been modeled by giving every cell in the raster image an altitude value. In Figure 18.2 the altitude values have been grouped and shaded to give the appearance of a contour map.

In the raster world the basic building block is the individual grid cell, and the shape and character ofan entity is created by the grouping of cells. The size of the grid cell is very important as it influences how an entity appears. Figure 18.3 shows how the spatial character of the Happy Valley road network changes as the cell size of the raster is altered.

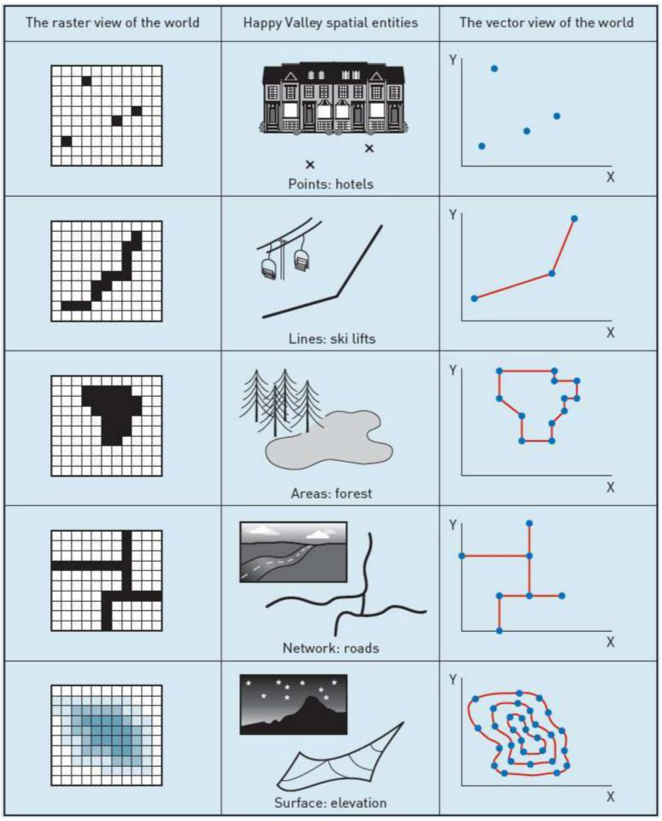


Figure 6.2 Raster and vector spatial data

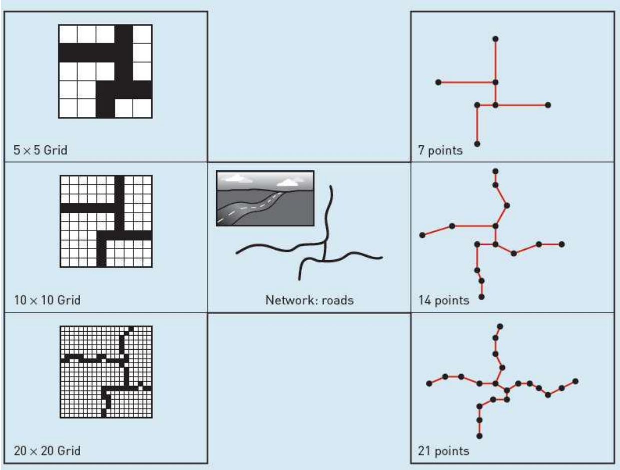


Figure 6.3 Effect of changing resolution in the vector and raster worlds

3.2 Vector Spatial Data Model

A vector spatial data model uses two-dimensional Cartesian (x,y) co- ordinates to store the shape of a spatial entity. In the vector world the point is the basic building block from which all spatial entities are constructed. The simplest spatial entity, the point, is represented by a single (x,y) co-ordinate pair. Line and area entities are constructed by connecting a series of points into chains and polygons. Figure 20.2 shows how the vector model has been used to represent various features for the Happy Valley ski area. The more complex the shape of a line or area feature the greater the number of points required representing it. Selecting the appropriate number of points to construct an entity is one of the major dilemmas when using the vector approach. If too few points are chosen the character, shape and spatial properties of the entity (for example, area, length, perimeter) will be compromised. If too many points are used, unnecessary duplicate information will be stored and this will be costly in terms of data capture and computer storage. Figure 20.3 shows how part of the Happy Valley road network is affected by altering the number of points used in its construction. Methods have been developed to automate the procedure for selecting the optimum number of points to represent a line or area feature

In the vector data model the representation of networks and surfaces is an extension of the approach used for storing line and area features. However, the method is more complex, and closely linked to the way the data are structured for computer encoding.

4 Comparison of Raster and Vector Models

The traditional advantages and disadvantages of raster versus vector spatial data structures have been documented by Kenndey and Meyers (1997). The basic issues include data volume, retrieval efficiency, data accuracy, data display, correctness to perturbation, and data manipulation, efficiency, and processing capabilities. Comparisons of data volume between raster and vector systems are entirely dependent upon the database elements, as well as considerations of accuracy and precision. Detailed comparisons between raster model and vector model are discussed below

Raster model	Vector model
Advantages	
It is a simple data structure. Overlay operations are easily and efficiently implemented. High spatial variability is efficiently represented in a raster format. The raster format is more or less required for efficient manipulation and enhancement of digital images.	It provides a more compact data structure than the raster model. It provides efficient encoding of topology, and, as a result, more efficient implementation of operations that require topological information, such as, network analysis. The vector model is better suited to supporting graphics that closely approximate hand-drawn maps.
Disadvantages	
The raster data structure is less compact. Topological relationships are more difficult to represent.	It is a more complex data structure than a simple raster.

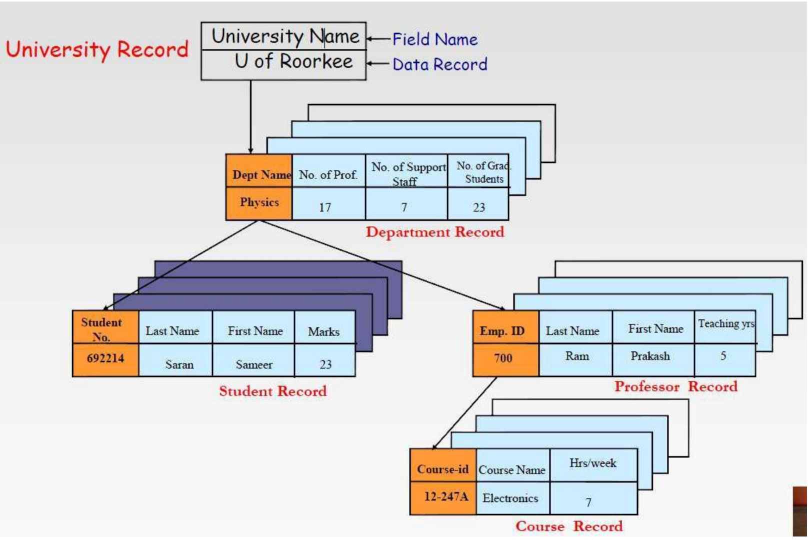
The output of graphics is less aesthetically pleasing because boundaries tend to have a blocky appearance rather than the smooth lines of hand-drawn maps. This can be overcome by using a very large number of cells, but it may result in unacceptably large files.	Overlay operations are more difficult to implement. The representation of high spatial variability is inefficient. Manipulation and enhancement of digital images cannot be effectively done in the vector domain.

5 Non spatial Data

Non spatial information, also known as attribute data, is the descriptive data that defines spatial data*. Data are raw material from which every land information system is built. They are gathered and assembled into records and files. A database is a collection of data that can be shared by different users. It is a group of records and files that are organized, so that there is little or no redundancy. A data base consists of data in many files, in order to be able to access data from one or more files easily, it is necessary to have some kind of structure or organization. The main kinds of data base structure are commonly recognized and termed as: *Hierarchical, Network and Relational.

5.1 Hierarchical database structure

A hierarchical file is a case of a tree structure. The tree is composed of hierarchy of nodes; the upper-most node is called the root. With the exception of this root, every node is related to a node at a higher level called its parent. No element though it can have more than one lower level element called children. A hierarchical file is one with a tree-structure relationship between the records for example a master detail file with two record types. Such a representation is often very convenient because much data tend to be hierarchical in nature or can easily be cast into this structure.

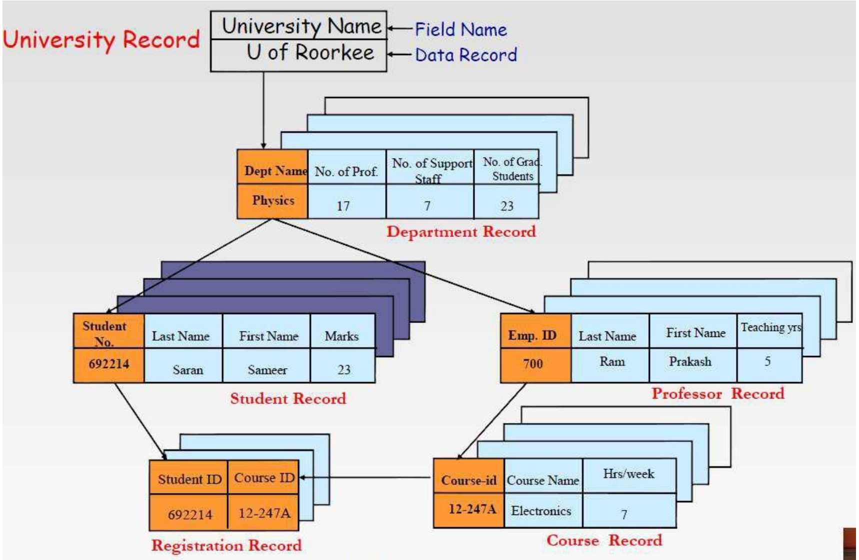


Hierarchical approach is very efficient if all desired access paths follow the parent child linkages. However, it requires a relatively inflexible structure to be placed on

the problem at the outset, when the record type consisting the tree structure is setup. The combination of inflexible structure is setups and the overheads of maintaining or changing pointer system makes extensive modification of the structure of hierarchical systems to meet new requirements, a resource intensive operation. These reasons have contributed to the lack of adoption of this type of DBMS for flexible GIS requirements.

5.2 Network structure

A network structure exists when a child in a data relationship has more than one parent. An item in such a structure can be linked to any other item. The physical data to support complex network structures is far more difficult to develop than for simple structures.



Organization of the Data Records

Each entity set with its attributes is considered to be a node in the network. Relationship sets are represented as linkages in the form of pointers between individual entities in different entity sets. As a result, all the different forms of mapping one-to-many, many-to-many, etc. can be handled directly with large number of pointers.

5.3 The Relational Model

The main data storage concept in the relational model is a table of records, referred to as a relation, or simply a table. The records in a table contain a fixed number of fields, which must all be different from each other, and all records are of identical format. There is, therefore, a simple row and column structure. In relational database terminology the rows, or records, are also referred to as tuples, while the columns of fields are sometimes referred to as domains. Each record of a table stores an entity or a relationship and is uniquely identified by means of a primary key which consists of one field, or a combination of two or more fields in the record. The need for composite keys, consisting of more than one field, arises if no one field can be guaranteed unique. The fields of an entity table store attributes of the entity to which the table corresponds. Table 6.1 illustrates an example for Settlement.

Table 6.1 Example of Relational Database

Settlement	Settlement	Settlement	County
Gittings	Village	243	Downshir
Bogton	Town	31520	Downshir
Puffings	Village	412	Binglia
Pondside	City	112510	Mereshire
Craddock	Town	21940	Binglia
Bonnet	Town	28266	Binglia
Drain	Village	940	Mereshire

In this type, data are organized in two-dimensional tables, such tables are easy for a user to develop and understand. This structure can be described mathematically, a most difficult task for other types of data structure. These structures are called relational structures because each table represents a relation.

7 RASTER DATA ANALYSIS

The raster data model uses a regular grid to cover the space and the value in each grid cell to represent the characteristic of spatial phenomenon at the cell location. The simple data structure of a raster with fixed cell location not only is computationally efficient, but also facilitates a large variety of data analysis operation.

In contrast to vector data analysis, which is based on the geometric objects of point, line and polygon, raster data analysis is based on cells and values (Digital Numbers). Raster data analysis can be performed at the level of individual cells, or group of cells, or cells within an entire raster. Some raster data operations use a single raster; others use two or more rasters. An important consideration in raster data analysis is the type of cell value. Statistics such as mean and standard deviation are designed for numeric values, whereas others such as majority (the most frequent cell value) are designed for both numeric and categorical values.

Various types of data are stored in raster format. Raster data analysis, however, operates only on software-specific raster data such as ESRI grids in ArcGIS. Therefore, to use digital elevation models (DEMs) and other raster data in data analysis we must process them first and convert them to software-specific raster data. The four basic types of raster data analysis are

1. Local functions: that work on every single cell,
2. Focal functions: that process the data of each cell based on the information of a specified neighbourhood,
3. Zonal functions: that provide operations that work on each group of cells of identical values, and
4. Global functions: that work on a cell based on the data of the entire grid.

The principal functionality of these operations is described here.

1 Local Functions

Local functions process a grid on a cell-by-cell basis, that is, each cell is processed based solely on its own values, without referring to the values of other cells. In other words, the output value is a function of the value or values of the cell being processed, regardless of the values of surrounding cells.

For *single layer operations*, a typical example is changing the value of each cell by adding or multiplying a constant. In the following example, the input grid contains values ranging from 0 to 4. Blank cells represent NODATA cells. A simple local function multiplies every cell by a constant of 3 (Fig. 7.1). The results are shown in the output grid at the right. When there is no data for a cell, the corresponding cell of the output grid remains a blank.

Input grid

2	0	1	1
2	3	0	4
4		2	3
1	1		2

Constant

x 3 =

Output grid

6	0	3	3
6	9	0	12
12		6	9
3	3		6

Figure 7.1 A local function multiplies each cell in the input grid by 3 to produce the output grid

Local functions can also be applied to *multiple layers* represented by multiple grids of the same geographic area (Figure 7.2).

Input grid

2	0	1	1
2	3	0	4
4		2	3
1	1		2

x

Multiplier grid

1	1	2	2
1	2	2	2
2	2	3	3
2	3	3	4

=

Output grid

2	0	2	2
2	6	0	8
4		6	9
2	3		8

Figure 7.2 A local function

Local functions are not limited to arithmetic computations. Trigonometric, exponential, and logarithmic and logical expressions are all acceptable for defining local functions.

2 Focal Functions (Neighbourhood functions)

Focal functions process cell data depending on the values of neighbouring cells. The surrounding cells are chosen for their distance and/or directional relationship to the focal cell. Common neighbourhoods include rectangles, circles, annulus or doughnut-shaped and wedges. For instance, computing the sum of a specified neighbourhood and assigning the sum to the corresponding cell of the output grid is the *"focal sum"* function (Figure 7.3). Neighbourhood is defined by a 3 x 3 kernel. For cells closer to the edge where the regular kernel is not available, a reduced kernel is used and the sum is computed accordingly. For instance, the upper left corner cell is adjusted by a 2 x 2 kernel. Thus, the sum of the four values, 2, 0, 2 and 3 yields 7, which becomes the value of this

cell in the output grid. The value of the second row, second column, is the sum of nine elements, 2, 0, 1, 2, 3, 0, 4, 2 and 2, and the sum equals 16.

Input grid

2	0	1	1
2	3	0	4
4	2	2	3
1	1	3	2

Focal Sum =

Output grid

7	8	9	6
13	16	16	11
13	18	20	14
8	13	13	10

Figure 7.3 A focal sum function

Another focal function is the mean of the specified neighbourhood, the *“focal mean”* function. In the following example (Figure 7.4), this function yields the mean of the eight adjacent cells and the center cell itself. This is the smoothing function to obtain the moving average in such a way that the value of each cell is changed into the average of the specified neighbourhood.

Other commonly employed focal functions include standard deviation *(focal standard deviation)*, maximum *(focal maximum)*, minimum *(focal minimum)*, and range *(focal range)*.

Input grid

2	0	1	1
2	3	0	4
4	2	2	3
1	1	3	2

Focal Mean =

Output grid

1.8	1.3	1.5	1.5
2.0	2.0	1.8	1.8
2.2	2.0	2.2	2.3
2.0	2.2	2.2	2.5

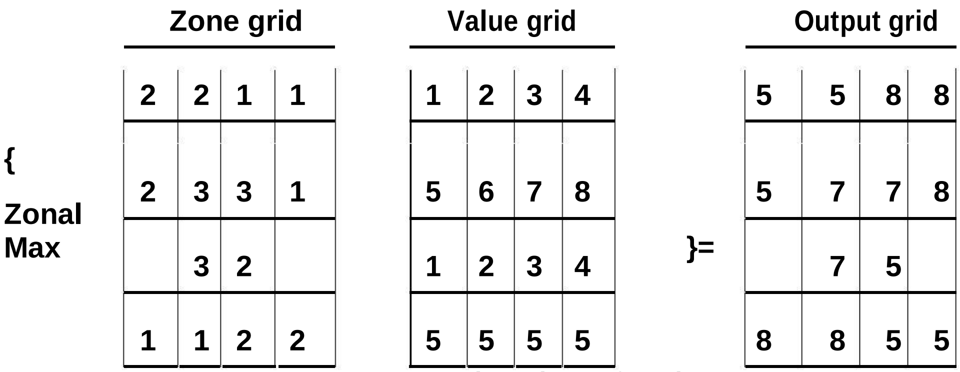
Figure 7.4 A Focal mean function

3 Zonal Functions

Zonal functions process the data of a grid in such a way that cell of the same zone are analyzed as a group. A zone consists of a number of cells that may or may not be contiguous. A typical zonal function requires two grids – a zone grid which defines the size, shape and location of each zone, and a value grid which is to be processed for

analysis. In the zone grid, cells of the same zone are coded with the same value, while zones are assigned different zone values.

