1 CONCEPTS AND FOUNDATIONS OF REMOTE SENSING

1.1 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. As you read these words, you are employing remote sensing. Your eyes are acting as sensors that respond to the light reflected from this page. The "data" your eyes acquire are impulses corresponding to the amount of light reflected from the dark and light areas on the page. These data are analyzed, or interpreted, in your mental computer to enable you to explain the dark areas on the page as a collection of letters forming words. Beyond this, you recognize that the words form sentences and you interpret the information that the sentences convey.

In many respects, remote sensing can be thought of as a reading process. Using various sensors, we remotely collect *data* that may be analyzed to obtain *information* about the objects, areas, or phenomena being investigated. The remotely collected data can be of many forms, including variations in force distributions, acoustic wave distributions, or electromagnetic energy distributions. For example, a gravity meter acquires data on variations in the distribution of the force of gravity. Sonar, like a bat's navigation system, obtains data on variations in acoustic wave distributions. Our eyes acquire data on variations in electromagnetic energy distributions.

This book is about *electromagnetic* energy sensors that are currently being operated from airborne and spaceborne platforms to assist in inventorying, mapping, and monitoring earth resources. These sensors acquire data on the way various earth surface features emit and reflect electromagnetic energy, and these data are analyzed to provide information about the resources under investigation.

Figure 1.1 schematically illustrates the generalized processes and elements involved in electromagnetic remote sensing of earth resources. The two basic processes involved are *data acquisition* and *data analysis*. The elements of the data acquisition process are energy sources (*a*), propagation of energy through the atmosphere (*b*), energy interactions with earth surface features (*c*), retransmission of energy through the atmosphere (*d*), airborne and/or spaceborne sensors (*e*), resulting in the generation of sensor data in pictorial and/or digital form (*f*). In short, we use sensors to record variations in the way earth surface features reflect and emit electromagnetic energy. The data analysis process (*g*) involves 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. This information is then compiled (*h*), generally in the form of hardcopy maps and tables or as computer files that can be merged with other "layers" of information in a *geographic information system* (*GIS*). Finally, the information is presented to users (*i*), who apply it to their decision-making process.

In the remainder of this chapter, we discuss the basic principles underlying the remote sensing process. We begin with the fundamentals of electromagnetic energy and then consider how the energy interacts with the atmosphere and with earth surface features. Next, we summarize the process of acquiring and interpreting imagery in both analog and digital formats. We also discuss the role that reference data play in the data analysis procedure and describe how the spatial location of reference data observed in the field is often determined using *Global Positioning System* (*GPS*) methods. These basics will permit us to conceptualize the strengths and limitations of "real" remote sensing systems and to examine the ways in which they depart from an "ideal" remote sensing system. We also discuss briefly the rudiments of GIS technology. At the end of this chapter, the reader should have a grasp of the general concepts and foundations of remote sensing and an appreciation for the close relationship among remote sensing, GPS methods, and GIS operations.

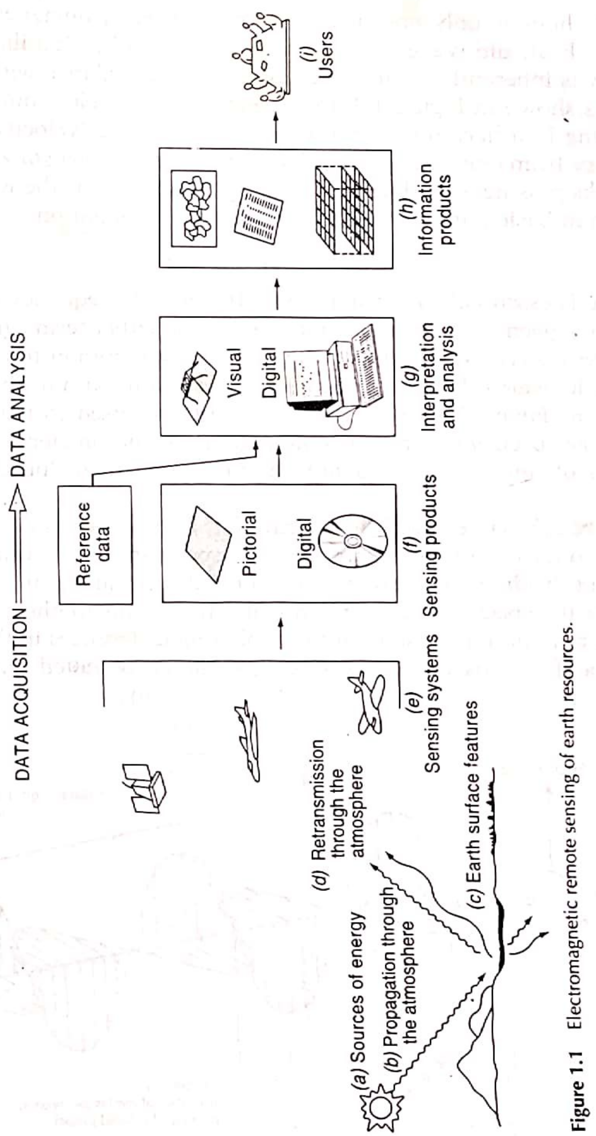

Figure 1.1 Electromagnetic remote sensing of earth resources.

