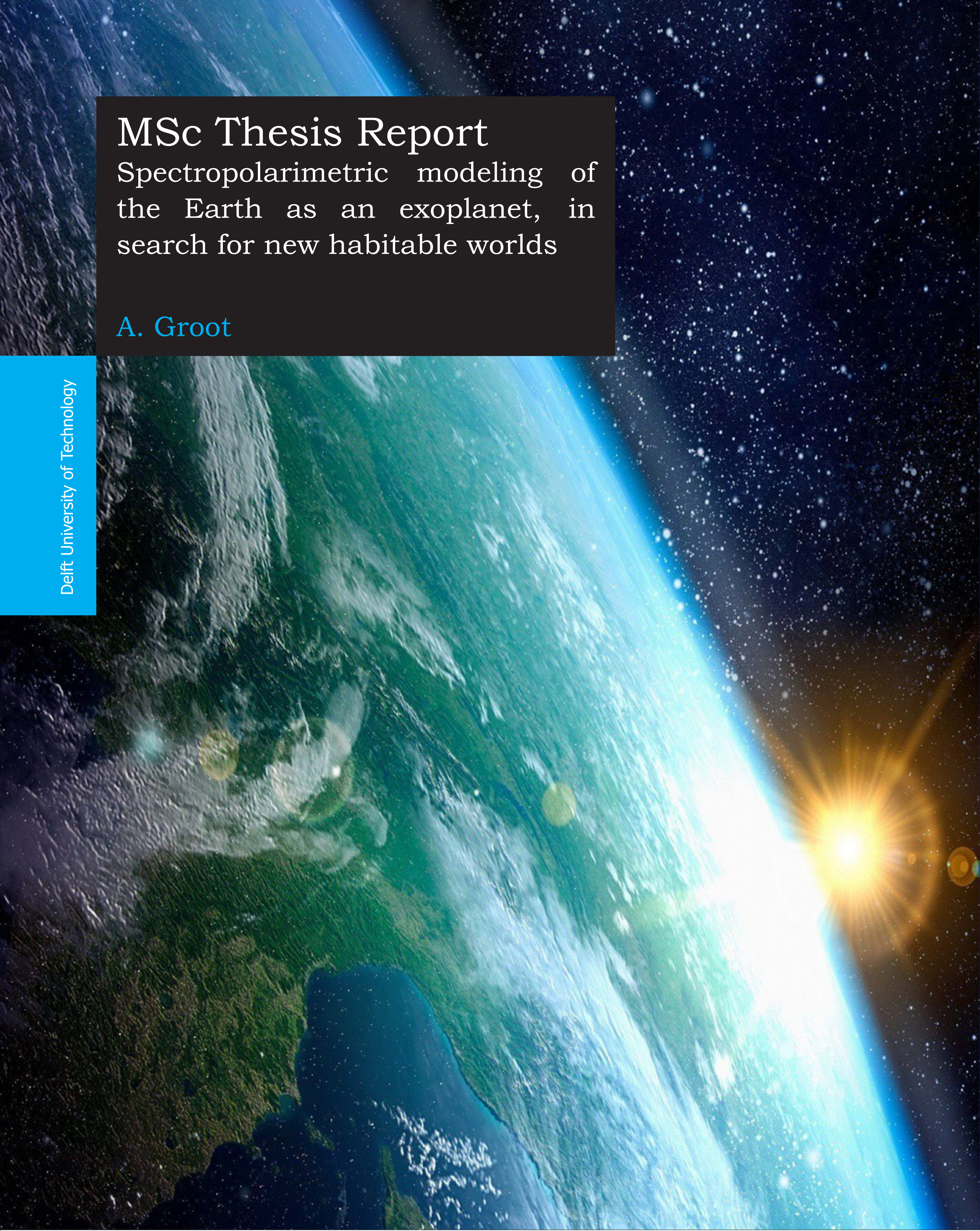


MSc Thesis Report

Spectropolarimetric modeling of the Earth as an exoplanet, in search for new habitable worlds

A. Groot

Delft University of Technology

TU Delft Delft University of Technology

Challenge the future

MSc Thesis Report

Spectropolarimetric modeling of the Earth as an exoplanet, in search for new habitable worlds

by

A. Groot

to obtain the degree of Master of Science,
at the Delft University of Technology,
to be defended publicly on Friday September 28, 2018 at 09:30.

Student number:	4233239	
Printed:	September 14, 2018	
Thesis committee:	Prof. Dr. L. L. A. Vermeersen	TU Delft, committee chairman
	Dr. L. C. G. Rossi	LATMOS, daily supervisor
	Dr. D. M. Stam	TU Delft, daily supervisor
	Dr. A. Menicucci	TU Delft, committee

An electronic version of this thesis is available at http://repository.tudelft.nl.

Cover page image is retrieved from https://xshyfc.com

Delft University of Technology

Abstract

Previous studies have investigated the remote appearance of Earth-like exoplanets in the prospect of retrieving biosignatures of planets orbiting extrasolar stars, utilizing the variation of (polarized) flux. However, these studies did not use horizontally inhomogeneous models that include (1) daily cloud observations of planet Earth accounting for different cloud parameters (i.e. using real data about the cloud optical thickness, cloud top pressure and effective size parameter of the cloud droplets) with (2) an underlaying Earth-like surface cover, for (3) a set of wavelengths covering the ultra-violet, visible and near-infrared spectral domain. We present simulations of spatially resolved disks and planetary phase curves of the total flux, degree of polarization and linearly polarized fluxes. We discuss the presence of spectropolarimetric signatures that can potentially be directly retrieved from future observations. Moreover, in the design of future telescopes the characteristics of these signatures may be considered.

The contribution of either the surface or cloud cover to the (polarized) reflected flux by the exoplanet depends on the considered wavelength. Hence, the signatures that may indicate the presence of liquid water particles suspended in the exoplanet atmosphere, namely the glory and primary rainbow, vary in strength. In particular, both features are visible in the total flux, degree of polarization and polarized flux Q, where the primary rainbow is the most likely candidate to be retrieved, however, the daily variation and seasonality in the cloud observations may suppress its enhancement in the total flux.

In a previous study, it was shown that ocean exoplanets may potentially be characterized by the color reversal in the planetary phase curves of the polarized flux Q. We show that, when Earth-like continents are introduced, this intersection in the planetary phase curves, corresponding to various wavelengths, may still be observed in the presence of an Earth-like ocean and is absent in the absence of an Earth-like ocean. We show that the continents do not affect the location of the intersection point but induce rapid oscillations in the planetary phase curves. Hence, we show that the cloud fraction can still roughly be estimated from the planetary phase angle where this intersection point is located. Alternatively, for the Earth-like vegetation and desert surfaces we are not successful in finding an unambiguous signatures in the planetary phase curves.

Using our planetary model, we attempted to fit Earthshine observations, i.e. measurement of the degree of polarization of the reflected Earthshine by the Moon. Our simulations show moderate agreement for all λ, which can be caused by (1) neglecting the presence of other aerosols, such as maritime aerosols, or (2) the approximation of the correction for the depolarizing behaviour of the Lunar surface.

We conclude that utilizing a set of wavelengths in the visible and near-infrared domain, could potentially allow one to retrieve information about the presence, abundance and micro-physical properties of clouds in the atmosphere of, and also the presence of an ocean cover on, an Earth-like exoplanet.

Contents

Acknowledgements

This document describes the spectropolarimetric signals originating from a potential Earth-like exoplanet. The research on this project was truly amazing, even though it required quite some self-discipline to stay focused on the main objective, and motivation to work till late and through the weekends to finish this research successfully. Of course, this was not possible without the help and support of some important people.

First of all, I would like to deeply thank Dr. Loïc. C. G. Rossi for supervising me during this thesis project as well as the preliminary literature study. He allowed me to be creative and independently manage the project, but gave critical remarks when necessary. Every time I went home after our long discussions I felt greatly inspired, contributing enormously to the success of my research. I also greatly thank Dr. Daphne. M. Stam for allowing me to carry out this topic, and most of all to supervise me after Loïc left the faculty. My discussions with her and her review on my thesis greatly aided in the overall success. As part of this research I required lots of computation effort, for which I want to thank Amanthla Biekman and Javier Berzosa Molina for using their server accounts. I greatly thank Prof. Dr. Bert L. A. Vermeersen and Dr. Alessandra Menicucci for their time and effort to review this thesis and for their assessment during my defence.

The road to finishing this masters has been long and exhausting, but also gave me lots of joy. I deeply thank my friends and fellow students Almamdouh Omar, Joey Cheung and Max Witteman for our incredible time in the last few years. Chasing my interests in science and especially space exploration would not have been possible without the incredible support of my parents, both financially, mentally, ..., actually in any way they could. I also deeply thank my girlfriend, Joycie Visser, who supported me by being extremely compassionate during periods that I devoted a lot of time to my education.

Aiding in the success of my research I want to greatly thank Joey Cheung and Victor Trees, for sharing the models and expertise they obtained in their research. Also, I thank Victor Trees for the long discussions we had about both our research topics and his critical review on my results. I am sure that sharing our thoughts made us think about our results in a different way, what I think is extremely valuable in this research area.

A. Groot
Delft, September 2018

1
Introduction

As long as civilized societies have existed, mankind has explored the boundaries of what was thought to be unreachable. In 1492 Christoffel Columbus explored the boundaries of Earth, which many people believed to be flat at the time, and found The New World, i.e. America. Although he thought that to have reached India, he crossed the entire Atlantic Ocean and enabled worldwide colonization. In the 19th century World War II brought harm to millions of people, nonetheless Germany was a pioneer in launching human made objects into space as they developed the V-2 rocket that reached 84.5 kilometres on October 3rd 1942. After WWII these series of ballistic missiles were used by the United States and the Soviet Union to further explore space. During the era of space exploration, the Soviets were the first to successfully cross the boundary of space as they launched the *Sputnik 1* in an orbit around Earth on the 4th of October 1957. Shortly after, this achievement was outdated by multiple American and Russian satellite programs. Although several animal space flights were already successful, the next step in exploration was the first human space flight. Yuri Gagarin left our atmosphere on the 12th of April 1961. After multiple interplanetary missions in our solar system, like the Venera, Mariner and Pioneer programs, Voyager 1 and 2 were developed to push the boundary of exploration to unimaginable distances.

The first telescope that allowed us to accurately picture the far boundaries of space is Hubble. Since Hubble was launched on April 24th 1990, remarkable pictures of other galaxies have been taken. In that same period the search for planets orbiting stars in extrasolar systems (i.e. exoplanets) started, when a Jupiter-like planet was found around the main sequence solar-type star *51 Pegasi*, using the *Radial velocity technique* (Mayor and Queloz 1995). As of today over 3700 exoplanets have been detected and confirmed in over 600 multi-planetary systems[1], with various types of observation methods. Additionally, almost 4500 exoplanet candidates have been identified by NASA's Kepler mission, with several of them being Earth-sized planets located in the habitable zone around Solar-type stars[2]. Of the confirmed exoplanets a majority is investigated via indirect methods providing important planet parameters such as its radius, its minimum mass, and its orbital period. Also, atmospheric components have been retrieved with spectroscopy during secondary eclipses and planetary transits (Swain et al. 2009, 2008; Tinetti et al. 2007). However, Earth-like exoplanets in habitable zones are relatively small and their transits so rare that no sufficient signal-to-noise ratio can be acquired (Kaltenegger and Traub 2009). One of the most recently and closest found Earth-like exoplanet is TRAPPIST-1e, orbiting around the star TRAPPIST-1. This potentially habitable planet is located 40 light years from Earth and has a similar radius and equilibrium temperature as Earth, thus located in the habitable zone. In particular, a planet is called habitable if it lies in the habitable zone with conditions similar to Earth, favouring the existence of water-based Earth-like life (Lammer et al. 2009).

At present and in the near future, direct observations of exoplanets are and will be spatially unresolved, i.e. one pixel images, because even the largest telescopes have insufficient spatial resolutions

[1]Obtained from: `http://exoplanetarchive.ipac.caltech.edu/`.
[2](*NASA Releases Kepler Survey Catalog with Hundreds of New Planet Candidates* n.d.)

at such large distances (Hoeijmakers et al. 2016). Ford et al. (2001); Stam et al. (2006) showed that spatially unresolved spectroscopy of reflected light from exoplanets by their parent star can characterize atmospheres and, if present, surfaces. A number of instruments that are being designed or used are able to obtain such observations, for example Earth based such as SPHERE (operational on the Very Large Telescope (VLT)), EPICS (which will be designed for European Extremely Large Telescope (E-ELT)) and GPI (operational on the Gemini South Telescope), and space based such as New Worlds Observer (NWO), Wide Field Infrared Survey Telescope (WFIRST) and James Webb Space Telescope (JWST) from NASA.

A promising addition to direct observations of reflected starlight is polarimetry. Early attempts to detect ”hot jupiters” type exoplanets have already been made (Hough et al. 2006; Lucas et al. 2006). Seager and Sasselov (2000); Stam (2008*a*); Stam et al. (2004*b*) show that a combination of reflected flux and linear polarized fluxes provides an extra dimension in retrieving signatures which can help in the characterization of these potential habitable worlds. In particular, polarized light carries information about the source from which it is scattered, such as dust, liquid water particles or an ocean surface. Additionally, as light from stars is naturally unpolarized and scattered light by an exoplanet surface and/or atmosphere generally is not, it is possible to improve resolving the planet from its parent star (Hoeijmakers et al. 2016). To obtain such observations, space based telescopes are most favourable because Earth’s atmosphere does influence the polarization of light, albeit that ground based telescopes with adaptive optics can also be used (Gisler et al. 2004; Saar and Seager 2003; Schmid et al. 2006).

1.1. Hypothetical relevance and contribution to the scientific community

In the last decade a significant amount of work has been published about the simulation results of Earth-like models. Most of these studies provide only photometric results and are focused on a signature specifically related to the presence of an ocean, vegetation or clouds individually. In recent years, the addition of polarization in models and observations, however, has shown to be an indispensable tool for exoplanet characterization.

In the characterization of oceans especially the presence of a glint in spectropolarimetric simulations seems to provide an unambiguous signature (Williams and Gaidos 2008). These models do, however, only only simple cloud models, no gaseous atmosphere and no ocean albedo, which might overestimate the strength of the ocean glint in reality. A more diverse study by Zugger et al. (2010*a*) provided the shift of the peak of polarization for cases with increasing optical depth of the gaseous atmosphere, the presence of ocean winds, the interference of clouds and the effect of maritime aerosols. The study is fully devoted to ocean planets, and the signatures found might not be present with an Earth-like continental surface distribution. Also, no variability in clouds is included and the clouds are modeled again as Lambertian reflectors without a gas layer on top.

A study on the presence of the vegetation’s ”red-edge”, an enhancement in the albedo of vegetation in the near-infrared, is modeled with a realistic cloud cover and an enhanced radiative transfer code by Montañés-Rodríguez et al. (2006). They find that enhancements are visible in the disk integrated spectrum, but cannot be associated with vegetation unambiguously without knowing the cloud distribution in advance. These results are obtained by photometric modelling and observations only. A similar study, based on four observations also shows that vegetation only induces a small enhancement in the photometric signal, especially when comparing its magnitude to absorption bands such as O_2 and O_3 gaseous absorption lines. A recent study by Berdyugina et al. (2016) shows that a similar feature of the ”red-edge” in photometric signals can also be observed in linearly polarized signals. These results were obtained by lab measurements and used later on in an Earth-like model that includes clouds and other surface types. In this configuration a rather unambiguous detection of photosynthetic pigments is found.

In general, one can approximate liquid cloud particles by spheres. If such spherical aerosols are present in an atmosphere, a glory, the primary and secondary rainbow may be present in both photometric and polarized signals. Bailey (2007) mainly investigated the strength and shape of the primary rainbow from single scattering and found that depending on the nature of particles the rainbow may shift to different phase angles. Also, he investigated the width and amplitude of this signature. The practical effect of spherical liquid water particles in the form of clouds are analyzed by Karalidi, Stam and Hovenier (2012); Karalidi et al. (2011). They used different cloud models in quasi homogeneous and horizontally inhomogeneous simulations with a realistic gaseous Earth-like atmosphere. They, however, did not model an Earth-like surfaces distribution, instead they used a black horizontally homogeneous surface. With these models they showed that variability in the rainbow feature is apparent for different particles sizes. The interesting question arises how spectropolarimetric signals evolve when applied to an inhomogeneous Earth-like model that has a variable cloud cover on top of an Earth-like surface distribution. Clouds in the form of patchy covers were modeled by Rossi and Stam (2017), who attempted to distinguish different types of cloud covers that can exist on exoplanets. The results show that a distinction between cloud covers can be made and an estimate of the total cloud fraction on the planetary disk can be retrieved with reduced ambiguities from the polarized signal.

In order to develop a more realistic Earth-like model, Muñoz (2015) simulated an Earth-like exoplanet with realistic surface albedos and cloud fractions according to MODIS data. The gaseous atmosphere is Earth-like and the clouds are modeled with a single cloud model that consists of a constant effective radius and optical thickness, but with a spectrally varying refractive index of the liquid water particles. With this model, multiple cases for cloud free, patchy-clouded and fully clouded atmospheres were computed for large sets of phase angles, sub observer longitude and wavelength regions covering the visible and infrared regions.

Based on photometric data obtained with the Deep Space Climate Observatory (DSCOVR), Jiang et al. (2018) recently simulated Earth as a proxy exoplanet. This data set comprises two years of observations for multi-wavelength in the ultra-violet, visible and near-infrared region. With these data an attempt is made to retrieve surface types, cloud patterns and the planetary rotation. The use of reflected light signals at multiple wavelengths, that evolve in time, shows to be a valuable tool in the characterization of Earth-like exoplanets. Moreover, polarization is not included in this study as it is based on photometric data only.

A more comprehensive and realistic model that includes the time evolution of clouds in an inhomogeneous 3D configuration for a full range of wavelengths in the ultra-violet, visible and near-infrared regions might provide stronger signatures that reveal important characteristics of Earth-like exoplanets. Utilizing a set of wavelengths allows for the analysis of the gradual spectral behaviour of scatterers. For example, Rayleigh scattering in a pure gaseous atmosphere is most effective at ultra-violet wavelengths, whereas Mie scattering in clouds is approximately equally effective at all wavelengths. Modeling time evolving clouds and a spatial inhomogeneous surface cover are expected to induce, while the planet rotates around its axis, rapid oscillations in the planetary phase curves. Combining the complexity of both cloud and surface covers based on Earth observations in combination with a gas atmosphere, at different wavelengths, will contribute to the understanding of the photometric and polarimetric signals originating from Earth-like exoplanets.

1.2. Lunar Observatory for Unresolved Polarimetry of Earth

The fact that polarization becomes an increasingly valuable tool in exoplanet characterization encouraged scientists to start the development of LOUPE, the Lunar Observatory for Unresolved Polarimetry of Earth (Karalidi, Stam, Snik, Bagnulo, Sparks and Keller 2012). This entails the placement of a polarimeter on the Lunar surface that faces Earth, allowing direct observation of the Earth. Hoeijmakers et al. (2016) presents the most recent and advanced design of LOUPE's polarimeter, which combines a spectral modulation and a micro-lens array. The micro-lens array essentially splits the observed object into multiple pixels, provided that the object is sufficiently close to the instrument. The spectral modulation is optimized for linear polarization alone (no circular polarization); more detailed information can be obtained from Snik et al. (2009).

The main objective of LOUPE for exoplanetary research is to generate disk-resolved spectra of the Earth, which can be disk-integrated, providing benchmark data for future exoplanetary observations.

In addition, the aim is to understand the different effects of Earth-like surfaces and atmospheric features in the flux and polarimetric spectra from these potential observations. This can now only be done by radiative transfer models, Earthshine observations and POLDER/PARASOL data. The latter two are far from ideal as Earthshine observations do not provide spectra for a full range of phase angles, require globally spread ground based telescopes, which introduce additional atmospheric interference and uncertainty in depolarization of the Moon's surface. The POLDER/PARASOL instruments are in low-Earth orbit, observing only small parts of Earth and thus cannot provide an image of the entire Earth disk (Hoeijmakers et al. 2016; Karalidi, Stam, Snik, Bagnulo, Sparks and Keller 2012). If successful, LOUPE will allow to observe Earth at all times, at all phase angles, during a full diurnal Earth rotation and possibly spanning over Earth's seasons as being in an almost edge-on orbit. In this research we will assess which wavelength regions are important for Earth-like exoplanet characterization and present results which can be compared to future observations.

1.3. Research framework

The research goals for this thesis research are based on research objective and research questions.

The research objective is stated as:

> *The research objective is to retrieve spectropolarimetric signals from an Earth-like exoplanet model in an edge-on configuration to be able to rationalize future disk integrated observations, by use of a radiative transfer algorithm in combination with Earth observations.*

From this research objective, we state the following *Central Research Questions*:

1. What are the spectropolarimetric signals for a resolved and unresolved Earth-like exoplanet?
2. Which signatures from spectropolarimetric signals can be identified such that Earth-like exoplanets can be characterized?

Each *Central Research Question* can be discriminated into *Sub Research Questions* as follows:

1. What is the spectropolarimetric signal for a resolved and unresolved Earth-like exoplanet?
   (a) How does light reflect from an Earth-like exoplanet?
   (b) How can an Earth-like exoplanet be modeled?
   (c) Which features characterize Earth and how can these features be used in future exoplanet characterization?
2. Which signatures from spectropolarimetric signals can be identified such that Earth-like exoplanets can be characterized?
   (a) Can Earth biomarkers be characterized in spectropolarimetric signals?
   (b) Can spectropolarimetric signatures characterize exoplanet surfaces?
   (c) Can spectropolarimetric signatures characterize exoplanet atmospheres?
   (d) How can spectropolarimetry be used to identify planetary and orbital elements?

The answer to these research questions will be found by taking the following steps:

1. Create a radiative transfer model for Earth-like exoplanets that incorporates Earth Observations.
2. Create an Earth-like model that allows one to efficiently and accurately compute spectropolarimetric signals.
3. Obtain and characterize the spectropolarimetric signals from the Earth-like model.
4. Assess the variability of the Earth observations in the spectropolarimetric signal.
5. Analyze and retrieve the presence of correlated parameters in the spectropolarimetric signal.
6. Extend the Earth-like model to incorporate polarizing surfaces.
7. Obtain and characterize the spectropolarimetric signals of the extended Earth-like model.

1.4. Outline of this thesis report

This thesis is separated into two parts. In the first part, we describe our radiative transfer model, supported by theoretical methods and algorithms and further practical considerations. In the second part, we use this model to present results and associated discussions.

In the first part, the principles of radiation and polarized light, and the methods to describe the transfer of light in a terrestrial atmosphere, are presented. These various techniques are combined in a radiative transfer code and developed further to be able to appropriately model Earth (*Chapter* 2). In *Chapter* 3 we describe our Earth model. This Earth model is based on high temporal observations and is created in such a way that we still provide accurate results at a reduced cost of computation effort.

In *Chapter* 4 we start the second part, were we present and discuss some straightforward results. The quest for finding correlations between the varying parameters in our Earth model and the retrieved spectropolarimetric signals is described in *Chapter* 5. These first two chapters of the second part are based on an Earth model that contains a simple surface approximation. In the third chapter of this part, *Chapter* 6, we introduce and incorporate anisotropic polarizing surfaces in our Earth model, to produce an even more comprehensive and detailed model. With this model we will simulate similar spectropolarimetric signals as in *Chapter* 4 to be able to compare and discuss the shortcomings of such a simplified model and the possible new signatures in the extended Earth Model. Finally, in *Chapter* 7, we compare our most comprehensive model to actual polarization observations of Earth, and discuss the possible limitations of our model and the observations.

In *Chapter* 8 we provide a review of our findings in the research on which we base the most important conclusions. Some points of improvements that came up during the thesis research will be provided as recommendations in *Chapter* 9.

2

Modelling scattered light curves of Earth-like exoplanets

Before we present the Earth model, a description of the theory that drives this model is presented. In this research, radiative transfer of reflected light by a terrestrial planet is computed using the doubling-adding method of de Haan et al. (1987*a*). First the basic principles of polarized light and some definitions are discussed, which will be used in the remainder of this report. Under the influence of a model atmosphere, light will or will not be scattered. Such an atmosphere is approximated by several layers that define the scattering behaviour of that single layer, essentially stacked together to form the entire model atmosphere. For the scattering behaviour of a specific layer we will describe the most general forms of scattering in a terrestrial atmosphere (*Section* 2.1). All this theoretical work is put together into a computing environment, which is called the Python Mie Doubling Adding Program (PyMieDAP) (Rossi et al. 2018). A description of this code is provided in *Section* 2.2. This code is used as a base for our computations, but is modified to include Earth modelling. It has to be noted that no alterations will be made in the radiative transfer routines, but only to the Python based interface (*Section* 2.3).

2.1. Radiative Transfer on Terrestrial Planets

Basic principles of light and radiation

Light consists of plane electromagnetic waves of which one is shown in *Figure* 2.1. Each of these waves is completely polarized and quasi-monochromatic as described by the solution of the *Maxwell's equations* (van de Hulst 1957). For each light ray we define the wavelength (λ) as the length between two subsequent peaks in the electric/magnetic field. The radiance and state of polarization that is

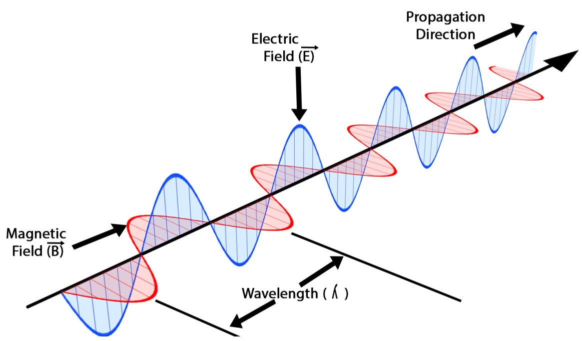


Figure 2.1: Representation of an electromagnetic wave. The red waves represent the magnetic field component and the blue wave represents the electric field component. The length of one period of either the magnetic or electric field is defined by λ.

reflected by an exoplanet is fully described by the Stokes vector I:

$$\mathbf{I} = \begin{bmatrix} I \\ Q \\ U \\ V \end{bmatrix} \tag{2.1}$$

where I is the total radiance, Q and U describe the linearly polarized radiance in mutually perpendicular directions with respect to a reference plane, and lastly V is the circularly polarized radiance. All quantities are expressed in $W\ m^{-2}\ sr^{-1}$ (or $W\ m^{-3}\ sr^{-1}$ to include the wavelength dependence). For all quantities expressed in $W\ m^{-2}$ (or $W\ m^{-3}$ to include the wavelength dependence), i.e. the irradiance or flux vector, we will use the following expression:

$$\pi\mathbf{F} = \pi \begin{bmatrix} F \\ Q \\ U \\ V \end{bmatrix} \tag{2.2}$$

For starlight that is integrated over the entire stellar disk, the electric vector does not have a preferred

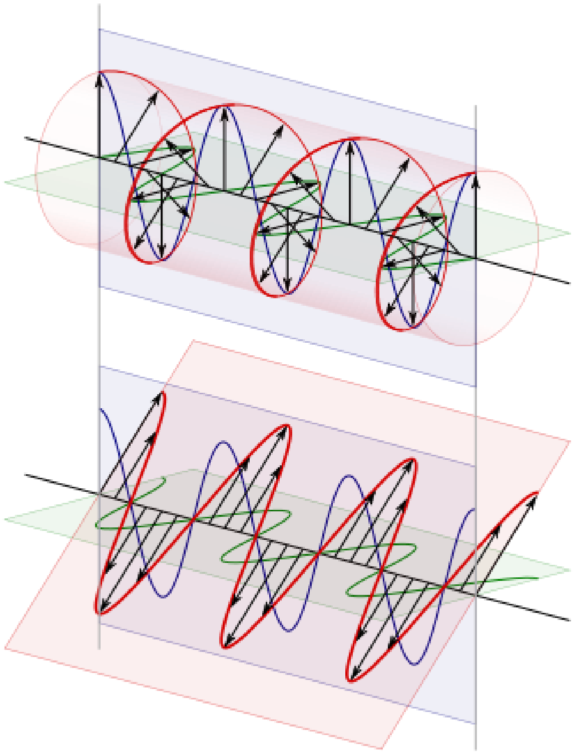


Figure 2.2: Illustration of polarized light. The top panel shows the principal characteristic of circularly polarized light, where the blue and green line represent the electric field of two light rays. The red line describes the superimposed electric field vectors of the two light beams. The lower panel shows the same curves for a linearly polarized light.

direction of vibration and is unpolarized (Hansen and Travis 1974*a*; Kemp et al. 1987). More specifically, if light is unpolarized the electric vector traces out no specific pattern in the plane perpendicular to the propagation direction. If light is polarized linearly or circularly, the electric field vector traces out a line or a circle, respectively. The concept of linear and circular polarized light is shown in the lower panel of *Figure* 2.2, where the blue and green curves represent the two different light waves. In case of linearly polarized light two orthogonal waves are in phase. If these waves are superimposed on each other, the resulting electric field vector traces out a linear pattern in time. When the phase between the two orthogonal light rays are exactly a quarter period out of phase, the resultant electronic field traces out a circular pattern in time (upper panel of *Figure* 2.2). To get more insight into the different Stokes

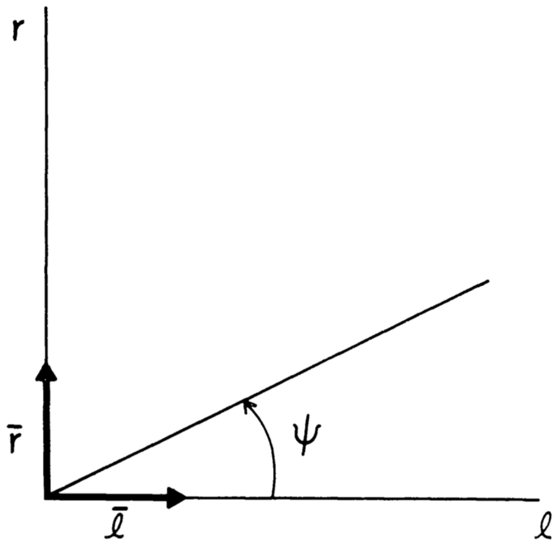


Figure 2.3: Representation of electric field in components of ***r*** and ***l***. Ψ is the angle between the unit direction ***l*** and the direction of vibration of the electric field vector (Hansen and Travis 1974*a*).

parameters, we will look at the definition of the electric field vectors. The electric field is described by unit vectors in the mutual perpendicular r- (E_r) and l-direction (E_l), shown in *Figure* 2.3. In single scattering events, the *l* unit vector is always oriented in the plane of scattering, i.e. the plane that contains the incident and scattered light beam. The electric field is described according to (Hansen and Travis 1974*a*)

$$\begin{aligned} E_l &= a_l e^{i(\omega t - kz - \epsilon_l)} \\ E_r &= a_r e^{i(\omega t - kz - \epsilon_r)} \end{aligned} \tag{2.3}$$

with ω the circular frequency of the light, *z* the distance oriented in the direction of propagation, *k* the wavenumber: $k = 2\pi/\lambda$, $i = \sqrt{-1}$, a_l and a_r the amplitude, ϵ_l and ϵ_r the phases retardation and *t* representing time. The Stokes vector is expressed in these terms of electric field components by (denoting the complex conjugate with an asterisk) (Hansen and Travis 1974*a*)

$$\mathbf{I} = \begin{bmatrix} < E_l E_l^* + E_r E_r^* > \\ < E_l E_l^* - E_r E_r^* > \\ < E_l E_r^* + E_r E_l^* > \\ i < E_l E_r^* - E_r E_l^* > \end{bmatrix} \tag{2.4}$$

Referring to *Figure* 2.3 the radiance, $I(\psi, \epsilon)$, is defined by ψ, the angle between the scattering plane (l-direction) and the direction of vibration, and ϵ a phase retardation. This phase retardation is thus zero for linearly polarized light and a quarter wave for circular polarized light. The Stokes vector can be measured according to (Hansen and Travis 1974*a*)

$$\mathbf{I} = \begin{bmatrix} I(0°,0) + I(90°,0) \\ I(0°,0) - I(90°,0) \\ I(45°,0) - I(135°,0) \\ I(45°,\pi/2) + I(135°,\pi/2) \end{bmatrix} \tag{2.5}$$

To further clarify the difference between the linearly polarized radiances *Q* and *U*, *Figures* 2.4 and 2.5 show the mutual different directions with respect to a horizontal reference frame. For *Q* we can see that in the directions of vibration, $0°$ and $90°$, a plane of symmetry is present, which we can also see for *U* at the inclined planes by $45°$ and $135°$. If one now considers a homogeneous planetary disk and a horizontal reference plane, when integrated over the disk *U* is zero. The same is true for *Q* if we rotate the reference plane by $45°$ in either direction. Because Earth can certainly not be considered as homogeneous, neither *U* and *Q* can be assumed to be zero. An effective way to transform between *Q*

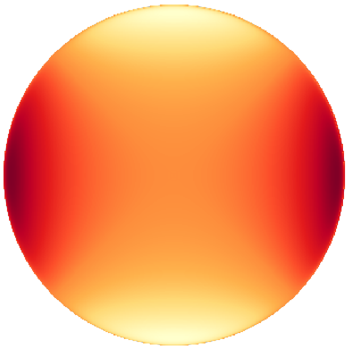


Figure 2.4: Illustration of linearly polarized flux Q on a spherical disk aligned to a horizontal scattering plane.

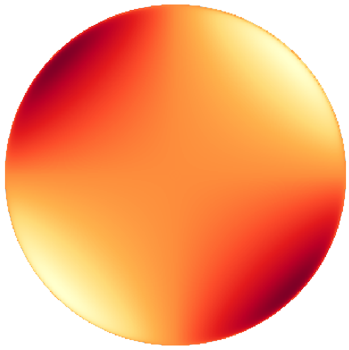


Figure 2.5: Illustration of linearly polarized flux U on a spherical disk aligned to a horizontal scattering plane.

and *U* is to use rotation matrix **L** (Eq. 2 of Stam et al. 2006)

$$\mathbf{L}(\beta) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos 2\beta & \sin 2\beta & 0 \\ 0 & -\sin 2\beta & \cos 2\beta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{2.6}$$

where β is the angle between the old and new reference frame defined in an anticlockwise direction, looking towards the observer. In the characterization of terrestrial exoplanets it is advantages to measure the degree of polarization (Hough et al. 2003; Saar and Seager 2003; Schmid et al. 2006; Seager and Sasselov 2000; Stam et al. 2004*a*)

$$P = \frac{\sqrt{Q^2 + U^2 + V^2}}{I} \tag{2.7}$$

The circular polarized part of light reflected by terrestrial planets is very small and can thus be neglected (Hansen and Travis 1974*a*)

$$Pl = \frac{\sqrt{Q^2 + U^2}}{I} \tag{2.8}$$

Ignoring the circular polarized light in the computation of Stokes elements *F*, *Q* and *U* does not induce significant errors according to Stam and Hovenier (2005). The situation we presented and which we will be using in the forthcoming simulations are upper limits to real future exoplanet observations. In reality, background stellar flux is present in spatially unresolved exoplanets and also to a lower degree in spatially resolved exoplanets. This additional unpolarized flux in spatially unresolved exoplanets causes the degree of linear polarization to be significantly lower.

2.1.1. Doubling-Adding Method

To describe the *Doubling-adding method*, first some definitions and parameters are provided. The incident and reflected light are expressed with respect to two types of reference planes:

1. Firstly, the planetary plane of scattering which coincides with the center of the planet, the observer and the parent star;
2. Secondly, the local plane of scattering that is defined by the local meridian plane and the local zenith direction.

For the *Doubling-adding method* the second reference plane will be used. In a subsequent section we will use the first reference plane to calculate the disk-integrated Stokes vector. In *Figure* 2.6 the local plane of scattering is illustrated. The local incident light is defined by: $\mu_0 = \cos\theta_0$, where θ_0 is the angle between the local zenith direction and the direction of incident light, where $\pi\mathbf{F}_0$ is the incident flux vector. Furthermore, for the internal radiation field the following is defined: $u = \cos\theta_{internal}$ or

$\cos^{-1} u = \theta_{internal}$ (illustrated in the right panel of *Figure* 2.6). The reflected radiance, $\mathbf{I}_r$ is reflected from the medium under an angle relative to the zenith direction: $\mu = |\cos\theta|$. Lastly, the azimuth angle or azimuthal difference (left panel), $\phi - \phi_0$, is used and defined counterclockwise looking downward to the local surface. The absolute azimuth angles of incident and outgoing light rays are not important as a locally rotational symmetric planet is assumed. The light that is reflected ($\mathbf{I}_r$) and transmitted ($\mathbf{I}_t$)

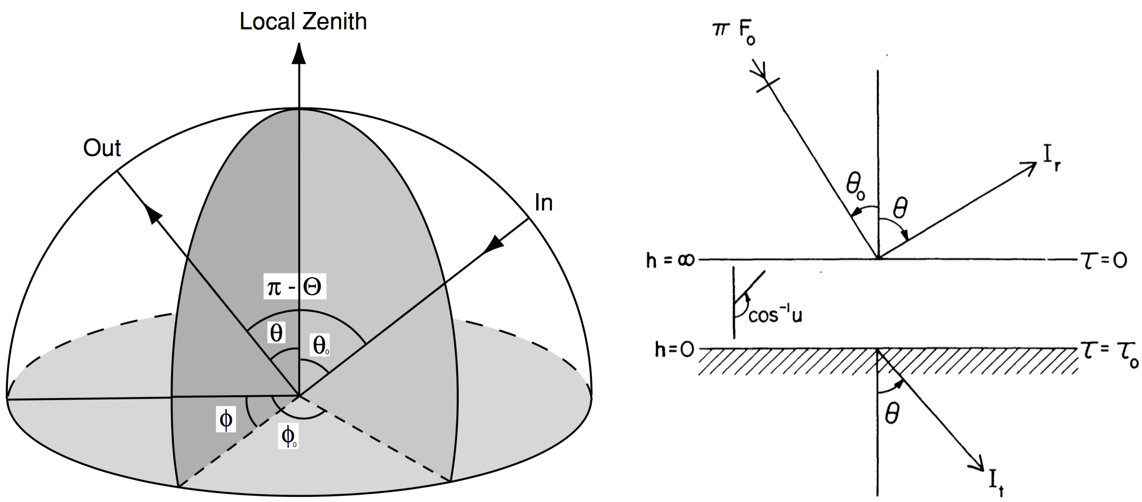


Figure 2.6: Left panel: Spherical representation of locally reflected flux with the azimuth angles ϕ and ϕ_0, zenith angles θ and θ_0, the scattering angle Θ, and the local zenith direction. Right panel: definition of light transmission and reflection on a single layer, with incidence irradiance of $\pi\mathbf{F}_0$, the same incidence angles as the left panel and an optical thickness of τ_0. (Hansen and Travis 1974*a*; Liou 1980).

by the atmosphere are described by the reflection ($\mathbf{R}$) and transmission ($\mathbf{T}$) $4x4$ matrices, respectively, as follows (Hansen and Travis 1974*a*)

$$\begin{aligned}
\mathbf{I}_r(\mu,\phi) &= \frac{1}{\pi}\int_0^1 \mu_0 d\mu_0 \int_0^{2\pi} d\phi_0 \mathbf{R}(\mu,\mu_0,\phi-\phi_0)\mathbf{I}_0(\mu_0,\phi_0), \\
\mathbf{I}_t(\mu,\phi) &= \frac{1}{\pi}\int_0^1 \mu_0 d\mu_0 \int_0^{2\pi} d\phi_0 \mathbf{T}(\mu,\mu_0,\phi-\phi_0)\mathbf{I}_0(\mu_0,\phi_0)
\end{aligned} \tag{2.9}$$

where $\mathbf{I}_0$ is the incident Stokes vector. The flux vector is related to the radiance by

$$\pi\mathbf{F} = 2\pi\int_{-1}^1 \mathbf{I}^0(u)u du \tag{2.10}$$

where $\mathbf{I}^0$ is the azimuth-independent term in the Fourier expansion of $\mathbf{I}(u,\phi-\phi_0)$. To approximate the incident Stokes vector as monodirectional, a dirac delta function (δ) is used accordingly

$$\mathbf{I}_0 = \delta(\mu-\mu_0)\delta(\phi-\phi_0)\pi\mathbf{F}_0 \tag{2.11}$$

The reflected and transmitted light are then defined by

$$\begin{aligned}
\mathbf{I}_r(\mu,\phi) &= \mu_0\mathbf{R}(\mu,\mu_0,\phi-\phi_0)\mathbf{F}_0, \\
\mathbf{I}_t(\mu,\phi) &= \mu_0\mathbf{T}(\mu,\mu_0,\phi-\phi_0)\mathbf{F}_0
\end{aligned} \tag{2.12}$$

Before we continue with the adding equations, some symmetry relations will be defined. Other symmetry relations that will not be used in this report can be found in Hansen and Travis (1974*a*). Light that is transmitted and reflected from below, resulting from reflection on lower layers or the surface, are labeled with a superscript asterisk (*). The equations for reflected light, that is traveling downwards, and the transmitted light, travelling upwards, for illumination from below are described by:

$$\begin{aligned}
\mathbf{I}_r^*(\mu,\phi) &= \mu_0\mathbf{R}^*(\mu,\mu_0,\phi-\phi_0)\mathbf{F}_0, \\
\mathbf{I}_t^*(\mu,\phi) &= \mu_0\mathbf{T}^*(\mu,\mu_0,\phi-\phi_0)\mathbf{F}_0
\end{aligned} \tag{2.13}$$

with the symmetric relations

$$\begin{aligned} \mathbf{R}^*(\mu, \mu_0, \phi - \phi_0) &= \mathbf{R}(\mu, \mu_0, \phi_0 - \phi), \\ \mathbf{T}^*(\mu, \mu_0, \phi - \phi_0) &= \mathbf{T}(\mu, \mu_0, \phi_0 - \phi) \end{aligned} \tag{2.14}$$

Adding method

In order to compute the reflected and transmitted light from a horizontally plane parallel system that is composed of different media, the reflection and transmission matrices for an arbitrary number of layers need to be defined. To explain the fundamentals of this method we use a representation that consists of two layers, provided with the appropriate reflection, transmission and layer characteristics in *Figure* 2.7. The transmission and reflection of the first layer are represented by: $\tilde{T}_1$ (diffusive and direct) and

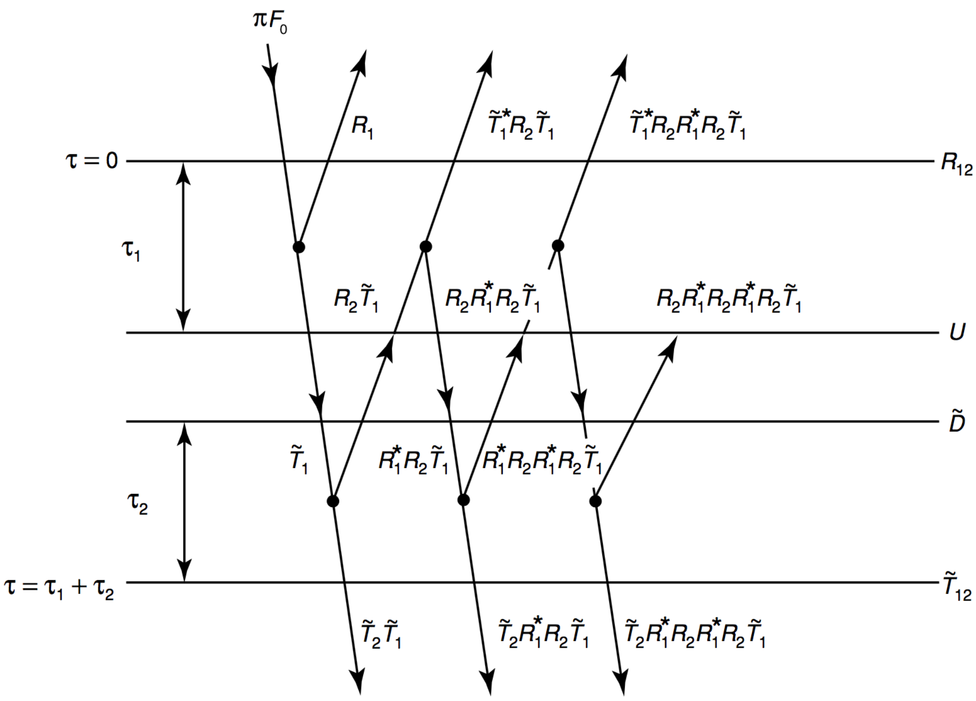


Figure 2.7: Schematic representation of adding method with two layers with different optical thicknesses stacked on each other. Incident irradiance of $\pi\mathbf{F}_0$, optical thicknesses of τ_1 and τ_2. (Hansen and Travis 1974*a*).

R_1, respectively; and for the second layer: $\tilde{T}_2$ (diffusive and direct) and R_2, respectively. The combined total reflection and transmission (diffusive and direct) between the two layers is represented by U and $\tilde{D}$, respectively. The optical thicknesses of the first and second layer are τ_1 and τ_2, respectively. Reflection and transmittance of a specific layer depend on its single scattering characteristics (R and T) and the optical thickness. The total reflection and transmittance of an arbitrary number of layers is mathematically described in de Haan et al. (1987*a*) and Hovenier et al. (2004). The doubling method is mathematically the same as the adding method but uses: $\tau_1 = \tau_2$, essentially doubling the layers. Local multiple scattering can now thus be numerically computed by only requiring the transmission and reflection matrices for single scattering of all layers that are considered in the local atmosphere. Furthermore, in the case of unpolarized incident sunlight, the adding equations can be simplified by using only the first column of the transmission and reflection matrices (see e.g. Rossi et al. 2018)

$$\begin{aligned} \mathbf{I}_r(\mu, \phi) &= \mu_0 \mathbf{R}_{4x1}(\mu, \mu_0, \phi - \phi_0)\mathbf{F}_0, \\ \mathbf{I}_t(\mu, \phi) &= \mu_0 \mathbf{T}_{4x1}(\mu, \mu_0, \phi - \phi_0)\mathbf{F}_0 \end{aligned} \tag{2.15}$$

In the subsequent two sections a detailed description of atmospheric scattering is provided, where the single scattering relations of Rayleigh and Mie scattering are defined.