Figure 7.5 illustrates an example of the zonal function. The objective of this function is to identify the zonal maximum for each zone. In the input zone grid, there are only three zones with values ranging from 1 to 3. The zone with a value of 1 has five cells, three at the upper right corner and two at the lower left corner. The procedure involves finding the maximum value among these cells from the value grid.



Zone grid

2	2	1	1
2	3	3	1
	3	2	
1	1	2	2

Value grid

1	2	3	4
5	6	7	8
1	2	3	4
5	5	5	5

Output grid

5	5	8	8
5	7	7	8
	7	5	
8	8	5	5

Figure 7.5 A Zonal maximum function

Typical zonal functions include *zonal mean, zonal standard deviation, zonal sum, zonal minimum, zonal maximum, zonal range, and zonal variety*. Other statistical and geometric properties may also be derived from additional zonal functions. For instance, the *zonal perimeter* function calculates the perimeter of each zone and assigns the returned value to each cell of the zone in the output grid.

4 Global Functions

For global functions, the output value of each cell is a function of the entire grid. As an example, the Euclidean distance function computes the distance from each cell to the nearest source cell, where source cells are defined in an input grid. In a square grid, the distance between two orthogonal neighbours is equal to the size of a cell, or the distance between the centroid locations of adjacent cells. Likewise, the distance between two diagonal neighbours is equal to the cell size multiplied by the square root of 2. Distance between non-adjacent cells can be computed according to their row and column addresses.

In Figure 7.6, the grid at the left is the source grid in which two clusters of source cells exist. The source cells labelled 1 are the first cluster, and the cell labelled 2 is a single-cell source. The Euclidean distance from any source cell is always equal to 0. For any other cell, the output value is the distance from its nearest source cell.

In the example given below, the measurement of the distance from any cell must include the entire source grid; therefore this analytical procedure is a global function.

Source grid

		1	1
			1
	2		

Euclidean distance =

Output grid

2.0	1.0	0.0	0.0
1.4	1.0	1.0	0.0
1.0	0.0	1.0	1.0
1.4	1.0	1.4	2.0

Figure 7.6 A Euclidean distance function

Another useful global function is the *cost path function* , which identifies the least cost path from each selected cell to its nearest source cell in terms of cost distance. These global functions are particularly useful for evaluating the connectivity of a landscape and the proximity of a cell to any given entities.

8 VECTOR DATA ANALYSIS

There is a wide range of functions for data analysis available in most GIS packages, including measurement techniques, attribute queries, proximity analysis, overlay operations and the analysis of models of surfaces and networks. The chapter begins by introducing methods for measurement and queries in GIS. Proximity, neighbourhood and then methods for integrating data using overlay functions are explained.

1 Measurements in GIS

Calculating lengths, perimeters and areas is a common application of GIS. However, it is possible that different measurements can be obtained depending on the type of GIS used (raster or vector) and the method of measurement employed. It is important to remember that all measurements from a GIS will be an approximation, since vector data are made up of straight line segments (even lines which appear as curves on the screen are stored as a collection of short straight line segments).

2 Queries

Performing queries on a GIS database to retrieve data is an essential part of most GIS projects. Queries offer a method of data retrieval, and can be performed on data that are part of the GIS database, or on new data produced as a result of data analysis. Queries are useful at all stages of GIS analysis for checking the quality of data and the results obtained. For example, a query may be used if a data point representing a hotel is found to lie erroneously in the sea after data encoding. A query may establish that the address of the hotel had been wrongly entered into a database, resulting in the allocation of an incorrect spatial reference.

There are two general types of query that can be performed with GIS: spatial and aspatial. Aspatial queries are questions about the attributes of features. ‘How many luxury hotels are there?’ is an aspatial query since neither the question nor the answer involves analysis of the spatial component of data. This query could be performed by database software alone. For the question ‘Where are the luxury hotels in Coimbatore?’, then the spatial component is added. Since this requires information about ‘where’ it is a spatial query. The location of the hotels will be reported and could be presented in map form. The method of specifying queries in a GIS can have a highly interactive mode. Users may interrogate a map on the computer screen or browse through databases with the help of prompts and query builders.

Queries can be made more complex by combination with questions about distances, areas and perimeters, particularly in a vector GIS, where these data are stored as attributes in a database. Individual queries can be combined to identify entities in a database that satisfy two or more spatial and aspatial criteria.

Boolean operators are often used to combine queries of this nature. These uses AND, NOT, OR and XOR, operations that are also used for the combination of different data sets by overlay. These are explained best with the help of Venn diagrams, where each circle in the diagram represents the set of data meeting a specific criterion (Figure 8.1). In the diagrams, A is the set of hotels that are in the ‘luxury’ category, and B is the set of hotels that have more than 20 bedrooms.

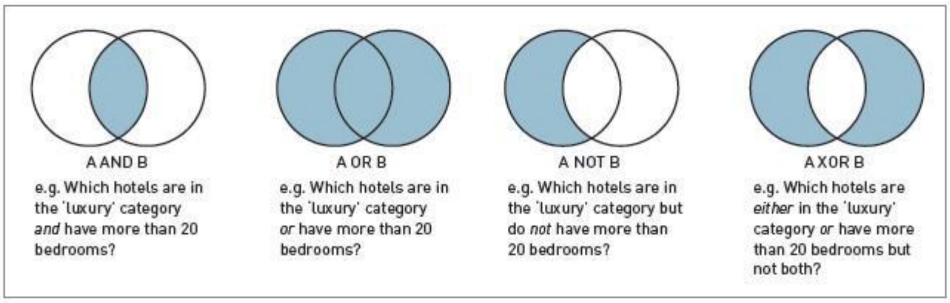


Figure 8.1 Boolean operators: Venn diagrams

3 Buffering Functions

Buffering is used to identify a zone of interest around an entity, or set of entities. If a point is buffered a circular zone is created. Buffering lines and areas creates new areas (Figure 21.2). Creating buffer zones around point features is the easiest operation; a circle of the required radius is simply drawn around each point. Figure 8.2 illustrates only the most basic set of buffer operations as there are many variations on this theme. For example, buffer zones may be of fixed or varying width according to feature attributes. When analyzing a road network, wide buffer zones could be attached to motorways and narrower buffer zones to minor roads to reflect traffic densities.

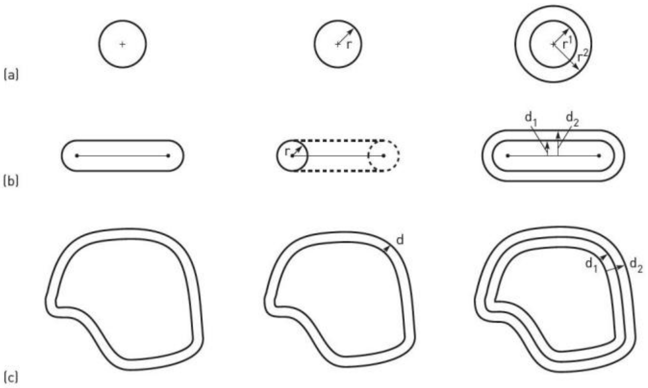


Figure 8.2 Buffer zones around (a) point, (b) line and (c) area features

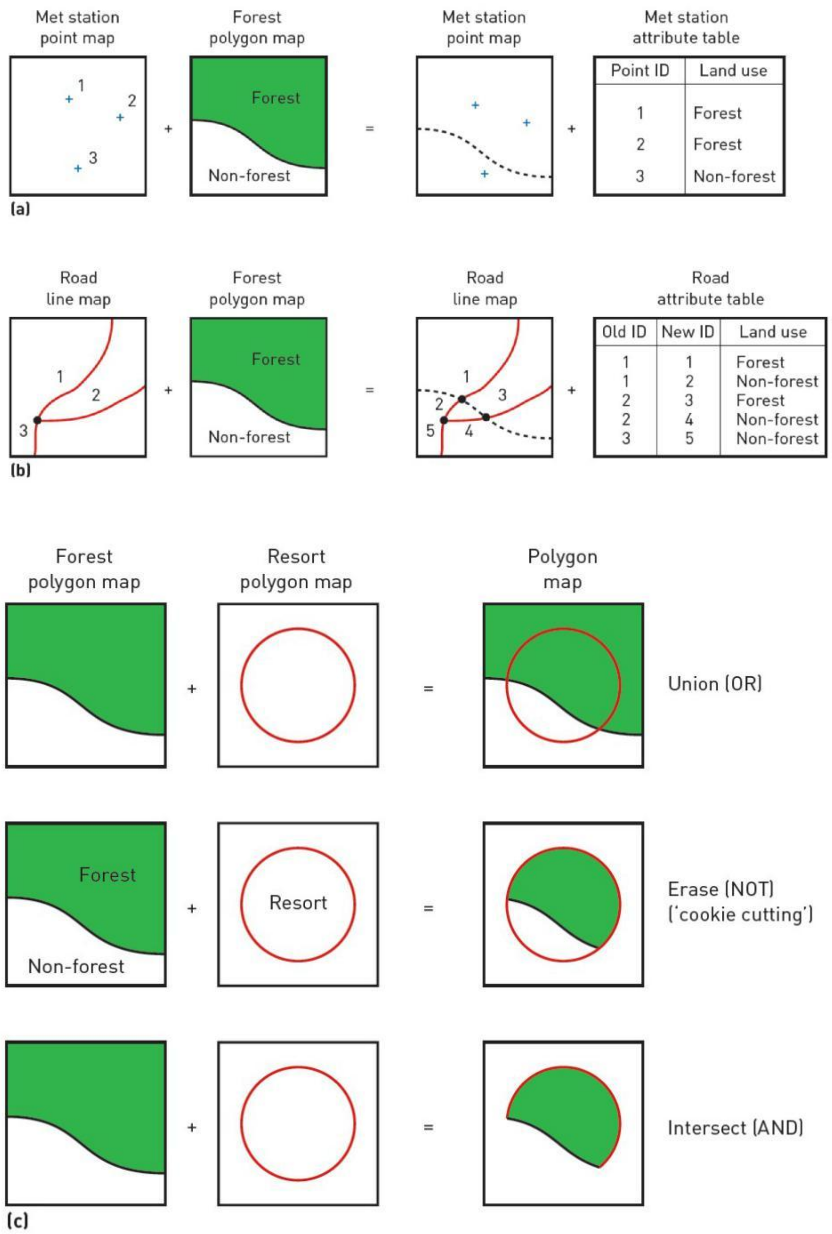
4 Map Overlay

The ability to integrate data from two sources using map overlay is perhaps the key GIS analysis function. Using GIS, it is possible to take two different thematic map layers of the same area and overlay them one on top of the other to form a new layer. Map overlay has many applications. At one level, it can be used for the visual comparison of data layers. Overlays where new spatial data sets are created involve the merging of data from two or more input data layers to create a new output data layer. This type of overlay may be used in a variety of ways.

Vector map overlay relies heavily on the two associated disciplines of geometry and topology. The data layers being overlaid need to be topologically correct so that lines meet at nodes and all polygon boundaries are closed. To create topology for a new data layer produced as a result of the overlay process, the intersections of lines and polygons from the input layers need to be calculated using geometry. For complex data this is no small task and requires considerable computational power. Figure 8.3 shows the three main types of vector overlay; point-in-polygon, line-in- polygon and polygon-on-polygon. This figure also illustrates the complexity of the overlay operations. The overlay of two or more data layers representing simple spatial features results in a more complex output layer. This will contain more polygons, more intersections and more line segments than either of the input layers.

***Point-in-polygon overlay* is used to find out the polygon in which a point falls. In the this example, meteorological stations are represented as points and land use as polygons. Using point-in-polygon overlay on these vector data layers it is possible to find out in which land use polygon each meteorological station is located. Figure 8.3(a) illustrates this overlay process. On the output map a new set of rain gauge points is created with additional attributes describing land use.**

***Line- in-polygon overlay* is more complicated. Imagine that we need to know where roads pass through forest areas to plan a scenic forest drive. To do this we need to overlay the road data on a data layer containing forest polygons. The output map will contain roads split into smaller segments representing 'roads in forest areas' and 'roads outside forest areas'. Topological information must be retained in the output map (Figure 8.3b), therefore this is more complex than either of the two input maps. The output map will contain a database record for each new road segment.**



Point ID	Land use
1	Forest
2	Forest
3	Non-forest

Old ID	New ID	Land use
1	1	Forest
1	2	Non-forest
2	3	Forest
2	4	Non-forest
3	5	Non-forest

Figure 8.3 Vector overlays: (a) point-in-polygon; (b) line-in-polygon; (c) polygon-on-polygon

Polygon-on-polygon **overlay could be used to examine the areas of forestry in the resort area. Two input data layers – a forest data layer containing forest polygons, and the resort boundary layer – are required. Three different outputs could be obtained (Figure 8.3c):**

1. **The output data layer could contain all the polygons from both of the input maps. In this case the question posed is 'Where are areas of forestry OR areas which are**

within the resort area?' This corresponds to the Boolean OR operation, or in mathematical set terms, UNION. This may be useful if the resort management committee was interested in buying new forest areas to extend the scenic forest drive.

2. The output data layer could contain the whole of the resort area, and forest within this. The boundary of the resort would be used as the edge of the output map, and forest areas would be cut away if they fall outside it. This operation is referred to as 'cookie cutting'. It is equivalent to the mathematical IDENTITY operation and the identity of the resort boundary is retained in the output. The questions being answered are 'Where the resort boundary is, and where areas of forest within this are?' This overlay might be used in preparation for calculation of the percentage of the area of the resort covered by forest.

3. The output data layer could contain areas that meet both criteria: that is, areas that are both forest and within the resort. An output map would be produced showing the whole of any forest polygons that are entirely within the resort boundary, and 'cut' away forest polygons that cross the resort boundary. This is the mathematical INTERSECT operation, and the output map shows where the two input layers intersect. 'Where are forest areas within resort' is the question being answered. As a thematic data layer showing forestry in the resort this may be useful for further analysis of the condition of the resort's forestry resources.

Overlay operations are seldom used in isolation. In practice, it is common to query a data layer first, and then perform an overlay. To obtain areas of forestry used in the examples above, it would be necessary to extract these areas from the land use data layer first using a query

5 Network analysis

A completely different set of analytic functions in GIS consists of computations on networks. A network is a connected set of lines, representing some geographic phenomenon, typically of the transportation type. The 'goods' transported can be almost anything: people, cars and other vehicles along a road network, commercial goods along a logistic network, phone calls along a telephone network, or water pollution along a stream/river network. Network analysis can be done using either raster or vector data layers, but they are more commonly done in the latter, as line features can be associated with a network naturally, and can be given typical transportation characteristics like capacity and cost per unit. One crucial characteristic of any network is whether the network lines are considered directed or not. Directed networks associate with each line a direction of transportation; undirected networks do not. In the latter, the 'goods' can be transported along a line in both directions. We discuss here vector network analysis, and assume that the network is a set of connected line features that intersect only at the lines' nodes, not at internal vertices. (But we do mention under- and overpasses.)

For many applications of network analysis, a planar network, i.e., one that is embeddable in a two-dimensional plane, will do the job. Many networks are naturally planar, like

stream/river networks. A large-scale traffic network, on the other end, is not planar: motorways have multi-level crossings and are constructed with underpasses and overpasses. Planar networks are easier to deal with computationally, as they have simpler topological rules. Not all GISs accommodate non-planar networks, or can do so only using trickery. Such trickery may involve to split overpassing lines at the intersection vertex and create four out of the two original lines. Without further attention, the network will then allow to make a turn onto another line at this new intersection node, which in reality would be impossible. Some GIS allow to associate a cost with turning at a node— see our discussion on turning costs below—and that cost, in the case of the overpass trick, can be made infinite to ensure it is prohibited. But, as we said, this is trickery to fit a non-planar situation into a data layer that presumes planarity.

The above is a good example of geometry not fully determining the network's behaviour. Additional application-specific rules are usually required to define what can and cannot happen in the network. Most GIS provide rule-based tools that allow the definition of these extra application rules.

Various classical spatial analysis functions on networks are supported by GIS software packages. The most important ones are:

Optimal path finding **which generates a least cost-path on a network between a pair of predefined locations using both geometric and attribute data.**

Network partitioning **which assigns network elements (nodes or line segments) to different locations using predefined critera.**

6 References

Raulph, A. De. 2001. Principles of Geographical Information System – An Introductory Textbook. ITC Educational Text book Series.

Ian Heywood, Sarah Cormelius and Steve Carber. 2006. An Introduction to Geographical Information System. Third Edition. Pearson Education Limited.

9 GEODESY

1 Geodesy

What is geodesy? Who needs it and why? These are some of the questions asked by many people. Actually, geodesy is nothing new having been around for centuries. It is a specialized application of several familiar facets of basic mathematical and physical concepts towards the understanding the shape and size of earth gravity and positioning. In practice, geodesy uses the principles of mathematics, astronomy and physics, and applies them within the capabilities of modern engineering and technology. A thorough study of the science of geodesy is not a simple undertaking. However, it is possible to gain an understanding of the historical development, a general knowledge of the methods and techniques of the science, and the way geodesy is being used today.

1.1 The shape and size of the earth

If an ellipse is rotated about its minor axis, it forms an ellipsoid of revolution. The customary ellipsoidal earth model has its minor axis parallel to the rotational axis of the earth. The size of such an ellipsoid is usually given by the length of the two semi-axes or by the semi-major axis and the flattening (Figures 9.1 and 9.2).

