1 Remote sensing and the cryosphere

Marco Tedesco

The City College of New York, City University of New York, New York, USA

Summary

This introductory first chapter provides a general overview on remote sensing and an introduction to the cryosphere, exposing the reader to general concepts. The chapter is mainly oriented toward those readers with minimal or no experience on either of the two subjects. Remote sensing can be defined as to that ensemble of techniques, tools, data and sensors that allow us to study the Earth and its processes from airborne, spaceborne and *in situ* sensors without being in physical contact with the object under examination.

In the first part of this chapter, a brief history of remote sensing is introduced, describing early tools and applications (such as the pioneering work from air balloons and cameras attached to pigeons), followed by a basic introduction to the electromagnetic spectrum and electromagnetic radiation. The reader is then presented with a description of remote sensing systems, divided into the categories of *passive* (aerial photography, electro-optical sensors, thermal systems, microwave radiometers and gravimetric systems) and *active systems* (LiDAR, radar). Concepts such as spatial, temporal, spectral and radiometric resolutions are also introduced. In the second part of the chapter, the several elements of the cryosphere are introduced, together with a description of their basic physical properties and a general overview of their spatial distribution and the impact on other fields (such as biology, ecology, etc.).

1.1 Introduction

This chapter contains a general overview on both remote sensing and the cryosphere and briefly introduces the reader to their general concepts. Both topics are vast, and it is not possible to cover them in their entirety here. Nevertheless, it is helpful to provide an introductory overview of the two fields, with the references in this chapter (and throughout the book) suggesting reading material for those interested in more details.

1.2   Remote sensing

Remote sensing is the collection of information about an object or phenomenon without physical contact with the object. For practical applications, throughout this book we will refer to remote sensing as that ensemble of techniques, tools, data and sensors that allow us to study the Earth and its processes from airborne, spaceborne and *in situ* sensors without being in physical contact with the object under examination.

Remote sensing of the Earth began with the development of flight. The first photographs of Paris were taken from air balloons as early as 1858 by Gaspard-Félix Tournachon, a French photographer known as Nadar (http://www.papainternational.org/history.asp). In the 1880s, Arthur Batut attached cameras to kites to collect pictures over Labruguière, France. The apparatus also included an altimeter so that the scale of the images could be estimated. At the beginning of 1900, the Bavarian Pigeon Corps had cameras attached to pigeons, taking pictures every 30 seconds (http://www.sarracenia.com/astronomy/remotesensing/primer0120.html; Jensen, 2006).

Systematic aerial photography began with World War I and was improved during World War II. At the end of the Wars, the development of artificial satellites allowed remote sensing to begin performing measurements on a large scale, leading to the modern remote sensing era. More information on the history of remote sensing can be found, for example, in Jensen (2006).

1.2.1   The electromagnetic spectrum and blackbody radiation

Remote sensing of the Earth is based on the interaction between electromagnetic waves and matter, with the exception of those approaches based on gravimetry. The interaction between materials and electromagnetic waves depends on both the characteristics of the electromagnetic radiation (e.g., frequency) and on the chemical and physical properties of the targets. In many cases, the source of the electromagnetic radiation is the sun, which can be approximated as a black body (an idealized body that absorbs all incident electromagnetic radiation, regardless of frequency) at a temperature of about 5800 K. Though a large number of remote sensing applications deal with the visible portion (400 – 700 nm) of the electromagnetic spectrum (Figure 1.1), visible light occupies only a fraction of it. Indeed, a considerable portion of the incoming solar radiation is in form of ultraviolet and infrared radiation, and only a small portion is in form of microwave radiation.

Before reaching a spaceborne or airborne sensor, the electromagnetic radiation propagates through the atmosphere, hence interacting with the different atmospheric components. For example, as the sunlight enters the atmosphere, it interacts with gas molecules, suspended particles and aerosols. Because of the preferential scattering and absorption of particular wavelengths and elements, the radiation reaching the Earth is a combination of direct filtered solar radiation and diffused light scattered from the sky.

Figure 1.1 Spectral regions used for thermal and passive microwave sensing (Adapted from Lillesand *et al.*, 2007).
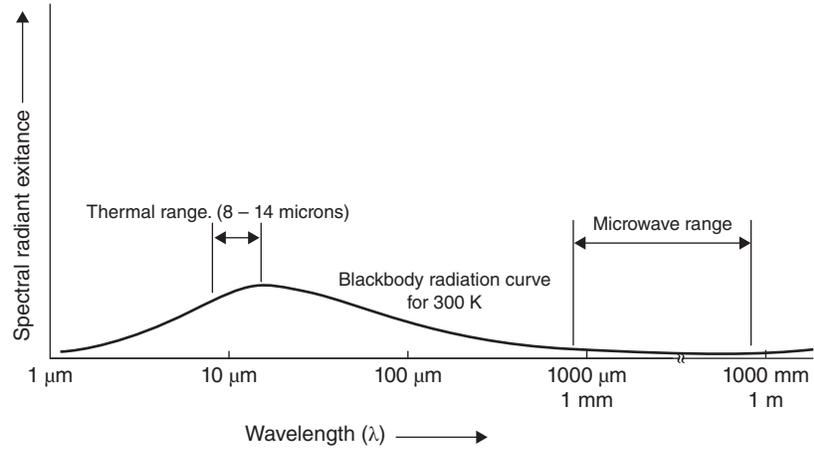
As the filtered and diffused solar radiation reaches the Earth, it interacts with surface targets (e.g., soil, snow, vegetation, ocean, etc.). Each of these materials interacts with the electromagnetic radiation through absorption, transmission and scattering, depending on its physical properties (e.g., leaves reflects back most of the radiation in the green regions, water reflects blue radiation more than red and green, etc.). Before reaching the sensors and being recorded by the instruments onboard, the upwelling radiation passes again through the atmosphere.

