# 1

# Introduction

## 1.1 Historical Background

Perhaps the theory of radiative transfer originated at the dawn of human civilization at the moment when a person thought about why the Sun warms. The gradual realization of what radiation is eventually led to the birth of quantum mechanics. Probably among the three types of heat transfer (thermal conduction, convection, and radiation), it was the phenomenon of radiation that was the most mysterious for ancient researchers. Hippocrates of Kos, an ancient Greek physician and philosopher (460–370 BC), said, "Light is a mystery to us. We do not understand what it is, or how it can be. We know only that it exists, and that it is beautiful."

Several early researchers made significant contributions to our understanding of radiative transfer. Their work continues to influence research in this field even today. Johann Lambert (1728–1777) studied how the incident light reflects from a surface depending on the angle of incidence. He found that the radiant intensity or luminous intensity observed from an ideal diffusely reflecting surface is directly proportional to the cosine of the angle $\theta$ between the observer's line of sight and the surface normal (Lambert, 1760):

$$I = I_0 \cos\theta, \tag{1.1}$$

where $I$ is the intensity of the reflected light, and $I_0$ is the intensity of the incident light. A surface that obeys Lambert's law is said to be Lambertian and exhibits Lambertian reflectance. Such a surface has the same radiance (luminance) when viewed from any angle. This means, for example, that to the human eye it has the same apparent brightness. It has the same radiance because, although the emitted power from a given area element is reduced by the cosine of the emission angle, the solid angle, subtended by surface visible to the viewer, is reduced by the very same amount. Because the ratio between power and solid angle is constant, radiance (power per unit solid angle per unit projected source area) stays the same. Despite its simplicity, the Lambert cosine law has found many practical applications, such as lighting design, solar energy collection, and computer graphics. In remote sensing, this law is used to model the reflection of electromagnetic radiation as it interacts with the Earth's surfaces. Lambert introduced two axioms of photometry: light travels in a straight line in a uniform medium and rays that cross do not interact. It was also shown that light intensity decays exponentially in an absorbing medium. All results of Lambert

*Analytical Methods in Radiative Transfer*, First Edition. Alexander Kokhanovsky, Vijay Natraj and Dmitry Efremenko.

were summarized in his book *Photometria,* which is considered a landmark in the field of photometry, as it introduced several important concepts and techniques that are still used today.

Pierre-Simon Laplace (1749–1827) made significant contributions to the theory of radiative transfer. Laplace was interested in understanding the nature of energy transfer in the Earth's atmosphere. He was one of the first scientists who realized the important role of radiative transfer in the energy balance of the Earth system. Laplace developed the theory of probabilities for describing the processes occurring in nature. He formulated what became known as the central limit theorem and proved that the distribution of errors in large data samples from astronomical observations can be approximated by a normal distribution. The description of the behavior of radiation, as it interacts with matter using probabilities and statistics, appeared to be very useful and became the fundamental concept of radiative transfer.

One of the first scientists during the late modern period who investigated the properties of radiation and conducted a quantitative study of radiative transfer was John Tyndall (1820–1893). One of his experiments involved measuring the infrared emission of a platinum filament and observing the corresponding color of the filament. A platinum wire was placed inside a glass tube filled with nitrogen gas. An electric current passed through the wire heated it, causing it to emit radiation. Tyndall used a thermopile to measure the intensity of the radiation emitted by the wire. He also observed that as the wire was heated, it first appeared red, then orange, yellow, white, and finally blue-white. This sequence of colors is referred to as the Tyndall effect. Tyndall measured the intensity of the radiation emitted by the wire at different wavelengths using a prism and a spectroscope. He found that as the wire was heated, the intensity of the radiation increased at shorter wavelengths, corresponding to the blue end of the spectrum. This experiment demonstrated the relationship between the color of a heated object and the wavelengths of radiation emitted by the object. It also provided important insights into the properties of infrared radiation, which is invisible to the human eye but plays a crucial role in many physical processes, such as heat transfer and climate change.

Based on experiments of Tyndall, Josef Stefan (1835–1893) found the proportionality of the intensity to the fourth power of the absolute temperature (see Figure 1.1). Later, Ludwig Boltzmann explained this result theoretically, what is now referred to as the Stefan–Boltzmann law (Boltzmann, 1884). It states that the total radiant heat power $E$ emitted from the surface of a blackbody across all wavelengths is proportional to the fourth power of its absolute temperature $T$:

$$E = \sigma T^4, \tag{1.2}$$

where $\sigma \approx 5.67 \cdot 10^{-8}\mathrm{W/(m^2K^4)}$ is the Stefan–Boltzmann constant. The discovery of Stefan–Boltzmann law made it possible to estimate quite accurately the temperature of the Sun's surface (Stefan, 1879). Now we know that any object emits radiation. It is even possible to visualize radiation using thermal imagers. As shown in Figure 1.2, faces and hands are the warmest "objects" and hence have the most intensive emission.

In 1900, Max Planck formulated his law as part of his efforts to solve the problem of blackbody radiation, which had puzzled physicists for decades. Planck's law is based on the

Um die Wärmemenge zu erhalten, welche dieselbe Fläche bei der Temperatur 100° ausstrahlt, hat man diese Zahl mit $a^{100} = 2{\cdot}146$ zu multipliciren und erhält 1·6644. Die Differenz der beiden 0·889 gibt die Wärmemenge, welche durch die Einheit der Oberfläche einer Glaskugel von 100° Temperatur in einer Glashülle von 0° per Minute abgegeben wird. Es stimmt diese Zahl mit der oben mit Zuziehung der Correction wegen der Wärmeleitung berechneten gut überein.

Wählt man für das Gesetz der Strahlung die Formel der vierten Potenzen der absoluten Temperaturen, so ist

$$H_1 = AT_1^4, \quad H_2 = AT_2^4$$

**Figure 1.1** The first reference to the fourth degree law – an excerpt from the article of Josef Stefan (Stefan, 1879). Translation: "To obtain the amount of heat which the same surface radiates at a temperature of 100°, this number must be multiplied by $a^{100} = 2.146$ and 1.6644 is obtained. The difference between the two 0.889 gives the amount of heat, which is given off per minute by the unit surface of a glass sphere at a temperature of 100° in a glass envelope at 0°. This number agrees well with the one calculated earlier with the correction for heat conduction. If one chooses the formula of the fourth power of the absolute temperatures for the law of radiation, then $H_1 = AT_1^4$, $H_2 = AT_2^4$."

**Figure 1.2** Example of an infrared image from a thermal imager. The skin has the highest temperature and hence the highest emission.

assumption that electromagnetic radiation is quantized, or composed of discrete packets of energy, rather than being continuous. According to Planck's law, the spectral density of energy emitted by a blackbody is proportional to the frequency of the radiation, and it is also affected by the temperature of the blackbody. This was a revolutionary concept in physics,

which required a revision of most fundamental concepts. The mathematical expression for the Planck's law is:

$$B(\lambda, T) = \frac{2hc^2}{\lambda^5} \frac{1}{e^{\frac{hc}{\lambda kT}} - 1}, \tag{1.3}$$

where $B$ is the spectral radiance, $\lambda$ is the wavelength of the radiation, $h$ is the Planck's constant, $c$ is the speed of light, and $k$ is the Boltzmann constant. It can be shown by integrating Eq. (1.3) across all wavelengths that the Stephan–Boltzmann law is a consequence of the Planck's law.

Wien's displacement law reveals that the peak wavelength of blackbody radiation shifts in a manner inversely proportional to the temperature, indicating that with higher temperatures, the radiation peak moves toward shorter wavelengths, namely

$$\lambda_{\max} T = b, \tag{1.4}$$

where $\lambda_{\max}$ is the wavelength at which the radiation curve has a peak, while $b \approx 2898\ \mu\text{m}\cdot\text{K}$ is the Wien's displacement constant. This fact is derived directly from the Planck radiation law. However, Wilhelm Wien (1864–1928) formulated his law years before Max Planck introduced his more comprehensive formula. In his derivation, Wien was based on the so-called thermodynamic argument.

Wien considered a cavity filled with electromagnetic radiation in thermal equilibrium at a certain temperature. He then imagined adiabatically compressing or expanding the cavity. According to the laws of thermodynamics, such an adiabatic process would change the frequency of the photons inside the cavity (due to the Doppler effect and the change in boundary conditions affecting the standing wave modes of the cavity), but it would not lead to any exchange of heat with the surroundings. The adiabatic compression or expansion of the cavity causes a proportional shift in the frequency (or wavelength) of the radiation inside. Because this process is adiabatic, the frequency shift is directly related to a change in the temperature of the radiation. The functional form of the radiation spectrum must therefore change in a specific way to maintain thermal equilibrium at a new temperature. Wien reasoned that the spectral distribution of radiation in the cavity must scale in a specific way with temperature to satisfy the requirements of thermodynamics. This led him to conclude that the product of the peak wavelength of the radiation and the temperature of the blackbody must be constant. This relationship is now known as Wien's displacement law. Peak wavelengths and temperatures for some objects are shown in Table 1.1.

Finally, in the 1860s, James Clerk Maxwell developed a theory that described the behavior of electric and magnetic fields and their interactions with matter. Maxwell's theory is considered one of the most significant achievements in the history of physics and laid the foundation for much of modern electromagnetism and telecommunications. Maxwell's theory is composed of four partial differential equations that relate the electric and magnetic fields to their sources, which can be electric charges or currents[1]:

---

1 Despite having six unknown components (three each for the electric and magnetic fields), the four Maxwell equations yield eight scalar equations. It has been shown that if a system meets the conditions of Faraday's law and Ampere's law, it will also satisfy both Gauss's laws, provided the system's initial conditions do so and under the assumptions of charge conservation and the absence of magnetic monopoles (Rosen, 1980).

