1

Background

1.1 Introduction

During the past forty years, rapid technological growth has advanced the ability of satellites to observe and monitor the global ocean and its overlying atmosphere. Because of similar advances in computer hardware and software, it is now possible to acquire and analyze, at short time delays, large satellite data sets such as the global distribution of ocean waves, the variations in sea surface height associated with large-scale current systems and planetary waves, surface vector winds and regional and global variations in ocean biology. The immediate availability of these data allows their assimilation into numerical models, where they contribute to the prediction of future oceanic weather and climate.

The ocean covers approximately 70% of the Earth's surface, is dynamic on a variety of scales, and contains most of the Earth's water as well as important marine ecosystems. The ocean also contains about 25% of the total planetary vegetation, with much of this restricted to a few coastal regions (Jeffrey and Mantoura, 1997). Regions of high biological productivity include the Grand Banks off Newfoundland, the Bering Sea and Gulf of Alaska, the North Sea and the Peruvian coast. Between 80% and 90% of the world's fish catch occurs in these and similar regions. For its role in climate, determination of the changes in ocean heat storage and measurement of the vertical fluxes of heat, moisture and CO_2 between the atmosphere and ocean are critical to understanding global warming and climate change.

Large-scale ocean currents carry about half of the heat transported between the equator and the poles; the atmosphere transports the remainder. Away from the polar regions, the combination of these transports with the large oceanic heat capacity relative to the atmosphere means that the ocean moderates the global climate and improves the habitability of the continents (Stewart, 1981; Chelton, 2001). For the polar regions, the recent increase in the melting of the Greenland and Antarctic icecaps and the dramatic decrease in the arctic summer sea ice cover (Comiso, 2010) show that the ability to monitor the extent and thickness of the Arctic and Antarctic ice covers is important both for short-term navigation needs and for long-term climate studies. All these examples illustrate the need to monitor and observe the ocean on a range of local to global scales.

The growth in satellite systems has been driven in part by technology and in part by societal concerns. Societal concerns include the importance of the ocean to national security and naval operations, global commerce, the prediction of severe storms and hurricanes, fisheries management, the extraction of offshore gas, oil and minerals, and public health and recreation. Regarding commerce, in 2012, there were about 100,000 ships engaged in commerce, oil, gas and mineral exploration, fisheries and recreation (Allianz, 2012). Increasingly, these concerns also include global sea level rise and the change in the areal extent of the Arctic and Antarctic sea ice. In addition, about half of the global population lives within 200 km of the coast, where fourteen of the seventeen largest cities are coastal. Of these, eleven are Asian, including Bangkok, Jakarta, Shanghai, Tokyo, Ho-Chi-Minh City, Calcutta and Manila (Creel, 2003). These populations are vulnerable to natural hazards such as the storm surge and flooding associated with the combination of sea level rise and hurricanes or typhoons. There are also public heath considerations associated with the oceanic disposal of urban runoff, sewage and garbage, and with the monitoring and prediction of the growth of pathogenic organisms such as red tides. Satellite observing systems and the interpretation of the resultant data play a central role in addressing these concerns.

In the 1970s, the United States launched the first ocean remote sensing satellites. Since that time, many countries have launched satellites that carry oceanographic instrumentation, and, as Section 1.8 describes, beginning in about 2002 there has been an international effort to organize satellites from different countries into what are called observing *constellations*. These constellations are made up of satellites that carry similar instruments, observe the same oceanic variables and fly in complementary orbits, so that the coverage by a single satellite is enhanced by observations from the other constellation members. The data from the constellation are then placed in a common format and distributed among the participants and other interested parties.

With these observations, there is an emphasis on the rapid dissemination of the data to the various government and private-sector users, and the use of this near-real-time data in numerical models and in other areas such as search-and-rescue, oceanographic research cruise support and the routing of cargo ships to avoid storms. Examples of the oceanic variables observed by these satellites include sea surface temperature (SST), the height and directional distribution of ocean swell, wind speed and direction, atmospheric water content and rain rate, the changes in sea surface height associated with ocean tides, currents and planetary waves, concentrations of phytoplankton, sediments and suspended and dissolved material, and the areal extent and types of polar sea ice.