It follows that the atmospheric components (such as water vapor, carbon dioxide and ozone, Figure 1.2) drive the design of the sensors used to study the Earth from space. On the other hand, in the case of thermal infrared radiation (e.g., 800–1400 nm), the sensors will detect the radiation emitted by the surface as a result of the solar heating. In the case of passive microwave remote sensing, the instruments will record the naturally emitted radiation by the objects, because the incoming solar radiation in the microwave region is negligible.

Features characterizing the data collected by remote sensing platforms are spatial, temporal, spectral and radiometric resolutions. Spatial resolution is a measure of the spatial detail that can be distinguished in an image, being, in turn, a function of the sensor design and its altitude above the surface. Temporal resolution is the frequency of data acquisition (e.g., how many acquisitions are collected within a day) or the temporal interval separating successive data acquisitions. Spectral resolution is the ability of the system to distinguish different parts of the range of measured wavelengths (e.g., the number of measured bands and how narrow each band is).

For recording purposes, the energy received by an individual detector in a sensor must be "quantized" (e.g., divided into a number of discrete levels that are recorded as integer values). Radiometric resolution quantifies the number of levels: the more levels that are recorded, the greater is the radiometric resolution. Many current satellite systems quantize data into 256 levels (8 bits of data in a binary encoding system, $2^8 = 256$), but other systems can have higher radiometric resolution (e.g., 12 bits, $2^{12} = 4096$ levels).
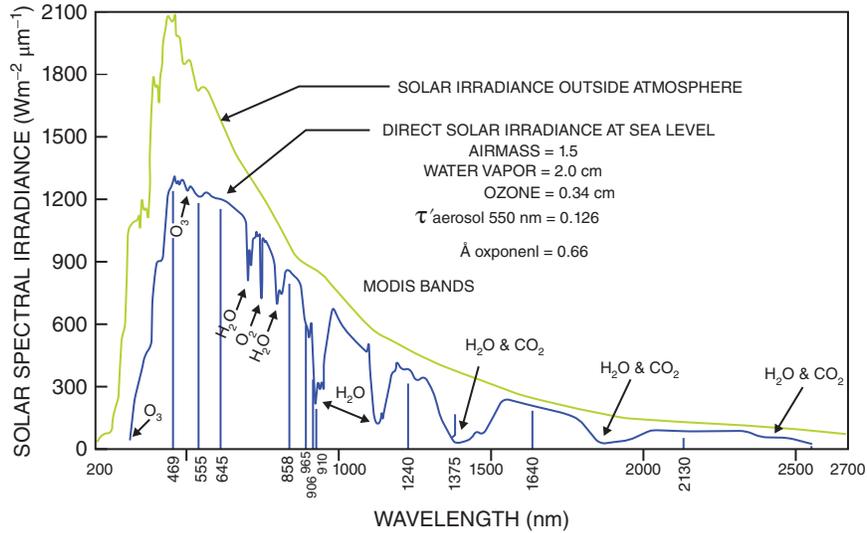
Figure 1.2 Solar spectral irradiance incident on the top of the atmosphere and transmitted through the atmosphere to the Earth's surface. Major absorption bands in the atmosphere are also shown (NASA).

Remote sensing instruments can be categorized into active and passive. Passive sensors measure the radiation that is naturally emitted or reflected by the target. For example, sensors operating in the visible range, measuring the solar radiation reflected by an object on Earth, are passive sensors. Other passive sensors are microwave radiometers, measuring the microwave radiation naturally emitted by the targets (in this case, as mentioned, the solar contribution is negligible). In this book, sensors collecting gravimetric data are also classified as passive. Active sensors emit energy and measure the amount of energy that is reflected or backscattered by the target. Examples of active sensors are radar (Radio Detection and Ranging) or LiDAR (Light Detection and Ranging).

1.2.2   Passive systems

1.2.2.1   Aerial photography

The first collection of aerial photography was performed by the French photographer and balloonist Gaspard-Félix Tournachon, known as Nadar, in 1858 over Paris (though it was destroyed, and the first surviving collection of aerial photography consists of a view of Boston taken from 630 m by James Wallace Black and Samuel Archer King in 1860). Aerial photography was crucial during World War I, supporting many strategic decisions for battlefronts, for example. After the wars, aerial photography became more and more accessible and was used by commercial companies and government agencies for many purposes. More information on the history of aerial photography can be found

at http://www.papainternational.org/history.asp (History of Aerial Photography, Professional Aerial Photographers Association, accessed December 28, 2012). Aerial photography is nowadays used for cartography, land use planning and management, archeology and environmental studies, to name just a few applications.