Rayleigh scattering

The most pronounced type of scattering in Earth's atmosphere is Rayleigh scattering, which was first described by Lord Rayleigh (1871). This type of scattering causes the clear sky to color blue for most part of the day and to color reddish at sunrise and sunset. Two conditions are associated with Rayleigh scattering. First, the wavelength of the incident light ray must be much larger than the size of a particle. Second, the wavelength must be much larger than the particle size after penetration of the incident ray. These conditions can be described as follows: the particle can be considered to be in an external homogeneous electric field, and as the incident ray penetrates the particle, the particles electric field arises instantly compared to the period of the light ray. The electric field that is produced by the incident radiation may be called the *applied field*, producing a dipole configuration on the particle. Following the *electrostatic formula*, the relation between the combined electric field (applied field plus particles electric field) and the induced dipole moment $\mathbf{P}^i$ becomes

$$\mathbf{P}^i = \alpha_p \mathbf{E}^i \tag{2.16}$$

where α_p denotes the polarizability of the particle. The polarizability is in general a tensor, caused by the misalignment of $\mathbf{E}^i$ and $\mathbf{P}^i$. When these vectors align α_p is a scalar and the particle has isotropic polarizability. We are most interested in the properties of the scattered electric field at a distance R away, namely the far-field ($R \gg \lambda$). Furthermore, the angle between the scattered dipole moment and the direction towards the observer is defined as γ, $\mathbf{P}^s$ is the scattered dipole moment and c the speed of light. A representation of these parameters is presented in *Figure* 2.8. The scattered electric field is described by

$$\mathbf{E}^s = \frac{\sin\gamma}{c^2 R}\frac{\partial^2 \mathbf{P}^s}{\partial t^2} \tag{2.17}$$

where the scattered dipole moment can be written as function of the induced dipole moment for an oscillating periodic field as

$$\mathbf{P}^s = \mathbf{P}^i e^{-ik(R-ct)} \tag{2.18}$$

where $ck = \omega$ represents the circular frequency, with k the wavenumber. Substituting *Equations* 2.16 and 2.18 in *Equation* 2.17, results in

$$\mathbf{E}^s = -\mathbf{E}^i e^{-ik(R-ct)}\frac{k^2\alpha\sin\gamma}{R} \tag{2.19}$$

To define the electric vector of a light ray, we employed the orthogonal representation (*section* 2.1, *Figure* 2.3). In this representation the scattered electric field is thus separated in the r and l direction. Based on *Figure* 2.8 and *Figure* 2.3 the following relations can be obtained

$$E_r^s = -E_r^i e^{-ik(R-ct)}\frac{k^2\alpha\sin\gamma_1}{R} \tag{2.20}$$

$$E_l^s = -E_l^i e^{-ik(R-ct)}\frac{k^2\alpha\sin\gamma_2}{R} \tag{2.21}$$

In *Figure* 2.8, Θ denotes the scattering angle defined in the scattering plane as being the plane on which the incident and scattered light wave travel. Furthermore, the l unit direction is defined in the same plane and constrains γ_1 to $\frac{\pi}{2}$. Subsequently, a system of scattering electric field components can be obtained

$$\begin{bmatrix} E_r^s \\ E_l^s \end{bmatrix} = -e^{-ik(R-ct)}\frac{k^2\alpha}{R}\begin{bmatrix} 1 & 0 \\ 0 & \cos\Theta \end{bmatrix}\begin{bmatrix} E_r^i \\ E_l^i \end{bmatrix} \tag{2.22}$$

The next step is to define the flux vector as function of the electric field: $F^s = C|E^s|^2$, with C a proportionality factor (Liou 1980), to obtain the polarized intensity components

$$F_r^s = F_r^i \frac{k^4\alpha^2}{R^2} \tag{2.23}$$

$$F_l^s = F_l^i \frac{k^4\alpha^2\cos\Theta}{R^2} \tag{2.24}$$

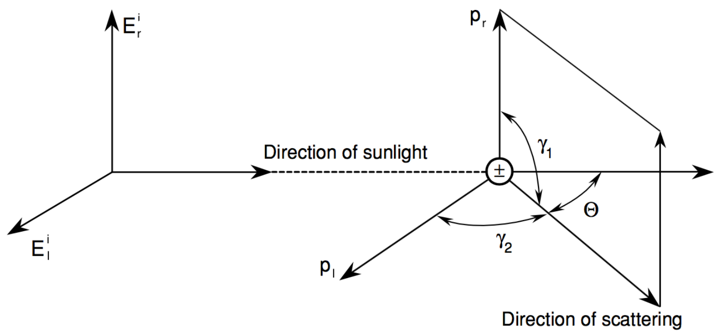


Figure 2.8: Schematic representation of Rayleigh scattering by a spherical particle, with electric field components $\mathbf{E}_r^i$ and $\mathbf{E}_l^i$, dipole moments $\mathbf{P}_r$ and $\mathbf{P}_l$, scattering angle θ, and dipole moment angles γ_r and γ_l. (Liou 1980)

The total scattered intensity is then simply: $F^s = F_l^s + F_r^s$. For unpolarized incident light the following is true

$$F^s = F^i \frac{(2\pi)^4 \alpha^2}{R^2 \lambda^4} \frac{(1 + \cos^2 \Theta)}{2} \tag{2.25}$$

The phase matrix for a full representation of light scattering follows accordingly (Hansen and Travis 1974*a*):

$$\mathbf{F}^m(\Theta) = \begin{bmatrix} \frac{3}{4}(1+\cos^2\Theta) & -\frac{3}{4}\sin^2\Theta & 0 & 0 \\ -\frac{3}{4}\sin^2\Theta & \frac{3}{4}(1+\cos^2\Theta) & 0 & 0 \\ 0 & 0 & \frac{3}{2}\cos\Theta & 0 \\ 0 & 0 & 0 & \frac{3}{2}\cos\Theta \end{bmatrix} \tag{2.26}$$

where the superscript *m* indicates molecular scattering. This phase matrix is representative for isotropic particles. In reality, however, atmospheric molecules show some degree of anisotropy, which is accounted for with a depolarization factor, δ, in the following form (Hansen and Travis 1974*a*):

$$\mathbf{F}^m(\Theta) = \frac{1-\delta}{1+\delta/2} \begin{bmatrix} \frac{3}{4}(1+\cos^2\Theta) & -\frac{3}{4}\sin^2\Theta & 0 & 0 \\ -\frac{3}{4}\sin^2\Theta & \frac{3}{4}(1+\cos^2\Theta) & 0 & 0 \\ 0 & 0 & \frac{3}{2}\cos\Theta & 0 \\ 0 & 0 & 0 & \frac{1-2\delta}{1-\delta}\frac{3}{2}\cos\Theta \end{bmatrix} + \frac{3\delta/2}{1+\delta/2} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{2.27}$$

A table with depolarization factors for various atmospheric molecules can be obtained from Hansen (1971). The scattering matrix elements can be expanded in generalized spherical functions, for which the expansion coefficients are provided in e.g. Stam et al. (2002). This approximation is used to save large amounts of computer storage and computing time. We now have the single scattering matrix of our molecular atmosphere. In the next paragraph we will define a single scattering matrix for the aerosols in our atmosphere. The conversion of this scattering matrix to a reflection matrix or transmission matrix is provided in Hovenier et al. (2004).

Mie scattering

Mie scattering is used to represent the single scattering behaviour of homogeneous isotropic spheres. Just as with Rayleigh scattering (*Paragraph* 2.1.1), the solution for Mie scattering in the far-field is required for exoplanetary reflection. Consider that an isotropic homogeneous spherical particle is illuminated by a light ray traveling in the positive z-axis (*Figure* 2.9). The scattered waves defined by Ru and Rv, where R denotes the distance to the far-field, are derived from reduced *Hankel functions* (described in more detail in Liou (1980)):

$$
\begin{aligned}
Ru^s &= -\frac{ie^{-ikR}\cos\phi}{k}\sum_{n=1}^{\infty}\frac{2n+1}{n(n+1)}a_n P_n^1(\cos\theta),\\
Rv^s &= -\frac{ie^{-ikR}\sin\phi}{k}\sum_{n=1}^{\infty}\frac{2n+1}{n(n+1)}b_n P_n^1(\cos\theta)
\end{aligned}
\tag{2.28}
$$

where k is the wavenumber, a_n and b_n are Mie coefficients and P_n^1 a Legendre polynomial. These waves are expressed in spherical coordinates, where the total system is provided in *Figure* 2.9. The components of the electric field vectors in spherical coordinates become (Liou 1980):

$$
\begin{aligned}
E_r^s &= 0,\\
E_\theta^s &= -\frac{ie^{-ikR}\cos\phi}{Rk}\sum_{n=1}^{\infty}\frac{2n+1}{n(n+1)}\left[a_n\frac{dP_n^1(\cos\theta)}{d\theta}+b_n\frac{P_n^1(\cos\theta)}{\sin\theta}\right],\\
E_\phi^s &= \frac{ie^{-ikR}\sin\phi}{Rk}\sum_{n=1}^{\infty}\frac{2n+1}{n(n+1)}\left[a_n\frac{P_n^1(\cos\theta)}{\sin\theta}+b_n\frac{dP_n^1(\cos\theta)}{d\theta}\right]
\end{aligned}
\tag{2.29}
$$

From *Equation* 2.29 two scattering functions can be defined

$$
\begin{aligned}
S_1(\Theta) &= \sum_{n=1}^{\infty}\frac{2n+1}{n(n+1)}[a_n\pi_n(\cos\Theta)+b_n\tau_n(\cos\Theta)],\\
S_2(\Theta) &= \sum_{n=1}^{\infty}\frac{2n+1}{n(n+1)}[b_n\pi_n(\cos\Theta)+a_n\tau_n(\cos\Theta)]
\end{aligned}
\tag{2.30}
$$

with the coefficients $\pi_n(\cos\Theta)$ and $\tau_n(\cos\Theta)$ being defined as:

$$
\begin{aligned}
\pi_n(\cos\Theta) &= \frac{P_n^1(\cos\Theta)}{\sin\Theta},\\
\tau_n(\cos\Theta) &= \frac{dP_n^1(\cos\Theta)}{d\Theta}
\end{aligned}
\tag{2.31}
$$

For consistency the scattered spherical waves are defined in the r- and l- unit direction (*Section* 2.1) (Liou 1980):

$$
\begin{aligned}
E_r^s &= -E_\phi^s,\\
E_l^s &= E_\theta^s
\end{aligned}
\tag{2.32}
$$

and from Liou (1980) the incident electric vectors are

$$
\begin{aligned}
E_r^i &= e^{-ikz}\sin\phi,\\
E_l^i &= e^{-ikz}\cos\phi
\end{aligned}
\tag{2.33}
$$

Using *Equations* 2.29, 2.30, 2.32 and 2.33 the fundamental system of equations for a light beam scattered by homogeneous spheres can be obtained (de Rooij and van der Stap 1984; Hansen and Travis 1974*a*; Liou 1980)

$$
\begin{bmatrix} E_l^s \\ E_r^s \end{bmatrix} = \frac{e^{-ikR+ikz}}{ikR}\begin{bmatrix} S_2(\Theta) & 0 \\ 0 & S_1(\Theta)\end{bmatrix}\begin{bmatrix} E_l^i \\ E_r^i \end{bmatrix}
\tag{2.34}
$$

More detailed derivations of the functions a_n and b_n are provided in de Rooij and van der Stap (1984). The next step is now to define the total single scattering matrix or phase matrix, which is proportional to the transformation matrix $\mathbf{T}(\Theta)$:

$$
\mathbf{F}^a(\Theta) = \frac{4\pi}{k^2\sigma_{sca}}\mathbf{T}(\Theta)
\tag{2.35}
$$

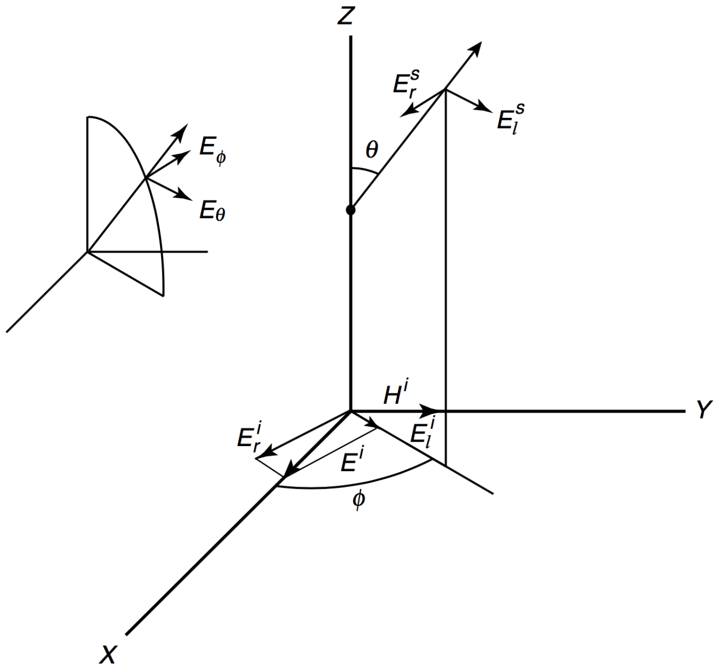


Figure 2.9: Incident and scattered electric vectors in Cartesian (analogous to r- and l-direction) and spherical coordinates (Liou 1980). Electric field components are thus provided by $\mathbf{E}_r^i$ and $\mathbf{E}_l^i$ (cartesian) and $\mathbf{E}_\theta$ and $\mathbf{E}_\phi$ (spherical). The direction of propagation is the positive Z-axis and a scattering angle of Θ. The scattered electric field components are denoted with the superscript *s*.

where the superscript *a* is a shortname for aerosols. Following van de Hulst (1957), the transformation matrix, $\mathbf{T}(\Theta)$, is defined as:

$$\mathbf{T}(\Theta) = \begin{bmatrix} \frac{1}{2}(S_1S_1^* + S_2S_2^*) & \frac{1}{2}(S_1S_1^* - S_2S_2^*) & 0 & 0 \\ \frac{1}{2}(S_1S_1^* - S_2S_2^*) & \frac{1}{2}(S_1S_1^* + S_2S_2^*) & 0 & 0 \\ 0 & 0 & \frac{1}{2}(S_1S_2^* + S_2S_1^*) & \frac{i}{2}(S_1S_2^* - S_2S_1^*) \\ 0 & 0 & -\frac{i}{2}(S_1S_2^* - S_2S_1^*) & \frac{1}{2}(S_1S_1^* + S_2S_2^*) \end{bmatrix} \tag{2.36}$$

where the asterisk denotes the complex conjugate. Similar to the Rayleigh scattering single scattering matrix, the Mie single scattering matrix is expanded into expansion coefficients according to the expansion provided in de Rooij and van der Stap (1984); Domke (1974). The conversion of this scattering matrix to a reflection matrix or transmission matrix is provided in Hovenier et al. (2004).

2.2. PyMieDAP code

Since the basic theory of radiative transfer was discussed in the former section, we can now have a closer look at the PyMieDAP code that we will use to model the Earth. Not only the atmosphere of our exoplanet will influence the reflection behaviour towards the observer, but also the surface, both of which implementations in PyMieDAP are covered. The radiative transfer of PyMieDAP is applied to an arbitrary amount of pixels. To understand how the reflection is computed for a grid of pixels the basic formulas for disk integrated and disk-resolved reflected spectropolarimetric signals are provided.

2.2.1. PyMieDAP's basic structure

The *PyMieDAP* radiative transfer code (Rossi et al. 2018) is used to model Earth as an exoplanet. This code allows the user to compute the Stokes vector reflected from exoplanets. The program consists of several modules/subroutines written either in *Python* or *Fortran*, with the *Fortran* subroutines being interfaced with *Python*. In *Figure* 2.10 an overview of the *PyMieDAP* code is provided. The core radiative transfer computations are housed by the *Fortran* subroutines, mainly covering the doubling-adding algorithm (see de Haan et al. 1987*a*), the Mie scattering and Rayleigh scattering and some geometrical conversion algorithms. The input to and output from these routines is housed in the Python environment.

The variables that an user can specify to model a certain locally plane horizontally homogeneous, but vertically inhomogeneous atmosphere are:

- a list containing the wavelengths of the incident light;
- gravitational acceleration of the planet;
- a reflection matrix or Fourier coefficients file to describe the planetary surface;
- the depolarization factor of the atmospheric gas;
- the molecular mass of the atmospheric gas;
- a confirmation to automatically compute the gaseous refractive index. Options are restricted to "air", CO_2, H_2, He and N_2. The relation for "air" is retrieved from Ciddor (1996), where "air" is defined as dry air at $15°C$, 1013.25 Pa, and with 443 ppm CO_2 content.
- temperature of the star;
- distance between the planet and star;
- the radius of the star;
- a sub class to define the planetary atmosphere with multiple layers. The input parameters for a single layer object are the following:
  - the bottom pressure of the layer;
  - the aerosol optical thickness of the layer;
  - the column density of the layer;
  - an option to compute the optical thickness for molecular (Rayleigh) scattering;
  - an user-defined optical thickness for molecular (Rayleigh) scattering;
  - optical thickness of gaseous absorption.
  - a subclass to define a type of aerosols in a layer. This aerosol object is described by the following input parameters, with additional input parameters for layered spherical particles:
    - the effective radius and variance of the aerosol;
    - the real and imaginary refractive index;
    - the type of distribution of particle sizes;
    - an user preferred label of the aerosol;
    - if layered spheres are present:
      - the real and imaginary refractive index of the inner core, the outer core is specified with the previous index;
      - the ratio between the radii of the outer and inner core.

Based on the input parameters for the aerosol properties and gaseous properties, Mie and Rayleigh scattering matrices are computed and combined for each distinct layer. Subsequently, the *doubling-adding method* computes the reflection matrix for the entire atmosphere, including the surface, and expands it into *Fourier coefficients*. These coefficients describe the Top Of Atmosphere (TOA) scattering behaviour of the entire system. The Stokes vector of the reflected light can then be computed at the desired geometry(ies) and wavelength(s). A more detailed description of the surface and atmospheric models are provided in the following paragraphs.

The arrangement of the pixels in a planetary disk can exhibit different structures, ranging from a full homogeneous disk based on one pixel model type to a patchy cloud coverage that is based on two different pixel model types: a model simulating clear sky conditions and one simulating cloudy conditions. More specifically, PyMieDAP calculates a gridded disk and allows one to mask regions of cloud cover on that disk that correspond to (an) a priori defined pixel model(s). To obtain disk-resolved or integrated spectropolarimetric signals from these grids a more detailed description is provided in the last paragraph of this section.

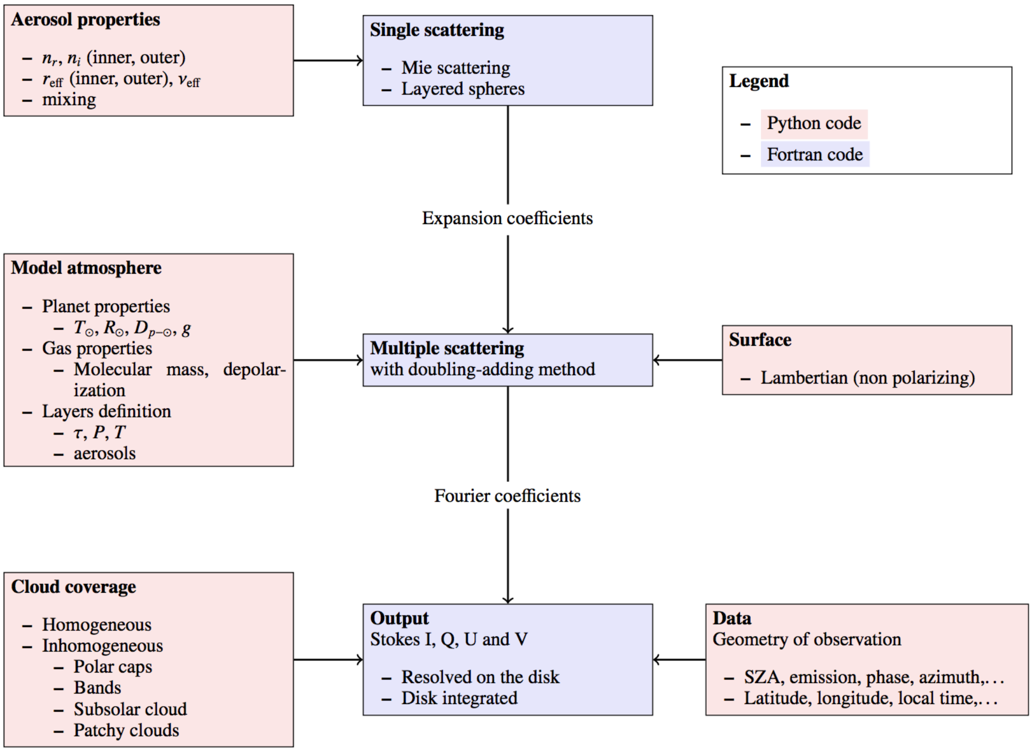


Figure 2.10: Schematic representation of PyMieDAP doubling-adding algorithm (Rossi et al. 2018).

Surface model

The surface model that is used in *PyMieDAP* is a Lambertian depolarizing reflecting surface, which is described by a reflection matrix. For a Lambertian surface, its albedo is specified in the $(1, 1)$ element of the reflection matrix, where other entries describe how incident light undergo changes in their polarized state, hence being zero (Rossi et al. 2018). Furthermore, due to its isotropic reflection, the matrix is geometrically independent. The surface layer is described by the lowest layer in the doubling adding method.

More detailed surface models can be implemented in PyMieDAP by transforming a geometric dependent surface model to "Fourier files" and providing it as a read file for the algorithm. These "Fourier files" need to be restricted to the same format as that generated by PyMieDAP itself. In short, Fourier files are data files that contain expansion coefficients, which describe the total reflection from the locally plane horizontally homogeneous model atmosphere for any desired geometry[1].

Atmospheric model

In *Section* 2.1 a detailed description of atmospheric scattering is provided. In this theoretical description Rayleigh and Mie scattering matrices are defined which can be described by expansion coefficients (de Rooij and van der Stap 1984; Domke 1974). In order to model a layer with multiple aerosols in combination with gaseous scattering, phase matrices for Rayleigh and Mie scattering thus have to be combined. Subsequently, a mixed layer is modeled by *PyMieDAP* as follows (see e.g. Stam 2008*a*)

$$\mathbf{F}(\Theta) = \frac{b^m_{sca}\mathbf{F}^m(\Theta) + b^a_{sca}\mathbf{F}^a(\Theta)}{b^m_{sca} + b^a_{sca}} \tag{2.37}$$

where $\mathbf{F}^m(\Theta)$ and $\mathbf{F}^a(\Theta)$ are the scattering matrices of molecular (Rayleigh scattering) and aerosol (Mie scattering) scattering; and b^m_{sca} and b^a_{sca} are the molecular (gaseous) and aerosol scattering optical thicknesses, respectively. For each layer there is no limitation for the number of aerosols that can be

[1]For more information about these coefficients one can consult *Section* 3 and 4 in de Haan et al. (1987*a*).

modeled. For each mixture of aerosols, the scattering matrix is computed and summed accordingly

$$\mathbf{F}(\Theta) = \frac{\sum_{i=0}^{n} b_{sca}^{a,i} \mathbf{F}^{a}(\Theta)}{\sum_{i=0}^{n} b_{sca}^{a,i}} = \frac{\sum_{i=0}^{n} b_{sca}^{a,i} \mathbf{F}^{a}(\Theta)}{b_{sca}^{a}} = \sum_{i=0}^{n} f_i \mathbf{F}_i^{a}(\Theta) \tag{2.38}$$

where *i* is the distinct type of aerosol, *n* is the number of aerosol types, b_{sca}^{a} the total aerosol scattering optical thickness of the layer, $b_{sca}^{a,i}$ the scattering optical thickness of aerosol type *i*, and $\mathbf{F}_i^{a}(\Theta)$ the scattering matrix of aerosol type *i*.

Integrating over the planetary disk

To compute the integrated signal we refer again back to the definitions of the two reference planes from *Section* 2.1:

1. Firstly, the planetary plane of scattering coincides with the center of the planet, the observer and the parent star;
2. Secondly, the local plane of scattering which is defined by the local meridians and the local zenith.

In that section, the second reference frame was used to define locally reflected light by the atmosphere and the surface. The first reference plane will be used to describe the disk-integrated and disk-resolved (*Section* 2.2.1) reflected light from a planet. The analytic formulation of the total reflected light over the illuminated and visible part of the planetary disk is given by (Eq. 17 of Stam et al. 2006)

$$\pi\mathbf{F}(\alpha, \lambda) = \frac{1}{d^2} \int_{\square} \mu\mu_0 \mathbf{R}_1'(\mu, \mu, \phi - \phi_0, \lambda) F_0(\lambda) dO \tag{2.39}$$

where *d* is the distance between the observer and planet, F_0 is the unpolarized incident flux vector, $\pi\mathbf{F}$ the flux vector of reflected starlight and $\mu dO/d^2$ is the solid angle from which the area dO is seen by the observer. For horizontally inhomogeneous planets, the reflection matrix is depending on the orientation between the planet and the observer. For all locally reflecting matrices a rotation is required, such that the reference plane is no longer the local meridian but the planetary scattering plane, where it is thus dependent on the location of surface area dO and the local viewing angle θ. The rotated reflection vector is obtained by (see Rossi et al. 2018)

$$\mathbf{R}_1' = \mathbf{L}(\beta)\mathbf{R}_1 \tag{2.40}$$

where $\mathbf{L}$ is the rotation matrix provided by *Equation* 2.6. In this case, β is the angle between the local meridian plane and the planetary scattering plane, being positive when looking towards the observer and rotating in the anti-clockwise sense from the old to the new plane. *PyMieDAP* calculates the reflected flux by replacing the integral in *Equation* 2.39 by a summation over all visible and illuminated pixels (Eq. 12 of Rossi et al. 2018)

$$\pi\mathbf{F}(\alpha, \lambda) = \frac{F_0(\lambda)}{d^2} \sum_{n=1}^{N} \mu_n \mu_{0n} \mathbf{L}(\beta_n) \mathbf{R}_1(\mu_n, \mu_{0n}, \phi_n - \phi_{0n}, \lambda) dO_n \tag{2.41}$$

In theory each pixel can have a different reflection matrix, depending on the local reflection properties, such that horizontal variations of Earth can be modeled. In practice this is computational very intensive though. *PyMieDAP* calculates the disk-integrated light in the following sequence:

1. The planetary disk is divided into a desired number of pixels in order to obtain a horizontally inhomogeneous model;
2. For each pixel μ, μ_0, $\phi - \phi_0$ and β are computed;
3. For each pixel the local reflection matrix $\mathbf{R}_1$ is computed;
4. *Equation* 2.41 is used in combination with previous computed parameters.

For the above process all local parameters (μ, μ_0, $\phi - \phi_0$, β and $\mathbf{R}_1$) are computed in the middle of the pixel. In order to get rid of the dependency on *d* we normalize the reflected flux vector to the geometric albedo:

$$A_G(\lambda) = \frac{\pi F(0^\circ, \lambda)}{\pi F_0}(\lambda)\frac{d^2}{r^2} \tag{2.42}$$

where $\pi F(0^\circ, \lambda)$ is the total reflected flux at phase angle $\alpha = 0^\circ$ and wavelength λ. The normalization on *Equation* 2.41 produces:

$$\mathbf{F}_{norm} = \frac{\mathbf{F}(\alpha, \lambda)}{F_0(\lambda)}\frac{d^2}{r^2} = \sum_{n=1}^{N} \mu_{0n}\mathbf{L}(\beta_n)\mathbf{R}_1(\mu_n, \mu_{0n}, \phi_n - \phi_{0n})\frac{dO_n\mu_n}{\pi r^2} \tag{2.43}$$

where $\frac{dO_n\mu_n}{\pi r^2}$ is equivalent to $\frac{1}{N}$ and $\mathbf{F}_{norm}$ is the normalized reflected flux.

Computing the resolved spectral disk is very similar to that of the disk-integrated process (*Section* 2.2.1). The only difference is that the summation of *Equation* 2.41 is left out and the reflected flux is calculated for each pixel separately:

$$(\mathbf{F}_{norm})_n = \mu_{0n}\mathbf{L}(\beta_n)\mathbf{R}_1(\mu_n, \mu_{0n}, \phi_n - \phi_{0n})\frac{dO_n\mu_n}{\pi r^2} \tag{2.44}$$

2.3. Modeling horizontally inhomogeneous Earth-like planets with PyMieDAP

PyMieDAP in its original form is able to create horizontally homogeneous and inhomogeneous cloud patterns above a homogeneous or inhomogeneous surface in the form of for example latitudinal bands, subsolar clouds, polar cups, and patchy clouds (Rossi and Stam 2017). These models are using only a small sequence of pixel models that are assigned to a "masked" grid that is created by the *Mask_planet* function. Due to the diversity of Earth, as well for its surface as its atmosphere, assigning a large sequence of pixel models a priori to a specific sequence for a specific observational day is very cumbersome. To create an inhomogeneous disk that is based on Earth observations the *Mask_planet* function is extended to do just that. In the upcoming section a description of this extension is provided. The term "pixel models" is used to denote the locally plane parallel horizontally homogeneous atmosphere and surface interface that describes the radiative transfer of a specific pixel. A general overview of the modified PyMieDAP routine is provided in *Appendix* A. Using PyMieDAP to model the Earth-like model requires us to define some core radiative transfer input parameters, which depend largely on the parameters that are extracted from the Earth observations. In *Section* 2.3.2, we provide a description of how we tune the PyMieDAP code to efficiently simulate the reflected Stokes parameters from an Earth-like exoplanetary model, to allow for reasonable computation effort. We also provide preliminary results which are valuable in the construction of a planetary model.

2.3.1. Mask_planet function extension

In this section we will provide a brief explanation of how the *Mask_planet* function uses the Earth observations to create an unique mask. The basic structure is provided in a small flow chart, *Figure* 2.11, and is referred to in the text.

Disk formation with **getgeos** function

The first step is to define the planetary disk in an arbitrary amount of pixels, on which we can apply a mask. The disk coordinates and geometric quantities are computed with the *getgeos*[2] function of PyMieDAP. The planetary disk is approximated by pixels with equal area. A planetary disk with 20×20 pixels is shown in the left panel of *Figure* 2.12.

Pixel coordinate transformation

To extract data for a single pixel we need to mask a specific region of longitudinal and latitudinal coordinates in the observation file(s). This is most efficiently done by converting the boundaries of the

[2]This function is part of the standard PyMieDAP tool (Rossi et al. 2018).

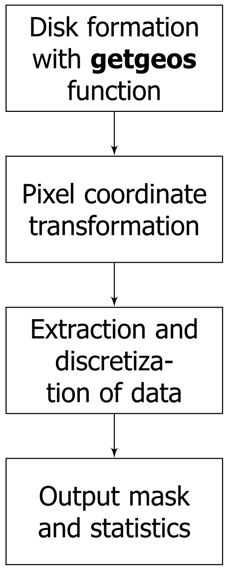


Figure 2.11: Flowchart of the general process with which an Earth-like disk mask is produced, based on Earth observations.

disk pixels to polygons as conversion to another map projection does not produce square shapes. The conversion is only valid for polygon coordinates that lie on the boundary or in the planetary disk. The left panel of *Figure* 2.12 shows that a region of some pixels lie outside the planetary disk, i.e. outside the trigonometric domain of conversion to longitudinal and latitudinal coordinates. On these pixels we apply a shaping method to replace the invalid coordinates by a section of the circle boundary before we convert it to the appropriate map projection. The result of this conversion is provided in the right panel of *Figure* 2.12, showing that some pixels are cut off at the edges. In this process, we took into account the resolution of the observations such that no data points are left out. The polygons are converted from a Vertical Perspective Projection to a Geographic Map Projection (an example of the latter is shown in *Figure* 2.14). The Vertical Perspective Projection approaches, in case of infinite distance between the observer and the object, the Orthographic Projection (shown in *Figure* 2.13). The conversion from a X,Y grid to a Longitude,Latitude (ϕ, λ) grid for a sphere is described by the following set of equations (Snyder 1987):

$$\begin{aligned}
\rho &= \sqrt{x^2 + y^2}, \\
c &= \arcsin \frac{D - \sqrt{1 - \rho^2 (D+1)/(R^2(D-1))}}{R(D-1)/\rho + \rho/(R(D-1))}, \\
\phi &= \arcsin \cos c \sin \phi_0 + (y \sin c \cos \phi_0 / \rho), \\
\lambda &= \lambda_0 + \arctan 2[(x \sin c), (\rho \cos \phi_0 \cos c - y \sin \phi_0 \sin c)]
\end{aligned} \tag{2.45}$$

where P is the distance between observer and object divided by the planet radius R, ϕ_0 the obliquity of the planet and λ_0 the sub-observer point. The radius of the planet is $R = 1$ to correspond with the grid coordinates X and Y. Verification of the coordinate conversion is provided in *Appendix* B.

Extraction and discretization of data

We now have created the pixel polygons. The next step is to use these polygons to extract a portion of the data. We achieve this by using the *Python*-package *Rasterio*. The "mask" function of *Rasterio* allows one to define a non-square geometry and read a two dimensional data set on which to extract the corresponding geometry. In *Figure* 2.14 it is shown how the geometries are applied on the grid of data, in this case Earth's land cover. From this representation one can see that at the edges some polygons seem to be missing. This is due to the square pixel approximation that we use in PyMieDAP, that can be more clearly seen from the disk in the right panel of *Figure* 2.12, where some regions of the planetary disk are not covered by a pixel.

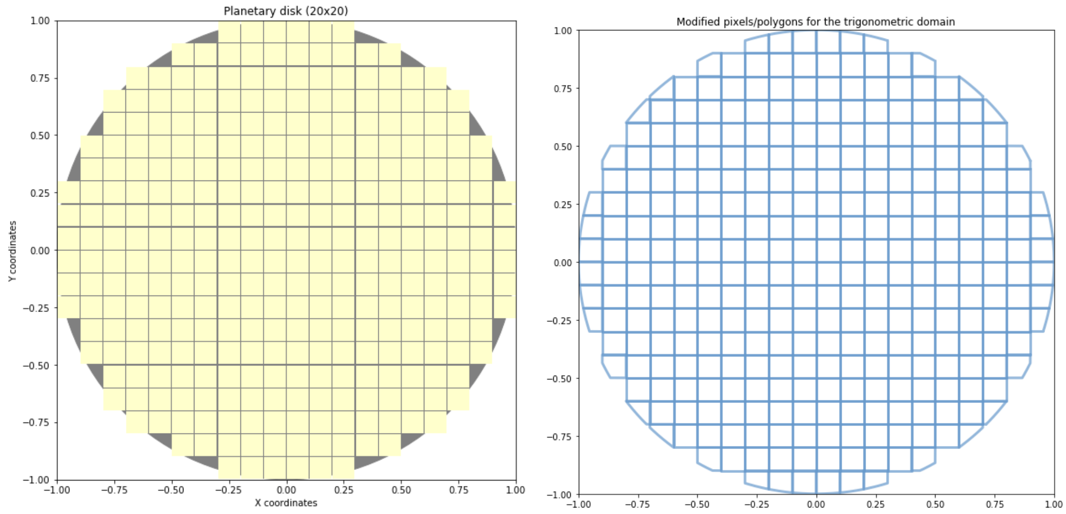


Figure 2.12: Illustration of a 20×20 pixel grid that approximates a planetary disk. In the left panel we see the pixel used to approximate the disk. In the right panel we see the constructed polygons that are still in a X,Y coordinate system.

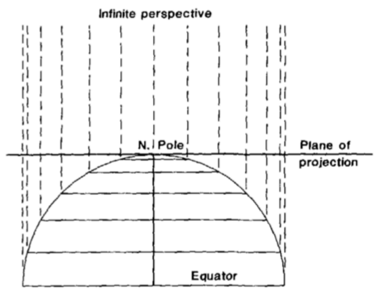


Figure 2.13: Schematic representation of disk coordinate conversion to geographic latitude and longitude coordinates (Snyder 1987).

Output mask and statistics

Now that we have acquired the (discretized) data for each pixel it is converted to an one dimensional bit-string. In *Figure* 2.15 one can see the concept of such a masked disk for a four parameter Earth model. In that case there was chosen to vary the surface type and three cloud parameters, where the first integer denotes the type of land cover, and the other three integers define a certain cloud parameter forming one type of cloud. Additionally, based on these parameters an additional set of statistics are computed to broaden the parameter space. The following set is provided:

1. Mean values of the extracted cloud properties;
2. Land cover fractions of each land cover type;
3. Cloud fraction, i.e. the fraction of pixels that have nonzero optical thickness to the total number of pixels;
4. Asymmetry of land cover and cloud properties, i.e. the fraction of pixels that have no equal land cover or cloud properties mirrored over the equator to the total number of pixels. This statistic thus has asymmetry values for each varying parameter on the disk;
5. Patchiness of each pixel.

2.3.2. PyMieDAP simulation strategy and pixel model results

In order to adequately implement and compute pixel models that can describe the Earth-like surface-atmosphere system, several core parameters in the PyMieDAP code need to be determined. In the first section we determine the required number of Gaussian abscissae based on the accuracy of simulated

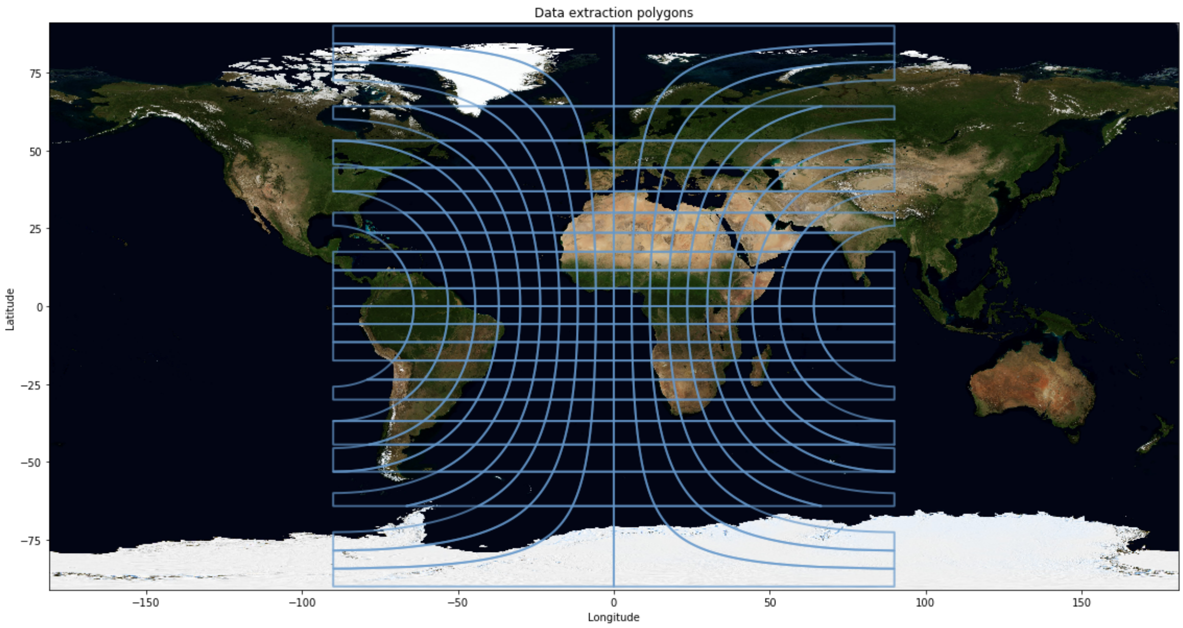


Figure 2.14: A representation of the method to extract data for a specific pixel. The data is provided in a Geographical Map Projection of which an example of Earth's land distribution is provided.

Stokes vectors and required computational effort. In the second section, based on the goal of this research and again computation effort, we determine the number of Stokes elements to be computed, where a lower number of Stokes elements yield a decrease in computation effort. Using these setting we present the reflected Stokes elements from an isolated cloud layer with Earth-like cloud parameters. Lastly, we will provide an estimation of the computation time required to model such cloud layers.

Number of Gaussian abscissae in pixel model computations

Earth-like clouds contain particles that have an effective particle size that is significantly higher compared to the wavelength of interest (Han et al. 1994*a*). In *Figure* 2.16 we provide the relative difference for F and Pl between a phase curve computed with 160 Gaussian abscissae and that with 110, 120, 130, 140 and 150 Gaussian abscissae. These simulations use a 20×20 pixel planetary disk covered by a horizontally homogeneous thick cloud composed of particles with a large effective radius, and located at mid altitudes: $\tau = 20$, $r_{eff} = 17.5$ and $P_c = 700mb$. The results that for 150 Gaussian abscissae we do not see any significant difference in accuracy compared to 160 Gaussian abscissae. For the total flux, we see a maximum relative difference of approximately ~ 0.003 and for the total degree of polarization a maximum relative difference of approximately $\sim 0.25\%$. When computing pixel models with 160 Gaussian abscissae as compared to 150 Gaussian abscissae, PyMieDAP will have to write an additional $\sim 13\%$ in file size and the computation time increases by $\sim 46\%$. For a single pixel model file this might not seem as a drastic increase, but when we want to compute an Earth-like exoplanet with 150 different models at 5 different wavelength, resulting in 750 different models, this increase is significant, without even considering the increase in time of reading the files. Lastly, for clouds with lower effective radii and optical thicknesses the convergence is more quickly and requires less Gaussian abscissae. We will thus compute all pixel models with 150 Gaussian abscissae.

Number of Stokes elements in pixel model computations

The size of the Stokes vector is limited to either three (F, Q, U) or four (F, Q, U, V) elements as we need to model at least linear polarized fluxes to retrieve Earth-like exoplanetary spectropolarimetric signals. From Hansen and Travis (1974*a*) we know that the circular polarization of reflected light from planetary atmospheres is very small. In this research the analysis will be restricted to total flux and linear polarized fluxes, so only requiring the computation of three Stokes elements. We omit the circular polarization, because (1) measurements of circular polarization reflected from terrestrial exoplanetary atmospheres show to be very small (Hansen and Travis 1974*a*). Although, some studies do argue that the circular polarization is a valuable tool to unambiguously retrieve the presence of homochirality,

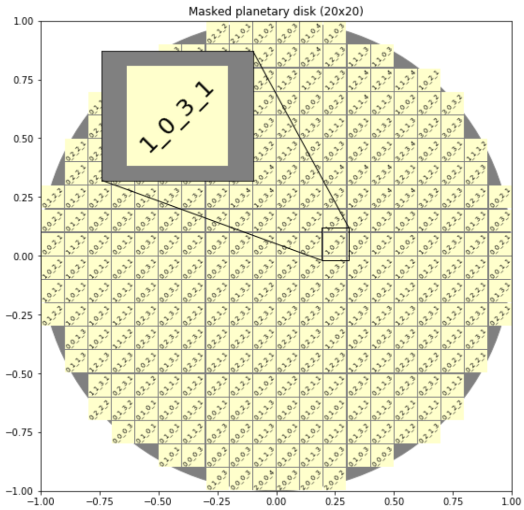


Figure 2.15: Representation of a 20×20 masked disk that is based on a four dimensional varying parameter space.

e.g. vegetation (Sparks et al. 2009), which we will not be able to retrieve because of the depolarizing Lambertian surface and the lack of a polarizing surface model anyway. (2) the computation time of the Fourier files is on average increased by a factor of ~ 3.

Reflection behaviour of cloud layers in the pixel models

We have now defined our core model parameters, which we can use to accurately compute our Earth-like pixel models. In this last section, we will look at some results from the different types of cloud models to identify the general behaviour. We will only model the layer with clouds and thus use a Rayleigh optical thickness of zero and a black surface. The Earth-like cloud layers will be modeled with four cloud particle effective radius values: $R_{eff} = 8, 12, 16, 20\ \mu m$, and 16 values of cloud optical thickness: $\tau = 0, 2, ..., 30$ (see Han et al. 1994*a*; Nakajima and King 1990). In *Figure* 2.17 the Stokes elements F, Q and U are provided for the geometrical angles: $\alpha = 0°, \theta_0 = 0.1° and\ \theta = 0.1°$, which is equivalent to a pixel located slightly above or below the planetary scattering plane. In the left panels the reflection is provided as function of the cloud optical thickness and in the right panel as function of the logarithmic value of optical thickness. With increasing optical thickness we see a logarithmic like increase in flux, which makes sense as eventually a cloud is so thick that no significant amount of light can travel through the layer and all is reflected back to space. Applying now a logarithm on the optical thickness provides a linear like behaviour. This logarithmic behaviour was also observed by Oreopoulos et al. (2007), who also argued that approximating a range of optical thickness values by one ”mean” can best be applied with the logarithmic mean. No tests were made for the polarized fluxes. The influence of the effective radius is very small, but seems to increase the reflected flux decreasing radius. In the lower two panels we provided the reflection behaviour of the two linearly polarized fluxes. We can see that the effect of the strong depolarizing Mie scattering of the clouds has a far more direct effect on the polarizing fluxes than on the total flux. Also, for different values of effective radius we see very different behaviour and an alteration in sign. In the right panels we can see that applying a logarithmic value on τ provides a very linearly like behaviour for both polarized Stokes elements, and thus shows that a logarithm mean can also better be applied for the polarized fluxes. The specific behaviour of the pixel model located at different locations on the planetary disk is provided in *Appendix* C. These different cases show the same behaviour as described in this section.

Computation effort for different cloud layer parameters

In the left panel of *Figure* 2.18 we present the computation time of pixel models with with a range of different optical thicknesses that shows an increasing computational effort with increasing optical thickness. In the middle panel we can see an increase in computation time for an increase in effective radii (see de Haan et al. 1987*a*). In the right panel we see an increasing computation time for a decreasing top pressure. This might not seem intuitively, but we will explain this by use of some computed pixel models. The source essentially lies in the complex reflection pattern of the clouds in

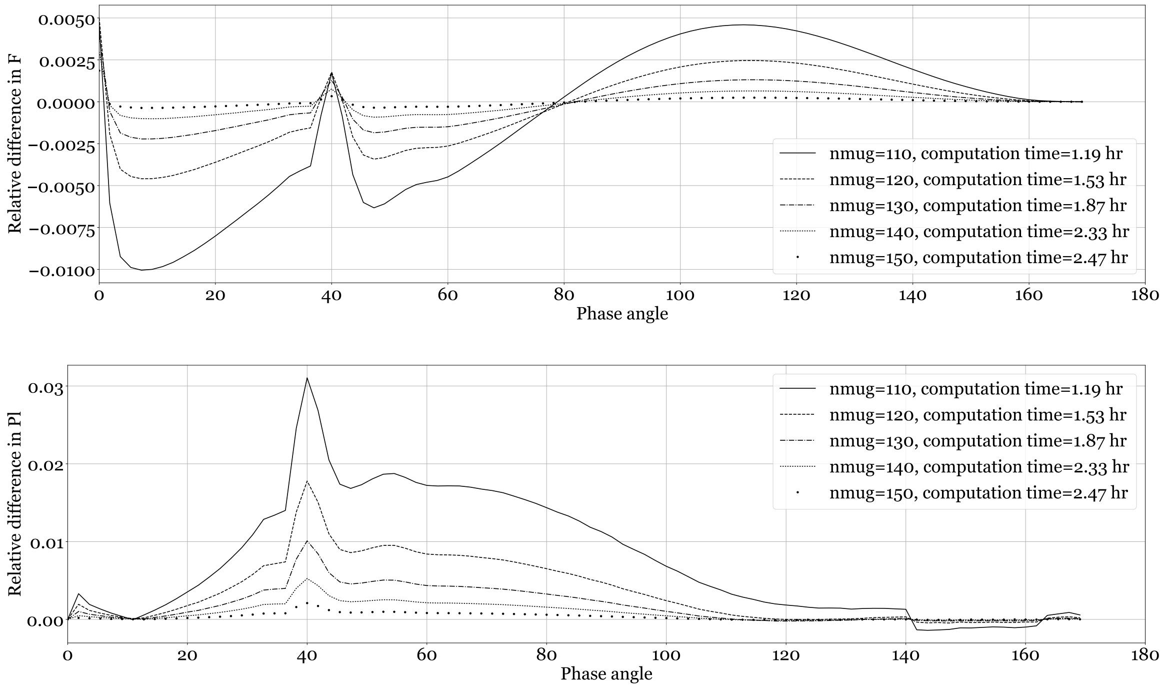


Figure 2.16: Relative difference in F (upper panel) and Pl (lower panel) as function of phase angle for different number of DAP Gaussian abscissae. All computations are made with 100 Mie Gaussian abscissae. Computation times are provided for all cases. The simulation is started at 110 Gaussian abscissae, the smallest number of points that provide a converging sequence.

the geometrical domain: θ_0, θ and $\phi - \phi_0$. Such a complex pattern is provided in the upper two panels of *Figure* 2.19, where we provided polar plots[3] of Q for incidence angles of $\theta_0 = 0, 20, 40, 60$ and $85°$. The emission zenith angle, θ, is provided radially and $\phi - \phi_0$ with the circle periphery. For completeness, we provided the polar plots for F, Pl and U in *Appendix* D, which essentially exhibit similar patterns as Q. These plots show that clouds exhibit ring like regions of high reflection that spread out for higher incident angles. Additionally, we know that the reflection pattern of molecular scattering (Rayleigh scattering) is rather easy (See Stam 2008*a*). If one puts a gaseous atmospheric layer on top of a cloud layer, the complex features are flattened, resulting in a less complex shape to be approximated by a Fourier expansion. By modelling the cloud layer at the top of the model atmosphere the TOA reflection patterns become more complex, requiring more Fourier terms. Similar behaviour also occurs when moving to longer wavelengths. We know that optical thickness of Rayleigh scattering approximately scales with $\frac{1}{\lambda^4}$. This means that with a constant layer division the gaseous layer on top of a cloud layer becomes optically thinner when moving to longer wavelengths, essentially allowing complex reflecting patterns of the clouds to dominate the overall reflection. This effect can also be seen in the lower three panels of *Figure* 2.19, where we modeled TOA reflection Q for $\lambda = 350, 550$ and 865 nm. In conclusion, the computation time increases with decreasing cloud top pressure and with increasing wavelength, due to the increase in complexity of the TOA reflection. Also note from the polar plots that by using 150 Gaussian abscissae the complex reflection patterns are reproduced well.

[3]These polar plots are not to be mistaken by resolved plots of homogeneous planetary disks. That is, the plots show the TOA reflection of a horizontally homogeneous, but vertically inhomogeneous atmospheric model with a low reflecting Lambertian surface.

