Inversion of the Earth's Bond albedo from space geodesy

Olli Ihalainen

Supervisors: Karri Muinonen
Olli Wilkman

Censors: Karri Muinonen
Thomas Hackman
Antti Penttilä

University of Helsinki
Department of Physics

Master's Thesis
Planetary Astronomy
Theoretical and Computational Methods

November 2019

Tiivistelmä - Referat - Abstract

The Earth's Bond albedo is the fraction of total reflected radiative flux emerging from the Earth's Top of the Atmosphere (ToA) to the incident solar radiation. As such, it is a crucial component in modeling the Earth's climate. This thesis presents a novel method for estimating the Earth's Bond albedo, utilising the dynamical effects of Earth radiation pressure on satellite orbits that are directly related to the Bond albedo. Where current methods for estimating the outgoing reflected radiation are based on point measurements of the radiance reflected by the Earth taken in the proximity of the planet, the new method presented in this thesis makes use of the fact that Global Positioning Satellites (GPS) together view the entirety of the ToA surface.

The theoretical groundwork is laid for this new method starting from the basic principles of light scattering, satellite dynamics, and Bayesian inference. The feasibility of the method is studied numerically using synthetic data generated from real measurements of GPS satellite orbital elements and the imaging data from the Earth Polychromatic Imaging Camera (EPIC) aboard the Deep Space Climate Observatory (DSCOVR) spacecraft.

The numerical methods section introduces the methods used for forward modeling the ToA outgoing radiation, the Runge-Kutta method for integrating the satellite orbits and the virtual-observation Markov-chain Monte Carlo methods used for solving the inverse problem. The section also describes a simple clustering method used for classifying the ToA from EPIC images.

The inverse problem was studied with very simple models for the ToA, the satellites, and the satellite dynamics. These initial results were promising as the inverse problem algorithm was able to accurately estimate the Bond albedo. Further study of the method is required to determine how the inverse problem algorithm works when more realism is added to the models.

Table of contents

Chapter 1

Introduction

Defined as the fraction of the total reflected incoming radiation to the incident solar radiation, the Earth's *Bond albedo*, also called the *spherical albedo*, is a crucial component of the Earth's radiation budget (ERB), along with the total outgoing emitted radiation. This energy budget sets critical constraints to climate models, making accurate estimates of the ERB of vital importance for studying the impact of anthropogenic forcing on the Earth system. As it stands, the primary means of estimating the ERB is based on point measurements of the outgoing radiance, taken in the proximity of the planet by the CERES instrument (Clouds and Earth's Radiant Energy System) [1, 2]. This measurement setting gives rise to uncertainties at the level of 1% to 3% [3, 4], that are difficult to evaluate, since, for example, the cloud cover evolves in a minute timescale while it takes several hours for the CERES instruments to scan the entire Earth system.

To significantly lower the uncertainties, this thesis introduces a novel method for estimating the Earth's Bond albedo in near-real time. The principal idea is to collect and to forward model radiometric imaging data from remote sensing satellites in tandem with simultaneous high precision space-geodetic observations of the satellite positions, which are affected by the emerging Earth radiation pressure. Using the forward model, the time evolution of the ERB along with uncertainties can then be resolved as a global statistical inverse problem. Ideally, an armada of satellites capable of simultaneous space geodetic observations and radiometric measurements would be used. But since such an armada does not yet exist, this thesis focuses on the options currently available. Promising candidates in this category are Global Positioning System satellites, the constellations of which are relatively homogeneously spread around the Earth (Fig. 1.1), and the Earth Polychromatic Imaging Camera (EPIC) aboard the Deep Space Climate Observatory (DSCOVR) spacecraft orbiting at the L1 point from where it always views almost the entire sunlit side of the Earth's surface (Fig. 1.2. Although the basis for utilising both radiometric and space geodetic

observations to solve the global inverse problem is laid in this thesis, only the space geodesy part of the inverse problem is explored in practice.
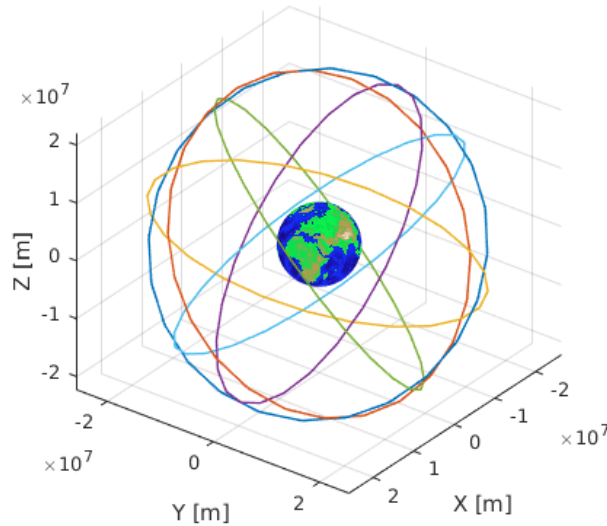
Fig. 1.1 GPS constellations.

To stress the importance of accurately estimating the ERB, nearly all of the radiant energy incident on the Earth, which drives the planet's climate and weather, comes from the Sun, with the average amount of solar irradiance on the Earth's Top of Atmosphere (ToA) being around 340.2 W/m^2. Of this incoming shortwave radiation, approximately 70% is absorbed by the Earth's surface, oceans or atmosphere, while the remaining 30% is reflected back to the surrounding space in all directions. As the Earth system is almost at radiative equilibrium, the absorbed energy is thermally emitted as outgoing longwave radiation back by the entire globe in the long-term. Hence, changes in ERB, determined by the balance between the received solar irradiance and the Earth's total outgoing radiation, i.e., emitted and reflected radiation have both direct and long-lasting consequences in the state of the planet's climate.

For space geodesy, an accurate model of the Earth's outgoing radiation can also improve the precision of satellite orbit estimation process. As the non-conservative forces acting on satellites, such as perturbations due to the Earth's non-spherical shape, tidal forces, or gravitational attraction of other solar system bodies are well known, a key factor for increasing the precision of orbit determination is accurate modeling of non-conservative forces. For satellites on Medium Earth Orbits (MEOs, 2,000 km to 35,786 km), these are mainly due to effects of radiation pressure, such as solar radiation pressure, Earth radiation pressure, thermal re-radiation of energy absorbed by satellite surfaces, or satellite antenna

thrust (Table 1.1). Besides solar radiation pressure, most orbit determination software rely on empirical force models to account for other non-conservative forces. However, these empirical models are dependent on continuous observations of the satellite positions giving them little or no predictive power over the satellite's future state.

Table 1.1 Major perturbations acting on GNSS satellites [5, 6]

Perturbation	Magnitude (m/s^2)
Earth's two-body attraction	0.59
Earth's oblateness	5×10^{-5}
Lunar gravitational attraction,	5×10^{-6}
Solar gravitational attraction,	2×10^{-6}
Other Earth's gravity field terms ,	3×10^{-7}
Solar radiation pressure,	5×10^{-7}
Earth radiation pressure,	2×10^{-9}

Fig. 1.2 EPIC view of the Earth at 2019-02-25 11:06:10 UTC. Credit: NASA

While using space geodesy for determining the Bond albedo as an inverse problem is a previously unstudied approach, the modeling of the radiation pressure forces has been

extensively studied previously [7–12]. As for utilizing imaging data to estimate the ERB, this approach has been used almost since the first satellites were launched. In addition to CERES, other means of estimating the outgoing radiation currently include using Geostationary satellites for radiometric measurements [13], and the Earthshine, the illumination of the dark part of the Moon due to Earth's reflected radiation [14]. Also worth mentioning is the National Institute of Standards and Technology Advanced Radiometer (NISTAR) aboard the DSCOVR spacecraft along with EPIC which is taking radiometric measurements of the global reflected radiation from the Earth's entire sunlit surface [15]. For future missions, there has already been interest for global multiangle measurements of the global outgoing radiation in utilizing CubeSats [16].

Although the discussion here has been about the Earth's Bond albedo, the methods developed as part of this research will naturally be applicable to studying any other planetary object with a well-known shape. In fact, the BepiColombo mission, arriving on its desired orbit around Mercury on December 5th, 2025, will provide an ideal case study for this method. The mission contains two probes: the Mercury Magnetospheric Orbiter (Mio) and the Mercury Planetary Orbiter (Bepi). The science payload on Bepi includes a laser altimeter for determining the surface shape, a high accuracy accelerometer designed for measuring the radiative forces due to radiation incident from the Sun and scattered by the planet, and imaging instruments observing the outgoing radiation over a wide spectral range. Despite the periherm altitude of the Bepi being only 590 km, the problems that were described with the CERES measurements do not apply for Bepi, since Mercury has no atmosphere and is tidally locked in 3/2 spin-orbit resonance with the Sun, that is, the changes in the measurement conditions are not nearly drastic in the case of the Earth.

To keep the scope of this study within reasonable limits, the following simplifications are made. Only the scattering, absorption and emission of electromagnetic radiation are considered. The reflected and emitted ToA radiation is considered be diffuse and isotropic, and the outgoing ToA radiation is considered to consist of two parts: shortwave reflected radiation and longwave emitted radiation. A static model of the ToA is used, meaning that it does not change over time when numerically simulating the inverse problem. The satellites are assumed to be spherical, and to emit all absorbed radiation immediately back as thermal re-radiation. When modeling the gravitational forces affecting the satellite motion, gravitational attraction of other solar system objects is not considered. Also, the model used for the Earth's gravitational potential considers only two terms that account for the variations in the Earth's density.

The remainder of this thesis is organized as follows. Chapter 2 presents the theory necessary for understanding the global inverse problem, while Chapter 3 introduces the

measurements used as data for the inverse problem. Chapter 4 covers the numerical methods and algorithms used in the software that solves the inverse problem. The results produced by the software are presented and discussed in Chapter 5, and the conclusions of this study are drawn in Chapter 6.

Chapter 2

Theory

2.1 Geometries and reference frames

A key element in solving the global inverse problem is determining the amount of electro-magnetic radiation received by a satellite orbiting the Earth from any point on the ToA. To fully describe the physics of the problem, a variety of geometries and frames of reference need to be defined. The rationale and order in which these frames are introduced in this section are as follows. First, to describe the position of a satellite with respect to the Earth and the Sun, a geocentric inertial frame is discussed. Second, to describe the ToA scenery as viewed by the satellite with global spherical coordinates (ϑ, φ), a geocentric frame that rotates along with the Earth's motion is used. Third, to describe the geometry of, e.g., a small planar segment of the ToA surface, local spherical coordinates (θ, ϕ) are used. The fourth and final frame is a satellite local orbital frame, which is used to describe orientation of a satellite with respect to the Earth.

2.1.1 Geocentric inertial frame

A geocentric inertial frame of reference is used to describe the motion of satellites orbiting the Earth. It is defined such that the origin lies at the Earth's center of mass, the z-axis points towards the celestial North Pole, while the xy-plane lies on the Earth's equatorial plane with the x-axis directed towards the vernal equinox, a reference point where the equatorial and ecliptic planes intersect (Fig. 2.1). More specifically, this thesis uses the J2000 frame, where the x-axis points towards the vernal equinox of the Julian date of January 1st, 2000 at 12:00 Terrestrial time.
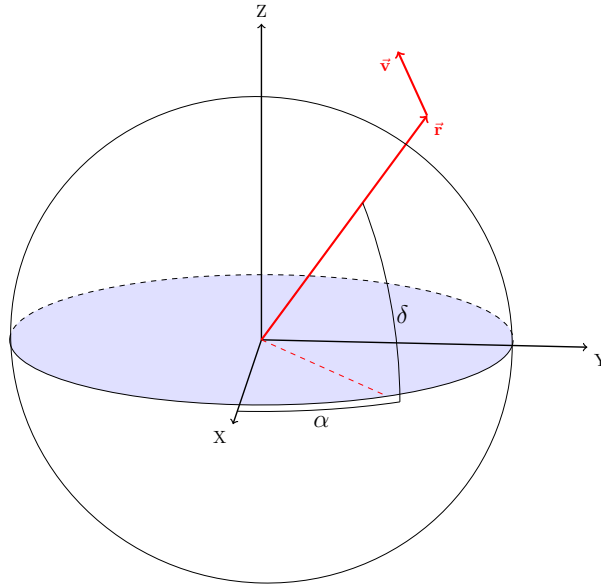
Fig. 2.1 Geocentric celestial inertial frame where the blue plane marks the Earth's equatorial plane, *x*-axis points towards the vernal equinox, *z*-axis is directed towards the celestial north pole. A satellite's position vector $\vec{\mathbf{r}}$ is given by right ascension α and declination δ.

2.1.2   Top of the Atmosphere geometry

The Top of the Atmosphere is a theoretical interface between the space and the Earth's atmosphere from which radiation enters and exits the Earth system. To describe the coordinates of a ToA scene seen by a satellite orbiting the Earth at a given moment, it is useful to define a coordinate system that rotates along with the Earth's surface. This Earth-fixed frame is centred on the Earth center of mass with the *z*-axis directed towards the celestial north Pole, and the *xy*-plane coinciding with the Earth's equatorial plane such that the *x*-axis points toward the Prime Meridian. Any point on the ToA can then be described using global spherical coordinates

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} r \sin \vartheta \cos \varphi \\ r \sin \vartheta \sin \varphi \\ r \cos \vartheta \end{pmatrix},\tag{2.1}$$

where r is the ToA radius, $\vartheta \in [0, \pi]$ is the polar angle, and $\varphi \in [0, 2\pi]$ is the azimuth angle. This study sets the ToA to be at 20 km height from the Earth's surface, as it has been estimated to be the optimal reference level for satellite-based measurements of the ToA [17]. As a consequence, an observer might see a different scenery when looking at the same point of the ToA, depending on their position, as illustrated in Fig. 2.2. However, the satellites that

are used for this thesis orbit the Earth at heights much greater than that of the ToA, around 2.0×10^4 km GPS satellites and 1.6×10^6 km for the DSCOVR satellite. At this scale the effect is considered to be fairly negligible insofar as the presented thesis is concerned.
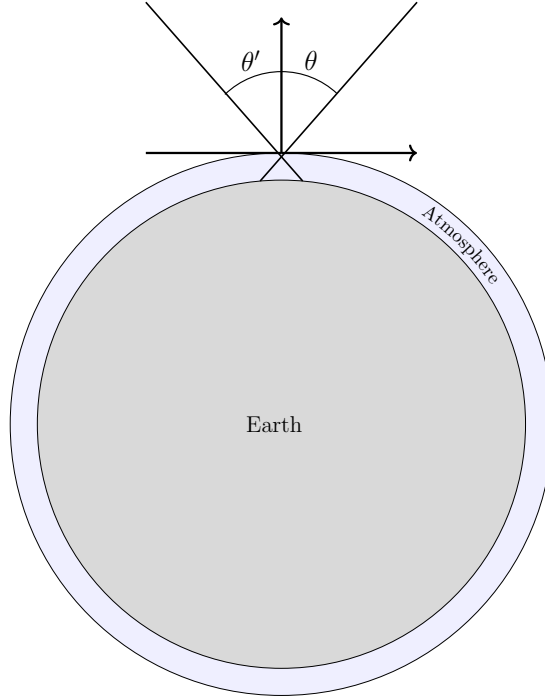
Fig. 2.2 Top of the Atmosphere Geometry. The height of the atmosphere is not to scale.