Traditional photographic systems make use of photographic films to collect information in the visible and near-infrared regions. Obviously, intrinsic limitations exist due to the need of solar illumination and the impossibility of collecting data in the presence of clouds. The film consists of a suspension of small crystals of silver halides in a gelatin matrix, and its exposure to light converts the silver ions into thermodynamically stable metallic silver. The development process removes the crystals that have not been exposed, generating the so-called *negative,* which is then used to print the photograph. Panchromatic films have a relatively flat response in the visible, while color films are sensitive to the three bands of red, green and blue, and infrared films are sensitive to the near-infrared region. The spectral response of the photographic system can also be changed through the deployment of filters.

The parameters characterizing the performance of a photographic system are resolution, speed and contrast. Spatial resolution is driven by the size of the grains in the film (e.g., the smaller the grain, the higher the resolution). The spatial resolution on the ground (R_g) can be computed from the film spatial resolution (R_f) from the knowledge of the focal length f and the height of the sensor H (e.g., Rees, 2005), as:

$$R_g \approx R_f \cdot H/f \tag{1.1}$$

Film speed provides a measure of the exposure (defined as the illuminance at the film multiplied by the exposure time). Finally, the contrast refers to the range of exposures to which the film responds.

In aerial photography, data collection can be performed either through vertical photography (in which all points in the ground, assumed to be a planar surface, are assumed to be at the same distance H from the camera) or through oblique photography, in which the optical axis is not vertical. Oblique photography allows the coverage of larger areas, but has the disadvantage that the ground plane is no longer constant. Other techniques aiming at covering larger areas include the use of panoramic, strip and reconnaissance cameras. The assumption of a planar surface is often not valid (and consequently the assumption of the ground surface perpendicular to the optical axis). This translates into a phenomenon known as *relief displacement,* in which a point in an image is displaced from the position it would have if it had a zero elevation above the surface. This aspect can be solved by using two vertical images taken over the same scene from two different positions (stereophotography).

More details about aerial photography can be found in Jensen (2006), Lillesand *et al.* (2007), Rees (2001), and Rees (2005), or at http://www.nrcan.gc.ca/earth-sciences/products-services/satellite-photography-imagery/aerial-photos/about-aerial-photography/884 and http://www.fas.org/irp/imint/docs/rst/Sect10/Sect10_1.html.

1.2.2.2   *Visible/near infrared electro-optical sensors*

Electro-optical sensors operating in the visible and near-infrared regions are similar to the ones above described using films. The major difference between film-based and electro-optical sensors consists in the way the data is detected and stored. In the case of electro-optical sensors, indeed, the radiation is collected by means of electronic systems rather than photochemical processes. Electro-optical sensors are similar to modern digital cameras, making use of either charge-coupled device (CCD) or complementary metal-oxide-semiconductor (CMD) sensors.

There are several advantages in using electro-optical systems with respect to traditional photography. Large quantities of data can be stored or transmitted to receiving systems, hence allowing the mounting of such instruments on unmanned airborne and spaceborne platforms. Electronic systems can be calibrated and controlled from remote areas, hence providing a better performance. The collection of images from electro-optical systems is performed through different techniques than those used in traditional photography, especially from space. The approach analogous to traditional photography, indeed, using a two-dimensional array of CCD, would require a large number of detectors. Therefore, in the *pushbroom* method, a one-dimensional CCD is used and the two-dimensional image is obtained through the motion of the platform. An extremely common method is *whiskbroom* imaging, in which a single detector is used and the scene is scanned by means of a rotating mirror.

The spatial resolution, in the case of electro-optical sensors, is driven by the size of the detector projected onto the ground through the instrument's optics. The corresponding feature in an image is called a pixel. The spatial resolution can, therefore, be calculated once the pixel is known from the scaling factor, as already discussed above in the case of traditional photography. The spatial coverage of systems employing digital sensors is generally larger than that of systems using traditional photography, and it can range from tens to thousands of kilometers, depending on the sensor.

The collection of images at multiple bands (usually more than those available through traditional photography) allows performing classification based on multispectral data. In this case, the information in all bands is used to classify a pixel into a land cover or type through its spectral signature, this being a characteristic spectral variation of pixel values at the different available bands. Spectral mixture modeling can be performed on multispectral images. In this case, the area in the image is assumed to be composed of a number of different classes (with known spectral signatures called end-members), and the overall goal is to determine the fraction of each class within each pixel. The number of classes cannot exceed the number of bands, and spectral mixing is extremely helpful when many bands are available (hyperspectral images).

Hyperspectral sensors are instruments that acquire images in many, narrow spectral bands, from the visible through the thermal infrared, up to more than 200 bands. This enables the reconstruction of the reflectance (or emittance) spectrum

for each pixel in the image, which can then be used to classify the features within the scene. The high number of points collected with hyperspectral sensors allows separation among different materials and targets. Applications include mineralogy, water quality, bathymetry, soil and vegetation types, and snow and ice properties.

1.2.2.3 Thermal infrared systems

Though the principles of detection in the case of thermal infrared (800–1400 nm) are similar to those of visible and near-infrared systems, the longer wavelength (and, consequently, the lower energy associated with the photons) in the case of the thermal infrared band imposes modifications on the detectors. Photon detectors employed in the 800–1400 nm range are the mercury cadmium telluride type (also known as MerCaT), or mercury-doped germanium sensors (e.g., Rees, 2005).