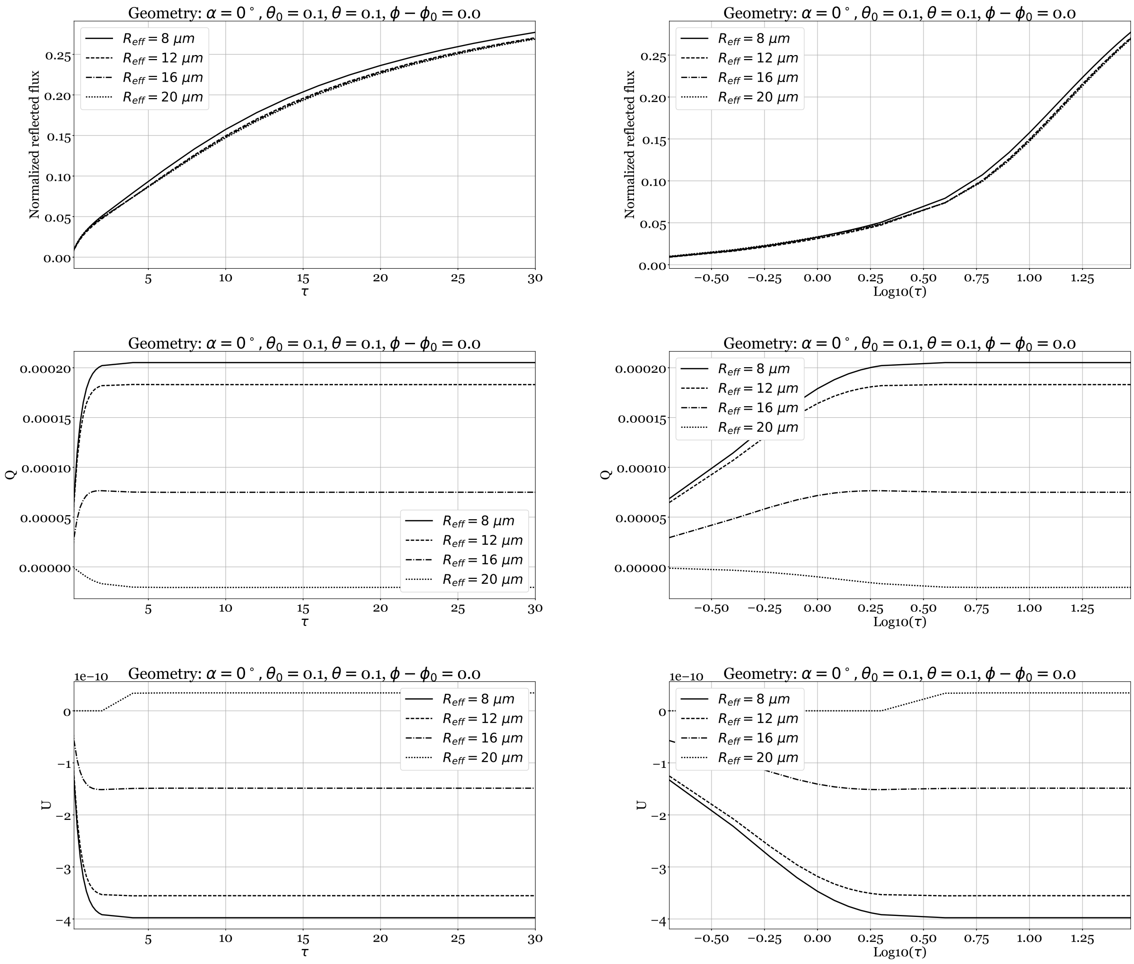


Figure 2.17: Normalized reflected flux (upper panel) and polarized fluxes Q (middle panel) and U (lower panel) as function of cloud optical thickness for four values of cloud particle effective radius. The computations are made for a pixel located at $\alpha = 0^\circ, \theta_0 = 0.1^\circ$ and $\theta = 0.1^\circ$. The left panels show the cloud optical thickness on the x-axis and the right panels the logarithmic value of cloud optical thickness.

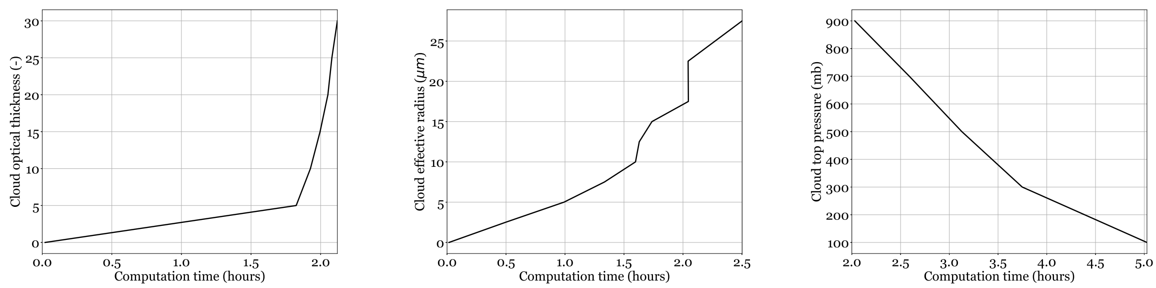


Figure 2.18: Computation time of Fourier files for different cloud optical thickness (left panel), cloud particle effective radius (middle panel) and cloud top pressure (right panel).

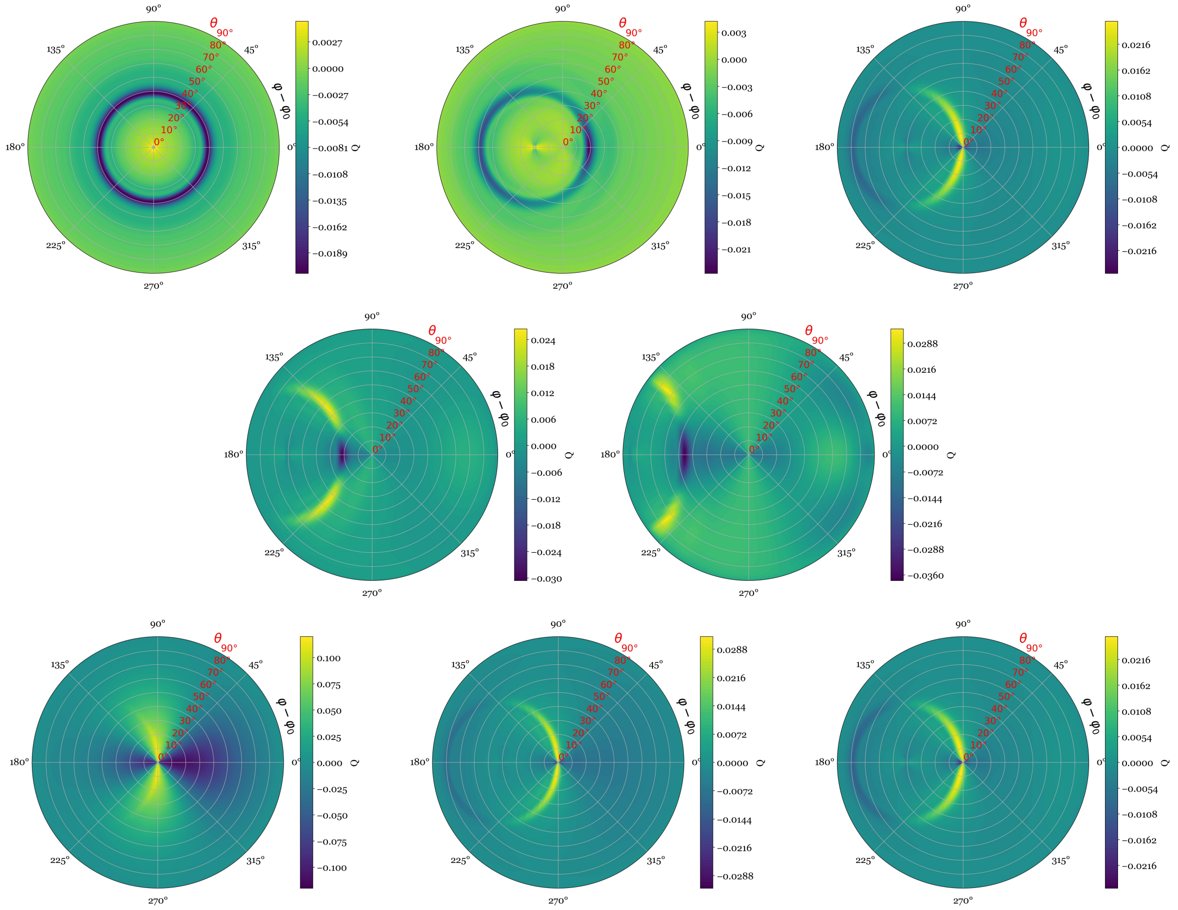


Figure 2.19: Polar plots representing the TOA reflection Q of a locally horizontally homogeneous, but vertically inhomogeneous atmosphere. The atmosphere includes a layer of clouds. The upper left to the middle right panels correspond to solar zenith angles of $\theta_0 = 0, 20, 40, 60$ and $85°$,respectively. In the bottom panel the polar plots for $\lambda = 350,\ 550$ and $865\ nm$ are provided.

3
The Earth model

In this chapter we present the Earth-like planetary model that is used in this research. If one looks at Earth the most distinctive inhomogeneities are the surface and clouds covers. To properly account for these inhomogeneities and still be able to run the PyMieDAP code with adequate computation times and storage a discretized (variable) parameter space is defined. The variable parameters are retrieved from Earth observations (*Section* 3.1). Due to the large computation times of pixel models a discretization will be applied on the model surface (*Section* 3.2.1) and model atmosphere (*Section* 3.3). Based on this discretization a sensitivity study is provided in *Section* 3.4. Further considerations and assumptions regarding our model are provided in the last section of this chapter (*Section* 3.5).

3.1. Earth Observations

PyMieDAP defines a reflecting terrestrial body by its surface, atmospheric molecules and atmospheric aerosols. The properties that described these three physical parts of Earth are extracted from the MOderate Resolution Imaging Spectroradiometer (MODIS) mission databases. In the remainder of this thesis we will refer to MODIS observations whenever we use observations. In the search for a suitable database, the major requirements were the following: quality of observations, preferable daily coverage for multiple years, and coverage of full parameter set.

Two other observational campaigns that provide global atmospheric observations are the International Satellite Cloud Climatology Project (ISCCP) and the Multiangle Imaging Spectroradiometer (MIRS). MIRS is not suitable as it only provides optical depth observations above the oceans and does not provide cloud effective radii observations. ISCCP is not able to retrieve the cloud particle effective radius and base their retrievals on a less accurate algorithm than MODIS (Marchand et al. 2010). Besides MODIS, land cover products can also be obtained from the GlobeCover (MERIS sensor) and GLC2000 (SPOT VEGETATION). These products are developed using unsupervised classification techniques and moreover are only restricted to land cover observations (Friedl et al. 2010).

The daily coverage over multiple years requirement is twofold; firstly, multiple years of observations are required if one wants to model inter-annual variability. As an example, Pallé et al. (2008) used multiple years to assess the accuracy of his rotation rate retrieval method. Secondly, if one wants to approximate observed data, such as Earthshine observations, daily coverage of that date is a minimum requirement, because of the daily changing cloud patterns on Earth.

Using MODIS data sets for every varying model parameter allowed us to simplify the extraction and processing procedures, because they consistently use the same data set structures. Simple examples are the naming conventions a specific data set uses, the file type in which observations are stored and the structure of the data. More specifically, data is most generally stored as an integer representation of a ”real” float value. This integer value then needs to be scaled and/or offsetted after the data is extracted from the file to its ”real” float value.

In the subsequent sections a description of the used MODIS data sets is provided as well as the discretization of the data. A detailed description of all MODIS data set structures, retrieval algorithms and additional information is provided in *EOS Data Products Handbook Volume 1 (2000)* (n.d.); *EOS Data Products Handbook Volume 2 (2000)* (n.d.).

MODIS Land Cover Type Data set from 2016 was retrieved from https://lpdaacsvc.cr.usgs.gov/appeears/, maintained by the NASA EOSDIS Land Processes Distributed Active Archive Center (LP DAAC) at the USGS/Earth Resources Observation and Science (EROS) Center, Sioux Falls, South Dakota. The Aqua/MODIS Cloud Top Pressure, Cloud Particle Effective Radius, Cloud Optical Thickness and Cloud Fraction Daily L3 Global 1 Deg. CMG datasets were acquired from the Level-1 and Atmosphere Archive & Distribution System (LAADS) Distributed Active Archive Center (DAAC), located in the Goddard Space Flight Center in Greenbelt, Maryland (https://ladsweb.nascom.nasa.gov/).

3.2. The model surface

In order to model the reflection by an Earth-like surface, different approximate methods can be applied to a specific type of surface, e.g. a *Fresnel's* reflecting surface to a fresh ice sheet and an ocean surface, or a Bidirectional Reflectance Distribution Function (BRDF) surface to that of vegetation (Coakley 2003). To be able to apply such methods to a specific surface type we need some indication of which surface type or land cover type corresponds to which pixel on the planetary disk. In similar studies that model Earth-like exoplanets, such as Cowan et al. (2009); Fujii et al. (2011, 2010); Kawahara and Fujii (2010, 2011); Oakley and Cash (2009), MODIS land cover observations are also used. To globally retrieve the land cover type, MODIS uses both the Aqua and Terra satellites. Five different classifications for land cover can be used: the International Geosphere-Biosphere Programme (IGBP), University of Maryland classification (UMD), MODIS LAI/FPAR algorithm (LAI/FPAR), Biome (BGC) and the Plant Functional Type (PFT) classification. Of these five classifications, the IGBP classification is used, being the most practical (Friedl et al. 2010). This classification consists of 17 classes, including eleven categories of natural vegetation discretized by life form, three classes of mosaic and developed covers and three non-vegetated classes. An example of an IGBP classified global land cover is provided in *Figure* 3.1. From this cover, one can clearly see the contributions of North and South Polar ice, the Amazone forests and the Sahara dessert. The land cover type algorithm development and validation efforts are based on a large global network of test sites that represent major biomes and cover types. The spatial resolution of the surface coverage data is $500km \times 500km$. On this grid a quality assurance data set is applied to only consider the confidently retrieved pixels in our surface model.[1] The overall accuracy of collection 5.1 is reported to be 74.8%, which is substantiated by several studies (*MODIS Land Validation Web Site* n.d.). In the newer collection 6 data, substantial upgrades have been applied. A study is in review that exactly describes the accuracy and development of this new data set that we applied in our model (*User Guide to Collection 6 MODIS Land Cover (MCD12Q1 and MCD12C1) Product* n.d.a).

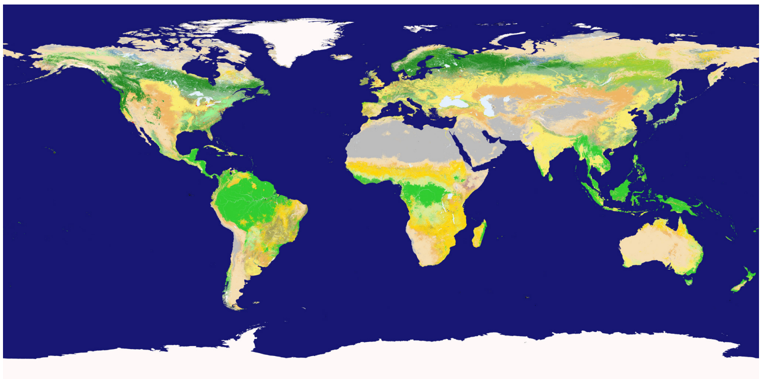


Figure 3.1: Global MODIS land cover type according to the IGBP classification.

[1]more information at: https://lpdaac.usgs.gov/sites/default/files/public/product_documentation/mcd12_user_guide_v6.pdf

3.2.1. Surface discretization

The surface is modeled with the land cover database of MODIS in combination with the Aster Spectral Library. The Aster Spectral Library (*User Guide to Collection 6 MODIS Land Cover (MCD12Q1 and MCD12C1) Product* n.d.*b*) provides Lambertian Equivalent Reflection (LER) albedos for the desired wavelength region, shown in *Figure* 3.2. Because of the high spatial resolution of the land cover observations we may end up with large sets of different land cover types in one disk pixel. To assign one value of land cover type to a pixel we will use the mode of the entire set of data that falls into that pixel. When using Lambertian albedos, a better approximation could have been to weight the land cover types by its reflection. But, if one were to use geometrical dependent surface models, i.e. anisotrope reflecting surfaces, such as a BRDF or a *Fresnel's* reflecting surface, this type of weighted processing of land cover becomes very complex.
For a first order approximation we will use four different surface types: Ocean, Vegetation, Desert and Snow/Ice. In *Table* 3.1 the IGBP classification for the four discretized surface covers is provided. Fujii et al. (2010) used the same discretization which showed to provide fairly accurate results in their attempt to retrieve land distributions from their simulated photometric signals.

Table 3.1: IGBP classification applied to the model surface

No.	IGBP Classification	Our Classification	Ocean/Land
0	Water	Ocean	Ocean
1	Evergreen needleleaf forest	Vegetation	Land
2	Evergreen broadleaf forest	Vegetation	Land
3	Deciduous needleleaf forest	Vegetation	Land
4	Deciduous broadleaf forest	Vegetation	Land
5	Mixed forest	Vegetation	Land
6	Closed shrubland	Vegetation	Land
7	Open shrubland	Soil	Land
8	Woody savannas	Vegetation	Land
9	Savannas	Vegetation	Land
10	Grasslands	Vegetation	Land
11	Permanent wetlands	Soil	Land
12	Croplands	Vegetation	Land
13	Urban and built-up	Soil	Land
14	Cropland/natural Vegetation mosaic	Vegetation	Land
15	Snow and Ice	Snow	Land
16	Barren or sparsely vegetated	Soil	Land

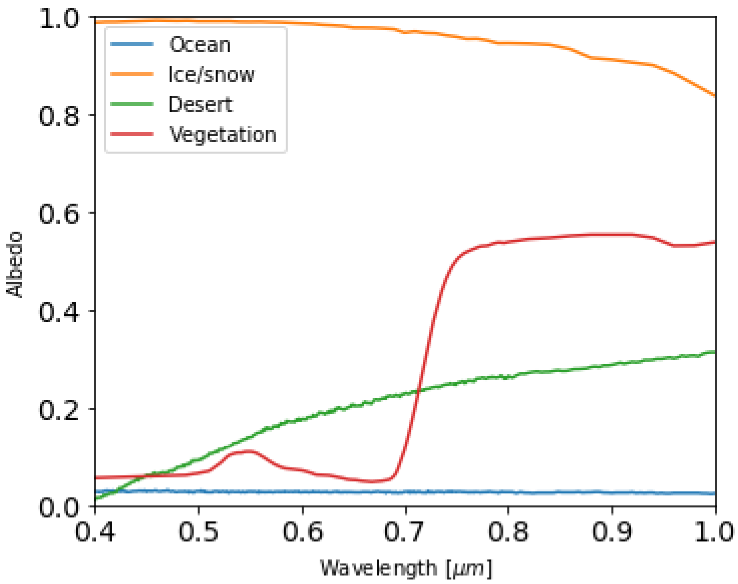


Figure 3.2: Lambertian Equivalent Reflection for Ocean, Ice/snow, Desert and Vegetation surface covers in the ultra-violet, visible and near infrared spectral domain (*User Guide to Collection 6 MODIS Land Cover (MCD12Q1 and MCD12C1) Product* n.d.*b*).

3.3. The model atmosphere

From a theoretical point of view the reflection behaviour of a cloud is depending almost exclusively on the cloud optical thickness and cloud particle effective radius (Fouquart et al. 1990; Nakajima and King 1990; Slingo 1989). Modeling clouds in an Earth-like atmosphere introduces two more variables: the vertical position and fraction of clouds. The effect of both parameters on reflected spectropolarimetric signals was thoroughly examined by Rossi and Stam (2017). Their results show a significant effect on the phase curves of F and P in an edge-on configuration. Furthermore, using PyMieDAP, one may also vary the effective variance and the refractive index of the cloud particles. From a practical view modeling a variable effective variance might not be ideal as global and daily observations are not available. On the other hand, does the variability induce significant changes in our spectropolarimetric signals? This question can also be asked about the variable particle refractive index. These questions will addressed in *Section* 3.3.2. Multi-layered liquid water clouds are not considered in this thesis. For a thorough analysis about multi-layered clouds one may consult Karalidi, Stam and Hovenier (2012).

In the following subsections we will describe the parameterised gaseous atmosphere (*Subsection* 3.3.1) and clouds (*Subsection* 3.3.2).

3.3.1. Parameterization of the molecular atmosphere

The spatial variation of the molecular atmosphere is kept constant in this analysis. A constant Earth-like gas mixture is assumed with a molecular mass of $29g/mol$, a depolarization factor of 0.03 and constant gravitational acceleration of $9.81m/s^2$ (Hansen and Travis 1974*a*; Rossi and Stam 2017). A wavelength dependent air refractive index is used to compute the scattering cross section of the gas molecules, according to the dispersion formula of Ciddor (1996). We will only use continuum wavelengths, thus no gaseous absorption is considered. The atmosphere is modeled by three horizontally homogeneous layers, being in hydrostatic equilibirum, on top of a horizontally homogeneous Lambertian surface at a surface pressure of $1bar$. If present, the cloud particles are placed in the middle atmospheric layer. The specification of such a cloud layer is given in the next section.

3.3.2. Parameterization of the clouds

Clouds are modeled as a horizontal homogeneous atmospheric layer filled with aerosols and a vertical extent of $100\ mb$. The aerosols are homogeneous spherical liquid water particles, with a constant refractive index of $n_r = 1.33 + 1e - 8i$ (Hale and Querry 1973) and a two-parameter gamma particle size distribution (Hansen and Travis 1974*a*). Daimon and Masumura (2007); Kokhanovsky (2004) show that the real part of the refraction index in the ultra-violet, visible and near-infrared wavelength regions is fairly constant, but the imaginary part varies significantly between $n_i = 10^{-5} - 8 \times 10^{-10}$ (Pope and Fry 1997). Hansen and Travis (1974*a*) showed the effect of n_i on the scattering properties of a terrestrial atmosphere for various effective particle size distributions ($x = 1 - 1000$). They provide that for $0.001 \geqq n_i$ the effect on the single scatter properties of cloud particles[2] is negligible.

A gamma size distribution is described by the particle effective radius and the effective variance. The effective variance is based on Earth clouds according to Han et al. (1994*b*); Nakajima and King (1990): $v_{eff} = 0.1$. In the ultra-violet, visible and near-infrared wavelength regions this effective variance varies between $\sim 0.05 - 0.2$ (Diem 1948; Platnick et al. 2015). Hansen and Travis (1974*a*) simulated the single scattering albedo as function of the effective size parameter (x) for various values of the effective variance. For values of $x \geqq 10$ the single scattering albedo is practically insensitive to the width of the particle size distribution. For Earth-like cloud effective radii and wavelength in the range of $0.3 - 1.0\mu m$ we obtain effective size parameters of $\gtrsim 30$. It is thus safe to assume a constant effective variance of $v_{eff} = 0.1$.
As mentioned before, the particle effective radius, optical thickness and top pressure are not constant in our model and will vary according to the MODIS observations. The horizontal position of clouds is modeled by the resolved planetary disk approximation. Additionally, in this research we will not consider ice particle clouds. Firstly, because they are hard to model (Liou and Yang 2016) and a method to do so is not implemented in PyMieDAP, although it is possible to import scattering matrices of ice crystals into

[2]The particle effective radius for cloud ranges from $\sim 5 - 30\ \mu m$ (Han et al. 1994*a*), resulting in a range of size parameters for the ultra-violet to near-infrared wavelength range of: $x =\sim 35 - 540$.

a layer. Secondly, it would increase the number of pixel models and thereby increasing computation time and storage. The effect of ice clouds on the spectropolarimetric signal in an Earth-like atmosphere is modeled by Karalidi, Stam and Hovenier (2012). In this paper, the effect of ice clouds on the liquid water cloud enhancements in flux and polarization, primarily the rainbow feature, shows to be fairly small; unless the atmosphere contains a very large number of ice clouds the primary rainbow feature of the liquid water clouds is still observable in the polarimetric signal. Karalidi, Stam and Hovenier (2012) also shows that the scattering properties of underlaying liquid water clouds are masked by ice clouds if their optical thickness exceeds approximately $\tau_{ice} \gtrsim 3$. From Rossow and Schiffer (1999) we can read that the most abundant ice clouds, i.e. Cirrus and Cirrostratus, have both annual mean optical thickness values of ~ 2.2. The third less abundant ice cloud type, the Deep convective cloud, has annual mean optical thickness values of over 30 but has only an approximate abundance of $\sim 2.6\%$. We therefore expect that ignoring ice clouds only insignificantly effects our spectropolarimetric signals, but greatly limiting computation times and storage.

In order to restrict the computation time and disk space for computing the pixel models even more, bins need to be created for the varying cloud properties. These restrictions are mainly based on the most abundant types of clouds in Earth's atmosphere and their position, and to second order based on computation time and storage. The histograms, which will show the abundance of each parameter, are based on the MODIS monthly averaged data sets: MYD08-M3 collection 6.1. A full year is chosen, such that any seasonal effects on the cloud properties are covered. To adequately represent the distributions we fitted 200 of the most commonly used continuous statistical distributions to the empirical data (Kotz 1994), from which we obtained a *Johnson SU Distribution*, *Johnson SB Distribution* and a *Generalized Normal Distribution* for the cloud optical thickness, cloud particle effective radius and cloud top pressure, respectively. From these distributions we can define confidence intervals, from which we strive to cover 95% of the data adequately. The distributions and the confidence intervals are shown in *Figure* 3.3.

The observations for cloud top pressure, cloud particle effective radius and cloud optical thickness are provided for probably cloudy and confidently cloudy pixels only, in time series of one day (solar day), eight days or a month (Platnick et al. 2015). In order to obtain the appropriate pixel cloud coverage we also need to apply a cloud fraction data set. In the subsequent sections we will provide a description and discretization of the used observations.

Cloud Top Pressure

The cloud top pressure is provided for the full spatial domain of latitude and longitude in a geographical format on a $1^\circ \times 1^\circ$ pixel sized grid. This daily data product is part of the MYD08 data set. A representation of such a grid is provided in *Figure* 3.4. In order to obtain the cloud top pressure of a planetary disk pixel, a weighted arithmetic mean will be applied on the cloudy MODIS pixels[3] only. The averaging routine is defined by:

$$\bar{P_c} = \frac{1}{N_c} \sum_{i=1}^{N_c} P_{c,i} * CF_i \tag{3.1}$$

where N_c are the number of cloudy MODIS pixels, $P_{c,i}$ the cloud top pressure for a MODIS pixel i, and CF_i the cloud fraction for a MODIS pixel i. MODIS uses the same averaging scheme (Hubanks et al. 2015), which is applied on the MOD06 data set from which the MYD08 data set is derived. The weights on the top pressure data are simply the collocated cloud fractions. The equatorial data gaps, which are caused by the limited spatial coverage of the satellite, are interpolated by averaging over only the valid pixels that fall into the disk pixel, and applying the same weighted arithmetic mean (Fujii et al. 2011).

Cloud Top Pressure Discretization

We will model the cloud top pressure on a relatively large range of top pressures as compared to the ISCCP classification (Rossow and Schiffer 1999). In *Figure* 3.5 we see that the mean cloud top

[3]Whenever we use the term MODIS pixels we refer to the pixels in the MODIS observation data sets.

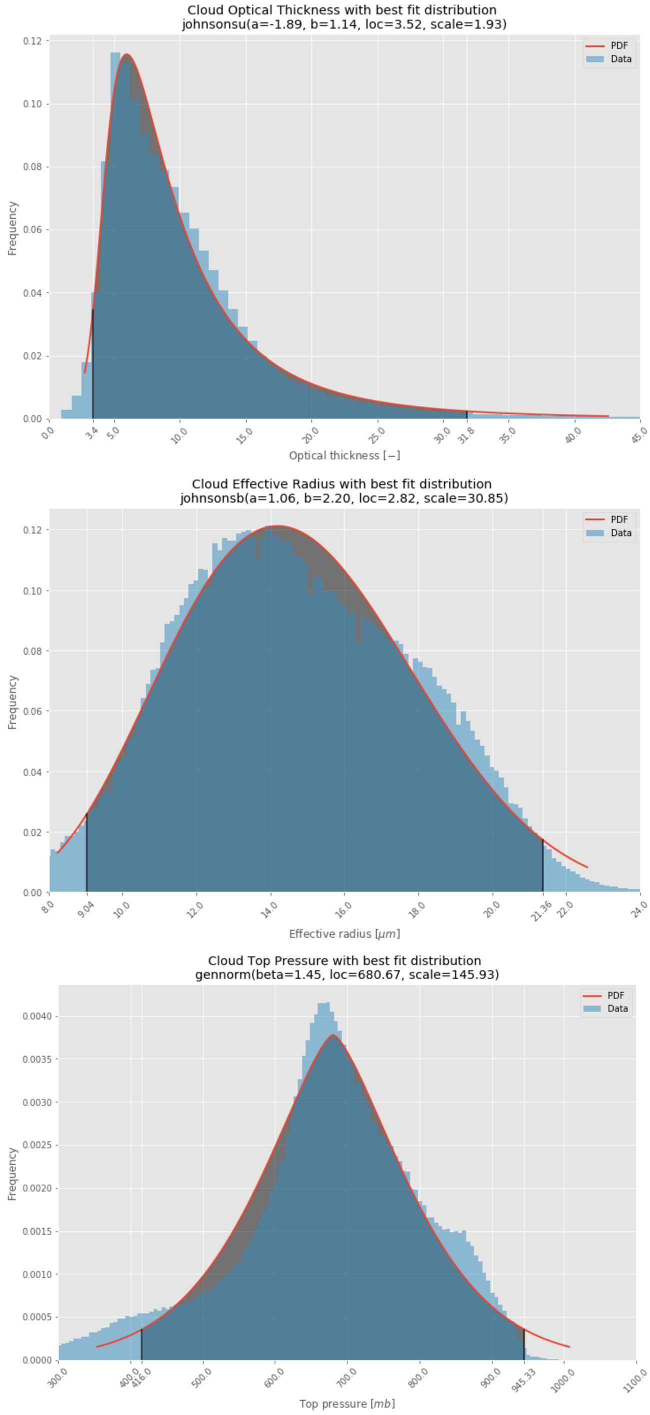


Figure 3.3: Histograms and fitted Probability Density Functions (PDF's) to the cloud optical thickness (upper panel), the cloud particle effective radius (middle panel) and the cloud top pressure (lower panel). For all PDF's a confidence interval of 95% is provided. On top of the figures the PDF fitting parameters are provided.

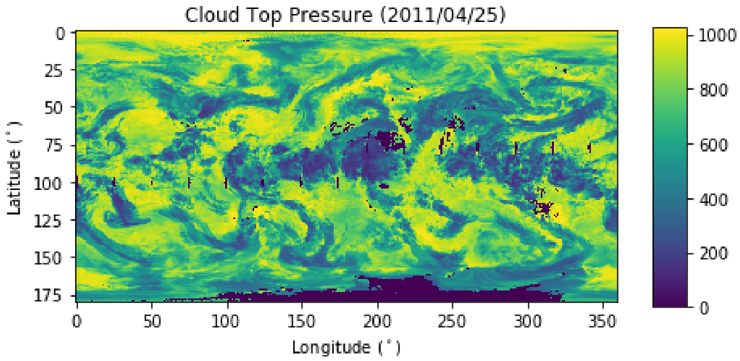


Figure 3.4: Daily averaged global MODIS cloud top pressure distribution.

pressure as function of latitude remains between approximately $\sim 550 - 800\ mb$, where this relation is based on monthly mean MODIS data from 2011 entirely. Also, from *Section* 2.3.2 we know that the computation time significantly increases with decreasing top pressure. From *Figure* 3.3(c) we see that the peak cloud top pressure is located at around $\sim 680\ mb$, which corresponds approximately with previous studies that show average cloud top pressures of $700\ mb$. We will center our bins on this average cloud top pressure of $700\ mb$. In the histogram distribution we see another local peak at around $850\ mb$, on which we center the second bin. The low cloud top pressures will be approximated by the third bin, centered at $500\ mb$, around which we see still some significant abundances. For the cloud top pressure we also considered the cloud top pressure - cloud optical thickness bins of ISCCP (Rossow and Schiffer 1999). In a later research by Hahn et al. (2001), however, a comparison with individual weather observations shows that the validity of this discretization is somewhat coarse for at least "low clouds". We do however cover approximately the same region of cloud top pressure that represent the liquid water clouds (Rossow and Schiffer 1999), but distinguishing it into three bins as compared to their two bin approximation. The three bins in *Table* 3.2 will be used to discretize the cloud top pressure. The bin boundaries show the interval in which a pixel parameter may fall. For every parameter that falls in such a bin, it is approximated by a bin value.

Table 3.2: Discretization bins of the cloud top pressure with associated bin values.

Bin number	1	2	3
Bin boundaries	0<CTP≦600	600<CTP≦800	800<CTP
Bin values	500	700	850

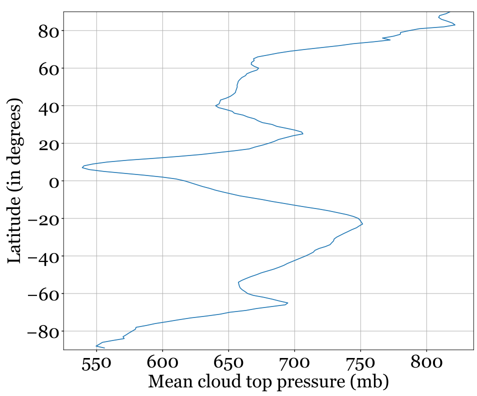


Figure 3.5: Latitudinal dependence on mean cloud top pressure derived from monthly data in 2011 entirely.

Cloud Particle Effective Radius

Cloud particle effective radius observations are restricted to daytime only as the retrieval is based on measurements of reflected sunlight. The observations are provided on a global grid of 1° by 1° pixels as shown in *Figure* 3.6. The disk pixel cloud particle effective radius is computed in a similar fashion as the cloud top pressure:

$$\bar{r_{eff,c}} = \frac{1}{N_c} \sum_{i=1}^{N_c} r_{eff,c,i} * CF_i \tag{3.2}$$

where N_c are again the number of cloudy pixels, $r_{eff,c,i}$ the cloud particle effective radius for a MODIS pixel i, and CF_i the cloud fraction for a MODIS pixel i. Interpolation of the latitudinal and orbital gaps is the same as that of the cloud top pressure, only for these daytime only observations we obtain a large gap at high latitudes. For the relatively coarse pixel disk sizes to the observations this type of interpolation is possible at these outer regions. If significantly more disk pixels are required, reprocessing of the data can be considered, or one could consider using 8-day or monthly averaged MYD08 data sets to fill the gaps.

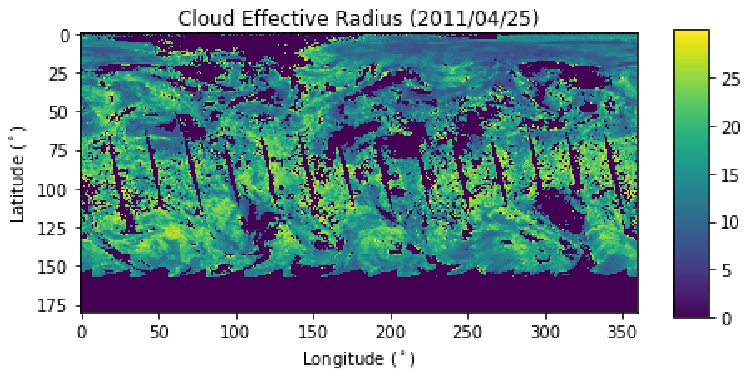


Figure 3.6: Daily averaged global MODIS cloud particle effective radius distribution.

Cloud Particle Effective Radius Discretization

The 95% interval of cloud particle effective radius lies between $9-21$ μm. From *Section* 2.3.2 we know that the computation time increases for increasing particle effective radii. We see, however, from *Figure* 3.7 that the largest effective radii are located near zero latitude. The histogram bars and the fitted distribution show that there are two regions of high abundance, one located near ~ 12.5 μm and one near ~ 17.5 μm. Two of the bins are centered at these regions. We will, however, not use a bin larger than 17.5 μm, due to the high computation time. A third bin is located at the lower boundary of the distribution, centered at 10 μm describing the fairly steady decreasing abundance. The fourth bin is centered at 15 μm that approximates the fairly steady varying abundance between the fist two bins. In total, the distribution is thus approximated by four equally spaced bins with corresponding bin values, provided in *Table* 3.3. A similar discretization is used by Nakajima and King (1990), also to

Table 3.3: Discretization bins of the cloud particle effective radius with associated bin values.

Bin number	1	2	3	4
Bin boundaries	0≦ CER<11.25	11.25≦ CER<13.75	13.75≦ CER<16.25	16.25≦ CER
Bin values	10	12.5	15	17.5

numerically simulate Earth clouds.

Cloud Optical Thickness

In *Figure* 3.8 a daily global observation of the cloud optical thickness is provided. The cloud optical thickness have the same spatial extend as the cloud particle effective radius, i.e. daytime only, and the same spatial resolution. For the cloud optical thickness two data sets are available: one based on

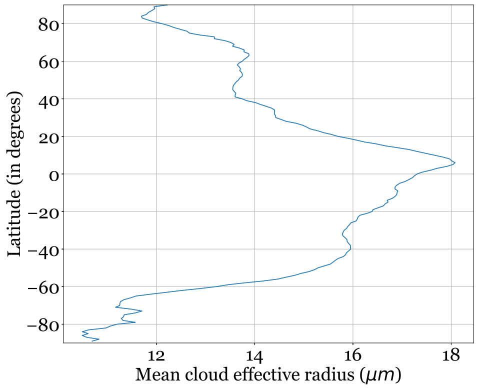


Figure 3.7: Similar to *Figure* 3.5, except for the cloud particle effective radius.

the "logarithmic mean", and one based on the "arithmetic mean". The "logarithmic mean" processed data set provides for a better approximation in the range of optical depths of $0.01 - 100$, according to Hubanks et al. (2015); Oreopoulos et al. (2007). In the analysis performed in *Section* 2.3.2 we observed the same results for total flux as well as for the polarized fluxes.
If we now zoom out to the planetary disk, we will not only be subjected to cloudy pixels, but also to clear pixels. These clear pixels will have a cloud optical thickness of zero. The problem that occurs is that for a single disk pixel that is patchy, we will have a fraction of cloudy pixels and clear pixels, τ^{cld} and τ^{clear} respectively.

This means that we need to define some sort of average disk pixel that includes both the radiative transfer behaviour of clear and cloudy pixels or more specifically, a Rayleigh scattering and a Rayleigh/Mie scattering behaviour. To approximate a patchy pixel by a cloudy and clear region their reflection matrices $\mathbf{R}_1^{cld}(\mu_n, \mu_{0n}, \phi_n - \phi_{0n}, \lambda)$ and $\mathbf{R}_1^{clear}(\mu_n, \mu_{0n}, \phi_n - \phi_{0n}, \lambda)$ need to be calculated. Ideally, to approximate one disk pixel by a cloudy and clear part, not only the reflection matrices should be case specific but also the geometrical parameters; the solar and emission zenith angles, the azimuthal difference angles, etc. In practice this is not possible because MODIS only provides which ratio of their pixels is cloudy and does not provide their specific location. We will approximate the single patchy disk pixel by calculating the reflection properties for a cloudy and clear case separately, at the center of the disk pixel, and weigh both pixels by the local cloud fraction. This method is similar to that presented by Stam (2008*a*), who approximated a horizontally inhomogeneous planet by weighting the results of horizontally homogeneous planets. In our case we will only use this approximation for one disk pixel. The total reflection of the patchy pixel will then be approximated by:

$$\mathbf{R}_{tot} = \mathbf{R}(\tau_{cld})\frac{\sum_{i=1}^{N} Fcld_i}{N} + \mathbf{R}(\tau_0)(1 - \frac{\sum_{i=1}^{N} Fcld_i}{N}) \tag{3.3}$$

where $\mathbf{R}_{tot}$ is the total reflection of the pixel, $\mathbf{R}(\tau_T^{cld})$ is the reflection of a pixel model that includes a cloud layer (with parameters according to the collocated MODIS cloud observations), $\mathbf{R}(\tau_0)$ the reflection of a pixel model that does not include a cloud layer, $\sum_{i=1}^{N} Fcld_i$ the total fraction of clouded MODIS pixels and N is the total number of MODIS pixels (cloudy+clear) that fall into the disk pixel. The reflection of the cloudy pixel will be computed with the logarithmic averaged optical thickness mentioned before. The only discrepancy of this approach is that the position of the cloudy and clear pixels cannot be precisely modeled, but we do not know these anyway. Also, this method allows the user to apply the exact factor of cloudy pixels. Furthermore, if one wants to incorporate more atmospheric scattering models in one disk pixel, such as models that describe scattering from ice, dust, biomass-burning, urban and/or maritime mineral aerosols, these can be easily included by incorporating them in the weighted sum.

Interpolation of the equatorial gaps is processed in the same sense as the other cloud parameters.

To minimize the effect of surface reflection, MODIS retrieves the optical thickness at 645, 858 and $1240\ nm$ MODIS bands for respectively land, ocean and snow/ice surfaces. So, in order to compute the TOA reflection of a cloudy pixels at a specific wavelength λ^* we will scale the cloud optical thickness assuming a constant column density. In our computations we do not consider any absorption by cloud particles, thus the scattering cross-section is equal to the extinction cross-section. We retrieve the extinction cross-section for a specific wavelength by use of the *Mie* subroutine in PyMieDAP. To scale the optical thickness the following formula is used

$$\tau^*(\lambda^*) = \frac{\sigma^*_{ext}(\lambda^*)}{\sigma_{ext}(\lambda)}\tau(\lambda) \tag{3.4}$$

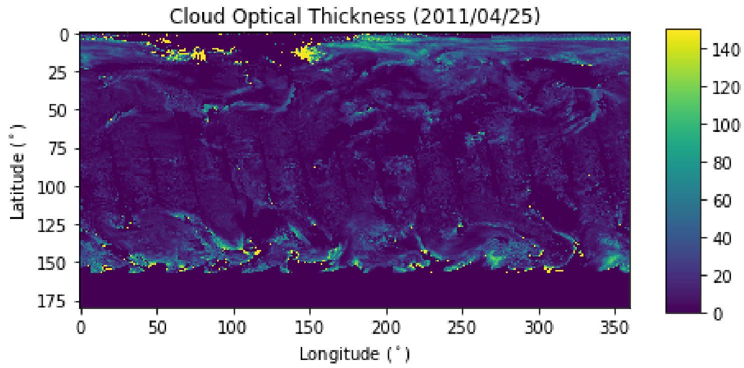


Figure 3.8: Daily averaged global MODIS cloud optical thickness distribution

Cloud Optical Thickness Discretization

The upper panel of *Figure* 3.3 shows that the cloud optical thickness peaks at approximately 5. We further see that 95% of the data lies between optical thickness values of $3 - 31$. From *Section* 2.3 we know that the computation time increases with increasing optical thickness, so we want to restrict high values of optical thickness as much as possible. To be able to further reduce the interval of interest we will look at the relation of the optical thickness as function of latitude (*Figure* 3.9). This relation is based on monthly mean data from 2011 entirely, and the mean value is calculated according to Hubanks et al. (2015); Oreopoulos et al. (2007). The effect of pixels at high latitude regions on the total exoplanet reflection properties is significantly lower than that of low-latitudes due to the combination of high solar and emission zenith angles. Furthermore, we can see from *Figure* 3.9 that the average optical thickness only just merely approaches 20 at $\sim 78^\circ$. A similar average maximum cloud optical thickness is observed by Hahn et al. (2001), for Nimbostratus clouds. Also, if we look at the reflection of cloudy pixels, we observe that for the polarized fluxes the reflection evens out fairly quickly and for the total flux starts to even out for optical values of 30 and higher values, thus not contributing significantly to the total reflected spectropolarimetric signal. We did also look at the basic ISCCP cloud type classification that is based on cloud top pressure and cloud optical thickness (Rossow and Schiffer 1999), but similarly as described in the discretization of cloud top pressure there occur ambiguities in these bins. The four bins in *Table* 3.4 will be used to discretize the cloud optical thickness. The first bin is used to extract the clear sky pixels from the observations. The second

Table 3.4: Discretization bins of the cloud optical thickness with associated bin values.

Bin number	1	2	3	4
Bin boundaries	COT=0	0.01≦ COT<7.5	7.5≦ COT<15	15≦ COT
Bin values	0	5	10	20

bin is centered at the peak abundance, to model the largest amount of clouds. For slightly higher optical depths we observe a small bumb at approximately a value of 10 on which we center the third bin. The last bin is centered at the maximum mean value of 20, modelling the very edge of the 95%

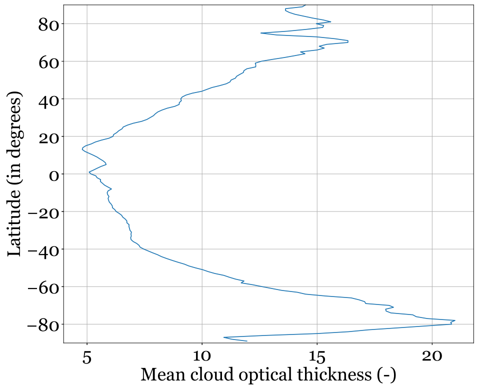


Figure 3.9: Similar to *Figure* 3.5, except for the cloud optical thickness. Mean values are computed according to Hubanks et al. (2015).

interval. The consequence of these bins is that some over and/or underestimation will occur, but this is inevitable when attempting to discretize the parameter space.

Cloud fraction

MODIS provides two cloud fraction data sets, namely: Cloud Fraction and Cloud Retrieval Fraction. According to Hubanks et al. (2015) does the Cloud Retrieval Fraction account to a better fit in regions with high amounts of aerosols such as dust, but performs worse in the interpretation of cloud-edges. For these areas MODIS uses an algorithm that identifies smoke contamination, partly clouded pixels, the sunglint, edges of clouds and heavy dust. These cases are expected to deviate from the homogeneous overcast cloudy 1-dimensional plane-parallel radiative transfer approximation for optical property retrieval and are assigned as a clear sky pixel. The Cloud Retrieval Fraction is thus derived from the Cloud Optical Properties retrieval algorithm and therefore assigns pixels with a clear sky label and fits the cloud particle effective radius and cloud optical thickness data exactly, whereas this is not the case for the Cloud Fraction (Platnick et al. 2015). However, when comparing the mean cloud fractions for both databases of the year 2011 to over a decade worth of ISCCP cloud data we see that the Cloud Fraction agrees significantly better: MODIS mean Cloud Fraction of 0.68, MODIS mean Cloud Retrieval Fraction of 0.27 and an ISCCP mean cloud fraction of 0.675 ± 0.012 (Rossow and Schiffer 1999). Thus the cloud retrieval fraction underestimates the ”real” cloud fraction in case of cloud-edges, as the confidence of cloud property retrieval at these edges is less confident. Furthermore, Ackerman et al. (1998, 2008); Li et al. (2004) reviewed the Cloud Fraction data set to other collocated observational data sets for which they retrieve good agreement. In this regard the choice is made to use the Cloud Fraction data set.

3.4. Sensitivity study

In *Table* 3.5 an overview of the entire surface-atmosphere system is provided. The values from this

Table 3.5: Overview of the discretized model atmosphere that will be used in the planetary model.

Parameter	Symbol	Value(s)
Surface (bounding) pressure [bar]	P_{surf}	1
Depolarization factor	δ	0.03
Mean molecular mass $[g/mol]$	m_g	29
Acceleration of gravity $[m/s^2]$	g	9.81
Cloud particle effective variance	v_{eff}	0.1
Cloud particle effective radius $[\mu m]$	r_{eff}	10;12.5;15;17.5
Cloud particle distribution	-	Two parameter gamma
Cloud particle refractive index	$n_c = n_r + n_i$	$1.33 + 1e - 08i$
Cloud optical thickness [-]	τ	0;5;10;20
Cloud top pressure [mb]	P_c	500;700;850
Cloud vertical extend [mb]	-	100

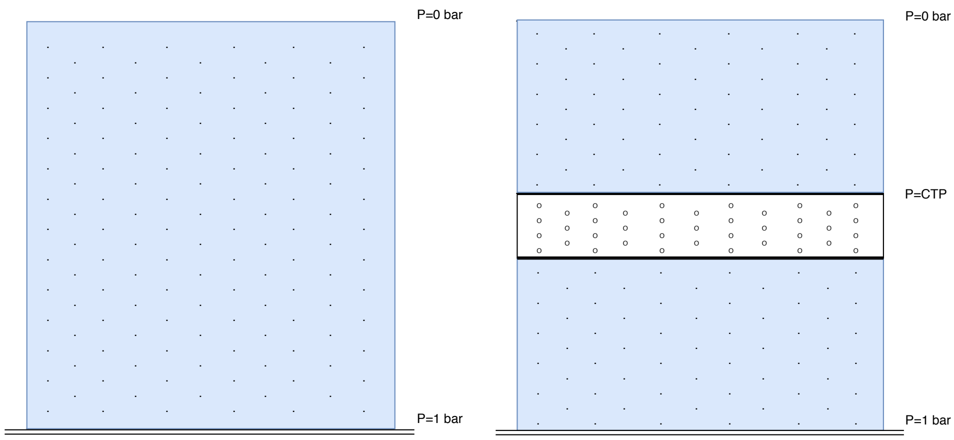


Figure 3.10: Sketch of the surface-atmosphere pixel model without a cloud (left) and with a cloud (right). A cloud layer also contains gaseous particles.

table are based on the schematic model atmospheres shown in *Figure* 3.10. We compute either clear sky pixel models, with no cloud layer, or a cloudy pixel model with a cloud layer. Both atmospheres are bounded by pressure levels of zero and one bar at which we model a reflecting surface layer. The cloud layer is defined by its top pressure; the size distribution of particles, i.e. the particle effective radius; and the optical depth of the layer. In this layer we do still consider Rayleigh scattering from molecular particles. Lastly, for the entire atmosphere the molecular mass and depolarization factor in the Rayleigh scattering computations is assumed constant.