2.1.3   Surface frame

To describe the directions of the incoming radiance and the satellite detecting outgoing radiance with respect to a ToA surface element, local spherical coordinates (θ, ϕ) of the surface are used. In this frame the surface normal, the unit vector pointing towards the sun, and the unit vector pointing towards the satellite are denoted by $\hat{\mathbf{n}}$, $\hat{\mathbf{n}}_{\odot}$, and $\hat{\mathbf{n}}_{\text{sat}}$. Then the cosines for the angles of incidence and emergence θ_0 and θ are given by

$$\mu_0 = \cos \theta_0 = \hat{\mathbf{n}} \cdot \hat{\mathbf{n}}_{\odot} \tag{2.2}$$

$$\mu = \cos \theta = \hat{\mathbf{n}} \cdot \hat{\mathbf{n}}_{\text{sat}}, \tag{2.3}$$

and the phase angle α is given by

$$\cos \alpha = \hat{\mathbf{n}}_{\odot} \cdot \hat{\mathbf{n}}_{\text{sat}}. \tag{2.4}$$

2.1.4 Local orbital frame

The local orbital frame of a satellite provides the orientation of the satellite with respect to the Earth. The frame is defined by radial, normal, and transverse directions, such that given the position and velocity vectors $\vec{\mathbf{r}}$ and $\vec{\mathbf{v}}$, the unit vectors corresponding to these directions are given by [18]

$$
\begin{aligned}
\hat{\mathbf{e}}_R &= \frac{\vec{\mathbf{r}}}{|\vec{\mathbf{r}}|} \\
\hat{\mathbf{e}}_N &= \frac{\vec{\mathbf{r}} \times \vec{\mathbf{v}}}{|\vec{\mathbf{r}} \times \vec{\mathbf{v}}|} \\
\hat{\mathbf{e}}_T &= \hat{\mathbf{e}}_N \times \hat{\mathbf{e}}_R.
\end{aligned}
\tag{2.5}
$$

2.2 Radiation model for Top of the atmosphere

The key radiometric quantity for studying the flux of energy through the Earth's ToA is the spectral *radiance L_λ*. It is defined as the differential amount of spectral *radiant flux* $d\Phi_\lambda$, or radiant power, passing through an infinitesimal surface element dA into or from an infinitesimal solid angle $d\Omega$

$$
L_\lambda(\theta,\phi) = \frac{d\Phi_\lambda}{dA \cos\theta \, d\Omega},
\tag{2.6}
$$

where $dA \cos\theta$ is the projected area of the surface element as seen from the direction (θ,ϕ) of the solid angle element. This thesis assumes that the outgoing radiance from any point in the ToA originates from a locally plane-parallel surface. The radiance over all wavelengths can be acquired simply by integrating over the whole spectrum

$$
L(\theta,\phi) = \int_0^\infty L_\lambda(\theta,\phi) \, d\lambda
\tag{2.7}
$$

Similarly, the *irradiance E*, or the flux density received by the surface element is given by

$$
E = \int_0^\infty E_\lambda \, d\lambda = \int_0^\infty \frac{d\Phi_\lambda}{dA} \, d\lambda.
\tag{2.8}
$$

where E_λ is the spectral irradiance. The disk-integrated brightness can be defined as the flux density from a surface per unit solid angle

$$
I = \int_0^\infty I_\lambda \, d\lambda = \int_0^\infty \frac{d\Phi_\lambda}{d\Omega} \, d\lambda.
\tag{2.9}
$$

If the outgoing radiance as a function over the entire ToA surface is known, the total outgoing radiant flux is then given by

$$\Phi = \int_A J(\vartheta, \varphi) \, dA, \tag{2.10}$$

where J is called *radiosity*, defined as a sum of the outgoing reflected, emitted, and transmitted flux densities as a function of the global spherical coordinates (ϑ, φ), the ToA surface.

As the largest components of the Earth's outgoing radiation are the reflected shortwave ($0.1 \, \mu$m to $2.5 \, \mu$m) radiation and the emitted longwave ($2.5 \, \mu$m to $100 \, \mu$m) radiation (Fig. 2.3), the transmitted fluxes are not considered in this study. The following subsections present the theoretical foundations needed for describing the angular characteristics of the reflected and emitted radiation from the ToA. Since the following analysis applies for both spectral and integral radiometric quantities, the word "spectral" and the subscript λ are omitted in the following unless distinguishing between the two is necessary.
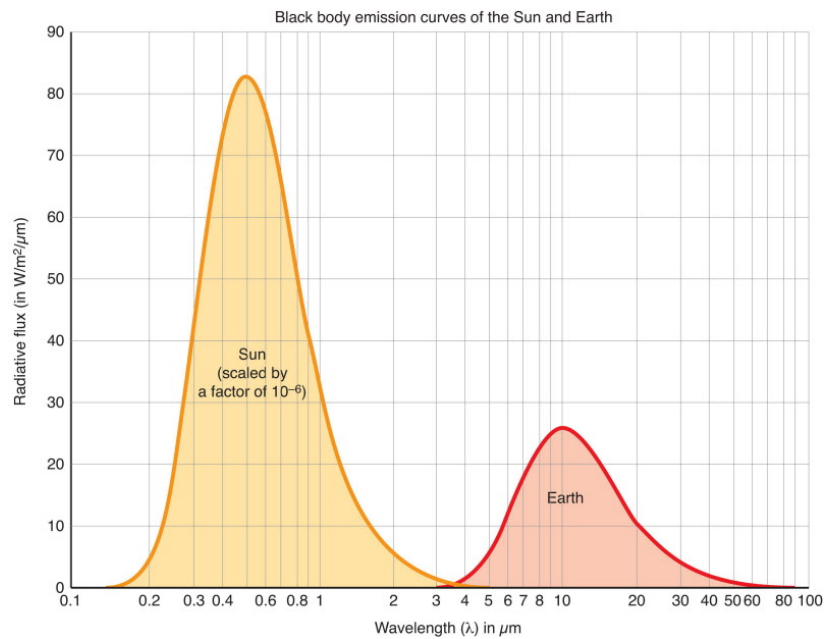
Fig. 2.3 Illustration of the blackbody emission spectra of the Sun's and the Earth's emitted radiation[1].

---

[1]Image source: `http://math.ucr.edu/home/baez/ecological/earth_and_sun_emission.jpg`
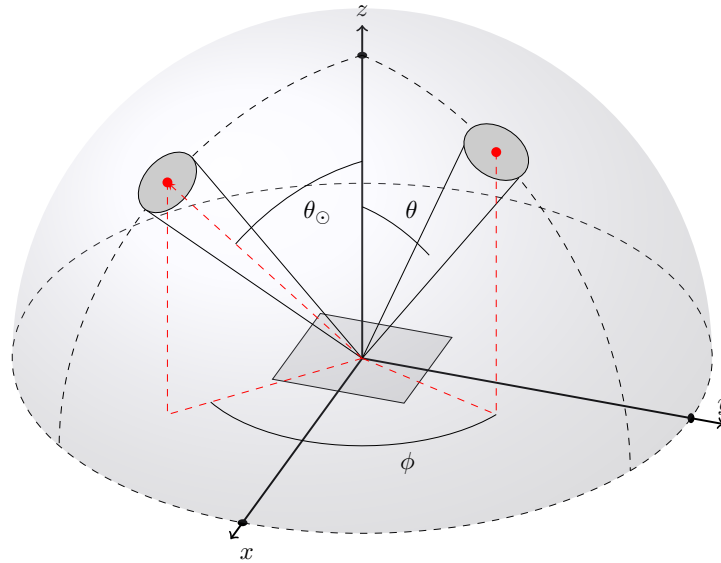
Fig. 2.4 The Geometry describing the BRDF.

2.2.1 Scattered radiation

Assuming that the incident solar radiation can be approximated as a collimated beam, the spectral irradiance is given by

$$E(\mu_0) = E_0\,\mu_0, \tag{2.11}$$

where E_0 is the solar irradiance, also called the solar "constant", and μ_0 is the cosine of the solar zenith angle θ_0. The radiance scattered diffusely from the surface element can be related to the irradiance via the *bidirectional reflectance distribution function* (BRDF) R as

$$L(\mu_0, \mu, \phi_0, \phi) = R(\mu_0, \mu, \phi_0, \phi)\,E. \tag{2.12}$$

In essence, the BRDF describes the angular anisotropy of the scattered radiance. Often, it is safe to assume that the BRDF of a surface has azimuthal symmetry, meaning that only the difference $\Delta\phi = \phi - \phi_0$ in the incident and emergent azimuth angles needs to be accounted for. By setting $\phi_0 = 0$, Eq. (2.12) can thus be re-written as

$$L_r(\mu_0, \mu, \phi) = R(\mu_0, \mu, \phi)\,E. \tag{2.13}$$

Within this framework, the phase angle α is given by

$$\alpha = \arccos(\cos\theta_0 \cos\theta + \sin\theta_0 \sin\theta \cos\phi). \qquad (2.14)$$

An example of a surface where the azimuthal symmetry assumption is invalid could be a crop field where the plants grow in rows. While such a surface can be relevant when modeling the different land types of the Earth, the resolution of the ToA used for this study is low enough to obscure such effects. Thus, this thesis considers all surfaces to follow Eq. (2.13).

A highly reflective surface can be approximated by a Lambertian surface, which reflects the incident radiation isotropically, meaning that the BRDF is a constant

$$R_L = \frac{p_L}{\pi} \qquad (2.15)$$

where p_L is a coefficient that describes the percentage of the reflected light. An ideal Lambertian surface would have $p_L = 1.0$, that is, all incident radiation would be reflected by it.

The Lommel-Seeliger scattering law can be used to describe the scattering behaviour of darker surfaces. As shown in [19], the Lommel-Seeliger scattering law assumes that incident radiation traversing in a semi-infinite layer of scattering media only scatters once and is attenuated exponentially. The resulting BRDF for this type of scattering is

$$R_{LS} = \frac{p_{LS}}{4\pi} \frac{1}{\mu + \mu_0} P(\alpha), \qquad (2.16)$$

where the reflection coefficient p_{LS} is called the single-scattering albedo, α is the phase angle, and $P(\alpha)$ is a unitless function normalized over 4π called the single-scattering phase function. For isotropic scattering, the Lommel-Seeliger BRDF simplifies to

$$R_{LS} = \frac{p_{LS}}{4\pi} \frac{1}{\mu + \mu_0}. \qquad (2.17)$$

In addition to the Lambert and Lommel-Seeliger BRDFs, there are of course plenty of more realistic scattering laws, such as the Torrance-Sparrow BRDF [20]. The Lambert and Lommel-Seeliger scattering laws do however offer a reasonable first approximation for bright and dark surfaces, making the more complicated BRDFs outside the scope of this thesis.

For the inverse problem, discussed in (see Subsection 2.4), the unknown parameters to be solved are the coefficients in the BRDFs, such as p_L, or p_{LS} in equations (2.15) and (2.17). But these parameters also include the coefficients of, e.g., the single-scattering phase function

in Eq. (2.16). However, in the following, all the coefficients are commonly referred to as *reflection coefficients*, and are denoted by p.

Having defined the reflected radiance in Eq. (2.13), the irradiance reflected by the surface can now be written as

$$E(\mu_0) = \int_{\phi=0}^{2\pi} \int_{\mu=0}^{1} L(\mu_0, \mu, \phi) \, \mu \, d\mu \, d\phi \qquad (2.18)$$

The ratio of reflected irradiance to solar irradiance is called the *plane albedo A_p* of the surface element, and with the use of (2.13) can be written as

$$A_p = \int_{\phi=0}^{2\pi} \int_{\mu=0}^{1} R(\mu_0, \mu, \phi) \, \mu \, d\mu \, d\phi. \qquad (2.19)$$

The geometric albedo can be defined as the ratio of the disk-integrated brightness of surface at phase angle $\alpha = 0°$ divided by the disk-integrated brightness of a purely Lambertian disk at the same phase angle and with the same cross-sectional area

$$A_g = \frac{I_\lambda(0°)}{I_L(0°)}. \qquad (2.20)$$

2.2.2 Emitted radiation

The spectral radiance emitted by a blackbody surface element at temperature T is described by Planck's law

$$B_\lambda(T) = \frac{2hc^2}{\lambda^5} \frac{1}{e^{\frac{hc}{\lambda k_B T}} - 1}, \qquad (2.21)$$

where h is the Planck constant, c is the speed of light, λ is the wavelength of the radiation, and k_B is the Boltzmann constant. A straightforward integration over the whole spectrum yields the blackbody radiance as

$$B(T) = \frac{2\pi^4 k_B^4}{15 \, c^2 \, h^3} T^4. \qquad (2.22)$$

The *radiant exitance*, or the flux density emitted by a blackbody is then simply

$$M_B(T) = \sigma T^4, \qquad (2.23)$$

where σ is the Stefan-Boltzmann constant.

A blackbody, however, is an idealisation of an object that absorbs all incident radiation and later emits it isotropically. This idealisation can be remedied by introducing a quantity called *directional emissivity* ε, which can be used to relate radiance emitted by a non-ideal surface $L_T(\mu, \phi)$ to the radiance emitted by a blackbody at the same temperature. Thus, the thermal radiation from a non-ideal surface can be written as

$$L(\mu, \phi) = \varepsilon(\mu, \phi) B(T).\tag{2.24}$$

In the case of a Lambertian surface, the thermal radiance is isotropic meaning that the emissivity is given by

$$\varepsilon_L = \frac{L_{T,L}}{B(T)} = \frac{q_L}{\pi},\tag{2.25}$$

where q_L is a coefficient for Lambertian emission. For this thesis, all surfaces modeled in the numerical applications are considered to emit radiation isotropically. Similar to the reflection coefficients p, for directional emissivity the unknown parameters are referred to as *emission coefficients*, and are denoted by q.

Using (2.24), the radiant exitance can now be defined as

$$M = \int_{\phi=0}^{2\pi} \int_{\mu=0}^{1} \varepsilon(\mu, \phi) B(T) \mu \, d\mu \, d\phi.\tag{2.26}$$

To describe how efficiently a surface element emits energy into all directions above the surface element, *hemispherical emissivity* ε_p is defined as the ratio of the radiant exitance of a non-ideal surface to the radiant exitance of a blackbody at the same temperature

$$\varepsilon_p = \int_{\phi=0}^{2\pi} \int_{\mu=0}^{1} \varepsilon(\mu, \phi) \mu \, d\mu \, d\phi.\tag{2.27}$$

2.2.3 Total outgoing radiation

Where the previous subsections (2.2.1) and (2.2.2) discussed the radiative properties of a single, locally plane-parallel ToA surface element, this section extends the analysis to the entire ToA surface. Essentially, this means that the BRDF and the directional emissivity vary as functions of the global spherical coordinates (ϑ, φ). As such, the total outgoing ToA

radiation described in Eq. (2.10) can now be expressed as

$$\Phi = \int_A \left[A_p(\vartheta, \varphi) E_0(\mu_0) + \varepsilon_p(\vartheta, \varphi) M_B(T, \vartheta, \varphi) \right] dA$$
$$= \Phi_r + \Phi_T, \tag{2.28}$$

where Φ_r is the total reflected radiant flux and Φ_T is the total emitted radiant flux. By denoting the incident solar radiant flux with Φ_0 and the emitted blackbody radiant flux by the entire globe with Φ_B at the same temperature as Φ_T, the *Bond albedo A_S* and *global emissivity ε_S* can then be defined as the fractions

$$A_S = \frac{\Phi_r}{\Phi_0}, \tag{2.29}$$

$$\varepsilon_S = \frac{\Phi_T}{\Phi_B}. \tag{2.30}$$

These fractions determine the Earth's radiation budget. Since this thesis assumes that all surfaces are Lambertian, the Bond albedo and the global emissivity are determined by how the reflection and emissivity coefficients p_L and q_L in equations (2.15) and (2.25) vary over the ToA. This thesis considers two approaches for this matter: one where the ToA surface is categorised into a set of classes by surface type and cloud cover, and a more analytical one, where the reflection and emission coefficients are given as square-integrable functions on the unit sphere.

ToA classification

As the BRDF and directional emissivity can vary depending on the surface type such as grasslands, deserts, or clouds, a natural approach for describing the ToA surface is to classify it according to each surface type. In this approach, each class is assigned with a corresponding BRDF and directional emissivity model. In practical applications, radiometric remote sensing measurements are used to classify the surface types and to interpolate the models describing the angular characteristics of the reflected and emitted radiation.

In more detail, the ToA surface is classified by dividing the surface to N_c subsets, or classes, such that the surface element belonging to the jth class have the same scattering and emission models and a set of coefficients. Then, the total number of parameters determining the ToA radiation field is given by

$$N_P = \sum_{j=1}^{N_c} (N_{p_j} + N_{q_j}), \tag{2.31}$$

where N_p, and N_q denote the number of reflection coefficients and emission coefficients, respectively. As an example, a surface element assigned with a combined Lambert and isotropic Lommel-Seeliger BRDF and a isotropic directional emission function would have $N_p = 2$ and $N_q = 1$.

There exist several land cover classification schemes with varying degrees of detail. However, for the purposes of this study, a simple classification algorithm was developed for detecting clouds, oceans, vegetated surfaces, and deserts from EPIC imaging data, which is presented in Chapter 4.