Prior to the 1980s, such properties were determined from dedicated and expensive ship expeditions, or in the polar regions from surveys made from aircraft, drifting ships and ice islands. This meant that the ocean could be surveyed only slowly and incrementally. At present, satellite imagers can make simultaneous observations of the desired variables with scales of 1–1,000 km that are difficult to observe even from multiple ships. For variables such as the near surface air temperature that are not retrievable by remote sensing, some satellites are designed to relay data from moored and drifting buoys that make direct measurements of such quantities to national data centers. Even for those ocean depths that are inaccessible to satellite observations, instruments called Argos floats are deployed in large numbers that profile the ocean interior and periodically come to the surface, where they report their observations by satellite.

Because satellites survey a variety of oceanic properties with near global coverage and at intervals of 1–10 days, then rapidly transmit these observations to national and international forecast centers, these data are of great operational importance. In addition, the observations contribute to long-term studies and numerical modeling of global climate change, sea level rise, and the decadal-scale atmospheric and oceanographic oscillations, including the Pacific Decadal Oscillation (PDO), North Atlantic Oscillation (NAO), El Niño/Southern Ocean oscillation (ENSO), and Arctic Oscillation (AO).

In the following, Section 1.2 defines remote sensing and describes its oceanographic applications. Section 1.3 describes the satellite orbits used in remote sensing and summarizes the hazards faced by satellites. Sections 1.4 and 1.5 describe the geosynchronous and Sun-synchronous satellites. Section 1.6 discusses the imaging techniques used by satellites in Sun-synchronous and other low Earth orbits. Section 1.7 describes the different processing levels of satellite image data and the NASA data archives. Section 1.8 gives a brief history of the changes in satellite remote sensing over the past forty years, describes the international context of these observations, and presents a table of past, present and pending satellite missions through 2015.

1.2 Definition of remote sensing

Earth remote sensing is primarily defined as the use of electromagnetic radiation to acquire information about the ocean, land and atmosphere without being in physical contact with the object, surface or phenomenon under investigation. Remote sensing is not unique to electromagnetic radiation, as this book shows, there are also techniques for studying changes in ocean circulation and ice sheet properties through remote observations of gravity anomalies. Unlike shipboard measurements of quantities such as SST or wind speed, which are direct measurements made at a point by a thermometer or anemometer, remote sensing measurements of such quantities cover broad areas and are indirect, in that the geophysical quantity of interest is inferred from the properties of the reflected or emitted radiation. The sensors can range from a radiometer mounted on a ship, oil platform or aircraft to a multispectral satellite imager. The following briefly describes the concepts behind remote sensing and the various observing bands.

Because the satellite instrument is not in physical contact with the phenomena under investigation, its properties must be inferred from the intensity and frequency distribution of the received radiation. This distribution depends on how the received radiation is generated and altered by its propagation through the atmosphere. This radiation has three principal sources: blackbody radiation emitted from the surface, reflected solar radiation, and, for the directed energy pulses transmitted by satellite radars, the backscattered energy received at the sensor. The properties of the received radiation also depend on the sensor, which must be designed so that its observing wavelengths are appropriate for the phenomenon in question. Finally, the received data must be organized into images or data sets so that the spatial distributions of the quantity under investigation can be viewed. This is the generally accepted definition of remote sensing; in the past decade, it has been expanded to include the use of satellite measurements of gravity to infer changes in land, ice sheet and ocean properties.

Because of the atmospheric contributions to the reflected and received radiation described in Chapters 4 and 9, there are three electromagnetic wavelength bands or windows, called the visible, infrared and microwave, through which the ocean is viewed. In the visible and extending into the near infrared, the observations depend on reflected sunlight and are restricted to daytime cloud-free periods. Because the visible spectrum contains the only wavelengths at which light penetrates to oceanic depths of order 10–100 m, visible observations yield the only information on the depth-averaged color changes associated with phytoplankton and sediment concentrations. In the infrared, the observations measure the blackbody radiation emitted from the top few micrometers of the sea surface. Although these observations are independent of daylight, infrared satellite observations are restricted to cloud-free conditions.