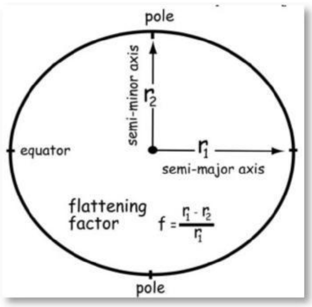
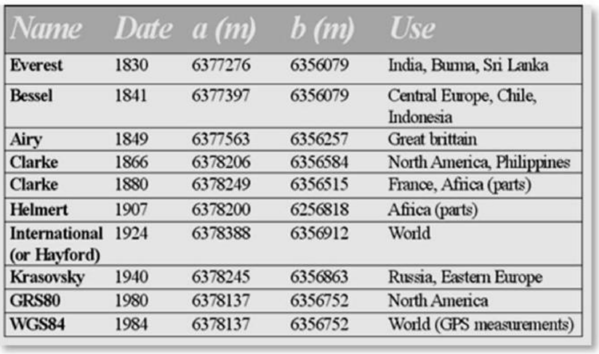


Figure 9.1 Ellipsoid



Name	Date	a (m)	b (m)	Use
Everest	1830	6377276	6356079	India, Burma, Sri Lanka
Bessel	1841	6377397	6356079	Central Europe, Chile, Indonesia
Airy	1849	6377563	6356257	Great brittain
Clarke	1866	6378206	6356584	North America, Philippines
Clarke	1880	6378249	6356515	France, Africa (parts)
Helmert	1907	6378200	6256818	Africa (parts)
International (or Hayford)	1924	6378388	6356912	World
Krasovsky	1940	6378245	6356863	Russia, Eastern Europe
GRS80	1980	6378137	6356752	North America
WGS84	1984	6378137	6356752	World (GPS measurements)

Figure 9.2 Example Ellipsoids

When the satellite data were analyzed further, it turned out that the nearest point in a satellite's orbit, the perigee, was always nearer to the earth when the satellite was over the northern hemisphere than when it was over the southern hemisphere. This indicates an asymmetry in the earth's shape. It is a little narrower in the north than in the south. Once, one had thought that the earth was a sphere, and then it seemed to be rather like a grapefruit. Now we found that it was slightly different from a grapefruit, rather like a pear.

Actually, things are quite complicated. When we talk about a pear-shape or an ellipsoid, we obviously do not mean the shape produced by the mountains and valleys, the topography. Since we can measure the elevations of places above sea level (this is what is recorded on topographic maps), we can discount them and inquire into the shape of what is left: that is, the sea-level surface itself, as if it were extended from the sea shore into the

land areas without those elevations above it. This sea-level surface is also called the GEOID (Figure 9.3).

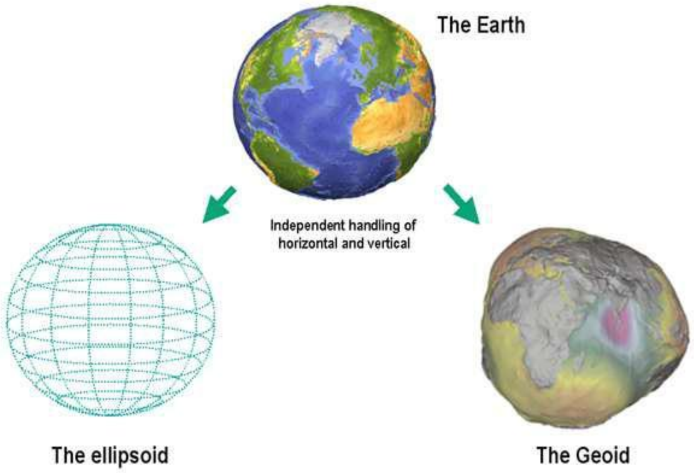


Figure 9.3 Geoidal Shape

The shape of the geoid is what we mean by the Figure of the Earth. We have found from many measurements that the shape of this geoid is very irregular as compared with an ellipsoid, and we describe these irregularities by the distances from the much smoother ellipsoid. These distances are called GEOIDAL HEIGHTS.

Thus we distinguish three surfaces: the topography, the geoid, and the ellipsoid (Figure 9.5) . Topographic maps give. the elevations above sea level (the geoid). Geoidal maps give the geoidal heights in relation to the ellipsoid. Both together give the total height of the topography above the ellipsoid at any point.

You have now learned the story of our developing knowledge of the size and shape of the earth. The shape of the earth is partly attributed to the force of gravity. The way we study the gravity field is the subject of the next topic.

1.2 The gravity field of the earth

The earth's attraction called gravity, causes things to fall. Remember the story about Newton sitting under an apple tree? When an apple fell and hit him, it started him thinking of a new theory of gravity. A heavy plumb bob, suspended by a string, is attracted by the earth and therefore pulls that string into a straight downward (vertical) direction.

Gravity holds the water to the earth and our feet to the ground so we don't fall off into space. About five-sevenths of the earth's surface is covered by oceans. The level surface which coincides with mean sea level is called the GEOID. Other similar level surfaces can be imagined at any elevation, for example, the water surface of mountain lakes. The higher

a level surface is above the geoid, the further removed it is from the irregularities 1n the earth's structure; thus the warping will be less pronounced.

1.3 Point Positioning

When driving along or hiking in unknown territory, your location can be of vital importance to you. An answer such as "You are right next to a huge anthill" will not be very helpful to find your way home, even if it is correct and useful from another aspect. You need an answer in relation to some known reference such as the nearest town or highway. You will want to know how far you are from that town; and in what direction you should turn off from your present path - or from the north direction.

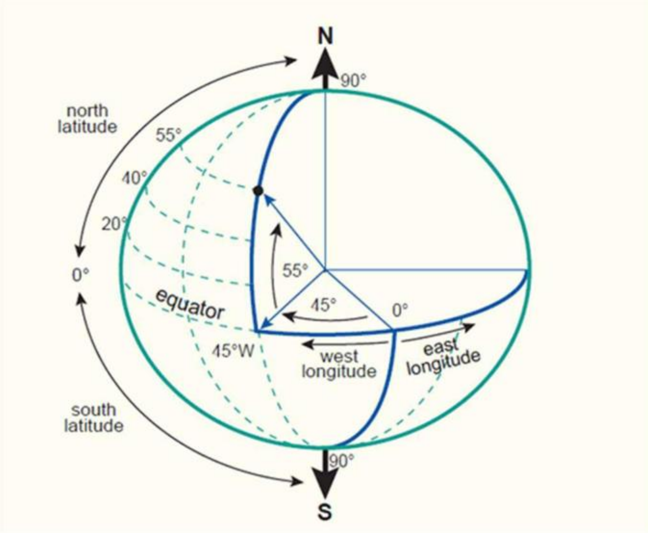
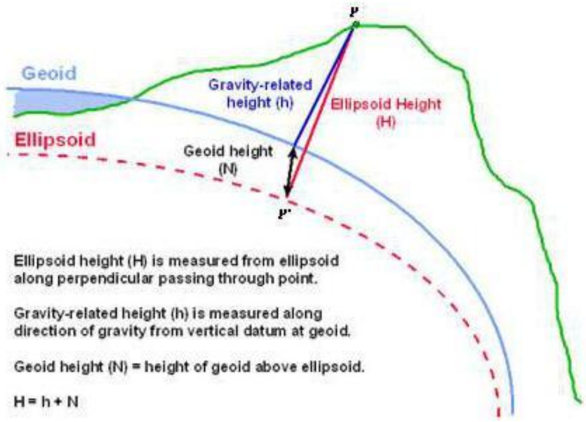


Figure 9.4 Latitude and Longitude

For geodetic systems large enough to be affected by the curvature of the earth, we use an ellipsoidal earth model and designate on it the position of any point, in terms of latitude, longitude, and height, with zero starting references for each. LATITUDE refers to a set of circles parallel to the equator, called parallels, the way you slice a tomato. The numbering starts at the equator and goes to 90_o north and 90_o south. LONGITUDE refers to a set of ellipses (or circles if the earth model is taken as a sphere) called meridians, the way you divide an orange. Their numbering starts customarily with the meridian through Greenwich in England and goes either to 360_o eastward or to 180_o east and 180_o west. The HORIZONTAL POSITION of a point is at the intersection of a parallel and a meridian and is therefore expressed in terms of latitude and longitude. The complete position must include a third value, the vertical position. This is the height of the point above or below the ellipsoid. It must also include the specifications of the ellipsoid itself. The latter, together with the complete posit10n of a particular point, the datum point, is called a GEODETIC DATUM.



Ellipsoid height (H) is measured from ellipsoid along perpendicular passing through point.

Gravity-related height (h) is measured along direction of gravity from vertical datum at geoid.

Geoid height (N) = height of geoid above ellipsoid.

H = h + N

Figure 9.5 Ellipsoidal and Geoidal Height

The figure above (Figure 9.5) is a meridional section through the earth. The height H of the point P above the ellipsoid is measured along the normal to the ellipsoid. The point P' on the ellipsoid itself has a zero height. P and P' have the same horizontal position (same latitude and longitude), but different vertical positions. The total height H of a point P above the ellipsoid is not directly observed. Leveling determines one part, the elevation h above mean sea level, which is also called the GEOID. The other part, the GEOIDAL HEIGHT N, that is the separation between the geoid and the ellipsoid, must be computed separately. Failure to compute the geoidal height will make the height coordinate of the point P incorrect. Theoretically, geoidal heights may vary between +100 meters and -100 meters, but in practice some geodetic systems may have several hundred meters of geoidal heights in some areas. For example, the old South American Datum had geoida1 heights of about 300 meters in Chile. Therefore, South America accepted a new continental datum in 1969, including a new reference ellipsoid, where geoidal heights are less than 50 meters. In Southeast Asia, the Indian Datum is still used with more than 300 m of geoidal heights, although better fitting datums have been computed. Geodesy can correct these problems.

Instead of describing the position of a point P in terms of latitude, longitude, and height as we do in surveying and practical applications in the field, we can also use a Cartesian coordinate system in x, y, z. The latter is often used within the process of computations, especially in satellite computations. One coordinate system can be converted into another by mathematical conversion formulas. But if one is incorrect, for instance, through the neglect of a large geoidal height, then the other will be incorrect also. Even if the positional inaccuracy of a specific point may seem tolerable in itself, it may snowball into significant errors in certain applications.

2 Map Projection

A map projection is a mathematically described technique of how to represent the Earth’s curved surface on a flat map. To represent parts of the surface of the Earth on a flat paper map or on a computer screen, the curved horizontal reference surface must be mapped onto the 2D mapping plane. Mapping onto a 2D mapping plane means transforming each

point on the reference surface with geographic coordinates (,) to a set of Cartesian coordinates (x,y) representing positions on the map plane (Figure 9.6)

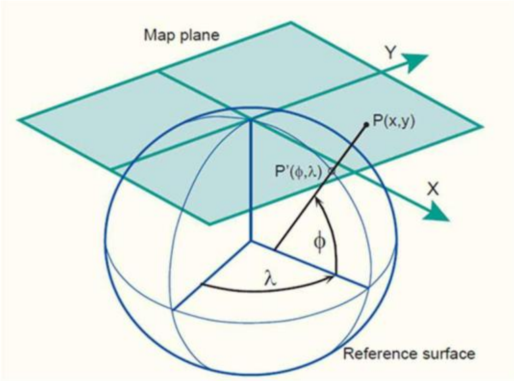


Figure 9.6 Map Projection

Most projections can be envisaged as analogous to shining a light through a scaled-down model of the Earth, known as the reference globe (or generating globe), onto a surface, referred to as a developable surface, which can be 'rolled out' into a flat plane. Different types of projection can be defined depending upon the source of the light (e.g. the centre of the Earth, or a point infinitely far away) and the location and shape of the projection surface (e.g. cone, cylinder or plane). The projected lines of latitude and longitude form a graticule. These lines are not necessarily equally spaced, may converge and may be curved, depending upon the projection (Figure 9.7). A rectangular co-ordinate system, known as a grid, is often superimposed for the purpose of providing grid references, etc.

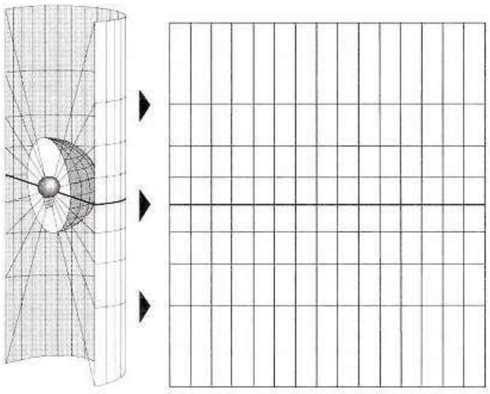


Figure 9.7 Lines of Projection

3 Classification of map projections

Map projections can be described in terms of their:

- (i). class (cylindrical, conical or azimuthal),
- (ii). point of secancy (tangent or secant),
- (iii). aspect (normal, transverse or oblique), and
- (iv). distortion property (equivalent, equidistant or conformal).

Class

The three classes of map projections are cylindrical, conical and azimuthal. The Earth's reference surface projected on a map wrapped around the globe as a cylinder produces a cylindrical map projection. Projected on a map formed into a cone gives a conical map projection. When projected directly onto the mapping plane it produces an azimuthal (or zenithal or planar) map projection. The figure below shows the surfaces involved in these three classes of projections.

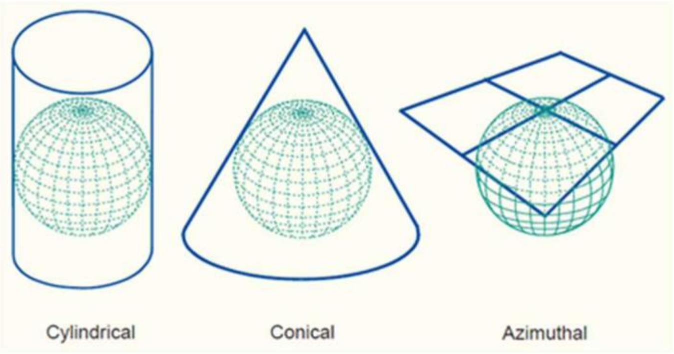


Figure 9.8 Class of Projections

Point of secancy

The planar, conical, and cylindrical surfaces in the figure above are all tangent surfaces; they touch the horizontal reference surface in one point (plane) or along a closed line (cone and cylinder) only. Another class of projections is obtained if the surfaces are chosen to be secant to (to intersect with) the horizontal reference surface; illustrations are in the figure below. Then, the reference surface is intersected along one closed line (plane) or two closed lines (cone and cylinder). Secant map surfaces are used to reduce or average scale errors because the line(s) of intersection are not distorted on the map (section 4.3 scale distortions on a map).

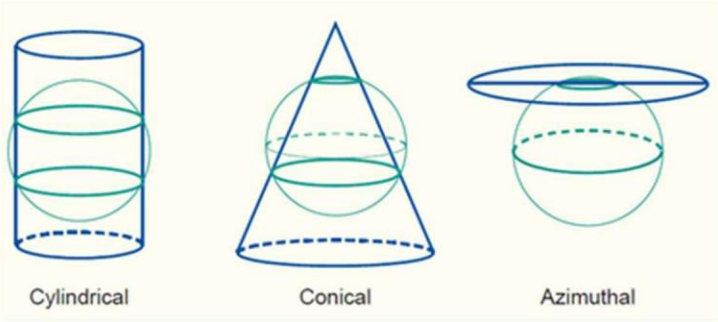


Figure 9.9 Secant projection classes

Aspect

Projections can also be described in terms of the direction of the projection plane's orientation (whether cylinder, plane or cone) with respect to the globe. This is called the aspect of a map projection. The three possible aspects are normal, transverse and oblique. In a normal projection, the main orientation of the projection surface is parallel to the Earth's axis (as in the figures above for the cylinder and the cone). A transverse projection has its main orientation perpendicular to the Earth's axis. Oblique projections are all other, non-parallel and non-perpendicular, cases. The figure below provides two examples.

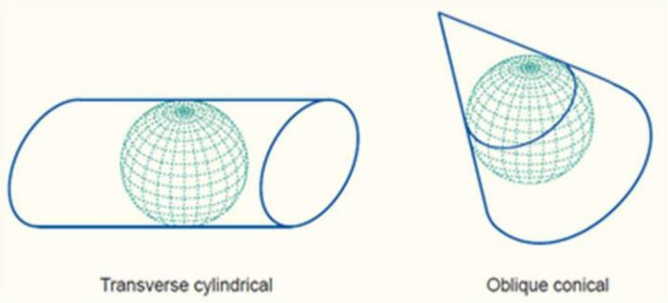


Figure 9.10 Transverse and oblique map

Distortion

So far, we have not specified how the Earth's reference surface is projected onto the plane, cone or cylinder. How this is done determines which kind of distortion properties the map will have compared to the original curved reference surface. The distortion properties of map are typically classified according to what is not distorted on the map:

In a conformal (orthomorphic) map projection the angles between lines in the map are identical to the angles between the original lines on the curved reference surface. This means that angles (with short sides) and shapes (of small areas) are shown correctly on the map.

In an equal-area (equivalent) map projection the areas in the map are identical to the areas on the curved reference surface (taking into account the map scale), which means that areas are represented correctly on the map.

In an equidistant map projection the length of particular lines in the map are the same as the length of the original lines on the curved reference surface (taking into account the map scale).

4 Choosing a map projection

Every map must begin, either consciously or unconsciously, with the choice of a map projection and its parameters. The cartographer's task is to ensure that the right type of projection is used for any particular map. A well chosen map projection takes care that scale distortions remain within certain limits and that map properties match to the purpose of the map.

Generally, normal cylindrical projections are typically used to map the world in its entirety (in particular areas near the equator are shown well). Conical projections are often used to map the different continents (the mid-latitudes regions are shown well), while the polar azimuthal projections may be used to map the polar areas. Transverse and oblique aspects of many projections can be used for most parts of the world, though they are usually more difficult to construct.