Spherical harmonics model

The following assumes that the BRDF for each ToA surface element is simple, in the sense that in can be written as a product between some reflection coefficient and a function describing the angular dependency of the reflected light. Examples of such BRDFs are the Lambertian BRDF and the Lommel-Seeliger BRDF for isotropic scattering, where the reflection coefficients are $p = p_L$ and $p = p_{LS}$, respectively. The same assumption applies to the directional emissivity of the surface elements. Then, assume that the BRDF and directional emissivity are the same over the entire ToA, but such that their associated reflection and emission coefficients p and q vary over the globe. If these coefficients are square-integrable functions on the unit sphere, they can be represented by the following series expansions

$$p(\vartheta, \varphi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} p_{lm} Y_{lm}(\vartheta, \varphi), \tag{2.32}$$

$$q(\vartheta, \varphi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} q_{lm} Y_{lm}(\vartheta, \varphi), \tag{2.33}$$

where the coefficients p_{lm} and q_{lm} are constant and $Y_{lm}(\vartheta, \varphi)$ are orthonormal basis functions on the unit sphere called *spherical harmonics*

$$Y_{lm}(\vartheta, \varphi) = \sqrt{\frac{2l+1}{4\pi} \frac{(l-m)!}{(l+m)!}} P_{lm}(\cos \vartheta) e^{im\varphi}. \tag{2.34}$$

Here the functions $P_{lm}(\cos \vartheta)$ are orthogonal functions called the associated Legendre functions.

The simplest model for the ToA is the *uniform Lambertian sphere*, that is, a model where only the $l = 0$ term is non-zero and all outgoing radiation is isotropic

$$p_L(\vartheta, \varphi) = p_{00} Y_{00}(\vartheta, \varphi), \tag{2.35}$$

$$q_L(\vartheta, \varphi) = q_{00} Y_{00}(\vartheta, \varphi). \tag{2.36}$$

Here $Y_{00} = \frac{1}{\sqrt{4\pi}}$, meaning that p and q are independent of the coordinates (ϑ, φ). the Bond albedo and the spherical emissivity are then simply $A_S = p_L$ and $\varepsilon_S = q_L$. In case the entire surface obeys the Lommel-Seeliger scattering law, the spherical albedo is given by $A_S = \frac{3}{2}(1 - \ln 2)p_{LS}$.

A semi-realistic model of the ToA can be achieved by assuming that the reflectivity and emissivity coefficients vary as functions of latitude, that is, only the $m = 0$ terms are non-zero.

$$p(\vartheta) = \sum_{l=0}^{\infty} p_l P_l(\cos \vartheta), \tag{2.37}$$

$$q(\vartheta) = \sum_{l=0}^{\infty} q_l P_l(\cos \vartheta), \tag{2.38}$$

where the subscript denoting $m = 0$ has been omitted. A version of this model has been presented in [9], dubbed here as the *polar cap model*, where the last non-zero term is $l = 2$

$$p(\vartheta) = p_0 + p_1 P_1(\cos \vartheta) + p_2 P_2(\cos \vartheta), \tag{2.39}$$

$$q(\vartheta) = q_0 + q_1 P_1(\cos \vartheta) + q_2 P_2(\cos \vartheta). \tag{2.40}$$

This model accounts for seasonal variability in the ToA reflectivity and emissivity by defining the first degree terms as

$$p_1 = c_0 + c_1 \cos(\omega(JD - t_0)) + c_2, \tag{2.41}$$

$$q_1 = k_0 + k_1 \cos(\omega(JD - t_0)) + k_2, \tag{2.42}$$

where the c_0, c_1, c_2, k_0, k_1, and k_2 are constants, $\omega = \frac{2\pi}{365.25}$, JD is the Julian date, and t_0 is the epoch of the periodic terms.

2.3   Satellite dynamics

Unperturbed orbit

Where previously the outgoing radiation from the planet Earth was discussed, this section describes how that radiation affects satellite orbits. But first, consider a satellite orbiting a perfectly spherically symmetric body. Assuming the satellite's size and mass are significantly smaller than that of the spherical body, it can be thought of as a point mass, and the gravitational attraction is described by Newton's law of gravitation

$$\ddot{\vec{\mathbf{r}}} = -\frac{GM}{r^3}\vec{\mathbf{r}}, \tag{2.43}$$

where G is the gravitational constant and M is the mass of the body, which for this thesis is the Earth's mass. The solution to Eq. (2.43) describes the motion of a satellite on an unperturbed orbit, or Keplerian orbit, which can be characterised in terms of the satellite's position $\vec{\mathbf{r}}(t)$ and velocity $\dot{\vec{\mathbf{r}}}(t)$ at an arbitrary epoch t.

Often, it is more convenient to describe the orbit using six parameters called the *Keplerian elements* $(a, e, i, \Omega, \omega, f(t))$, illustrated in Fig. 2.5. The first two of the Keplerian elements describe the shape of the orbit, with a being the *semi-major axis* of the orbit and e is its *eccentricity*, i.e. how much the orbit deviates from being perfectly circular.

The next three of the Keplerian elements describe various angles relating the orbit to a given frame of reference, which for this thesis is the J2000 frame, as discussed in section 2.1.1. The *inclination i* gives the angle between the satellite's orbital plane and the Earth's equatorial plane, marked by two points of reference called ascending nodes. The *right ascension of the ascending node* Ω is the angle between some point of reference and the orbital node which the satellite passes northwards. The vernal equinox is used as the point of reference in the GCRF frame. The *argument of perigeum* ω is the angle between the ascending node and the perigee, the point at which the satellite is closest to the Earth.

Finally, the *true anomaly* $f = f(t)$ is used to describe the position of the satellite at time t as the angle between the satellite and the perigee. To describe the geometry between the satellite, the Earth, and the Sun, it is useful to define the satellite's *argument of latitude* $u = \omega + f$, which gives the angle between the ascending node and the satellite. Then, the Sun's argument of latitude, which is determined from its apparent motion around the Earth, is projected on the satellite's orbital plane, and denoted by u_s. The relative position between the satellite and the Sun is then given by the difference $\Delta u = u - u_s$, and the Sun's elevation angle to the satellite's orbital plane β_s, as shown in Fig. 2.6. These angles are especially useful when examining the effects of radiation pressure on a satellite's trajectory.
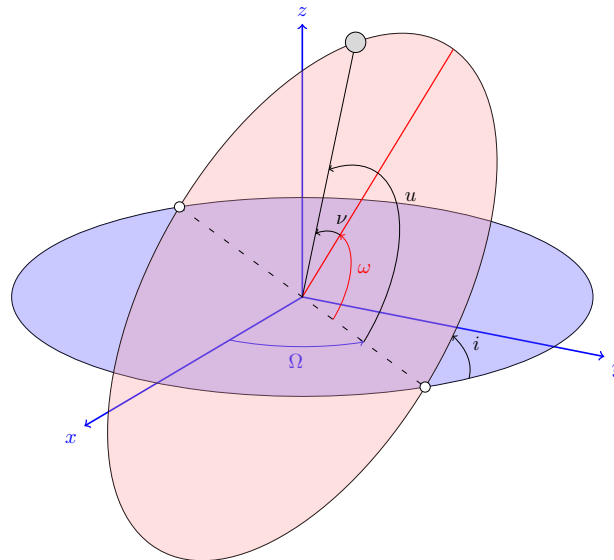
Fig. 2.5 The orbital elements of a satellite orbiting the Earth.
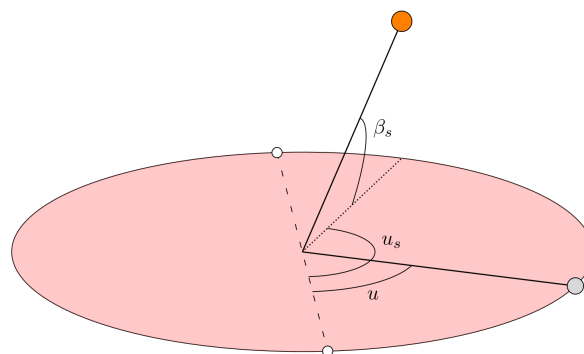

Fig. 2.6 Geometry between the Sun, the Earth and the satellite.

Osculating orbit

In reality, the satellite's equations of motion are far more complicated than what is described in Eq. (2.43) as there are various other forces affecting the satellite's motion. These forces can be thought of as small perturbation terms $\vec{\mathbf{a}}$ in addition to the dominating two-body attraction term

$$\ddot{\vec{\mathbf{r}}} = -\frac{GM}{r^3}\vec{\mathbf{r}} + \vec{\mathbf{a}}(t, \vec{\mathbf{r}}, \dot{\vec{\mathbf{r}}}, c_0, \dots, c_n), \tag{2.44}$$

where $c_0, \dots, c_n$ are parameters that characterize the perturbations [5]. The perturbations mean that the orbit cannot be determined by the same orbital elements at any given epoch. Instead, osculating orbital elements are used, which can be defined as follows. Given a solution to Eq. (2.44), obtained via numerical integration over some time period, a set of osculating orbital elements $\{a(t), e(t), i(t), \Omega(t), \omega(t), f(t)\}$ can be used to determine the satellite's trajectory at each epoch t within the integration time interval.

In general, the perturbations can be split into two categories: conservative forces and non-conservative forces. The first category accounts for all the gravitational effects acting on the satellite. Of this category, only the perturbations due to the Earth's non-spherical shape are considered for the scope of this thesis. As for the non-conservative forces, this thesis considers the perturbations due to radiation pressure from various sources. Unlike with gravitational attractions, the shape of the satellite becomes relevant, when calculating the effects of radiation pressure on the satellite's trajectory because how each surface of the satellite absorbs, scatters, and emits radiation can significantly affect the magnitude of the perturbation. This thesis limits its scope to using a so-called cannonball satellite design, which assumes that the satellite is shaped as a sphere.

2.3.1 Complex gravitational field

Since the Earth obviously is not perfectly spherically symmetric, but rather resembles an oblate spheroid (Fig. 2.7), a more realistic model than the one assumed in the two-body problem is needed. The gravitational potential of an arbitrarily shaped object, provided a continuous density distribution $\rho(\vec{\mathbf{r}'})$ at some point $\vec{\mathbf{r}'}$ within the object, is given by

$$U(\vec{\mathbf{r}}) = -G \int \frac{\rho(\vec{\mathbf{r}'})}{|\vec{\mathbf{r}} - \vec{\mathbf{r}'}|} d^3\vec{\mathbf{r}'}. \tag{2.45}$$

As shown in [21], Eq. (2.45) can be expanded using associated Legendre functions $P_{lm}(x)$
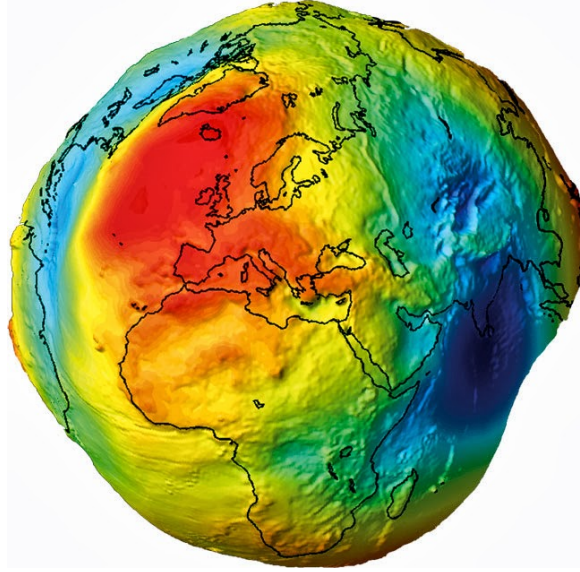
Fig. 2.7 Earth's exaggerated deviations from a spherical shape, where red areas are outside of the sphere and blue areas are inside the sphere. Credit: ESA

$$U(r, \theta, \phi) = -\frac{G}{r} \left\{ 1 - \sum_{l=1}^{\infty} \left(\frac{R}{r}\right)^l J_l P_{l0}(\cos\theta) \right.$$
$$\left. = -\sum_{l=1}^{\infty} \sum_{m=1}^{l} \left(\frac{R}{r}\right)^l P_{lm}(\cos\theta)\left[c_{lm}\cos(m\phi) + s_{lm}\sin(m\phi)\right] \right\}, \qquad (2.46)$$

where θ and ϕ are the polar angle and the azimuth angle, R is some scale factor and J_l, c_{lm}, and s_{lm} are dimensionless coefficients defined as

$$J_l = -\frac{1}{MR^l} \int P_{l0}(\cos\theta')r'^l \rho(\vec{\mathbf{r}'})\,d^3\vec{\mathbf{r}'}$$
$$c_{lm} = -2\frac{(l-m)!}{(l+m)!}\frac{1}{MR^l} \int P_{lm}(\cos\theta')r'^l \cos(m\phi')\rho(\vec{\mathbf{r}'})\,d^3\vec{\mathbf{r}'}, \qquad (2.47)$$
$$s_{lm} = -2\frac{(l-m)!}{(l+m)!}\frac{1}{MR^l} \int P_{lm}(\cos\theta')r'^l \sin(m\phi')\rho(\vec{\mathbf{r}'})\,d^3\vec{\mathbf{r}'},$$

where M is the total mass of the object. When $l = m = 1$ each coefficient can be set to zero by choosing the center-of-mass frame as the frame of reference, as these coefficients only describe the object's center of mass. The coefficients are called zonal coefficients when $m = 0$, tesseral coefficients when $m < l$, and sectorial coefficients when $m = n$. An example of each case is depicted in Fig. 2.8.

Mainly three types of indirect observations are used in combination to determine the coefficients: satellite tracking, surface gravimetry, and altimetric measurements [21]. The
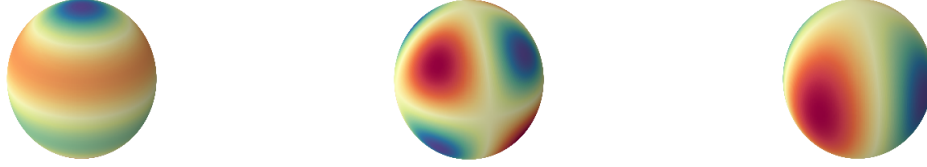
Fig. 2.8 From left to right: illustration of zonal ($l = 3$), tesseral ($l = 3$, $m = 2$), and sectorial ($l = m = 2$) coefficients.

state-of-the-art Earth gravity models such as the Earth Gravity Model 2008 use coefficients of degree 2190 and order 2159 [22].