In the microwave and especially at the longer microwave wavelengths, the surface can be viewed through clouds and is obscured only by heavy rain. Microwave observations divide into passive and active. Passive microwave instruments observe the naturally emitted blackbody radiation, which can be used to retrieve such atmosphere and ocean surface properties as the areal extent of ice cover, the atmospheric water vapor and liquid water content, sea surface temperature (SST), salinity, and, through the directional dependence of the sea surface roughness, the vector wind speed.

In contrast, different kinds of radars make *active* measurements; these instruments transmit pulses of energy toward the ocean, then receive the backscatter, so that they provide their own illumination. The active microwave instruments include imaging radars (the Synthetic Aperture Radar or SAR), directed, pulsed vertical beams (altimeter), several pulsed fan beams at oblique angles to the satellite orbit (scatterometer), and an oblique rotating pulsed beam (also scatterometer). The scatterometers are highly directional radars that receive the backscatter from relatively small surface areas. Together, these instruments provide information on the roughness and topography of the sea surface, wind speed and direction, wave heights, directional spectra of ocean surface waves and the distribution and types of sea ice.

1.3 Satellite orbits

The orbit of an Earth-observing satellite divides into two parts, the satellite motion in its orbit plane relative to the Earth's center of mass, and the satellite position relative to the rotating Earth. The time-dependent position of the satellite in its orbit is called the satellite ephemeris. For the rotating Earth, the orbit is frequently described in terms of its ground track*,* which is the time-dependent location of the surface intersection of the line between the satellite and the Earth's center of mass. The point directly beneath the satellite is called the satellite nadir. The first of the following sections considers the theoretical case of satellite motion in its orbit plane, and describes how the addition of the Earth's rotation determines the satellite ground track; the second considers the actual space environment of these satellites, and the constraints imposed on the satellites and their instruments by space debris and uncontrolled satellites, gravity-induced orbit perturbations, solar storms and radiation, and radio-frequency interference (RFI).

1.3.1 Satellite orbits and their applications

Rees (2001, Chapter 10), Elachi (1987, Appendix B) and Duck and King (1983) survey the commonly used, near circular orbits used in remote sensing. These orbits are described in a rectangular coordinate system with its origin at the Earth's center of mass. The z-axis is in the northerly direction and co-located with the Earth's rotation axis, the x-axis is in the equatorial plane and points in the direction γ of a star in the constellation Aries, and the y-axis is in the direction appropriate for a right-handed coordinate system. Relative to these axes, the six Keplerian orbital elements describe the satellite location. Because two of these are specific to elliptical orbits, for circular orbits, the six elements are reduced to four.

As Figure 1.1 shows, these four elements are as follows. First, the right ascension of the ascending node, or simply the ascending node Ω, is the angle between the x-axis and the point at which the orbit crosses the equator. Second, the radial distance H is the height of the satellite above the Earth's center of mass. Third, the true anomaly θ is the angular position of the satellite in its orbit relative to Ω. Fourth, the inclination I is the angle between the Earth's axis and the normal to the orbit plane with the convention that I is always positive. Of these variables, I and Ω specify the orientation and position of the orbit plane relative to the fixed stars; H and θ specify the satellite position within the orbit plane. The advantage of this description is that I, Ω and H are either fixed or slowly varying, so that, over short periods, θ describes the instantaneous satellite position. Based on the magnitude of I, there are three kinds of orbits. If $I = 90°$, the orbit is polar; if $I < 90°$, the orbit is prograde and precesses in the same direction as the Earth's rotation as in Figure 1.2; if $I > 90°$, the orbit is retrograde and precesses in the opposite direction.