1.2 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. As shown in Figure 1.2, 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.

From basic physics, waves obey the general equation

$$c = v\lambda \tag{1.1}$$

Since c is essentially a constant (3×10^8 m/sec), frequency v and wavelength λ for any given wave are related inversely, and either term can be used to characterize a wave. 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* (μm). A micrometer equals 1×10^{-6} m. (Tables of units used frequently in this book are included inside the back cover.)

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 from 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

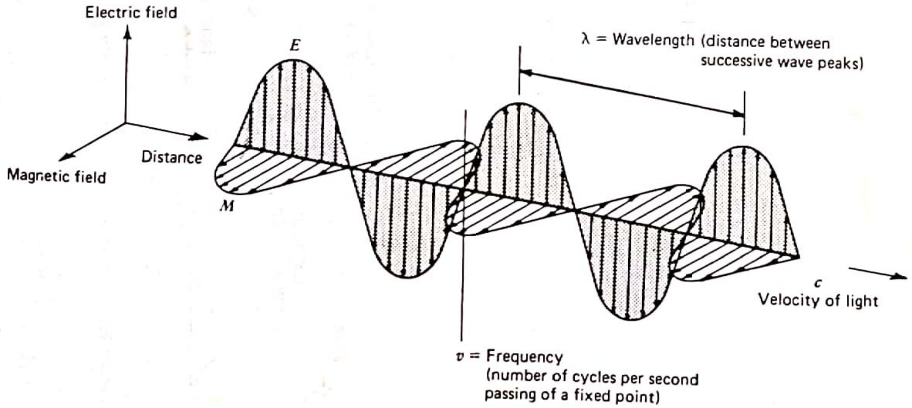

Figure 1.2 Electromagnetic wave. Components include a sinusoidal electric wave (*E*) and a similar magnetic wave (*M*) at right angles, both being perpendicular to the direction of propagation.

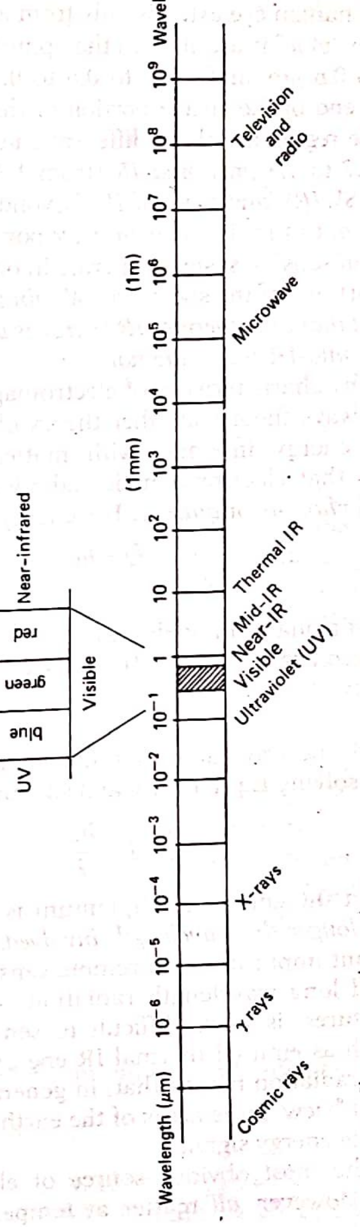

Figure 1.3 Electromagnetic spectrum.

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; also referred to as *shortwave IR* or *SWIR*), and *thermal IR* (beyond 3 to 14 μm). At much longer wavelengths (1 mm to 1 m) is the *microwave* portion of the 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 are 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

$$Q = h\nu \tag{1.2}$$

where

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

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

$$Q = \frac{hc}{\lambda} \tag{1.3}$$

Thus, we see that the energy of a quantum is inversely proportional to its wavelength. *The longer the wavelength involved, the lower its energy content.* This has important implications in remote sensing from the standpoint that naturally emitted long wavelength radiation, such as microwave emission from terrain features, is more difficult to sense than radiation of shorter wavelengths, such as emitted thermal IR energy. The low energy content of long wavelength radiation means that, in general, systems operating at long wavelengths must "view" large areas of the earth at any given time in order to obtain a detectable energy signal.