One crucial aspect is to minimize the sensors' own thermal emission. The sensors are, therefore, cooled down to temperatures that are as close as possible to the absolute zero and are surrounded by a dewar containing liquid nitrogen at a temperature of 77 K. Because of the larger size of detectors, the spatial resolution in the case of thermal infrared sensors is generally coarser than that of visible or near-infrared sensors. Spectral resolution of thermal sensors for earth surface studies range between 0.1 and 10 μm, depending on the application. Radiometric resolution characterizes the temperature resolution, in view of the possibility of discriminating between small changes in the brightness temperature of the incident radiation, and present-day systems can reach a temperature resolution of the order of 0.1°C.

1.2.2.4 Microwave radiometers

Microwave radiometers measure the energy naturally emitted by the Earth and its atmosphere in the microwave region (0.3–300 GHz). In this regard, passive microwave radiometers are similar to thermal radiometers and scanners. However, unlike the visible and thermal cases, the contribution from the sun is negligible with respect to that from the energy naturally emitted by the earth materials. The quantity measured by microwave radiometers is called brightness temperature (Tb), and is defined as the temperature that a black body in thermal equilibrium with its surrounding would have to be to match the intensity of a grey body at a given frequency. For the temperature range of objects on Earth and in the microwave band, it is possible to adopt the so-called Rayleigh-Jeans law, so that:

$$Tb = \varepsilon T \tag{1.2}$$

with ε being the emissivity and T the body's kinetic temperature.

Unlike the case of visible and near-infrared bands, it is also possible to acquire microwave data through clouds in many of the frequencies available on board spaceborne sensors. The possibility of collecting data in all-weather conditions,

8 | **Marco Tedesco**

and without the need of solar illumination, allows spaceborne passive microwave remote sensors to collect data at high temporal resolution. The relatively coarse spatial resolution (of the order of several tens of kilometers) of data collected by spaceborne microwave sensors is compensated for by the large spatial coverage, due to the large swath.

As an example, Figure 1.3 shows spaceborne brightness temperatures collected at 19.35, 37.0 and 85.5 GHz horizontal polarization, over the northern hemisphere by the Special Sensor Microwave Imager (SSM/I) on November 1, 2006. Although the spatial patterns of the data at different frequencies are similar (e.g., land is warmer than ocean), differences exist due to the different interaction of electromagnetic waves with the earth materials at the different
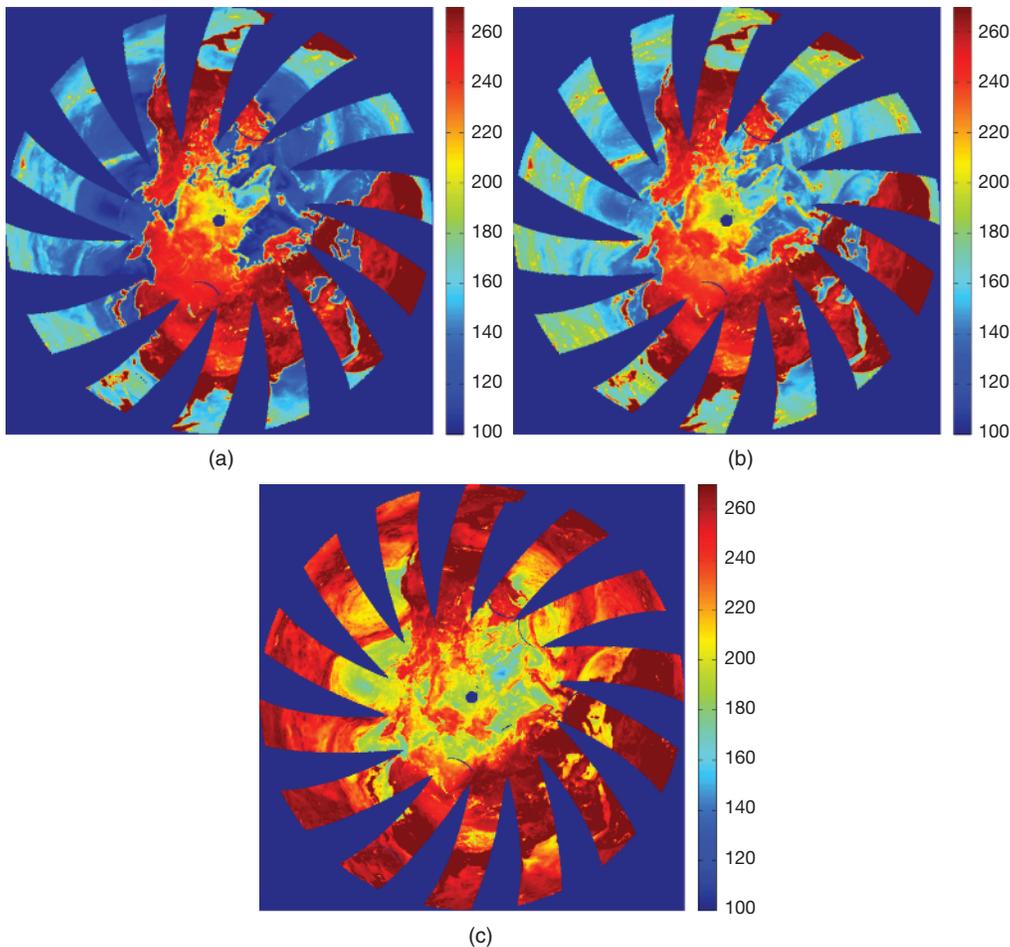
Figure 1.3 Spaceborne brightness temperatures [K] at (a) 19.35, (b) 37.0 and (c) 85.5 GHz horizontal polarization, collected over the Northern Hemisphere by SSM/I on November 1st, 2006 (Data from National Snow and Ice Data Center).