**Table 1.1** Objects, their temperatures, and peak wavelengths of thermal emission.

| Object | $T$ (K) | $\lambda_{\max}$ (μm) |
|---|---|---|
| Cosmic microwave background | 2.725 | 1062.88 |
| Neptune | 59 | 49.15 |
| Earth | 288 | 10.06 |
| Human body | 310 | 9.35 |
| Cool red star (Proxima Centauri) | 3050 | 0.95 |
| Venus | 737 | 3.93 |
| Sun | 5778 | 0.50 |
| Hot blue star (Rigel) | 11000 | 0.26 |
| White dwarf (Sirius B) | 25200 | 0.11 |

- Gauss's Law for Electricity:

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}, \tag{1.5}$$

  where $\mathbf{E}$ is the electric field, $\rho$ is the electric charge density, and $\epsilon_0$ is the permittivity of free space. This equation states that the electric flux out of any closed surface is proportional to the charge enclosed.
- Gauss's Law for Magnetism:

$$\nabla \cdot \mathbf{B} = 0, \tag{1.6}$$

  where $\mathbf{B}$ is the magnetic field. This equation implies that magnetic monopoles do not exist, as the net magnetic flux out of any closed surface is zero.
- Faraday's Law of Electromagnetic Induction:

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}. \tag{1.7}$$

  Here $t$ represents time. This law indicates that a changing magnetic field induces an electric field.
- Ampére's Law with Maxwell's Addition:

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J} + \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}, \tag{1.8}$$

  where $\mu_0$ is the permeability of free space and $\mathbf{J}$ is the current density. This law shows that magnetic fields can be generated by electric currents and by changing with time electric fields.

Maxwell's equations are significant because they unify the fields of electricity and magnetism, which were previously considered to be separate phenomena. They also predict the existence of electromagnetic waves, which are disturbances in the electric and magnetic fields that propagate through space at the speed of light. In principle, Maxwell's equations

give a theoretical framework for describing the radiative transfer phenomena from the basic principles. However, this approach appeared to be very complex from the computational point of view. Therefore, its applications are very limited up to now. In this regard, the radiative transfer theory evolved as a separate discipline in physics.

## 1.2 What Is Radiative Transfer About and What Is It Good For?

Radiative transfer refers to the transfer of electromagnetic radiation (such as light) as it interacts with matter in a medium (such as a gas, liquid, or solid). It is a fundamental process that occurs in many natural and artificial systems, including the Earth's atmosphere, hydrosphere, cryosphere, oceanic and other terrestrial surfaces, stars and galaxies, and many industrial processes. The study of radiative transfer involves understanding how radiation is emitted, absorbed, and scattered by matter, and how it is influenced by factors such as temperature, pressure, and composition. Radiative transfer has many applications in science and engineering. In atmospheric science, radiative transfer models are used to understand the behavior of radiation in the Earth's atmosphere and to study the effects of atmospheric composition and climate change (Liou, 2002b). In astrophysics, radiative transfer models are used to study the properties of stars and galaxies and to understand the evolution of the universe (Rybicki and Lightman, 1985). In industrial processes, radiative transfer models are used to design and optimize energy-efficient systems for heating, cooling, and lighting.

Radiative transfer is also important in many fields of engineering, including aerospace, optics, and materials science. For example, in aerospace engineering, radiative transfer models are used to understand the behavior of radiation in space and to design spacecraft that can withstand extreme temperatures and radiation levels. In optics, radiative transfer models are used to design and optimize optical devices such as lenses and filters and to understand the behavior of light in complex media such as biological tissues. In materials science, radiative transfer models are used to study the thermal properties of materials and to design materials with specific optical and thermal properties (Chen et al., 2016). Table 1.2 summarizes the applications of radiative transfer theory.

**Table 1.2** Applications of radiative transfer theory.

| Field | Context | Applications |
|---|---|---|
| Climate Science and Meteorology | Understanding the Earth's atmosphere's interaction with solar radiation, influencing climate patterns, weather forecasting, and climate change studies. | Global climate models, weather prediction models |
| Astronomy and Astrophysics | Interpreting light from stars and galaxies to deduce compositions, temperatures, motions, and ages; studying star formation and galaxy structures. | Star and galaxy analysis, cosmic background studies |

| | | |
|---|---|---|
| Remote Sensing and Satellite Imaging | Interpreting Earth observation satellite data for mapping surfaces, monitoring vegetation, assessing pollution, and observing planetary bodies. | Land and ocean mapping, environmental monitoring |
| Optical Engineering | Designing optical devices by optimizing light propagation through components like lenses and filters. | Telescope, camera, and laser design |
| Biomedical Applications | Applications in medical imaging technologies (e.g. magnetic resonance imaging, computer tomography scans) and therapeutic interventions (photodynamic therapy, laser surgery) by understanding light-tissue interactions. | Medical diagnostics and treatments |
| Nuclear Reactor Design | Analyzing neutron flux distributions in reactor designs, where radiative transfer principles apply analogously to neutron propagation. | Safety and efficiency optimization in nuclear reactors |

## 1.3 Phenomenological Radiative Transfer

The study of radiative transfer can be approached in two different ways, namely, the phenomenological approach and the microphysical approach.

Radiative transfer theory was initially developed as a phenomenological approach that relied heavily on intuition and experience-based shortcuts (Kravtsov and Apresyan, 1996). It focuses on describing how radiation interacts with matter at a macroscopic level, using empirical data and observations to formulate models and theories. The main assumptions of this approach are the following:

1. The use of the ray approximation, which assumes that radiation propagates in the form of rays and is valid in the limit of geometrical optics;
2. The assumption that rays arriving at a given point are incoherent and can be algebraically summed up without regard for phases or interference terms;
3. The use of time- and space-averaged squared field characteristics instead of local or instantaneous values;
4. Principle of ergodicity, which means that averaging over time for a given realization of a random scattering process is equivalent to ensemble averaging;
5. The medium is assumed to consist of randomly distributed scatterers, and the interaction between the ray and the medium is statistically independent of the outcome of subsequent scattering events, which is known as a Markov process. The particles are randomly oriented.

While the phenomenological approach is advantageous due to its simplicity and efficient numerical models for simulating scattered radiance, it is limited in its accuracy

when the wave properties of light are important. In particular, this approach cannot accurately predict coherent backscattering (Barabanenkov et al., 1991), i.e. constructive interference of multiple scattered paths that have the same length but are reversed in direction. Coherent backscattering results in an enhancement of radiance in the backscattering direction observed in an extremely sharp cone around that direction.

## 1.4 Microphysical Approach

In the microphysical approach, radiative transfer is viewed as the propagation of electromagnetic waves in a discrete medium comprising a set of scatterers. The radiative transfer theory is considered as a consequence of Maxwell's equations (Ishimaru, 1978). As a matter of fact, in the microphysical approach, it is essential to determine conditions that allow us to define the quantities used in the radiative transfer theory and derive the radiative transfer equation using Maxwell's equation as a starting point. A comprehensive review of discussions on microphysical and phenomenological approaches can be found in Mishchenko (2013). Rigorous derivations of the microphysical approach are outside the scope of this book, but can be found elsewhere (Mishchenko, 2008; Doicu and Mishchenko, 2018, 2019). Here, we provide a brief outline of the main ideas.

The medium is assumed to consist of discrete particles; each particle scatters the electromagnetic wave, and in the case of a set of particles, the radiation is scattered multiple times before it is measured by the detector at the observation point $r'$. The total electric field $\mathbf{E}$ at $r'$ is the superposition of the respective incident $\mathbf{E}^{\text{inc}}$ and $N$ partial scattered fields $\mathbf{E}_i^{\text{sca}}$:

$$\mathbf{E}\left(r',t\right)=\mathbf{E}^{\text{inc}}\left(r',t\right)+\sum_{i=1}^{N}\mathbf{E}_i^{\text{sca}}\left(r',t\right). \tag{1.9}$$

Here $\mathbf{E}_i^{\text{sca}}$ corresponds to the wave, scattered by $j$-th particle. For the magnetic field $\mathbf{H}\left(r',t\right)$, the equation looks analogical. To proceed with mathematical derivations, the microphysical approach usually assumes that the particles are randomly and uniformly distributed, and their positions are uncorrelated. While an ensemble average over all position realizations would typically be required, assuming the property of ergodicity allows for time averaging instead. As the scattered wave travels a sufficient distance from the scattering particle, it becomes spherical. The far-field approximation assumes that all particles are in the far-field of each other, thereby making the medium sparse.