In remote sensing, interest is generally not in the satellite position in its orbit, but rather in its location on its surface ground track. For a non-rotating spherical Earth, the orbit track is a great circle, or, on the Mercator map shown in Figure 1.2(a), a simple sine wave (Elachi, 1987, Section B-1–4). Because of the Earth's rotation, the orbit track is steadily displaced to the west, yielding the succession of tracks shown in Figure 1.2(b). On the tracks, the numbers i, ii, iii mark the beginning and end of each orbit, where, for example, the points marked ii are at the same time and geographic location. Another orbit property

Cambridge University Press
978-1-107-01938-6 - An Introduction to Ocean Remote Sensing: Second Edition
Seelye Martin
Excerpt
More information

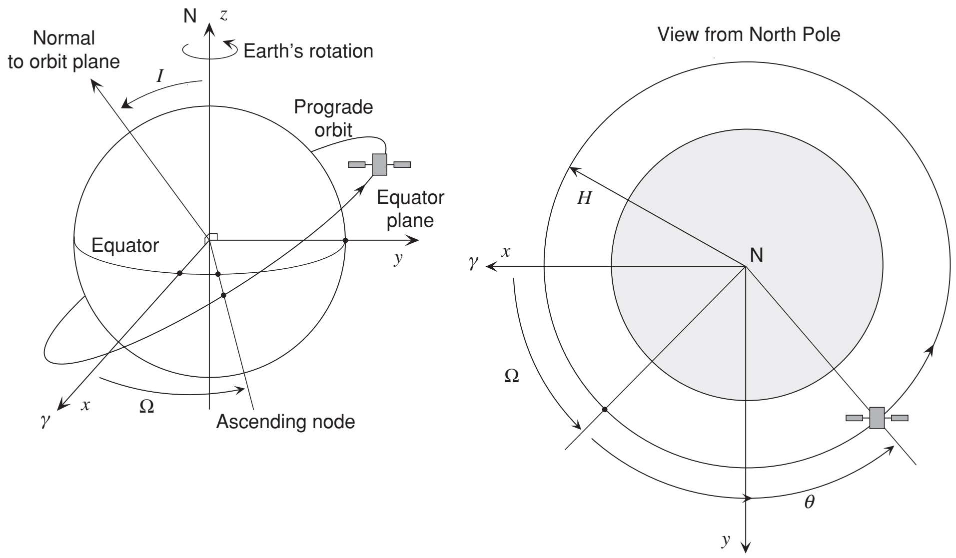

Fig. 1.1. For a circular orbit, the Keplerian parameters used to describe the orientation of the orbit plane and the satellite position along the orbit.

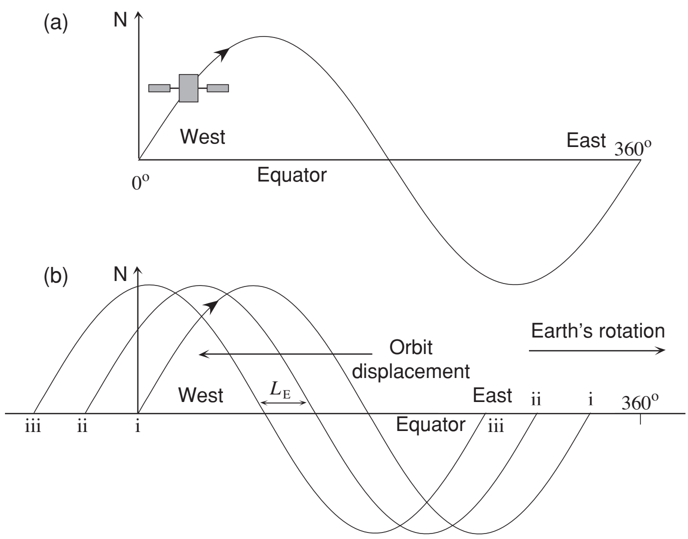

Fig. 1.2. Mercator map of the satellite ground track for the orbit shown in Figure 1.1 and for a (a) non-rotating Earth and (b) rotating Earth. See the text for further description. (Adapted from Elachi (1987, Figure B-6).