Zonal Earth gravity model

By far the largest influence on the Earth's gravitational field is due to the J_2 term which describes the Earth's oblateness, whereas coefficients of higher order and degree are several orders of magnitude smaller [21]. This is to be expected since the difference between the Earth's nominal equatorial radius and nominal polar radius is around 21 km while other deviations from a perfectly spherical shape are much smaller. For this study, a simplified zonal Earth gravity model, including only the terms J_2 and J_3, is used, as it provides a somewhat accurate model of the Earth's gravitational field, and is easy to implement numerically. Such a gravitational potential is longitudinally symmetric, meaning that the c_{lm} and s_{lm} terms vanish, reducing Eq. (2.46) to

$$U(r, \theta) = -\frac{GM_\oplus}{r}\left[1 - \sum_{l=2}^{\infty}\left(\frac{R}{r}\right)^l J_l P_{lm}(\cos\theta)\right], \tag{2.48}$$

where R is now set to be the Earth's nominal equatorial radius, $M_\oplus$ is the Earth's total mass, and the J2000 frame is chosen as the frame of reference. The potential corresponding to the J_2 term is

$$U_{J_2}(r, \theta) = \frac{GM_\oplus}{r^3}\frac{J_2 R^2}{2}(3\cos^2\theta - 1) \tag{2.49}$$

The corresponding accelerations in the J2000 frame are

$$\ddot{r}_{J2,x} = \mu J_2 R^2 \frac{x}{r^7} \left(6z^2 - \frac{3}{2}(x^2 + y^2) \right)$$

$$\ddot{r}_{J2,y} = \mu J_2 R^2 \frac{y}{r^7} \left(6z^2 - \frac{3}{2}(x^2 + y^2) \right) \tag{2.50}$$

$$\ddot{r}_{J2,z} = \mu J_2 R^2 \frac{z}{r^7} \left(3z^2 - \frac{9}{2}(x^2 + y^2) \right)$$

The potential corresponding to the J_3 term is given by

$$U_{J_3}(r, \theta) = \frac{GM_\oplus}{r^4} \frac{J_3 R^3}{2} \cos\theta \left(5\cos^2\theta - 3 \right), \tag{2.51}$$

and the corresponding accelerations are

$$\ddot{r}_{J3,x} = \mu J_3 R^3 \frac{xz}{r^9} \left(10z^2 - \frac{15}{2}(x^2 + y^2) \right)$$

$$\ddot{r}_{J3,y} = \mu J_3 R^3 \frac{yz}{r^9} \left(10z^2 - \frac{15}{2}(x^2 + y^2) \right) \tag{2.52}$$

$$\ddot{r}_{J3,z} = \mu J_3 \frac{R^3}{r^9} \left(4z^2 \left(z^2 - 3(x^2 + y^2) \right) + \frac{3}{2}(x^2 + y^2)^2 \right).$$

2.3.2 Radiation pressure

Direct source of radiation

The energy of a particle E_e is related to its mass m and momentum p via the relativistic dispersion relation as

$$E_e = \sqrt{(pc)^2 + (mc^2)^2} \tag{2.53}$$

where c is the speed of light in vacuum and subscript "e" is used here to mark the difference between the symbols for energy and irradiance. Eq. 2.53 shows that even a massless particle, such as a photon, has momentum. Thus, when interacting with massive particles, a photon must then transfer its momentum, and as a result of this transfer, the photon mediates an electromagnetic force on the particle it interacts with. On a macroscopic scale this interaction causes a radiation pressure P_{RP} on a surface, which for a cross-sectional area depends on the

irradiance as

$$P_{RP} = \frac{E}{c}.$$ (2.54)

Now, consider a satellite of mass m. Following the analysis in [23], if the surface of the satellite *absorbs* part of the radiation, scatters part of it *diffusely*, and part of it *specularly*, the corresponding radiation pressure acceleration is given as

$$\ddot{\vec{r}}_{RP} = -\frac{E}{mc} \int_A \hat{\mathbf{e}}_S \cdot \hat{\mathbf{e}}_N \left[(1-\rho)\hat{\mathbf{e}}_S + 2\left(\frac{1}{3}\delta + \hat{\mathbf{e}}_S \cdot \hat{\mathbf{e}}_N \rho \right)\hat{\mathbf{e}}_N \right] dA,$$ (2.55)

where $\hat{\mathbf{e}}_N$ is the unit normal of the surface element dA and $\hat{\mathbf{e}}_S$ is a unit vector pointing towards the source of the radiation. The parameters α, δ, and ρ are, respectively, the fractions of absorbed, diffusely scattered, and specularly scattered radiation, with

$$\alpha + \delta + \rho = 1.$$ (2.56)

When the integral in Eq. (2.55) is carried out over a spherical surface, the radiation pressure acceleration for a cannonball-shaped satellite of radius R is given by the expression

$$\ddot{\vec{r}}_{RP} = -\frac{E}{mc}\pi R^2 \left[1 + \frac{4}{9}\delta \right]\hat{\mathbf{e}}_S.$$ (2.57)

If the surface of a satellite consists of N planar surfaces, the total acceleration due to radiation pressure is given by a sum of the accelerations over each surface,

$$\ddot{\vec{r}}_{RP} = -\sum_{j=1}^{N} \frac{E_j}{mc} A_j \hat{\mathbf{e}}_{S,j} \cdot \hat{\mathbf{e}}_{N,j} \left[(1-\rho_j)\hat{\mathbf{e}}_{S,j} + 2\left(\frac{1}{3}\delta_j + \hat{\mathbf{e}}_{S,j} \cdot \hat{\mathbf{e}}_{N,j}\rho_j \right)\hat{\mathbf{e}}_{N,j} \right].$$ (2.58)

The numerical methods used to evaluate the total force due to radiation pressure on a satellite are covered in Chapter 4.

Thermal re-radiation

As part of the incoming radiation is absorbed by the satellite, this energy will later on be re-emitted to the surrounding space, which causes a recoil force on the satellite. Modeling this thermal re-emission would require detailed description of the heat transfer within the satellite, and therefore of the satellite's material properties. As this would be computationally expensive and difficult to implement, some simplifications need to be made. This thesis assumes that the satellite surfaces immediately re-radiate the absorbed energy isotropically

into all directions, which results in a thermal recoil force defined as [8]

$$\vec{\mathbf{r}}_{\text{TR}} = -\alpha \frac{P_{RP}}{m} A \frac{2}{3} (\hat{\mathbf{e}}_S \cdot \hat{\mathbf{e}}_N) \hat{\mathbf{e}}_N. \tag{2.59}$$

Antenna thrust

Another significant radiation induced effect on the satellite is caused by the antenna emission. This is simply given by

$$\vec{\mathbf{a}}_a = \frac{\Phi_a}{mc} \hat{\mathbf{e}}_N, \tag{2.60}$$

where Φ_a is the radiative power of the antenna, and $\hat{\mathbf{e}}_N$ points towards the center of the Earth. E.g. for a GPS satellite the antenna power is around 77 W [8].

2.4 The inverse problem

Bayesian inversion

The goal of an inverse problem is to find an estimate of some unknown parameters described by $\vec{\mathbf{P}}$, given some indirect observations $\vec{\mathbf{m}}$ and a forward model $\mathbf{M}$ describing the relation between the measurements and the unknowns,

$$\vec{\mathbf{m}} = \mathbf{M}(\vec{\mathbf{P}}) + \varepsilon, \tag{2.61}$$

where ε accounts for the measurement uncertainty in the model. This thesis treats the inverse problem from a Bayesian viewpoint, where the measurement, the unknown parameters, and the model uncertainty are considered to be random variables. Thus, instead of trying to find an exact values for the unknown, as done in the framework of regularization theory, the Bayesian approach is to find an *a posteriori* probability density of $p_p(\vec{\mathbf{P}})$, which is a probability of the unknown parameters given an outcome of an experiment. This *posterior* distribution is given by the Bayes' formula

$$p_p(\vec{\mathbf{P}}) = \frac{p_{pr}(\vec{\mathbf{P}}) \, \pi_\varepsilon(\Delta \mathbf{M}(\vec{\mathbf{P}}))}{\pi(\vec{\mathbf{m}})}. \tag{2.62}$$

Here, $p_{pr}(\vec{\mathbf{P}})$ and $\pi(\vec{\mathbf{m}})$ are the *prior* and the *marginal* probability densities of the unknowns and the measurement, respectively. The likelihood density $\pi_\varepsilon(\Delta \mathbf{M}(\vec{\mathbf{P}}))$ expresses the probability of the experiment outcomes given a realization of the unknown parameters, and $\Delta \mathbf{M}(\vec{\mathbf{P}})$

is the residual

$$\Delta \mathbf{M}(\vec{\mathbf{P}}) = \vec{\mathbf{m}} - \mathbf{M}(\vec{\mathbf{P}}). \tag{2.63}$$

Since the marginal distribution acts just as a normalization factor, the behaviour of the posterior distribution depends on the likelihood and the prior distribution. Indeed, it is often enough to only consider the following relation

$$p_P(\vec{\mathbf{P}}) \propto p_{pr}(\vec{\mathbf{P}})\, \pi_\varepsilon(\Delta \mathbf{M}(\vec{\mathbf{P}})). \tag{2.64}$$

In practice, the role of the prior is to indicate the *a priori* degree of belief about the unknown parameters while the likelihood quantifies the results of the experiments. The beauty of the Bayesian analysis lies in the fact that the Bayes' formula allows for the degree of belief of the unknowns to be updated with the results of the experiment and that a previous experiment can serve as the *a priori* information for a subsequent experiment.

2.4.1 Inversion of the Earth's Bond albedo

Having described all the necessary theory for posing the inverse problem, the statistical inversion of the Earth's energy budget can finally be formulated properly. Suppose that at times $\{t_j\}_{j \in S_t}$, with $S_t = \{1, 2, \ldots, N_s\}$, a set of observations of the position and velocity of N_s satellites have been made

$$S_s = \{\{[\vec{\mathbf{r}}(t_j)_1, \dot{\vec{\mathbf{r}}}(t_j)_1]^T, \ldots, [\vec{\mathbf{r}}(t_j)_{N_s}, \dot{\vec{\mathbf{r}}}(t_j)_{N_s}]^T\}\}_{j \in S_t} \tag{2.65}$$

In addition, assume that these satellites simultaneously collect a set of radiometric imaging observations of the Earth

$$S_f = \{\{[f(t_j)_{1,1}, \ldots, f(t_j)_{N_{pix},1}]^T, \ldots, [f(t_j)_{1,N_s}, \ldots, f(t_j)_{N_{pix},N_s}]^T\}\}_{j \in S_t}, \tag{2.66}$$

where $f(t_j)_{i,k}$ denotes the ith imaging data pixel from the kth satellite at time t_j. Altogether these measurements are described by the vector

$$\begin{aligned}
\vec{\mathbf{m}} = [&\vec{\mathbf{m}}(t_1)_{s,1}, \ldots, \vec{\mathbf{m}}(t_1)_{s,N_s}, \vec{\mathbf{m}}(t_1)_{f,1}, \ldots, \vec{\mathbf{m}}(t_1)_{f,N_s}, \ldots, \\
&\vec{\mathbf{m}}(t_{N_t})_{s,1}, \ldots, \vec{\mathbf{m}}(t_{N_t})_{s,N_s}, \vec{\mathbf{m}}(t_{N_t})_{f,1}, \ldots, \vec{\mathbf{m}}(t_{N_t})_{f,N_s}]^T,
\end{aligned} \tag{2.67}$$

where $\vec{\mathbf{m}}(t_j)_{s,k}$ represents the space geodetic observation of the kth satellite at time t_j, and $\vec{\mathbf{m}}(t_j)_{f,k}$ is its radiometric counterpart. The total length of the measurement vector is then $(6 + N_{pix})N_s N_t$.

Now, let the vector $\vec{\mathbf{P}}$ describe the reflection coefficients $p(\vartheta, \varphi)$ and emission coefficients $q(\vartheta, \varphi)$ of a ToA radiation model at time t_j, as defined in (2.2). The observation uncertainty ε is assumed to be a multivariate Gaussian random variable with zero mean and a covariance matrix Λ_ε, the off-diagonal elements of which contain the correlations between the observation uncertainties. Furthermore, let $\mathbf{M}(\vec{\mathbf{P}})$ represent the forward model which maps the unknowns p and q to the measurement space. On the basis of the central limit theorem, the likelihood density is often assumed to be Gaussian, and that assumption is made here as well. Thus, Eq. (2.64) becomes

$$p_p(\vec{\mathbf{P}}) \propto p_{pr}(\vec{\mathbf{P}}) \exp\left[-\frac{1}{2}\chi(\vec{\mathbf{P}})^2\right], \tag{2.68}$$

where χ^2 is the squared residual of the model and measurement scaled by the model uncertainty, defined as

$$\chi^2 = \Delta\mathbf{M}(\vec{\mathbf{P}})^T \Lambda_\varepsilon^{-1} \Delta\mathbf{M}(\vec{\mathbf{P}}). \tag{2.69}$$

To sample the posterior distribution, one can use Markov-chain Monte Carlo (MCMC) methods.

2.4.2 Markov-chain Monte Carlo methods

To study the posterior distribution given in Eq. (2.62) various numerical methods have been developed, one of which is the Markov-chain Monte Carlo (MCMC) method. The term Monte Carlo refers to the numerical methods that rely on the generation of random variables from well-known probability distributions and a Markov-chain is a sequence of random variables $\{X_j\}_{j \in \mathbb{Z}^+}$ where each state X_{j+1} depends only on the previous state X_j, i.e. the transition probability satisfies

$$\begin{aligned} &\mathbb{P}(X_{j+1} = v \,|\, X_j = u_j, X_{j-1} = u_{j-1}, \ldots, X_0 = u_0) \\ &= \mathbb{P}(X_{j+1} = v \,|\, X_j = u_j), \end{aligned} \tag{2.70}$$

where u and v are here used to denote realization of the random variables.

This study assumes that the *Markov kernel P*, which dictates the probability of transitioning from one state to another, does not depend on time:

$$\mathbb{P}(X_{j+1} = v \,|\, X_j = u) = P(u, v). \tag{2.71}$$

To sample a target distribution $p(X)$ with a MCMC algorithm, the Markov kernel has to be *invariant, irreducible*, and *aperiodic*. Invariance of the Markov kernel with respect to the target distribution $p(u)$ means that it satisfies

$$\int p(u) P(u, v) \, du = p(v). \tag{2.72}$$

Irreducibility means that the probability of reaching any state of the target distribution from some other state in the chain with an arbitrary number of state transitions is non-zero. Aperiodicity means that the Markov kernel does not generate a chain that remains in a loop forever.

Metropolis-Hastings kernel

A frequently used method of ensuring invariance is requiring that the Markov kernel satisfies so-called detailed balance

$$p(u) P(u, v) = p(v) P(v, u). \tag{2.73}$$

One of the most common MCMC method is the Metropolis-Hastings (MH) algorithm, which samples the target distribution $p(u)$ using the *Metropolis-Hastings kernel*

$$K(u, v) = a(u, v) \, q(u, v), \tag{2.74}$$

where the Markov kernel q is known as the *proposal distribution* and the transition kernel a is defined as

$$a(u, v) = \min\left(\frac{p(v)q(v, u)}{p(u)q(u, v)}, 1 \right). \tag{2.75}$$

Now the Metropolis-Hastings kernel satisfies the detailed-balance condition in Eq. (2.73). If the proposal distribution q is symmetric, $q(v, u) = q(u, v)$, then the transition kernel further simplifies to

$$a(u, v) = \min\left(\frac{p(v)}{p(u)}, 1 \right). \tag{2.76}$$

The numerical implementation based on the MH algorithm is described in Chapter 4.

Chapter 3

Measurements and observations

3.1 Earth Polychromatic Imaging Camera images

The EPIC instrument, mounted on the DSCOVR spacecraft by the National Oceanic and Atmospheric Administration (NOAA), is a 2048×2048 resolution CCD camera with 10 narrowband filters ranging from the ultraviolet spectral band (317 nm) to the near-infrared wavelengths (780 nm). Located at the Lagrange 1 point from where it can see nearly the entire sunlit surface of the Earth at relatively small phase angles ($4.5°$ to $11.5°$), the spatial resolution of a pixel at nadir is around 18 km. The EPIC instrument takes a new set of images of the Earth covering its spectral regime 22 times a day during the summer period and 13 times in the winter. These image sets, along with the corresponding geolocation data and additional metadata, are available on the Atmospheric Science Data Center (ASDC) archive.

The raw EPIC per pixel imaging data $ct_j \in \mathbb{R}^{2048 \times 2048}$ for each EPIC wavelength band is in units of counts s^{-1}, which need to be converted to reflectances for the data to be of use for this study. Here the indexing $j = 1, \ldots, 10$ corresponds to wavelength bands from the smallest (317 nm) to the highest (780 nm) wavelength. These reflectances are the geometric albedo A_g (Eq. 2.9) values for each pixel. This conversion of ct_j to reflectance ρ_j for each EPIC wavelength band $j = 1, \ldots, 10$ can be done using radiometric calibration factors K_j [24]

$$A_{g,j} = ct_j K_j. \tag{3.1}$$

The calibration factors acquired, e.g., by comparing the EPIC images to observations made by other remote sensing satellites, and they are listed in the Table A.

This EPIC reflectance data is used in the ToA radiation model in two ways. First, by assuming that each pixel in the image is a Lambertian surface, the EPIC reflectance maps

are used to produce a quasi-realistic ToA model after normalising the values to the range of $[0, 1]$. This model is ideal for generating synthetic data, as discussed in Section 3.3. Second, to produce much cruder model of the ToA, the imaging data is used for ToA scene type classification. The EPIC images as well as all other DSCOVR mission data products are available for the public at NASA Langley Atmospheric Science Data Center (LASDC). The EPIC level 1B image dataset `epic_1b_20170101112452_02` was used for this thesis.

Table 3.1 EPIC spectral bands, calibration factors, and spectral solar irradiances[1].