Furthermore, the $i$-th particle scatters waves that may have previously been scattered by other particles. Thus, the field that strikes the $i$-th particle reads as

$$\mathbf{E}\left(r_i,t\right)=\mathbf{E}^{\text{inc}}\left(r_i,t\right)+\sum_{\substack{j=1\\ i\neq j}}^{N}\mathbf{E}_j^{\text{sca}}\left(r_i,t\right). \tag{1.10}$$

In each turn, $\mathbf{E}_j^{\text{sca}}$ is a result of scattering of the striking wave $\mathbf{E}\left(r_j,t\right)$. The resulting electric and magnetic fields at the observation point can be expressed as order-of-scattering series (Foldy, 1945; Lax, 1951). There the resulting field is represented as a sum of fields being scattered by all possible sequences of particles. However, Mishchenko (2014) warned

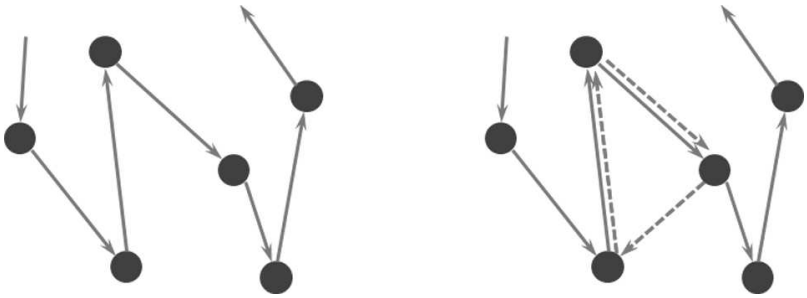

**Figure 1.3** Scattering sequences: (left) self-avoiding scattering path, which is allowed by the Twersky approximation; (right) non-self-avoiding scattering path, which is not allowed by the Twersky approximation. The dashed line marks a closed loop in the trajectory.

against perceiving this scheme as a physical picture and not a mathematical construction since all mutual excitations occur simultaneously.

To evaluate **E** and **H** at the observation point, contributions from all sequences of scattering have to be summed up. To perform analytical summation, the Twersky approximation (Twersky, 1964) is used. Specifically, it is assumed that the radiation follows self-avoiding scattering paths, where particles appear only once. The examples of self-avoiding and non-self-avoiding trajectories are shown in Figure 1.3. In the Twersky approximation, the coherent component of radiation is neglected, and the contributions from different scattering sequences are performed disregarding the phases of waves.

The Poynting vector is a fundamental concept in electromagnetism that describes the flow of energy in electromagnetic waves. It is defined as:

$$\mathbf{S} = \mathbf{E} \times \mathbf{H}, \tag{1.11}$$

In the context of radiative transfer, the Poynting vector describes the flow of energy through a medium that scatters or absorbs radiation. As electromagnetic waves propagate through the medium, they transfer energy to the particles in the medium. The scattered or absorbed radiation can be quantified by integrating the Poynting vector over a surface enclosing the medium. This integration yields the total power that the electromagnetic field conveys to the medium, effectively capturing the dynamic interaction between the electromagnetic waves and the medium at a macroscopic level. Since the Poynting vector **S** oscillates with a period of about $\sim 10^{15}$ seconds at optical frequencies and optical detectors are characterized by a time constant, which is much larger than such a period, the time-averaged value of the Poynting vector is measured. To compute a time-averaged Poynting vector, it is convenient to consider **E** and **H** as complex fields. In this case, the time-averaged Poynting vector can be represented as

$$\langle \mathbf{S} \rangle = \frac{1}{2} \mathrm{Re}\{\mathbf{E} \times \mathbf{H}^*\}, \tag{1.12}$$

where "*" denotes complex conjugate.

When considering the scattering of electromagnetic waves by a collection of particles, the same particle can appear in both the **E** and **H** scattering sequences, since the electric and magnetic fields are scattered independently. To address this, the ladder approximation is employed, which assumes that the orders of common particles are the same in both

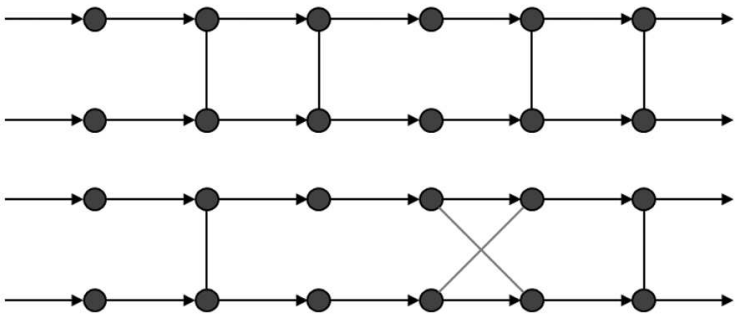

**Figure 1.4** Feynman diagrams: (top) in the case of the ladder approximation, (bottom) with crossing connectors.

sequences. This approximation is valid for systems with a large number of particles, similar to the Twersky approximation. By making these assumptions, it becomes possible to perform an analytical summation of the scattering sequences, leading to the derivation of the radiative transfer equation. This is an important result, as it provides a means of quantifying the transport of energy through a medium in terms of the Poynting vector.

Since **E** and **H** are scattered independently by an ensemble of particles, there may be common particles in the **E** and **H** scattering sequences. Let us consider these particles on the Feynman diagrams (Figure 1.4), wherein the arrows indicate the order of scattering while the vertexes connect the same particles. In the so-called ladder approximation, the orders of common particles are the same in both sequences for **E** and **H**. By making these assumptions, it becomes possible to perform an analytical summation of the scattering sequences, leading to the derivation of the radiative transfer equation. For mathematical details, interested readers are encouraged to refer to recent reviews (Doicu and Mishchenko, 2018, 2019).

## 1.5 Atmospheric Remote Sensing

Remote sensing is the collection of information about an object or environment without direct physical contact. This is achieved through the use of various sensors that capture and measure electromagnetic radiation emitted or reflected by the target. In particular, atmospheric remote sensing is a technique for observing and measuring the physical properties of the Earth's atmosphere from a distance using sensors and instruments mounted on satellites or airplanes. It has a wide range of applications, including weather forecasting, climate research, air quality monitoring, and disaster management.

Satellite remote sensing has revolutionized our ability to observe Earth, providing a comprehensive view that was once impossible. This technology has yielded critical insights into the physical and chemical processes of the atmosphere, impacting areas such as climate change, air quality, and weather forecasting. Tools like NASA Worldview[2] and Google Earth[3] enhance our understanding by offering real-time environmental data and detailed

2 https://worldview.earthdata.nasa.gov/

3 https://earth.google.com/web/

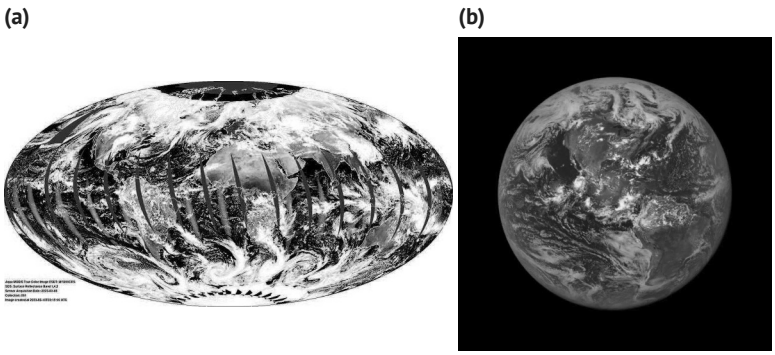


**Figure 1.5** Example of an image (originally in color; rendered in grayscale for print) provided by (a) MODIS Aqua on February 08, 2023. (b) DSCOVR EPIC from the Lagrange point. *Source*: (a) Credits: (a) NASA / Public domain (b) Marshall Sutton / NASA / Public domain.

global views, respectively, making the vast data collected by satellites accessible and usable for experts and the public alike. Additionally, the DSCOVR EPIC camera,[4] positioned at the Lagrangian Point 1 (L1),[5] offers unique perspectives on the Earth's atmosphere and surface, enhancing our understanding of environmental and atmospheric dynamics. As illustrated in Figure 1.5, satellites are indispensable for atmospheric remote sensing, enabling consistent, systematic global coverage. Satellite-based remote sensing of the atmosphere has provided significant insights into the physical and chemical processes occurring in the atmosphere, including those related to climate change, air quality, and weather forecasting.

Satellites used in atmospheric remote sensing are equipped with sensors that measure the interaction of the atmosphere with electromagnetic radiation in different spectral regions. The sensors can measure UV, visible, infrared, and microwave radiation, among other wavelengths. One of the key advantages of satellite-based remote sensing is the ability to collect data over large areas, including remote and inaccessible regions such as the polar regions or oceans. This data can be used to better understand global processes, such as the transport of pollutants or the movement of weather systems. Satellites also provide temporal coverage, allowing for the monitoring of changes in the atmosphere over time. This is particularly important for tracking long-term changes in the atmosphere, such as those related to climate change or the effects of human activity on the environment.

Remote sensing can be passive and active. Passive remote sensing involves measuring the radiation emitted or reflected by the Earth's atmosphere and surface in various spectral regions, such as visible, infrared, and microwave. Examples of passive remote sensing instruments include the Moderate Resolution Imaging Spectroradiometer (MODIS) on NASA's Terra and Aqua satellites and the Atmospheric Infrared Sounder (AIRS) on NASA's

4 https://earthobservatory.nasa.gov/images/86257/an-epic-new-view-of-earth

5 Lagrangian Point 1 is one of five special points in a two-body system (such as the Earth and the Sun) where the gravitational forces of the two large bodies, along with the centrifugal force due to the orbit of a smaller object, balance out. This balance allows a smaller object to remain in a stable or quasi-stable position relative to the two larger bodies. L1 is located on the line connecting the two large bodies (e.g. the Earth and the Sun) and lies between them. For the Earth–Sun system, L1 is about 1.5 million kilometers from Earth, or about 1/100th the distance between the Earth and the Sun.