frequencies. This provides complementary information that can be use for retrieval of quantitative information on the physical properties of the different targets. Moreover, the large swath of passive microwave sensors allows covering large areas of the Earth on a daily basis.

1.2.2.5 Gravimetric systems

Gravimetry is the measurement of the strength of a gravitational field, and it has been recently used in remote sensing applications. In this regard, the Gravity Recovery And Climate Experiment (GRACE), a joint mission of NASA and the German Aerospace Center, has been measuring the Earth's gravity field and its changes since March 2002 (http://www.csr.utexas.edu/grace/). GRACE is the first earth-monitoring mission in the history of space flight whose key measurement is not derived from the interaction of electromagnetic waves with the target. The system is composed of two identical spacecraft flying in a polar orbit about 220 kilometers apart at an altitude of 500 kilometers. A microwave-ranging system is used to measure changes in the speed and distance between the two spacecraft with an accuracy as small as 10 μm.

As the twin satellites orbit over the Earth, they encounter gravity anomalies and, from the measurements of the distance between the two satellites and the support of precise positioning measurements from through Global Positioning System (GPS) instruments, scientists can construct a detailed map of Earth's gravity. Accelerometers are also used on each instrument to measure their own movement. Finally, GPS receivers, magnetometers and star cameras are also used to establish baseline positions.

GRACE data have been used for many applications, from the flow of water through aquifers to the hydrological cycle in the Amazon, to ocean circulation. In the case of the cryosphere, GRACE has provided an invaluable contribution in assessing the mass of both the Greenland and Antarctica ice sheets, for example, and seasonal accumulation of snow. A mission similar to GRACE, the Gravity Field and Steady-State Ocean Circulation Explorer (GOCE) by ESA, consisting of a satellite carrying a highly sensitive gravity gradiometer, was launched on March 17, 2009 (http://www.esa.int/Our_Activities/Observing_the_Earth/GOCE). A major difference between GRACE and GOCE is that GRACE measures time variations, while GOCE measures the static gravity field. More information on GRACE is reported in Chapters 8 and 15.

1.2.3 Active systems

1.2.3.1 LiDAR

LiDAR stands for LIght Detection And Ranging, and it makes use of transmitted pulses of laser light, measuring the time of flight to the return of the reflected laser energy. The recorded time is then used to calculate the distance between the sensor and the target. The use of LiDAR for elevation terrain determination and surface

feature mapping began in the late 1970s. Modern LiDAR applications make use of aircraft equipped with GPS and sensors for measuring the angular orientation of the sensor with respect to the ground, called Inertial Measurement Unit (IMU), beside a pulsing laser (from 10 KHz to 800 KHz). Other components include an extremely accurate clock, a reliable data storage system and electronics.

Most airborne LiDAR systems record multiple returns per pulse, allowing discrimination not only between, for example, forest and bare ground, but also surfaces in between, and allowing the construction of digital elevation models of bare earth (e.g., when vegetation is removed). Some newer LiDAR systems record the time series of the return energy waveform rather than the time of several discrete peaks. This full waveform analysis provides an effectively infinite number of returns, and it allows an improved characterization of the physical properties of the medium and of the surface, as well as a more reliable retrieval of the ground surface in vegetated environments. Besides recording the three-dimensional coordinates, LiDAR systems also typically record the energy level of the returning pulse, referred to as *lidar intensity*. This information can be helpful for the identification of surface types, and it can be used as ancillary data or for visualization.

The primary data product from a LiDAR survey is a geolocated point cloud. The point cloud can be characterized by the spacing of points on the target surface, quantified by the ground point spacing (distance) or ground point density (number of points per unit area), which are analogous to the concept of resolution in other technologies. The point spacing is a function of the interaction of laser system properties and flight/orbit parameters with the properties of the target area. For airborne surveys, parameters such as *pulse repetition frequency* (PRF), scan angle, scan frequency, flight speed, and swath overlap are adjusted to achieve the desired ground point spacing as allowed by the terrain relief, target complexity, and surface reflectivity.

The first satellite carrying a LiDAR system for cryospheric applications was the NASA Ice, Cloud, and Land Elevation Satellite (ICESat). ICESat was launched on January 12, 2003 and carried the Geoscience Laser Altimeter System (GLAS), operating at 1.064 and 0.532 μm wavelengths. ICESat operated for seven years and was retired in February 2010. Its purpose was to collect precise altimetry measurements of the polar ice sheets in order to allow mass balance estimates and studies of the impact of Earth's climate on ice sheets and sea level. More information about ICESat, and its planned successor ICESat-2, is provided throughout the book and in Chapter 15.