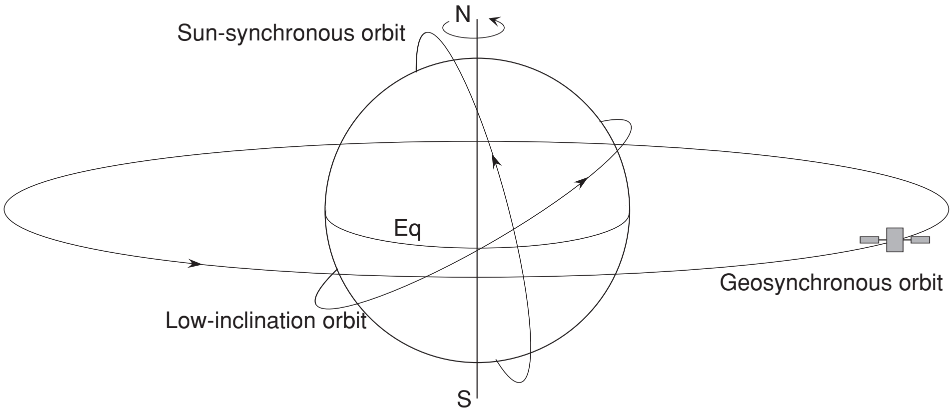

Fig. 1.3. Examples of the Sun-synchronous, geosynchronous and low-inclination orbits, where "Eq" is the equator. (Adapted from Asrar and Dozier (1994), Figure 3).

concerns the equatorial separation L_E between successive orbits. If division of a multiple of the equatorial circumference by L_E is an integer, the orbit is an exact repeat orbit, so that, after a given period of time, the satellite repeats the same track lines. This property is particularly valuable for instruments such as the altimeter, since it allows successive measurements of sea surface height along the same ground track.

The three common Earth observation orbits are called geosynchronous, Sun-synchronous, and near equatorial low inclination (Figure 1.3). There is also a fourth altimeter orbit used for observations of sea surface topography that is at a slightly higher altitude than the Sun-synchronous orbits, and there are also various low-altitude non-Sun-synchronous orbits used for observations of phenomena such as winds and rainfall. The following summary shows that each particular orbit has advantages and disadvantages. Because no single orbit allows coverage of all space and time scales, there is no such thing as a "perfect" satellite orbit or system. Instead, the choice of orbit depends on the phenomenon under investigation.

The geosynchronous orbits are located at an altitude of 35,800 km above the equator. The geostationary orbit is a special case; it lies in the Earth's equatorial plane ($I = 0°$). In this orbit, although the satellite is orbiting the Earth such that it moves in and out of the Earth's shadow, its position remains over a fixed equatorial location so that it continuously observes the same surface area. The plane of the more general geosynchronous orbit is tilted relative to the equator ($I \neq 0°$), so that, although the mean surface position of this satellite is stationary, its ground path is described by a figure eight centered on the equator (Elachi, 1987). The period of a geosynchronous satellite is 23.93 hours, which is the time in which the Earth rotates around its axis relative to the fixed stars. In contrast, the 24-hour day is the time between successive noons, defined as when the sun is directly overhead, so that the length of day is determined from a combination of the Earth's rotation about its axis and the Earth's rotation in its orbit.

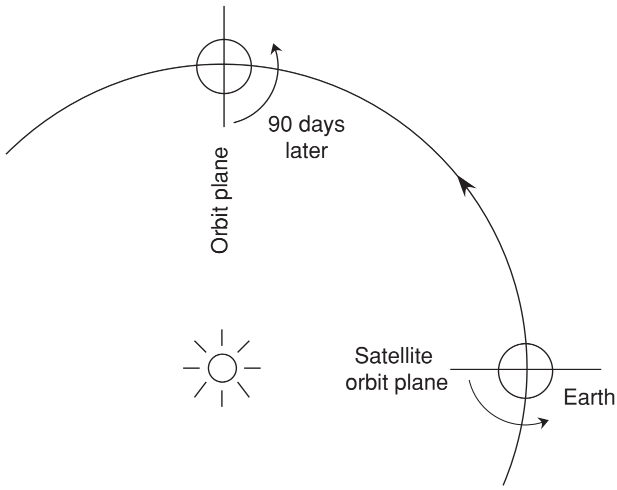

Fig. 1.4. Rotation of the plane of a Sun-synchronous orbit in the Earth–Sun orbit plane.