Aqua satellite. A schematic illustrating passive remote sensing is shown in Figure 1.6. Active remote sensing involves transmitting electromagnetic energy into the atmosphere and measuring the backscattered signals. CALIPSO (Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation)[6] uses lidar technology to analyze the vertical structure of clouds and aerosols, providing essential data for studying their properties and distribution. CloudSat[7] employs radar to penetrate cloud layers and yield detailed information about cloud structure and composition. The scheme of passive remote sensing is depicted in Figure 1.7.

The acquired remote sensing data can then be used to retrieve a variety of information related to the Earth's atmosphere, including temperature profiles, air pressure profiles,

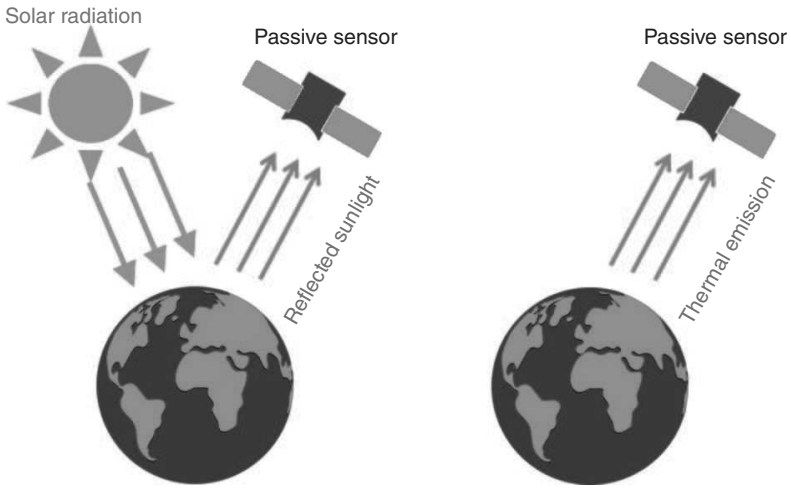


**Figure 1.6** Schematic illustrating passive remote sensing: (left) a sensor onboard a satellite measuring reflected sunlight; (right) a space-based sensor measuring terrestrial thermal emission.

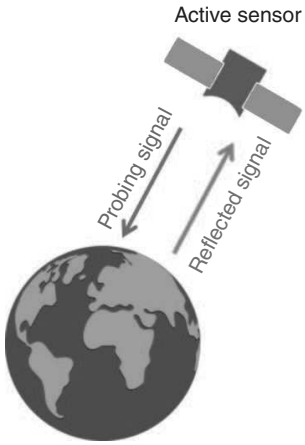


**Figure 1.7** Schematic illustrating active remote sensing.

6 https://www-calipso.larc.nasa.gov/
7 https://cloudsat.atmos.colostate.edu/

water vapor content, cloud properties, aerosol characteristics, and trace gas concentrations. This information is essential for understanding the physical and chemical processes that occur in the atmosphere and how they impact climate, weather, and air quality. Additionally, remote sensing data can be used for monitoring natural disasters, such as wildfires and volcanic eruptions, and for assessing the impact of human activities on the environment, such as air pollution from industrial emissions and deforestation.

In remote sensing, the quantity of interest is not directly measured by the instrument. The received signal undergoes several stages of processing before the necessary information about a certain parameter can be extracted and used in various applications. To describe these processing stages, a standardized set of "levels" has been adopted. Each level represents a sub-process with well-defined inputs and outputs that can be considered independently. The levels are as follows:

**Level 0:** This is the raw, unprocessed data collected by the sensor, usually in the form of counts or volts.

**Level 1:** At this stage, the Level 0 data is converted into energy units (typically radiance).

**Level 2:** The retrieved geophysical parameters, such as aerosol/cloud parameters and trace gas concentrations, are obtained at the same resolution and location as the Level 1 data.

**Level 3:** The Level 2 data is mapped onto uniform space/time grid scales, allowing data from different sources to be combined.

**Level 4:** The final stage involves using the Level 3 data to derive conclusions, results, and other derived data.

## 1.6 Radiative Transfer Models in Atmospheric Remote Sensing

Forward and inverse problems represent fundamental concepts in mathematical modeling and computational sciences, including areas like physics, engineering, remote sensing, and medical imaging. Understanding the distinction between these two types of problems is crucial for developing models and algorithms to interpret data and infer properties of physical systems (see Figure 1.8).

Generally speaking, forward problems start with the causes (the parameters and conditions) and compute the effects (the observables). The forward problem is generally well-posed; given the input parameters, the outcome is predictable and unique. This process is foundational in simulation studies where the aim is to understand how a system behaves under specific conditions. A forward model simulates observables, such as measured spectra, for a given set of model parameters. This approach allows researchers and scientists to

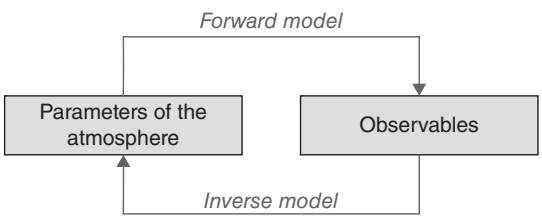


**Figure 1.8** Forward and inverse problems.

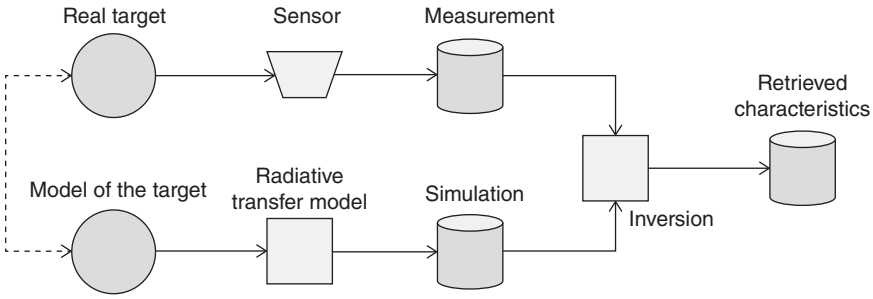


**Figure 1.9** Radiative transfer model in remote sensing.

predict the outcome of a system under specific conditions, effectively turning theoretical model parameters into directly comparable real-world measurements. In remote sensing, forward modeling involves computing the radiances for a given atmospheric state, typically by converting physical parameters of the atmosphere to optical parameters and then using a radiative transfer model to compute the scattered radiance field (Ustinov, 2014). The forward modeling process is used to generate simulated data that can be compared (fitted) to actual measurements to infer properties of the underlying system. The key component of the forward model in remote sensing is a radiative transfer model (RTM) (see Figure 1.9). RTMs are used to calculate the scattered radiance field and thereby estimate the signal received by a sensor.

Inverse problems, on the other hand, seek to determine the unknown parameters or characteristics of a system from observed outcomes or data (Rodgers, 2000; Doicu et al., 2010). Inverse problems are often much more challenging than forward problems because they involve inferring causes from effects. Inversion can be a challenging problem, as it often requires the use of specialized algorithms to solve the inverse problem, and may require additional information or assumptions to be made about the system being studied. This process can be ill-posed, meaning that the solution might not be unique or may not depend continuously on the data; small changes in the observed data can lead to large changes in the inferred parameters. In atmospheric remote sensing, the extraction of information about atmospheric constituents and surface properties is referred to as retrieval.

RTMs can be used to infer the concentrations of various atmospheric constituents such as water vapor, carbon dioxide, ozone, and various atmospheric pollutants. The estimation of aerosol optical properties, including aerosol optical depth, size distribution, and single-scattering albedo, is crucial for understanding aerosol effects on climate, air quality, and visibility. RTMs can be used to retrieve cloud optical thickness, cloud top height, cloud particle size distribution, and cloud phase (liquid, ice) from remote sensing observations. Cloud properties are important for understanding cloud radiative effects and their influence on climate. RTMs can infer surface properties such as surface reflectance, emissivity, and temperature. This information is essential for various applications, including land surface monitoring, climate modeling, and agriculture. To obtain surface imagery without atmospheric effects, it is necessary to perform atmospheric correction (Richter, 1990) by removing the effects of the atmosphere from the signal obtained through remote sensing.

## 1.7 Electromagnetic Spectrum

Taking the curl of Eq. (1.7), and substituting Eq. (1.8) into it, along with the application of the vector identity

$$\nabla \times (\nabla \times \mathbf{E}) = \nabla(\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E}, \tag{1.13}$$

leads us to derive the wave equation for the electric field:

$$\nabla^2 \mathbf{E} - \epsilon_0 \mu_0 \frac{\partial^2 \mathbf{E}}{\partial t^2} = 0. \tag{1.14}$$

A typical solution to Eq. (1.14) is a plane wave expressed as:

$$\mathbf{E}(\mathbf{r}, t) = \mathbf{E_0} \cos(2\pi\nu t - \mathbf{k} \cdot \mathbf{r}), \tag{1.15}$$

where $\mathbf{E}_0$ is the amplitude, $\mathbf{r}$ is the position vector in space, $t$ is the time, $\nu$ is the frequency, and $\mathbf{k}$ is the wave vector, indicating the direction and wavelength of the wave.