In order to investigate how sensitive the spectropolarimetric signals are to the binned cloud parameters, we will generate some cases where we restrict the bins. For all cases we model the phase curves with low temporal resolution: 3.059 days. With this resolution the Earth rotates four times as seen from the observer. The reference case is the model utilizing all bin values. The acronyms CTP, COT and CER represent the cloud top pressure, cloud optical thickness and cloud particle effective radius, respectively. The following cases are computed:

1. Without CTP of 850 mb -> pixels with CTP of 850 mb are considered as 700 mb;
2. Without CTP of 500 mb -> pixels with CTP of 500 mb are considered as 700 mb;
3. Without CTP of 850 and 500 mb -> both type of pixels are considered as 700 mb;
4. Without COT of 20 -> pixels with COT of 20 are considered as an optical depth of 10;
5. Without COT of 10 -> pixels with COT of >10 are considered as 20 and <10 as 5;
6. Without COT of 20 and 10 -> all pixels are considered as an optical depth of 5;
7. Without CER of $10\mu m$ -> pixels with CER of $10\mu m$ are considered as pixels with a CER of $12.5\mu m$;
8. Without CER of $15\mu m$ -> pixels with CER of $> 15\mu m$ are considered as $17.5\mu m$ and $15 >, > 11.25\mu m$ are considered as $12.5\mu m$;
9. Without CER of 10 and $15\mu m$ -> pixels with CER of $> 15\mu m$ are considered as $17.5\mu m$ and $15 > \mu m$ are considered as $12.5\mu m$;
10. Without CER of $17.5\mu m$ -> pixels with CER of $17.5\mu m$ are considered as pixels with a CER of $15\mu m$;
11. Without CER of $10,$ 15 and $17.5\mu m$ -> all pixels are assigned a CER of $12.5\mu m$.

For the first case we neglect low altitude clouds and consider them as the most abundant cloud type in our observations. In case two the same is true for high clouds. Then for the third case we consider only an average cloud top pressure of $700\ mb$. In the discretization of the COT we saw that the average optical thickness values were only significant for high latitudes. By considering optical values of 20 as 10 we will see how much effect the abundance of thick clouds have. Furthermore, we see in the histogram of the optical thickness that there was a small bump at optical values of 10. By neglecting this bin value we will be able to observe its effect on the planetary phase curves. For the last case of COT we approximate every pixel by the optical depth with the highest abundance: 5. The highest abundance of CER lies approximately at the bin value of $12.5\ \mu m$. The last cases show the individual sensitivity of neglecting the other effective radii values and a combination of these values.

From *Figure* 3.11 one can see the effect of using more heavily discretized bins for the three cloud parameters relative to a reference phase curve. This reference phase curve is computed with the same observations and temporal resolution of 3.059 days. From the top left panel one can see that for low phase angles there is a significant disagreement when we ignore bin values for COT of 10 and 20. Also when only ignoring the COT 10 bin we see a relative difference near 60 degrees of $\sim 5\%$. At crescent phases we see a large disagreement for a discretized bins with bin values of CTP $700\ mb$ and CTP $700, 850\ mb$. A much smaller, but still significant difference is induces by using bin values of CTP $700, 500\ mb$, but this is much smaller than the other CTP bin cases. The heavily discretized bins for cloud particle effective radius induce only minor differences, for which the heaviest discretized bin CER $12.5\ \mu m$ shows the largest disagreement. These results show that for the normalized reflected flux we certainly need to use the full bins for at least the cloud optical thickness and top pressure. We also see that if we use all the bin values in our bins we get a behaviour of convergence.
In the upper right and middle left panel we see the effect on the degree of polarization and normalized polarized flux Q, respectively. The behaviour of both is almost identical: the largest errors, up to $300 - 400\%$, are induces at low phases by a more heavy CER bin discretization and both CER and CTP bin discretizations at crescent phases. For the discretized COT bins we see only a maximum relative difference of $\sim 15\%$. Again, if we approach our maximum bin values we see a convergence behaviour to our reference phase curve. Lastly, in the middle right panel we see that the different bins induce very large errors, up to $3000 - 4000\%$ for heavy discretization on the CTP and COT bins. The errors induced by the heaviest CER bin discretization are up to $\sim 350\%$. These major disagreements lie in the region of $\alpha = 40 - 60°$. The reason for this large disagreement is because the absolute values of U are really small, and badly discretized bins induces relatively large differences in U. We do, however, again see some sort of convergence when we decrease the discretization on all the bins. For example, the case "Wo CER 10" we see an overall agreement within 15%, except at $\alpha =\sim 50°$. For the case "Wo CER 15" we see an overall agreement within 40%, except at $\alpha =\sim 50°$. Our full bin Earth-like planetary model will be more accurate and produce smaller errors, which for the normalized reflected flux, degree of polarization and Q will not be significant anymore, but for U, especially around $\alpha = 40 - 60°$, will still produce a disagreement of at least $\lesssim 15\%$. This sensitivity study is simulated at $\lambda = 550\ nm$. Simulations at $\lambda = 350$ and $865\ nm$ show overall better agreement.

3.5. Observation strategy

The orbital plane of the exoplanet around its parent star can be inclined with respect the observer, e.g. Earth. In *Figure* 3.12 the definition of this inclination is shown. The inclination is defined as the angle between the total angular momentum vector of the extrasolar system and a line connecting the observer (Earth) and the star. All computation in this research are conducted with a spherical Earth in a $90°$ inclined circular orbit, with the orbital period the same as an Earth year with the definition of a solar day as a diurnal rotation period. With such an inclination and an obliquity of zero degrees the entire surface of the exoplanet can in theory be observed.
In earlier attempts to map the surface of Earth-like planets, Fujii et al. (2011, 2010); Kawahara and Fujii (2011) considered an orbital geometry in a face-on sense, i.e. $i = 0°$, with an obliquity of $90°$. This allowed them to continuously observe the entire planet during its full rotation about the parent star. However, in this orientation it would not be possible to observe the phase angle dependency of (polarized) flux, which proved to be a valuable tool in among other things: characterization of oceans (Williams and Gaidos 2008; Zugger et al. 2010*a*) and clouds (Bailey 2007; Karalidi et al. 2011; Rossi and Stam 2017). It has to be noted that in practice one can not recover reflected starlight at certain phase angles due to the lack of spatial separation of the exoplanet and its star at full and new phase

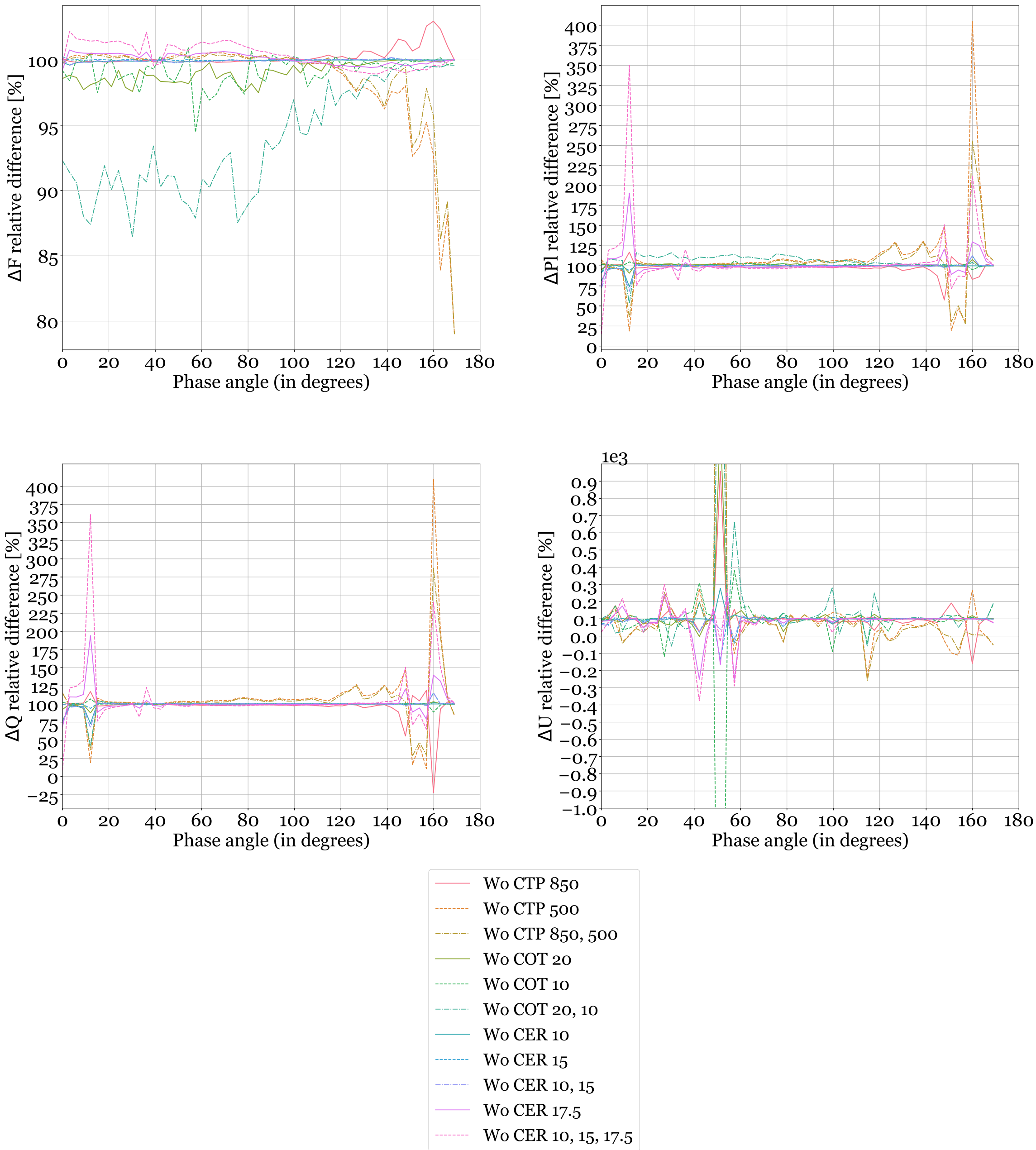


Figure 3.11: Sensitivity study of different bin parameters on our Earth-like models. The different cases are compared to a case which uses all bin parameters. For every run we use the same starting day: 1st of January 2011. In the upper left panel the relative difference in F is provided. In the upper right panel the relative difference in degree of polarization is provided, and the two normalized polarized fluxes Q and U are provided in the bottom left and right panel, respectively.

at least for $i = 90°$.

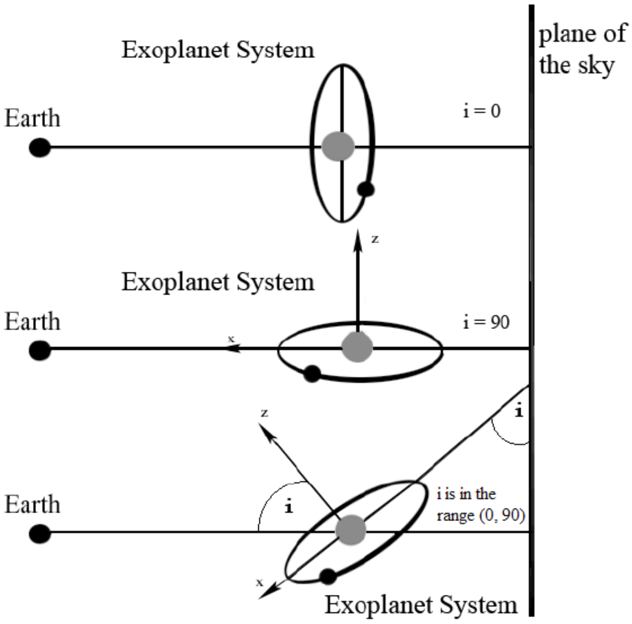


Figure 3.12: Interstellar representation of an exoplanet for different inclination angles i (Todorov 2008).

We have now described the basic geometry of our extrasolar system. However, it is interesting to model an Earth-like exoplanet at an Earth-like obliquity. Obliquity angles larger than that of Earth will not be considered, because higher obliquity angles violate the physical limit of the land distributions and cloud distributions, or in other words extremely deviate from our current climate. For example, Williams and Pollard (2003) simulated obliquity's between $0 - 85$ degrees by using a three-dimensional general-circulation climate model, showing that for obliquity's $\geqq 54°$ some seasonal ice and snow covers the equatorial regions, for which our land cover distribution would not be realistic anymore. Additionally, Williams and Pollard (2003) simulated Earth at an obliquity of $0°$. These simulations show that the seasonal changes are small, and the land coverage stays fairly constant. In one way this strengthens the choice of using a yearly constant land cover (*Section* 3.2.1), but weakens the use of daily cloud observations that include seasonal effects. The reversed logic is true for Earth-like obliquity angles. The effects for these low obliquity angles, however, do not show to have a major effect on the land cover distribution and cloud seasonality as compared to modeling obliquity's greater than Earth's (Williams and Pollard 2003).

As one can read from *Section* 3.3.1 is that only continuum wavelengths are considered as no absorption for as well the molecular atmosphere as the cloudy atmospheric layers is considered. As a result of the large computation times and required disk space we will have to limit the number of wavelengths on which we will perform computations. Ideally we want to cover the domain of the ultra-violet, visible and near-infrared wavelengths as these show significant alterations in the strength of Rayleigh scattering and Mie scattering. Also, we want to avoid absorption wavelength bands for the gaseous atmosphere and liquid water particles. Bogumil et al. (2003); Lacis and Hansen (1974) and King et al. (1990); Stephens and Tsay (1990) provide a parameterization of spectral absorption by the gaseous atmosphere and the cloud particles, respectively. The following wavelengths will be used: $350, 443, 500, 550, 670, 750$ and 865 nm, where we mainly focus on providing results for $350, 443, 550, 670$ and $865 nm$ due to the high computation times. At 350 nm we will be able to produce simulations where Rayleigh scattering is strongest and no Ozone absorption is present. At $865nm$ we simulate our outer boundary in the near-infrared where Rayleigh scattering is very ineffective. Regarding the vegetation green-bump and the red-edge we will use 550 and $750, 865$ nm, respectively. Additionally, the trough in the vegetation spectrum is accounted for with 670 nm.

4

The scattered light curves from an Earth-like exoplanet

In this chapter, we present simulations of disk-resolved and disk integrated planetary phase curves of our Earth-like planetary model. First, in order to investigate how the different cloudy and clear pixel models contribute to a disk integrated signal we will investigate the locally reflected light and present the results as spatially resolved planetary disks (*Section* 4.1). We compare the (polarized) flux and degree of polarization at different wavelengths and various orbital geometries to provide a comprehensive overview of the possible features in an Earth-like planetary disk. In *Section* 4.2 we will present the planetary phase curves of our planetary model. A more thorough analysis on how these phase curves are built up from different components of our model, such as the pure gaseous atmosphere, cloud layers and surface reflection, is provided. An Earth-like phenomenon that everyone encounters in their day-to-day life are weather changes and seasons. We discuss the variability that is induced on the Earth-like phase curves due to the seasonality in the observations. In the second to last part of this section, we provide the wavelength dependent Earth-like phase curves. In the last section, the diurnal variations are addressed. That is, horizontal inhomogeneities on the planet, in combination with the rotation of Earth around its axis, induce major oscillations in the reflected (polarized) flux and degree of polarization.

4.1. Resolved planetary disks

This section acts as an introduction to the reflection of light from an Earth-like exoplanet that is modeled as a horizontally inhomogeneous disk. That is, we will show that depending on the local properties and position of a pixel the reflection behaviour can be very different. Moreover, the reflection of the pixels also depend on the orientation of the exoplanet-star-observer system and the wavelength considered. The disk-resolved cases are simulated at phase angles $0°, 40°, 90°$ and $135°$ for $\lambda = 350,\ 550$ and $865\ nm$. We present F, Q, Pl and U for every wavelength in *Figure* 4.2, 4.3 and 4.4. Also, we present the associated land cover, cloud top pressure, cloud optical thickness, cloud particle effective radius and cloud fraction distributions on the corresponding planetary disks in *Figure* 4.1. One can observe that the parameters in the disks attain only a specific set of values, being in accordance to the discretization in *Chapter* 3. Thus, we use the planetary model described in *Chapter* 3: surfaces with Lambertian depolarizing reflection, variable cloud layers that are described using the complete set of bins and a constant gaseous atmosphere. All disks are simulated with 100×100 pixels at the same sub-observer longitude. Our findings for F at all phases and wavelengths are itemized as follows (cf. *Figure* 4.2, 4.3 and 4.4):

- At $350\ nm$ the features in the disks of F are dominated by the cloud fraction and cloud optical thickness, which are directly related to each other. Because of the dominant Rayleigh scattering at this wavelength, we see that only pixels with high cloud optical thickness exhibit high values of reflection. The area of low reflection at the Saharan Desert/Atlantic Ocean is caused by (1) the weak reflecting Lambertian desert and ocean surfaces at $350\ nm$ and (2) the low cloud fraction

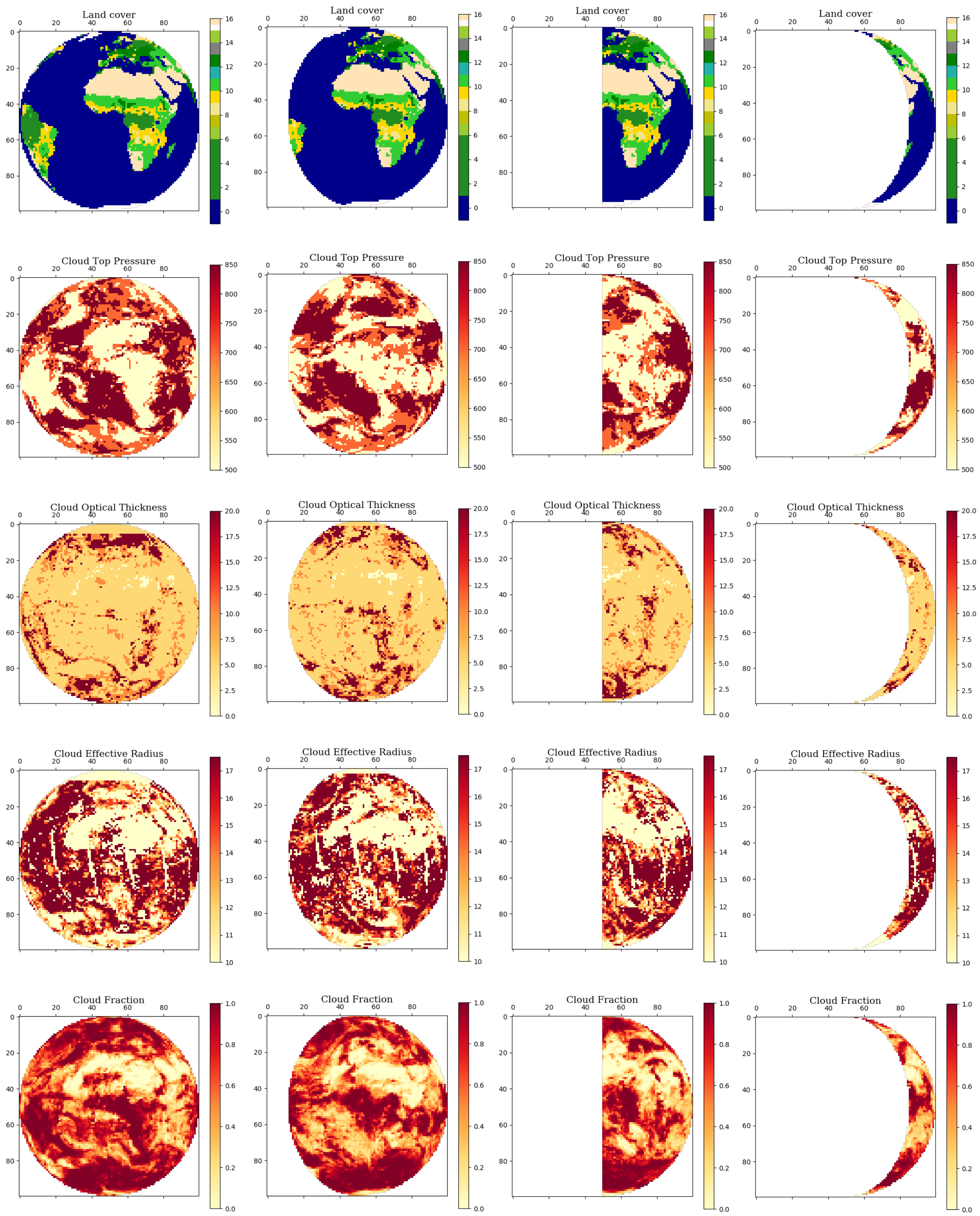


Figure 4.1: From top to bottom we present the land cover, cloud top pressure, cloud optical thickness, cloud particle effective radius and cloud fraction distributions are provided for $\alpha = 0,\ 40,\ 90,\ 135°$ from left to right, respectively. All disks are simulated with 100×100 pixels.

in this region. Albeit, Rayleigh scattering is highly efficient, we are still able to see the effect from the surface. For larger phases the effect of surface reflection is not apparent and the distribution of all cloud types, i.e. the cloud fraction, seem to agree well with the anomalies in the disk.

- At $550\ nm$ we start to recognize the spatial distribution of the clouds and the surface clearly. More specifically, the regions of large cloud fraction and cloud optical thickness correspond to high reflection, but one can also vaguely see the contribution of the bright vegetation[1] and desert as compared to the dark ocean. For higher phases we remain to see these features.
- At $865\ nm$ the spatial distribution of the continents are clearly visible. This is caused by (1) the increasing albedo of the desert and vegetation at the near-infrared[2] (*Figure* 3.2) and (2) the light beams are less effectively scattered by gas in the atmosphere (Rayleigh scattering) allowing them to penetrate through the atmosphere and getting reflected on the surface and clouds. Hence, we also recognize the contribution of the clouds, especially those on top of the oceanic regions. More specifically, highly reflecting cloudy pixels correspond to high values of optical thickness. For $\alpha = 40$ and $90°$ the dominance of the surface and cloud distribution is still well observable. At crescent phases the presence of Saudi-Arabia is still visible, but only barely, where only the dominance of the clouds is clearly visible.
- For crescent phases at $350\ nm$ the reflection looks homogeneous. These light beams a more effectively scattered in the gaseous atmosphere than those of longer wavelengths. Hence, a much smaller fraction of the light beams are able to reflect from the cloud layers and/or surface. At longer wavelengths the distribution of the clouds becomes more apparent, because the penetration depth of light beams penetrating in the gas layer on top of the clouds is much longer as Rayleigh scattering becomes much less effective, allowing more light beams to scatter on the highly reflecting cloud layer.

Our findings for the polarized flux Q at all phases and wavelength are itemized as follows:

- At $350\ nm$ and full phase the disk closely resembles the reflection of a homogeneous disk, like the disk in *Figure* 2.4, which is caused by the constant gaseous atmosphere that effectively scatters light beams by Rayleigh scattering. The anomalies in the disk correspond fairly well to the spatial distribution of optically thick clouds. At $\alpha = 40°$, we can not only recognize optically thick clouds, but also the vertical position of the clouds. The pixels with optically thick clouds cause a decrease in the state of polarization, because in general clouds cause depolarization due light that is scattered multiple times in the cloud layer Stam (2008*a*). At quadrature and crescent phases the disk becomes increasingly homogeneous-like, because of the optically thick gas layer on top of the clouds and surface.
- At $550\ nm$, Q shows a far weaker homogeneous-like disk as Rayleigh scattering becomes less effective, essentially increasing the contribution of the clouds. For Q at full phase the distribution of different particle effective radii show correspondence to reflection pattern in Q. For bigger phases angles the distribution of the entire cloud cover becomes more apparent, again most dominantly at $40°$.
- At long ($865\ nm$) wavelength the clouds dominate the polarized light Q at especially full phase and $40°$, where large particle effective radii are correlated to high polarized fluxes. When moving to crescent phases we can see some correlation with the cloud top pressure and cloud fraction. At quadrature, the relative difference in reflection between regions of different cloud properties and fraction is very small. This is caused by the fact that at a scattering angle of $90°$ the spherical liquid particles induce low polarization on the scattered light beams as compared to $\alpha = 40°$ (see *Figure* 1b of Stam 2008*a*).
- At full phase and $\alpha = 40°$ we saw that the contribution of the clouds to the polarized signal was apparent, where in general clouds are depolarizing because of the high degree of multiple scattering. The disks at full phase are not modeled exactly at $\alpha = 0°$, but at $\alpha = 3°$. If one again

[1] In *Figure* 3.2 one can see that at $550\ nm$ vegetation exhibits an increase in reflection, namely the green bump.

[2] For vegetation this is also known as the red-edge, i.e. an enhanced reflection caused by the presence of chlorophyll contained in vegetation (Horler et al. 1983).

consults *Figure* 1 of Stam (2008*a*) or *Figure* 3 of Bailey (2007) it is apparent that at $0 < \alpha <\sim 5°$ and α around $40°$ single scattering of spherical liquid particles exhibit high polarized intensities. These enhancements are generally known as the glory and the primary rainbow, respectively. In total intensity these phenomenon also cause increased reflection from single scattering (see Bailey 2007; Hansen and Travis 1974*a*; Karalidi, Stam and Hovenier 2012). In a realistic Earth-like atmosphere this enhancement is much smaller due to the occurrence of multiple scattering. Bailey (2007) states that polarization is generally suppressed by multiple scattering, so that the polarized intensity is dominated by single scattering from the top layers of the clouds. Consequently, the shape of the polarized intensity curve in a multiple scattering medium is very similar to that of the single scattering curve. In contrast, for the total intensity unpolarized light from multiple scattering is added to the total signal, essentially diluting the polarization and thus reducing the rainbow peak. In a later section we will illustrate this in more detail.

Our findings for the polarized flux U at all phases and wavelength are itemized as follows:

- At $350\ nm$ and $550\ nm$ a homogeneous like reflection behaviour of a gaseous atmosphere is prominent, comparable to *Figure* 2.5. Small anomalies are caused by optically thick clouds. At $550\ nm$ the clouds become more apparent at the edges of the disks, where the absolute magnitude of U is largest.
- At $865\ nm$ U again shows some homogeneous-like patterns, but with more anomalies induced by cloudy pixels and the depolarizing surfaces. At bigger phases angles we can hardly relate any features in the disks, but we see the largest agreement with the cloud fraction.
- Although in magnitude the resolved pixels of U show to be only one order of magnitude lower than Q the integration over the disk results in very small values of U, which will be explained in the next section. As described in *Section* 2.1 U, is defined as: $I(45°, 0) - I(135°, 0)$. For a homogeneous disk this results in four quadrants that are in magnitude symmetric over the plane of scattering, but in sign opposite. This essentially means that with small asymmetric[3] deviations (anomalies) from this homogeneity a non zero disk integrated polarized flux U can be measured. The larger the anomalies, the larger U becomes in magnitude. Although, U can also be zero if the anomalies in the disk are spatially symmetric with respect to the planetary scattering plane.

Our findings for Pl at all phases and wavelength are itemized as follows:

- Recall from *Section* 2.1 that the degree of polarization is defined as the polarized flux divided by the total flux. Since U is very small compared to Q, which was explained above, the degree of polarization is mainly determined by the ratio between F and Q. At $350\ nm$ F and Q were mainly affected by the cloudy pixels and more specifically the optically thick cloudy pixels. Consequently, also Pl shows to be mostly affected by the clouds. For short wavelength we thus recognize no effect from the surface cover.
- For Pl at $550\ nm$ we can clearly observe the land distribution of Africa and the clouds at $40°$ phase. Low values of polarization occur at the Saharan Desert and small parts of Southern Africa. Stam (2008*a*) already showed that Lambertian surface reflection with increasing surface albedo results in a lower degree of polarization. This is caused by the increased flux from the reflection by the surface, which is, in case of a Lambertian (i.e. non-polarizing) reflecting surface, completely unpolarized. An increased total flux compared to the polarized flux thus results in a low degree of polarization. At $90°$ high degree of polarization is caused by the clouds and, in particular, clouds that have large particle effective radii. At quadrature and crescent phases any presence of the clouds lower the degree of polarization, i.e. agreeing well to the cloud fraction distribution.
- Pl at $865\ nm$ shows regions of low magnitude that correlate to the African continent, Southern America, parts of Azia/Europe and even Antarctica. This is similar to what we found at $550\ nm$, but is stronger due to the increase in reflection of the vegetated and desert surface. At quadrature and crescent phases we can not clearly extract any contribution from the surface cover or clouds, although in F the contribution of the land cover was still significant. This is caused by the small variations in Q.

[3]More specifically, asymmetric around the planetary scattering plane.

In conclusion, we showed that the dominance of inhomogeneous surface and cloud distributions become more apparent for longer wavelength as Rayleigh scattering becomes much less effective, allowing the light beams to penetrate further into the atmosphere. The contribution of the clouds to Q is largest at $\alpha = 40^{\circ}$ (The location of the primary rainbow). U is dominated by Rayleigh scattering, but small anomalies exhibited by optically thick pixels are present. The depolarizing Lambertian surface approximation exhibits regions of low Pl that correspond fairly well to the land cover distribution of high reflecting surfaces. Moreover, the increased contribution from the (non-polarizing) Lambertian surfaces at long wavelengths is caused by the fact that the light beams are able to penetrate furthest through the gaseous atmosphere.

4.2. Disk integrated light curves

In this section, disk integrated planetary phase curves are presented to provide insight into the reflection behaviour of an Earth like exoplanet in an edge-on orbit around its star. If not stated otherwise, the planetary model for our simulations in this section is exactly the one used in *Section* 4.1, except for the disk size which is 20×20 pixels. In *Section* 4.2.1 we present the general form of the planetary phase curves. We also show the seasonal effect of the clouds, and the effect of the obliquity on the planetary phase. Then, we present the planetary phase curves at multiple wavelengths ranging from the ultra-violet to the near-infrared domain (*Section* 4.2.2). Lastly, in *Section* 4.2.3, we show what the rapid oscillations in the planetary phase curves look like at a diurnal time scale.

4.2.1. Dependence on phase angle

Figure 4.5 shows the phase curves of F, Q, U and P at $\lambda = 550nm$. The observation that is used at full phase is that of January 1st, 2011, and consecutive observation days for consecutive solar days for half a year. The temporal resolution is 2 hours. At a phase angle of 0°, the exoplanet and its parent star are in line with the observer. In practice, this means that one cannot measure reflected light as it is blocked by the star. At a phase angle of 180°, no light can be scattered from the planet to the observer as the planet is exactly in front of the star. For all phases many variations occur in F that are caused by the rotation of the planet in combination with the inhomogeneous surface and cloud cover. The absolute amplitude of the variations tend to decrease with increasing α as the continuum of F gradually tends to zero. That is, for increasing α the visible and illuminated part of the planetary disk becomes smaller, thus reflecting less light to the observer. Despite these daily variations we are still able to observe the enhancement in total intensity due to the spherical liquid water particles in our clouds. The glory, near full phase, seems less apparent, but near full phase we observe an enhancement in F. For Q the clearly visible glory produces a sign difference that is also retrieved in simulations by Bailey (2007); Karalidi, Stam and Hovenier (2012); Stam (2008*a*). When comparing the absolute magnitude to that of F this enhancement is small. The daily variations present in F show to be much less in Q, except for a small region of phases around the primary rainbow. This primary rainbow is much stronger than in F as was also suggested by Bailey (2007), caused by the depolarization of light due to multiple scattering. Also, we observe a second bump near $\alpha = 56^{\circ}$, corresponding to the secondary rainbow.

From Pl the primary rainbow at $\alpha = 40^{\circ}$ and glory near full phase are also clearly visible. The secondary rainbow is mostly suppressed by the daily variations. These daily variations are induced by F and are maximum in a large region around quadrature.

The variability due to Earth's rotation can be clearly seen in U. On a similar scale as F or even Q these variations, however, would be extremely small. Consequently, the contribution of U to Pl is negligible. Nevertheless we will analyze our findings for U for all our simulations as from a theoretical point of view it might show us information about the asymmetry in the planetary disk. The overall trend in U seems to oscillate around zero at low phase angles, shifts to slightly more positive values at higher phase, where after it decreases to zero at new phase. This behaviour might be due to some asymmetry between the Northern and southern Hemispheres of Earth, and will be studied more deeply in a following section.

The significance of the primary rainbow feature shows that measuring Pl or Q is a powerful tool in characterizing especially terrestrial atmospheres. It should be noted that the rainbow features exhibits

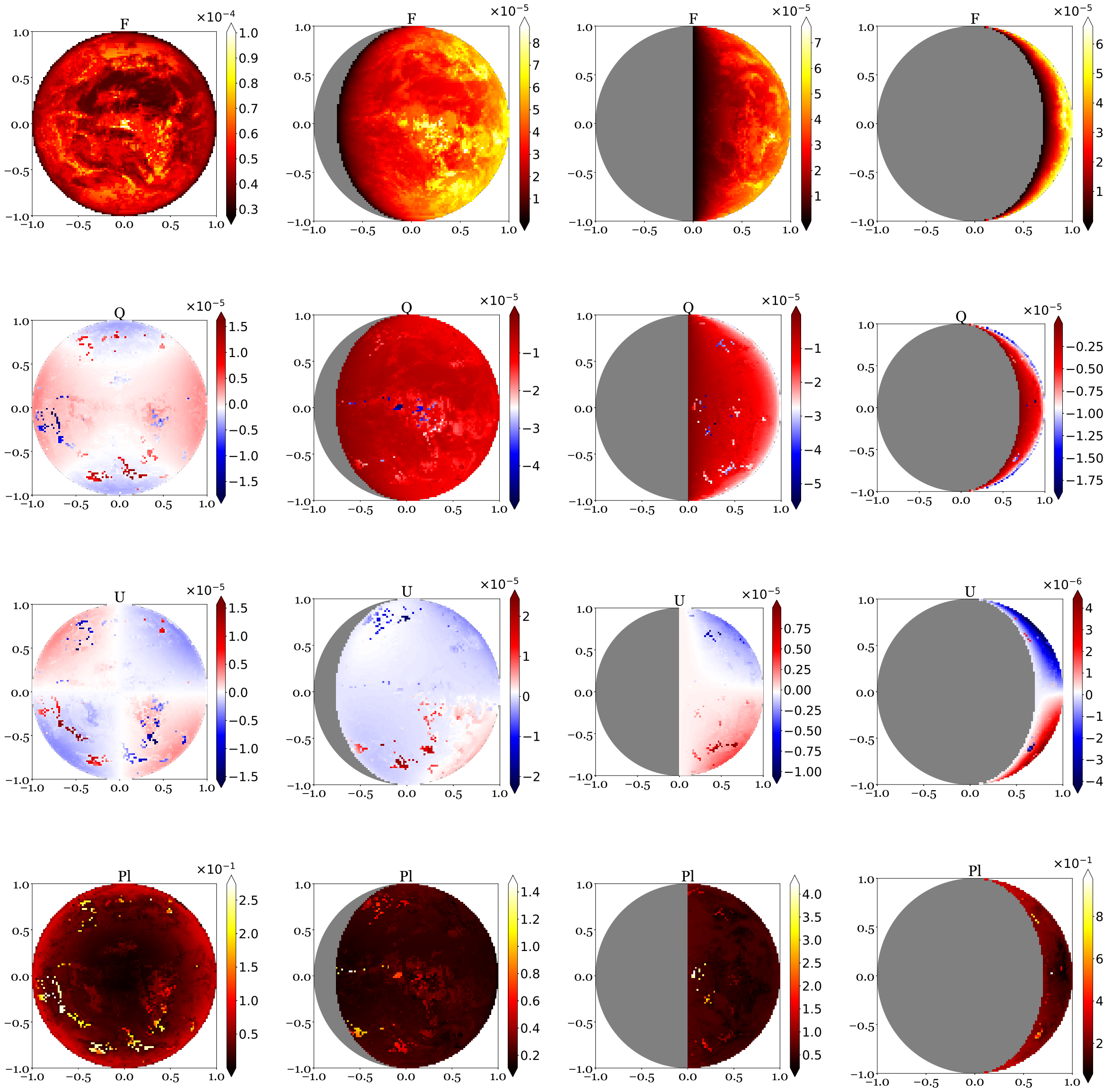


Figure 4.2: Resolved disk for $\lambda = 350\ nm$ at $\alpha = 0°, 40°, 90°$ and $135°$. In the upper four panels the Stokes elements F, Q and U, and Pl are provided, respectively. The corresponding disk properties are provided in *Figure* 4.1. All disks are simulated with 100×100 pixels.

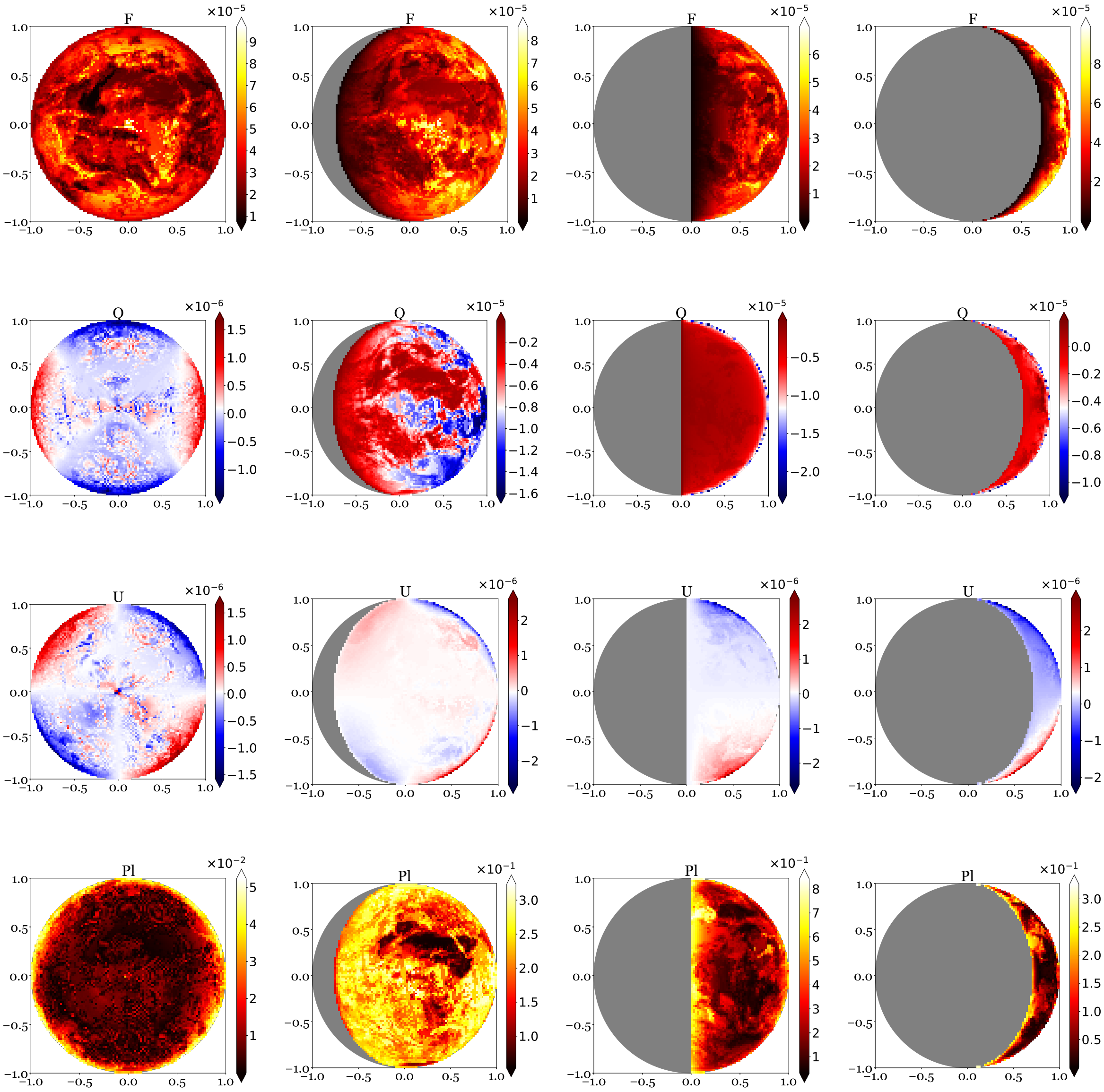


Figure 4.3: Similar to *Figure* 4.2, except for λ = 550 nm.

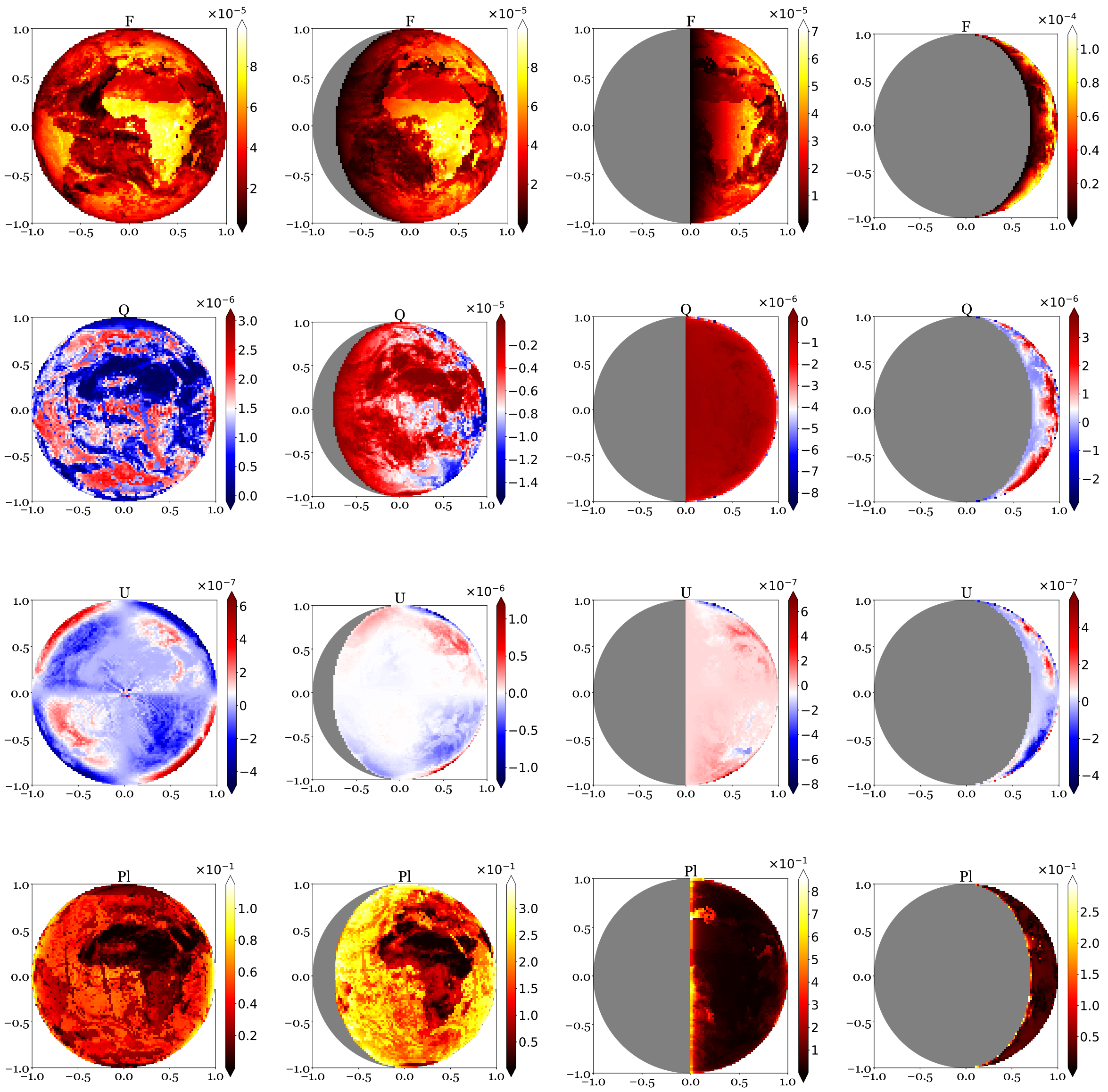


Figure 4.4: Similar to *Figure* 4.2, except for $\lambda = 865\ nm$.

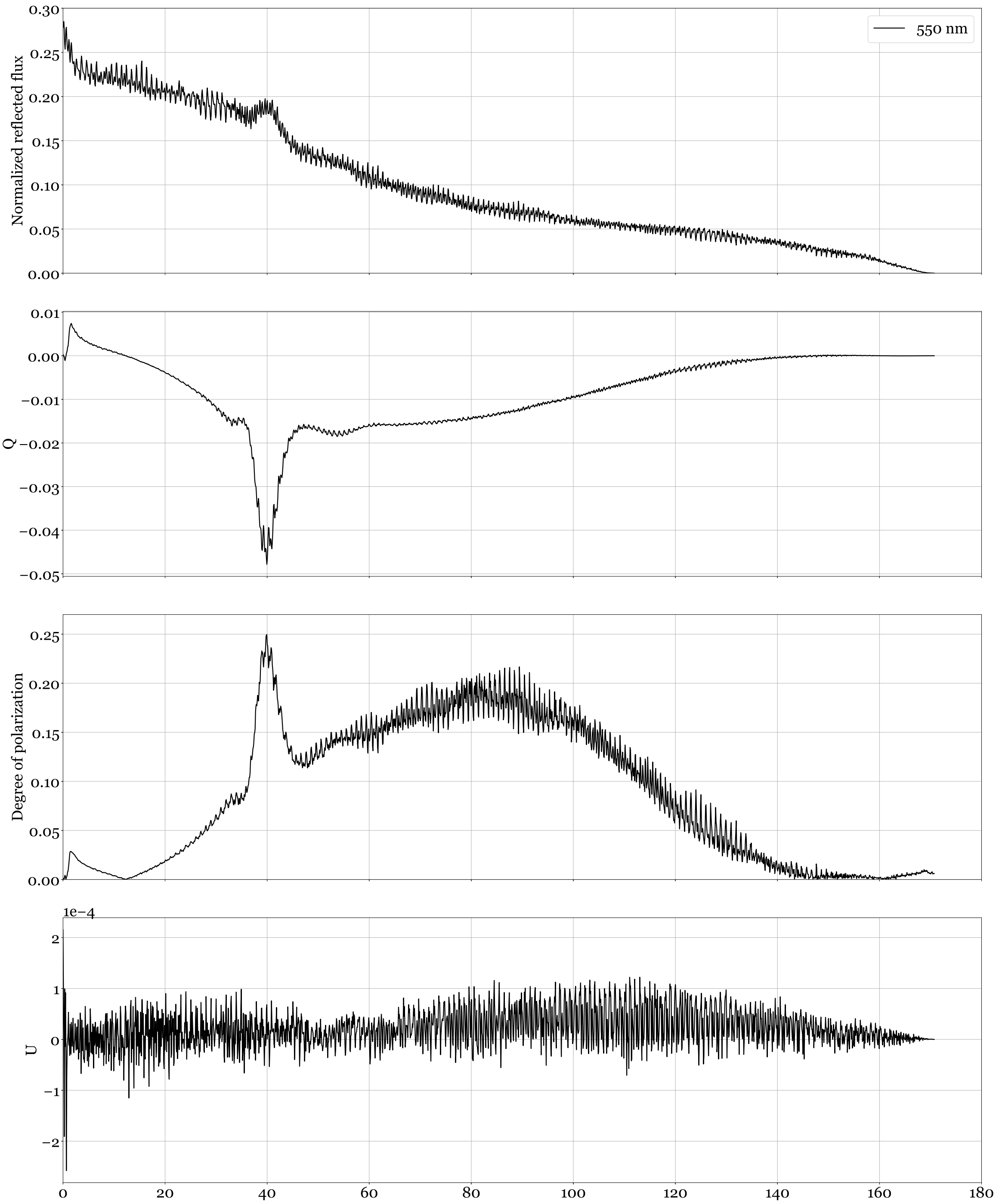


Figure 4.5: Disk integrated planetary phase curves from an Earth-like exoplanet using MODIS data from 2011 at a wavelength of $550nm$. From top to bottom: F, Pl, Q and U as functions of the phase angle.

a major enhancement in our signal on the assumption that our cloud particles are spherical. A slight deviation from this sphericity has been shown to strongly affect the strength of these rainbow features (Bailey 2007). More specifically, they modeled a prolate spheroid with axis ratio[4] of 0.8, an oblate spheroid with axis ratio of 1.2 and a cylindrical particle with length equal to its diameter simulated by using the T-matrix method (see Waterman 1971) to calculate the scattering properties of a size distribution of randomly oriented axially symmetric particles. In their comparison they used the same size distribution as those for spherical particles: effective radius of $5\ \mu m$ and effective variance 0.1. For all slightly non-spherical particles the peak of a clear rainbow peak is lost.