In theory, the selection of a map projection for a particular area can be made on the basis of:

the shape of the area,

the location (and orientation) of the
area, and the purpose of the map.

10 GPS - COMPONENTS AND FUNCTIONS

1 Introduction

Space-based geodetic observations can be categorized into four basic techniques: positioning, altimetry, interferometric synthetic aperture radar (InSAR), and gravity studies.

Precise positioning is the fundamental geodetic observation required for surveying and mapping. Instead of the traditional triangulation and levelling networks that require line of sight (LOS) between measurement points, space geodetic methods use LOS between the measurement points and celestial objects or satellites.

Building on this idea, scientists have developed advanced positioning techniques through Global Navigation Satellite Systems (GNSS) . GNSS encompasses the various satellite navigation systems, such as the United States' GPS, Russia's Globalnaya Navigatsionnaya Sputnikovaya Sistema (GLONASS), Japan's Quazi- Zenith Satellite System (QZSS), India's Indian Regional Navigation Satellite System (IRNSS), China's Beidou and Europe's Galileo. Although these satellite systems were designed mainly for navigation, they were found to be very useful for precise positioning, with accuracy levels of less than a centimeter. GNSS also provides very high temporal resolution measurements (second by second, or even faster), yielding key observations of time-dependent processes in the lithosphere, atmosphere, and ionosphere.

2 GNSS Architecture

A GNSS basically consists of three main segments: the space segment, which comprises the satellites; the control segment (also referred to as the ground segment), which is responsible for the proper operation of the system; and the user segment, which includes the GNSS receivers providing positioning, velocity and precise timing to users.

2.1 Space Segment

The main functions of the space segment are to generate and transmit code and carrier phase signals, and to store and broadcast the navigation message uploaded by the control segment. These transmissions are controlled by highly stable atomic clocks onboard the satellites. The GNSS space segments are formed by satellite constellations with enough satellites to ensure that users will have at least four satellites in view simultaneously from any point on Earth's surface at any time.

The GPS (US NAVSTAR) satellites are arranged in six equally spaced orbital planes surrounding Earth, each with four `slots' occupied by baseline satellites. This 24-slot arrangement ensures there are at least four satellites in view from virtually any point on the planet. The satellites are placed in a Medium Earth Orbit (MEO) orbit, at an altitude of 20 200km and an inclination of 55° relative to the equator. Orbits are nearly circular, with an eccentricity of less than 0.02, a semi-major axis of 26,560km and a nominal period

of 11 hours, 58 minutes and 2 seconds (12 sidereal hours), repeating the geometry each sidereal day.

The nominal GLONASS constellation consists of 24 MEO satellites deployed in three orbital planes with eight satellites equally spaced in each plane. The orbits are roughly circular, with an inclination of about 64.8°, and at an altitude of 19,100km with a nominal period of 11 hours, 15 minutes and 44 seconds, repeating the geometry every eight sidereal days. Due to funding problems, the number of satellites decreased from the 24 available in 1996 to only 6 in 2001. In August 2001, the Russian government committed to recover the constellation and to modernise the system, approving new funding. A total of 24 operational satellites plus 2 in maintenance were again available in December 2011, restoring the full constellation.

The Galileo constellation in Full Operational Capability (FOC) phase consists of 27 operational and 3 spare MEO satellites at an altitude of 23,222 km and with an orbit eccentricity of 0.002. Ten satellites will occupy each of three orbital planes inclined at an angle of 56° with respect to the equator. The satellites will be spread around each plane and will take about 14 hours, 4 minutes and 45 seconds to orbit Earth, repeating the geometry each 17 revolutions, which involves 10 sidereal days. This constellation guarantees, under nominal operation, a minimum of six satellites in view from any point on Earth's surface at any time, with an elevation above the horizon of more than 10°.

The Beidou (Compass) constellation (Phase III) will consist of 35 satellites, including 5 Geostationary Orbit (GEO) satellites and 30 non-GEO satellites in a nearly circular orbit. The non-GEO satellites include 3 Inclined Geosynchronous Satellite Orbit (IGSO) ones, with an inclination of about 55°, and 27 MEO satellites orbiting at an altitude of 21,528km in three orbital planes with an inclination of about 55° and with an orbital period of about 12 hours and 53 minutes, repeating the ground track every seven sidereal days. The GEO satellites, orbiting at an altitude of about 35 786 km, are positioned at 58.75°E, 80°E, 110.5°E, 140°E and 160°E, respectively, and are expected to provide global navigation service by 2020. The previous Phase II involves a reduced constellation of four MEO, five GEO and five IGSO satellites to provide regional coverage of China and surrounding areas. The initial Phase II operating service with 10 satellites started on 27 December 2011.

The Indian Regional Navigation Satellite System (IRNSS) consists of a constellation of seven satellites (IRNSS-1A, IRNSS 1-B, IRNSS 1-C, IRNSS 1-D, IRNSS 1-E, IRNSS 1-F and IRNSS 1-G). IRNSS 1-A was launched in 2013 and the last one of the series IRNSS 1-G was launched on April 28, 2016. This is an independent Indian Satellite based positioning system for critical National applications. The main objective is to provide Reliable Position, Navigation and Timing services over India and its neighbourhood, to provide fairly good accuracy to the user. The IRNSS will provide basically two types of services, viz., Standard Positioning Service (SPS) and Restricted Service (RS). Space Segment consists of seven satellites, three satellites in GEO stationary orbit (GEO) and four satellites in Geo Synchronous Orbit (GSO) orbit with inclination of 29° to the equatorial plane. This constellation of seven satellites was named as "NavIC" (Navigation with Indian Constellation)

2.2 Control Segment

The control segment (also referred to as the ground segment) is responsible for the proper operation of the GNSS. Its basic functions are to:

1. control and maintain the status and configuration of the satellite constellation;
2. predict ephemeris and satellite clock evolution;
3. keep the corresponding GNSS time scale (through atomic clocks); and
4. update the navigation messages for all the satellites.

2.3 User Segment

The user segment is composed of GNSS receivers. Their main function is to receive GNSS signals, determine pseudoranges (and other observables) and solve the navigation equations in order to obtain the coordinates and provide a very accurate time. The basic elements of a generic GNSS receiver are: an antenna with preamplification, a radio frequency section, a microprocessor, an intermediate-precision oscillator, a feeding source, some memory for data storage and an interface with the user. The calculated position is referred to the antenna phase centre.

Various GNSS receivers are available in the market, from chips on watches and mobile phones, to tracking devices, amateur receivers with small antenna, mapping receiver with single or dual frequency capable antenna, survey grade dual or triple frequency receivers, geodetic survey receivers with special antenna and high data rate, mentioned in increasing order of price and accuracy. They may cost from about Rs. 3,000 to about Rs. 30,00,000.

3 GNSS SIGNALS

GNSS satellites continuously transmit navigation signals at two or more frequencies in L band. These signals contain ranging codes and navigation data to allow users to compute both the travel time from the satellite to the receiver and the satellite coordinates at any epoch. The main signal components are described as follows:

Carrier: Radio frequency sinusoidal signal at a given frequency.

Ranging code: Sequences of zeros and ones which allow the receiver to determine the travel time of the radio signal from the satellite to the receiver. They are called PRN (Pseudo Random Noise) sequences or PRN codes.

Navigation data: A binary-coded message providing information on the satellite ephemeris (pseudo-Keplerian elements or satellite position and velocity), clock bias parameters, almanac (with a reduced-accuracy ephemeris data set), satellite health status and other complementary information.

The current `legacy' Navigation Message (NAV) is modulated on both carriers at 50 bps. The whole message contains 25 pages (or `frames') of 30 s each, forming the master frame that takes 12:5 min to be transmitted. Every frame is subdivided into five subframes of 6 s each; in turn, every subframe consists of 10 words, with 30 bits per word (figure above of NAVSTAR GPS). Every subframe always starts with the telemetry word TLM, which is necessary for synchronisation. Next, the transference word (HOW) appears. This word provides time information (seconds of the GPS week), allowing the receiver to acquire the week-long P(Y) code segment.

4 The Position Fix By Trilateration

As soon as a receiver is powered on it starts searching for satellites. However ignorance of satellites' approximate position delays the time taken for the first position fix. Therefore an almanac is needed to speed up this process. The almanac is a small file that provides the positions of the GNSS satellites to a certain degree of accuracy for a 48 hours period. The tracking stations monitor the satellites and pass the information to the master control station where the information is used among other things to generate the almanac file and upload them to each satellite. The user receivers while powered on can download this file from the satellite in a matter fo12.5 minutes of through the internet.

Then receivers then lock on to each satellite and receive the ephemerides from each satellite. The ephemerides provide the current information about the satellites. The receiver must then align signals sent from the satellite to an internally generated version of a pseudorandom binary sequence, also contained in the signal. Since the satellite signal takes time to reach the receiver, the two sequences do not initially coincide; the satellite's copy is delayed in relation to the local copy. The receiver generates the pseudorandom sequence, but they do not match. By increasingly delaying the local copy, the two copies can eventually be aligned. The correct delay represents the time needed for the signal to reach the receiver, and from this the distance from the satellite can be calculated (Figure 10.1).

Propagation Time = Time Signal Reached Receiver - Time Signal Left Satellite.

Multiplying this propagation time by the speed of light gives the distance to the satellite.

Distance or Pseudo Range 'D' = Speed of light in vacuum × Propagation Time

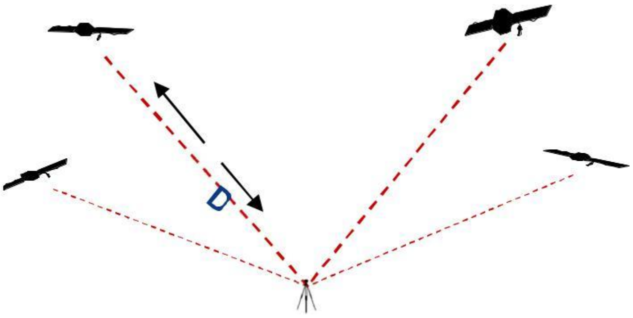


Figure 10.1 Distance calculation

Knowing the position of the satellites from their ephemerides, the receiver calculates its position. The receiver knows that the reason the pseudoranges to the three satellites are not intersecting is because its clock is not very good and apply an ingenious techniques to correct its clock error. The receiver is programmed to advance or delay its clock until the pseudoranges to the three satellites converge at a single point as seen in the following figure.

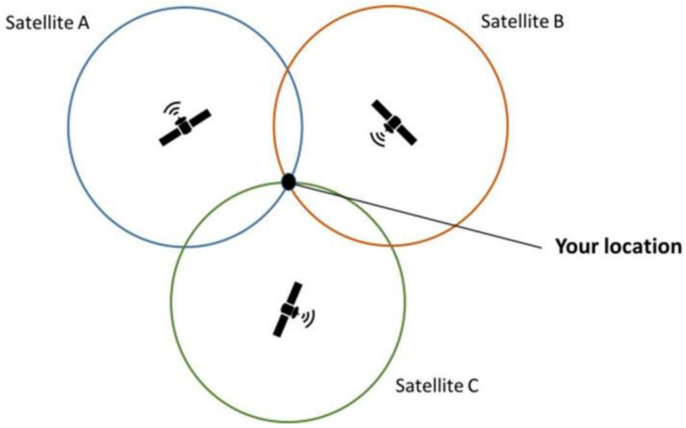


Figure 10.2 Concept of Trilateration

The accuracy of the resulting range measurement is essentially a function of the ability of the receiver's electronics to accurately process signals from the satellite, and additional error sources such as non mitigated ionospheric and tropospheric delays, multipath, satellite clock and ephemeris errors, etc.

5 **Errors in Position**

5.1 Clock Errors

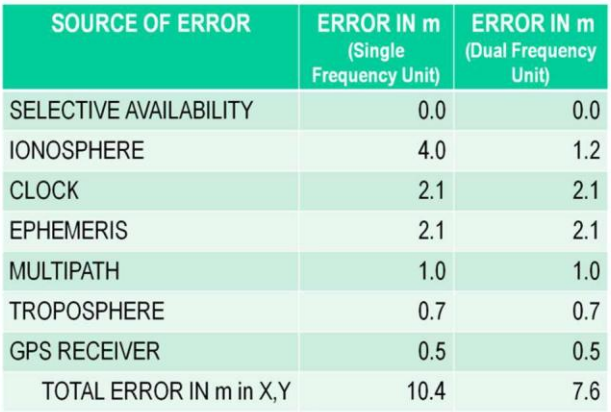
Fundamental to GNSS is the one-way ranging that ultimately depends on satellite clock predictability. These satellite clock errors affect both the C/A- and P-code users in the same way. This effect is also independent of satellite direction, which is important when the technique of differential corrections is used. All differential stations and users measure an identical satellite clock error. The ability to predict clock behaviour is a measure of clock quality. The GPS uses atomic clocks (cesium and rubidium oscillators) which have stability of about 1 part in 10E13 over a day. If a clock can be predicted to this accuracy, its error in a day (~10E5 s) will be about 10E- 8 s or about 3.5 m.

5.2 Ephemeris Errors

Ephemeris errors result when the GNSS message does not transmit the correct satellite location. Because satellite errors reflect a position prediction, they tend to grow with time from the last control station upload. These errors were closely correlated with the satellite clock, as one would expect. Note that these errors are

the same for both the P- and C/A- codes. Each satellite has a unique Precision (P) and Coarse Acquisition (CA) codes that distinguish between the different satellites comprising the GNSS.

Table 10.1 The Various sources of Error



SOURCE OF ERROR	ERROR IN m (Single Frequency Unit)	ERROR IN m (Dual Frequency Unit)
SELECTIVE AVAILABILITY	0.0	0.0
IONOSPHERE	4.0	1.2
CLOCK	2.1	2.1
EPHEMERIS	2.1	2.1
MULTIPATH	1.0	1.0
TROPOSPHERE	0.7	0.7
GPS RECEIVER	0.5	0.5
TOTAL ERROR IN m in X,Y	10.4	7.6

5.3 Multipath errors

Multipath is the error caused by reflected signals entering the front end of the receiver and masking the real correlation peak. These effects tend to be more pronounced in a static receiver near large reflecting surfaces. Monitor or reference stations require special care in siting to avoid unacceptable errors. The first line of defense is to use the combination of antenna cut-off angle and antenna location that minimizes this problem. A second approach is to use so-called "narrow correlator" receivers, which tend to minimize the impact of multipath on range tracking accuracy.

5.4 Ionospheric errors

Because of free electrons in the ionosphere, GPS signals do not travel at the vacuum speed of light as they transit this region. The modulation on the signal is delayed in proportion to the number of free electrons encountered and is also (to first order) proportional to the inverse of the carrier frequency squared (1/f2). The phase of the radio frequency carrier is advanced by the same amount because of these effects. Carrier-smoothed receivers should take this into account in the design of their filters. The ionosphere is usually reasonably well-behaved and stable in the temperate zones; near the equator or magnetic poles it can fluctuate considerably. Due to the above the delays range from a few meters at night to a maximum of 10 or 20 m at about 1400 hrs.

5.5 Troposphere errors

Another deviation from the vacuum speed of light is caused by the troposphere. Variations in temperature, pressure, and humidity all contribute to variations in the speed of light and radio waves. Both the code and carrier will have the same delays.

5.6 Dilution of Precision

The geometry formed by the observed positions of satellites by a receiver at a point in time can present an estimate of the achievable accuracy. Any receiver will try to use signals from satellites in a manner that reduces the DOP value. A value of 6 or less is regarded acceptable. DOPs can change with time and space. The DOP can be further defined as separate elements as Horizontal DOP (HDOP), Vertical DOP (VDOP) and Position DOP (PDOP).

6 Differential Correction

Standalone GNSS receivers are prone for the errors discussed above. Hence a DGNNS receiver is positioned at a known location (reference/base station) and coordinates computed and errors determined. This error can then be applied as a correction to nearby rover stations surveyed in the project area within a vicinity of about 50 km. It should be noted however that the farther the rover from the base, more the error. It is assumed that environmental factors are similar at base and rover locations.

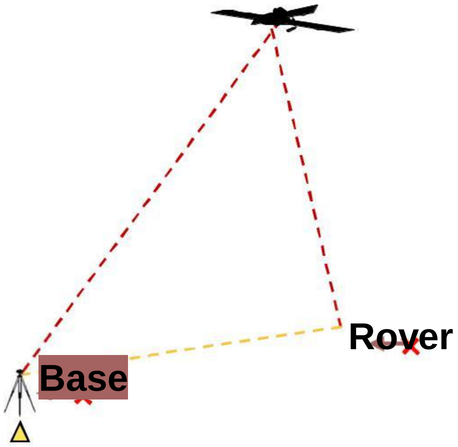


Figure 10.3 Differential Correction

Data collected at Rover stations should overlap in both TIME and GNSS Satellite Vehicle so that corrections for the exact same satellites at the exact same time can be applied. Data from rovers can be brought to the office at the end of the survey day and processed in a software along with the base station data. This is referred as the classical DGNSS operation the Static Post Processed, and gives the best accuracies. However, it requires longer observation times than Real Time Kinematic discussed below.

Where a project dictates the availability of corrected position value is real time, the corrections can be broadcast from the base over a radio link and rovers receiving them in real time for applying the corrections. This is referred as Real Time Kinematic as the corrections are applied on the go. The method takes advantage of the slow variation with time and user position of the errors due to ephemeris prediction, residual satellite clocks, ionospheric and tropospheric delays. Starting from the reference station, the system computes and broadcasts either correction to the GNSS position or to the pseudorange measurements to the DGNSS users. Other uncorrelated errors (e.g. multipath) cannot be corrected by this method and specific techniques have to be applied to mitigate them.