In remote sensing applications, these principles of wave propagation are utilized extensively. Various parts of the electromagnetic spectrum are used to gather data about the Earth's surface and atmosphere. Each segment of the spectrum offers distinct characteristics and is suitable for different applications, as detailed in Table 1.3. At one end of the spectrum are the high energy, short wavelength gamma rays, and X-rays, which are typically not used in terrestrial remote sensing due to atmospheric opacity. These high-energy wavelengths can be useful for studying crystal structures and surface states in laboratory measurements, or for studying planetary surfaces with no atmosphere, such as the Moon. X-rays are used for medical diagnostics. One of the primary reasons for the widespread use

**Table 1.3** Electromagnetic spectrum regions.

| Name | Wavelength | Frequency | Photon energy |
|---|---|---|---|
| Gamma ray | 0.01 nm | 30 EHz | 124 keV |
| X-ray | 0.01–10 nm | 30 PHz–30 EHz | 1.24 eV–124 keV |
| Extreme ultraviolet | 10–120 nm | 2.5–30 PHz | 10.35–124 eV |
| Far ultraviolet | 120–200 nm | 1.5–2.5 PHz | 6.21–10.35 eV |
| Middle ultraviolet | 200–300 nm | 1.0–1.5 PHz | 4.14–6.21 eV |
| Near ultraviolet | 300–400 nm | 750 THz–1.0 PHz | 3.1–4.14 eV |
| Visible | 400–700 nm | 420–750 THz | 1.77–3.1 eV |
| Near-infrared | 700–3500 nm | 86–420 THz | 0.35–1.77 eV |
| Thermal infrared | 3.5–100 $\mu$m | 3–86 THz | 0.012–0.35 eV |
| Microwaves | 1 mm–1 m | 300 MHz–300 GHz | 1.24 $\mu$eV–1.24 meV |
| Radio waves | 1 mm–10000 km | 30 Hz–300 GHz | 1.24 meV–124 feV |
| Extremely low frequency | 10000 km | 30 Hz | 12.4 feV |

of X-rays in medical imaging is their ability to penetrate the body without scattering, thus producing highly detailed and focused images. While X-rays can be absorbed by certain materials, their ability to travel through soft tissue and bone allows them to provide a clear and sharp view of the internal structures of the body.

Going further along the spectrum, we find the ultraviolet (UV) region, which can be divided into three subregions: extreme UV, far UV, and near UV. Extreme UV is generated by the solar corona and is highly absorbed in air, ionizing $O_2$ and $N_2$, while far UV is absorbed by $O_2$ and middle UV is mostly absorbed by the ozone layer, which is located around 25 km above sea level. The near UV can reach Earth's surface and can have harmful effects on human health, causing skin damage, cataracts, immune system suppression, and genetic damage, resulting in skin cancer. However, moderate sun exposure is important for human health, as it stimulates the synthesis of vitamin D in the skin.

The visible range (400–700 nm) is the part of the electromagnetic spectrum that humans can see. It is divided into seven subregions: violet, indigo, blue, green, yellow, orange, and red. This region is used extensively in imaging techniques, mapping, and multispectral photography. Clouds and aerosols have a strong impact on the spectrum in this region, and this feature can be used for both passive and active remote sensing techniques.

The near-infrared (NIR) region, which spans from approximately 700 to 1400 nm, contains strong molecular absorption bands of water vapor, oxygen, and carbon dioxide. The solar light reflectance in the gaseous absorption bands can be used for various applications, including the estimation of cloud/aerosol layer altitudes. The thermal infrared (TIR) region is used for thermal imaging and remote sensing of land and surface water temperatures. The emission of thermal radiation is an important mechanism of radiation cooling, and so, this part of the spectrum has a strong impact on the global climate.

At the other end of the spectrum, we find the microwave region, which spans from approximately 1 mm to 1 m in wavelength. Microwave radiation can penetrate through clouds, making it useful for remote sensing applications in all weather conditions and also at night. Passive microwave remote sensing detects the naturally emitted microwave energy, which is small compared to the optical region. The field of view of the instrument should be large enough so that the spatial resolution of the microwave sensors is lower than those operated in the visible region. Active microwave sensors can be imaging and non-imaging, with imaging sensors called RADAR (radio detection and ranging), which can be used for imaging at any time, day or night, and in most weather conditions. The time between the transmitted and reflected signals, and the corresponding phase difference, determines the distance to the target. Non-imaging sensors, such as the Meteorological Temperature Profiler MTP-5, serve the same purpose.

In our book, we consider radiative transfer models that are typically used in the UV, VIS, NIR, and SWIR regions of the electromagnetic spectrum. These spectral regions are used extensively in remote sensing applications, and understanding the unique characteristics and applications of each region is crucial for developing effective remote sensing methods. The spectra at the top of the atmosphere and at the sea level are different due to the attenuation of radiation as it passes through the atmosphere, as shown in Figure 1.10. The top of the atmosphere spectrum is the amount of energy per unit area per unit of time that is incident on the outermost layer of the Earth's atmosphere. This spectrum includes all types of radiation, such as ultraviolet, visible, and infrared radiation, as well as X-rays and gamma

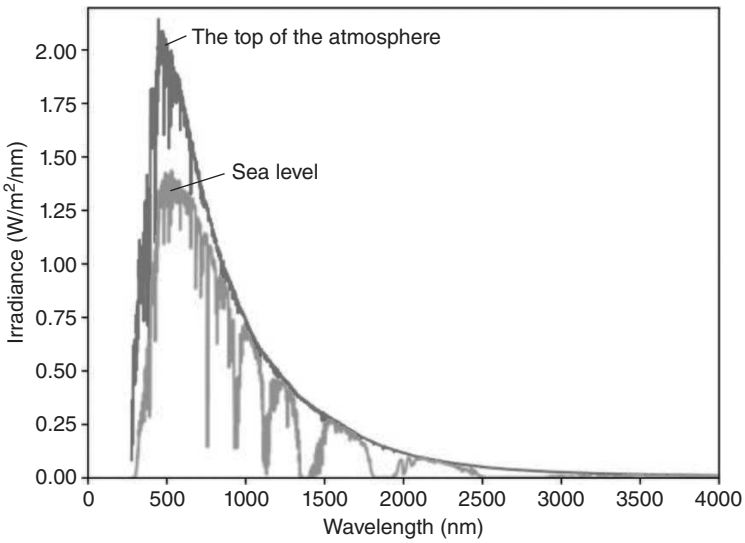


**Figure 1.10** Examples of solar spectra at the top of the atmosphere and at sea level.

rays. As the radiation travels through the atmosphere, it encounters various molecules such as oxygen, water vapor, ozone, carbon dioxide, and aerosols, which can scatter and absorb the radiation. As a result, the spectrum of radiation at sea level is different from the spectrum at the top of the atmosphere. The amount of absorption and scattering varies with wavelength, so different regions of the electromagnetic spectrum are affected differently. For example, ultraviolet radiation is mostly absorbed by ozone in the upper atmosphere, so very little ultraviolet radiation reaches the Earth's surface. In contrast, visible light is not significantly absorbed or scattered, so it can penetrate the atmosphere and reach the surface relatively unimpeded. The essential portion of infrared radiation, on the other hand, is absorbed by water vapor and carbon dioxide, which contributes to the greenhouse effect and can lead to warming of the lower atmosphere. Radiative transfer models can help to predict how the spectra will be modified as radiation passes through different layers of the atmosphere, taking into account factors such as the composition and density of atmospheric gases, as well as the presence of aerosols and clouds.

## 1.8 Why Do We Need Analytical Models in Radiative Transfer?

Recent developments in optics, sensor design, and measurement techniques have greatly enhanced the features of remote sensing sensors, in particular, the spatial resolution of the measurements. The spatial resolution of remote sensing images refers to the smallest object that can be detected on the Earth's surface. Advances in optics, sensor technologies, and measurement techniques have led to a significant reduction in the size of the smallest detectable feature. This improvement in spatial resolution means that finer-scale changes can be monitored over time, enhancing applications in urban planning, agriculture, forestry, disaster management, and environmental monitoring.

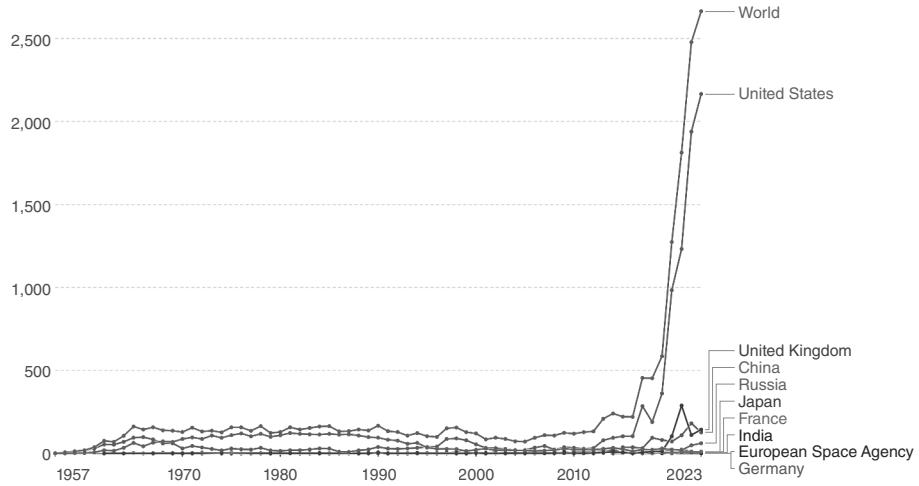


**Figure 1.11** Annual number of objects launched into space. *Source*: Data taken from Mathieu and Roser (2022) (licensed under CC BY 4.0).