Decomposition of Earth-like phase curve

In order to show the effect of different components of our planetary model, we computed the phase curves for four different end cases: the planetary model with a homogeneous black surface and no cloud layers (pure Rayleigh scattering of the gas), the planetary model without cloud layers but with an Earth-like surface distribution (No Earth clouds), the planetary model with a black homogeneous surface but with Earth-like distributed cloud layers (Black surface Earth), and lastly the complete planetary model (Earth-like) as we use it in general (*Figure* 4.6). The cases are computed at a wavelength of $550\ nm$. For this simulation we used a temporal resolution of 2 hours over half a year of observations from 2011, where full phase corresponds to January 1st 2011.

The Rayleigh scattering curves for F and Pl show a very similar shape with plots in Stam (2008*a*). From a theoretical point of view the polarized flux Q is zero at full and new phase as a result of the homogeneity of the disk. U is zero for all phases owing to the fact that the planet is homogeneous. The maximum Pl is located at $\alpha =\sim 90^\circ$. The maximum polarized flux Q is located near $\alpha =\sim 70^\circ$. The slight asymmetry in the curve of Pl is caused by the low occurrences of multiple scattering in the pure gaseous atmosphere.

If we now add an Earth-like surface major variations occur in F and Pl. Also, the peak of polarization in Pl moves to larger phases. Similar behaviour can also be observed in results from Stam (2008*a*). For all phases Pl decreases as a result of the overall increase in F. The increase in F is attributed to the increase of surface albedo over the entire planetary disk. When we look at the behaviour of Q there is virtually no difference with respect to the Rayleigh curve, because the Lambertian reflecting surface does not add polarized light. The small oscillations that are present are caused by the fact that the generally unpolarized light that gets reflected by the surface can get polarized in the atmosphere while traveling to space. For U we retrieve similar extremely small oscillations. In conclusion, the Lambertian depolarizing surface alone in a gaseous atmosphere has no significant effect on the linearly polarized fluxes.

A major effect on the reflected signals is induced by the addition of an Earth-like cloud cover (The Earth-like case). Due to the high occurrence of (multiple) scattering in the clouds, more light is reflected from the exoplanet causing a significant increase in F. On the other hand, multiple scattering depolarizes light, decreasing Pl significantly. In the addition of liquid water clouds we can also see the primary rainbow clearly and even the secondary rainbow. Furthermore, the daily variations are affected in amplitude and become less smooth, potentially affecting the periodicity of the constant surface cover. For Q the clouds only significantly induced a higher polarized reflection near the two rainbows and the glory, whereas for other phases the effect is minimal. This is in accordance with what we expect in theory as the clouds mostly induce low amounts of polarization on reflected light except near these known regions. At the primary rainbow daily variations of the clouds are visible, whereas for other phases the variability in the signal remains very low. We state the following hypothesis about adding polarized surfaces:

> *When including a horizontal inhomogeneous planetary surface (i.e. continents and oceanic regions) that polarize reflected light beams, the smooth curve of Q will exhibit major variability, while the planet rotates around its rotation axis.*

This variability will, however, not be solely attributed to the spatial inhomogeneity of the surface cover, because (1) polarized reflected light from the surface gets depolarized when a sufficiently thick cloud is

[4]Ratio between minor and major axis.

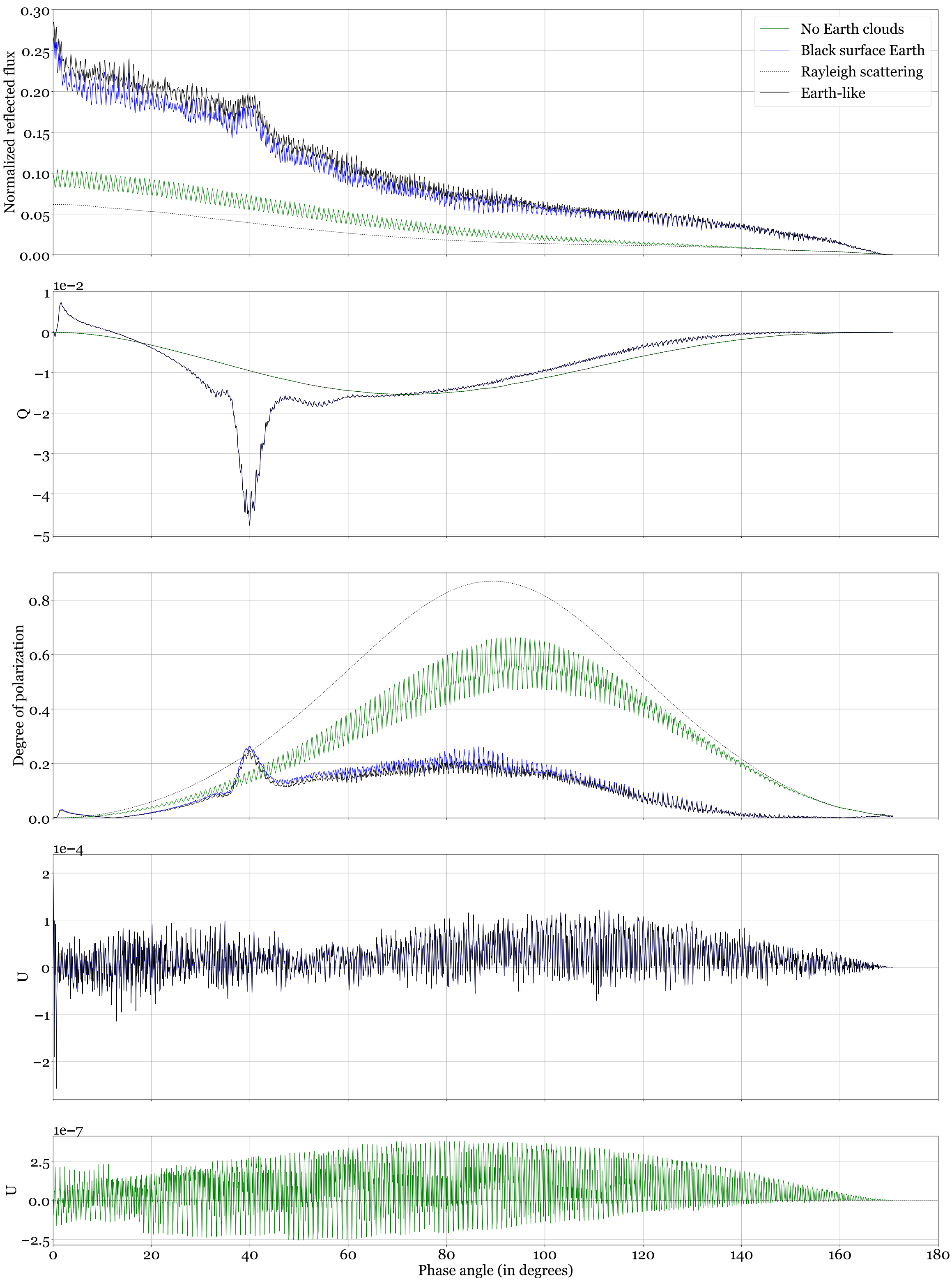


Figure 4.6: Disk integrated planetary phase curves computed at $\lambda = 550\ nm$. From top to bottom we provide F, Q, U and Pl. All subplots show four different cases: the planetary model (Earth-like), the planetary model without cloud layers and a homogeneous black surface (Rayleigh scattering), the planetary model with a black homogeneous surface (Black surface Earth), and the planetary model without cloud layers (No Earth clouds).

overcast, where the presence of such a cloud can vary daily, and (2) the vertical variability in the clouds induce changes in the vertical extend of the top pure gas layer, yielding variability in the strength of polarized reflected light. Lastly, from *Figure* 4.6 we observe a significant increase in the strength of U, where the clouds induce large oscillations.

To further discriminate the effect of the surface we applied a black surface to our Earth-like atmosphere (Black surface Earth). When we omit the surface reflection, we see an overall decrease in F. Similar to the "No Earth clouds" end case the fully absorbing surface causes a decrease in overall brightness of the planetary disk. The primary rainbow shows to be relatively untouched, although there seems to occur more fluctuations around $\sim 50°$ phase. A similar behaviour is observed for Pl at the edge of the primary rainbow. Overall Pl increases, because less unpolarized light is scattered back into space. For Q and U the absence of a reflecting surface has negligible effect. For U small increases occur in amplitude. The extremely small effect on Q shows again that the assumption of Lambertian surfaces has a negligible contribution to the polarized reflection. This strengthens the hypothesis that was introduced in the former paragraph. We will further investigate this hypothesis in *Chapter* 6.

Phase curves for cloud end cases

In the former section we showed that the clouds induce major oscillations in the planetary phase curves, but what is the effect of the different cloud parameters? To investigate the effect of a single cloud type as function of phase angle we will replace any type of cloud that we retrieve from the observations by one specific cloud type, e.g. a cloud with $P_c = 700\ mb$, $\tau = 10$ and $r_{eff} = 12.5\ \mu m$. These end cases allow us to present the pure effect of each parameter in an Earth-like cloud cover distribution. The cloud fraction for every pixel is not altered and we use the Earth-like surface distribution and gaseous atmosphere from our standard planetary model.
The first day of observations again corresponds to the first of January 2011. The different cloud types are labeled in the lower panel of *Figure* 4.7, where CTP resembles the cloud top pressure, COT the cloud optical thickness and CER the cloud particle effective radius.

Figure 4.7 shows the results of these end case simulations. From inspection of the upper panel we retrieve that the dominant factor in F is the cloud optical thickness. A significant difference in the reflection of F is observed until approximately a phase angle of $120 - 130°$. The highest reflection occurs for optically thick clouds, as less light is transmitted through the clouds. Furthermore, we observe slight differences in the cloud particle effective radii and the cloud top pressure. An increase in cloud particle effective radius relates to a decrease in F, agreeing to the results from *Section* 2.3.2 for the single pixel TOA reflection. For increasing cloud top pressures F increases, due to the thicker gas layer above the clouds that reflect more light and decreases the penetration depth into the atmosphere, so that less light reaches the surface and getting partly absorbing.

The second panel shows the dependence of Pl on the end cases. At phases higher than the primary rainbow and around crescent phases, the effect of the cloud optical thickness and cloud top pressure on the degree of polarization are most prominent, where the cloud optical thickness exhibits the largest differences. We also see a small effect of the cloud particle effective radius in this range of phase angles. So, it seems that high Pl occur in the presence of a large abundance of optically thin clouds in the planetary disk. Also, for increasing P_c Pl increases. By increasing P_c the column of gas on top of the cloud layer increases, essentially allowing more light to be polarized by Rayleigh scattering. At the primary rainbow peak we see an interesting behaviour regarding the cloud particle effective radius. Although we still observe the dominant effect of the cloud optical thickness, higher values of r_{eff} yield higher Pl. Bailey (2007) modeled the strength of the rainbow peak for $r_{eff} = 1-100\ \mu m$ with $v_{eff} = 0.1$. They also found that the strength of the primary rainbow peak in polarized flux increases with increasing r_{eff}. A similar behaviour for the secondary rainbow can not be observed.

To be able to retrieve the effect of the Earth-like clouds solely on the polarized fluxes we also computed and plotted the effect on Q and U. This effect can be particularly interesting, because we saw in *Figure* 4.6 that the Earth-like clouds seem to be only significantly effective on Q near the primary rainbow feature. In the third panel we indeed retrieve that the different cloud types induce significant varia-

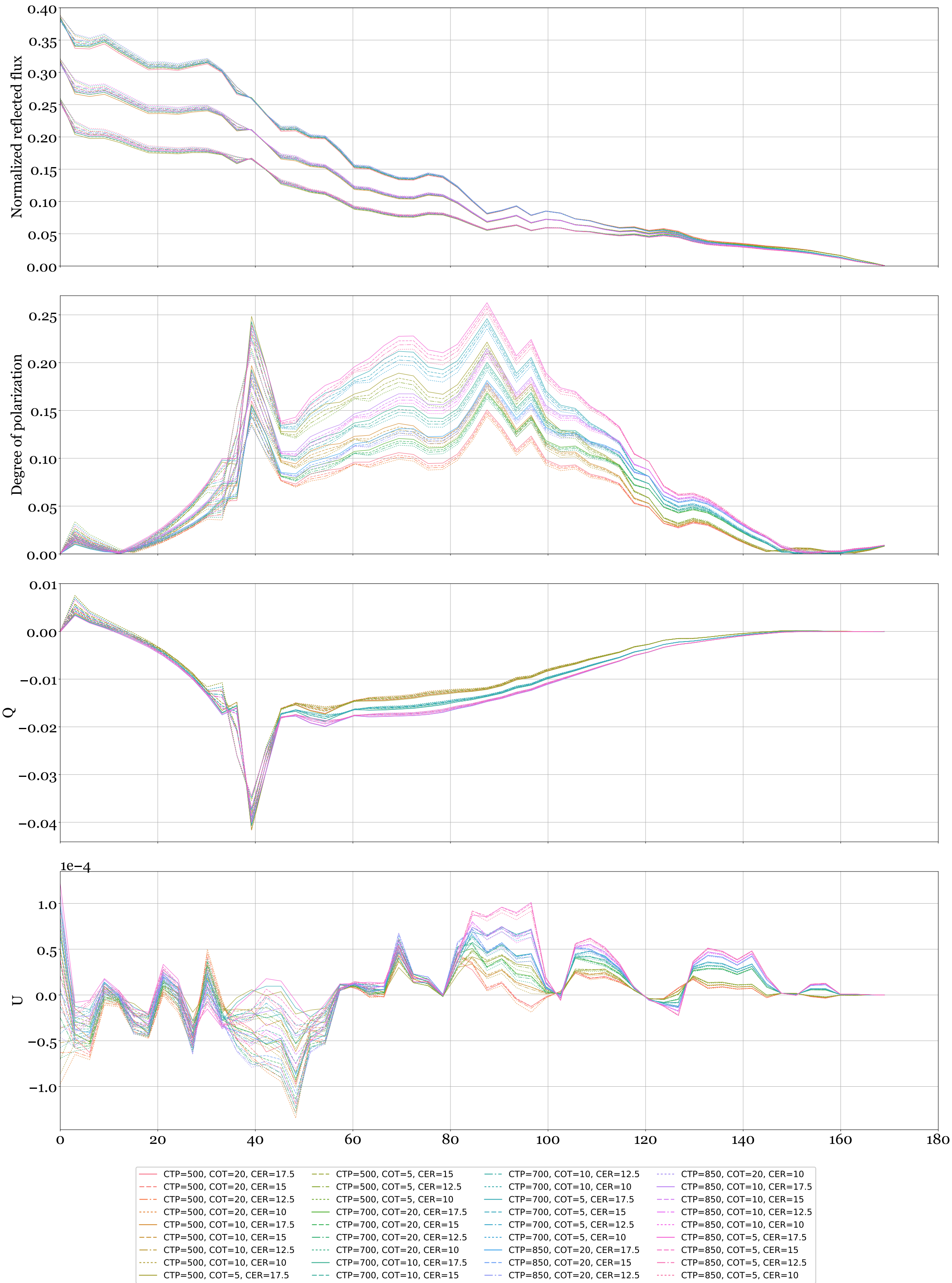


Figure 4.7: Multiple phase curves with different cloud end cases. The temporal resolution is 3.059 days, and correspond to the 1st of January 2011 at full phase. The upper panel shows the F, the second panel Pl, and the third and lower panels show the polarized fluxes Q and U, respectively. The acronyms in the legend are: CTP is cloud top pressure, COT is cloud optical thickness and CER is cloud particle effective radius.

tions near the primary rainbow, but also at quadrature. For phases near the quadrature the dominant parameter is the cloud top pressure. In essence this is caused by the gaseous atmosphere on top of the clouds, allowing more or less light to be scattered by the pure gaseous top layer. At the primary rainbow, the dominating parameter is the cloud particle effective radius, where the cloud optical thickness shows almost no dominance. The dominance of the cloud optical thickness on the rainbow in the degree of polarization is thus mostly induced by F(again in agreement with Bailey (2007)).

The lower panel shows the Stokes vector U. Similarly to *Figures* 4.5 we observe for all cases large fluctuations, but for small absolute values of U. Near full phase and before the rainbow phase angle a small dominance of cloud optical thickness and cloud top pressure is observed. This trendy behaviour seems to be somewhat distorted near the primary rainbow, where after U shifts to positive values. At quadrature different cloud top pressures and optical thicknesses produce clear distinctive trends. At higher phase angles this can be seen even more clearly agreeing well to what we found for the disk-resolved simulations in *Section* 4.1.

At different wavelength, the behaviour of the previously presented results will be more pronounced or less so, for example in the blue it is expected that the clouds are far less dominant and we wouldn't retrieve the dominance of the cloud optical thickness. The cloud top pressure might show the largest dominance at this wavelength region as Rayleigh scattering is most efficient in the blue. Provided that significant light can still reach the top of the clouds. At red wavelengths, Rayleigh scattering is less effective and reflection by the clouds is more dominant. We then expect the cloud optical thickness and particle effective radius to show clear distinctive trends in our signals. In *Section* 5.2 we will show whether any of the cloud parameters from our Earth-like planetary model is correlated to the reflection of the Stokes parameters.

Effect of Earth's seasonality and temporal sampling

For the previously presented phase curves, full phase always corresponded with the 1st of January 2011 and consecutive solar days with consecutive observation dates. In this section, we show the variability in the phase curves if we simulate different starting days. This type of analysis will enable us to retrieve how the seasonal variability in the Earth observations affects F, Pl, Q and U. In *Chapter* 3 it was provided that our planetary model requires quite some pixel models for every time step already for a planetary disk of 20×20 pixels. To save time, we will therefore not compute the phase curves for all 365 possible starting days. We provide two simulations with slightly different temporal resolutions: 3.059 and 3.0 days. For these two resolutions the longitude/phase relations are provided in *Figure* 4.8. With a resolution of 3.0 days we basically simulate an Earth-like exoplanet that is artificially phase locked to its parent star as seen from the observer. With a resolution of slightly more than three days we simulate the Earth-like planet to artificially rotate four times in half an orbit as seen from the observer. This will also give us a general idea of how different types of sampling can effect the variety in our signals and the retrieval of Earth signatures. The observation date at full phase is shifted every 14 days over the full year of 2011, resulting in 53 different runs.

In *Figure* 4.9 the planetary phase curves of with a resolution of 3.0 and 3.095 days are shown. The solid lines corresponds to the mean and the shaded areas corresponds to the minimum and maximum values of the 53 different simulations. The surface cover and gaseous atmosphere are invariant in time, so the seasonal variations are solely due to the variability in cloud observations.

In the upper left panel we provide F. Similarly to *Figure* 4.7 much variations occur around $\alpha = 5 - 35^{\circ}$ and decreases with for larger phase angles. The shape of the primary rainbow is maintained for the 3 day resolution, but not for 3.095 day resolution. Overall we see that by artificially rotation Earth in our observations the continuum is less smooth continuum. For both simulations the secondary rainbow is not clearly present. In the upper right panel, Pl shows to be relatively insensitive to the changing cloud cover at low phase angles and at the primary rainbow feature for both resolutions. Albeit, the magnitude of the primary rainbow peak varies with the temporal resolution, being caused by the daily variations at this peak, which was shows in *Figure* 4.5. A bump, corresponding to the secondary rainbow, can be observed, but the amplitude does not exceed that of the seasonal variability, which makes it hard to retrieve it unambiguously. For higher phases, near quadrature and crescent phases much more variations are induced from the seasonality in the cloud observations. For Q we only see a slight variation near the peak of the primary rainbow. As compared to the middle left panel of *Figure* 4.7, we see a major difference in variability at phases other than the rainbow: very low variability induced by the dynamic Earth-like seasonal cloud cover at especially quadrature and crescent phases, showing again that Q is particularly insensitive to the clouds. Also, we see that changing the sub-observer longitude also has minimal effect on Q. Albeit, this may not be the case by introducing polarizing surface models. For U we see a lot of variation for the entire phase angle region, but no pattern similar to that in the middle right panel of *Figure* 4.7. We do retrieve that for other sub-observer longitude U seems to consistently attain higher positive values. This can be caused by (1) the different continental distribution that is not facing the observer for an artificially phase locked Earth or (2) the presence of a substantially different cloud cover at other sub-observer longitude.

From both simulations we conclude that Q is the least sensitive to the seasonally changing cloud cover, except near the primary rainbow. The amplitude of the variation remains the same for different types of sampling, and glimpses of the primary rainbow in F depend on the sampling. For both simulations U still induces the most complex behaviour, but is also very small. Lastly, it may be noted that all the seasonal variations are in the order of relative magnitude of the diurnal variations, so it would be very difficult to identify seasonality from such curves. Also, we have seen that with a temporal resolution of ~ 3 days we are still able to observe the gross shape of the curves including important features as the rainbow and the glory for Pl and Q. However, the exact width and amplitude of this primary rainbow may not be retrieved properly.

Effect of different orbital geometries

In this research we mainly model our Earth-like planet with an obliquity of zero. Here, we will show results of simulations with different values of the obliquity: $-23.4, -15.4, -7.8, 7.8, 15.4$ and 23.4 degrees. In *Figure* 4.10, the angle of obliquity is directed away (negative) or towards (positive) the observer for all phases. In *Figure* 4.11, the angle of obliquity is directed to the left (negative) or the right (positive) of the observer for all phases. Due to computation times we only present these result for $\lambda = 550\ nm$. In the upper left, upper right, middle left and middle right panels of *Figure* 4.10 we provide F, Pl, Q and U, respectively. Every case includes 27 runs to incorporate the effect of Earth's seasonality as explained in *Section* 4.2.1, with a temporal sampling of 3.059 days (*Figure* 4.8).

By rotating the north pole of an Earth-like exoplanet away from the observer (see the lower right panel of *Figure* 4.10, F increases due to higher reflection from the increased visible region of the

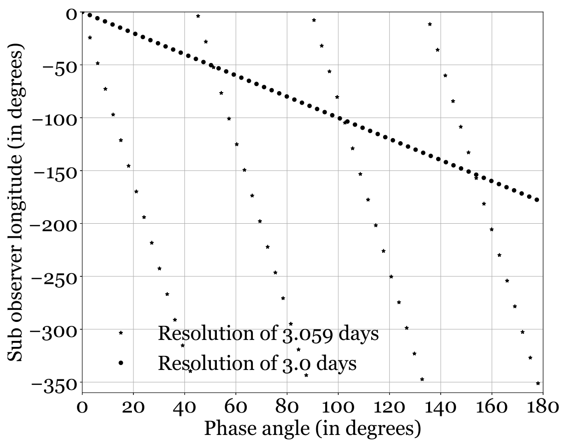


Figure 4.8: Two different observing geometries with different temporal sampling. The stars denote the phase angle and sub observer longitude of all data points with a temporal sampling of 3.059 days. The dots denote the phase and sub observer longitude for all data points with a temporal resolution of 3.0 days.

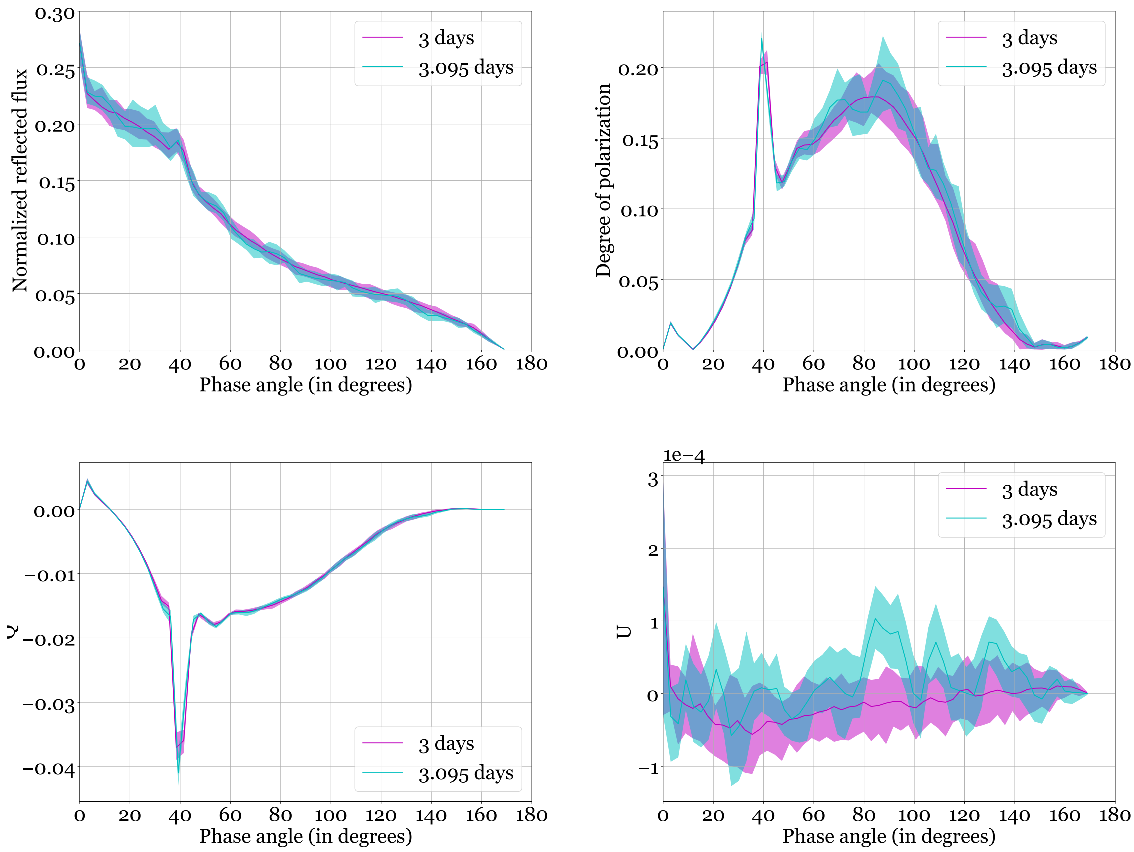


Figure 4.9: From top to bottom: F, Pl, Q and U, all as functions of the phase angle, at $\lambda = 550\ nm$. The shaded areas show the minimum and maximum values and the solid lines the mean values. The plots are build up of 53 runs with different starting days. The temporal resolution of each run is 3.0 days.

Antarctica continent or due to an annually present cloud deck on the southern mid latitudes. This relative increase in F becomes smaller with increasing phase angle as is illustrated in the upper left panel of *Figure* 4.10. With increasing phase angle the illuminated and visible region of Earth becomes smaller, but more importantly the poles becomes less visible. For Pl we see the same effect for large phase angles. For phase angles near quadrature negative obliquity angles (away from the observer) yield lower values of Pl. At the primary rainbow and smaller phases we do not see a significant difference between the different values of obliquity. Also, for both F and Pl, we can see that negative values of obliquity show much more variations than the positive values. This may again be related to the ice caps on our poles. The ice abundance on the North Pole is much smaller than on the South Pole, essentially providing much less variation. If we look at the polarized flux Q, as illustrated in the middle left panel of *Figure* 4.10, we see that with positive obliquity values the averaged reflected light is less polarized, but the overall difference is very small. Because the ice surface reflection as well as the reflection from clouds is generally depolarizing, we can expect variations in Q to be minimal. For U, see the middle right panel of *Figure* 4.10, one can see that the different types of obliquity exhibit very different patterns, especially for the mean values of U. In the first half of the phase curve it seems that negative values of obliquity induce more negative values of U and that this behaviour is clearly reversed for the other half of the phase angle range, vice versa for positive values of obliquity.

In the upper left, upper right, middle left and middle right panels of *Figure* 4.11 one can see F, Pl, Q and U, respectively for the six different cases of obliquity, for the same planetary model. For F we observe that there is no significant difference between the different values of obliquity from full phase to approximately $\alpha = 60°$. For quadrature and crescent phases we see a consistently higher total reflection from negative obliquity's. If we look at the orientation of the land distributions in the lower panel one can see that for negative obliquity's and high phases the Antarctic ice sheet is more dominantly visible, whereas for negative obliquity's the North Pole is more dominantly visible. As mentioned before the abundance of ice, and thus the abundance of highly reflecting surfaces, is significantly higher for the Antarctic continent, essentially inducing the increase in reflection. For Q we again see almost no effect from the different orientations. So, any variation from the six cases in Pl are induced by F. Hence, the continuous lower degree of polarization for negative obliquity's at quadrature and crescent phases. For U we see a clear distinction between negative and positive obliquity values for phases until $\alpha =\sim 80°$, where negative obliquity's seem to induce negative values of U and vice versa for positive obliquity's. This clear division might be induced by the fact that with relative large visible and illuminated regions of the disk, for negative obliquity's, a higher portion of the continents is located on the norther hemisphere, whereas for positive values a larger portion is located on the southern hemisphere. At high phases the different obliquity values can not be clearly distinguished.

In conclusion, different orientations and angles of the Earth's obliquity influence especially F, Pl and U. The polarized flux Q seems to be relatively insensitive even at the rainbow feature, essentially implying that the different cloud distributions on the disk do not show much differences. The main influence appears to come from the North and South Polar ice caps, whose spatial extent plays a roll in the total increase of reflection. The different distributions of the continents on the upper and lower part of the planetary disk have effect on U, at least for small to moderate phase angles.

4.2.2. Phase curves at multiple wavelengths

In the former section, we provided the main characteristics of an Earth-like polarimetric signal at $\lambda = 550\ nm$. In *Section* 3.2.1, we discussed the wavelength dependency of surface albedos. Here, we will show phase curves computed at wavelength: $\lambda = 350, 443, 550, 670$ and $865\ nm$. *Figure* 4.12 shows the phase curves F, Pl, Q and U. We simulate the planetary disk with 20×20 pixels. We use the planetary model described in *Chapter* 3: surfaces with Lambertian reflection, variable clouds that are described by using all bins and a constant gaseous atmosphere. Our findings for F are itemized as follows:

- The daily variations increase with increasing wavelength, due to the decreasing effectively of Rayleigh scattering, essentially exposing the spatially inhomogeneous daily varying clouds.
- For smaller phase angles, F attains high values at $350\ nm$ because (1) more light is reflected from the optically thick gaseous layer on top of the clouds and (2) if light is able to travel through

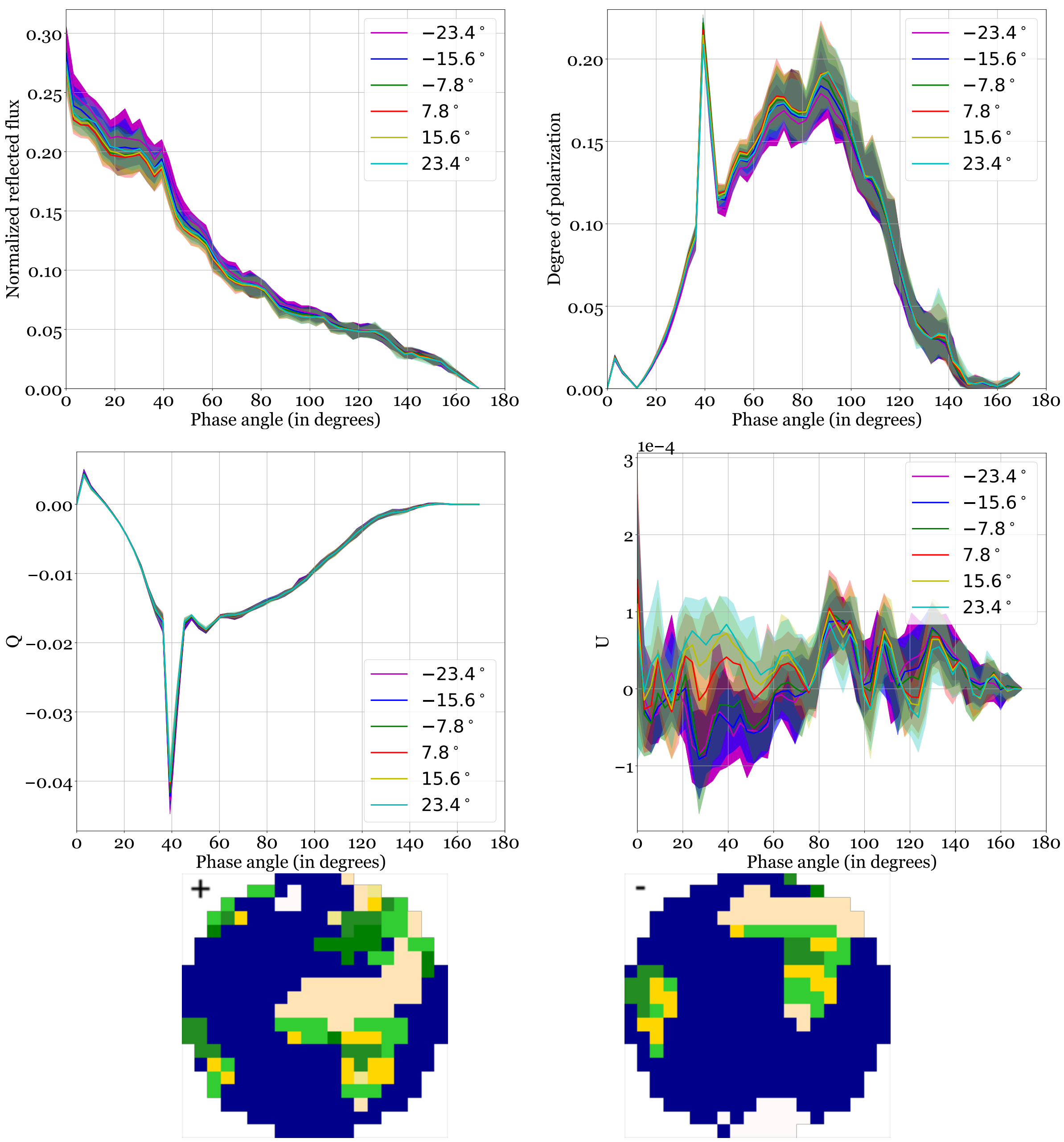


Figure 4.10: Six runs with different angles of obliquity oriented towards (positive) or away (negative) from the observer at full phase. In each run the seasonality is included, similar to *Figure* 4.9 and 4.15. The top left panel provides F, the top right panel Pl, the bottom left panel Q and the bottom right panel U.

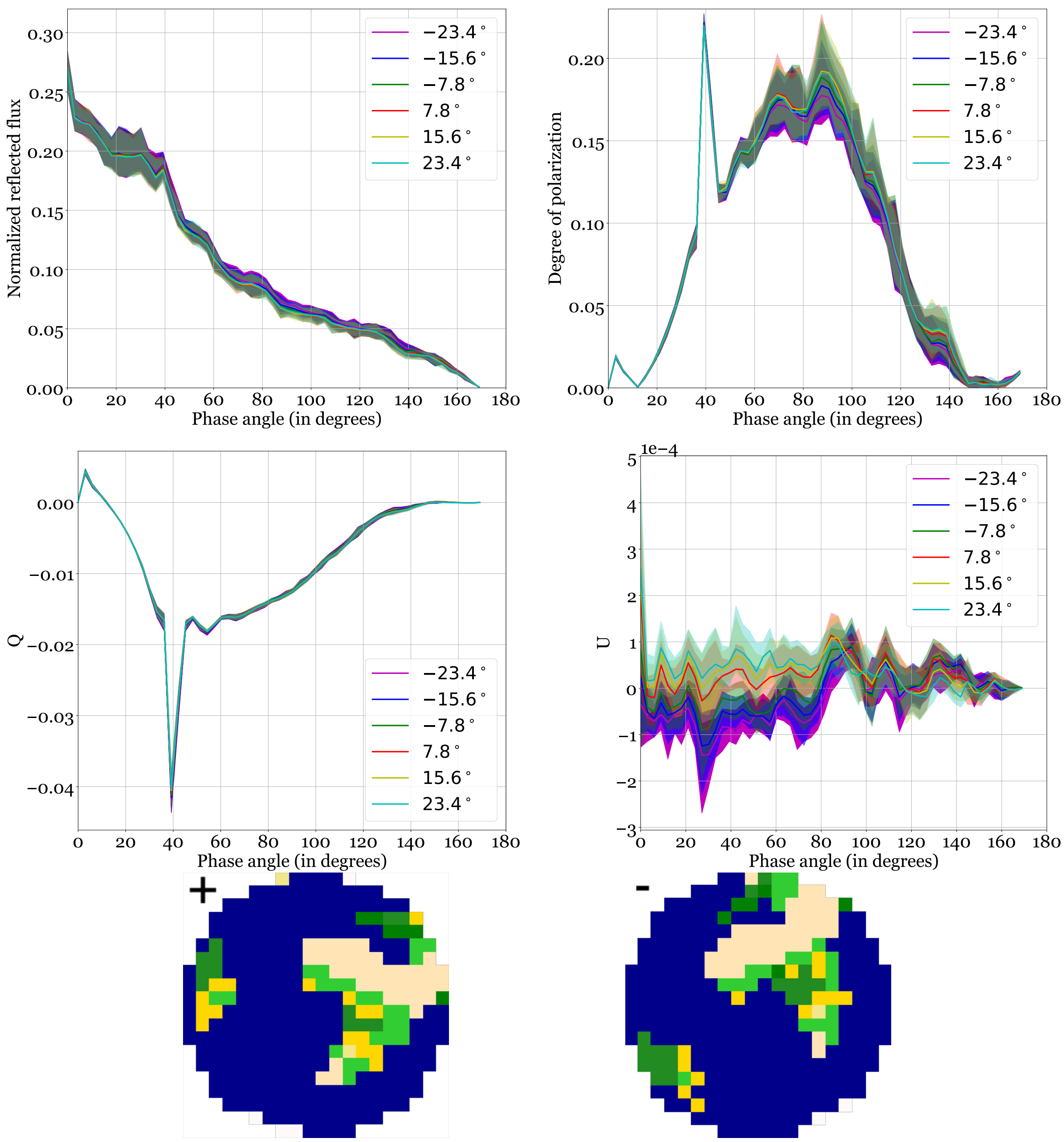


Figure 4.11: Six runs with different angles of obliquity oriented with the North Pole to the left (positive) or to the right (negative) as seen from the observer at full phase. In each run the seasonality is included, similar to *Figure* 4.9 and 4.15. The top left panel provides F, the top right panel Pl, the bottom left panel Q and the bottom right panel U.

that gaseous layer it is likely to be reflected by an underlaying cloud layer, if present. Hence, due to the small penetration depth of the light beams in the atmosphere only a small portion of the light reaches the surface, and getting partly absorbed. Thus, for longer wavelength light is more likely to travel to the absorbing surface, essentially decreasing the overall reflection. However, at $865\ nm\ F$ attains a higher reflection than at 670 and $550\ nm$, which is caused by (1) the spectral increase of the surface albedo of vegetation and desert, where the albedo of the ocean and ice only vary slightly, and (2) the increase in reflection of the liquid water particles for longer wavelength as is shown in *Figure* 1a from Stam (2008*a*). Albeit, this increase in cloud particle reflection is only minor and not the case for all scattering angles.

- The primary rainbow is visible for all wavelengths. The strength of this enhancement relative to the continuum of the phase curve increases for longer wavelengths. The magnitude of the daily variations does not seem to be affected much at the primary rainbow.
- At $\alpha =\sim 120°$ there occurs a color reversal in a particularly clean intersection point. For large α at long wavelengths the clouds reflect light more intensively to the observer, because (1) the optically thin gaseous atmosphere on top of these clouds allow more light to travel relatively undisturbed to the clouds and back through the atmosphere to the observer, and (2) because the clouds reflect more light in a forward scattering direction (see *Figure* 1a in Stam (2008*a*)). At short wavelengths the atmosphere is much thicker, i.e scattering light in a diffuse manner.

Our findings for Q are itemized as follows:

- The absolute magnitude of the rainbow peak decreases for longer wavelengths, but the magnitude relative to the continuum increases with increasing wavelengths. The enhancement centered at $\alpha =\sim 56°$ shows up for all wavelengths except $350\ nm$, where reflection from the clouds are mostly suppressed by the thick gaseous atmosphere.
- Without considering phases near both rainbows and the glory, Q decreases for longer wavelengths. Also for longer wavelengths the variability in Q decreases. The overall decrease in Q is caused by (1) the fact that Rayleigh scattering becomes less effectively allowing (2) the clouds to scatter more light multiple times, essentially depolarizing it, and allowing (3) the more accessible surfaces to completely depolarize the reflecting light. The latter reasoning also explains why the daily variations are suppressed for longer wavelengths. That is, other than gaseous medium in our model other sources of scattering mostly or fully depolarize the reflected light (see (2) and (3)) thus inducing no variability in the polarized flux Q where the gas is optically thin.
- The fact that almost no variability is present at all wavelengths, could be a valuable tool when adding polarized surfaces to our model, as we expect that these will cause large oscillations at longer wavelengths, due to (1) the increase in albedo of some surface types and (2) the larger penetration depth of the atmosphere, allowing more light to be reflected from the surface.
- The daily variation on the primary rainbow is present for all wavelengths, being caused by the fact that, as we know, the cloud particles polarize light more strongly and thus induce variability in Q.

Our findings for Pl are itemized as follows:

- For all wavelengths the daily variability is small for $\alpha <\sim 35°$, where the maximum variations occur, except $865\ nm$, in a large region of phases around quadrature. For $865\ nm$ this maximum variability is observed around the primary rainbow.
- The absolute magnitude of the primary rainbow peak in Pl is relatively constant, albeit the magnitude relative to the continuum increases for longer wavelengths. At $350\ nm$ the peak only slightly exceeds the continuum. The secondary rainbow is only barely visible for 670 and $865\ nm$, but its presence is mostly suppressed by the daily variations.
- Other than near the glory and primary rainbow the order of colors are maintained for all phases, except near crescent phases where long wavelength cross other phase curves due to attaining negative values in Q.

Our findings for U are itemized as follows:

- At short wavelengths, U shows large variations with high magnitude relative to longer wavelengths. In absolute magnitude, however, the variations of U are very small. The relatively high oscillations are caused by the high effectively of Rayleigh scattering. For longer wavelengths, this source of polarized light decreases as the clouds are mostly depolarizing and the surface completely depolarize.
- At short wavelengths the variations quickly diminish for phases larger than quadrature, which is not the case for longer wavelengths.

In comparison to the results of Karalidi, Stam and Hovenier (2012), *Figures* 5, 6, 7 and 8, we found that the strength of the rainbow in F relative to the continuum is much stronger. Also we found that the relative magnitude increases for longer wavelengths, whereas in their analysis the rainbow completely vanishes. This major difference is caused by the fact that they use $r_{eff} = 0.2; 6.0\ \mu m$, $v_{eff} = 0.1; 0.4$ and $\tau = 2.0$, whereas we use $r_{eff} = 10; 12.5; 15; 17.5\ \mu m$, $v_{eff} = 0.1$ and $\tau = 5.0; 10; 20$. It is thus apparent that the strength and possible retrieval of the primary rainbow signature greatly depends on the size and distribution of the particles and the optical thickness of the clouds. Lastly, we find for F and Q that the peak of the primary rainbow moves to higher phase angles for decreasing wavelength. This is completely trivial as one remembers the visual effect of an actual cloud bow on liquid water clouds in our atmosphere, which does not appear white, but exhibits different colors. In the next section we will continue this discussion. In the last section we will provide the seasonal effect of the clouds for all wavelengths considered here.

Color decomposition of the spectropolarimetric signal

In the former section we saw that in F the reflection from a "cloudy" Earth-like exoplanet is blueish for $\alpha <\sim 120°$ and reddish for larger phase angles. That is, we saw an alternation between the different wavelengths in a clear intersection point where for $\alpha <\sim 120°$ the light beams with short wavelengths are scattered more intensely and for larger phase angles light beams with longer wavelengths. If we compare this to the appearance of clouds in our day to day life we would expect that clouds appear white and thus reflect light beams at all wavelengths with the same intensity. As we mentioned before, the cloud fraction is in the range of ~ 0.68. So, how do we explain that the reflected light appears blueish and reddish with such a high amount of clouds? To answer this question we will simulate the color of reflected light with a weighted additive color mixing model in combination with phase curves at $\lambda = 443,\ 550$ and $670\ nm$, acting as the primary colors.

In *Figure* 4.13 we simulated different homogeneous cloudy planets with a black depolarizing surface and a gaseous Earth-like atmosphere. More specifically, in the upper left panel we model a cloudy planet with $P_c = 700\ mb$ and $\tau = 20$, in the upper right panel $P_c = 0\ mb$ and $\tau = 20$ and in the lower panel $P_c = 0\ mb$ and $\tau = 5$, where with the latter two cases we simulate a cloud layer at the top of the atmosphere. In all simulations we use $r_{eff} = 10$, where for other values no significant difference was observed on the phase curves. Also, we provide RGB plots under every plot. We observe the following:

- With a gaseous atmosphere on top of a thick cloud we still observe a weak blueish and reddish reflection from our planet as we saw in the planetary phase curves of our Earth-like model.
- By placing a thick cloud layer at the top of the atmosphere, both colors disappear and we obtain a white reflection from the exoplanet at all phase angles.
- If we model the cloud at the top of the atmosphere, with an optically thin cloud layer we see that at large phase angles the cloudy planet still appears white, but becomes blueish at quadrature and smaller phase angles.

We can conclude the following: (1) the red appearance at large phases emerges if a gaseous layer is present of top of a cloud layer, and (2) the blue appearance at small phases emerges from the gas molecules on top, but also in and under the cloud layer, being more pronounced for lower optical depth of the cloudy layer.

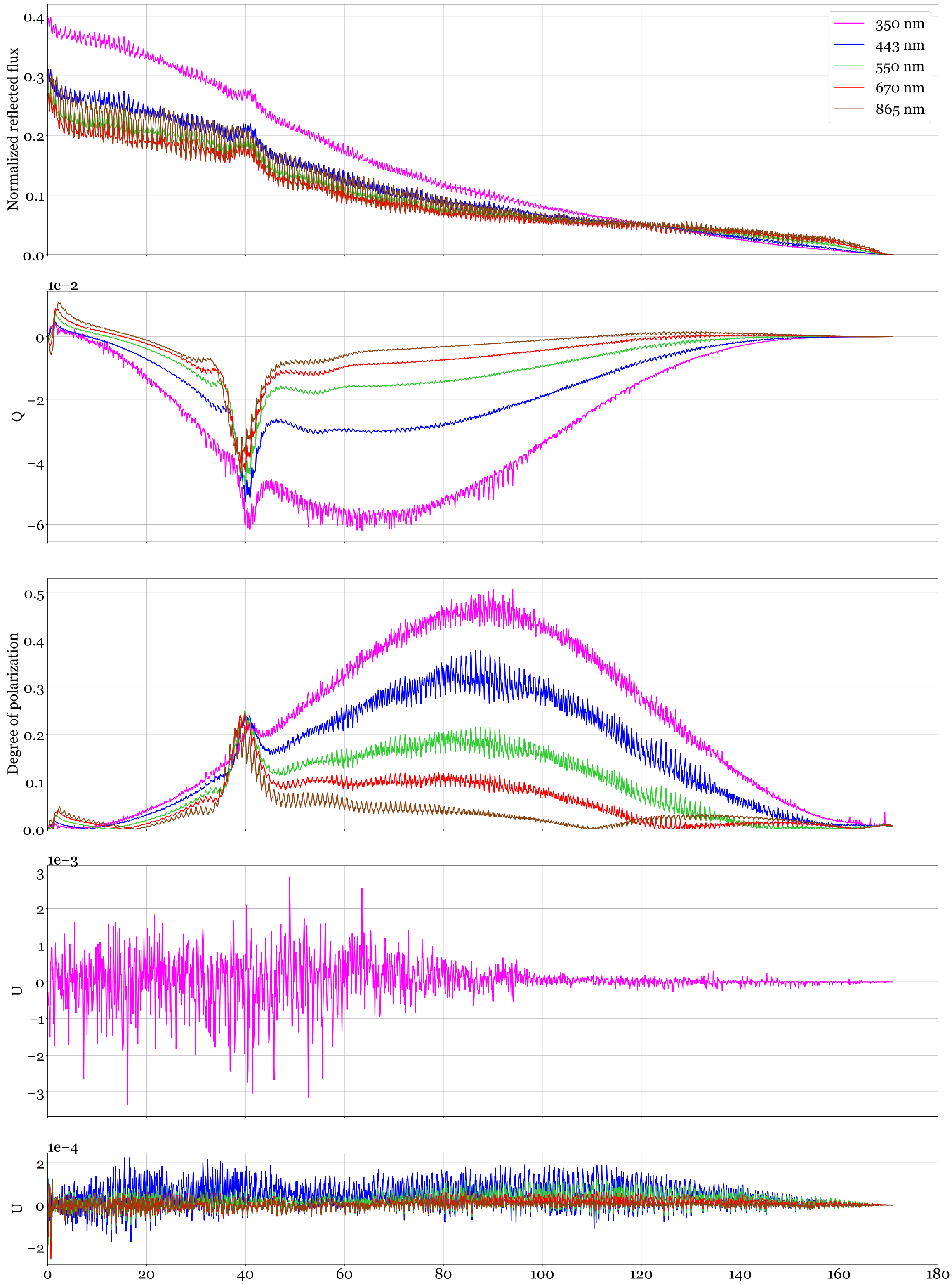


Figure 4.12: Planetary phase curves for F, Pl, Q and U in the top to bottom panels, respectively. These phase curves are provided for five wavelengths: 350, 443, 550, 670 and 865 nm. Lambertian surface models are used and the gaseous atmosphere is kept constant. The cloud layers vary according to MODIS data.