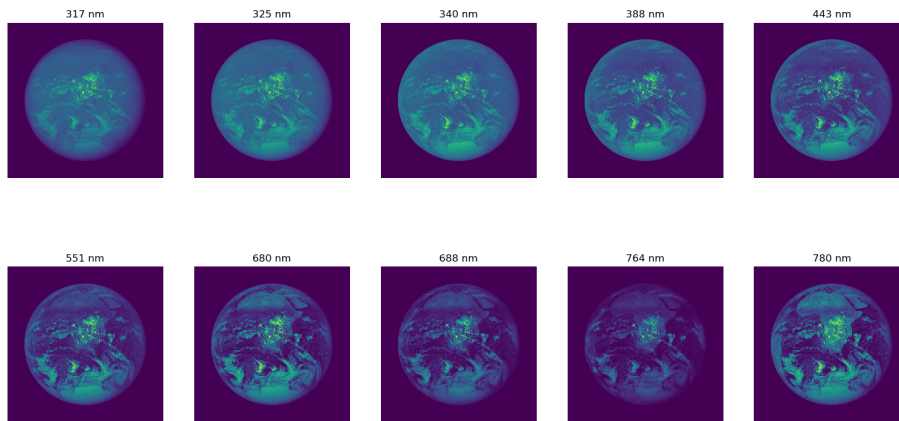| Wavelength (nm)	Calibration factor K_j	Solar irradiance (W m^{-2} nm^{-1})
318	1.216E-04	0.810
325	1.111E-04	0.651
340	1.975E-05	0.965
388	2.685E-05	0.939
443	8.340E-06	1.945
551	6.660E-06	1.865
680	9.300E-06	1.495
688	2.020E-05	1.465
764	2.360E-05	1.230
780	1.435E-05	1.190

Fig. 3.1 EPIC images from each band on 2017-01-01 11:12:52 UTC.

[1]Source: `https://eosweb.larc.nasa.gov/project/dscovr/dscovr_table`

3.2   Space geodesy data

The GPS satellite orbit data products provided by the International GNSS Service (IGS) are used to serve as the initial observations when modeling the satellite orbits. The IGS has a network of over 400 globally distributed permanent ground stations for making observations. The final data product is produced by the Bernese GNSS Software, and consists of sets of osculating orbital elements for each of the 31 active GPS satellites at the desired epochs along with their root-mean-square (RMS) deviations. Each of the GPS satellites has a unique space vehicle number (SVN), which is used for identification. The orbital elements of SVN 63 produced by the Bernese software are shown in Table 3.2.

The software accounts for complex gravitational potential due to the Earth's non-uniform density and non-spherical shape up to degree 2190 and order 2159, changes in the potential due to ocean tides, solid Earth tides along with the pole tides, and general relativistic corrections. It also includes the gravitational forces due to Mars, Venus, Jupiter, and Sun. An empirical force model is used to account for non-conservative forces such as the Earth's and the Sun's radiation pressure [5].

Table 3.2 Orbital elements and their uncertainties of SVN 63 at 2017-01-01 00:00:00 UTC.

Orbital element	Value	RMS
a (m)	26560285.85945	5×10^{-3}
e	0.0060969206	0.0
i (degree)	55.469926174	9×10^{-9}
Ω (degree)	107.195868066	13×10^{-9}
ω (degree)	31.800351190	1.558×10^{-6}
$u(t_0)$ (degree)	231.865348301	19×10^{-9}

3.3   Synthetic data

To test the performance of the inverse problem algorithm, synthetic datasets was generated, as these allow for full control over every parameter in the problem. However, care must be taken to avoid committing an *inverse crime* when solving the inverse problem from synthetic data. In simple terms, an inverse crime is committed if the same model is used to both generate the data and to solve the inverse problem, as this results in overly optimistic or even unrealistic results. To avoid such felonies, two steps were taken in this thesis. The first and the most obvious step was to use a different forward model in generating the data and

in inverting it. The second step is to use a different discretisation in the inversion and data generation.

As mentioned in Section 3.1, EPIC reflectance data is ideal for generating a forward model, as it provides a complex map of the ToA, that cannot be matched with the classification methods used in this thesis. Of the image dataset depicted in Fig. 3.1, the band 388 nm image was selected since it shows clouds quite prominently, while oceans and land surfaces are obscured due to Rayleigh scattering. The reasoning behind this choice was that even simple ToA models with few to no surface details included could perform well when solving the inverse problem from data generated using this image. For outgoing longwave radiation, the Earth was assumed to emit radiation isotropically into all directions with a radiant exitance of 240 W/m^2. Since this thesis uses a static model for the ToA (see Chapter 1), only a single EPIC image was used when modeling the ToA for. To produce a discrete model of the ToA, HEALPix scheme (Sec. 4.1.1) was used to subdivide the ToA surface into 12288 segments, while the inverse problem was solved with a ToA model consisting of 768 segments.

With this ToA model, the orbits of six non-absorbing identical cannonball-shaped satellites were simulated using the GPS orbital elements acquired from the Bernese software as initial positions, and for modeling measurement uncertainty. The synthetic data consisted of 30 sets of simulated observations of the orbital elements for each satellite taken every 24 minutes, corresponding to one revolution around the Earth for the satellites, i.e., 12 hours.

Chapter 4

Numerical methods

4.1 Forward model

To solve the inverse problem, a description of the numerical methods for implementing the forward model $\mathbf{M}(\vec{\mathbf{P}})$ in practice is needed. As part of this study, a software to do so was developed using Python. This section describes the components of the forward model in the following order. Discretising and classifying the ToA is presented first, followed by ray tracing of the radiances from surface to surface. Finally, the numerical integration to obtain the Bond albedo, and to solve the satellite's equations of motion is described.

4.1.1 Discretised Top of the Atmosphere

The discrete ToA model is achieved, using the *Hierarchical Equal Area isoLatitude Pixelization* (HEALPix) scheme. The HEALPix tessellation of the sphere partitions it into twelve diamond-shaped, equal-area base surfaces that can be recursively subdivided into four similarly shaped equal-area surfaces. These surfaces are distributed along iso-latitude rings, which allows for a fast evaluation of spherical harmonics [25, 26]. This scheme is a natural choice for the purpose of this study, as it provides fast analysis of high resolution spherically mapped data and can easily be implemented with various programming languages. An additional benefit comes from each surface having an equal area, which speeds up the computation of irradiances from the discretised ToA to satellite surfaces. The implementation of the HEALPix scheme for this study was done using `healpy` Python package.

In practical terms, a HEALPix map is an array of N_{pix} elements, or pixels, such that the spherical coordinates of the ith pixel (ϑ_i, φ_i) are at the center of a HEALPix surface element (Fig. 4.1). These surface elements are assumed to be locally plane-parallel. The numerical ToA radiation model can hence be implemented by assigning each pixel with a BRDF and
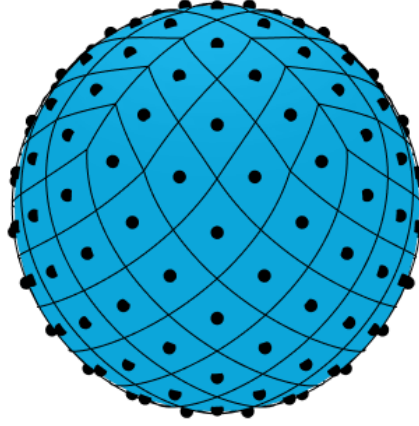
Fig. 4.1 HEALPix discretisation of a sphere into 192 pixels and the corresponding pixel centers.

a directional emissivity model with coefficients $\{p_{ij}\}$ and $\{q_{ik}\}$, where j and k index the number of scattering and emissivity coefficients N_{p_i} and N_{q_i} of the ith pixel.

4.1.2   Top of the Atmosphere classification scheme

To provide a classified model of the ToA surfaces, an unsupervised classification scheme was developed. This two-fold scheme first extracts features (clouds, water, etc.) from the images using independent component analysis (ICA) and then classifies the features using Gaussian Mixture Model (GMM) clustering. Both of these methods are included in the Python `scikit learn` library, which was used for practical implementation of the classification scheme. In the absence of proper training data, the end product of this classification scheme needs to be validated manually.

Independent component analysis

In ICA, the observations $\vec{\mathbf{x}} = [x_1, \ldots, x_D]^T$ are assumed to arise from hidden source signals $\vec{\mathbf{s}} = [s_1, \ldots, s_L]^T$, defined by the relation

$$\vec{\mathbf{x}} = \mathbf{W}\vec{\mathbf{s}} + \eta, \tag{4.1}$$

where $\mathbf{W}$ is called the mixing matrix, and η is a non-zero Gaussian random variable. A key aspect in ICA is that the source signals are assumed to have independent *non-Gaussian* probability distributions p_j

$$p(\vec{\mathbf{s}}) = \prod_{j=1}^{L} p_j s_j. \tag{4.2}$$

This assumption makes it possible for ICA to uniquely recover the source signals, which is not the case if p_j is a Gaussian distribution [27].

Gaussian mixture model

GMM can be used to classify features, such as land cover types and clouds in the case of EPIC data, from observations. The premise behind GMM clustering is to use Gaussian statistics for assigning data points $\vec{\mathbf{x}}_i$ in a given dataset $\mathbf{X}$ into K clusters. For a set of $k \in \{1, \ldots, K\}$ clusters, a Gaussian mixture model is parameterised by a set of weights, means, and covariances $\theta = \{\pi, \mu, \Sigma\}$ and is defined as

$$p(\vec{\mathbf{x}}_i \,|\, \theta) = \sum_{k=1}^{K} \pi_k \, p(\vec{\mathbf{x}}_i \,|\, \mu_k, \Sigma_k), \tag{4.3}$$

where π_k are mixing weights, $p(\vec{\mathbf{x}}_i \,|\, \mu_k, \Sigma_k)$ are Gaussian probability densities with means μ_k, and covariances Σ_k. The mixing weights satisfy $\sum_{k=1}^{K} \pi_k = 1$, and $0 \leq \pi_k \leq 1$. The GMM method uses latent variables z to determine the probability γ that a data point $\vec{\mathbf{x}}_i$ belongs to kth cluster

$$p(z_{ik} = 1 \,|\, \vec{\mathbf{x}}_i) = \gamma_k, \tag{4.4}$$

where obviously $0 \leq \gamma_k \leq 1$. The latent variable has value 1 when $\vec{\mathbf{x}}_i$ belongs to the kth cluster, and value 0 otherwise. To find the optimal parameters, the GMM method uses an iterative approach called expectation–maximisation (EM) algorithm. The end product of the GMM clustering algorithm is the probabilities at which the data points belong each cluster. Finally, the assignment of a given data point to a cluster is determined by the highest probability γ_k [27].

4.1.3   View factor ray-tracing

Consider now the spectral radiant flux from such a surface element dA_i to a surface element dA_j, which could be, e.g., a satellite surface element. The solid angle corresponding to dA_j

as viewed from the surface element dA_i is

$$d\Omega_j = \frac{dA_j\,\mu_j}{r_{ij}^2}, \tag{4.5}$$

where μ_j is the cosine of the angle between the surface normal of dA_j and the incident radiation and r_{ij} is the distance between the two surface elements (Fig. 4.2). The spectral radiant flux can thus be written as

$$d\Phi_{\lambda,ij} = L_\lambda(\mu_i,\phi_i)\mu_i\,dA_i\,d\Omega_j. \tag{4.6}$$

The spectral irradiance from the entire ToA intercepted by dA_j is then given by

$$E_{\lambda,j} = \int_{A_i} L_{\lambda,i}(\mu_i,\phi_i)\frac{\mu_i\,\mu_j}{r_{ij}^2}dA_i, \tag{4.7}$$

where A_i is the entire surface radiating towards $d\Omega_j$. One can recognize Eq. (4.7) as a *Fredholm integral equation of the first kind*, the solution of which is the bread and butter of inverse problems.
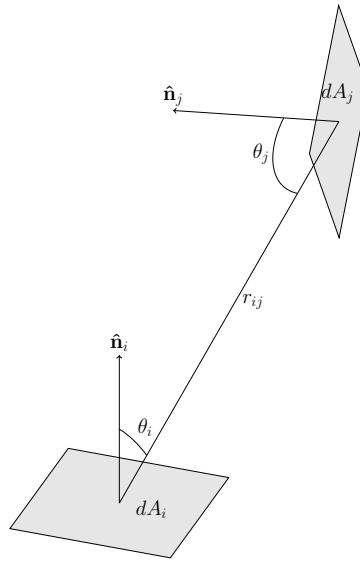

Fig. 4.2 Geometry of the view factor for surfaces i and j.

To numerically determine the incident and emergent ToA radiation, the entire surface is discretized into N_{pix} small surfaces, or pixels. Then Eq. (4.7) can be approximated as a sum

over the pixels

$$E_{\lambda,j} = \sum_{i=1}^{N_{pix}} L_{\lambda,i}(\mu_i, \phi_i) \pi F_{ji}. \tag{4.8}$$

where F_{ij} is called the view factor, defined as [28]

$$F_{ji} = \frac{\mu_j \mu_i}{\pi r_{ij}^2} dA_i. \tag{4.9}$$

ToA irradiance over the entire spectrum towards dA_j can be written as a sum of the scattered shortwave radiation and the emitted longwave radiation as

$$E_j = \sum_{i=1}^{N_{pix}} \pi F_{ji} \left[R_i(\mu_{0,i} \mu_i, \phi_i) \mu_{0,i} E_0 + \varepsilon_i(\mu, \phi_i) B(T_i) \right]. \tag{4.10}$$

One can easily recognize the expression (4.10) as a dot product between two vectors $\vec{L}$ and $\vec{F}_j$, or as a product between a matrix $\mathbf{F}$ and a vector $\vec{L}$ in the case of multiple surface elements

$$\begin{aligned} E_j &= \pi \vec{F}_j \cdot \vec{L}, \\ \vec{E} &= \pi \mathbf{F} \vec{L}. \end{aligned} \tag{4.11}$$

4.1.4 Numerical integration

Quadrature on a sphere

To evaluate the integrals over spherical surfaces numerically, such as Eq. (2.28), a spherical quadrature scheme is used [29]. The idea is to approximate the integral of a function $f : \mathbb{S}^2 \to \mathbb{R}$ over a domain Ω on the unit sphere $\mathbb{S}^2$ as a weighted sum over a finite set of nodes

$$\int_\Omega f(\vec{r}) \, d\Omega \approx \sum_{i=0}^{N-1} w_i f(r_i), \tag{4.12}$$

where N is the number of nodes r_i within the integration domain, and $i = 1, 2, \ldots, N$. One approach is to assume equal weights for all nodes, provided that the nodes are uniformly spread out over the sphere. If the quadrature is exact for f, the integral becomes

$$\int_\Omega f(\vec{r}) \, d\Omega = \frac{4\pi}{N} \sum_{i=0}^{N-1} f(r_i), \tag{4.13}$$

Since the distribution of nodes in the HEALPix scheme is seemingly uniform over a spherical surface, it is applied in this study in the numerical implementation of Eq (4.13).

Runge-Kutta 4 method

Since finding an analytical solution to a satellite's equations of motion described in Eq. (2.44) in most cases would be impossible, numerical methods have to be relied upon to solve the equation. A popular choice for numerical integration is the Runge-Kutta 4 method (RK4) due to its stability at large time-steps, ease of implementation, and relatively good computational efficiency. Applying RK4 to orbit integration requires separating the second-order differential equation in Eq. (2.44) into two first-order differential equations [30]

$$\dot{\vec{r}} = \vec{v}, \tag{4.14}$$

$$\dot{\vec{v}} = \vec{a}. \tag{4.15}$$

The RK4 algorithm calculates the values of $\vec{r}$ and $\dot{\vec{r}}$ at subsequent time steps as

$$\vec{v}_{i+1} = \vec{v}_i + \frac{h}{6}(\vec{k}_{1,i+1} + 2\vec{k}_{2,i+1} + 2\vec{k}_{3,i+1} + \vec{k}_{4,i+1}), \tag{4.16}$$

$$\vec{r}_{i+1} = \vec{r}_i + \frac{h}{6}(\vec{l}_{1,i+1} + 2\vec{l}_{2,i+1} + 2\vec{l}_{3,i+1} + \vec{l}_{4,i+1}), \tag{4.17}$$

where h determines the size of the time step. The coefficient vectors $\vec{k}_1, \ldots, \vec{k}_4$ and $\vec{l}_1 \ldots, \vec{l}_4$ estimate the slopes of the vector functions $\vec{a}$ and $\vec{v}$ at different points within the time step

$$
\begin{aligned}
\vec{k}_{1,i+1} &= \vec{a}(\vec{r}_i), & \vec{l}_{1,i+1} &= \vec{v}_i, \\
\vec{k}_{2,i+1} &= \vec{a}(\vec{r}_i + \tfrac{h}{2}\vec{l}_{1,i+1}), & \vec{l}_{2,i+1} &= \vec{v}_i + \tfrac{h}{2}\vec{k}_{1,i+1}, \\
\vec{k}_{3,i+1} &= \vec{a}(\vec{r}_i + \tfrac{h}{2}\vec{l}_{2,i+1}), & \vec{l}_{3,i+1} &= \vec{v}_i + \tfrac{h}{2}\vec{k}_{2,i+1}, \\
\vec{k}_{4,i+1} &= \vec{a}(\vec{r}_i + h\vec{l}_{3,i+1}), & \vec{l}_{1,i+1} &= \vec{v}_i + h\vec{k}_{3,i+1}.
\end{aligned}
\tag{4.18}
$$

In order for the algorithm to work, it is imperative that the coefficients are calculated in the order shown in Eq (4.18).