Operators and managers of geosynchronous satellites work in terms of a "geosynchronous belt", defined as the region extending 200 km above and below the geosynchronous altitude and ±15° in latitude (IADC, 2007; Weeden, 2010). Within this belt, the satellites occupy slots that measure about 2° in longitude, where their operators try to maintain the satellite within a 0.1° box (Weeden, 2010). In Earth observations, geosynchronous satellites provide observations of weather, SST and ocean color, and provide data relay services.

The Sun-synchronous orbit is retrograde with $I > 90°$, and has an altitude of about 800 km, or a much lower altitude than the geosynchronous orbits. The Sun-synchronous period is about 90 minutes, corresponding to about sixteen orbits per day. The reason why this orbit is called Sun-synchronous is that throughout the year each orbit crosses the equator at the same local time of day. Consequently, Ω is not constant, but changes slowly with time. The drift occurs because of the Earth's equatorial bulge, which causes the plane of a near polar orbit to rotate slowly around the pole (Rees, 2001). For a retrograde orbit, the inclination and orbit height can be set so that the orbit rotates about 1° per day in the ecliptic or Earth–Sun plane, and in an equal but opposite direction to the orbital motion of the Earth around the Sun. Relative to the fixed stars, the Sun-synchronous orbit plane rotates once per year, so that its orbit plane remains at a constant angle to the line between the Sun and Earth. Figure 1.4 shows the change in the angular position of the orbit in the Earth–Sun plane as the Earth moves an angular distance of 90° in its orbit, during a period of approximately 90 days.

Sun-synchronous satellites are the most common of the ocean-observing satellites and are often referred to as polar orbiters. Their orbits are described in terms of their daytime equatorial crossing times, as in a 0730-descending or a 1330-ascending orbit, where descending refers to a southward satellite velocity, ascending refers to a northward velocity, and the crossing time is local. The orbits are also described in terms of their crossing times, as "early morning", "mid-morning" and "early afternoon". Because the Sun-synchronous equator crossings always occur at the same local time of day, satellites in this orbit can make daily observations of SST or ocean chlorophyll at the same time in their diurnal cycle. Since cloudiness over the ocean generally increases throughout the day, the crossing time can be chosen to minimize cloudiness under the satellite.

One difficulty with this orbit is that, because of the tilted orbit plane, the satellite does not pass directly over the poles. This means that the regions around the poles may be excluded from instrument coverage; this lack of coverage is called the *hole at the pole*. Figures 4.2 and 9.18 give examples of the swath coverage for this orbit, and show that, depending on the instrument, a single Sun-synchronous satellite can provide near global coverage at 1–2-day intervals.

The near-equatorial low-inclination orbit used for missions such as the Tropical Rainfall Measuring Mission (TRMM) is circular with an altitude of 350 km and an inclination angle of 35°. This orbit covers approximately half the globe, and, in a one-month period, observes any specific area at every hour of the day with a sampling rate that is roughly twice that of a polar orbiter. The advantage of this orbit is that it allows TRMM to determine the variability of tropical rainfall throughout its diurnal cycle. The successor to this mission is the joint US/Japanese Global Precipitation Measurement (GPM) Core mission, with a greater inclination angle of 65° that is scheduled for launch in 2014. Another member of the GPM constellation in a similar orbit is the Indian/French Megha-Tropiques rainfall mission with an inclination angle of 22° that was launched in 2010.

Finally, the altimeter occupies an orbit designed to measure sea surface height. Because the tidal bulge associated with the 12- and 24-hour tides always lies directly beneath a satellite in a Sun-synchronous orbit, some altimeters operate at a higher non-synchronous altitude of 1200–1400 km. Consequently, the orbit is not in phase with the tides and the satellite experiences a smaller atmospheric drag. Altimeter satellites in this orbit include the US/French TOPEX/POSEIDON JASON-1, JASON-2 and the forthcoming JASON-3 mission discussed in Chapter 12.