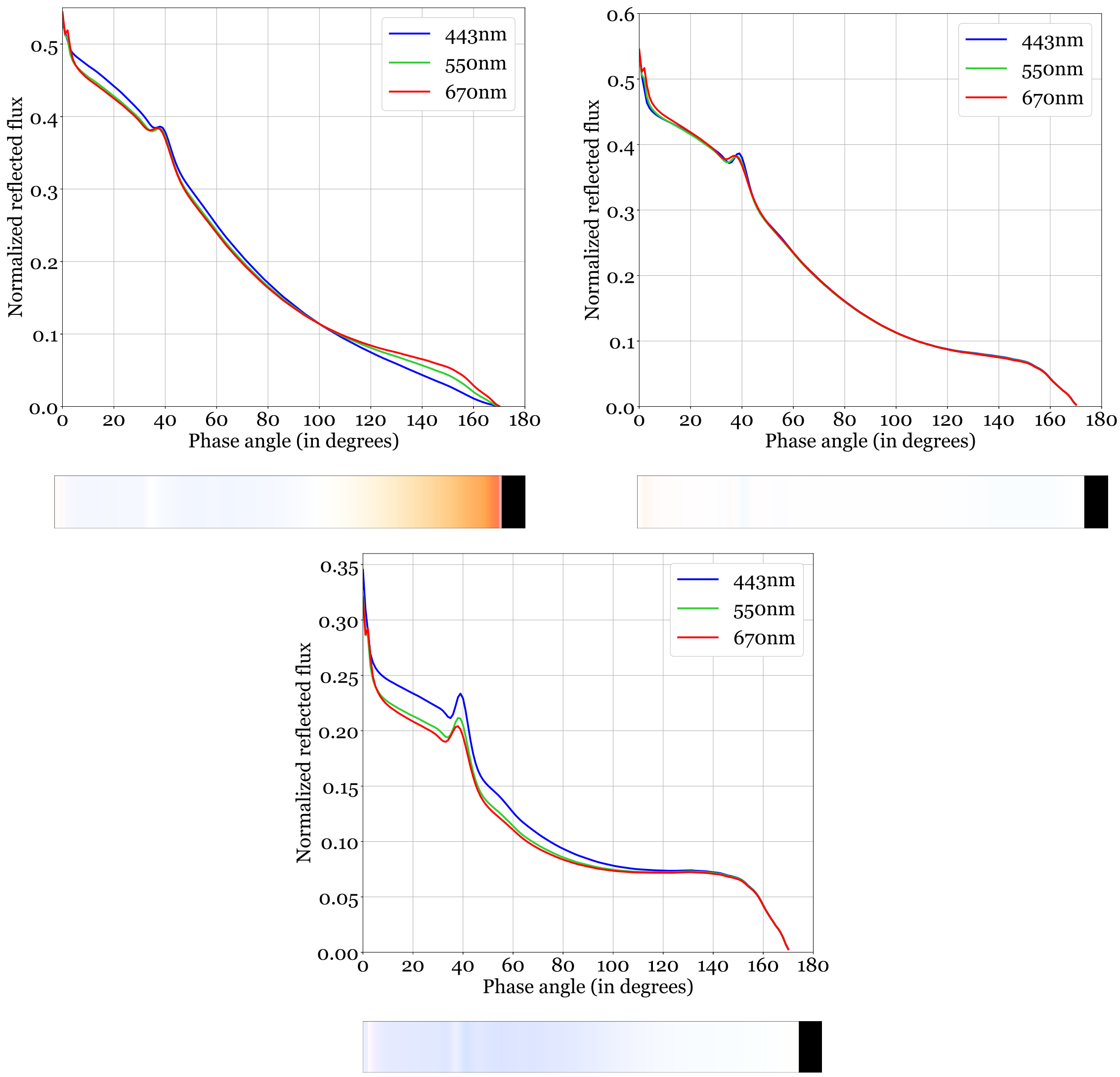


Figure 4.13: Upper left panel: homogeneous planet with black surface; clouds of $\tau = 20$, $P_c = 700\ bar$ and $r_{eff} = 10\ \mu m$. Upper right panel: homogeneous planet with black surface; clouds of $\tau = 20$, $P_c = 0\ bar$ and $r_{eff} = 10\ \mu m$. Lower panel: homogeneous planet with black surface; clouds of $\tau = 50$, $P_c = 0\ bar$ and $r_{eff} = 10\ \mu m$. All have a gaseous atmosphere like we used for all our simulations. RGB color strokes are provided for all α.

By inspection of *Figure* 4.12 it may be observed that the primary rainbow peak shifts to higher phases for longer wavelengths. We argued that this is completely trivial as one sees multiple colors from a cloud bow in our atmosphere. *Figure* 4.14 shows the RGB colors for F, Q and Pl from the planetary phase curves in *Figure* 4.12 at $\lambda = 443,\ 550$ and $670\ nm$. Despite the large daily variability's in our simulated phase curves we retrieved a clear color alternation near $\alpha = 40°$ that corresponds fairly well to that of a cloud bow. The color alternation is most clearly retrieved for Pl for which we saw that the primary rainbow peak attain approximately the same magnitude. For Q the rainbow peak attained different absolute magnitude for different λ, but still provides a clear color alternation. For both parameters this feature is not observable by a human eye, because it can both be observed only by using a linear polarization filter. For F the variation of absolute magnitude of the rainbow causes the cloud bow to be hardly visible, but if one looks closely it is present. This is caused by the gaseous atmosphere in our model that was show in the former discussion. In conclusion, the shift in position of the primary rainbow results in a cloud bow not only for F, but also for Q and Pl, unless the presence a of "blue" gaseous atmosphere in the planetary model.

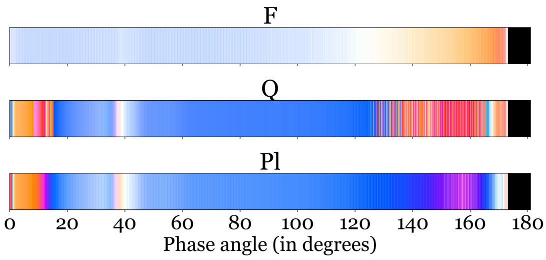


Figure 4.14: RGB color strokes according to the phase curves F, Q and Pl at λ = 443, 550 and 670 nm provided in *Figure* 4.12.

Seasonality at different wavelengths

In this section we provide phase curves, which are simulated in the exact same manner as described in *Section* 4.2.1, but at $350,\ 443,\ 550,\ 670$ and $865\ nm$ and only for a temporal resolution of 3.095 days. The phase curves are provided in *Figure* 4.15. The shaded areas and solid lines represent the same statistics as in *Figure* 4.9. The main feature that we observe for these phase curves are already addressed in the first part of this entire section.

The seasonal variability in F shows to increase for increasing wavelengths. These variations have similar or less magnitude as compared to the daily variations in *Figure* 4.12. The primary rainbow feature is barely visible at the short wavelengths, but is clearly visible at the near-infrared wavelengths. Also, similar to *Figure* 4.12 the absolute seasonal variability decreases with phase angle. Q appears to be relatively insensitive to the seasonal variability for wavelengths longer than $350\ nm$. At the rainbow peak we see small variations for $\lambda = 443,\ 550,\ 670$ and $865\ nm$. One could expect this result as we have seen that Q is virtually insensitive to the cloud and surface cover. For $\lambda = 350\ nm$ we see much variations for phase angles near quadrature and the primary rainbow. We do, however, again see that these variations have the same relative magnitude as the daily variations. For Pl the relative magnitude in variability, induced by the different seasons, also closely resembles that of the daily variations in *Figure* 4.12. As for the previous parameters also U shows similar orders of magnitude as the general phase curves presented in the previous section.

In conclusion, Q is least sensitive to the seasonally changing cloud cover for long wavelengths. For all wavelengths the seasonal variations are in the order magnitude of the diurnal variations, so it would be very difficult to identify seasonality at any wavelength region directly from the planetary phase curves.

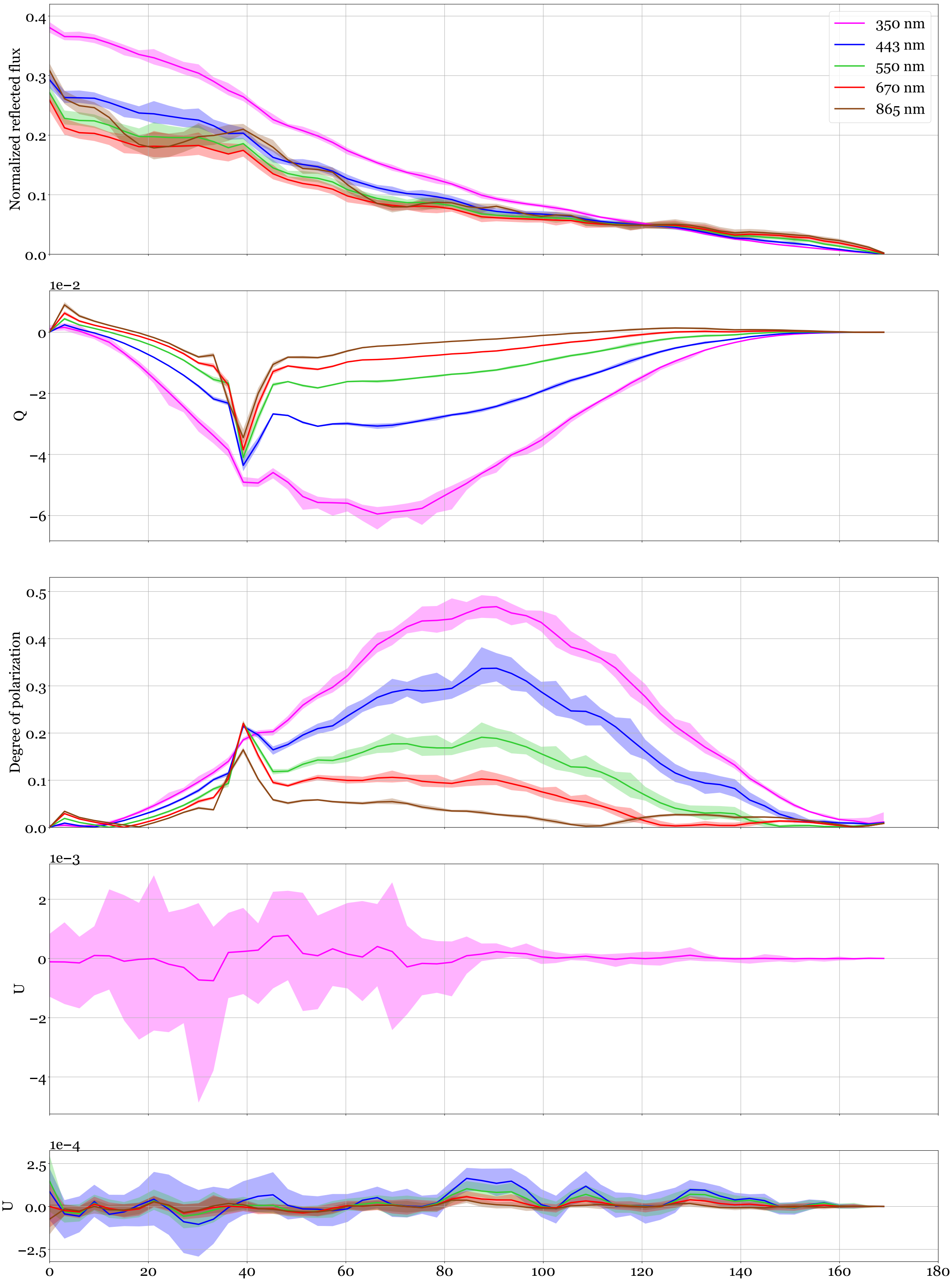


Figure 4.15: From top to bottom we provide F, Pl, Q and U, all as function of phase angle. The shaded areas show the minimum and maximum values and the solid line shows the mean value. The plots consist of 53 runs with different starting days and are simulated at λ = 350(magenta), 550(green) and 865 nm(brown). The temporal resolution of each run is 3.059 days.

4.2.3. Diurnal light curves

The major oscillation that we observed in *Figure* 4.12 occurs due to (1) the rotation of Earth around its own axis in combination with (2) the horizontally inhomogeneous cloud and surface cover. The analysis for diurnal light curves is limited to a number of interesting phase angles, which are: $\alpha =\sim 0°$, $\sim 40°$, $\sim 90°$ and $\sim 135°$; with a temporal resolution of one hour. *Figure* 4.16 shows the variations in F, Pl, Q, and U as functions of the sub-observer longitude for various wavelengths and phase angles, normalized by subtracting the mean value of each diurnal curve. In the bottom panels we present the land cover, cloud top pressure, cloud optical thickness, cloud particle effective radius and cloud fraction distributions at full phase. If one refers to $\alpha = 90°$ only the right half side of the disk would contribute to the signal, etc. The planetary model the same as that used in *Section* 4.1.

For F at $350nm$ we see an overall agreement between the different phases, except for crescent phases, with peak values corresponding the vegetated continents of Southern America and Asia, even though the surface albedo is very low. Because various cloud patches are correlated to the shapes of the ocean/land, during a full rotation, parts of the variability in our curves can be attributed to both clouds and/or surface features. From the different diurnal curves we can observe a slight shift to lower longitudes with increasing phase angle, which is caused by the fact that we see a decreasing visible and illuminated region that shifts to the right side of the planetary disk, essentially delaying the diurnal variation to smaller longitudes. At $865nm$ the variation for $\alpha = 40$ and $90°$ become stronger, where for other phases it remains approximately the same. For example, at $90°$ phase there are three peaks that seem to correspond to the highly reflecting vegetation covering Africa, America and Asia, where at this wavelength the vegetated surface covers are highly reflecting. In all cases the low values of reflection near $0°$ and $180°$, corresponding to the Atlantic and Pacific Ocean respectively.

For Pl at $350\ nm$ we see only significant variations at quadrature. This trend shows two minima for the vegetated land covers of Southern America and Asia. For crescent we observe a lot of variations, whereas for full phase these are relatively smooth. At 443 and $550\ nm$, there are strong variations near zero longitude that seem to be caused by a major region containing clouds with high cloud top pressure. For longer wavelengths we start to see much more variation at other phases, especially at the primary rainbow.

Q shows much variations for all phase except crescent phases at $350\ nm$. As shown in *Section* 4.2 we would expect the strongest variations for $\alpha = 40$, which is clearly visible. For increasing wavelengths these variations tend to smooth out. At quadrature and crescent phases no clear variations are visible in agreement to our phase curves in *Figure* 4.12.

For U we can see some major variations for all wavelengths regions and phases. In magnitude, however, the strongest variations occur at $350\ nm$. As seen before, it is hard to retrieve an unambiguous correlation with cloud parameters or surface types. In the presented analysis, it was found that U is affected by the clouds, and mostly by the cloud optical thickness (*Section* 4.1). For full phase at $350\ nm$ we do see that with increasing longitude a higher abundance of high optical thickness values near the center of the disk occurs, but for longer wavelengths this correlation seems to vanish, which is in agreement for what we found in the resolved disks. When referring back to *Section* 4.1 we saw that U is not only affected by the abundance of cloud parameters over the center of the disk, but more so by extreme cloud parameters in one of the four quadrants of the circle, such that an asymmetry over the planetary scattering plane occurs. Although we may know how U is affected, it is still hard to retrieve some kind of cloud distribution from the disk integrated signal.

In conclusion, the shape of the diurnal light curve of F as function of wavelength show to be fairly constant. The variations of Pl are small near full phase and crescent phase of Q for quadrature and crescent phases. For U we found major variation, but can not be unambiguously correlated these to any temporal variation in the surface or cloud distributions. The fact that observations on the level of the diurnal period might not tell us much directly is not necessarily a loss, because in reality integration time depends on both the telescope and polarimeter capabilities, the planetary and star properties, the orbital geometry, distance to the extrasolar system, etc, and may exceed a sufficient temporal resolution to map the diurnal variability accurately.

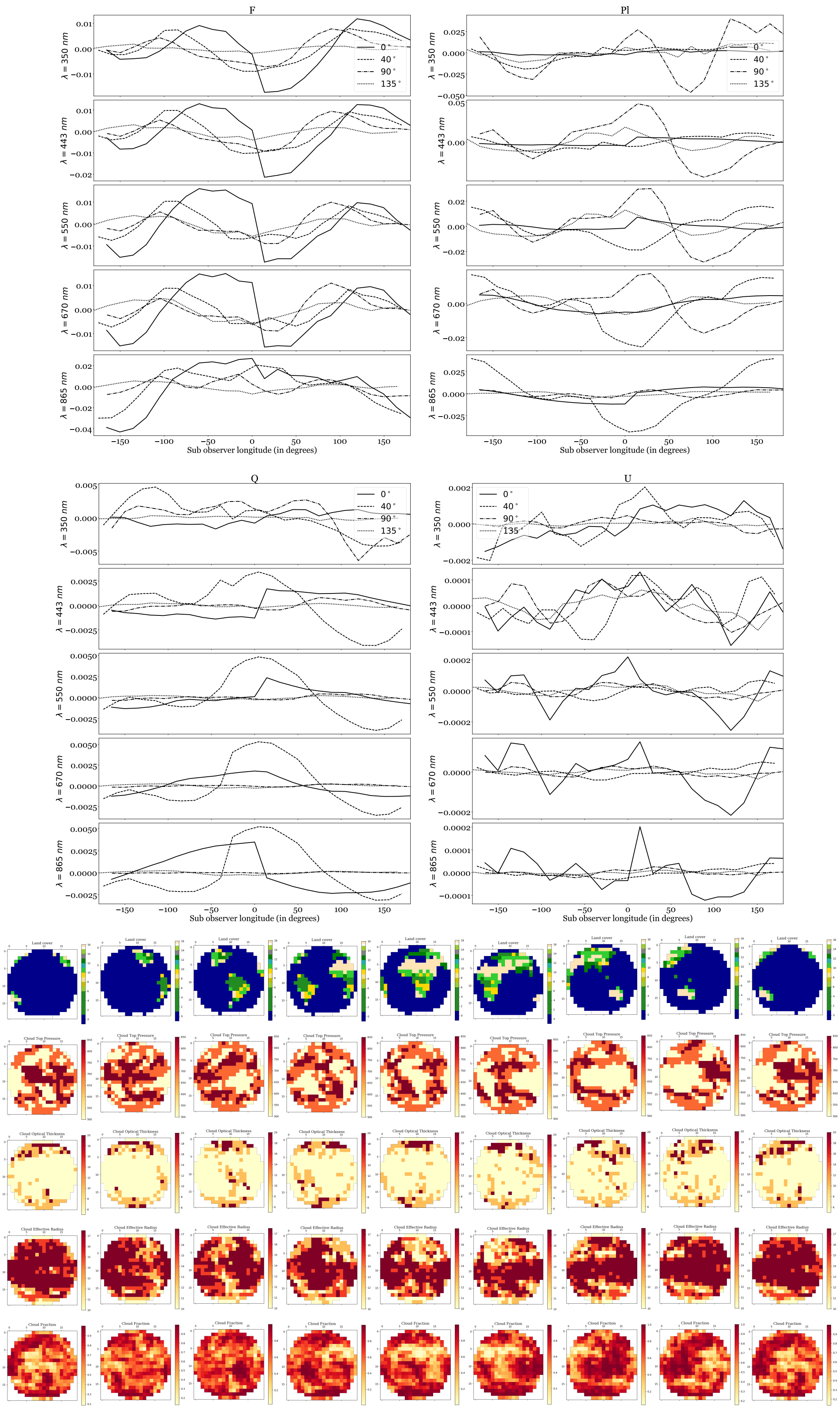


Figure 4.16: Diurnal variations for F, Pl, Q and U in the upper left, upper right, bottom left and bottom right panel, respectively. In each sub panel we present diurnal variations as function of sub observer longitude for $\alpha = 0°, 40°, 90°$ and $135°$. Each sub panel corresponds to a specific wavelength, from top to bottom: $\lambda = 350, 443, 550, 670$ and 865 nm. As an addition we provide the land cover distribution, cloud top pressure, cloud optical thickness, cloud particle effective radius and cloud fraction.

5

Characterization of Earth clouds from the scattered light curves

A parameter that potentially can be retrieved from photometric signals is the rotation rate of an exoplanet around its own axis. Pallé et al. (2008) used Fourier power analysis and an autocorrelation to retrieve the rotation rate of Earth around its axis by using a synthetic planetary model, for different sub-observer views, exposure times, signal-to-noise ratios and observation periods. These cases were applied to a data set of 21 years and percentages of success for the 24 hour, 12 hour and other periods were documented. Based on these success rates, Pallé et al. (2008) concluded that the autocorrelation method is more accurate and robust in the characterization of the rotation period using photometric time-series data. In a later study by Oakley and Cash (2009), similar results were obtained also based on photometric signals only. The question remains: can polarimetry provide additional confidence in retrieving this period? Also, can we retrieve the presence of dynamic cloud cover on an Earth-like exoplanet by comparing the confidence in retrieving the rotation rate for F, Pl and Q together? Both questions will be answered in the first section (*Section* 5.1). In *Section* 5.2, we will investigate the variability of the spectropolarimetric signals at two fixed phase angles and assess whether there exists a correlation with any of the cloud parameters.

5.1. Periodicity analysis on spectropolarimetry

Periodic patterns in our spectropolarimetric signals can be retrieved by transforming them into the frequency domain with a Discrete Fourier Transform (DFT). We consider the following spectropolarimetric signal:

$$x(n), \qquad n = 0, 1, ..., N-1 \tag{5.1}$$

where $X(n)$ is a specific data point and N the total number of data points. The principal idea of the Fourier transform method is that the original signal is expressed by a linear combination of periodic components, more specifically the following complex sinusoid (Vlachos et al. 2005):

$$s_f(n) = \frac{e^{j2\pi f\ n/N}}{\sqrt{N}} \tag{5.2}$$

The DFT on the sequence of data points X provides us (Vlachos et al. 2005):

$$X(f_{k/N}) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x(n) e^{-\frac{j2\pi k}{N}} \tag{5.3}$$

where k/N is the frequency of a specific coefficient and $k = 0, 1, ...N-1$. In order to retrieve the periodic components of the signal, the Power Spectral Density (PSD) or power spectrum needs to be examined, which essentially provides the power of the signal at a specific frequency, or in other words the most dominant time period. Estimators of this PSD are the periodogram or the autocorrelation function.

In *Sections* 5.1.1 and 5.1.2 we provide an introduction to the periodogram and autocorrelation function. Next, we will assess what type of noise is generally induced the photometric and polarimetric signals (*Section* 5.1.3). In *Section* 5.1.4 we attempt to retrieve the rotation period of Earth for various level of noises, temporal intervals and temporal resolutions. Lastly, we use the autocorrelation method to retrieve the presence of the dynamic cloud cover in our planetary model (*Section* 5.1.5). If not stated otherwise, we use the phase curves presented in *Figure* 4.6, *Section* 4.2. These phase curves were computed with the planetary model described in *Chapter* 3 that consists of a Lambertian surface model, a temporally invariant gaseous atmosphere and the cloud distribution according to MODIS data.

5.1.1. Periodogram

The periodogram $\mathcal{P}$ is produced by using the DFT from *Equation* 5.3 and computing the length of each frequency component (Vlachos et al. 2005):

$$\mathcal{P}(f_{k/N}) = ||X(f_{k/N})||^2, \qquad k = 0, 1, ..., [\frac{N-1}{2}] \tag{5.4}$$

where the power signal can only be computed up to half the maximum signal frequency, limited by the Nyquist theorem. For large periods the accuracy of the power spectrum deteriorates for two reasons: (1) the frequency component $X(f_{k/N})$ corresponds to the period interval $[\frac{N}{k}...\frac{N}{k-1})$, which increases in width with increasing period; (2) Spectral leakage, which occurs due to bad sampling of frequencies in the DFT bins, resulting in a dispersion over the entire spectrum. The periodogram is thus particularly useful for small to medium periods.

5.1.2. Autocorrelation function

The circular autocorrelation function, or more commonly known as the autocorrelation function (ACF) is an estimator of the dominant periods, by examining the similarity between sequences of data separated by different lags (τ) (Vlachos et al. 2005):

$$ACF(\tau) = \frac{1}{N} \sum_{n=0}^{N-1} x(\tau) \cdot x(n+\tau) \tag{5.5}$$

A more convenient expression can be exploited by using the DFT of the signal:

$$ACF = \mathcal{F}^{-1} < X, X^* > \tag{5.6}$$

where $*$ denotes the complex conjugate. The ACF is a better periodicity detector than the periodogram and can detect larger periods more accurately. This difference was also observed by Oakley and Cash (2009); Pallé et al. (2008) for Earth-like photometric signals. However, the ACF on itself is not sufficient for automatic periodicity detection. According to Vlachos et al. (2005) the ACF requires a manually set threshold, and the method introduces many false alarms that need to be removed manually. Also high frequency periodicity events with low amplitude appear less strongly in the ACF than in the periodogram.

5.1.3. Noise in the spectropolarimetric signal

Until now, we provided simulated spectropolarimetric signals without attenuation of e.g. interstellar dust and noise from e.g. instrumentation. Here, we take a brief look at how our periodicity analysis is affected when we add noise to our simulated signals. We believe that this is appropriate for the current analysis, because we will attempt to retrieve the rotation period of Earth based on the techniques proposed by Oakley and Cash (2009); Pallé et al. (2008), ultimately linking these results to the presence of dynamic cloud covers. Examples of dominant non-instrumental noise sources are: direct star light from the parent star and exo-zodiacal/zodiacal light from for example dust (Oakley and Cash 2009; Traub et al. 2006). Examples of instrumental noise sources are: throughput, quantum efficiency, readout noise and dark current.

For the current test case we will ignore non-instrumental noise. White Gaussian noise will be added to the fluxes F, Q and U separately to mimic instrumental noise (Snyder et al. 1995). This consideration is based on a recently developed polarimeter: DIPOL-2 (Piirola et al. 2014). This polarimeter is designed to observe polarized light for three wavelength passbands simultaneously by used of three CCD's (Charged Coupled Device). The two orthogonal polarized light beams are split by a calcite analyzer. The polarized light beams are thus separately detected on one CCD and independently receive instrument noise. It has to be noted that currently and in the future other types of polarimeters or spectrometers might be employed that induce noise on the spectropolarimetric signals differently, like if for example the degree of polarization is measured directly. In this case, the White Gaussian noise should have been applied directly on the simulated parameter Pl, whereas in this case we compute Pl from F, Q and U that contain already some degree of noise.

5.1.4. Retrieval of rotational period

In this section we will review the ability to reproduce the rotation rate Earth, i.e. a solar day. As we described in the previous section, we will apply some degrees of White Gaussian noise to our spectropolarimetric signals. This is mainly done to assess with what level of noise the diurnal rotation period cannot be retrieved anymore. Additionally, we try to retrieve the period with different temporal resolutions and intervals. The White Gaussian noise that will be used has a mean of zero and a variable standard deviation. We start by analyzing a full set of data that consists of a simulation done with a diurnal resolution of data 12 points and an observation interval between and including 01/01/2011-06/30/2011. We increase the standard deviation of noise until we see no significant peak in either the correlogram or periodogram. Values of noise are labeled in the lower panel of *Figure* 5.2.

In *Figure* 5.1 one can observe the periodogram for F, Pl, Q and U in the upper left, upper right, bottom left and bottom right panel, respectively. For F we can see that there are significant peaks at frequencies that correspond to 24 and 12 hours. Lower peaks can be observed for 6 hours, and even less strong peaks at 8 and 4.8 hours. The maximum level of noise is simulated with a standard deviation of 2% and shows to completely dilute any possible retrieval of any dominant frequency. With a standard deviation of 1% we are able to retrieve peaks that correspond to 24 and 12 hours, where there is no significant difference between the two. By decreasing the noise level we start to be able to retrieve strong peaks for 8, 6 and 4.8 hours. Also, with decreasing noise level the peak at 12 hours is stronger than that at 24 hours.

When we apply the Fourier technique on Pl we see that by adding 0.1% noise the major peaks at 24 and 12 hours are still visible, but by increasing the noise to 0.2% we completely lose any dominant frequency. We have seen in *Chapter* 4 that Q contains barely any significant variability, or diurnal variations, in the simulated signal. With no noise we see that the strength of the peaks especially at 24 and 12 hours are already an order of magnitude lower in strength than that of F or Pl. It is therefore no surprise that when adding only a small portion of noise any dominant frequency is immediately lost. For U the magnitude is even lower. Without noise we however see significant peaks with respect to the continuum, which might indicate that U carries the rotation period quite well. Albeit, if we add only a small portion of noise any dominant frequency is immediately lost. These results show that the rotation rate can only be confidently retrieved from F, and that Q and U are a relatively weak tool if we consider some noise in our signals.

Now that we have seen what periods we can retrieve with the Fourier based technique, we will analyze the results from the autocorrelation method. In *Figure* 5.2 the correlogram of F, Pl, Q and U are provided in the upper left, upper right, middle left and middle right panels, respectively. We again simulated several cases with different standard deviations to include White Gaussian noise. The corresponding values are provided in the lower panel. With the autocorrelation function, one is able to retrieve periods on a longer timescale relative to the Fourier based technique.

The correlogram of F shows a significant bump/peak at 12 legs, which corresponds to 24 hours. One can also see a smaller bump, which corresponds to 12 hours. Similar bumps can be observed that correspond to consecutive rotation periods up to 96 hours or four solar days. If we increase the noise in these signals, one can see that the dominant peak of Earth's rotation period is maintained fairly well, and the peak corresponding to half of Earth's rotation's rate decreases much more rapidly. Similar to the Fourier based technique, with $\sigma = 2\%$ noise, the rotation rate cannot be unambiguously re-

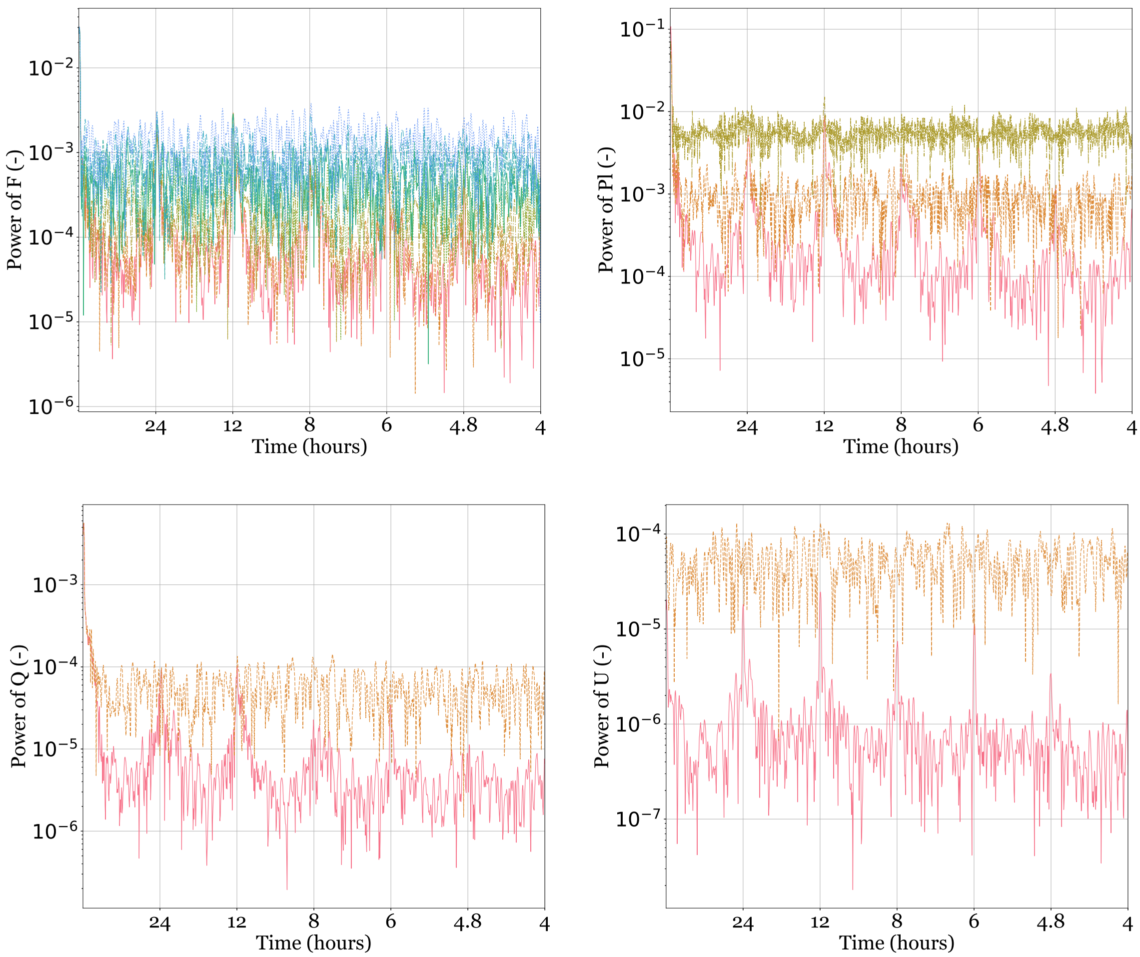


Figure 5.1: Periodogram of F, Pl, Q and U in the upper left, upper right, bottom left and bottom right panel, respectively. Different line colors correspond to different levels of White Gaussian noise. The red curve corresponds to our simulated spectropolarimetric signal without noise. The corresponding noise values are provided in a legend in *Figure* 5.2. On the x axis the time in hours is provided.

trieved. In these results for the correlogram and periodogram we see a similar behaviour as Pallé et al. (2008) identified, namely that the autocorrelation provides a much more confident retrieval of the rotation rate, whereas the Fourier based technique often retrieves 12 hours as the most dominant period.

If we add only a small portion of noise to Pl any significant correlation is immediately lost, essentially oscillating around zero. This effect was not as strong for the Fourier based technique, where we were still be able to retrieve some peaks with $\sigma = 0.1\%$. The lack of daily variability in Q can also be retrieved from the correlogram of Q, where we see hardly any highly correlated period, without even considering noise. In this case, we see that in an ideal case the Fourier based technique would have been more appropriate for the low fluctuations in Q. For U we again easily retrieve a significant correlated period at 24 hours. However, by adding a small amount of noise any highly correlated period vanishes. This is again in close correspondence to what was found with the Fourier based technique.

In conclusion, we found that the autocorrelation technique applied on F will provide a much more confident retrieval of an Earth-like exoplanet rotation period than that with the Fourier based technique, where the latter method retrieves the 12 hour period more confidently. For both methods we saw that the retrieval of any dominant period in Pl with the addition of some noise deteriorates fast. This was also observed for Q and U. In the case of Q this might be totally different if we only observe the region of the primary rainbow, where we see lots of variations. In the following section we will investigate different temporal intervals.

Periodicity analysis on different observation intervals

In theory we can simulate our spectropolarimetric signal in any way we want, but in reality an observer is unlikely to observe an exoplanet for half an Earth year. In this section we assess the periodicity retrieval at four different observation intervals with a total observation time of 10 days and an observational resolution of two hours: $\alpha =\sim 35-45^\circ$, $\alpha =\sim 85-95^\circ$ and $\alpha =\sim 130-140^\circ$. In the past sections we have included some deficiencies of present and possible future polarimeters or spectrometers that limit our ability to retrieve characteristics from an exoplanet. In this section the phase angle range to $30^\circ - 150^\circ$, similar to Oakley and Cash (2009). The range of observable phase angles depends largely on the Inner Working Angle (IWA) of the observing system, but also the inclination that have assumed as 90° (*Section* 3.5). For a more detailed discussion about this IWA one can review Rossi and Stam (2017). In this paragraph we will not show any results regarding periodicity's found in U, because they did not provided any substantial difference to what we saw in the previous section.

In *Figure* 5.3 we have provided the correlograms of F, Pl and Q in the left, middle and right panel, respectively. In the upper panel we provide the correlograms of the light signal for $\alpha =\sim 35-45^\circ$, in the middle for $\alpha =\sim 85-95^\circ$ and in the lower panel for $\alpha =\sim 130-140^\circ$. The associated noise values are provided in the lower panel of *Figure* 5.2. Overall we observe that the correlation of intermediate lags, i.e. lags that do not correspond to Earth's rotation rate, fractions of this period or consecutive Earth periods, are significantly lower. For F at the rainbow peak we see a stronger identification of the rotation period and consecutive periods, even for high values of noise. For Pl and Q no significant correlation is present at all, which is fairly surprising especially for Q. By inspection of the phase curves in *Chapter* 4 we know that at the rainbow large variations are induced by the polarizing clouds. The bad autocorrelation retrieval can be caused by two things: (1) The clouds that dominate the fluctuations in the signal have a large temporal variation, causing a lot of anticorrelation in the signal; (2) The highly non-linear continuum causes large anticorrelation in the signal. If we look at the retrieval around 90° we observe major peaks that correspond to the rotation period and consecutive periods for F and Pl. The retrieval of dominant periods in F, however, performs worse with increasing noise. We also see a consistent trend of slightly shorter periods than the real period at 24 hours, 12 hours or periods of consecutive days for the ideal as well as the noisy retrievals. For Q we find small peaks for the ideal case, but when adding a small amount of noise, no significant correlated period is retrieved. At crescent phases we see that F performs slightly worse than at $\alpha = 90^\circ$ in terms of absolute values of correlation and resilience to noise. For both Pl and Q we see that some addition of noise completely dilutes any possible correlated period.

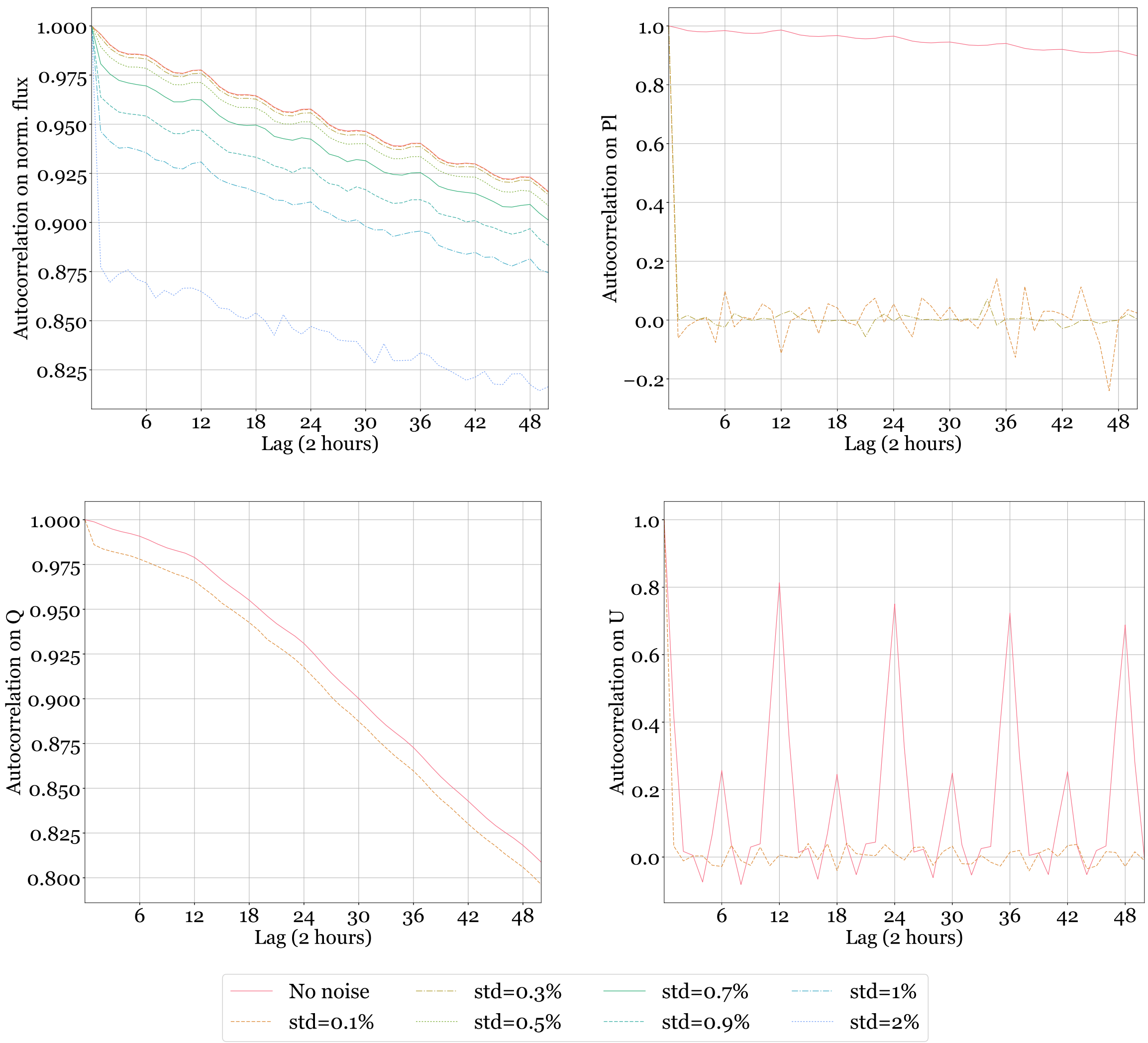


Figure 5.2: Correlogram of F, Pl, Q and U in the upper left, upper right, bottom left and bottom right panel, respectively. Different line colors correspond to different levels of White Gaussian noise. The red curve corresponds to our simulated spectropolarimetric signal without noise. The corresponding noise values are provided in a legend in the lower panel. On the x axis the number of lags is provided, where 1 leg corresponds to 2 hours in a solar day.

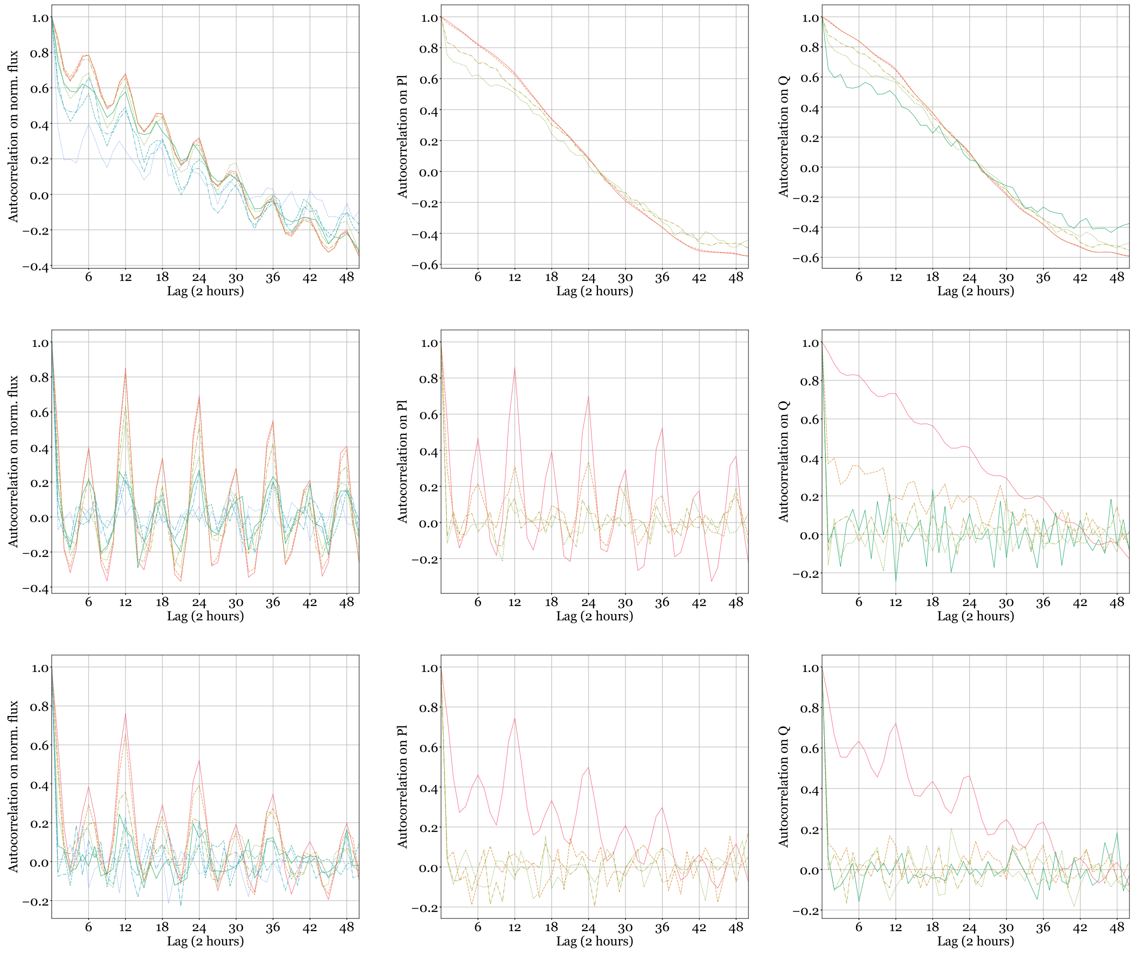


Figure 5.3: Autocorrelation on F, Pl and Q in the left, middle and right panel, respectively. In the upper panel we provide the correlograms of the light signal for $\alpha =\sim 35-45°$, in the middle for $\alpha =\sim 85-95°$ and in the lower panel for $\alpha =\sim 130-140°$. All data ranges comprise of a 10 day data set with a diurnal variation of 12 data points.

The Fourier based technique results are provided in *Figure* 5.4. Results of F show that at the rainbow phases half an Earth rotation period is dominant, and is significant up to noise with a standard deviation of 2%. The peak corresponding to 24 hours is significantly lower, and as dominant as the peak at 8 hours. Furthermore, we can see that the peaks are a bit blunt and do not correspond exactly to the expected periods, again being short of or longer than the real rotation rate. In Pallé et al. (2008) similar behaviour was identified, on which they provide that depending on the resolution of the data one has to take into account a delta equivalent to the exposure time, in this case two hours ($\Delta = 2hr$). For the retrieval at $\alpha =\sim 90°$ and $\alpha =\sim 135°$ any dominant peak deteriorates for lower values of noise than at the rainbow, where at crescent phases the effect of noise most prominent. This is also true for the signals Pl and Q. For Pl and Q we moreover find that again any retrieval is diluted quickly by noise if we compare their performance to F.

In Pallé et al. (2008) several interesting findings were made that can identify dynamic cloud cover on Earth-like exoplanets. By simulating different sub sets of a full data set, the presence of dynamic weather can be related to ”apparent” rotational periods that are consistently short of or longer than the real rotation period. These shifts are completely gone when considering only a clear atmosphere on which they argue that these are produced by the dynamic cloud cover. More specifically, shorter periods might correspond to cloud decks that move westward (in the direction of Earth’s rotation) and longer periods to cloud decks that move eastward. It is hard to confidently retrieve whether this is

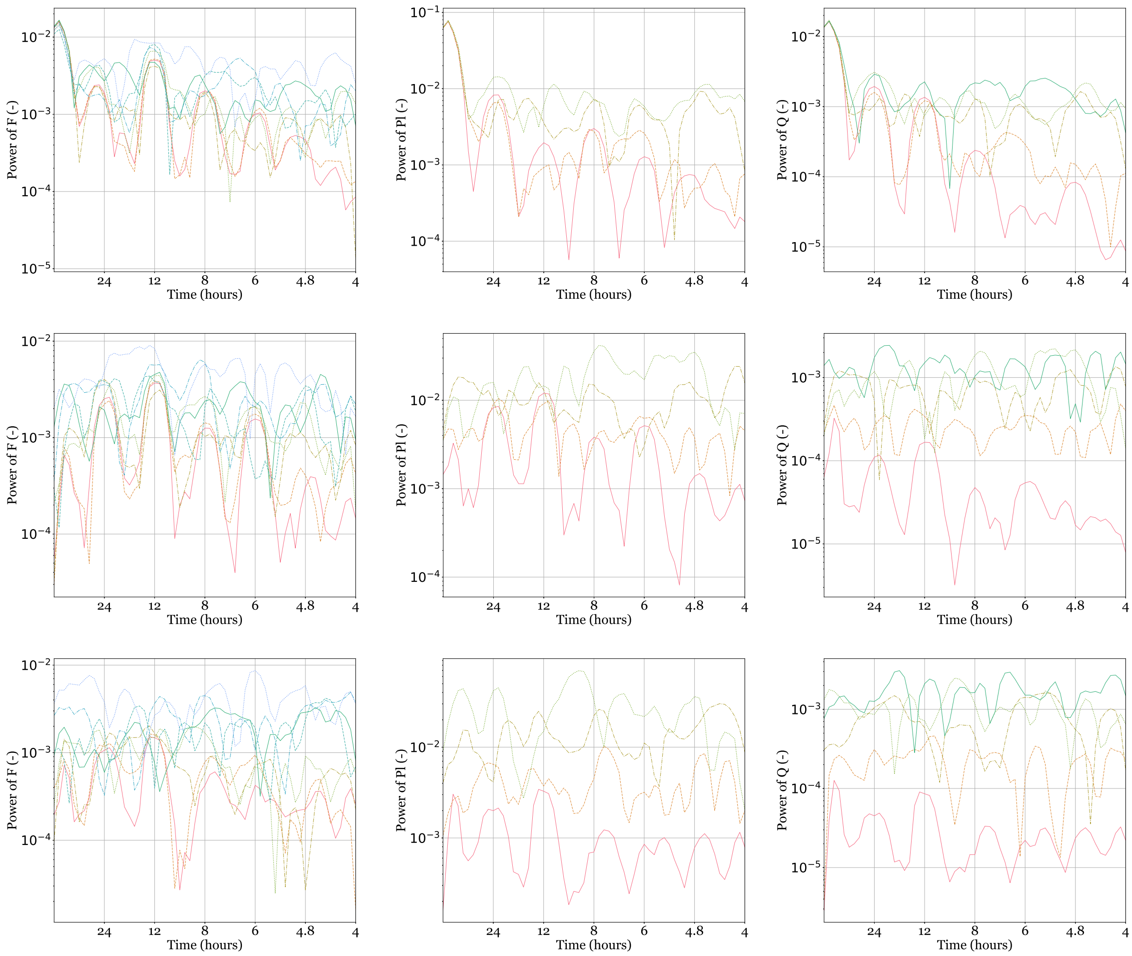


Figure 5.4: Fourier based analysis on F, Pl and Q in the left, middle and right panel, respectively. In the upper panel we provide the periodograms of the light signal for $\alpha =\sim 35-45°$, in the middle for $\alpha =\sim 85-95°$ and in the lower panel for $\alpha =\sim 130-140°$. All data ranges comprise of a 10 day data set with a diurnal variation of 12 data points.

exactly the cause, because in reality Earth houses different global weather patters like the Intertropical Convergence Zone, the Horse Latitudes and the Polar Fronts. In our results, we can identify similar slightly shorter apparent rotation periods like Pallé et al. (2008) found.