There is a significant change taking place in the industry regarding the number of satellites and their market value. There is a forecasted increase of 4.5 in satellite demand over the next 10 years, which means an average of about 1800 satellites will be launched each year, compared to about 400 over the last decade.[8] According to the Union of Concerned Scientists (UCS) Satellite Database,[9] the number of satellite launches increases exponentially, as shown in Figure 1.11.

The high volume of data generated by remote sensing can be overwhelming, making it challenging to analyze and interpret. Traditional approaches are often inadequate to deal with large amounts of data, as the amount of satellite data is growing faster than the computational power (Ma et al., 2015). The performance bottleneck in the processing chain is radiative transfer calculations that are computationally expensive and time-consuming. The remote sensing data is recognized as Big Data (Liu, 2015) since it satisfies Doug Laney's 3V criterion (Patgiri and Ahmed, 2016): significant growth in volume, velocity, and variety. Volume refers to the physical amount of data, velocity to the speed at which the data is created, and variety to the different forms of data. New computational paradigms are required to cope with these high-efficiency requirements, including integrating data from different sources in various locations, distributing data across massively parallel platforms for processing, and providing fast access to remote sensing data to users worldwide.

To address this challenge, analytical models can provide useful information while reducing the computational burden. Approximate models are simplified versions of the original

8 https://digital-platform.euroconsult-ec.com/product/satellites-to-be-built-launched-new/
9 https://www.ucsusa.org/resources/satellite-database

models that capture the essential features of the physical system. They are designed to provide a reasonable approximation of the physical model while reducing the complexity and computational requirements. The so-called surrogate models can be applied (Brence et al., 2022). Essentially, a surrogate model is a computationally efficient approximation of a complex physical model or a function. This method is particularly beneficial for critical applications, as it accelerates the simulation process, freeing up more computational time to explore a larger parameter space or improve the accuracy of methods through more iterations. Approximate radiative transfer models can provide a relatively fast and accurate approximation of the radiation field, which can then be used to derive insights and understanding of the problem (Doicu et al., 2022). Moreover, they can provide a solution in an analytical form, making it easier to analyze and understand the behavior of the radiation field under different conditions, such as varying atmospheric parameters or surface properties. This can be particularly important in fields such as climate science or astrophysics, where small changes in atmospheric or environmental conditions can have significant effects on the radiation field. Some approximate models can significantly improve the accuracy of the dimensionality reduction and machine learning models, as discussed further in this book.

## 1.9 Radiative Transfer and Climate Modeling

Radiative transfer plays a fundamental role in climate modeling by simulating the transfer of electromagnetic radiation through the Earth's atmosphere and surface. Climate models utilize RTMs to understand how incoming solar radiation, outgoing terrestrial radiation, and energy exchanges within the atmosphere influence Earth's climate system. The Earth's surface absorbs solar radiation and re-emits it as longwave (infrared) radiation. Radiative transfer models calculate the absorption and emission of terrestrial radiation by greenhouse gases such as water vapor, carbon dioxide, methane, and ozone. These models also consider the scattering and absorption of infrared radiation by clouds and aerosols. Understanding the balance between incoming solar radiation and outgoing terrestrial radiation is crucial for determining Earth's energy budget and establishing the global temperature distribution. Radiative transfer models simulate the greenhouse effect, which is the process by which certain atmospheric gases trap heat and warm the Earth's surface. These models quantify the radiative forcing exerted by greenhouse gases, which is the imbalance between incoming solar radiation and outgoing terrestrial radiation caused by the presence of these gases. By incorporating information on greenhouse gas concentrations, radiative properties, and atmospheric dynamics, climate models can assess the magnitude of the greenhouse effect and its impact on global temperature changes.

Radiative transfer models also account for feedback mechanisms that can amplify or dampen climate change. Climate models simulate these feedback processes to project future climate scenarios under different greenhouse gas emission scenarios and assess the uncertainties associated with climate projections. Two distinct types of feedback loops can be identified: positive and negative. Positive feedback amplifies changes, driving systems away from equilibrium and often leading to exponential growth or instability. In contrast, negative feedback mitigates changes, promoting balance and stability by counteracting deviations and steering systems back toward equilibrium. Positive feedback mechanisms within

the atmosphere are those that magnify initial alterations in the climate system, thereby intensifying either warming or cooling trends. Conversely, negative feedback mechanisms serve to mitigate climate perturbations, often restoring the climate system to its original state by counteracting the initial changes. The concept of positive and negative feedback loops in atmospheric science is closely linked to the work of early climate scientists and systems theorists. One of the most notable figures in formalizing these ideas within the context of climate dynamics is Hansen et al. (1981). His work on the effects of increasing greenhouse gases emphasized the role of feedback mechanisms in climate systems, especially how positive feedback loops could amplify global warming. Several feedback mechanisms are important for climate modeling, especially in the context of analysis of the robustness of the climate system, including:

**Water vapor feedback:** The atmosphere warms due to increased greenhouse gas concentrations. This leads to higher evaporation rates from the oceans and other water bodies. Since water vapor is itself a greenhouse gas, the increased atmospheric water vapor content further enhances the greenhouse effect, causing additional warming. This feedback loop amplifies the initial temperature increase, contributing to enhanced global warming (Hall and Manabe, 1999).

**Cloud feedback:** The atmosphere warms due to increased greenhouse gas concentrations. This leads to higher evaporation rates from the oceans and other water bodies. Since water vapor is itself a greenhouse gas, the increased atmospheric water vapor content further enhances the greenhouse effect, causing additional warming. This feedback loop amplifies the initial temperature increase, contributing to enhanced global warming (Hall and Manabe, 1999).

**Ice-albedo feedback:** The reflectivity of the Earth's surface, known as albedo, plays a crucial role in regulating the planet's temperature. Ice and snow have high albedo, meaning they reflect a significant portion of incoming solar radiation back to space. However, as temperatures rise, ice and snow cover decrease, exposing darker surfaces such as ocean or land. This reduces the overall albedo, leading to increased absorption of solar radiation and further warming. The loss of Arctic sea ice and the melting of glaciers are examples of ice-albedo feedback mechanisms that contribute to amplified warming in polar regions (Budyko, 1969). The darkening of ice and snow surfaces due to different pollutants, including dust and black carbon, contributes to the global albedo reduction as well (Hansen and Nazarenko, 2003; Flanner et al., 2007).

**Permafrost feedback:** Permafrost, frozen soil found in Arctic and subarctic regions, contains vast amounts of organic matter that can decompose and release greenhouse gases like methane and carbon dioxide when thawed. As temperatures rise, permafrost thaws, accelerating the decomposition of organic material and releasing additional greenhouse gases into the atmosphere. This process further enhances the greenhouse effect, leading to more warming and additional permafrost thawing (Schuur et al., 2015) – a positive feedback loop.

**Vegetation feedback:** Changes in temperature and precipitation patterns can affect vegetation growth, distribution, and productivity. For example, warmer temperatures may increase the rate of photosynthesis and plant growth in some regions, leading to greater carbon uptake by vegetation. However, in other cases, warming and changes in

precipitation may lead to drought stress, wildfires, and dieback of forests,[10] reducing carbon sequestration capacity. These changes in vegetation cover and carbon cycling can influence atmospheric carbon dioxide levels, contributing to positive feedback in the climate system (Cheng et al., 2021).

**Water vapor and lapse rate feedback:** While increased atmospheric temperatures can lead to higher water vapor content due to enhanced evaporation, this also leads to greater cloud formation. Clouds have both warming (by trapping outgoing longwave radiation) and cooling (by reflecting incoming solar radiation) effects. Additionally, as the atmosphere warms, the lapse rate (the rate of temperature decreases with altitude) may change, which can influence cloud formation and distribution. Changes in cloud cover and the lapse rate can act as negative feedback, ultimately counteracting some of the initial temperature increases.

**Cloud albedo effect:** Changes in cloud properties, such as cloud cover, altitude, and particle size, can influence the amount of solar radiation reflected back to space. Low-level clouds, such as stratocumulus, have a cooling effect because they reflect a significant portion of incoming solar radiation. As temperatures rise, increased cloud cover or changes in cloud properties may enhance this cooling effect, offsetting some of the initial warming (Mueller et al., 2011).

## 1.10 Remote Sensing of Trace Gases

Remote sensing of trace gases is a critical area of study in atmospheric science, environmental monitoring, and climate research. It involves the detection and quantification of gases present in the atmosphere in small quantities, which, despite their low concentrations, can have significant impacts on the environment and human health. The concentrations of different gases in the Earth's atmosphere are listed in Table 1.4.

Despite the relatively low concentrations of trace gases in the atmosphere, they can be effectively detected by analyzing the spectrum of electromagnetic radiation at specific wavelengths where the absorption by these gases is particularly strong. Each gas is characterized by a unique set of absorption lines, often referred to as its spectral "fingerprints." These unique patterns are critical in identifying and quantifying different gases, as they allow for precise discrimination between various atmospheric components.

To illustrate this, let's conduct a thought experiment. Consider the passage of light through the atmosphere, but imagine that the atmosphere consists solely of one gas. Then, depending on the wavelength, a different fraction of radiation will pass through for each gas. This fraction is called the transmittance. The corresponding transmittance spectra for four gases are shown in Figure 1.12. Regions with transmittance close to 1 are called atmospheric windows because the atmosphere is transparent at these wavelengths, allowing radiation to pass through with minimal absorption. Conversely, regions where the transmittance is significantly lower indicate wavelengths that are strongly absorbed by the atmospheric gases.