4.2 Albedo inversion algorithm

Using the forward model described above, the inverse problem can now be solved numerically with Markov-chain Monte Carlo methods. In the following, the virtual-observation MCMC algorithm used to do so is described, following [31, 32].

4.2.1 Virtual-observation MCMC

As mentioned in Section 2.4.2, a common MCMC technique for sampling the posterior distribution in (2.62) is the Metropolis-Hastings algorithm which simulates successive transitions in the state space from state $\vec{\mathbf{P}}_i$ to state $\vec{\mathbf{P}}'$ as

$$
\vec{\mathbf{P}}_{i+1} =
\begin{cases}
\vec{\mathbf{P}}' \sim p_t(\vec{\mathbf{P}}', \vec{\mathbf{P}}), & \text{if } u \leq \min(1, a(\vec{\mathbf{P}}_j, \vec{\mathbf{P}}')), \\
\vec{\mathbf{P}}_i, & \text{otherwise,}
\end{cases}
\tag{4.19}
$$

where $p_t(\vec{\mathbf{P}}', \vec{\mathbf{P}}_i)$ is the proposal distribution, $u \in [0, 1]$ is a uniform random variable and a is the Metropolis-Hastings transition kernel

$$
a(\vec{\mathbf{P}}_i, \vec{\mathbf{P}}') = \frac{p_p(\vec{\mathbf{P}}') \, p_t(\vec{\mathbf{P}}_i, \vec{\mathbf{P}}')}{p_p(\vec{\mathbf{P}}_i) \, p_t(\vec{\mathbf{P}}', \vec{\mathbf{P}}_i)}.
\tag{4.20}
$$

The Metropolis-Hastings algorithm however can be inefficient for multivariate probability distributions in that it may require a large number of iterations to converge in the posterior distribution. To make sampling of the posterior more efficient, virtual-observation MCMC algorithm, based on the MH algorithm, is utilized in this thesis. In the algorithm, virtual observations $\vec{\mathbf{m}}_v$ are generated from the original observations $\vec{\mathbf{m}}$ as

$$
\vec{\mathbf{m}}_v = \vec{\mathbf{m}} + \varepsilon_v,
\tag{4.21}
$$

where $\varepsilon_v \sim \mathcal{N}(0, \Lambda_v)$ with Λ_v being the covariance matrix for the virtual observations. This is followed by finding virtual least-squares parameters

$$
\vec{\mathbf{P}}_v = \operatorname{argmin}_{\vec{\mathbf{P}}_v} ([\vec{\mathbf{m}}_v - \mathbf{M}(\vec{\mathbf{P}})]^T (\Lambda_v + \Lambda_\varepsilon)^{-1} [\vec{\mathbf{m}}_v - \mathbf{M}(\vec{\mathbf{P}})]),
\tag{4.22}
$$

For this thesis, a direct search algorithm called the Nelder-Mead method was used, since it does not require the derivatives of the objective function to be known and is available for implementation in the Python `scipy` library.

By denoting the Gaussian probability density function for the real observations as $p(\vec{\mathbf{m}})$, the probability density function for the virtual observations can be expressed as a marginal distribution

$$
p_v(\vec{\mathbf{P}}) = \int d\vec{\mathbf{m}} \, \delta_D(\vec{\mathbf{P}} - \vec{\mathbf{P}}(\vec{\mathbf{m}})) \, p(\vec{\mathbf{m}}),
\tag{4.23}
$$

where δ_D is the Dirac delta function. The proposal distribution for for the parameter difference $\Delta\vec{\mathbf{P}}$ can now be expressed as

$$
\begin{aligned}
p_t(\Delta\vec{\mathbf{P}}) &= \int\int d\vec{\mathbf{P}}\,d\vec{\mathbf{P}}'\,\delta_D(\Delta\vec{\mathbf{P}} - (\vec{\mathbf{P}}_v - \vec{\mathbf{P}}'_v))\,p_v(\vec{\mathbf{P}}_v)\,p_v(\vec{\mathbf{P}}'_v) \\
&= \int d\vec{\mathbf{P}}_v p_v(\vec{\mathbf{P}})\,p_v(\vec{\mathbf{P}}_v - \Delta\vec{\mathbf{P}}),
\end{aligned}
\tag{4.24}
$$

which is a symmetric distribution, meaning that the Metropolis-Hastings kernel simplifies to

$$
a(\vec{\mathbf{P}}_i, \vec{\mathbf{P}}') = \frac{p_p(\vec{\mathbf{P}}')}{p_p(\vec{\mathbf{P}}_i)}
\tag{4.25}
$$

By producing $N_v \gg 1$ virtual observations and incidentally solving for $\vec{\mathbf{P}}_v$ before running the MCMC algorithm, the parameter differences can be obtained from

$$
\Delta\vec{\mathbf{P}}_{jk} = \vec{\mathbf{P}}_{v,j} - \vec{\mathbf{P}}_{v,k},
\tag{4.26}
$$

where $j, k = 1, 2, \cdots, N_v$, $j \neq k$. The transitions in the state space are then readily available from these pre-computed parameter differences during the MCMC sampling, such that any pair of virtual parameters $\vec{\mathbf{P}}_{v,j}$ and $\vec{\mathbf{P}}_{v,k}$ are discarded after being used for calculating the parameter difference. As the number of parameter differences is of the order of N_v^2, a relatively small number of virtual observations, such as 1000 is needed to generate a large number of parameter differences.

Chapter 5

Results and discussion

5.1 Forward model verification

To verify the performance of the ToA model for producing radiation pressure induced accelerations on satellites against previous studies, the radial, transverse, and normal components of the Earth radiation acceleration calculated for the geodetic satellite LAGEOS in Knocke et al. [9] were reproduced. The ToA model used to calculate the accelerations was the polar cap model defined in Eqs. (2.39) and (2.40) with Lambertian reflection coefficient

$$p_0 = 0.34, \quad c_0 = 0.0, \quad c_1 = 0.10, \quad c_2 = 0.0, \quad p_2 = 0.29,$$
$$q_0 = 0.68, \quad k_0 = 0.0, \quad k_1 = -0.07, \quad k_2 = 0.0, \quad q_2 = -0.18,$$

with the epoch t_0 being December 22, 1981. The HEALPix maps of the polar cap models are shown in figures 5.1 and 5.2. The properties of the LAGEOS are found in Table 5.1, where δ is the fraction of diffusely reflected radiation by the satellite surface, as defined in Eq. (2.56). The reproduced acceleration components are shown in Fig. 5.3, which matches well with

Table 5.1 LAGEOS properties

altitude (km)	eccentricity	inclination (deg)	area / mass (m^2/kg)	δ
5893.0	0.004	109.7	0.0007	0.13

the corresponding plot in [9]. The radial acceleration component in Fig. 5.3 demonstrates how the acceleration depends on the satellite position with respect to the Sun and the Earth. At $\Delta = 0°$ the satellite is the closest to the Sun, where it receives the maximum amount of reflected and emitted radiation from the ToA, which shows as the high peak in the figure. The two minima at $\Delta = 90°$ and $\Delta = 270°$ show when the satellite passes the Earth's terminator.

Finally, when the satellite is the farthest from the Sun, it receives mostly or only emitted radiation, depending on the inclination of its orbit. This shows as the smaller peak at at $\Delta = 180°$.
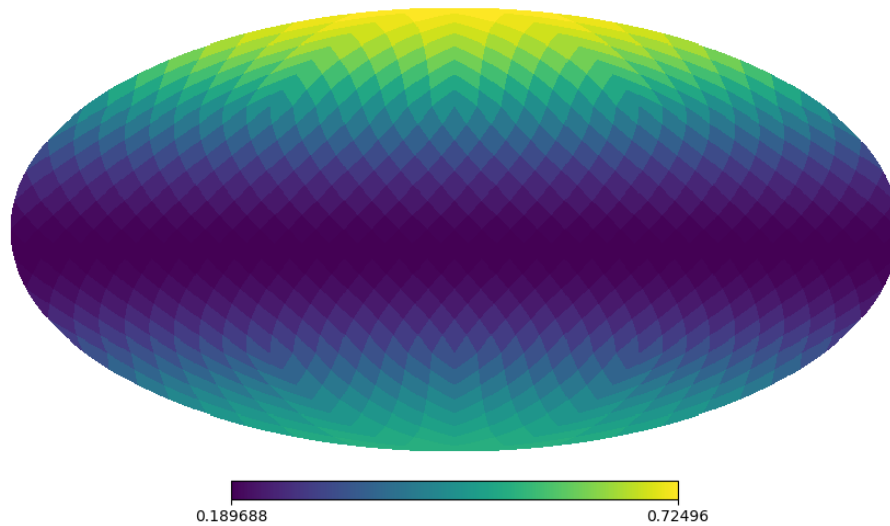

Fig. 5.1 Mollweide projection of the HEALPix reflectance map used to calculate the Earth radiation acceleration for the LAGEOS satellite.
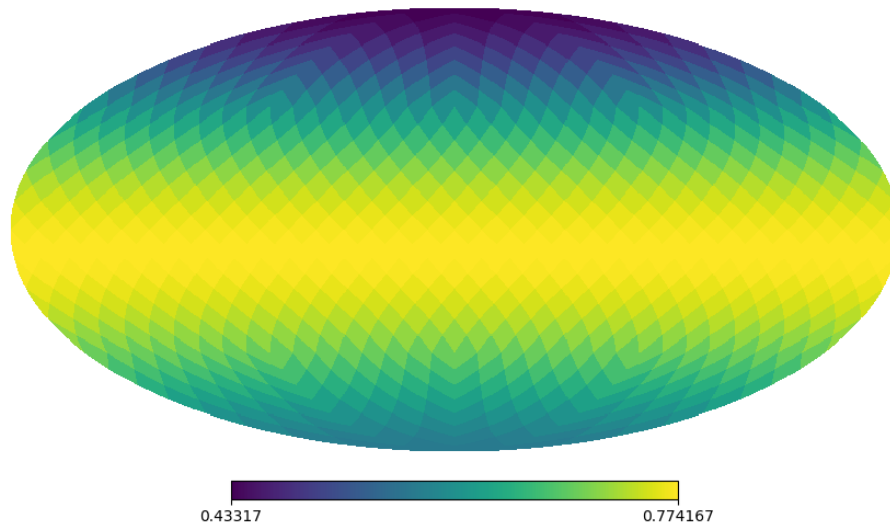

Fig. 5.2 Mollweide projection of the HEALPix emissivity map used to calculate the Earth radiation acceleration for the LAGEOS satellite.
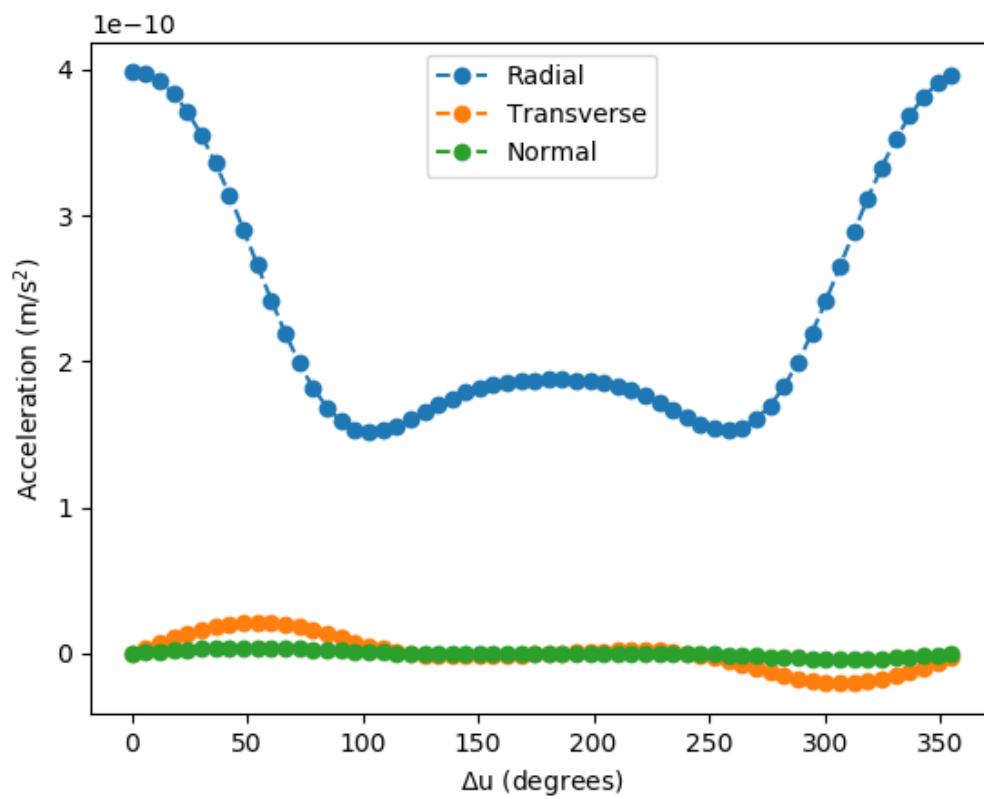
Fig. 5.3 The radial, transverse and normal components of the Earth radiation acceleration on LAGEOS.

5.2   Classified Top of the Atmosphere models

A total of five different classified ToA models were used for solving the inverse problem. These models, labeled M1–M5 in the order of increasing complexity, were created by applying the unsupervised classification scheme introduced in Section 4.1.2 on the `epic_1b_20170101112452_02` dataset. First, ICA was applied to the 10 images of the EPIC dataset (Section 3.1) for feature extraction. As some of the ICA components contained noise or were otherwise irrelevant for this level of classification, four components depicting different surface types and cloud types were selected (Fig.5.4). Then, GMM clustering was applied to the four ICA components to provide classes for the water (W), overcast clouds (C1), dim clouds (C2) and bright clouds (C3), vegetated surfaces (V), savannas (S), and bare soil (B). These classes constituted the most complex model, M5.

To create the simpler ToA models M1–M4, all the land surfaces were combined to create a single class (L). To create a class for all non-cloudy surfaces (N), class L was combined with class W, and to create a class representing all cloudy surfaces, the classes C1–C3 were combined into C0. All the classes in these five models are shown in Fig. 5.4, and the classes each model contains are detailed in Table 5.2. Although the classes in each model represent physical objects, such as clouds or water, the classification itself is not based on physics but rather on common features in the imaging data that the ICA and GMM methods recognise. This is notable especially in the case of cloud classification, where there is no simple way of verifying whether a pixel is correctly or incorrectly identified as part of a cloud. Yet, since the classification algorithm could identify several classes that visually matched features such as clouds, oceans or vegetated surfaces in the images, it offered good grounds for investigating the performance of the albedo inversion algorithm on multiple semi-realistic ToA models.

Table 5.2 Classes used in ToA models M1–M5. The different cloud types are represented by labels C0–C3 such that the classes C1–C3 are merged into C0. The bare soil surface types are represented by B, vegetated surfaces by V, savannas by S, and water by W. The label L represents the merged classes B, V, and S, while the label N stands for non-cloudy surfaces, i.e., the combination of W and L.