5.1.5. Using polarized surfaces to detect the presence of dynamic weather

In this section we attempt to retrieve the presence of dynamic weather from the correlogram of F, Pl and Q. In the previous section it was mentioned that dynamic cloud cover can be related to the "apparent" rotation rate of different data sub sets. In this thesis we will not delve any further into this theory, but investigate a possible new signature that can be retrieved from the periodicity analysis, essentially employing our ability to process polarized light curves. Our hypothesis:

> *The polarized flux Q shows to be relatively insensitive to the varying cloud cover, except around the primary rainbow and glory. The strength of the autocorrelation over multiple consecutive Earth rotation periods from Q is higher than for F and Pl. That is, the surface cover is annually invariant, but the spatial distribution of the clouds depend on daily observations (Chapter 3).*

More illustrative, *Figure* 4.6 shows that Q is barely affected by introducing clouds in our model for $\alpha \gtrsim 60°$, whereas F and Pl are affected significantly for the full range of α. We employ the anisotropic "polarizing" surfaces that are extensively introduced and analyzed in *Chapter* 6[1]. Also, this analysis is provided for signals without noise.

Figure 5.5 shows the correlogram of F, Pl and Q on various data sub sets. We provided the auto-correlation up to 88 lags, where 84 lags corresponds to a period of 7 solar days. The specific sub set that is used is illustrated in the panels below the correlograms. Our findings for the different sub sets are itemized below:

- For $\alpha < 60°$ the phase curve of Q showed to be affected by clouds. The corresponding correlogram in the upper left sub figure shows that the correlation in the signal Q exhibits no clear peak at the solar day nor at consecutive solar days. After a three day period the correlation of Q is consistently lower than that of F and Pl.

- For $35 < \alpha < 150$ we include both the region of phase angles that cover the primary rainbow and the region where Q showed to be fairly insensitive to the clouds (upper middle panel). Here, we restrict the phase angle range to $< 150°$ (or equivalently $> 30°$), which for example corresponds to a system at 40 light-years, with the planet at $1AU$ from its star and an IWA of $40\ mas$[2]. In this phase angle range a correlation at a solar day or consecutive days is more distinctive than the former case. The correlation for consecutive days becomes largest for F and approximately similar for Q and Pl.

- In the upper right panel we see a similar case the the former, but now without the primary rainbow phases. Overall the correlation for all parameters is higher than the former case, and we see that Q exhibits the highest correlated periods, with no false positives that relate to half an Earth day. After multiple consecutive solar days there occurs a clear divergence in the autocorrelation of Q compared to F and Pl.

- In the bottom left panel we show the correlogram that includes the "unobservable" phase angles near new phase. The correlation for high number of lags increases for Q, and less so for Pl and F, showing even a more clear diverging trend than in the former case.

- For $\alpha > 60°$ in the lower middle panel we model the region which is least affected by the variant cloud cover. As compared to the former case we still see a clear divergence, but for all parameters the correlation at multiple consecutive solar days is less.

[1] For our Lambertian surface approach we saw that no significant period was retrieved. By analyzing phase curves with a planetary model that include polarizing surfaces a much stronger period can be retrieved. We provided a figure equivalent to *Figure* 4.6, except with anisotropic "polarizing" surfaces, in *Appendix* F *Figure* F.1, to show that these polarizing surface models induce major oscillations on Q.

[2] See *Figure* 14 of Rossi and Stam (2017).

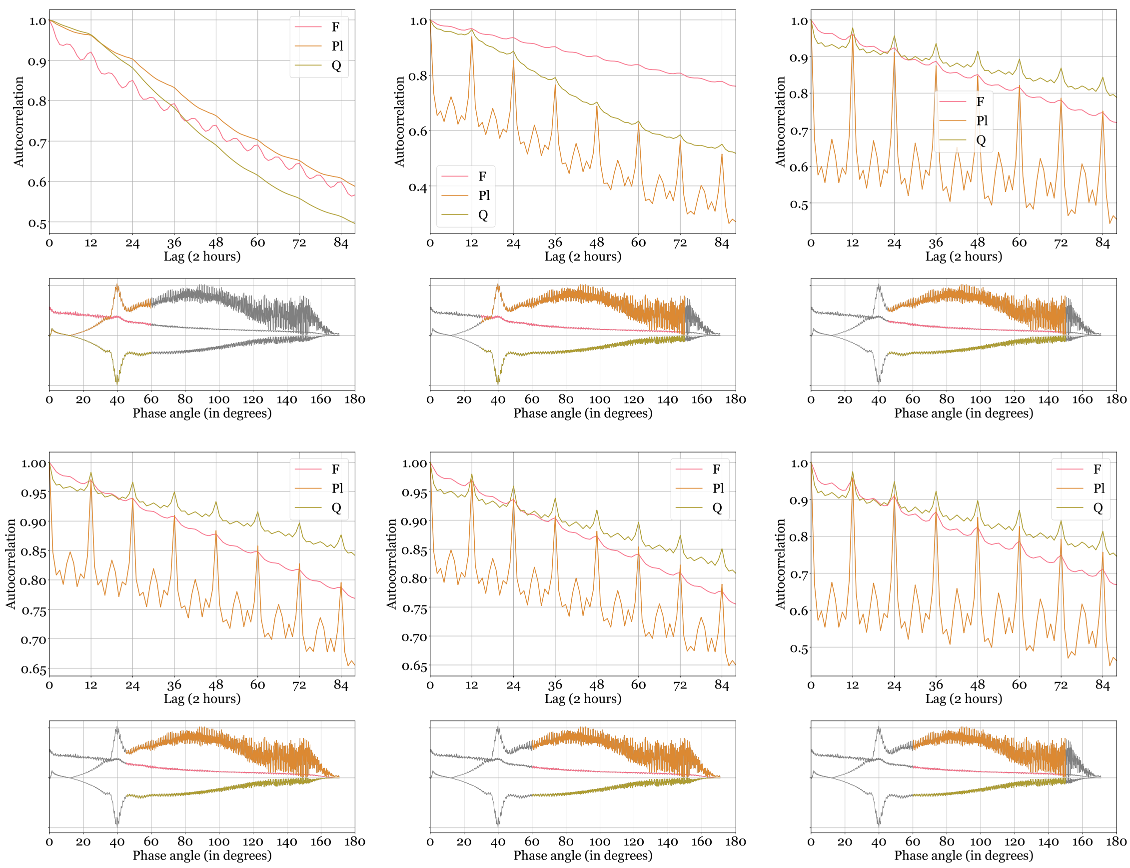


Figure 5.5: Correlogram of F, Pl and Q on various data sub sets, illustrated in the lower panel of each sub plot. On the x axis the number of lags is provided, where 1 leg corresponds to 2 hours in a solar day.

- If we restrict us to the "observable" part of phase angles that show low sensitivity to the variant clouds the diverging trends are less strong. The peaks at an Earth day or consecutive days is still clearly retrievable and also is the diverging trend.

In conclusion we showed that for small phases $\alpha < 45°$ the presence of a dynamic cloud deck can not be retrieved from the correlation in the signals of F, Q and Pl. Considering only $\alpha > 45°$ we retrieved a clear diverging trend in the autocorrelation of Q at multiple consecutive rotation periods for all cases as compared to F, but less so to Pl. This is only considered for the reflection of light at $\lambda = 550\ nm$ and for a temporal resolution of 2 hours, but it might provide as a valuable tool in retrieving the presence of dynamic clouds in an exoplanet atmosphere.

5.2. Dependency of total flux and polarized light on Earth's dynamic clouds

In the previous section we saw that the dynamic cloud cover of Earth has quite some effect on the diurnal variations for different days in the year. In this section we will look if there exists some correlations in the spectropolarimetric signals with the mean cloud properties of Earth. From *Chapter* 4 we know some interesting regions which we will further investigate in this section: the approximate rainbow peak, $40°$ phase angle, and the approximate peak of max polarization, at $90°$ phase (which may not be true for all wavelengths). For these two phase angles the diurnal variations with a 2 hour temporal resolution are computed for 53 different observation days in 2011, spaced by 2 weeks. This means that the simulation includes the full diurnal rotation of Earth and thus captures different continental distributions, as well as the seasonality in the clouds. Furthermore, we will look at three wavelengths, namely 350, 550 and $865\ nm$. In the subsequent section we will look at signals of F, Pl and Q only, because U does not provide any prominent correlation with any of the cloud parameters. Also, some cloud properties are left out for the same reason.

5.2.1. Correlations of Stokes elements with cloud parameters

We start with the behaviour of F. In *Figure* 5.6 we plotted F as function of the mean cloud fraction (CF), particle effective radius (CER) and top pressure (CTP) for $\lambda = 350, 550$ and $865\ nm$. The behaviour of the cloud optical thickness does not provide any prominent correlation and is therefore not included in this analysis. At short wavelengths we see a strong relationship for CF at both the rainbow peak and quadrature. For CER and CTP we do not observe such strong trends. With reasonable confidence we can say that an increase in F at both phases correlate to an increase in CF, regardless of the position of the cloudy pixels on the disk. Increasing the number of cloudy pixels increases the brightness of the disk, because as we know clouds are white and highly reflecting. At the rainbow the correlation is more pronounced, which may be due to the fact that cloud particles reflect light less intensively at phases around quadrature (see *Figure* 1 in Stam 2008*a*). We do, however, observe a small trend for CTP, at short wavelengths, of decreasing F for increasing CTP, caused by the high effectiveness of Rayleigh scattering, such that light has a shorter path length and thus the reflects more intensively from the clouds if the gaseous layer is geometrically thinner.
At green wavelengths the dominant correlation from CF gets somewhat diluted. Although, in theory we would expect this correlation to strengthen with decreasing effectiveness of Rayleigh scattering, we observe the opposite. For CTP we observe very minor differences with respect to short wavelengths and for CER we again see no clear trends.
At long wavelengths we observe a significant trend for CER. At quadrature a decreasing trend of F occurs with increasing CER, whereas at the rainbow peak the shape of this correlation seems to tend to a logarithmic relation. In *Section* 2.3.2 we saw a similar effect for a single pixel TOA reflection, where for increasing CER a continuously lower value of reflected flux was simulated. For CF we observe that the degradation of the trend at short wavelengths continued onward from $550\ nm$. If one looks closely to this anti correlation at both phases, one can see some division of data points into three areas that might suggest a major degenerescence of another cloud property at this wavelength. For CTP the small correlation that was present is completely gone at this wavelength.

The interesting part of this analysis is to look at the degree of polarization, because Rossi and Stam (2017) provided an extensive study, with slightly idealized Earth-like clouds in the form of a patchy cloud cover, into the variation of Pl with different values of CF. *Figure* 5.7 shows the variation of Pl as function of CF, CER and CTP. The relation with the cloud optical thickness is not provided, because of no clear trends. At $350\ nm$ we only see a significant correlation with CTP at the rainbow. This decreasing trend of Pl versus an increasing CTP is similar to that of F, but much more significant. Also, we observe a small correlation with the cloud particle effective radius at the rainbow, but much less significant.
If we move to $550\ nm$ a clear trend at the rainbow for CER is observed. At quadrature it is less prominent. CF shows different trends at the rainbow and quadrature, where at quadrature an increasing CF provides a lower Pl and at the rainbow this is correlated to a high Pl. An increase in CTP at quadrature causes and increase in Pl, whereas at the rainbow the data is randomly scattered.
For long wavelengths CER seems again to be the most dominant correlated parameter as we also saw for F. For most of the parameters at both phases we again see some division into three or more regions. We also find that CF at the rainbow is correlated more clearly than for F. CTP shows only randomly scattered data, but this can be easily related to the optically thin gas layers.

As we have seen in *Chapter* 4, Q shows limited variability induced by the clouds, except at the primary rainbow. In *Figure* 5.8 the behaviour of Q as function of CF, CER and CTP is provided. For $350\ nm$ at quadrature we see some significant trends for CF and CER. It has to be noted however that the variability in Q is relatively small near these phases. At the rainbow a less clear trend for both CF and CER is retrieved, but a clear trend for CTP. This trend is again related to the optical properties of the top gas layer, but in this case Q decreases in magnitude for increasing CTP. When referring back to the fact that Mie scattering of the clouds is generally strongly depolarizing this might seem odd, but at the rainbow we know that the clouds polarize reflected light more intensively. This causes the increase in

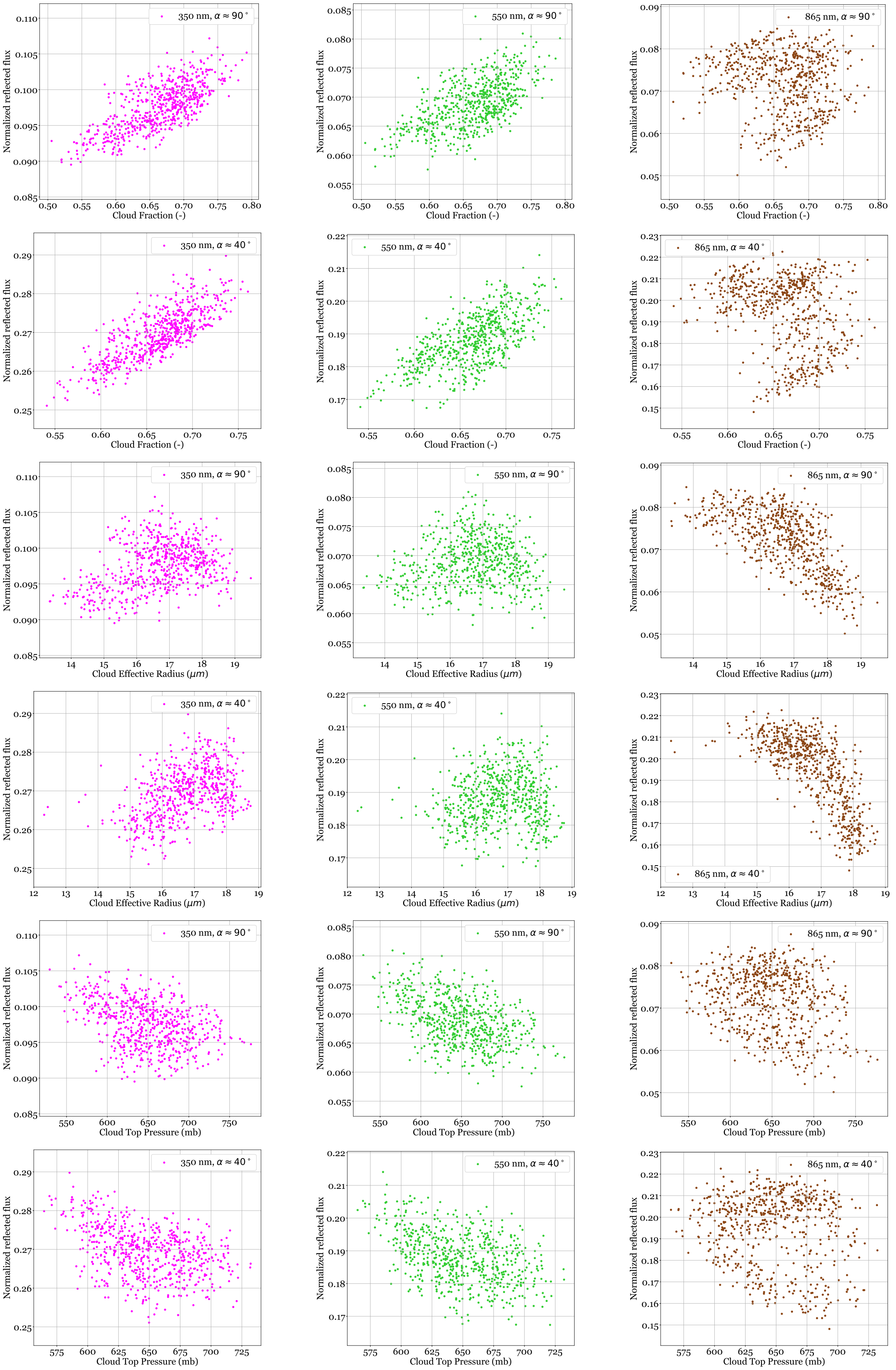


Figure 5.6: Relation between cloud parameters and total normalized reflected flux, for $\lambda = 350, 550$ and $865\ nm$ in the left, middle and right panel respectively. Alternatively from top to bottom we see F computed near $\alpha = 90^\circ$ and $\alpha = 40^\circ$, the approximate point of maximum polarization (in general) and peak of the primary rainbow, respectively. For each of these two cases we see from top to bottom the cloud fraction, cloud particle effective radius and cloud top pressure.

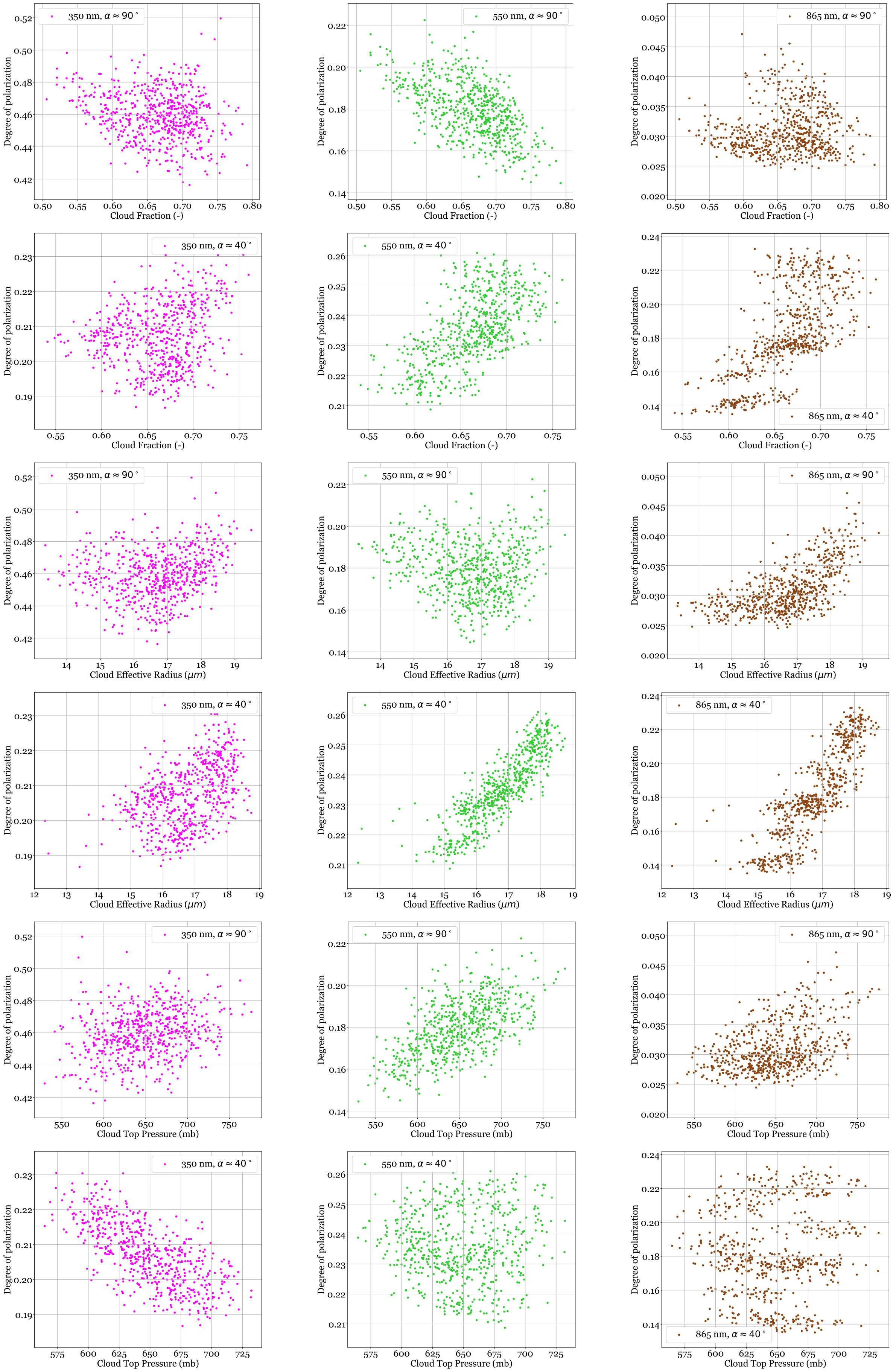


Figure 5.7: Similar to *Figure* 5.6, except for Pl.

Q with increase of CF.
At $550\ nm$ the variability near quadrature are very small, where the effect becomes stronger for longer wavelengths. For CF and CER at the rainbow we retrieve clear trends. These trends strengthen for longer wavelengths, because of the more dominant clouds. The correlation of CTP deteriorates for longer wavelengths. We see thus that Rayleigh scattering dilutes the correlation of the cloud fraction at short wavelengths and and its absence weakens the trend found for CTP at long wavelengths.

5.2.2. Retrieval of the shape of the light curves

From the former section we have seen several trends that can explain some behaviour in F, Pl and Q. Rossi and Stam (2017) (RS2017) showed that the shape of Pl exhibits some strong variability as function of CF. We can not directly compare our results, because RS2017 fixed the CF for a full phase curve and in our case this fraction varies constantly according to the daily observations. Also, RS2017 introduced a region of variability in which different spatial distributions of cloudy pixels in their planetary disk are modeled. In our case, we attempt the same by simulating different observation days for the two phase angles. Even with this region of variability, RS2017 could still distinguish between different cloud fractions by observing a monotone decrease of Pl at $\alpha =\sim 50-140^{\circ}$ for increasing CF, where the largest distinction can be made with values of 0.1-0.4. In our case we observe a values of $\sim 0.5-0.8$. For these fractions RS2017 could also observe a relative difference, but much smaller in absolute magnitude. Also, RS2017 only modeled $\lambda = 300$ and $500\ nm$, where we also introduced a near-infrared wavelength. For $\lambda = 500\ nm$ they found that different cloud fractions do not induce changes in the strength of the primary rainbow. In *Chapter* 4 and the former section we saw that in our case some variability is induced on this peak in the presence of varying cloud parameters. More specifically, this variability showed to be mostly induced by the different values of CER that we introduced in our model, but the CER is fixed in the analysis by RS2017. In the following part of this section we will look at the percentage of F, Pl and Q at 90° phase with respect to that at our reference phase of 40°, and attempt to retrieve changes in the shape of the phase curve that are correlated to CF, CER or CTP. 90° phase is chosen because this point can in most cases be considered as the maximum point of Pl. This can slightly deviate, but in our case we are mostly interested in the shape of the curves and not so in the exact position of the maximum polarization peak. The shape of the curve is more clearly defined in *Figure* 5.9, where in the left panel we observe a shape that corresponds to a fraction below 100%, in the middle panel a shape that corresponds to a fraction slightly higher than 100%, and in the right panel that of a fraction significantly higher than 100%. For the following plots we basically use the data from the scatter plots we presented in the former section and look at the relation between every possible combination of data points. In the following we refer to the fractional percentage by the fraction of F, Pl or Q.

Fraction of F

Despite the fact that the rainbow is much more varying in F than in Pl, we find some correlation between the fraction of F at the rainbow and quadrature phase, which are provided in *Figure* 5.10. In the upper three panels we observe clearly a decreasing correlation between CF and the fraction of F. In this case we expect the fraction to be higher than 100 as F decreases continuously with increasing phase angle. We see that for small CF at quadrature and high values at the rainbow we see a large fraction of F. In this region of the plot we reasonably constrain the high and low values of CF at the rainbow and quadrature, respectively, at $350\ nm$. For longer wavelengths we still see a significant correlation at $550\ nm$, but at the red wavelengths no clear trend is prominent and any fraction of F can be induced by a large set of CF at both phases.
For CF at 350 and $550\ nm$ one can not confidently constrain any value of CER at both the rainbow or quadrature. At $865\ nm$, however, we see a strong trend that constraints at least the mean CER at quadrature and to a lesser extend CER at the rainbow, because the red region extends more to lower effective radii at this phase.
In the lower panel we plotted all combinations of CTP, which show that there are slight trends visible. At short wavelengths low values of CTP at quadrature seem to correspond to a low fraction of F, whereas high values correspond to low CTP at the rainbow. For $550\ nm$ the trend strengthens somewhat, where high fractions move also to high top pressures at quadrature, whereas at $865\ nm$ this trend looks more diluted. For the fraction of F we observed that we can reasonably constraint high and

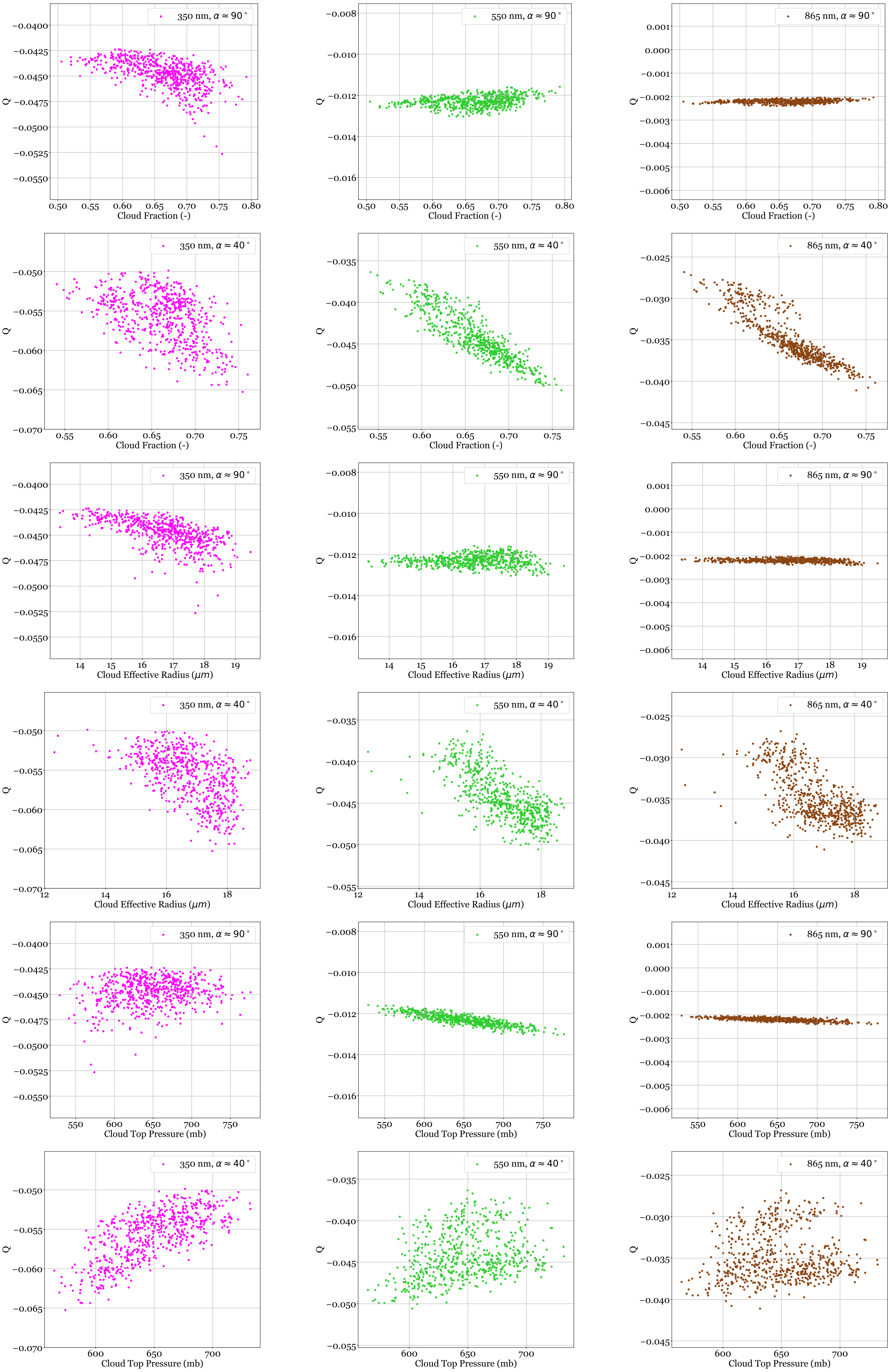


Figure 5.8: Similar to *Figure* 5.6, except for Q.

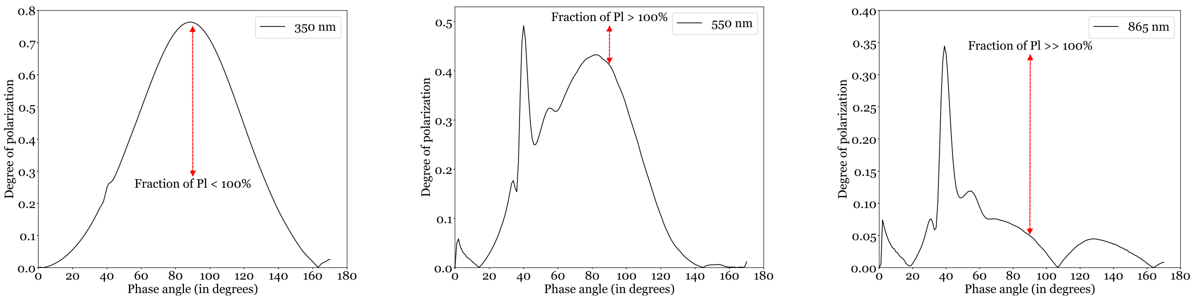


Figure 5.9: Definition of the fraction of Pl, similarly applied to F and Q.

low values of CF at both phases, and with less confidence the high and low values of CER at both phases.

Fraction of Pl

In *Figure* 5.11 the fraction of Pl is provided. For short wavelengths the fraction less than 100%, meaning that Pl at quadrature is higher than at the rainbow. At 550 and 865 nm we observe a fraction $> 100\%$, meaning that the "peak of maximum polarization" is less in absolute magnitude than the rainbow peak. In this case low and high values of CF at both phases are clearly restricted at 350 and 550 nm. At 865 nm any unique correlation of CF at any phase with the fraction of Pl is lost.
CER at both phases are not related to any unique fraction of Pl at 350 nm. By moving to longer wavelengths we start to see a clear trend that constraints in particular CER at the rainbow. It has to be noted that for good visibility the y-axis is inverted for the 3d plot at 865 nm.
For both phases we are able to constraint values of CTP reasonably well. At 350 nm we observe that CTP at the rainbow has a linearly like trend, where low CTP correspond to high fractions of Pl. At 550 nm the trend is more clear and the pattern is rotated to constrain CTP at quadrature, where no unique value of CTP at the rainbow can be related to a fraction of Pl. For red wavelengths we see a similar pattern, but much more diluted.

Fraction of Q

All fractions of Q are larger than 100%. For long wavelengths this fraction increases, and at red wavelengths we see fractions as high as 1800. For the fraction of F and Pl we were able to constraint values CF reasonably well. The fraction of Q shows very clear trends at 550 and 865 nm, essentially constraining high and low values of CF at the rainbow reasonably well. For 350 nm a similar pattern is observed, but includes much more dilution.
For CER heavily diluted patterns are found. They do show, especially for 550 nm an increasing trend that is only related to values of CER at the rainbow. So, from these results we could distinguish with some restricted accuracy high mean values of CER from low values at the rainbow for mutual different observations.
In case of CTP we only see a clear trend at 350 nm. For the fraction of Pl we saw that we could constrain CTP at quadrature with observations at 550 nm, but this trend is somewhat diluted. In case of the fraction of Q we a similar trend at 350 nm but much stronger and less diluted. For longer wavelengths this pattern vanishes quickly, from which we can hardly retrieve constraints for CTP.

Comparison to previous study

We compare our results with RS2017. In *Figure* 6 and 7 of RS2017 they show that at $\lambda = 500$ nm an increasing cloud fraction correlates with a decreasing degree of polarization at quadrature. For $\lambda = 350$ nm they also found such a trend, but much weaker to where it is almost not distinguishable anymore. As was described before these trends of decreasing Pl with increasing CF corresponds to an increase in fraction of Pl. In the upper left and middle panel of *Figure* 5.11 one can observe the corresponding results in our simulations. Despite the fact that the wavelengths do not agree exactly, we expect that this small difference does not change the conclusions qualitatively. In the

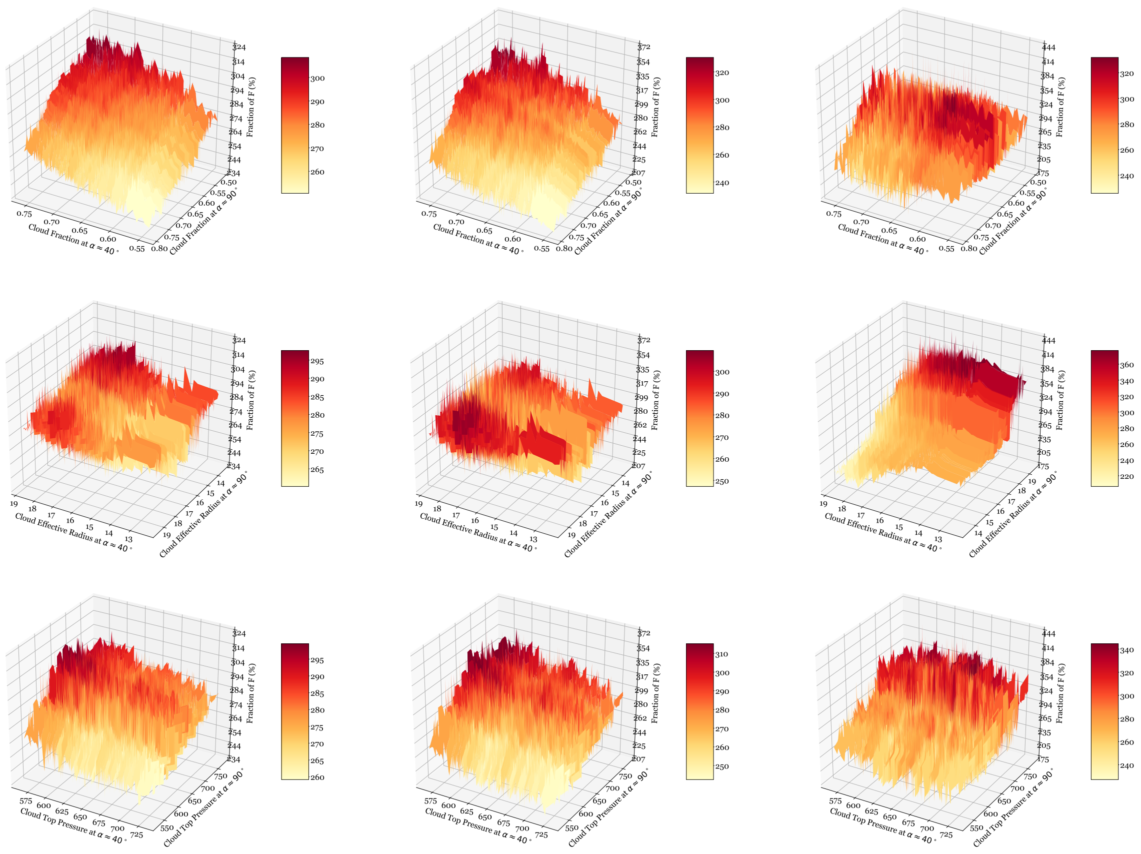


Figure 5.10: The fraction of F (z-axis) as function of all combinations between the cloud fraction, cloud particle effective radius and cloud top pressure at $\sim 40°$ phase (x-axis) and $\sim 90°$ phase (y-axis). The cloud parameters are retrieved from 53 simulations of diurnal variations with a resolution of 2 hours to fully include any seasonality in the clouds. In the top panel, middle and bottom panel one can see the cloud fraction, cloud particle effective radius and cloud top pressure, respectively. In the left, middle and right column we distinguish the wavelengths $350, 550$ and 865 nm.

upper left panel, which corresponds to *Figure* 7 in RS2017 shows a very ambiguous trend from which we can not qualitatively conclude that the fraction of Pl increases for increasing CF. We do however see that the largest fractions tend to the upper far corner in the plot, which corresponds to high cloud fractions at both phases. Alternatively, minimum values seem to be mostly present at the closest corner, corresponding to low cloud fractions at both phases. Also, one can see from the color scale that the fractions only change by $\sim 6 - 7\%$. Thus, unless the fact that the trend is not strong, the limiting cases do correspond to the results that RS2017 found. *Figure* 6 of RS2017 presents the same case at $\lambda = 500$ nm. From our results (the upper middle panel of *Figure* 5.11) we retrieve a clean trend that shows an increasing fraction of Pl to be correlated to increasing values of CF at quadrature. These results agrees well with the results of RS2017.

5.2.3. Parameter retrieval overview

To conclude the information in the former two sections we provide a cloud parameter retrieval overview in the form of a summary that entails the main correlations that we found and if and how they are influenced by other cloud parameters. An analysis of the latter is provided in *Appendix* E. In *Table* 5.1 this summary is provided. As a last note we must mention that by assuming Lambertian reflecting surfaces the true variability in Q might not be simulated properly. The consequence might be that some major dilution from the variability in the inhomogeneous surface distribution can occur in the relatively clear correlations that we found for Q.

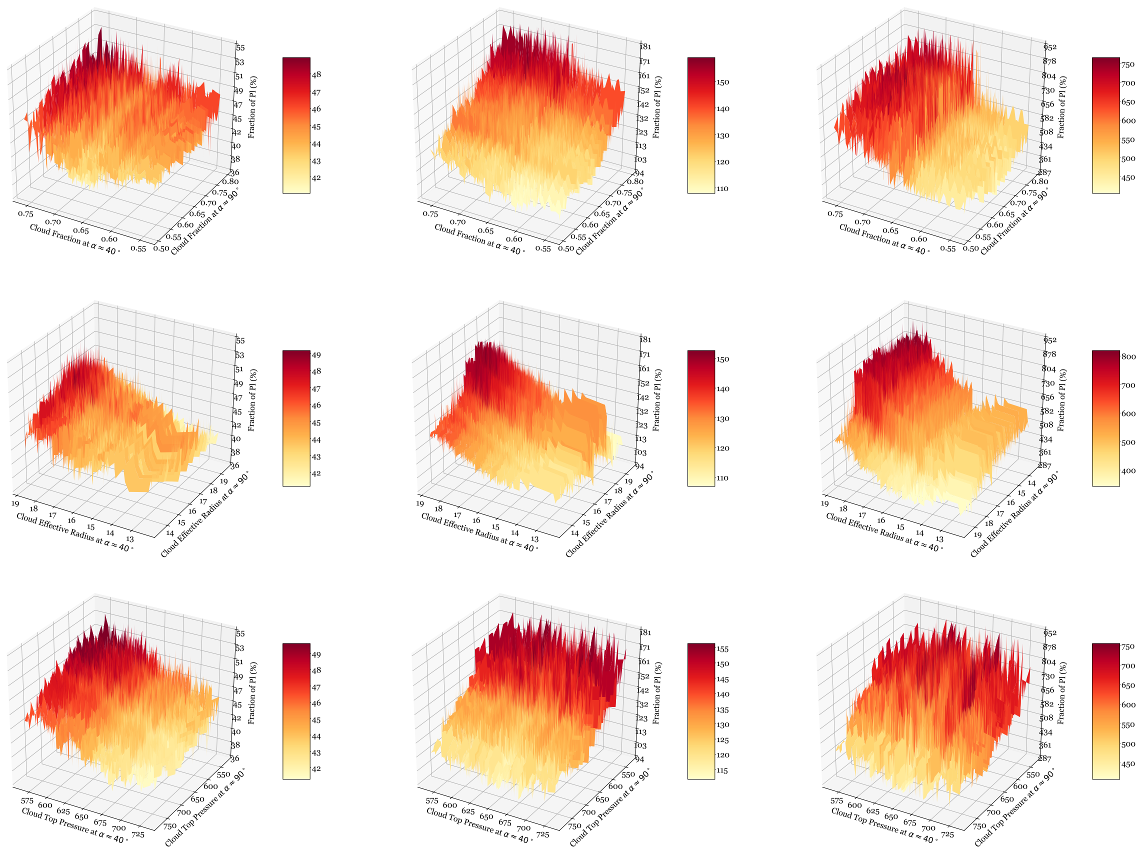


Figure 5.11: Similar to *Figure* 5.10, except for the fraction of Pl.

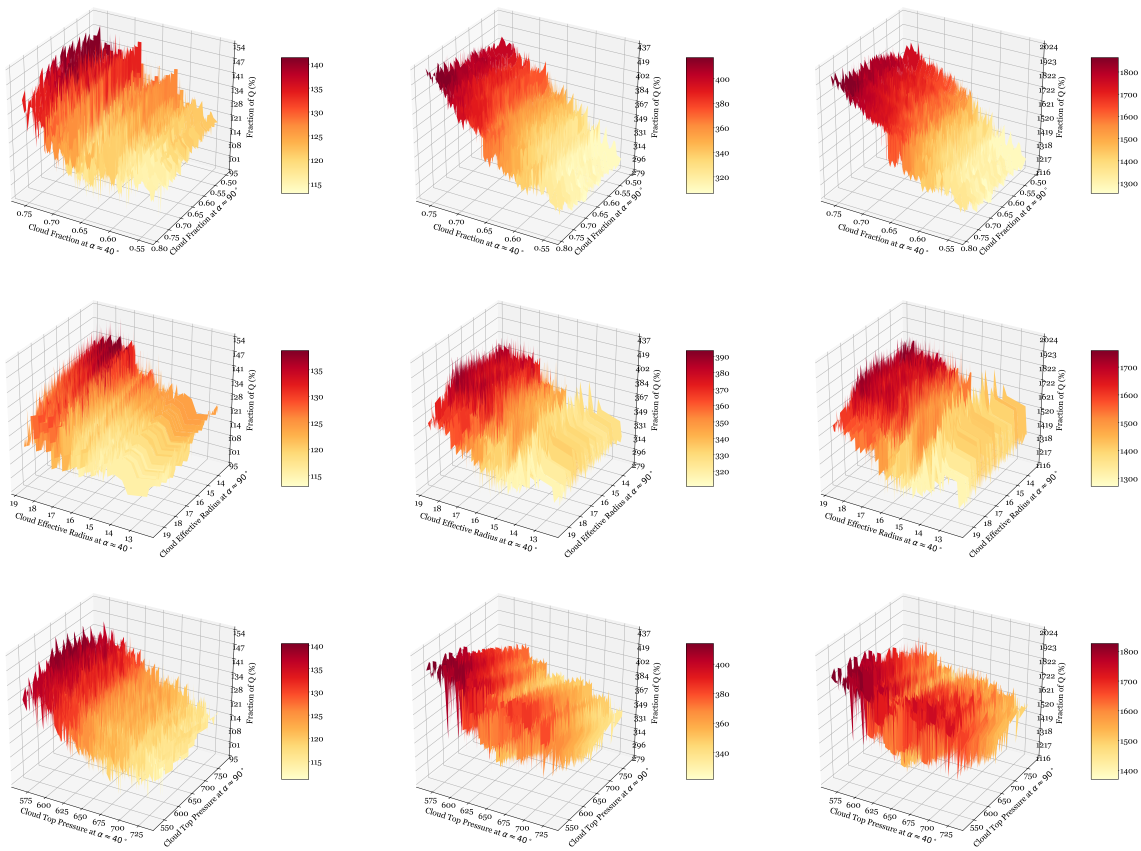


Figure 5.12: Similar to *Figure* 5.10, except for the fraction of Q.

Table 5.1: Overview of the retrieval strategy of mean cloud parameters for the normalized reflected flux, degree of polarization and normalized reflected polarized flux Q for $\alpha = 40$ and 90 degrees at $\lambda = 350,\ 550$ and $865\ nm$.

	Normalized reflected flux (F)
Cloud fraction	• The cloud fraction can be retrieved at $\lambda = 350nm$. • For both phases we retrieve a highly correlated trend. • From the shape of the curve we can confidently constraint high and low cloud fractions, but the absolute difference in the shape parameter is very minor. • Limited degenerescence from other cloud parameters. • The cloud fraction can also be retrieved at $\lambda = 550nm$, but with less confidence and more degenerescence than at 350 nm.
cloud particle effective radius	• The cloud particle effective radius can be retrieved at $\lambda = 865nm$. • For both phases we retrieve a highly correlated trend. • From the shape of the phase curve we can confidently constraint high particle sizes at $\alpha = 90°$ and low particle sizes at $\alpha = 40°$, where other sizes can be retrieved with more ambiguity. • Limited degenerescence from other cloud parameters.
Cloud top pressure	• The cloud top pressure can be retrieved at $\lambda = 550nm$. • For both phases we retrieve a moderately correlated trend. • From the shape of the phase curve we can only constraint high and low cloud top pressures with some ambiguities. • degenerescence from the cloud fraction.
	Degree of polarization (Pl)
Cloud fraction	• The cloud fraction can be retrieved at $\lambda = 550nm$. • For both phases we retrieve a correlated trend with some dilution. • From the shape of the curve we can confidently constraint high and low cloud fractions. • degenerescence from cloud particle effective radius at $\alpha = 40°$, and moderate degenerescence from both cloud particle effective radius and top pressure at $\alpha = 90°$. • The cloud fraction can also be retrieved at $\lambda = 865nm$, but with less confidence and more degenerescence than at 550 nm. • From the shape of the curve we can constraint high and low cloud fractions at $\alpha = 40°$. • High degenerescence from cloud particle effective radius at $\alpha = 90°$
cloud particle effective radius	• The cloud particle effective radius can be retrieved at $\lambda = 550$ and $865\ nm$. • For $\alpha = 40°$ we retrieved a highly correlated trend for both wavelengths. • From the shape of the phase curve we can confidently constraint particle sizes at $\alpha = 40°$, but ambiguities arise for high values at $\alpha = 90°$. • An increasing degenerescence of the cloud fraction at $\alpha = 40°$ for longer wavelengths. At $\alpha = 90°$ we see a continuous degenerescence from the cloud fraction.
Cloud top pressure	• The cloud top pressure can be confidently retrieved at $\lambda = 550nm$. • For $\alpha = 90°$ we retrieve a moderately correlated trend. • From the shape of the phase curve we can only confidently constraint the cloud top pressure at $\alpha = 90°$ • Some moderate dilution of the cloud particle effective radius at $\alpha = 40°$ is observed. • At $\lambda = 350\ nm$ one is able to constraint high and low top pressures from the shape of the phase curve, but the absolute difference in the shape parameter is very minor.
	Polarized flux Q
Cloud fraction	• The cloud fraction can be retrieved at $\lambda = 550$ and $865\ nm$. • For $\alpha = 40°$ we retrieve highly correlated trends. • From the shape of the curve we can confidently constraint the cloud fraction at $\alpha = 40°$ at both wavelengths considered. • At both phases and wavelengths the cloud particle effective radius only slightly dilutes the shape parameter. • Moderate degenerescence from the cloud particle effective radius.
cloud particle effective radius	• The cloud particle effective radius can be retrieved with some ambiguity at $\lambda = 550nm$. • We only find a moderate trend for $\alpha = 40°$. • From the shape of the phase curve we can identify the low and high values only at $\alpha = 40°$. • Significant degenerescence form the cloud fraction.
Cloud top pressure	• The cloud top pressure can be retrieved at $\lambda = 350$. • For $\alpha = 40°$ we retrieve a correlated trend. • From the shape of the curve we can confidently constraint the cloud top pressure at $\alpha = 40°$. • Some dilution from the cloud particle effective radius.