Airborne LiDAR surveys are seeing increasing application for mapping of snow depth. In 2012, the NASA Airborne Snow Observatory (ASO) mapped snow depth distribution over an entire California river basin on a nominally weekly basis, providing unprecedented scientific opportunities and immediately valuable data products in support of operational water supply forecasting efforts. LiDAR applications such as ASO are likely to be a model for future snow mapping endeavors, providing high-resolution, multitemporal data products, with the flexibility to respond to rapidly changing conditions or extreme events.

1.2.3.2   Radar

Radar is an acronym for Radio Detection and Ranging. Unlike passive sensors, radar generates its own illumination by sending pulses of electromagnetic energy in the microwave region. A fraction of the energy reflected by the target returns to the radar's receiving antenna. The data recorded by the sensor can contain information about the target shape and its physical properties, both at the surface and below. The components of a radar system are a pulse generator, a transmitter, a duplexer (separating the outgoing and returned pulses), a directional antenna that focuses the pulses into a beam and records the returned pulses, a receiver and a recording device.

In the late 1940s, radar was an ideal reconnaissance system for military purposes, as it provides information independently from weather conditions and during both day and night times. The declassification of military technology led to civilian applications, which include terrain mapping, as well as forestry resources, water supplies, and monitoring of ocean surface to study winds and waves, for example.

Spaceborne radar remote sensing began in the late 1970s with the launch of Seasat, followed by the Shuttle Imaging Radar (SIR) and the Soviet Cosmos experiments during the 1980s. In the 1990s, the number of spaceborne radar missions increased, with agencies from different countries launching numerous satellites (such as ERS-1 by ESA, JERS-1 by Japan, Almaz-1 by the former Soviet Union and Radarsat by Canada). Following the 1990s, a number of spaceborne radar missions were launched (such as Envisat by ESA, PALSAR by Japan, the Shuttle Radar Topography mission, SRTM, Radarsat-2 by Canada). A summary of spaceborne radar missions is reported, for example, in Lillesand *et al.* (2007).

The *azimuthal direction* is defined as the forward direction of the aircraft at the flight altitude, while *range direction* is the one where the pulses spread out. Any line-of-sight from the radar to points on the ground defines the *slant range* to that point. The distance between the aircraft nadir (directly below) line and any ground target point is called *ground range*. The duration of a single pulse determines the resolution at a given slant range. The range resolution can be seen as the minimum distance between two reflecting points along the look direction, at that range at which these may be sensed as separate and distinct, and it can be written as:

$$R_r = \tau c/2 \cos \beta \tag{1.3}$$

where:

τ is the pulse length (in microseconds)
c is the speed of light
β (beta) is the depression angle, defined as the angle between a horizontal plane and a given slant range direction.

The *azimuth resolution*, defined as the minimal size of an object along the direction of the flight path that can be resolved, is given by:

$$R_a = S\gamma \tag{1.4}$$

where:

S is the slant range
γ is the angular beamwidth, defined as the wavelength λ divided by the effective
  length of the antenna.

It follows that, for a given wavelength, the antenna beamwidth (and conse-
quently the azimuth resolution) can be controlled by either the physical size of
the antenna or by synthesizing a virtual antenna length. Systems in which the
beamwidth is controlled by the physical size of the antenna are called *real aperture*
or *non-coherent* radars.

Radar antennas on aircraft can be mounted below the platform and direct their
beam to the side of the airplane in a direction normal to the flight path. For air-
craft, this mode of operation is implied in the acronym SLAR, for Side Looking
Airborne Radar. Antennas can reach a size up to 5–6 m. Unlike SLAR, the Syn-
thetic Aperture Radar (SAR) uses small antennas to simulate a large one. This is
achieved by sending pulses from different positions as the platform advances, and
integrating the pulse echoes into a composite signal. With SAR, it is possible to sim-
ulate effective antenna lengths up to 100 m or more. The *Doppler* effect (apparent
frequency shift due to the target's or the radar's velocity) is used in SAR. Indeed, as
coherent pulses from the radar reflect from the ground to the moving platform, the
target acts as if it were in apparent (relative) motion. This translates into changing
frequencies, which give rise to variations in phase and amplitude in the returned
pulses. This data is recorded by the radar system and is recombined to synthesize
signals equivalent to those from a real-aperture system with a larger antenna.

1.3 The cryosphere

The cryosphere is the portion of the Earth where water is in its solid form, either
seasonally or annually. The term comes from the fusion of the Greek words *cryos*
(meaning cold, icy) and *sphaira* (ball, globe). According to Barry *et al*. (2011),
the term was first introduced in 1923 by Antoni Boleslaw Dobrowolski, a Polish
scientist, geophysicist and meteorologist. The components of the cryosphere are
snow cover, glaciers, ice sheets and ice shelves, freshwater ice, sea ice, icebergs, per-
mafrost and ground ice (Figure 1.4). In this book, we will study how remote sensing
is applied to the components of the cryosphere, with the exception of icebergs (see,
for example, Rees (2005) for remote sensing of icebergs).