The difficulty in making an RTK system is properly aligning the signals. The navigation signals are deliberately encoded in order to allow them to be aligned easily, whereas every cycle of the carrier is similar to every other. This makes it extremely difficult to know if you have properly aligned the signals or if they are "off by one" and are thus introducing an error of 20 cm (approximate wave length of the carrier), or a larger multiple of 20 cm. This integer ambiguity problem can be addressed to some degree with sophisticated statistical methods that compare the measurements from the C/A signals and by comparing the resulting ranges between multiple satellites. However, none of these methods can reduce this error to zero.

11 PRECISION AGRICULTURE

Precision agriculture or Precision farming also known as site-specific management refers to the practice of applying agronomic inputs across a farm, mainly fertilizers and other chemicals, at variable rates based on soil nutrients or chemical tests, soil textural changes, weed pressures and yield maps for each field in the farm.

Precision farming also has the potential for improving water use efficiency on large fields provided there is a quantitative understanding of what factors and where in the field they affect crop-water use. In most fields (eg: > 40 ha) crop yields are noticeably variable. The sources of this variation are related to the physical and chemical properties of the soil, pests, microclimate, genetic and phonological responses of the crop and their interactions. The technology for crop yield mapping and understanding the soil process that explain the crop yield variability must be done at the landscape level and by using appropriate statistical and remote sensing tools for large scale mapping (Lascano, 2002).

1 Why Precision Agriculture?

There are several reasons that precision farming has come about as a management method in the recent past:

- **High cost of crop inputs including seed, fertilizer, pesticides and fuel**
- **Environmental concerns about fertilizers and pesticides near sensitive areas, runoff and de‑nitrification**
- **The technology has become available and economically feasible**

2 Many names of Precision Agriculture

Precision agriculture goes by many names but they all refer to managing variability:

- **Precision farming**
- **GPS farming**
- **Prescription farming**
- **Farming by satellite**
- **Spatially variable agriculture**
- **Farming by the foot**
- **Site specific management**
- **Variable rate application**

3 The Basic Components of Precision Farming

Precision farming basically depends on measurement and understanding of variability, the main components of precision farming system must address the variability. Precision farming technology enabled, information based and decision focused, the components include, (the enabling technologies) Remote Sensing (RS), Geographical Information System (GIS), Global Positioning System (GPS), Soil Testing, Yield Monitors and Variable Rate Technology.

4 Application of Remote Sensing, GIS and GPS in Precision Farming

Remote sensing is a potential tool in providing spatial and temporal information on soil and crop variables which could be related with crop growth and yield models. The kind of information on soil and crop variables largely depends on the remote sensing platforms (ground based, aircraft and satellite) and the remote sensors (camera radiometers and scanners with different resolutions).

GIS is a computer – based technology capable of gathering, storing, analyzing and retrieving geographically referenced data. GIS combines different kinds of data (map, tables, digital data and point data). It integrates layers of information about to give better understanding of that place. For example, it can combine soil maps, rainfall maps, topographic maps and land use maps to show areas where there is high danger of soil erosion. GIS requires suitable software for the data analysis and integration.

GIS is an invaluable tool in planning and monitoring of natural resources like soils, lad use etc., at a regional or national level. It can be used as decision making tool in agriculture. It can take in to account of soil fertility, gradient of lands, annual rainfall, availability of rural labour and access to markets.

Without having a reliable method of locating equipment and items in a field, it is difficult to manage in-field variability. A crude method might be to stake out the field to show areas that require different treatment, but this is not practical on large fields. A reliable positioning method is needed to accurately locate field features to make precision agriculture work. Some local positioning systems were developed but not successfully commercialized. The advent of GPS allowed for low-cost, reliable positioning of equipment in the field. Data from other sensors could be tied to a specific point in the field with precision.

Role of GPS in precision agriculture:

- **Yield mapping**
- **Variable rate control**
- **Field mapping**
- **Asset tracking**
- **Irrigation**
- **Tracking livestock**
- **Aerial spraying**
- **Autosteering**
- **Drainage**
- **Guidance**

5 Steps in Precision Farming

The basic steps in precision farming are,

(i). Assessing variability
(ii). Managing variability and
(iii). Evaluation

i). Assessing variability

Assessing variability is the critical first step in precision farming. It is clear that one cannot manage what one does not know. Factors and the processes that regulate or control the crop performance in terms of yield vary in space and time. Quantifying the variability of these factors and processes and determining when and where different combinations are responsible for the spatial and temporal variation in crop yield is the challenge for precision agriculture.

ii). Managing variability

Once variation is adequately assessed, farmers must match agronomic inputs to known conditions employing management recommendations. Those are site specific and use accurate applications control equipment. For successful implementation, the concept of precision soil fertility management requires that within-field variability exists and is accurately identified and reliably interpreted, that variability influences crop yield, crop quality and for the environment. Therefore inputs can be applied accurately.

The higher the spatial dependence of a manageable soil property, the higher the potential for precision management and the greater its potential value. The degree of difficulty, however, increases as the temporal component of spatial variability increases. Applying this hypothesis to soil fertility would support that Phosphorus and Potassium fertility are very conducive to precision management because temporal variability is low. For N, the temporal component of variability can be larger than its spatial component, making precision N management much more difficult in some cases.

iii). Evaluation

There are three important issues regarding precision agriculture evaluation. 1.Economics, 2. Environment and 3. Technology transfer

The most important fact regarding the analysis of profitability of precision agriculture is that the value comes from the application of the data and not from the use of the technology. Potential improvements in environmental quality are often cited as a reason for using precision agriculture. Reduced agrochemical use, higher nutrient use efficiencies, increased efficiency o f managed inputs and increased production of soils from degradation are frequently cited as potential benefits to the environment. Enabling technologies can make precision agriculture feasible, agronomic principles and decision

rules can make it applicable and enhanced production efficiency or other forms of value can make it profitable.

6 Soil Test Crop Response (STCR) Studies – Concept and Methodology

The first phase of soil testing laboratories Advisory Service in India was initiated in 1955-56 under the TCM-USAID Programme. The soil test calibrations (low, medium and high) and fertilizer recommendations advocated to the then tall varieties of food crops were qualitative in nature.

With the introduction of high yielding varieties and hybrids of crops triggering green revolution, a new phase of quantitative soil test - crop response correlation (STCR) studies were conceived and implemented through an All India Co-ordinated STCR Project by Dr. B.Ramamoorthy in 1967 under the aegis of the Indian Council of Agriculture Research (ICAR). A novel field experimentation methodology was devised and soil test - based fertilizer dose calibrations were derived by creating a macrocosm of soil fertility variability within a microcosm of the experimental field.

The various aspects of soil test calibration and soil fertility management themes covered under the project include:

1. **Efficient fertilizer recommendation according to the investment capacity of the farmer Fertilizer allocations under conditions of fertilizer/credit shortage**
2. **Fertilizer recommendation for targeted yields and maintenance of soil fertility**
3. **Prediction of post- harvest soil test values from initial soil test values in multiple cropping system**
4. **Apportioning the fertiliser application between crops in a multiple cropping system for increased fertiliser use efficiency**
5. **Area wise fertiliser recommendation based on yield targeting and nutrient index of soil fertility**
6. **Limitations to extensive use of fertilisers and the possible methods of overcoming them**
7. **The need for maintenance of optimum C/N ratio for increased nutrient efficiency when both organic manures and fertilisers are used.**

The soil test based fertiliser recommendation for targeted yield of crops under fertilizer resource constraints. The "Law of optimum" as propounded by Ramamoorthy and Velayutham (2011) is an experimentally proven concept of soil test based major plant nutrients (N, P and K) applications to crops for desired targeted yields, based on the derivation of three parameters from standard soil test - crop response factorial field experiments, namely 1) per cent contribution (efficiency) from the soil available nutrients (Cs) as estimated by chemical soil tests in soil testing laboratory, 2) per cent contribution of nutrients from added fertilisers (Cf) and manures (Cm) and 3) nutrient requirement of the crop (Kg/ton) as estimated from yield and plant nutrient uptake data from STCR field experiments.

12 VARIABLE RATE TECHNOLOGY

There are a number of questions that must be answered before establishing a site-specific, or precision, management program for crop production. Many of those questions are economic, some are agronomic, and others are technology-related. One important technology-related question is: "What methods of variable-rate application of fertilizer, crop chemicals, and seed are available?"

There are two basic methods of implementing site-specific management (SSM) for the Variable-Rate Application (VRA) of crop production inputs: *map-based and sensor-based*. While each method has unique benefits and limitations, some SSM systems have been developed to take advantage of the benefits of both methods.

The first site-specific management method is based on the use of maps to represent crop yields, soil properties, pest infestations, and variable-rate application plans. The map-based method can be implemented using a number of different strategies. Crop producers and consultants have crafted strategies for varying inputs based on: soil type, color and texture, topography (high ground, low ground), crop yield, field scouting data, remotely sensed images, and a host of other sources. Some strategies are based on a single information source while others involve a combination of sources. Regardless of the actual strategy, the user is in control of the development process.

To develop a plan for variable-rate fertilizer application in a particular field, the map-based method could include the following steps:

- perform systematic soil sampling (and lab analysis) for the field;
- generate site-specific maps of the soil nutrient properties of interest;
- use some algorithm to develop a site-specific fertilizer application map; and
- use the application map to control a variable-rate fertilizer applicator.

A positioning system is used during the sampling and application steps to continuously know or record vehicle location in the field. Differentially-corrected Global Positioning System (DGPS) receivers are the most commonly used positioning devices. The process of map-based, variable-rate application is illustrated in Figure 12.1.

The second SSM method provides the capability to vary the application rate of crop production inputs with no mapping involved. The sensor-based method utilizes sensors to measure the desired properties, usually soil properties or crop characteristics, on the go. Measurements made by such a system are then processed and used immediately to control a variable- rate applicator (Figure 12.2) . This second method doesn't necessarily require the use of a DGPS system. Nor does it require extensive data analysis prior to making variable-rate applications.

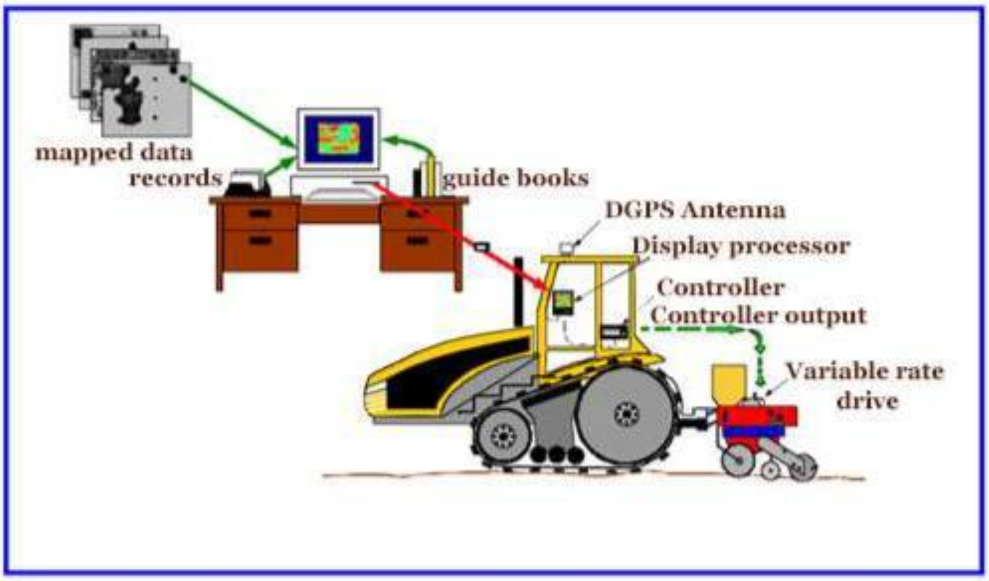


Figure 12.1 An illustration of a map based system for varying crop inputs

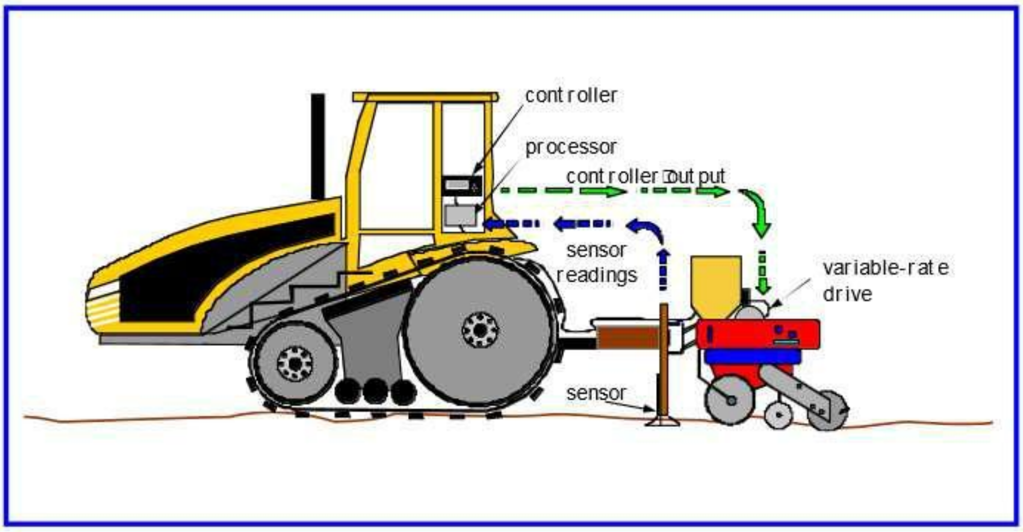


Figure 12.2 An illustration of a sensor based system for varying crop inputs

1 Map-based Technologies

Currently, the majority of available technologies and applications in site-specific farming utilize the map- based method of sampling, map generation, and variable- rate application. This method is more popular due to the scarcity of sensors for rapidly monitoring soil and crop conditions. Also, laboratory analysis is still the most trusted and reliable method

for determining most soil and plant properties. Once field data have been collected and assigned position coordinates (e.g. latitude and longitude), mapping is easily performed using a computer program (usually a geographic information system (GIS) program). Such programs can use mathematical techniques for “smoothing” or interpolating the data between sampling points. However, some site-specific practitioners choose to use a constant value for the measured property over each sampling area or grid cell (Figure 12.3). As illustrated in Figure 3, the level indicated by each grid cell is determined by analyzing samples collected from the center of each cell. This represents the common practice of using software to divide a field into a set of imaginary, equal-sized rectangles or grid cells, identifying the center of each cell, and directing that samples be collected from the vicinity of each grid cell center.

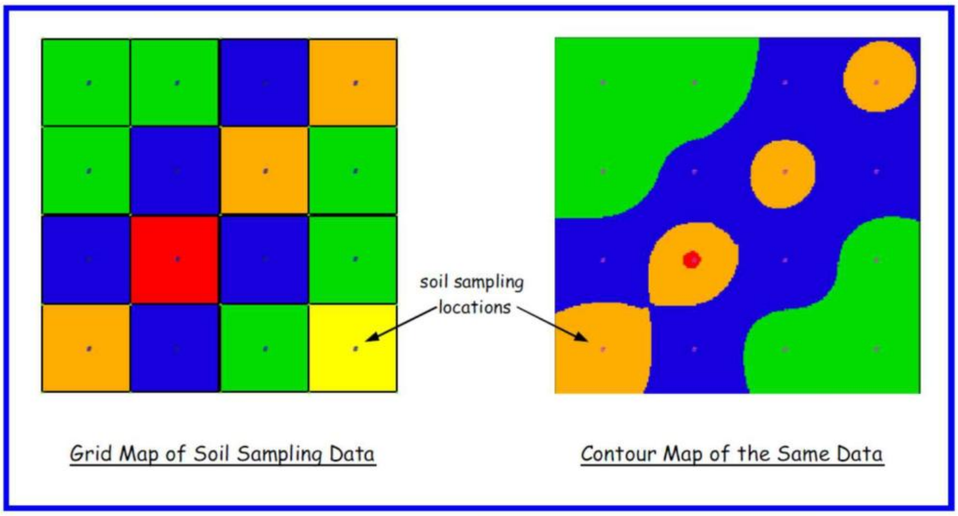
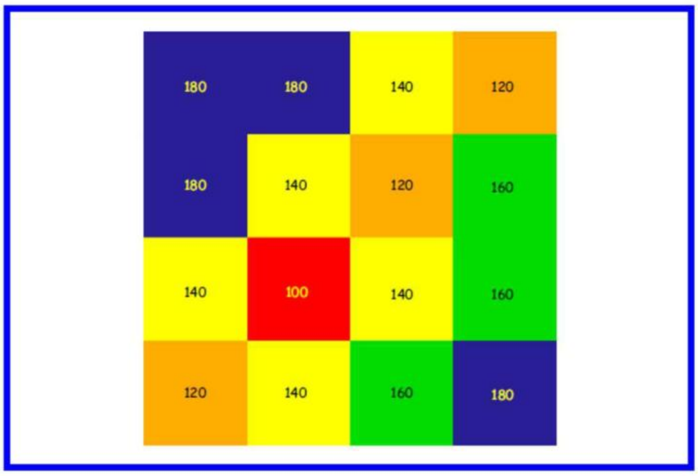


Figure 12.3 Two alternatives for representing systematic soil sampling data

Regardless of how data are represented, the mapping facilitates long-term planning and analysis. It provides an opportunity to make decisions regarding the selection and purchase of crop production inputs well in advance of their use. Maps are especially good for collecting and interpreting data for soil properties that do not fluctuate greatly from year to year. Properties such as organic matter content and soil texture tend to change quite slowly, if at all. Soil fertility, on the other hand, may change more quickly.