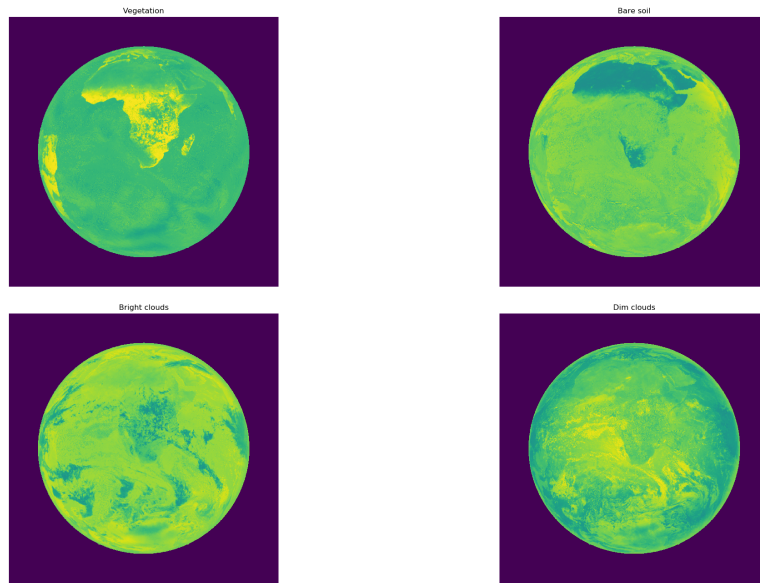| Model	Class									
M1	C0	-	-	-	N	-	-	-	-	-
M2	C0	-	-	-	-	L	W	-	-	-
M3	-	C1	C2	C3	N	-	-	-	-	-
M4	-	-	-	-	N	-	W	V	-	B
M5	-	C1	C2	C3	-	-	W	V	S	B

Fig. 5.4 Four ICA components showing the features for vegetated surface types (up to the left), bare soil (up to the right), bright clouds (down to the left), and dim clouds (down to the right).

5.3   Inverse problem solutions

To solve the inverse problem (see Subsection 2.4.1) from the synthetic dataset (SD, see Section 3.3), the virtual-observation MCMC algorithm was used on the models M1–M5 to infer the reflection coefficients corresponding to each class of a given model. The uncertainty used to forward model the observations was assumed to be Gaussian noise with standard deviations in the same range as the uncertainties in Table 3.2 from Section 3.2. As an example, the model M2 has only two Lambertian reflection coefficients (Eq. (2.15)) $\vec{\mathbf{P}} = [p_N, p_{C0}]^T$. For each model, the algorithm sampled the posterior distribution 5000 times. The run-time had a clear dependence on the number of unknown coefficients, as it increased by half an hour each time a new coefficient was added to the model, such that the run-time for in the case of M1 was one hour while for M5 it was around 3.5 hours. This long run-time was the reason for choosing to use only six satellites to study the inverse problem and to use 768 pixels in the ToA models when solving the inverse problem. The MCMC-derived distributions for the reflection coefficients and the Bond albedos for M1 – M5 are shown in Figs. 5.6 – 5.10, while the corresponding estimates are listed in Tables 5.3 – 5.8.

The uncertainties of the Bond albedo distributions as well as the reflection coefficient distributions of M1, M2, and M3 are relatively similar, whereas for M4 and M5 the width of the distributions clearly increases. But overall, the MCMC estimates obtained for the Bond
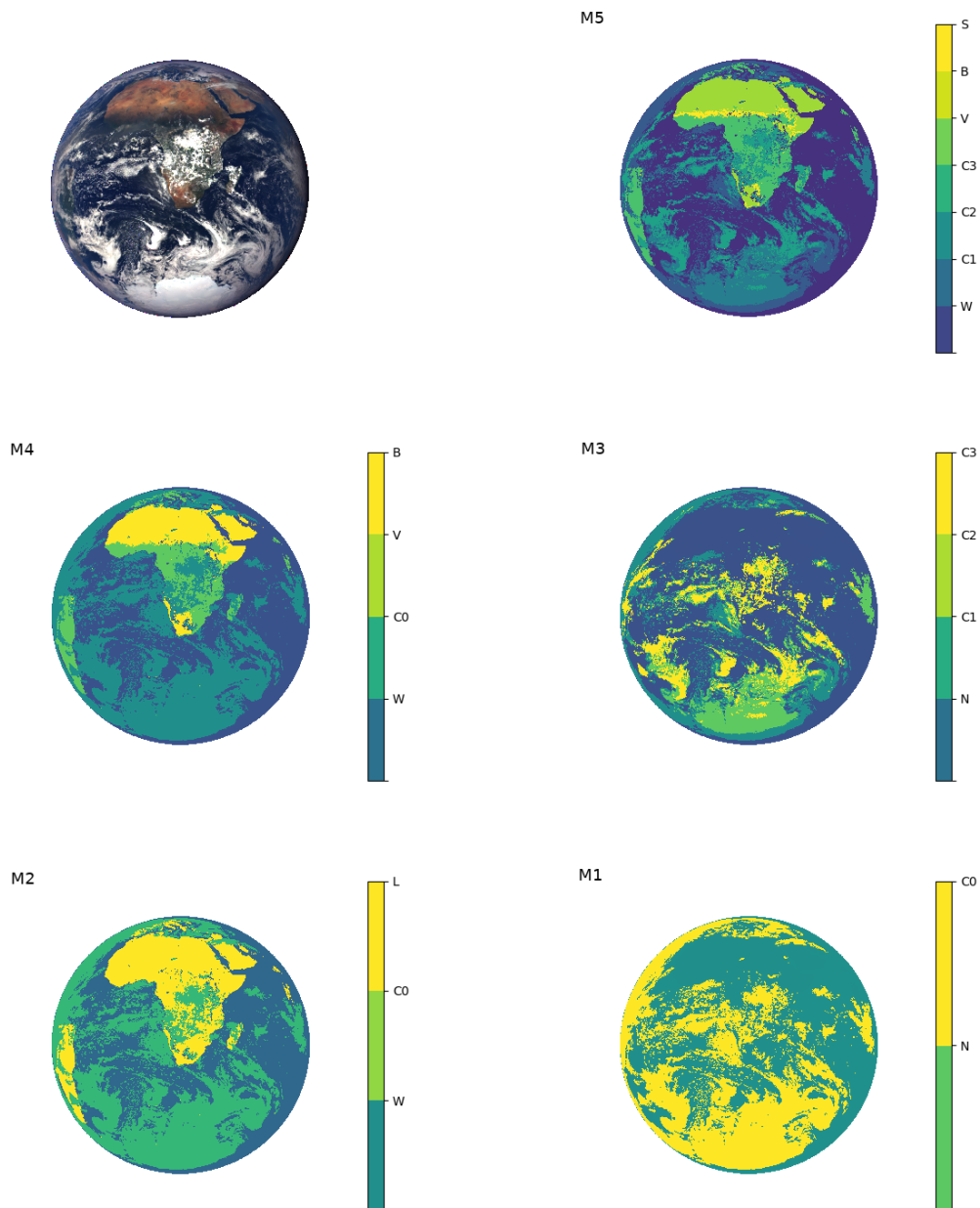
Fig. 5.5 EPIC RGB image at 2017-01-01 11:24:52, and the classification schemes for models M5 to M1 in decreasing order from top to bottom.

albedo are remarkably similar between the models and close to the true value. The similarity between the estimates can be attributed to the similarity between the models as well as their close resemblance to the ToA model used to generate the synthetic data. In simple terms, the classification algorithm finds a clear difference between bright surfaces and dark surfaces in the EPIC data. And while the number of classes varies with each model, the fraction of the number of bright surfaces to the number of dark surfaces remains relatively constant between the ToA models.

What is surprising, however, is how well the albedo inversion algorithm can identify various classes. For nearly every model, the albedo inversion algorithm estimates the reflection coefficients of the classes representing the clouds to have a high value, whereas for the classes representing the various other surfaces the estimated reflection coefficients are much lower, as one would expect. The largest exception to this is M3, which has a relatively low value for the reflection coefficient of the class C1. This can be explained by considering that most of the pixels belonging to this class were located at the edges of the Earth's disc, where the limb effects affect their brightness. For M4 there is a clear overlap in the reflection coefficient distributions between classes W, V, and B, as well as a surprisingly high mean value for each coefficient. The overlap is even stronger for M5, with the reflection coefficient distributions of V and S coinciding almost completely. On the other hand, many of estimated reflection coefficients had a significantly high uncertainty, e.g., for M3, $p_N = 0.045 \pm 0.042$. These uncertainties could indicate that the modeled measurements are not highly sensitive to the values of the reflection coefficients. These results can be due to the low uncertainty of the simulated measurements, the low number of unknown coefficients to be solved in the inversion, and the various simplifications made in this study.

These simplifications naturally raise a question of how realistic the results are? In this regard, there are several factors to be considered, the study of which is left for future research. By far the largest discrepancy between reality and the model was that the ToA was assumed to be static over the 12-hour simulated observation period. In reality, the ToA changes continually due to e.g. the Earth's rotation and the evolution of the cloud cover. A good start towards trying to remedy this discrepancy would be to utilise consequent EPIC images to create a ToA model that updates at the rate of one to two hours.

However, a ToA model constructed from the EPIC images accounts only for the reflected shortwave radiation, while the emitted longwave radiation is left unknown. And as it stands, there is no counterpart for EPIC that provides similar imaging data for longwave radiation. This study assumed a simple Lambertian model for the ToA emissivity that does not depend on latitude, longitude, temperature or time. But the emitted radiation is also a major component of the Earth radiation pressure acting on the satellite, since it perturbs

satellite motion on both sides of the Earth's terminator, as seen in Fig. 5.3. Hence, while this uniform time-independent emissivity assumption simplifies the inversion the Earth's Bond albedo, adding more complexity to the emissivity model could make the inversion significantly harder. Not to even mention the difficulty of trying to invert the total ToA emissivity along with the Bond albedo.

This study also assumed all properties of the satellite to be known, as the goal of this thesis is to act as a feasibility study (see Chapter 1). But, even if all the properties of the satellite materials and the satellite geometry were known in the fullest detail, it is still very difficult to build a complete physical model. For one, neglecting even the tiniest detail can lead to biases whose source is difficult to identify [11, 10]. Secondly, this requires immense computational resources, which for the purposes of solving the inverse problem is highly undesirable [21].

Another factor is the use of a simple gravitational model in this study, which only consists of the two-body term, the J_2 term, and the J_3 term. While significantly more stable than radiation pressure effects in terms of temporal evolution, the lack of more comprehensive modeling of gravitational perturbations does raise the question about their impact on the results of the Bond albedo inversion.

Table 5.3 MCMC-derived Bond albedos for models M1–M5 and the corresponding value from the synthetic dataset (SD).

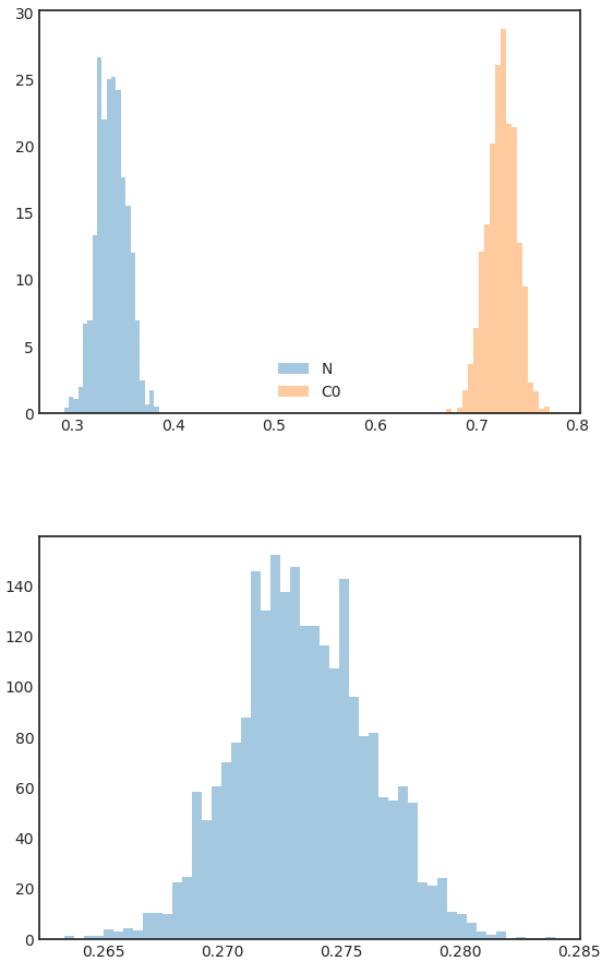| Model	Bond albedo	Uncertainty
SD	0.27335	–
M1	0.27338	0.00748
M2	0.27340	0.00765
M3	0.27326	0.00752
M4	0.27397	0.01393
M5	0.28340	0.01566

Fig. 5.6 M1: reflection coefficient distribution (top), Bond albedo distribution (bottom).

Table 5.4 MCMC-derived reflection coefficients for M1

Reflection coefficient	Value	Uncertainty
p_N	0.339	0.043
p_{C0}	0.725	0.047

Fig. 5.7 M2: reflection coefficient distribution (top), Bond albedo distribution (bottom).

Table 5.5 MCMC-derived reflection coefficients for M2

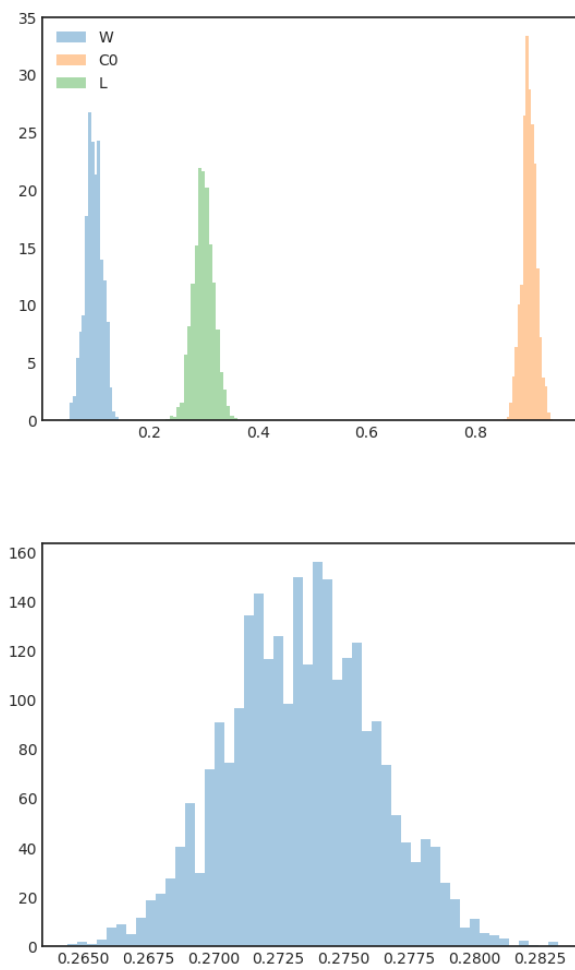| Reflection coefficient	Value	Uncertainty
p_W	0.098	0.041
p_{C0}	0.898	0.039
p_L	0.300	0.046

Fig. 5.8 M3: reflection coefficient distribution (top), Bond albedo distribution (bottom).

Table 5.6 MCMC-derived reflection coefficients for M3

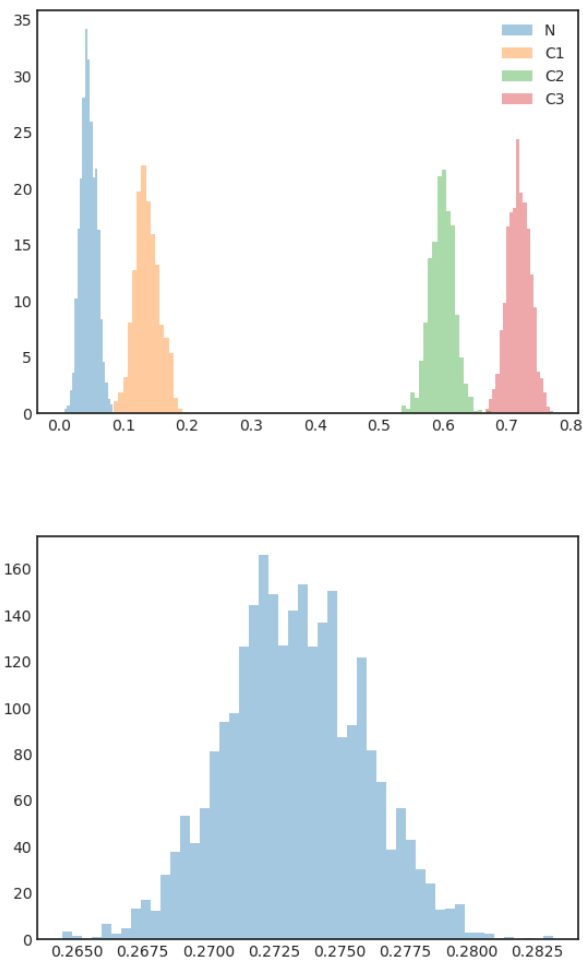| Reflection coefficient	Value	Uncertainty
p_N	0.045	0.042
p_{C1}	0.137	0.051
p_{C2}	0.599	0.052
p_{C3}	0.719	0.049

Fig. 5.9 M4: reflection coefficient distribution (top), Bond albedo distribution (bottom).