The cryosphere plays a key role in the global climate system, with important
linkages and feedbacks with the atmosphere and hydrosphere through its impact
on surface energy and moisture fluxes, the release of large amounts of freshwater
into oceans and the lockup of water during the freezing season. Atmospheric,
geophysical, ecological, biological, chemical and geological processes, to name
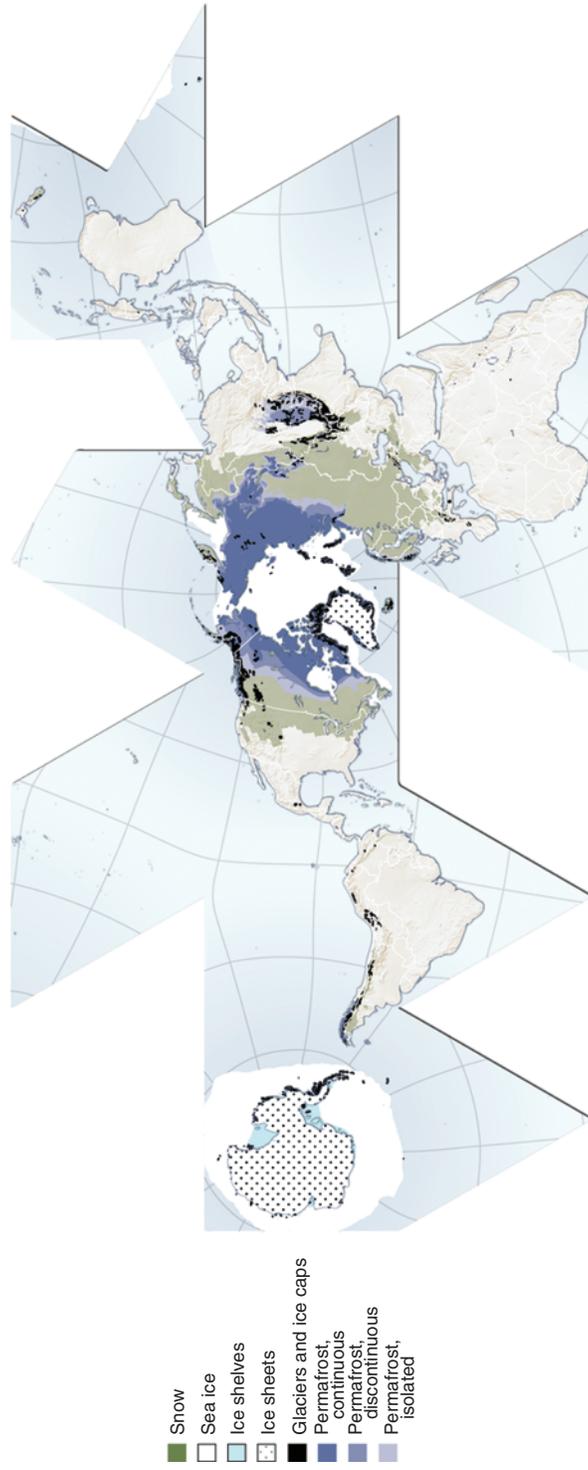
Snow

Sea ice

Ice shelves

Ice sheets

Glaciers and ice caps

Permafrost, continuous

Permafrost, discontinuous

Permafrost, isolated

Figure 1.4 Overview of the cryosphere and its larger components, from the UN Environment Program Global Outlook for Ice and Snow (Fraxen at en.wikipedia. CC-BY-SA-3.0/GFDL).

a few, are impacted by the cryosphere. A brief overview of the components of the cryosphere follows. Readers interested in a more detailed description can refer to Barry and Gan (2011), Marshall (2012) or http://en.wikipedia.org/wiki/cryosphere and related sources and links.

Frozen ground (permafrost and seasonally frozen ground) is the most extensive component of the cryosphere, with 55 million km^2 in the northern hemisphere (of which 22 million km^2 is due to permafrost; Barry and Gan, 2011). Permafrost might occur when the mean annual air temperature (MAAT) is lower than $-1°$C and it is usually continuous when MAAT is below $-7°$C. The thickness of frozen ground can exceed ≈ 600 m along the Arctic coasts of Siberia and Alaska, becoming thinner toward the margins. The active layer (defined as the top layer of the soil that thaws during the summer and freezes again in autumn) plays a key role on the hydrologic and geomorphic regimes.

The second most extensive component of the cryosphere is snow, with a maximum cover extent of ≈ 47 million km^2 in January and a mean annual area of ≈ 26 million km^2 (Barry and Gan, 2011). Most of the snow cover is located in the northern hemisphere, with the maximum snow cover extent of ≈ 0.85 million km^2 in the southern hemisphere occurring in July. Though relatively limited in extent, snow stored in high mountain areas provides the major source of runoff for streamflow and groundwater recharge in many mid-latitude areas, and many regions of the Earth rely on snow accumulated during the winter for water resources and management. The high albedo of snow (e.g., the ratio between the incoming and outgoing solar radiation, which can be up to $0.8–0.9$ in the case of fresh snow) regulates the amount of solar energy that is absorbed by snow-covered areas in the northern hemisphere during winter.

Sea ice, formed by the freezing of the ocean water, is the third most extensive component of the cryosphere, with maximum winter extent of $\approx 14–16$ million km^2 in the northern hemisphere and $\approx 17–20$ million km^2 in the southern hemisphere. The extent in the northern hemisphere reduces down to less than $4–6$ million km^2 during the summer and $3–4$ million km^2 in the southern hemisphere.