Particular nutrients such as phosphorous and potassium may change from year to year, but one can probably obtain benefits from sampling only every two to three years. Levels of other nutrients may vary considerably even during a single season. For instance, the forms and concentrations of nitrogen in the soil are greatly affected by temperature and moisture conditions and can fluctuate rapidly. Nitrogen is an example of an important soil fertility factor that doesn’t lend itself to a typical site-specific management program based on soil sampling data due to delays between sampling and fertilizer application. Nitrogen management approaches that rely on other, more stable information such as crop yield potential have been developed. In order to use computer-generated maps, they must be converted to a form that can be used by a variable-rate applicator. The conversion process is performed using specialized software that applies user-selected algorithms (mathematical recipes). Algorithms are usually based on standard fertilizer recommendation formulas. The application map contains application rate information for all locations within a field. A rate map such as the one illustrated in Figure 12.4 is typically generated by software running on a desktop computer. The application map is then transferred to a data card that is read by a drive in the in-cab application system processor, then used by application software acting through a controller onboard an applicator to deliver the proper rate at each location in the field. Again, a DGPS system must be used to continuously correlate the vehicle's location in the field with a coordinate on the map and the desired application rate for that coordinate.



180	180	140	120
180	140	120	160
140	100	140	160
120	140	160	180

Figure 12.4 Example of an application rate map

Most variable- rate controllers are designed to synchronize the application rate with the position in the field by "looking ahead" on the map for the next change in rate. This takes into account the ground speed of the vehicle and the time required to change the rate of product coming out of the applicator. A fertilizer spreader truck may operate at field speeds exceeding 25 km per hour. Without the "look ahead" feature, if the applicator took only one second to respond to a rate change command from the controller, an area at least 10 m long and as wide as the spread pattern would be treated at the wrong rate. With the "look ahead" feature available in map-based application systems, it is not necessary to reduce travel speeds to accomplish accurate variable-rate applications.

One benefit of the map-based method is the knowledge of the needed amounts of chemicals, or inputs, for the operations prior to entering a field. This knowledge can aid in managing field operations. The multiple sources of data that are necessary to facilitate map-based applications can also be used in other decision-making processes for a farming operation. A farm manager using GIS software can examine all yield, soil property, pest, and as-applied data.

With typical map-based variable-rate application systems, the high cost of the soil analysis limits the number of samples that a farmer can afford to test. There is currently much

discussion on the optimum number of acres represented by each sample and the location of those samples.

2 Sensor-Based Technologies

While knowing how much product will be needed is a benefit of map-based systems, sensor-based systems hold a significant advantage in sampling density. A typical map-based application program is based on a single sample or small set of samples from 2.5-acre areas within a field. A sensor- based system can collect dozens of "samples" from each acre. This increase in sampling density should produce a more accurate depiction of within-field variability.

At this point, the major challenge is to develop sensors that will work accurately in field conditions at realistic working speeds. Sensor-based application systems must be capable of accomplishing the sensing, data processing, and application rate adjustment steps in one machine pass. Speed, both in regard to sensing and processing, is a major requirement of true sensor-based systems. There is lag time between sensing a soil or crop property and converting the sensor signal to information that can be used by the system to change the rate of application. Developers of sensor-based systems must synchronize the sensor measurement site with the desired application rate for that same site. In some instances, the sensor may have to be mounted on the front of the tractor, or applicator truck, to give the variable rate controller enough time to adjust the rate accordingly before it passes the sensed location.

In order to effectively accomplish this on-the-go control, the sensors must respond almost instantaneously to changes in the soil or crop characteristics. One component of an on-the-go control system that has been developed is a soil organic matter sensor. This sensor is designed to facilitate the variable-rate application of dry soil-applied herbicides and/or blended fertilizer on the go, without a map. The organic matter sensor consists of a light sensor (photodiode) surrounded by six light sources (light emitting diodes or LEDs). The light sensor measures the amount of light reflected by the soil. This reflection signal is related to the amount of organic matter in the soil. High organic matter content results in dark soil color and a reduction in light reflectance. Moisture can also affect the sensor but as long as the soil is uniformly moist, the effects are small.

Some technologies for on-the-go sensing and variable-rate control are already on the market. One such system is the Soil Doctor® (Crop Technology, Inc., Houston, TX). The system uses pairs of ground-engaging rolling electrodes to examine soil type, organic matter, cation exchange capacity, soil moisture, and nitrate nitrogen levels in the soil volume between electrode pairs. By sensing these properties on the go, the need for a positioning system is eliminated and the data processing is greatly reduced because no maps are required. And, if the operator desires to record the sensor outputs and use this information for other operations, the system is capable of interfacing with a GPS receiver and generating site-specific maps.

Researchers around the world are actively developing additional sensors for on-the-go soil property measurements including: nitrate nitrogen, pH, potassium, phosphorous, and soil

texture. Application systems that use variations in plant canopy color as the basis for varying nitrogen fertilizer application rates are being developed. So are post-emergence sprayers that can distinguish between weeds and crops. When these research and development efforts succeed, site-specific farming will become more economical – possibly even automatic. In the meantime, there are approaches available to take advantage of sensors within a more traditional map-based variable rate application program. There are sensor-based tools that takes advantage of a high-rate sampling to create data-dense soil property maps. Many private companies manufacture devices that measure soil electrical conductivity (EC). They offer vehicle-drawn units that use rolling electrodes (coulters) as sensing elements and combine EC data with GPS-supplied position data. Data collected by the units can be used to produce highly-detailed maps of soil electrical conductivity. EC information can then be related to soil physical characteristics such as texture and topsoil depth. This information can then be used to produce variable-rate application plans.

Both map- and sensor-based variable-rate application systems are available to the site specific farmer. There are also VRA strategies that incorporate aspects of both sensing and mapping. Each variable-rate application method holds advantages and disadvantages. Strong points of each system are summarized below:

3 Advantages of Map-Based Variable-Rate Application

- systems are already available for most crop production inputs
- the user has a database that can be useful for a number of management-related activities
- the user can employ multiple sources of information in the process of formulating a variable-rate application plan
- the user has significant control regarding the function of such systems because of the involvement in application rate planning
- field travel speeds need not be reduced

4 Advantages of Sensor-Based Variable Rate Application

- pre-application data analysis time requirements can be eliminated
- sensors produce far higher data resolution than traditional sampling methods
- no time delay between measurement and application with real-time systems
- systems are self-contained

It is important to match the application system with the objectives of the overall site-specific management program in which it will be used. Producers should expect an increasing number of options for both map-based and sensor-based site-specific operations as research and development efforts continue.

13 CROP DISCRIMINATION AND YIELD MONITORING USING REMOTE SENSING

Identifying, discriminating and mapping crops is important for a number of reasons. Maps of different crop types are necessary to prepare an inventory of what was grown in which area and when. This serves the purpose of forecasting grain supplies (yield prediction), collecting crop production statistics, assessment of crop damage due to storms and drought and monitoring farming activity

The two most important data requirement activities related to agriculture are crops grown with their extent and yield monitoring. Traditional methods of obtaining these information is to take ground survey, which consumes more time and labour. Also, the accuracy of these data is also under close scrutiny.

Why remote sensing?

Remote sensing offers an efficient and reliable means of collecting information required for crop mapping and area estimation. Remote sensing enables collection of data in inaccessible areas and over large areas in short time. Besides providing a synoptic view, remote sensing can provide additional information on the status, stage and health of the crop under consideration.

1 Crop mapping / discrimination

Crop mapping or Crop discrimination is the preparation of maps with crops grown in the area of interest using spectral differences of multi data remote sensing data.

The data (historical or in-season data) used for crop mapping depends on the scale of the project. Generally, for crop mapping projects, multi temporal (multi date) data will be used to identify the different stages of crops and thereby identifying the crop itself. Ancillary data (rainfall data, collection in addition to remote sensing data is also of paramount importance

Crop mapping and acreage estimation through remote sensing (Figure 13.1) broadly consists of identifying representative sites (sampling plan) of various crops / land cover classes on the image based on the ground truth collected. Ground truth planning and collection is another important step. While collecting ground truth points, care should be taken so that no classes / crops are left out. Also, number of samples to be collected are to be estimated in prior before going for ground truth collection. Ground truth collection should be done during different stages of crop growth. Once ground truth collection is completed, generation of signatures for different training sites and classifying the image using training statistics has to be done. Finally, accuracy of classification is to be estimated and the user's, producer's and overall accuracy is to be estimated. Based on the crop concentration statistics, agrophysical and / or agroclimatic conditions, the study area is divided into homogenous strata and sample segments from each stratum are analysed.

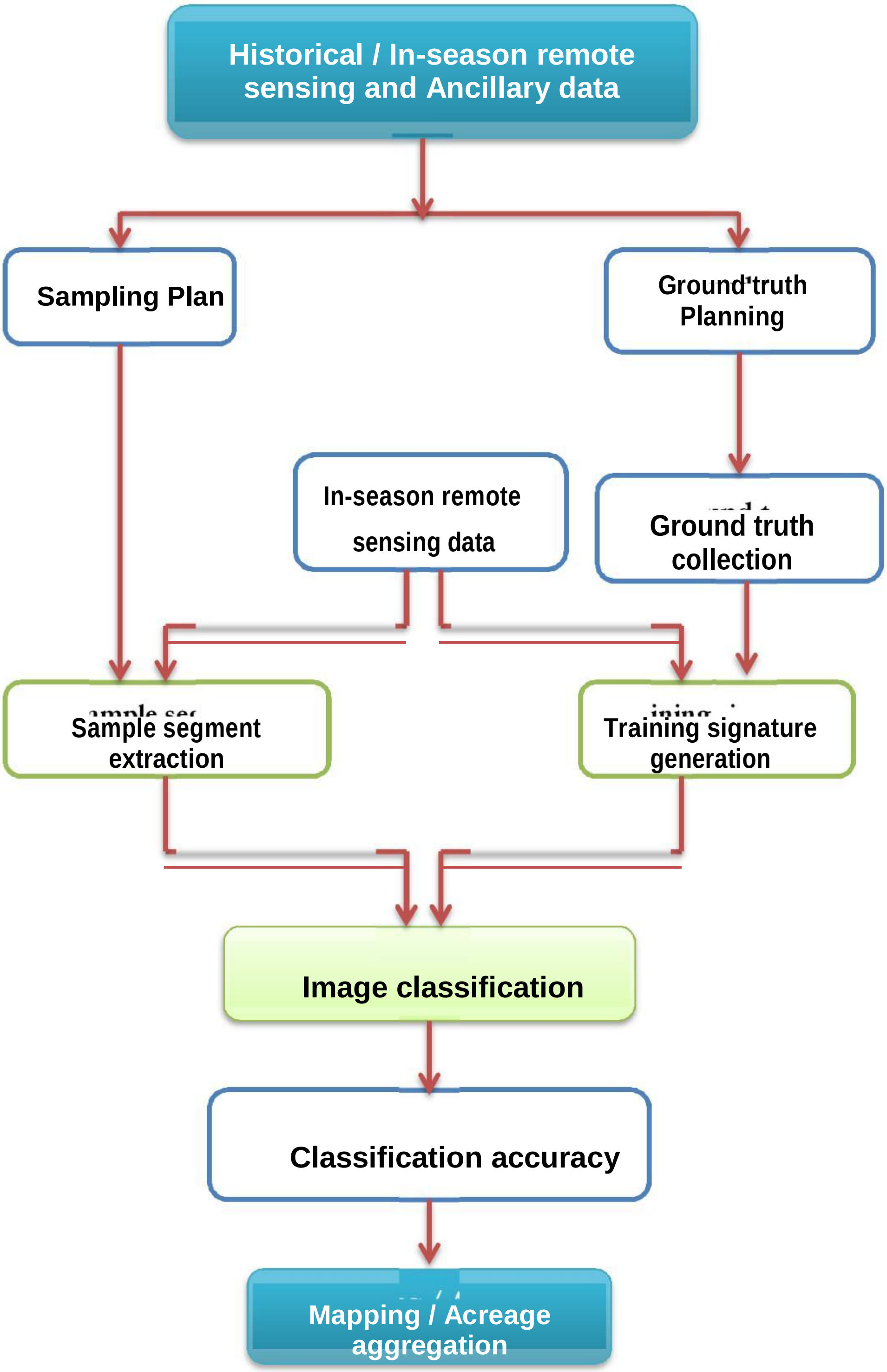


Figure 13.1 Methodology for Crop Mapping and Acreage estimation

During the Kharif season, the availability of cloud free data of optical sensors are difficult and thereby the potential of microwave sensor operated in C-band is utilized for acreage estimation and crop monitoring.

2 Yield monitoring

Yield is influenced by a large number of factors such as crop genotype, soil characteristics, cultural practices adopted (e.g. irrigation, fertilizer), weather conditions, and biotic influences, such as weeds, diseases, pests, etc.

The procedure for yield mapping is explained in the Figure 13.2. The classified data with different crops is the input required. A crop mask highlighting specific crop can be prepared using any GIS software. Remote Sensing data is used to estimate some of the biometric parameters, which in turn are input parameters to a yield model. Spectral index of the crop canopy (NIR / Red, Greenness, NDVI) at any given point of time reveals the crop growth and its decay as affected by various factors in the time domain. These inputs are provided to any crop model and the output is overlaid on the crop map to arrive at the yield map.

The crop health/condition is affected by factors such as supply of water and nutrients, insect/pest attack, disease out-break and weather conditions. These stresses cause physiological changes which may alter the optical properties of leaves and bring about changes in crop geometry. The regular monitoring of crop health helps in differentiation of stressed crops from the normal crop at a given time, quantification of extent and severity of stress and ultimately assessment of production loss, if any.

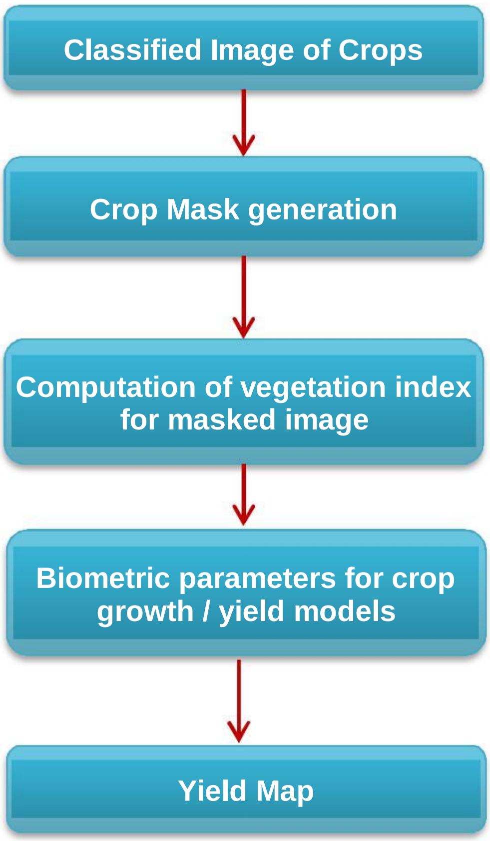


Figure 13.2 Methodology for Yield mapping

14 SOIL MAPPING AND FERTILIZER RECOMMENDATION USING REMOTE SENSING & GIS

The database required for Farm-Level Planning can be obtained by carrying out detailed characterization and mapping of all the existing land resources like soils, climate, water, minerals and rocks, vegetation, crops, land use pattern, socio-economic conditions, infrastructure, marketing facilities and various schemes and developmental works of the government. From the data collected at farm level, the problems and potentials of the area can be highlighted, conservation measures required for the area can be indicated, suitability of the area for various uses can be worked out and finally viable and sustainable land use options suitable for each and every land holding can be prescribed to the farmer and other land users of the area. The soil survey are of different types based on the level of mapping and the required details,

1 Detailed Soil Survey (Conventional method)

Detailed Soil Survey (usually at 1:4000; 1:5000; 1:8000 or 1:10,000 scales) provides the information necessary to identify homogeneous management units in the field at village level. The methodology followed in the field and laboratory is briefly indicated below.

1.1 Base Maps Used

The detailed survey of the villages can be carried out by using cadastral maps as a base. The cadastral map shows all the field boundaries with their survey numbers, location of tanks, streams and other permanent features of the area.

Remote sensing data products (1:12,500 scale) were used in conjunction with the cadastral maps to identify the landforms and surface features of the area. The Imageries helped in the delineation of the boundary between the uplands and lowlands, water bodies, forest and vegetated areas, salt affected lands, roads, habitations and other cultural features of the area.

Survey of India toposheets at 1:50,000 scale can be also used as a base for initial traversing, identification of geology and landform, drainage features, present land use and for the selection of transects at block level.

1.2 Field Investigation

Preliminary traverse can be carried out by using 1:50,000 scale toposheets. During the traverse, geological formations, drainage patterns, surface features, slope characteristics, types of land use and landforms were identified.

In the selected transect, profiles can be located at closely spaced intervals to take care of any change in the land features like break in slope, erosion, gravel, stones etc. In the selected sites, profiles (vertical cut showing the soil layers from the surface to the rock) were opened up to 200 cm or to the depth limited by rock or hard substratum and studied

in detail for all their morphological and physical characteristics. The soil and site characteristics can be recorded for all profile sites on a standard proforma.

Based on the soil-site characteristics, the soils were grouped into different soil series (soil series is the most homogeneous unit having similar horizons and properties and behaves similarly for a given level of management). Soil depth, texture, colour, amount and nature of gravel present, calcareousness, presence of limestone, nature of substratum and horizon sequence were the major identifying characteristics of soil series in the area. Phases of soil series (phase is a subdivision of a soil series based mostly on surface features that affect its use and management. For example, slope, texture, erosion and stoniness) were separated and their boundaries delineated on the cadastral maps based on the variations observed in the surface texture, slope, erosion, presence of gravels, salinity, sodicity etc.