10 The phenomenon of a stand of trees losing health and dying without an obvious cause.

**Table 1.4** Gaseous concentrations in the Earth's atmosphere.

| Gas | Concentration | Comments |
|---|---|---|
| Nitrogen ($N_2$) | 78.08% | Most abundant atmospheric gas, inert |
| Oxygen ($O_2$) | 20.95% | Essential for respiration and combustion |
| Argon (Ar) | 0.93% | Inert |
| Carbon dioxide ($CO_2$) | 420 ppm | Greenhouse gas, important for photosynthesis |
| Water vapor ($H_2O$) | Variable | Critical for weather systems, greenhouse gas |
| Neon (Ne) | 18 ppm | Inert |
| Helium (He) | 5.2 ppm | Light, inert |
| Methane ($CH_4$) | 1.8 ppm | Potent greenhouse gas affects climate change |
| Krypton (Kr) | 1 ppm | Inert |
| Hydrogen ($H_2$) | 0.5 ppm | Lightest gas |
| Xenon (Xe) | 0.09 ppm | Inert |
| Ozone ($O_3$) | 0.1–10 ppb | Absorbs UV radiation, pollutants at ground level |

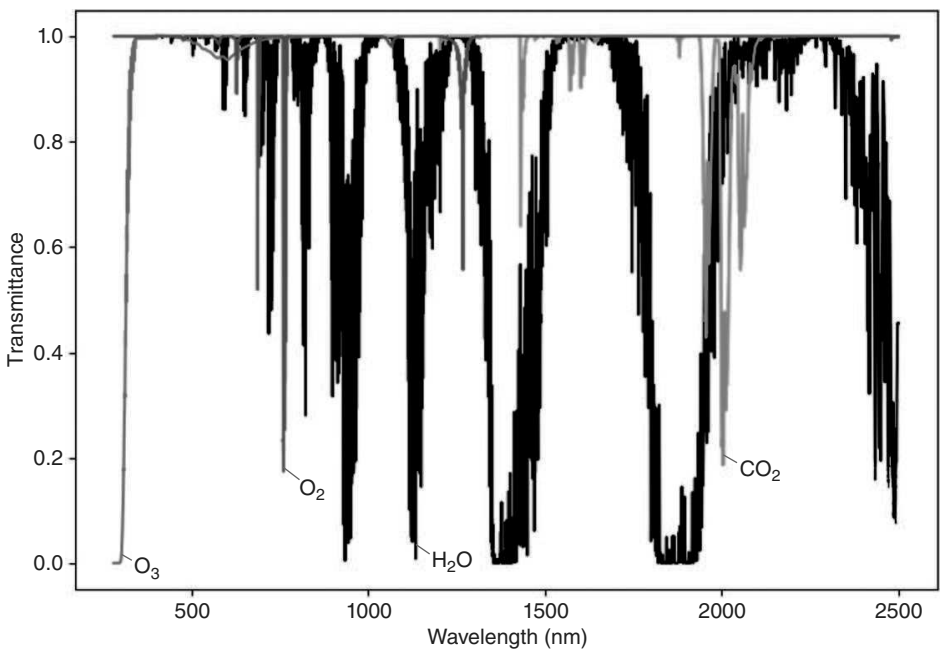


**Figure 1.12** Transmittances for different gases in the spectral range 400–2500 nm.

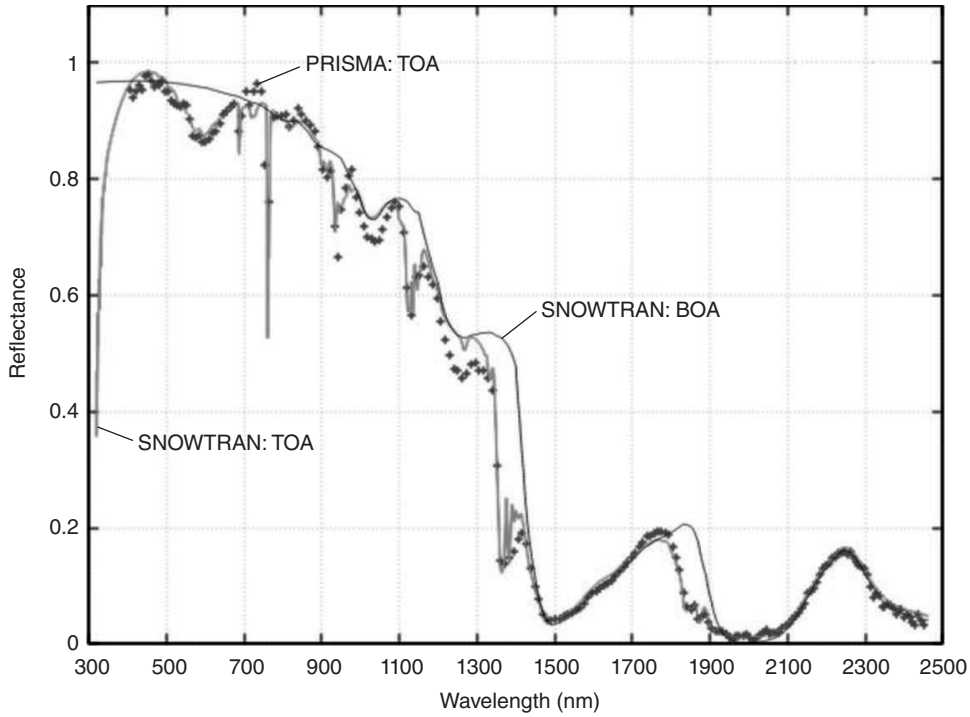


**Figure 1.13** Measured PRISMA top-of-atmosphere (TOA) spectral reflectance (crosses) for a solar zenith angle of 57.6° and nadir observation over the underlying snow surface in Antarctica. The lines show results from theoretical modelling of TOA and bottom of atmosphere (BOA) reflectances with a spectral resolution of 1 nm using SNOWTRAN software. *Source:* Adapted from Kokhanovsky et al. (2024) (licensed under CC BY 4.0).

Measurements taken at characteristic wavelengths are extremely sensitive to the presence of specific gaseous species. This sensitivity enables the detection and analysis of trace gases even at minute concentrations, making it an invaluable tool in environmental monitoring, atmospheric chemistry, and climate science. The technique not only helps in tracking pollution and studying atmospheric conditions but also plays a crucial role in monitoring changes related to global warming and other environmental issues.

We show the top-of-atmosphere reflectance measured by PRISMA hyperspectral spaceborne instrument[11] on December 21, 2022 (00:14UTC) at the point with coordinates (76.14N, 129.93E) over underlying snow surface in Figure 1.13. The absorption of reflected light by various molecules ($O_2$, $O_3$, $H_2O$) is clearly identified. The depth of absorption lines as seen in the reflected light can be used to retrieve the total columns of respective gases.

11 https://www.eoportal.org/satellite-missions/prisma-hyperspectral

## 1.11 Remote Sensing of Clouds

Clouds, composed of microscopic water droplets and/or ice crystals, play a crucial role in Earth's hydrological cycle and climate system. They significantly influence radiation processes, atmospheric dynamics, chemistry, and thermodynamics. Clouds play a crucial role in weather prediction and climate models. Understanding their types and behaviors helps meteorologists predict weather changes and climatologists understand the Earth's climate system.

Clouds are classified into various types based on their appearance, altitude, and the physical processes that lead to their formation. Table 1.5 provides an overview of common cloud types alongside typical models used to study them. Clouds can be classified based on their altitude, which also corresponds to differences in cloud top pressure and thickness. These classifications include:

1. **High-level clouds** (5–13 km): Typically found at altitudes of 5–13 km, these clouds include types such as cirrocumulus, cirrus, and cirrostratus, which are composed mainly of ice particles.
2. **Mid-level clouds** (2–7 km): Existing between 2 and 7 km, these clouds include altocumulus, altostratus, and nimbostratus, which may be composed of water droplets, ice particles, or a mixture of both.
3. **Low-level clouds** (0–2 km): Found at altitudes up to 2 km, these include stratus, cumulus, cumulonimbus, and stratocumulus, which are generally water cloud types.

It is important to note that these classifications are not universal, and different meteorological sources may use alternative schemes based on regional climatic conditions and research needs.

The optical properties of clouds are primarily determined by the characteristics of the particles they are composed of. This includes the particle size, shape, and the complex refractive index, which affect how light interacts with cloud particles.

**Table 1.5** Common cloud types and corresponding models.

| Cloud type | Description | Model used |
|---|---|---|
| Cumulus | Fluffy, heap-like clouds with flat bases | Cumulus parameterization models |
| Stratus | Layered, uniform clouds covering the sky | Stratiform cloud models |
| Cirrus | Thin, wispy clouds at high altitudes | High-altitude cloud models |
| Nimbus | Rain-bearing clouds that are thick and dark | Precipitation models |
| Cumulonimbus | Tall, dense clouds with significant vertical growth | Severe weather models |

Water clouds are often modeled as collections of spherical water droplets, adhering to specific size distributions. One common model used is the modified gamma distribution, represented mathematically as follows (Deirmendjian, 1969):

$$f(r) = ar^{\alpha}e^{-br^{\gamma}}, \tag{1.16}$$

where $a$, $\alpha$, $b$, and $\gamma$ are parameters that define the shape of the size distribution. This distribution helps in describing the variability in droplet sizes within a cloud.