The first sea ice is made of small (0.3 cm) separate disk-shaped crystals floating on the surface, then turning into crystals with a hexagonal shape and arms stretching out over the surface. As these arms break, the turbulence of the water causes a suspension of randomly shaped crystals of increasing density in the surface water, an ice type called *grease ice* or *frazil*. In quiet conditions, the crystals freeze together, forming a thin layer of ice, called *nilas* when it is still transparent. First-year ice (FYI) is then formed by the *congelation growth*, in which water freezes at the bottom of the ice layer. In rough waters, fresh new ice is formed by the cooling of the ocean as heat is released into the atmosphere. The upper layer of the ocean is supercooled and frazil ice forms, which then turns into grease ice, a mushy surface layer. Waves and wind then compress this ice into plates of several meters in diameter, called pancake ice, which will form upturned edges as the plates collide with one other while floating on the ocean surface. Eventually, the pancake ice plates will freeze together to form a consolidated pancake ice.

First-year sea ice is ice that is thicker than *young ice* but has no more than one year of growth. Simply put, first-year ice grows during the fall and winter, but melts during the summer. *Old sea ice* is sea ice that has survived at least one melting season and is generally divided into s*econd-year ice* (which has survived one melting season) and *multiyear ice* (which has survived more than one melting season). *Leads* and *polynyas* are areas of open water that occur even though air temperatures are below freezing. Leads are narrow and linear, while polynyas are larger, and both provide a direct interaction between the ocean and the atmosphere.

Glaciers and ice sheets are large bodies of ice that form where accumulation of snow exceeds its ablation (e.g., melting and sublimation) over time scales of many years (decades, centuries and millennia). Under the pressure of the layers above, snow and firn fuses into denser and denser firn. These layers undergo further compaction over many years and turn into glacial ice. By definition, ice sheets cover areas larger than $50,000$ km^2. At least 0.1 km^2 and 50 m thick, a glacier deforms and flows because of the stress due to its own weight.

Ice sheets hold 77% of the world's freshwater, of which Antarctica accounts for 90% and Greenland for almost 10% (and the remaining ice caps and glaciers $\approx 0.5\%$). The ice sheets contain an ice volume of 2.85 million km^3 (Greenland) and 24.7 million km^3 (Antarctica), compared to an estimated ice volume of 0.24 million km^3 in the case of glaciers and ice caps, of 0.66 million km^3 in the case of ice shelves and only 0.002 million km^3 in the case of snow on land in late January.

The ablation zone of a glacier is the region where there is a net loss in glacier mass (e.g., negative surface mass balance). The equilibrium line altitude (ELA) is the elevation where the surface mass balance turns from negative to positive, separating the ablation zone from the accumulation zone, where snowpack or superimposed ice accumulation persists. Within the accumulation zone, the dry snow zone is the region where no melt occurs (even during summer) and the percolation zone is where melt occurs, causing meltwater to percolate into the snowpack.

Ice forms over rivers and lakes during the winter. The effects of such components of the cryosphere are mostly local, in view of the small extent. However, because the freeze/thaw cycles respond to large-scale and local weather factors, measurements of lake and river ice can support climate studies. Because of the absence or low concentration of salt in lake and river waters, the ice occurring from their freezing can be considered as pure ice. In the case of lakes, the freezing begins on the surface with millimeter-size crystals (frazil). Because of their preferential growth along the crystallographic *c*-axis, a continuous layer of crystals with the *c*-axis horizontal forms. This ice is termed *black ice*, and it is represented by a smooth transparent slab of freshwater ice.

If snow deposits on the frozen lake, the weight of the snowpack might create depressions and, eventually, might crack the surface. Liquid water can then fill the cracks, saturating the snow overlying the ice. Once refrozen, this layer is called *white ice*. The formation on rivers differs slightly from that of lakes, because of the presence of river flow. In general, for slow-flowing rivers, the conditions are similar to those for a lake; but for fast-flowing rivers, the frazil can be distributed along the water column.

References

Barry, R.G. & Gan, T.Y. (2011). *The Global Cryosphere*. Cambridge University Press, 1st Edition, pp. 498, ISBN: 0521156858.

Barry, R.G., Jania, J., & Birkenmajer, K. (2011). A.B. Dobrowolski – the first Cryospheric scientist – and the subsequent development of Cryospheric science. *Hist. Geo-Space Sci.* **2**, 75 – 79.

Jensen, J. (2006). *Remote Sensing of the Environment: An Earth Resource Perspective.* Prentice Hall, pp. 608, ISBN-10: 0131889508.

Lillesand P., Kiefer, R.W., & Chipman, J. (2007). *Remote Sensing and image interpretation*. Wiley, 6th Edition, pp. 804.

Marshall, S.J. (2012). *The Cryosphere*. Princeton University Press, pp. 288, ISBN 9780691145259.

Rees, W.G. (2001). *Physical Principles of Remote Sensing*, Second edition. Cambridge, Cambridge University Press.

Rees, G.W. (2005). *Remote Sensing of Snow and Ice*. CRC Press. pp. 312.