The delineated mapping units occurring in each village are shown on the soil map in the form of symbols. In arriving the phases, a combination of letters, both in upper and lower case, and numerals were used. For example, the map unit SvcB3 occurring in Illuppakudi village is a phase of Sivagangai series. In this, the first two letters indicate the name of the soil series, the third lower case letter indicates the texture of the surface soil, the fourth upper case indicates the slope of the land and the fifth numeral indicates the severity of the soil erosion.

Similarly other features like gravelliness, salinity, sodicity, etc which affects the use of the land can be indicated in phases. Wherever gravelly phase occurs, the upper case letter G follows the name of soil series in the third position

1.3 Laboratory analysis

For the soil series identified, soil samples has to be collected from representative pedons for laboratory analysis. The soil samples are to be analysed for various physio, chemical characteristics by following standard procedures.

1.4 Finalisation of Soil Maps

The soil map for each village can be finalized in the field itself after thorough checking of soil and site characteristics. The village soil maps can be generated using GIS without any generalization and loss of information and presented for each village separately.

From the village maps, the soil map of the Block can be prepared by combining all the soil maps of the villages through the use of GIS software. Since the village maps were at larger scale with phases as mapping units, they were subjected to both cartographic and categorical generalization to prepare the soil map of the block.

1.5 Generation of Thematic Maps

The soil map and other resources database can be interpreted for identifying the constraints and potentials, and evaluated for land capability, land irrigability, fertility capability and land suitability for various crops and other uses for each village separately by using GIS software.

By using thematic information table for all the mapping units occurring in a village, required thematic maps can be generated by using GIS software at any time based on the needs of the farmers or any other users.

The major problem faced in conventional soil survey and cartography is the accurate delineation of boundary. Field observations based on conventional soil survey are tedious and time consuming. The remote sensing data in conjunction with ancillary data provide the best alternative, with a better delineation of soil mapping units (Karale 1992; Kudrat et al.,1990; 1992; Mulders & Epema 1986; Sehgal 1995).

Soil surveyors consider the topographic variation as the basis for depicting the soil variability. Even with the aerial photographs only physiographic variation in terms of slope, aspects and land forms are identified for delineating the soil boundary. Multispectral satellite data are being used for mapping soil up to family association level (1:50,000). The methodology in most of the cases involves visual interpretation (Karale et al., 1981). However, computer aided digital image processing technique has also been used for mapping soil (Epema 1986; Korolyuk & Sheherbenko 1994; Kudrat et al., 1990) and advocated to be a potential tool (Kudrat et al., 1992; Lee et al., 1988).

2 Visual image interpretation

Visual interpretation is based on shape, size, tone, shadow, texture, pattern, site and association. This has the advantage of being relatively simple and inexpensive. Soil mapping needs identification of a number of elements. The elements which are of major importance for soil survey are land type, vegetation, landuse, slope and relief. Soils are surveyed and mapped, following a 3 tier approach, comprising interpretation of remote sensing imagery and/or aerial photograph (Mulder 1987), field survey (including laboratory analysis of soil samples) and cartography (the art and science of map making)(Sehgal et al., 1989). Several workers (Karale 1992; Kudrat & Saha 1993; Kudrat et al. 1990; Sehgal 1995) have concluded that the technology of remote sensing provides better efficiency than the conventional soil survey methods (USDA 1951) at the reconnaissance (1:50,000) and detailed (1:10,000) scale of mapping (Figure 14.1).

3 Computer-aided approach

Numerical analysis of remote sensing data utilizing the computers has been developed because of requirement to analyze faster and extract information from the large quantities of data. The computer aided techniques utilize the spectral variations for classification. The pattern recognition in remote sensing assists in identification of homogeneous areas, which can be used as a base for carrying out detailed field investigations, and generating models between remote sensing and field parameters. However, there is a need to have an automated method for accurate soil boundary delineation with a transdisciplinary and integrated approach (Figure 14.2).

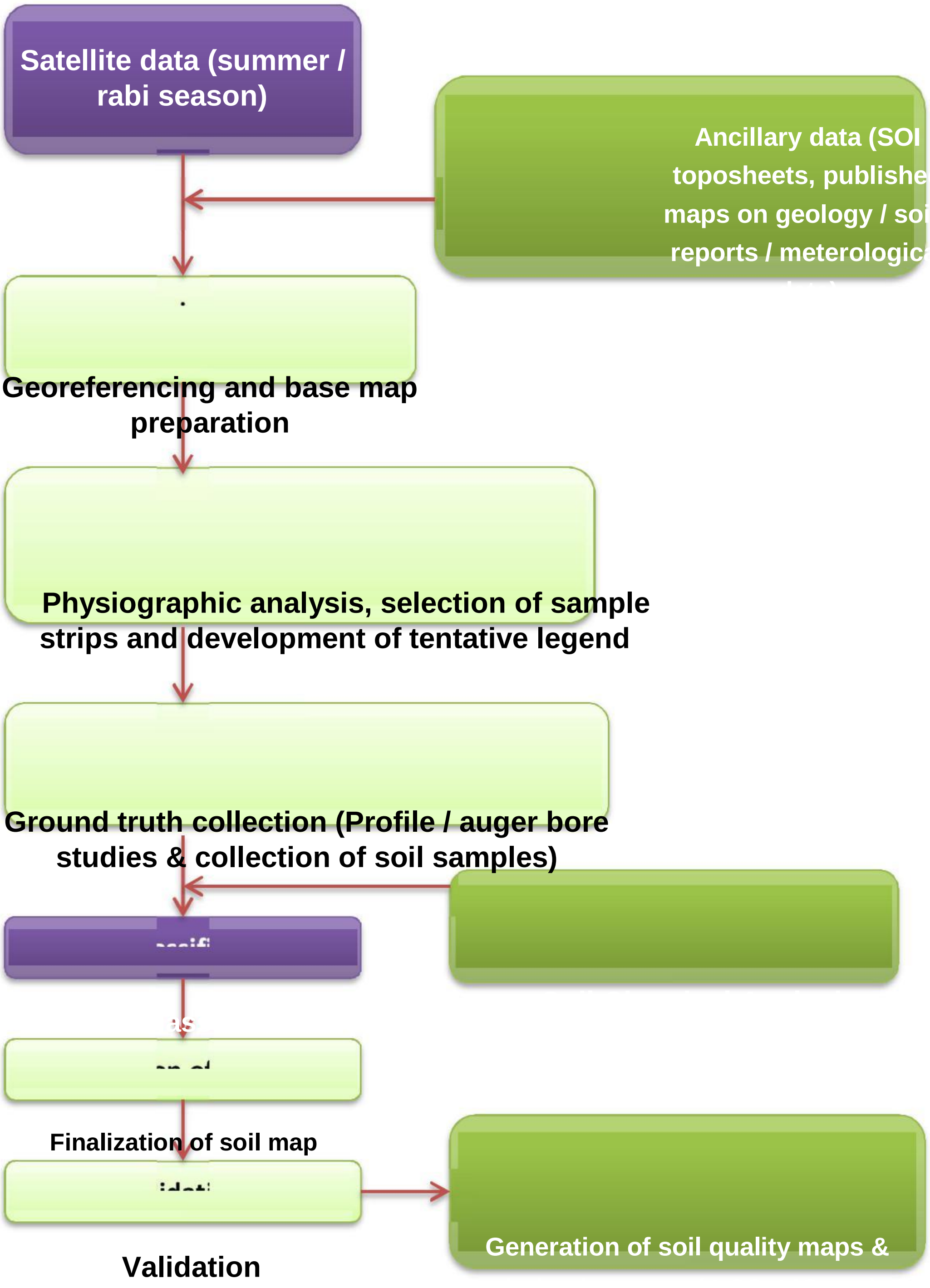


Figure 14.1 Methodology for soil mapping through visual interpretation approach

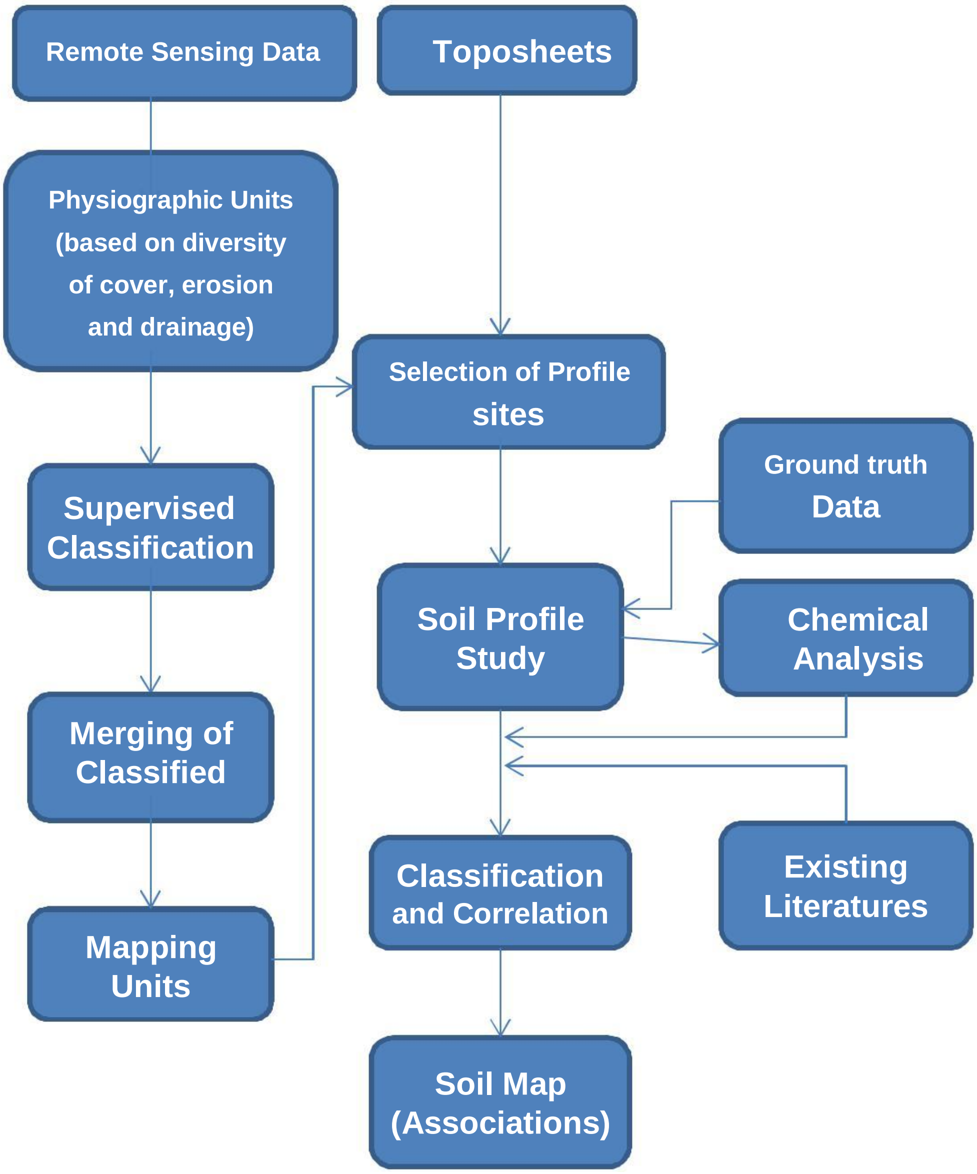


Figure 14.2 Methodology for soil mapping through computer aided approach

4 Fertilizer Recommendation Tool

The database generated from the above soil mapping process can be linked and integrated with a software tool called 'Fertilizer recommendation tool' to generate soil and crop specific fertilizer recommendation based on well-established mathematical functions (Mitscherlich-Bray approach). The major factors to be considered are soil fertility levels (macro nutrients like N, P & K and micro nutrients like zinc, copper, manganese and Iron) and soil problems like sodicity and calcarousness. The parameters viz., doses of nitrogenous, phosphatic and potassium fertilizers based on the type of crop and area of crops to be cultivated.

This information are useful for the farmers, extension workers, crop advisors, and also the researchers. The data base includes soil survey details, soil analytical data and crop suitability and are available for each survey number on the village map.

15 CROP SIMULATION MODELS

Before going on to Crop simulation models, one should understand the concepts of System and Model.

1 Terminologies related to

Model System

A system is a part of reality that contains

interrelated elements Model

A model is a simplified representation of

a system Simulation

The building of mathematical models and the study performance of their behaviour in reference to those of the systems

Crop Simulation Model

A crop simulation model is a simple representation of crop that aims to study crop growth and development and to compute their responses to the environment.

In essence, they are computer programs that mathematically simulate the growth of a crop in relation to its environment. They often operate at time steps one or two orders of magnitude below the duration of the growing season and provide output data to describe attributes of the crop at different points in time on the web. A dynamical systems model is a mathematical description of a system, and mathematical modeling similarly focuses on understanding the system and identifying opportunities for better management.

The main advantages of using crop models are linked with the possibility to overcome the limitations of classic experimental approach (i.e. extrapolating the results in different conditions) and to provide information to the end-users.

Crop model can be used:

- **– at field and regional scales,**
- **– under different weather regimes,**
- **– in different conditions, cultivars, cropping systems, etc.**

2 Why Model?

Used for manipulations and experiments that are impractical, too expensive, too lengthy or impossible (in real-world social and economic systems)

Address dynamic complexity ("emergent properties") of systems in a way that reductionist science may not be able to do

Identify "best management" strategies (through optimization)

Study the long-term effects of options (predictions, projections)

Allow the researcher to control environmental and experimental conditions

Allow hypothetical and exploratory situations to be investigated

Allow insight to be gained into the relative importance of different system elements Assemble and synthesise what is known about particular processes

3 What models can produce?

Models produce Predictions and Understanding

"Predictions"
Point prediction: temperature in Coimbatore tomorrow
Behaviour: trends, patterns in space and time
Differences: system response with/without an intervention

"Understanding"
Best bet: optimised performance of the system (N application rate)
Trade-offs: household income and range condition
Syntheses: what do we know about these processes, and which are still black boxes?

4 Some important crop models

The following Table displays some important crop models:

Table 15.1 Important Crop Models

Name	Crop and goal
APSIM	Modelling framework for a range of crops
GWM	General weed model in row crops
CROPSYST	Wheat & other crops
SIMCOM	Crop (CERES crop modules) & economics
SIMPOTATO	Potato
INFOCROP	Effects of weather, soils, agronomic management and pests on crop growth and yield.
WOFOST	Wheat & maize, Water and nutrient
ORYZA1	Rice, water
SIMCOY	Corn
GRAZPLAN	Pasture, water, lamb
EPIC	Erosion Productivity Impact Calculator

CERES	Series of crop simulation models
DSSAT	Framework of crop simulation models including modules of CERES, CROPGRO and CROPSIM
CANEGRO	Sugarcane, potential & water stress conditions
SWAT	Soil and Water Assessment Tool

5 DSSAT

DSSAT is a software application program that comprises crop simulation models for 28 crops. The program integrates the effects of soil, crop phenotype, weather and management options and allows users to ask "what if" questions and simulate results by conducting experiments on a computer in minutes that would otherwise consume a significant part of an agronomist's career.

DSSAT has been in use for more than 20 years by researchers in over 100 countries

DSSAT is one of the principal products developed by the International Benchmark Sites Network for Agrotechnology Transfer (IBSNAT) project supported by the U.S. Agency for International Development (USAID) from 1983 to 1993. It has subsequently continued to be developed through collaboration among scientists from IFDC, the International Food Policy Research Institute, University of Florida, University of Georgia, University of Guelph, University of Hawaii, USDA's Agricultural Research Service, Universidad Politecnica de Madrid, Washington State University and other scientists associated with the International Consortium for Agricultural Systems Applications (ICASA).

6 ORYZA: A crop growth simulation model for rice

ORYZA2000 is a growth model for lowland rice (Oryza sativa L.) developed by the International Rice Research Institute (IRRI) and Wageningen University.

ORYZA is an ecophysiological model which simulates growth and development of rice including water, C, and N balance (Bouman et al., 2001; IRRI, 2013) in lowland, upland, and aerobic rice ecosystems.

It works in potential, water-limited, nitrogen-limited, and NxW-limited conditions, weather, irrigation, nitrogen fertilizer, general management, variety characteristics, soil properties.

This model has been evaluated extensively in a wide range of environments. The model ORYZA2000 simulates the growth and development of rice under conditions of potential production and nitrogen limitations.

Model ORYZA2000 was sufficiently accurate in the simulation of leaf area index (LAI) and biomass of leaves, panicles, and total above ground biomass yield under nitrogen limit conditions (Tayefe et al., 2013).

Why ORYZA?

It has strong ability on estimating weather constrained rice growth and yield – the potential growth and yield;

It has good ability on estimation of actual growth and yield under water- and/or nitrogen-limited conditions;

It can be used to study rice cropping management on water (irrigation), nitrogen fertilizer, sowing/transplanting date, etc.

It can be used in application-oriented research such as the design of crop ideotypes, the analysis of yield gaps, the optimization of crop management, the ex-ante analysis of the effects of climate change on crop growth, and agroecological zonation;

It was calibrated and validated for 18 popular rice varieties in 15 locations throughout Asia.