The sun is the most obvious source of electromagnetic radiation for remote sensing. However, *all* matter at temperatures above absolute zero (0 K, or −273°C) 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 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

$$M = \sigma T^4 \tag{1.4}$$

where

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

The particular units and the value of the constant are not critical for the student to remember, yet 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. Also, it should be noted that this law is expressed for an energy source that behaves as a *blackbody*. A blackbody is a hypothetical, ideal radiator that totally absorbs and reemits all energy incident upon it. Actual objects only approach this ideal. We further explore the implications of this fact in Chapter 5; suffice it to say for now that the energy emitted from an object is primarily a function of its temperature, as given by Eq. 1.4.

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 $m^{-2}\mu m^{-1}$) express the radiant power coming from a blackbody per 1-μm 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. The *dominant wavelength*, or wavelength at which a blackbody radiation curve reaches a maximum, is related to its temperature by *Wien's displacement law*,

$$\lambda_m = \frac{A}{T} \tag{1.5}$$

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 process from start to finish, users of remote sensing systems need to keep in mind the following factors:

1. **The energy source.** All passive remote sensing systems rely on energy that originates from sources other than the sensor itself, typically in the form of either reflected radiation from the sun or emitted radiation from earth surface features. As already discussed, the spectral distribution of reflected sunlight and self-emitted energy is far from uniform. Solar energy levels obviously vary with respect to time and location, and different earth surface materials emit energy with varying degrees of efficiency. While we have some control over the sources of energy for active systems such as radar and lidar, those sources have their own particular characteristics and limitations, as discussed in Chapter 8. Whether employing a passive or active system, the remote sensing analyst needs to keep in mind the nonuniformity and other characteristics of the energy source that provides illumination for the sensor.
2. **The atmosphere.** The atmosphere normally compounds the problems introduced by energy source variation. To some extent, the atmosphere always modifies the strength and spectral distribution of the energy received by a sensor. It restricts "where we can look" spectrally, and its effects vary with wavelength, time, and place. The importance of these effects, like source variation effects, is a function of the wavelengths involved, the sensor used, and the sensing application at hand. Elimination of, or compensation for, atmospheric effects via some form of calibration is particularly important in those applications where repetitive observations of the same geographic area are involved.
3. **The energy–matter interactions at the earth's surface.** Remote sensing would be simple if every material reflected and/or emitted energy in a unique, known way. Although spectral response patterns such as those in Figure 1.9 play a central role in detecting, identifying, and analyzing earth surface materials, the spectral world is full of ambiguity. Radically different material types can have great spectral similiarity, making differentiation difficult. Furthermore, the general understanding of the energy–matter interactions for earth surface features is at an elementary level for some materials and virtually nonexistent for others.
4. **The sensor.** An ideal sensor would be highly sensitive to all wavelengths, yielding spatially detailed data on the absolute brightness (or radiance) from a scene as a function of wavelength, throughout the spectrum, across wide areas on the ground. This "supersensor" would be simple and reliable, require virtually no power or space, be available whenever and wherever needed, and be accurate and economical to operate. At this point, it should come as no surprise that an ideal "supersensor" does not exist. No single sensor is sensitive to all wavelengths. All real sensors have fixed limits of spectral sensitivity and spatial resolution.

   The choice of a sensor for any given task always involves trade-offs. For example, photographic systems generally have very good spatial resolution characteristics, but they lack the broad spectral sensitivity obtainable with nonphotographic systems, which usually have poorer spatial resolution characteristics. Similarly, many nonphotographic systems (and some photographic systems) are quite complex optically, mechanically, and/or electronically. They may have restrictive power, space, and stability requirements. These requirements often dictate the type of *platform*, or vehicle, from which a sensor can be operated. Platforms can vary from stepladders to helicopters, conventional aircraft, uninhabited aerial vehicles (UAVs), and satellites. Depending on the sensor–platform combination needed in a particular application, the acquisition of remote sensing data can be a very expensive endeavor, and there may be limitations on the times and places that data can be collected. Airborne systems require detailed flight planning in advance, while data collection from satellites is limited by the platform's orbit characteristics.
5. **The data processing and supply system.** The capability of current remote sensors to generate data far exceeds the capacity to handle these data. This is generally true whether we consider "manual" image interpretation procedures or digital analyses. Processing sensor data into an interpretable format can be—and often is—an effort entailing considerable thought, hardware, time, and experience. Also, many data users would like to receive their data immediately after acquisition by the sensor in order to make the timely decisions required in certain applications (e.g., agricultural crop management, disaster assessment). Fortunately, the distribution of remote sensing imagery has improved dramatically over the past two decades. Some sources now provide data within 24 hours of the image acquisition, with data downloaded over the Internet. In other cases, however—particularly for highly specialized types of imagery, or for experimental or newly developed remote sensing systems—it may take weeks or months before data are made available. Finally, as discussed in Section 1.6, most remote sensing applications require the collection and analysis of additional reference data, an operation that may be complex, expensive, and time consuming.
6. **The users of remotely sensed data.** Central to the successful application of any remote sensing system is the person (or persons) using the remote sensor data from that system. The "data" generated by remote sensing procedures become "information" only if and