1.3.2 The satellite environment: Solar storms, radiation pressure, the South Atlantic Anomaly, gravitational perturbations, space debris, graveyard orbits and radio frequency interference (RFI)

In space, various factors perturb the satellites, their orbits and their instruments. First, the lunar and solar gravity fields and radiation pressure from the solar wind exert forces on the satellites and perturb their orbits. Second, there are two bulges in the Earth's gravity field called libration points, one over India (105° W) and the other at the longitude of the US Rocky Mountains (75° W), that also affect the orbits (Weeden, 2010). For this reason, all satellites have engines and carry fuel so that they can maintain their desired orbits. Third, the satellite can be damaged or destroyed by collisions with space debris or other, sometimes decommissioned, satellites.

The NASA Orbital Debris Program Office (NASA, 2012a) monitors space debris; ESA (2012a) describes the ESA monitoring of debris. As of 2009, ESA (2012a) states that there were 14,000 catalogued pieces of space debris, and approximately 600,000 uncatalogued pieces of debris with dimensions greater than 1 cm. Depending on their relative velocity, even a small object can damage or destroy a satellite. In the low Earth orbits (LEO), the maximum amount of debris occurs at two altitudes: the polar orbit altitudes at 800–1000 km and the altimeter satellite altitude of 1400 km. For the geosynchronous belt, the amount of debris is about two orders of magnitude less than in LEO.

ESA (2012a) describes the growth in the amount of debris and its sources. For example, in January 2007, the Chinese use of an anti-satellite missile to destroy the Sun-synchronous Feng-Yun 1C satellite led to a 25% increase in catalogued debris. In February 2009, the first accidental collision of two satellites occurred in LEO when the American commercial satellite, Iridium-33, collided with a Russian military satellite, Kosmos-2251, destroying both satellites and generating a large amount of debris. For the rest of 2009, five satellites, namely the remote sensing satellites AQUA and Landsat-7 at altitudes of about 700 km, the Space Station and Space Shuttle at an altitude of 400 km, and a NASA Tracking and Data Relay Satellite (TDRS-3) in geosynchronous orbit, maneuvered to avoid collisions with debris (David, 2010). Based on the current growth in satellite debris, Donald Kessler has forecast the occurrence of what is called a *Kessler* syndrome or cascade, where the frequency of collisions will increase at such a rate and generate so much debris that all of the satellites in LEO would be destroyed (Kessler interview in David, 2010).

For geosynchronous satellites, Weeden (2010) states that, in 2010, there were 1238 catalogued objects in the geosynchronous belt, of which 391 were under control, 594 were drifting, 169 had been captured by the libration points, and the remainder were lost or undocumented. He also describes the fate of the Intelsat Galaxy-15 satellite that, during a solar storm in April 2010 when the satellite was positioned at 130° W, lost contact with its ground controllers. Because of this, it drifted east toward the North American libration point, and received the nickname *Zombiesat*. As it drifted east, its transponders continued to receive and transmit data broadcast from the ground, causing both radio interference and hazards to other satellites. This situation continued until January 2011, by which time the satellite had passed through the orbital slots of about fifteen communication satellites, when Intelsat restored communications with Galaxy-15, and returned it to a safe position (Space News, 2011).

Given these problems with space debris, 11 nations with space programs and ESA formed the 12-member Inter-Agency Space Debris Coordination Committee (IADC, 2012). The IADC recommends that, to avoid further generation of debris, two protected regions be established. The first contains the LEO, which IADC defines as the global region extending in altitude from the surface to 2000 km, and covering the Sun-synchronous and altimeter orbits; the second contains the geosynchronous orbits (GEO). For LEO, IADC (2007) recommends that, when the satellite approaches the end of its lifetime, it be deorbited into the atmosphere. For GEO, IADC recommends that a satellite approaching its end of service should be placed into a graveyard orbit located at an altitude of about 100–200 km above