7 Benefits of Crop Simulation Models

1. Reduction in time required for experimentation and observation
2. Increased control over environmental variability
3. Provision of safe learning environment
4. Provision of opportunity to undertake undesirable experiments
5. Transferal of expert knowledge and research experience
6. Elucidation of complex plant environment mathematical descriptions
7. Synthesis of fragmented knowledge
8. Integration of different but associated topic areas
9. Focus for peer experience
10. Promotion of heuristic learning
11. Facilitates distance education and education at a distance
12. Gives greater control of learning to the student

8 Limitations of Crop Simulation Models

1. Loss of field and laboratory skills
2. Separation from the subject of study
3. Development of belief that CSMs are reality
4. Frustrating and boring
5. Experimentation and observation outside model range

9 Integration of remote sensing data with crop growth model

The integration of remotely sensed data with a crop growth model can be achieved by using two distinct methods. In the first method, model initialization is done by estimating crop parameters from remote sensing data and using these parameters as a direct input to the growth model (Maas, 1988). Crop parameters successfully used in this method are measures of light interception by the canopy, namely, leaf area index (LAI) and crop canopy cover.

In a second method, a time series of remotely sensed measurements is used to calibrate the crop growth model.

16 DRONES FOR PRECISION AGRICULTURE

1 Introduction

Unmanned Aerial Vehicles (UAVs) or Drones have seen unprecedented levels of growth in military and civilian application domains. When initially introduced during World War I, UAVs were criticized heavily as being unreliable and inaccurate, and only a handful of people recognized at that early stage their potential and (future) impact on changing the battlefield

2 What is an Unmanned Aerial Vehicle (Drone)?

The term unmanned aerial vehicle (also known as a drone) refers to a pilotless aircraft, a flying machine without an on-board human pilot. As such, 'unmanned' refers to total absence of a human who directs and actively pilots the aircraft. Control functions for unmanned aircraft may be either onboard or off-board (remote control).

3 Fixed Wing vs Copter

A fixed- wing UAV (Figure 161) refers to an unmanned airplane that requires a runway to take -off and land, or catapult launching. A helicopter (Figure 16.2) refers to an aircraft that takes off and lands vertically; it is also known as a rotary aircraft with the ability to hover, to fly in very low altitudes, to rotate in the air and move backwards and sideways. It is capable of performing non-aggressive or aggressive flight

Figure 16.1 Fixed wing UAV

Figure 16.2 Quodcopter

4 Terms related to

Drones 4.1 Quadcopter

A four bladed drone, the most common basic type, because that number of blades gives more stability

4.2 Payload

Anything the UAV / drone carry other that required for its flight

like a camera 4.3 Attitude

This is the orientation of UAV, whether its tilting forward or flying upside down. Includes Pitch, roll and Yaw

4.4 Pitch

It represents the orientation of UAVs in association with Roll and Yaw.. Pitch says wheher the UAV is tilted up or down

4.5 Roll

It represents the orientation of UAVs in association with Pitch and Yaw. Roll is when you twist the drone as if you intend to twist it all the way around its control axis.

4.6 Yaw

It represents the orientation of UAVs in association with Pitch and Roll. Yaw is when the drone is turning slightly left or right.

4.7 Gyroscope

It detects the whether the flying is at level

4.8 Gimbal

The type of mount that lets a camera stay steady on a UAV while turning and when in high wind

5 Advantages of Drones

Images of any area can be obtained at any time / season / date.

By virtue of their small size and easy operation, drones are cheaper and more efficient than manned aircrafts or satellite imaging.

They provide cheaper imaging, greater precision and Drone cameras can take centimetre-level images

Earlier detection of problems is possible.

6 What does Drones do?

Drones in agriculture are simply a low-cost aerial camera platform, equipped with an autopilot using GPS and sensors for collecting relevant data, like a regular point-and-shoot camera for visible images. While a regular camera can provide some information about plant growth, coverage and other things, a multi-spectral sensor expands the utility of the

technique and unleashes its full potential. It allows you to see things which you cannot see in the visible spectrum, such as moisture content in the soil, plant health, stress levels and fruits.

The basic principle of NDVI relies on leaves reflecting a lot of light in the near -infrared, in stark contrast to most non-plant objects. Leaves are green in colour due to the presence of a pigment called chlorophyll, which strongly absorbs almost all non-green light from the visible spectrum of sunlight and reflects mostly green light back to our eyes. Live green plants absorb solar radiation in the photosynthetically active region (PAR) and leaf cells re-emit the solar radiation in the near-infrared spectral region. Thus, a healthy plant appears dark in PAR and bright in near infrared.

On the other hand, in an unhealthy or stressed plant, the leaves reflect less near-infrared light even if its emissions in the visible spectrum remain unchanged. Tucker found that combining these two signals can help differentiate plants from non-plants and a healthy plant from a sickly plant. This work gave rise to indices like the NDVI, which is now used to assess plant health.

With the advancement of technology, it is now relatively inexpensive to modify a consumer camera to collect infrared bands and to fly it aboard a small drone. The ground resolution of UAV imagery is more than one thousand times higher as the reflected radiation does not have to travel through the entire atmosphere to be collected, and the incident light is dramatically more varied.

Using near -infrared, you can identify stress in a plant, ten days before it becomes visible to the eye. When a plant goes into stress, it’s either due to a water or fertilizer shortage, or because it’s being attacked by a pest. Photosynthetic activity decreases and that affects the chlorophyll. That’s what the near-infrared sensor can detect, but our human eye can’t see it until it’s more advanced.

7 How does Drones work?

To survey crop fields with a drone, you start by planning the flight path of the drone that will best cover the plot. Many of the latest agricultural drones come with flight-planning software that let you outline a box around the field you want to survey on Google Maps. The flight plan is then automatically computed. The drone then flies over the field in a pattern while taking pictures with one or more cameras with special light sensors. These pictures are geo-tagged and overlap each other.

After landing, special software is used to stitch together the geotagged photos into a large mosaic and processed to interpret the amount of light that is reflected in different wavelengths. Processing this data makes areas of poor growth or stressed plants easy to identify. Generating this data immediately and quickly opens to doors to better interventions and decision-making. The last step in this process is reviewing and taking remedial action. Prescriptive software packages also come up with comprehensive recommendations based on the field survey but these are not completely reliable yet.

Processing images is the most challenging part of any drone-based agricultural operation. Here, downstream software packages use the high-resolution images and data from the different sensors on a drone to generate a meaningful and insightful image. Most agricultural drone operators use a tool like Pix4D or Correlator3D to turn these aerial images into useful data. Some others use proprietary software packages custom-built for their devices. The challenge however is that, being computation intensive, they rely on cloud services that are not always available in agricultural areas, especially in the context of emerging economies like India.

While some other general-purpose programs (like Pix4D) are available, they involve a steep learning curve. In developed economies of the west, companies also provide only a back-end processing service that will analyse the data once uploaded and generate reports for you. Other smaller start-ups have developed their own sensors, software packages and extensive back-end analytics support that can be used with any UAV.

Some of this software is also built to avoid reliance on cloud-computing; all analytics can be performed on a local computer once downloaded and set up. Some companies, like Slantrange, have also developed core intellectual property enabling the identification of weeds from crop plants using multi-parametric image analysis. This further extends the capabilities of drone-based agriculture systems.

8 Challenges of Drone usage

Outdoor use is highly weather dependent

Imaging can vary depending on sunlight and cloud cover although one can account for ambient lighting conditions

Limited internet access and cellular infrastructure can make it harder to rely on cloud-based computing services

Higher costs especially for small landholders in

emerging economies Limited flight times

Maintenance costs and resources

The need for skilled operators

Uncertain government regulation that need to be overcome before this technology can be widely applied.

9 Application of Drones in Agriculture

The following are some of the applications of Drones in agriculture and in its allied fields:

Water stress detection.

Estimation of nitrogen

level. Pathogen detection.

Aerobiological sampling.

Plant health monitoring.

Mapping invasive weeds.

Monitoring herbicide applications

Forest fire monitoring

Monitoring biodiversity in forests

Assessing erosion in agricultural fields

AUTHORS

Pazhanivelan,S. Professor and Head, Department of Remote Sensing and GIS, TNAU, pazhavelans@gmail.com

Sivasamy, R, Professor (SS&AC), Dept of Remote Sensing and GIS, TNAU, Coimbatore 641003, siva_tnau@yahoo.com

Balaji Kannan, Associate Professor (SWC&E), Dept. of Remote Sensing and GIS, TNAU, Coimbatore 641003, balajikannan73@gmail.com

Jagadeeswaran, R, Assistant Professor (SS&AC), Dept. of Remote Sensing and GIS, TNAU, Coimbatore 641003, jagawaran@yahoo.com / jagawaran@tnau.ac.in

Kumaraperumal, R, Assistant Professor (SS&AC), Department of Remote Sensing and GIS, TNAU, kumaraperumal.r@gmail.com

Ragunath, K.P, Assistant Professor (SS&AC), Department of Remote Sensing and GIS, TNAU, ragunathkp@gmail.com

APPENDIX

Appendix 1 Selected list of sensors and their missions

Name	Description	No of Bands	Mission
ASTER	Advanced Spaceborne Thermal Emission and Reflection Radiometer	14	Terra
AVHRR/2	Advanced Very High Resolution Radiometer	5	NOAA-10, NOAA-11, NOAA-12, NOAA-14
AVHRR/3	Advanced Very High Resolution Radiometer	6	Metop-A, Metop-B, Metop-C, NOAA-15, NOAA-16, NOAA-17, NOAA-18, NOAA-19
AWiFS	Advanced Wide Field Sensor	4	RESOURCESAT-1 (IRS-P6), RESOURCESAT-2, RESOURCESAT-2A
Kompsat-MSC	B&W panchromatic, MSS and merged 1m resolution images	4	KOMPSAT-2
LISS 1	Linear Imagine Self Scanning System	4	IRS-1A, IRS-1B
LISS 2	Linear Imagine Self Scanning System	4	IRS-1A, IRS-1B, IRS-P2
LISS 3 (IRS1C/1D)	Linear Imagine Self Scanning System	4	IRS-1C, IRS-1D
LISS-3 (RESOURCESAT)	Linear Imagine Self Scanning System	4	RESOURCESAT-1 (IRS-P6), RESOURCESAT-2, RESOURCESAT-2A
LISS-4	Linear Imagine Self Scanning System	3	RESOURCESAT-1 (IRS-P6), RESOURCESAT-2, RESOURCESAT-2A
MSS (LS 1-3)	Multispectral Scanner - Landsat 1,2,3	5	Landsat 1, Landsat 2, Landsat 3

Name	Description	No of Bands	Mission
MSS (LS 4-5)	Multispectral Scanner - Landsat 4,5	4	Landsat 4, Landsat 5
OCM	Ocean Color Monitor	8	OCEANSAT-1, OCEANSAT-2
OCO-2	Spectrometers	3	OCO-2
OLI	Operational Land Imager	9	Landsat 8
OLI-2	Operational Land Imager 2	9	Landsat 9
OSA	Optical Sensor Assembly	5	IKONOS, IKONOS-1
PALSAR	L-band Synthetic Aperture Radar	1	ALOS
PALSAR-2	L-band Synthetic Aperture Radar-2	1	ALOS-2
PALSAR-3	L-band Synthetic Aperture Radar-3	1	ALOS-4
PAN (Cartosat-2 Series)	PAN	1	CartoSat-2 Series (2C), CartoSat-2 Series (2D), Cartosat-2 Series (2E), Cartosat-2 Series (2F
PAN (IRS)	Panchromatic camera	1	IRS-1C, IRS-1D
PAN C (Cartosat)	Panchromatic camera	1	CartoSat-2A
PAN Cartosat-2B	Panchromatic camera	1	CartoSat-2B
PAN-Aft	Panchromatic-Aft pointing	1	CartoSat-1 (IRS-P5)
PAN-Fore	Panchromatic-Forward pointing	1	CartoSat-1 (IRS-P5)
PR	Precipitation Radar	2	TRMM
Quickbird	High resolution: Pan: 61 cm (nadir) to 72 cm (25° off-nadir), MS: 2.44 m to 2.88 m	4	Quickbird
RADARSAT 2	Radar		RADARSAT-2
RBI	Radiation Budget Instrument (cancelled Jan 2018)	3	JPSS-2
SAR 2000	RADAR	1	COSMO-SkyMed
SAR-C Radarsat1	Synthetic Aperture Radar on RADARSAT-1	1	RADARSAT-1

Name	Description	No of Bands	Mission
SAR-C Sentinel1	C-band SAR on Sentinel-1A/Sentinel-1B	1	Sentinel-1A, Sentinel-1B
SIR-C	Spaceborne Imaging Radar-C	1	SRTM
SpaceView 110 Imaging System	formerly GIS-2, GeoEye Imaging System-2	5	WorldView-4
TIRS	Thermal Infrared Sensor	2	Landsat 8
TM	Thematic Mapper	7	Landsat 4, Landsat 5
WiFS	Wide Field Sensor	3	IRS-1C, IRS-1D, IRS-P3

Source: For more details on characteristics of each sensor check the following link - https://www.itc.nl/Pub/sensordb/AllSensors.aspx

Appendix 2 List of Indian Earth Observation Satellites

Satellite	Launch Date	Launch Mass	Launch Vehicle	Orbit Type	Application
Cartosat-2 Series Satellite	Jan 12, 2018	710 Kg	PSLV-C40/Cartosat-2 Series Satellite Mission	SSPO	Earth Observation
Cartosat-2 Series Satellite	Jun 23, 2017	712 kg	PSLV-C38 / Cartosat-2 Series Satellite	SSPO	Earth Observation
Cartosat -2 Series Satellite	Feb 15, 2017	714 kg	PSLV-C37 / Cartosat -2 Series Satellite	SSPO	Earth Observation
RESOURCESAT-2A	Dec 07, 2016	1235 kg	PSLV-C36 / RESOURCESAT-2A	SSPO	Earth Observation
SCATSAT-1	Sep 26, 2016	371 kg	PSLV-C35 / SCATSAT-1	SSPO	Climate & Environment
INSAT-3DR	Sep 08, 2016	2211 kg	GSLV-F05 / INSAT-3DR	GSO	Climate & Environment, Disaster Management System
CARTOSAT-2 Series Satellite	Jun 22, 2016	737.5 kg	PSLV-C34 / CARTOSAT-2 Series Satellite	SSPO	Earth Observation
INSAT-3D	Jul 26, 2013	2060 Kg	Ariane-5 VA-214	GSO	Climate & Environment, Disaster Management System
SARAL	Feb 25, 2013	407 kg	PSLV-C20/SARAL	SSPO	Climate & Environment, Earth Observation
RISAT-1	Apr 26, 2012	1858 kg	PSLV-C19/RISAT-1	SSPO	Earth Observation
Megha-Tropiques	Oct 12, 2011	1000 kg	PSLV-C18/Megha-Tropiques	SSPO	Climate & Environment, Earth Observation
RESOURCESAT-2	Apr 20, 2011	1206 kg	PSLV-C16/RESOURCESAT-2	SSPO	Earth Observation
CARTOSAT-2B	Jul 12, 2010	694 kg	PSLV-C15/CARTOSAT-2B	SSPO	Earth Observation
Oceansat-2	Sep 23, 2009	960 kg	PSLV-C14 / OCEANSAT-2	SSPO	Climate & Environment, Earth Observation
RISAT-2	Apr 20, 2009	300 kg	PSLV-C12 / RISAT-2	SSPO	Earth Observation
CARTOSAT – 2A	Apr 28, 2008	690 Kg	PSLV-C9 / CARTOSAT – 2A	SSPO	Earth Observation

Satellite	Launch Date	Launch Mass	Launch Vehicle	Orbit Type	Application
IMS-1	Apr 28, 2008	83 kg	PSLV-C9 / CARTOSAT – 2A	SSPO	Earth Observation
CARTOSAT-2	Jan 10, 2007	650 kg	PSLV-C7 / CARTOSAT-2 / SRE-1	SSPO	Earth Observation
CARTOSAT-1	May 05, 2005	1560 kg	PSLV-C6/CARTOSAT-1/HAMSAT	SSPO	Earth Observation
IRS-P6 / RESOURCESAT-1	Oct 17, 2003	1360 kg	PSLV-C5 /RESOURCESAT-1	SSPO	Earth Observation
The Technology Experiment Satellite (TES)	Oct 22, 2001		PSLV-C3 / TES	SSPO	Earth Observation
Oceansat(IRS-P4)	May 26, 1999	1050 kg	PSLV-C2/IRS-P4	SSPO	Earth Observation
IRS-1D	Sep 29, 1997	1250kg	PSLV-C1 / IRS-1D	SSPO	Earth Observation
IRS-P3	Mar 21, 1996	920 kg	PSLV-D3 / IRS-P3	SSPO	Earth Observation
IRS-1C	Dec 28, 1995	1250 kg	Molniya	SSPO	Earth Observation
IRS-P2	Oct 15, 1994	804 kg	PSLV-D2	SSPO	Earth Observation
IRS-1E	Sep 20, 1993	846 kg	PSLV-D1	LEO	Earth Observation
IRS-1B	Aug 29, 1991	975 kg	Vostok	SSPO	Earth Observation
SROSS-2	Jul 13, 1988	150 kg	ASLV-D2		Earth Observation, Experimental
IRS-1A	Mar 17, 1988	975 kg	Vostok	SSPO	Earth Observation
Rohini Satellite RS-D2	Apr 17, 1983	41.5 kg	SLV-3	LEO	Earth Observation
Bhaskara-II	Nov 20, 1981	444 kg	C-1 Intercosmos	LEO	Earth Observation, Experimental
Rohini Satellite RS-D1	May 31, 1981	38 kg	SLV-3D1	LEO	Earth Observation
Bhaskara-I	Jun 07, 1979	442 kg	C-1 Intercosmos	LEO	Earth Observation, Experimental

LEO stands for Low Earth Orbit; SSPO Stands for Sun Synchronous Polar Orbit; GSO – Geo Synchronous Orbit

Source: For more details on characteristics of each sensor check the following link - https://www.isro.gov.in/spacecraft/list-of-earth-observation-satellites

INDEX