when someone understands their generation, knows how to interpret them, and knows how best to use them. *A thorough understanding of the problem at hand is paramount to the productive application of any remote sensing methodology. Also, no single combination of data acquisition and analysis procedures will satisfy the needs of all data users.*

Whereas the interpretation of aerial photography has been used as a practical resource management tool for nearly a century, other forms of remote sensing are relatively new, technical, and "unconventional" means of acquiring information. These newer forms of remote sensing have had relatively few satisfied users until recently. However, as new applications continue to be developed and implemented, increasing numbers of users are becoming aware of the potentials, as well as the limitations, of remote sensing techniques. As a result, remote sensing has become an essential tool in many aspects of science, government, and business alike.

1.9 SUCCESSFUL APPLICATION OF REMOTE SENSING

The student should now begin to appreciate that successful use of remote sensing is premised on the *integration* of multiple, interrelated data sources and analysis procedures. No single combination of sensor and interpretation procedure is appropriate to all resource inventorying and environmental monitoring applications. In fact, *many inventorying and monitoring problems are not amenable to solution by means of remote sensing at all.* Among the applications that are appropriate, a wide variety of data acquisition and analysis approaches exist. Conceptually, however, all designs of successful remote sensing efforts involve, at a minimum, (1) clear definition of the problem at hand, (2) evaluation of the potential for addressing the problem with remote sensing techniques, (3) identification of the remote sensing data acquisition procedures appropriate to the task, (4) determination of the data interpretation procedures to be employed and the reference data needed, and (5) identification of the criteria by which the quality of information collected can be judged.

All too often, one (or more) of the above components of a remote sensing application is overlooked. The result may be disastrous. Many resource management programs exist with little or no means of evaluating the performance of remote sensing systems in terms of information quality. Many people have acquired burgeoning quantities of remote sensing data with inadequate capability to interpret them. Many occasions have occurred when remote sensing has *or* has not been used because the problem was not clearly defined. A clear articulation of the information requirements of a particular

problem and the extent to which remote sensing might meet these requirements is paramount to any successful application.

The success of many applications of remote sensing is improved considerably by taking a *multiple-view* approach to data collection. This may involve *multistage* sensing, wherein data about a site are collected from multiple altitudes. It may involve *multispectral* sensing, whereby data are acquired simultaneously in several spectral bands. Or, it may entail *multitemporal* sensing, where data about a site are collected on more than one occasion.

In the multistage approach, satellite data may be analyzed in conjunction with high altitude data, low altitude data, and ground observations (Figure 1.24). Each successive data source might provide more detailed information over smaller geographic areas. Information extracted at any lower level of observation may then be extrapolated to higher levels of observation.

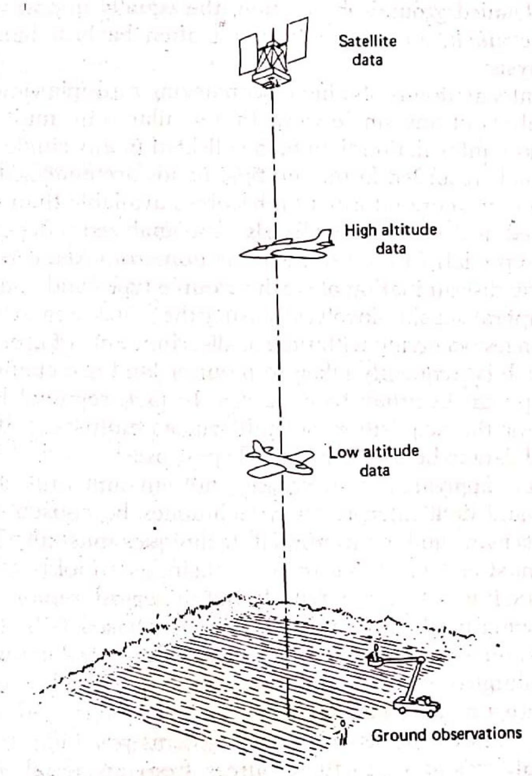

Figure 1.24 Multistage remote sensing concept.