Table 5.7 MCMC-derived reflection coefficients for M4

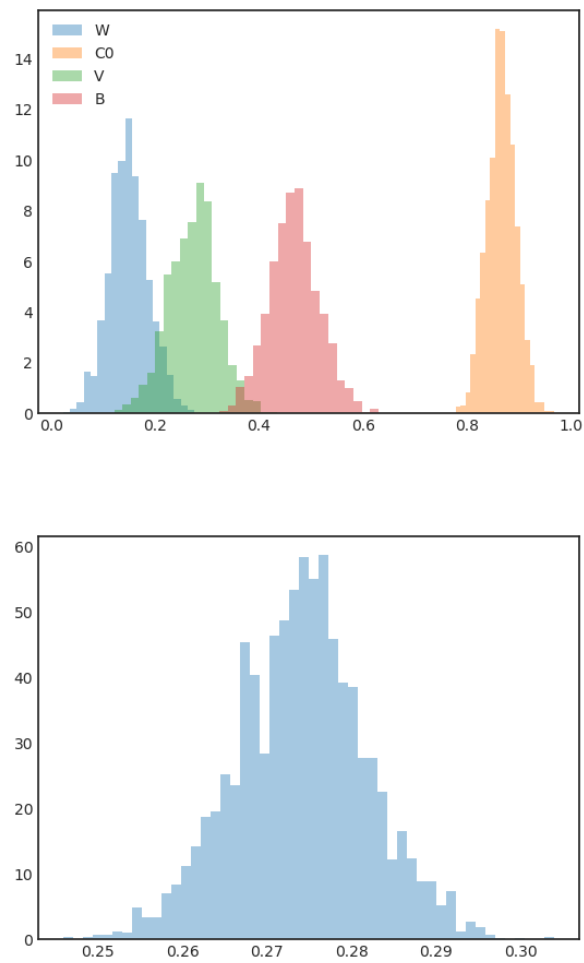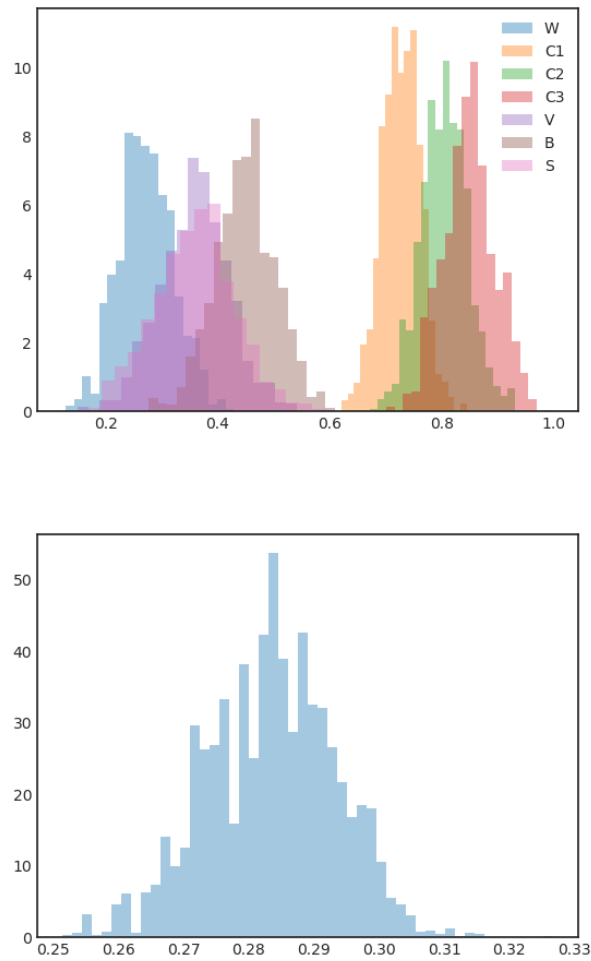| Reflection coefficient	Value	Uncertainty
p_{C0}	0.886	0.073
p_W	0.150	0.103
p_V	0.273	0.110
p_B	0.467	0.112

Fig. 5.10 M5: reflection coefficient distribution (top), Bond albedo distribution (bottom).

Table 5.8 MCMC-derived reflection coefficients for M5

Reflection	Value	Uncertainty
p_{C1}	0.727	0.105
p_{C2}	0.805	0.111
p_{C3}	0.849	0.110
p_W	0.272	0.123
p_V	0.359	0.217
p_S	0.360	0.216
p_B	0.448	0.149

Chapter 6

Conclusions

While multiple attempts have been made to estimate the Earth's total outgoing radiation, the measurements behind these estimates heavily rely on modeling and averaging over long time periods, and as such, pose uncertainties that are difficult to quantify. This thesis provides a novel approach that can significantly improve the uncertainty quantification related to Bond albedo determination and, as far as is known, is the first of its kind to be studied. The strength of this new method lies in that it uses mathematical, not physical modeling for the ToA scattered and emitted radiation properties, combines two different types of data that provide observations from multiple viewing geometries, and use Bayesian inference in the hunt for a single number and its uncertainty.

Although this first study is a rather theoretical one, as it relies on synthetic data, the theoretical principles built from basic concepts of electromagnetic radiation as described in this thesis can be applied to realistic data, provided a software with a more complete forward model is developed. Among a number of things, this requires modeling the gravitational forces acting on the satellites at least on the same level of detail as is done by the Bernese GNSS software along with having a more detailed design for the satellite in the forward model. Then, almost as a side product of this thesis, space-geodetic positioning could be improved from cm-level precision to mm-level precision using this state-of-the-art model.

As for the imaging component of this method, the future software versions would create mathematical model for the outgoing ToA radiance from imaging observations by EPIC or possible future remote sensing missions, making use of spherical wavelets for, e.g., efficient representation of the BRDFs or fast integration of functions on a sphere [33]. For classifying the different ToA scene types, there already exist various remote sensing data products for land type classification. However, for real-time cloud classification there are not as many options. From pre-existing cloud classification data, one could easily train a convolutional neural network to recognise clouds in new imaging data. But with the lack of pre-classified

data, one could also use some threshold value for subtracting clouds from other features in the case of multispectral imaging data such as EPIC [24].

In testing the performance of the inverse problem algorithm on synthetic data, inverse crimes were avoided by solving the inverse problem using different forward model with lower discretisation resolution than what was used for generating the synthetic data. Although the models used for this study were rather simplistic, these tests provided promising first results to the extent that the Bond albedo uncertainties were within reasonable limits compared to estimates by, e.g., CERES. In more detail, all the ToA radiation models that were used to solve the inverse problem had Lambertian surfaces for the overall surface ToA with one to six unknown parameters. The simplest model had uniform reflectivity over the entire globe while the next simplest model had varying reflectance along the latitudes, and the most complex models classified the ToA scenery into dim and bright clouds, oceans, deserts, and areas with vegetation. The virtual-observation MCMC algorithm efficiently inferred the unknown parameters for each of these models, providing accurate estimates for the Bond albedo.

When it comes to software numerical efficiency, the most demanding numerical component is by far the satellite ERP model since it requires ray tracing ToA radiance from tens of thousands of surface elements to each satellite surface at each integration step. This can act as a limiting factor for the model accuracy regarding satellite designs as well as ToA resolution. The first steps towards improving the software performance would be to analyse and optimise numerically heaviest parts of the code, along with including the option for parallel computing to the software.

The method presented in this thesis is of course limited by the quality of the measurement data. In this sense, the current sources of measurement data could be improved upon by designing future satellite missions where an armada of satellites collectively observe the entire ToA from multiple viewing geometries with each satellite capable of simultaneous collection of imaging data on a wide spectral range and precise space-geodetic observations. In addition to this ambitious future prospect for Earth observations, the method described in this thesis could be applied to any other planetary object. A promising future opportunity is the BepiColombo mission to Mercury, starting its observations in 2025, which collects both radiation pressure and imaging data.

References

[1] Bruce A Wielicki, Bruce R Barkstrom, Edwin F Harrison, Robert B Lee III, G Louis Smith, and John E Cooper. Clouds and the Earth's Radiant Energy System (CERES): An Earth observing system experiment. *Bulletin of the American Meteorological Society*, 77(5):853–868, 1996.

[2] Seiji Kato, Norman G Loeb, David A Rutan, and Fred G Rose. Clouds and the Earth's Radiant Energy System (CERES) data products for climate research. *Journal of the Meteorological Society of Japan. Ser. II*, 93(6):597–612, 2015.

[3] Graeme L Stephens, Denis O'Brien, Peter J Webster, Peter Pilewski, Seiji Kato, and Jui-lin Li. The albedo of Earth. *Reviews of geophysics*, 53(1):141–163, 2015.

[4] Norman G Loeb, David R Doelling, Hailan Wang, Wenying Su, Cathy Nguyen, Joseph G Corbett, Lusheng Liang, Cristian Mitrescu, Fred G Rose, and Seiji Kato. Clouds and the earth's radiant energy system (CERES) energy balanced and filled (EBAF) top-of-atmosphere (TOA) edition-4.0 data product. *Journal of Climate*, 31(2):895–918, 2018.

[5] R. Dach, S. Lutz, P. Walser, and P. Fridez. *Bernese GNSS Software Version 5.2. User manual*. Astronomical Institute, University of Bern, Switzerland, 2015.

[6] A. Milani, A. M. Nobili, and P. Farinella. *Non-gravitational perturbations and satellite geodesy*. Adam Hilger, 1987.

[7] Marek Ziebart. Generalized Analytical Solar Radiation Pressure Modeling Algorithm for Spacecraft of Complex Shape. *Journal of Spacecraft and Rockets*, 41(5):840–848, 2004.

[8] Sima Adhya. *Thermal Re-Radiation Modelling for the Precise Prediction and Determination of Spacecraft Orbits*. PhD thesis, University College London, 2005.

[9] P. C. Knocke, J.C. Ries, and Tapley B. D. Earth radiation pressure effects on satellites. *Proceedings of AIAA/AAS Astrodynamics Conference*, pages 577–587, 1988.

[10] Laurent Olivier Froideval. *A Study of Solar Radiation Pressure Acting on GPS Satellites*. PhD thesis, The University of Texas at Austin, 2009.

[11] C. J. Rodriguez-Solano, U. Hugentobler, and P. Steigenberger. Adjustable box-wing model for solar radiation pressure impacting GPS satellites. *Advances in Space Research*, 49(7):1113 – 1128, 2012.

[12] C. J. Rodriguez-Solano, U. Hugentobler, P. Steigenberger, and S. Lutz. Impact of Earth radiation pressure on GPS position estimates. *Journal of geodesy*, 86(5):309–317, 2012.

[13] S Dewitte, L Gonzalez, N Clerbaux, A Ipe, C Bertrand, and B De Paepe. The Geostationary Earth Radiation Budget edition 1 data processing algorithms. *Advances in Space Research*, 41(11):1906–1913, 2008.

[14] E. Pallé, P. R. Goode, V. Yurchyshyn, J. Qiu, J. Hickey, P. Montañés Rodriguez, M.-C. Chu, E. Kolbe, C. T. Brown, and S. E. Koonin. Earthshine and the Earth's albedo: 2. Observations and simulations over 3 years. *Journal of Geophysical Research: Atmospheres*, 108(D22), 11 2003.

[15] Barbara Carlson, Andrew Lacis, Christopher Colose, Alexander Marshak, Wenying Su, and Steven Lorentz. Spectral Signature of the Biosphere: NISTAR Finds it in our Solar System from the Lagrangian L-1 Point. *Geophysical Research Letters*, 2019.

[16] William H Swartz, Steven R Lorentz, Stergios J Papadakis, Philip M Huang, Allan W Smith, David M Deglau, Yinan Yu, Sonia M Reilly, Nolan M Reilly, and Donald E Anderson. RAVAN: CubeSat demonstration for multi-point Earth radiation budget measurements. *Remote sensing*, 11(7):796, 2019.

[17] Norman G Loeb, Seiji Kato, and Bruce A Wielicki. Defining top-of-the-atmosphere flux reference level for Earth radiation budget studies. *Journal of climate*, 15(22):3301–3309, 2002.

[18] O Montenbruck, R Schmid, F Mercier, P Steigenberger, C Noll, R Fatkulin, S Kogure, and AS Ganeshan. GNSS satellite geometry and attitude models. *Advances in Space Research*, 56(6):1015–1029, 2015.

[19] Maxwell B Fairbairn. Planetary photometry: The Lommel-Seeliger law. *Journal of the Royal Astronomical Society of Canada*, 99:92, 2005.

[20] Rosana Montes and Carlos Ureña. An overview of BRDF models. *University of Grenada, Technical Report LSI-2012*, 1, 2012.

[21] Oliver Montenbruck and Eberhard Gill. *Satellite orbits: models, methods and applications*. Springer Science & Business Media, 2012.

[22] Nikolaos K Pavlis, Simon A Holmes, Steve C Kenyon, and John K Factor. The development and evaluation of the Earth Gravitational Model 2008 (EGM2008). *Journal of geophysical research: Solid Earth*, 117(B4), 2012.

[23] Andrea Milani and Giovanni Gronchi. *Theory of orbit determination*. Cambridge University Press, 2010.

[24] Jonathan H. Jiang, Albert J. Zhai, Jay Herman, Chengxing Zhai, Renyu Hu, Hui Su, Vijay Natraj, Jiazheng Li, Feng Xu, and Yuk L. Yung. Using Deep Space Climate Observatory measurements to study the Earth as an exoplanet. *The Astronomical Journal*, 156(1):26, jun 2018.

[25] Robert W Youngren and Mikel D Petty. A multi-resolution HEALPix data structure for spherically mapped point data. *Heliyon*, 3(6):e00332, 2017.

[26] Krzysztof M Gorski, Eric Hivon, Anthony J Banday, Benjamin D Wandelt, Frode K Hansen, Mstvos Reinecke, and Matthia Bartelmann. HEALPix: a framework for high-resolution discretization and fast analysis of data distributed on the sphere. *The Astrophysical Journal*, 622(2):759, 2005.

[27] Kevin P Murphy. *Machine learning: a probabilistic perspective*. MIT press, 2012.

[28] Michael F Modest. *Radiative heat transfer*. Academic press, 2013.

[29] Casper HL Beentjes. Quadrature on a spherical surface. *Working note available on the website http://people. maths. ox. ac. uk/beentjes/Essays*, 2015.

[30] C. J. Voesenek. Implementing a Fourth Order Runge-Kutta Method for Orbit Simulation, 2008.

[31] Karri Muinonen, Mikael Granvik, Dagmara Oszkiewicz, Tuomo Pieniluoma, and Hanna Pentikäinen. Asteroid orbital inversion using a virtual-observation Markov-chain Monte Carlo method. *Planetary and Space Science*, 73(1):15–20, 2012.

[32] Karri Muinonen, Olli Wilkman, Alberto Cellino, Xiaobin Wang, and Yibo Wang. Asteroid lightcurve inversion with Lommel–Seeliger ellipsoids. *Planetary and Space Science*, 118:227–241, 2015.

[33] G. Xu, M. Gritsevich, Penttilä A. Peltoniemi, J. and, O. Wilkman, O. Ihalainen, and K. Muinonen. Spectral reflectance processing via local wavelength-direction correlations. *IEEE Geoscience and Remote Sensing Letters*, in press.

Appendix A

Conventions

Quantities

Table A.1 Spectral radiometric terms and symbols according to International Organization for Standardization (ISO)

Quantity	Symbol	Units
Radiance	L	$\text{W m}^{-2}\,\text{sr}^{-1}$
Radiant flux	Φ	W
Flux density	F	W m^{-2}
Irradiance	E	W m^{-2}
Radiant exitance	M	W m^{-2}
Radiosity	J	W m^{-2}
Disk-integrated brightness	I	W sr^{-1}

Mathematical conventions

Table A.2 Mathematical notation used in this thesis

Symbol	Object
$\vec{\mathbf{a}}$, $\vec{\mathbf{F}}$	vector
$\hat{\mathbf{e}}$	unit vector
$\mathbf{A}$,	Matrix
$[]^{T}$	Matrix transpose
$(\dot{\,})$	Time derivative $\frac{d}{dt}$