Cirrus clouds contain nonspherical ice particles. This introduces significant computational challenges in simulating how radiation interacts with these clouds. Nonspherical particles require more complex models to accurately represent their interaction with light, impacting the accuracy of remote sensing data and climate models (Yang and Liou, 2016). Retrieving cloud properties using the molecular oxygen absorption bands is a critical technique in remote sensing, especially in atmospheric and environmental sciences. This approach utilizes the strong absorption features of molecular oxygen in the optical range to assess cloud characteristics such as cloud top pressure, cloud cover, and optical thickness. Several Earth-observing satellites and instruments are equipped to analyze the $O_2$ A-band for the retrieval of cloud properties, such as MODIS, GOME-2, and OMI. These instruments provide global data that are crucial for daily weather forecasting, climate modeling, and atmospheric research.

Retrieving cloud properties using the $O_2$ A-band typically involves analyzing the changes in the depth and shape of the absorption band as light passes through the atmosphere and is reflected back by clouds. This technique is based on the principle that the amount of light absorbed by oxygen varies with the path length:

- **Spectral analysis:** Instruments measure the spectral characteristics of sunlight that has traveled through the atmosphere and reflected off the cloud tops. The absorption features in the $O_2$ A-band are analyzed to determine the path length of light.
- **Cloud top pressure:** The amount of absorption in the $O_2$ A-band can be used to estimate cloud top pressure. Light interacting with higher clouds (at lower pressures) will show different absorption characteristics compared to light interacting with lower clouds (at higher pressures).
- **Optical thickness and cloud cover:** By analyzing the intensity and width of the $O_2$ A-band, researchers can infer the cloud optical thickness and overall cloud cover. Thicker clouds and denser cloud cover affect the absorption features differently compared to thin or sparse clouds.

Figure 1.14 illustrates the variations in radiance at the top of the atmosphere within the $O_2$ A-band in response to changes in cloud optical thickness and cloud top height, with a consistent cloud geometrical thickness set at 1 km. The wings of the band are predominantly sensitive to variations in cloud optical thickness, whereas the central portion of the band is more responsive to changes in the cloud's vertical position.

Radiative transfer simulations that incorporate scenarios with clouds can benefit from considering the medium as optically thick. In such cases, the cloud acts as a barrier that obscures the underlying surface, along with the spectral signatures of trace gases located beneath it. Often, clouds can be approximated as diffuse reflective surfaces, simplifying the computational models used to analyze and predict their impact on observed radiation from above.

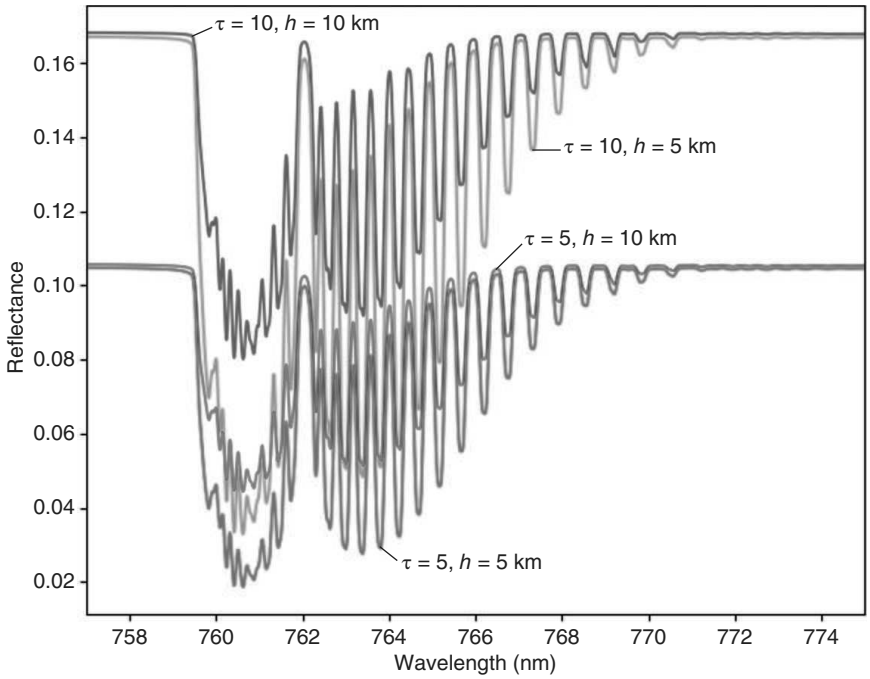


**Figure 1.14** TOA radiance in the $O_2$ A-band for different values of cloud optical thickness $\tau$ and cloud top height $h$. The spectra are convolved with a slit function that has a full width at half maximum of 1 nm. The viewing zenith angle is 40°, and the solar zenith angle is 0°. The computations are performed using the PYDOME radiative transfer model (Efremenko et al., 2023).

## 1.12 Remote Sensing of Atmospheric Aerosol

Atmospheric aerosol is composed of liquid and solid microscopic particles with average radii in the range of 0.1–3 μm. These tiny particles are responsible for cloud formation and have important climatic impacts. The aerosol optical thickness is usually small (in the range 0.1–0.5) in the visible spectrum, with smaller values in the polar regions and extremely large values during heavy forest fires, volcanic eruptions, and dust storms. The single-scattering albedo is close to 1 in the visible range in the majority of cases. The parameters of scientific interest are particle size distributions (e.g. in fine and coarse modes) and also complex spectral refractive index of particles, the columnar concentration of particles, the position of aerosol layer in atmosphere, fine particulate matter at the ground ($PM_{2.5}$), spectral aerosol optical thickness, and single-scattering albedo. We show the global spatial variation of the parameter $PM_{2.5}$ derived from satellite top-of-atmosphere measurements before and during COVID-19 pandemic in Figure 1.15. The decrease in the value of fine particulate matter with diameters below 2.5 μm at the ground surface is clearly seen. It is related to the decline in industrial production during COVID-19 pandemic. This also means that cloud nuclei concentrations have been reduced, which have reduced cloud cover and resulted in heating of the Earth system with important climatic impacts due to both atmospheric aerosol and clouds.

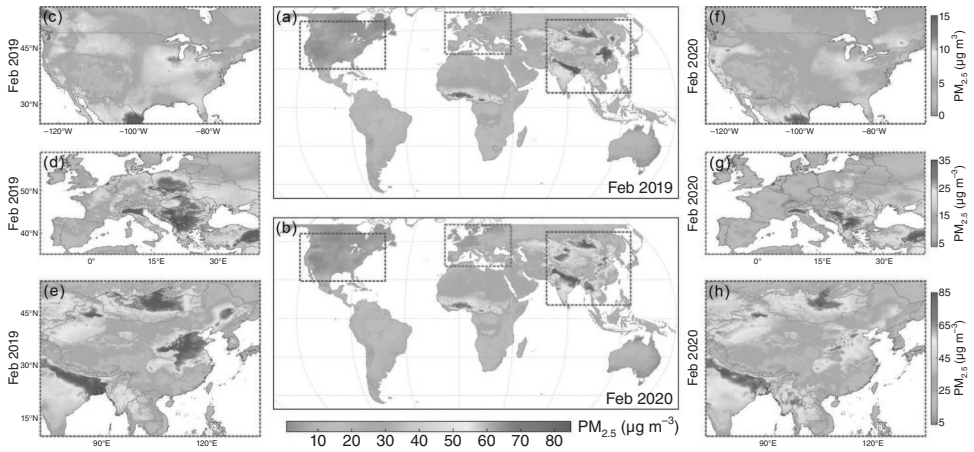


**Figure 1.15** Global distribution of fine particulate matter at the surface before and during the COVID-19 pandemic. *Source*: Adapted from Bai et al. (2024) (licensed under CC BY 4.0).

**Table 1.6** Types of aerosols, their sources, and radiative effects.

| Type | Source | Radiative effect |
|---|---|---|
| **Oceanic aerosol** | Sea spray from the ocean surface, salt particles, and organic matter | Primarily negative forcing: reflects sunlight |
| **Dust aerosol** | Wind-blown dust from deserts, arid regions, and land surface disturbances | Mixed forcing: reflects and absorbs sunlight |
| **Volcanic aerosol** | Eruptions release sulfur dioxide ($SO_2$), forming sulfate aerosols | Strong negative forcing: reflects sunlight |
| **Smoke aerosol** | Biomass burning, forest fires, and human activities like agricultural burning | Mixed forcing: absorbs and reflects sunlight |
| **Biogenic aerosols** | Emissions from plants, trees, and other biological processes | Negative forcing: reflects sunlight |
| **Background water-soluble aerosol** | Naturally occurring water droplets and dissolved compounds in the atmosphere | Negative forcing: reflects sunlight |

Maps, as shown in Figure 1.15, can be derived from the analysis of reflected solar light intensity and polarization in different regions of the planet with the use of radiative transfer models to be discussed in this work. This underlines the importance of the topic considered in this book.

The atmospheric aerosol is usually classified in terms of its origin. Their radiative effect is also different. Some types of aerosols together with their effects are summarized in Table 1.6.

Usually it is assumed that there are two fractions of atmospheric aerosol: fine mode and coarse mode aerosol fractions, which are characterized by the lognormal volume particle size distribution with the median radius in the range 0.12–0.15 μm for the fine mode and 1.9–3.3 μm for the coarse mode (Dubovik et al., 2002). The best retrievals of aerosol properties are achieved if multi-angular spectropolarimetry is used (Dubovik et al., 2011; Kokhanovsky et al., 2015). The aerosol top height position can be derived from lidar measurements or oxygen A-band spectrometry (Jänicke et al., 2023; Huang et al., 2015).