6

The scattered light curves from an Earth-like exoplanet with polarizing surfaces

In this chapter, we present simulations of disk-resolved and disk integrated light signal of the planetary model presented in *Chapter* 3 in combination with more realistic surface models for oceanic, vegetated and desert covers. We start by introducing the model atmosphere that is used in these simulations (*Section* 6.1). The anisotropic polarizing surface models that are incorporated in the planetary model are introduced in *Section* 6.2. In *Section* 6.3 we present resolved disks at the same geometry and at the same spectral region as the disks presented in *Section* 4.1. This allows us to provide a more illustrative explanation of how these new surface models impact our simulated spectropolarimetric results. In the last section, *Section* 6.4, we provide the spectropolarimetric phase curves of this extended model, and a thorough discussion on the impact of all introduced surface models.

6.1. The model atmosphere

We model the gaseous atmosphere in the same fashion as described in *Section* 3.3.1 and the clouds as described in *Section* 3.3.2. The gaseous atmosphere in the atmosphere-ocean system uses a wavelength dependent air refractive index, which is based on the dispersion formula of Peck and Reeder (1972) (Trees 2018). The different dispersion formulas of Ciddor (1996) and Peck and Reeder (1972) only show minor differences in the ultra-violet, visible and near-infrared wavelength domain (see Ciddor 1996). Additionally, scattering cross sections were compared and showed good agreement.The pixel models with underlaying vegetation cover are directly computed with scattering cross sections from the PyMieDAP code. The pixel models with underlaying desert surface are computed with PyMieDAP.

6.2. The polarizing Earth-like surface models

In this section we provide a brief explanation of the anisotropic polarizing surface types that are used in the planetary model. In this regard also a description of how the models are implemented in our Earth-model is provided. This is especially important to restrict the total computation time, by for example only adding these polarized surfaces under clear and thin clouds as one might expect that under thick clouds the reflection of these surface are insignificant. The discretization of the model surfaces is provided in *Table* 6.1. In the following sections we will describe in detail the models for the vegetated, oceanic and desert surface types.

6.2.1. The polarizing vegetation model

The polarizing vegetation code developed by Cheung (2018) that we use in our planetary model is a combination of a Bidirectional Reflectance Distribution Function (BRDF) model and a Bidirectional Polarized Distribution Function (BPDF) model. Several BRDF and BPDF models exist that can be used. In order to find the most suitable BRDF and BPDF models the following criteria were set:

Table 6.1: IGBP classification applied to the model surface

No.	IGBP Classification	Our Classification	Ocean/Land
0	Water	Ocean	Ocean
1	Evergreen needleleaf forest	Vegetation	Land
2	Evergreen broadleaf forest	Vegetation	Land
3	Deciduous needleleaf forest	Vegetation	Land
4	Deciduous broadleaf forest	Vegetation	Land
5	Mixed forest	Vegetation	Land
6	Closed shrubland	Vegetation	Land
7	Open shrubland	Soil	Land
8	Woody savannas	Vegetation	Land
9	Savannas	Vegetation	Land
10	Grasslands	Vegetation	Land
11	Permanent wetlands	Soil	Land
12	Croplands	Vegetation	Land
13	Urban and built-up	Soil	Land
14	Cropland/natural Vegetation mosaic	Vegetation	Land
15	Snow and Ice	Snow	Land
16	Barren or sparsely vegetated	Soil	Land

1. Anisotropic reflection must be accounted for;
2. Different types of vegetation must be able to be simulated;
3. The total flux and/ or polarized flux need to be modeled in the spectral domain of the visible and near-infrared wavelength regions.

From these criteria one BRDF and one BPDF model is chosen to be implemented into the vegetation code. The BRDF model developed by Roujean et al. (1992) is used to simulate the anisotropic reflected flux of several types of vegetation, where mutual shadowing is not taken into account, i.e. the shadows projected by the trees do not overlap each other. With the use of the k-parameters provided in the table of Roujean et al. (1992), different types of vegetation, such as deciduous forest, pine forest, steppe, grass lawn and more as well as plowed fields can be simulated. These k-parameters are given in the near-infrared (NIR) and visible (VIS) wavelength region, i.e. $580-680\ nm$ and $730-1100\ nm$. A smoothstep interpolation method is used to obtain the k-parameters between $680-730\ nm$ as discussed in Cheung (2018). Extrapolation is used to account for the ultra-violet wavelength region.

In order to account for the polarization induced by vegetation, the polarized reflection model developed by Maignan et al. (2009) is used. According to Schaepman-Strub et al. (2006), the BPDF model is then obtained by dividing the polarized reflection model with π. This BPDF model is a linear, one parameter-model, which is a simplification of the non-linear, two-parameter model from Nadal and Breon (1999). It assumes that the polarized reflection is reflected specularly, which introduces the polarized Fresnel function component to the *Maignan model*. By using the normalized difference vegetation index (NDVI) and the α-parameter provided in the table of Maignan et al. (2009), the polarized flux of different types of vegetation is calculated. The parameters that are used are listed in *Table* 6.2. As seen from *Table* 6.2 only the vegetation types deciduous forest and steppe are taken into account.

Table 6.2: Parameters of our standard atmosphere and vegetation. Unless stated otherwise, the values listed in this table are used.

Deciduous Forest Parameters	Symbol	Value(s)
k-parameters (NIR)	k_0, k_1, k_2	40.0; 4.0; 29.5
k-parameters (VIS)	k_0, k_1, k_2	3.0; 0.0; 8.7
α-parameter	α	6.87
NDVI	v	0.8
Steppe Parameters		
k-parameters (NIR)	k_0, k_1, k_2	35.6; 5.6; 21.7
k-parameters (VIS)	k_0, k_1, k_2	26.6; 5.0; 5.9
α-parameter	α	6.66
NDVI	v	0.3

The vegetation code also follows the doubling-adding method from de Haan et al. (1987*b*). Since the doubling-adding method only considers mirror symmetric functions with respect to the planetary scattering plane, the *Roujean model* is altered to satisfy this condition. A detailed description is given in Cheung (2018). In order to implement a fully vegetated surface into the doubling-adding method, the *Roujean model* and *Maignan model* are inserted into the reflection matrix $\mathbf{R}(\mu_0, \mu, \Delta\phi)$, which then simulates the reflection of light that is reflected anisotropically and gets polarized from vegetation. The *Roujean model* is put on the (1,1) element and the *Maignan model* in the elements (2,1) and (3,1) of the reflection matrix, with the relationships (Litvinov et al. 2011)

$$Q = -\text{BPDF}\cos(2\beta) \tag{6.1}$$

$$U = \text{BPDF}\sin(2\beta) \tag{6.2}$$

where BPDF is the *Maignan model* and β the rotation angle from the planetary scattering plane to the local meridian plane. Since no rotation is required $\beta = 0$. For more details about the vegetation reflection model, we refer to Cheung (2018).

Figure 6.1 shows the phase curves of the F, Pl and Q in the upper left, upper right and bottom panel, respectively, for the $\lambda = 350,\ 443,\ 550,\ 670$, and $865\ nm$ of a planet covered fully with deciduous forest. For these curves a purely gaseous atmosphere is assumed, i.e. there are no cloud layers present. F is lowest at $670\ nm$ for $0° < \alpha < 100°$ and increases with decreasing wavelengths except at $865\ nm$, which results in the highest reflected F. From the spectrum of vegetation, provided in *Figure* 3.2, one sees that at $670\ nm$ vegetation has the lowest albedo, which explains why at this λ the F attains the lowest reflection. The second, third and fourth lowest surface albedo is, respectively, observed at $350,\ 443$ and $550\ nm$. One would therefore expect the same relationship in the flux-phase curve. However, as seen in *Figure* 6.1, the second lowest reflected F is observed at $550\ nm$ followed by 443 and $350\ nm$. These differences are caused by the gaseous atmosphere that is accounted for on top of the vegetated surface. Due to the high albedo of vegetation at $865\ nm$ we obtain a high reflection in F.

For Q the polarized reflection is lowest at $865\ nm$ due to ineffective Rayleigh scattering, which thus results in the lowest Pl at $865\ nm$. At $350\ nm$ one sees that Q is the highest. However, due to also having a high reflected F, Pl is not largest at $350\ nm$ but at $443\ nm$. One also sees that at $350,\ 443,\ 550$ and $670\ nm$ the peak in Pl occurs around $\alpha = 90°$, while at $865\ nm$ the peak is shifted towards $\alpha = 130°$. This shift in peak is caused by the diffuse scattering of light in the atmosphere at $\alpha > 100°$. Furthermore, in Q there is no color alternation, whereas for F we did see some alternation because of the high red-edge reflection. Pl exhibits several color alternations, but all occurring at different α.

6.2.2. The polarizing ocean model

For the ocean pixels we use the atmosphere-ocean model of Trees (2018). That is, the ocean consists of a wind-ruffled air-water interface, whose roughness as a function of wind speed is determined by the isotropic wind slope distribution model of Cox and Munk (1954). We use the shadowing function of Smith (1967) and Sancer (1969) to account for the energy abundance across the rough air-water interface at grazing angles caused by neglecting wave shadowing (see also Tsang et al. 1985; Zhai et al. 2010). The elements of the reflection matrix of the air-water interface are verified with the bidirectional reflection code of Mishchenko and Travis (1997), which is available on `https://www.giss.nasa.gov/staff/mmishchenko/brf/`. [1] From the reflection and transmission matrices, we computed the energy balance across the air-water interface and obtained a similar plot as Fig.4 in Nakajima (1983). The interface matrices are normalized through the division by the remaining energy deficiency, which is caused by the ignorance of multiple scattering of light between the wave facets (Nakajima 1983). The ocean parameters are listen in *Table* 6.3.

The water-leaving reflection accounting for the ocean albedo is computed with the adding-doubling method of de Haan et al. (1987*b*), analogous to scattering computations in the gaseous atmospheric

[1]The analytic solution of the elements of the reflection matrix of the air-water interface for illumination from below, and the elements of the transmission matrices for illumination from above and below are verified with the equations of Zhai et al. (2010).

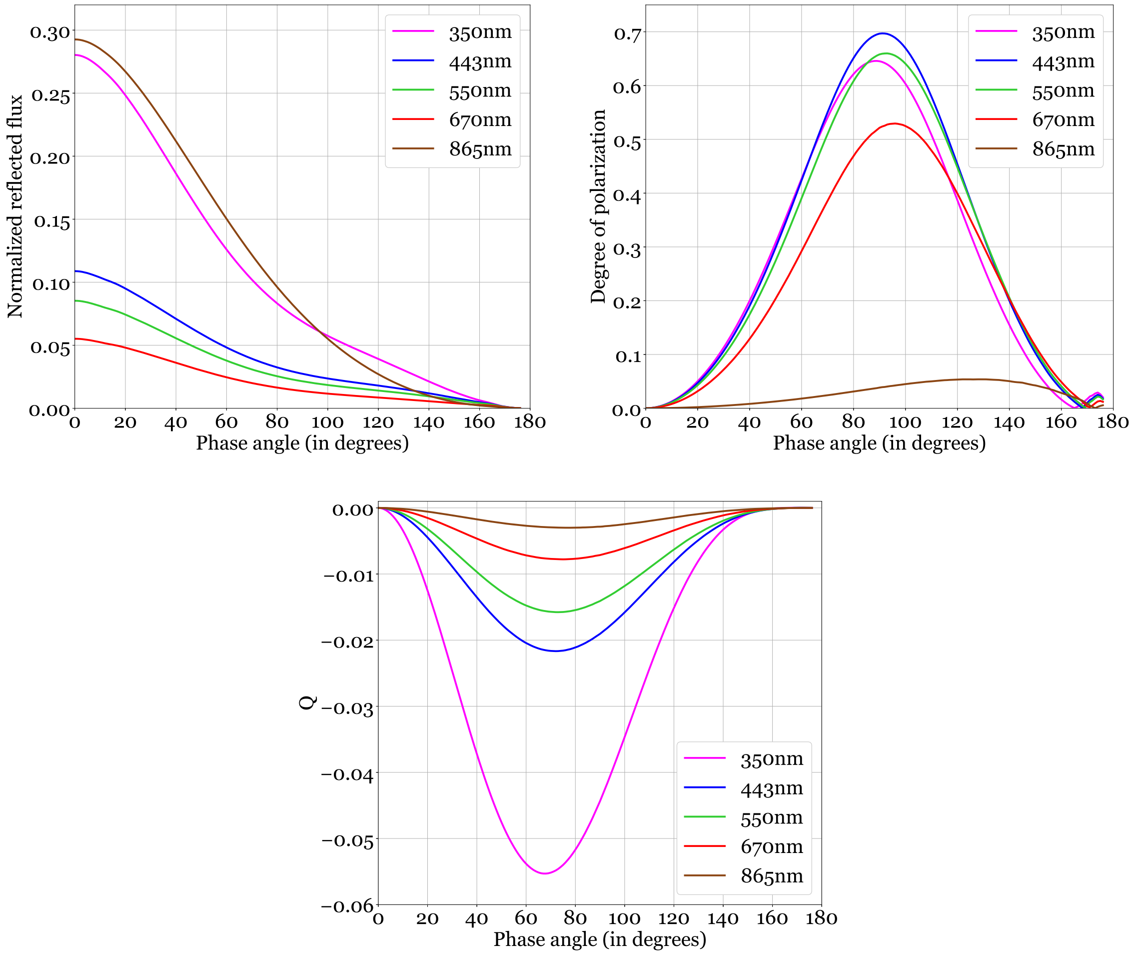


Figure 6.1: Phase curves from a homogeneous planet fully covered with a ”polarizing” vegetated surface cover, more specifically that of deciduous forests. On top of this surface we modeled only a gaseous atmosphere and no clouds. We provide the normalized reflected flux, degree of linear polarization and polarized flux in the left, middle and right panel, respectively.

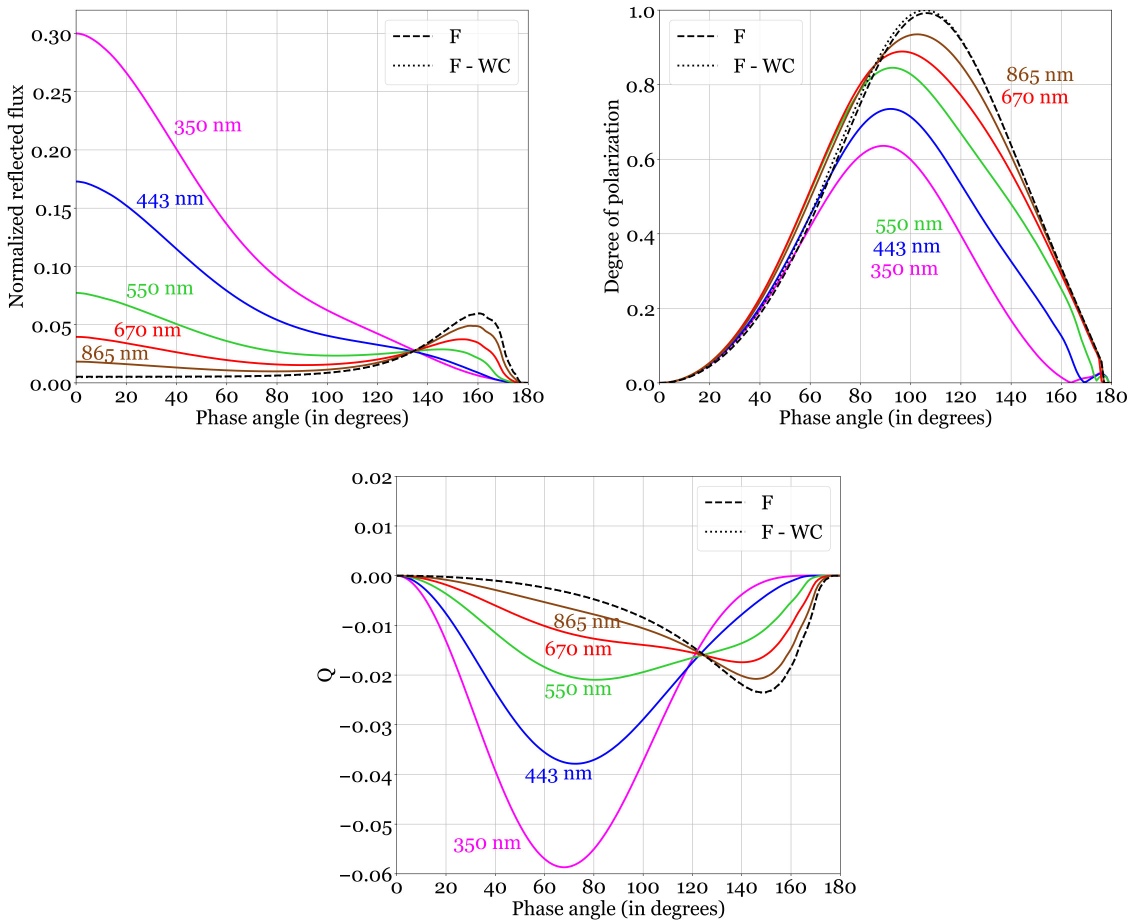


Figure 6.2: *Figure* 7.6 of Trees (2018). Disk-integrated flux, degree of polarization and polarized flux in terms of Stokes parameter Q as a function of planetary phase angle α of light reflected by the ocean planet with a gas atmosphere. Also, the rough Fresnel interface (F) and rough Fresnel interface without whitecaps (F - WC) are drawn, where the surface pressure is set equal to 0 bar and the subinterface ocean scattering is neglected. The wind speed is 7 m/s for all curves. The lines for F and $F - WC$ can hardly be distinguished.

Table 6.3: Parameters of our standard atmosphere and ocean. Unless stated otherwise, the values listed in this table are used.

Ocean Parameters	Symbol	Value(s)
Wind speed [m/s]	v	5.0; 7.0
Foam albedo	a_{foam}	0.22
Depolarization factor	δ_{w}	0.09
Air refractive index just above air-water interface	n_1	1.0
Water refractive index just below air-water interface	n_2	1.33
Chlorophyll concentration [mg/m^3]	[Chl]	0
Ocean depth [m]	z	100
Ocean bottom surface albedo	a_{bs}	0

layer. Thus, the ocean water is divided into a stack of homogeneous layers, and we assume an ocean depth of 100 m with a black Lambertian surface underneath representing the ocean bottom. Following Chowdhary et al. (2006), we use the scattering matrix for anisotropic Rayleigh scattering of Hansen and Travis (1974*b*), and a depolarization factor for pure seawater of 0.09, which was measured by Morel (1974). We use the wavelength dependent scattering coefficient of pure seawater tabulated in the work of Smith and Baker (1981) and the wavelength dependent absorption coefficients for pure seawater of Pope and Fry (1997).[2] We do not include hydrosols (e.g. phytoplankton, detrituts and bubbles) in the water, which would require a proper determination of the hydrosol scattering matrix elements (see Chowdhary et al. 2006). Our ocean thus belongs to the clearest natural waters. However, we compared the recomputed subinterface ocean albedo with the ocean albedos computed with the bio-optical model for Case 1 Waters of Morel and Maritorena (2001) for various chlorophyll concentrations and find realistic ocean albedos that correspond to low chlorophyll concentrations.[3] The final ocean reflection is a weighted sum of the clean ocean reflection (as described in this paragraph) and the Lambertian (i.e. isotropic and non-polarizing) reflection from wind-generated foam, also known as whitecaps, with an effective foam albedo as a function of wind speed taken from Koepke (1984). The weighting factor for the whitecap reflectance is determined by the whitecap fraction of Monahan and Ó Muircheartaigh (1980). Note that the air refractive index used for the Fresnel computations is set equal to 1.0, while the air refractive index to compute the scattering cross section of the gas molecules varies with wavelength, as explained in *Section* 6.1. For more details about the ocean reflection model, we refer to Trees (2018).

The upper, middle and lower panel of *Figure* 6.2 show the F, Pl and Q as a function of planetary phase angle α of light reflected by the ocean planet with a gaseous atmosphere for a wind speed of 7 m/s. As explained in Trees (2018), the increased flux F at big phase angles ($\alpha \approx$ 90°-180°) increases with increasing wavelength because of the glint in the ocean. That is, at big phase angles the Fresnel reflection from the wave facets is stronger due to the bigger reflection angles and the glint fraction of the illuminated disk, which has a crescent shape at these phase angles, is increased. Because the light beams of longer wavelengths are less effectively scattered by Rayleigh scattering, they can penetrate deeper into the atmosphere as compared to light beams of shorter wavelengths. Thus, the beams of longer wavelengths are more likely to penetrate through the atmosphere at these big phase angles, to be reflected by the ocean, and to reach the top of the atmosphere again without being multiply scattered in the atmosphere.

Because the ocean reflection contribution is more dominant at long wavelengths, a shift of the peak of the degree of polarization towards bigger phase angles (in the direction of two times the Brewster angle) may be observed (middle panel of *Figure* 6.2), as explained in Trees (2018), and was also found by Zugger et al. (2010*b*, 2011*a*,*b*). Trees (2018) showed that this shift is sensitive to surface pressure and can hardly be detected for a surface pressure of 5 bar when the wavelengths $350,\ 443,\ 550,\ 670$ and $865\ nm$ are used. However, the degree of polarization is not a measure of the actual polarized flux. Trees (2018) mentioned that the intersection point in the polarized flux Q could potentially be

[2]The lower limit of the range from 380 nm to 350 nm is extended by using the additional data of Sogandares and Fry (1997). Between 727.5 nm and 800 nm we use the values of Smith and Baker (1981). Above 800 nm, we assume there is no water-leaving light and the ocean reflection reduces to the reflection by the interface only.

[3]The recomputed ocean albedos at a solar zenith angle of 30° are 0.0920, 0.0870, 0.0058, 0.0032 and 0 at the wavelengths $350,\ 443,\ 550,\ 670$ and $865\ nm$. These values correspond to chlorophyll concentrations between 0 and 0.1 mg/m^3 in Morel and Maritorena (2001).

used for detecting oceans on exoplanets. More specifically, this intersection point resembles a point at which the (polarized) reflected light at different λ alternate in order of absolute strength, i.e. the color of the exoplanet changes. This signature could be observed in the presence of an ocean for surface pressures op to 10 bar (provided that at least 2 significantly different wavelengths are used, preferably of which one is in the near-infrared) and cloud fractions up to 95%, while it never occurred in the absence of an ocean. Trees (2018) also found an apparent intersection of the phase curves of the degree of polarization when modelling clouds with different cloud fractions above the ocean, however, such an intersection was also found in case of a substellar cloud with a black Lambertian surface underneath the atmosphere,[4] although the color reversal was not very apparent for the substellar cloud model. In this chapter, we will investigate whether the intersection in the phase curves of the Stokes parameter Q still occurs for an Earth-like exoplanet model, thus including Earth-like continents with appropriate land cover types.

Wind speed statistics are based on half year's worth of data from 01/01/2011 till 30/06/2011, the same interval as the observation data files that we use for our cloud observations. The wind speed data is provided as a near global six hour data set every day at an altitude of 10 meters above the ocean surface, retrieved from *Cross-Calibrated Multi-Platform (CCMP) Wind Vector Analysis Product*[5]. An example of this data set is provided in the right panel of *Figure* 6.3. The spatial extend of the data set is limited in latitude by $\pm 80°$, with a spatial resolution of $0.25°$. In the left panel we provide the latitudinal dependence of the mean wind speed in m/s, showing an approximate mean between 7-8 m/s. The reflection behaviour of the glint, more specifically the spatial extend of the glint, which has no real physical boundary, varies with the wind speed as mentioned before and as is extensively described in Trees (2018). In *Figure* 6.4 the spatial extend of the glint is presented for various phase angles and a wind speed of $7\ m/s$. These homogeneous disks are computed with a pixel model consisting of an Earth-like gaseous atmosphere without clouds on top of the polarizing ocean model. Outside the glint the effect of wind speed is negligible. From these disks we can see that the significant reflection from the glint does not exceed $\pm 30°$ latitude. To discretize the wind speed we therefore focused on this region. Form the left panel of *Figure* 6.3 we see that for $\pm 30°$ latitude the mean wind speed varies between 5 and 7 m/s. The difference in magnitude and spatial extend of the glint for 5-6 or 6-7 m/s is small. Considering also computation time and storage we will limit the number of bins to two. Hence, we will discretize the wind speed by 5 and 7 m/s (*Table* 6.4).

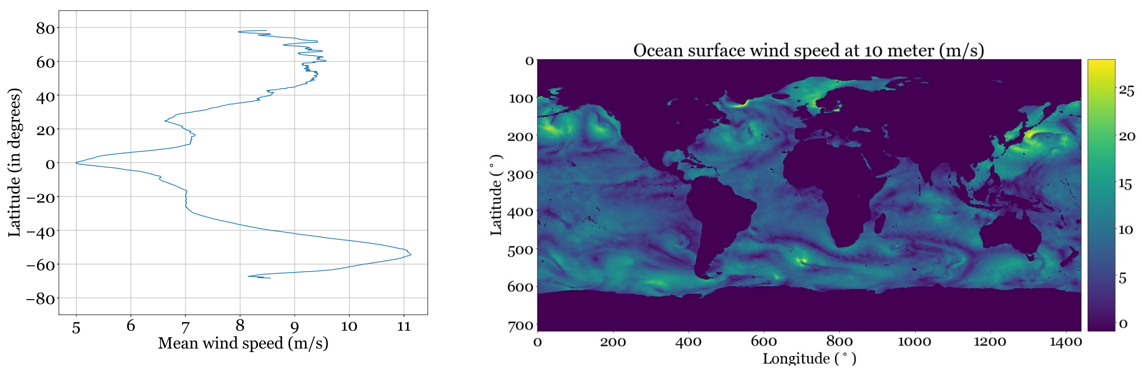


Figure 6.3: Left panel: mean wind speed as function of latitude. This mean is computed from all data spanning 01/01/2011 till 30/06/2011. Right panel: wind speed observation.

6.2.3. The polarizing desert model

For the vegetation and ocean models we were able to use semi-empirical models. For the desert no such model has been created that characterizes both the anisotrope reflection as well as polarized

[4]A black Lambertian surface underneath the gaseous atmosphere allows for a high degree of polarization in the red part of the visible spectrum when the substellar cloud has been rotated away from the observer, because of the singly scattered (and limited multiply scattered) light at these long wavelengths, see also Stam (2008*b*).

[5]Further details can be found on: `http://www.remss.com/measurements/ccmp/`

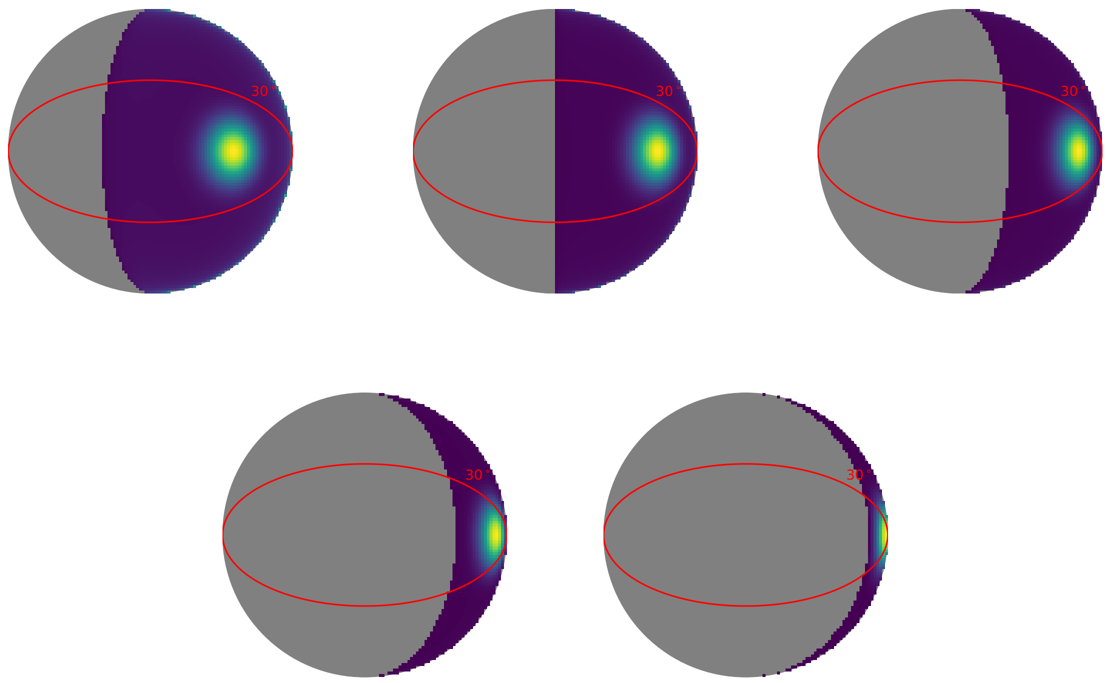


Figure 6.4: Spatial extend of the glint for $\alpha = 70, 90, 110, 130$ and 150° from left to right, respectively. The atmosphere is Earth-like and without clouds. The red line represents the small circles at $\pm 30^{\circ}$ latitude.

Table 6.4: Discretization bins of the ocean wind speed (Wspd) with associated bin values in m/s.

Bin number	1	2
Bin boundaries	0<Wspd≦ 6	6<wspd
Bin values	5	7

reflection. In the attempt to obtain at least some approximation for the reflected polarized fluxes we use a scattering matrix from an "Olivine S" sample. The scattering matrix elements are measured as a function of a finite number of scattering angles, which are freely available in the *Amsterdam light scattering database* (Munoz et al. 2012). In (Moreno et al. 2006) they attempted to approximate the scattering matrix of the "Olivine S" sample for a distribution of irregularly shaped compact particles using a Discrete Dipole Approximation (DDA) method. They found that the synthetic scattering matrix from a size distribution of synthetic compact irregularly shaped particles fits the measured scattering matrix quite well. The approximate scattering matrix for incident light beams at $\lambda = 663\ nm$ is available in the form of *Expansion Coefficients*. In PyMieDAP one is able to load such *Expansion Coefficients* into an user defined layer. As Moreno et al. (2006) did not consider absorption of the particles, the provided single scattering albedo (SSA) is set to 1. In order to model a desert surface with the scattering matrices of irregularly shaped Olivine S particles, we will define the following

1. The *Expansion Coefficients* are loaded into the bottom layer of the pixel model;
2. The bottom pressure of the layer is $1.001\ bar$ and the top pressure is $1\ bar$. This top pressure is in accordance with the planetary model we presented in *Chapter* 3;
3. We model no Rayleigh scattering in this layer;
4. To approximate this layer as a surface, we define it as an extremely thick aerosol layer, that only comprises of the Olivine S particles. We will use an optical thickness of 100. For all wavelengths considered any higher optical thickness provided the same reflection properties;
5. We model the wavelength dependence by altering the SSA.

We changed the SSA to account for the wavelength dependent reflection of the anisotrope surface. By changing the SSA, the hemispherical albedo of our local atmosphere changes. Ultimately we fitted this albedo to the albedo of Entisol that we obtained earlier from the Aster spectral database. This Lambertian Equivalent Reflection (LER) is defined as the total hemispherical reflection at an incidence angle of $10°$. In *Table* 6.5 the fitted SSA's are listed for the corresponding wavelengths.

Table 6.5: Fitted Single Scattering Albedos (SSA's) to approximate dust aerosols with empirically obtained scattering coefficients by Moreno et al. (2006). The scattering coefficients were fitted to Lambertian Equivalent Reflectance (LER) albedos from Aster Spectral Database.

Wavelength $[nm]$	SSA [-]	LER [-]	Model albedo [-]
350	0.235	0.010604	0.010613
443	0.6103	0.05319	0.053186
550	0.82353778	0.139387	0.139390
670	0.89436	0.215352	0.215391
865	0.9279	0.279383	0.279351

In *Figure* 6.5 we present the disk integrated phase curves for a homogeneous planet fully covered with the polarizing desert model with a gaseous atmosphere as described in *Chapter* 3, but without clouds. We provide F, Pl and Q in the upper left, upper right middle and bottom panel, respectively. For 350 nm one can see that F at $0 \leq \alpha \lesssim 60$ exceeds that of the longer wavelengths, which is caused by the highly effective Rayleigh scattering of the gaseous layer on top of the surface. At big phases angles, near new phase, F at 350 nm becomes less than at longer wavelengths. This is mainly caused by the diffuse scattering of light as a result of the gaseous atmosphere. For longer wavelengths, except for 443 nm and 550 nm near full phase, we see a clear separation in strength, which is maintained for all phases. For the longer wavelengths this is caused by the LER albedo that increases almost monotonously with wavelength in the ultra-violet, visible and near infrared region (see *Figure* 3.2). For 443 nm the effectiveness of Rayleigh scattering is still significant to produce a higher reflection near full phase as opposed to 550 nm. When comparing our results to that of a homogeneous planet with a Lambertian surface (see the left panel of *Figure* 4 in (Stam 2008*b*)), one can see that due to the anisotropy in the modeled scattering matrix there exists an enhancement of flux for high phase angles. For Pl one can see a rather symmetric behaviour around $90°$ phase, which looks very similar to that of a homogeneous planet with a Lambertian surface (See the right panel of *Figure* 4 in (Stam 2008*b*)). We also know that with increasing surface albedo a shift to higher phase angles is apparent. For the polarizing desert, however, we see no significant shift. Furthermore, the order of colors is approximately maintained for all phases, except for 350 nm and 443 nm.
For Q the order of colors is preserved even more clearly for all phases, even for 350 nm and 443 nm as opposed to what was seen for Pl. The peak of maximum polarized fluxes shifts slightly to higher phases with increasing wavelength, but is very insignificant.

In the planetary model we will model the polarizing surface models for clear pixels and the optically thinnest clouds, i.e. with a cloud optical thickness of 5. For these clouds we do model variant cloud particle effective radii and cloud top pressures as we presented in *Chapter* 3. For values of optical thickness of 10 or 20 we use the Lambertian surface approximation from our initial planetary model. This was decided upon the fact that we are limited by computation time and storage.

6.3. Spectrally and geometrically varying disk-resolved simulations

The disk-resolved cases of an Earth-like exoplanet are simulated at phase angles $0°, 40°, 90°$ and $135°$ for wavelengths $350, 550$ and 865 nm. We present F, Pl, Q and U for every wavelength in *Figure* 6.7, 6.8 and 6.9. Also, we present the associated land cover, cloud top pressure, cloud optical thickness, cloud particle effective radius and cloud fraction distributions on the corresponding planetary disks in *Figure* 6.6. All disks are simulated with 100×100 pixels with the same sub-observer longitude. The planetary model is that described in the former sections.

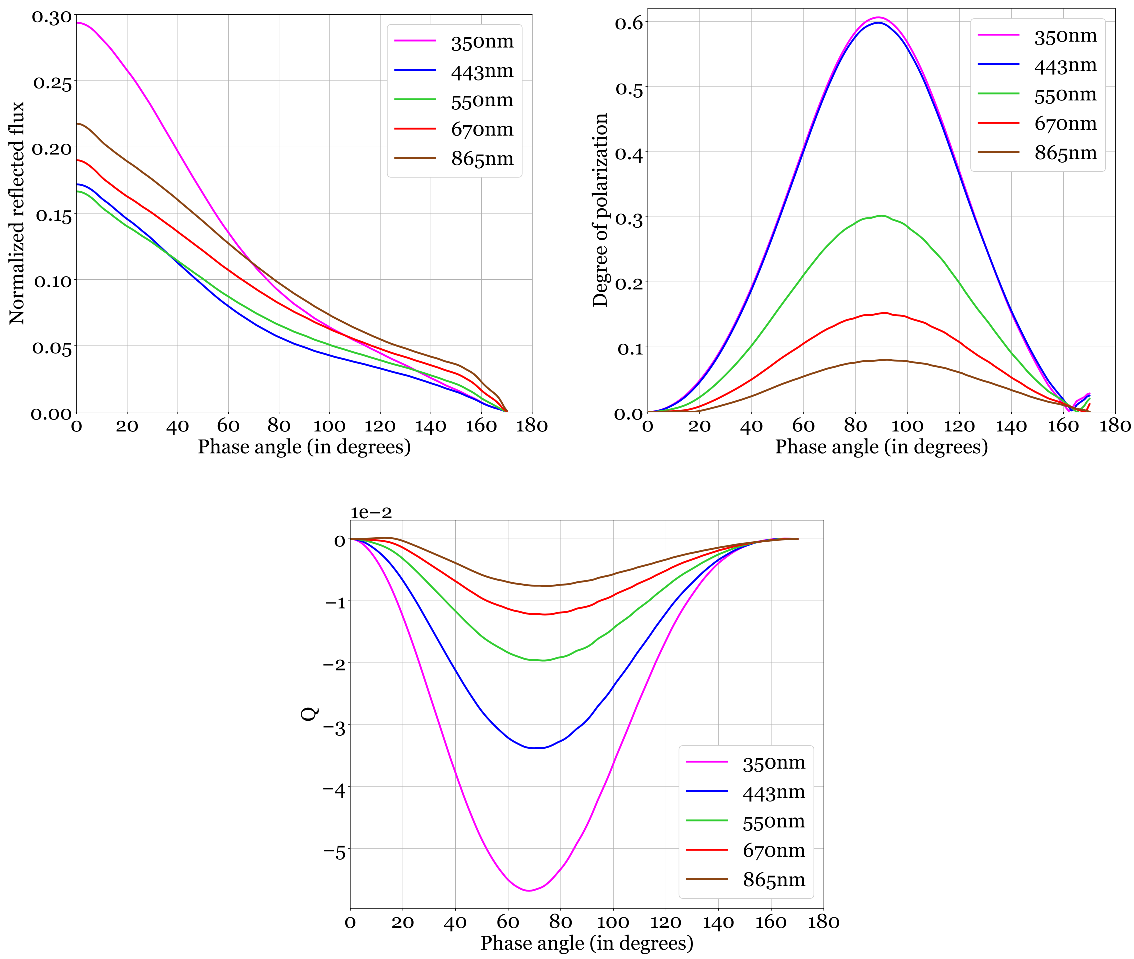


Figure 6.5: Phase curves from a homogeneous planet fully covered with a "polarizing" desert surface. On top of this surface we modeled only a gaseous atmosphere and no clouds. We provide the normalized reflected flux, degree of linear polarization and polarized flux in the left, middle and right panel, respectively.

In this section we will highlight what the main contributions from the ”polarizing” surfaces are with respect to the Lambertian surfaces we used in *Section* 4.1. The former model will be referred to as ”Model 2” and the later model as ”Model 1”. Because we use exactly the same atmosphere, i.e. pure gas with cloud layers, any changes in the disks will be attributed to the anisotropic polarizing reflection models for the ocean, vegetation and desert.

Our findings for F at all phases and wavelengths are itemized as follows:

- At 350 nm, and for all phases considered, the surfaces that are covered by clear sky pixels or pixels with a low cloud fraction are more recognizable, where especially pixels with desert cover exhibit lower values of reflection, where for *Model 1* we only saw low values of reflection corresponding to low values of cloud fraction. The correlation between optically thick clouds and high values of reflection is still clearly visible.
- At 550 nm we observe the same as at short wavelengths, but with a more clear spatial distribution of the different surface covers, owing due to the anisotrope reflection in the surface models. At $\alpha = 135°$ we observe a major increase in reflection, which is concentrated around the planetary scattering plane, where the underlaying surface is that of ocean. This increase is directly related to the specular reflection of the ocean, i.e. the glint. For an illustration of where exactly the region of high reflection is located we refer *Appendix* G where we modeled a homogeneous ocean planet with a pure gaseous atmosphere at $\alpha = 90$ and $135°$, at 865 nm.
- At 865 nm the disks at $\alpha = 0$ and $40°$ show no clear difference to that of *Model 1*. At quadrature we see a small enhancement in reflection to the right of Africa on the planetary scattering plane, corresponding to a small glimpse of the glint. At $\alpha = 135°$ the contribution of the glint to the total reflection is clearly visible.

Our findings for Q at all phases and wavelengths are itemized as follows:

- As we know from *Model 1* is that Q is dominated by highly effective Rayleigh scattering at 350 nm. Hence, we do not see any substantial differences between both models. However, at quadrature and crescent phases we do see an enhanced region of polarized fluxes above the oceanic surface coverage. More specifically, the specular reflection on the ocean highly polarizes scattered light beams (see Trees 2018, for a more thorough analysis on the ”polarizing” glint).
- At 550 nm a small effect of the surface can be seen near Saudi Arabia at full phase. At quadrature and crescent phase we can clearly retrieve the presence of the glint. At quadrature, this glint is partly masked by the presence of the African continent. On this continent, the vegetated and desert surface covers can be distinguished from each other where the latter exhibits a lower magnitude of polarization. At both phases we also see that glint is slightly suppressed by regions of high cloud fraction.
- At 865 nm we basically see an enhanced case of what we observed at 550 nm, but with a stronger magnitude of polarized reflection from the glint, caused by the lower effectiveness of Rayleigh scattering.

Our findings for U at all phases and wavelengths are itemized as follows:

- At 350 nm we do not see any substantial difference between *Model 1* and *Model 2*. At 550 nm we observe the appearance of several surface types for all phases, for example Saudi Arabia and Madagascar. The strongest appearance of different surface covers is shown for $\lambda = 865$ nm, whereas with *Model 1* no such dominance was retrieved. In U we do not retrieve the specular reflection of the glint, but we do not expect that anyway. The disk-resolved polarized flux U of a homogeneous ocean planet is provided in *Appendix* G. This simulation only models a gaseous atmosphere on top of the ocean surface, thus without any clouds. These disk show no enhancement whatsoever from the ocean cover.

Our findings for Pl at all phases and wavelengths are itemized as follows:

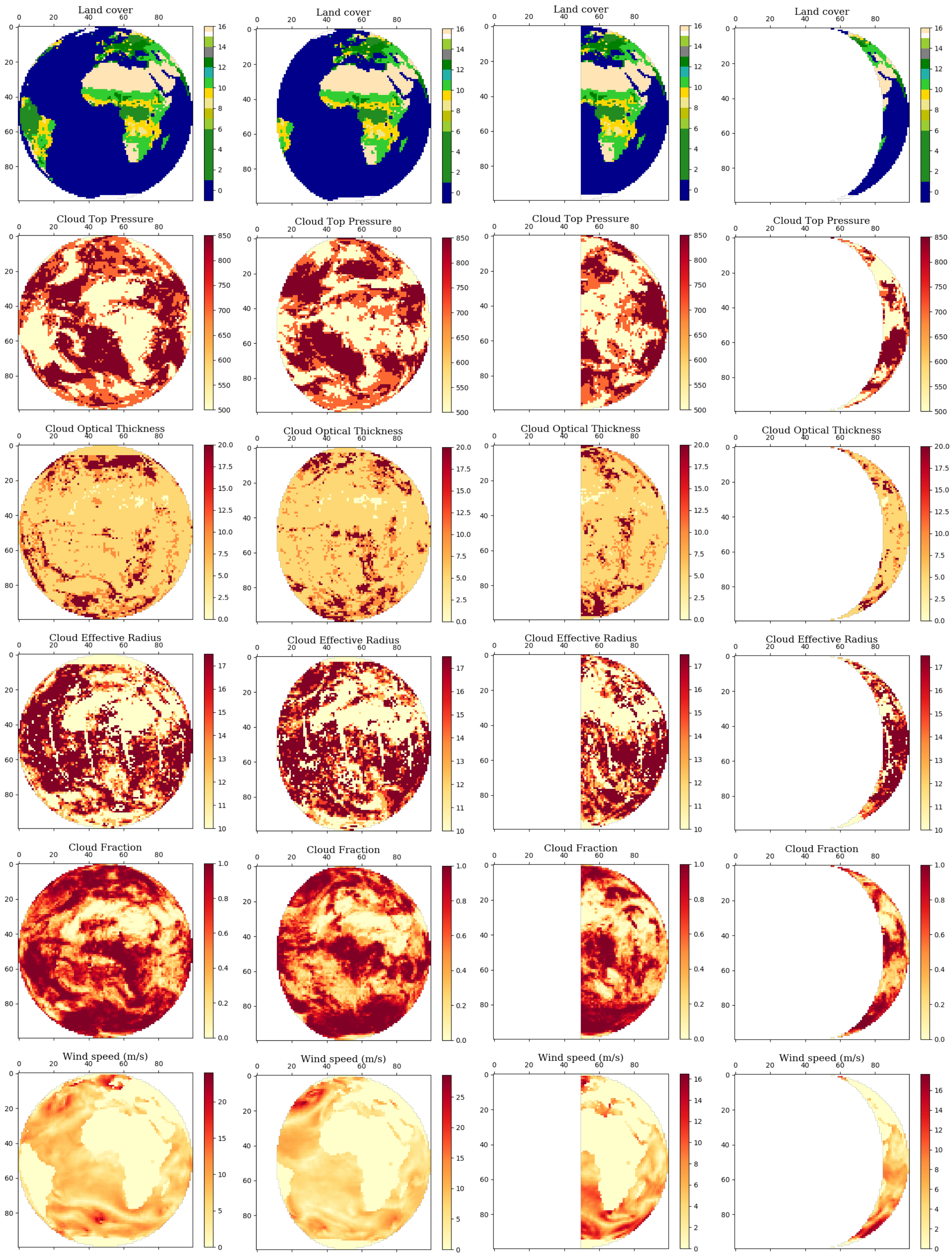


Figure 6.6: From top to bottom we present the land cover, cloud top pressure, cloud optical thickness, cloud particle effective radius and cloud fraction distributions are provided for $\alpha = 0,\ 40,\ 90,\ 135°$ from left to right, respectively. All disks are simulated with 100×100 pixels.

- At $350\ nm$ we only observe a more apparent surface coverage in regions of low cloud fraction, caused by more clear surface distinction found in F. We do not recover any presence of the glint at large phases.
- At $550\ nm$ we observe the same as at short wavelengths for full phase. At $\alpha = 40°$ we observe a much clearer land cover distribution of vegetated and desert covers. At quadrature we observe the same, but mainly the presence of the glint. At crescent phase the glint is even more pronounced, but for both phases still affected by the over-laying clouds.
- In the near-infrared at full phase and $\alpha = 40°$ we observe not much difference with respect to *Model 1*. At quadrature and crescent phase we observe an overwhelming appearance of the glint. As compared to the appearance of the glint in F and Q, for Pl the spatial extend is much larger at both quadrature and crescent phase. This can also be observed in *Appendix* G.

In conclusion, we have seen that the vegetation and desert models induce slight differences in F that provide more distinguishable surface covers. Also, by including polarization in the model a more clear distinction in the polarized fluxes Q and U for the different surface types is observed. By introducing the anisotropic polarizing ocean model we observed a strong presence of the glint in F, Q and Pl. For F and Q the glint is already present at $350\ nm$, where at long wavelengths the glint is clearly visible for F, Q and Pl. The width of the glint is largest for Pl, but the presence is most clearly observed in Q.

6.4. Spectrally and geometrically varying disk integrated simulations

The disk integrated signals of an Earth-like exoplanet are simulated for wavelengths 350, 443, 550, 670 and $865\ nm$. We present F, Pl, Q and U for every wavelength in *Figures* 6.10, 6.11, 6.12 and 6.13. All disks are simulated with a 20×20 pixel grid. The planetary model is again that described in the first two sections of this chapter.

In this section we will highlight what the main contributions are from the model with "polarizing" surfaces to that with Lambertian surfaces we used in *Section* 4.2.2. Again, we refer to the former model as "Model 2" and the later model as "Model 1". The results of *Model 2* are provided in *Figure* 6.10. We also provide the absolute difference with respect to *Model 1*. Our findings for F at all phases and wavelengths are itemized as follows:

- At $\lambda = 350\ nm$ the magnitude in daily variability is only slightly affected. For $\alpha \lesssim 120°$ the absolute magnitude of F at $\lambda = 443°$ increases, whereas at $\lambda = 865°$ it slightly decreases, but the daily variability decreases significantly.
- For $\alpha \gtrsim 120°$ F increases stronger in absolute magnitude for longer wavelengths.
- The intersection point where the colors alternate is still located near $\alpha = 120°$.

Our findings for Q at all phases and wavelengths are itemized as follows:

- For all wavelengths the absolute magnitude of the reflected polarized flux increase, owing to the introduction of our polarized surface models. Furthermore, the daily variability increases for all wavelengths and most strongly for shorter wavelengths.
- At crescent phases *Model 1* showed no alternations in colors, whereas for *Model 2* we observe an intersection point at $\alpha \approx 135°$ in which the colors clearly alter.

Our findings for Pl at all phases and wavelengths are itemized as follows:

- For all wavelengths, Pl increases in absolute magnitude for all phases. At crescent phases we observe a strong bump, which becomes stronger for longer wavelengths. For $\lambda > 670\ nm$ this bump does not increase anymore, because Rayleigh scattering is already very ineffective. More evidently, the daily variability also increases for longer wavelengths. As a consequence, we also observe an intersection point in Pl, but less clear due to the major daily variations.

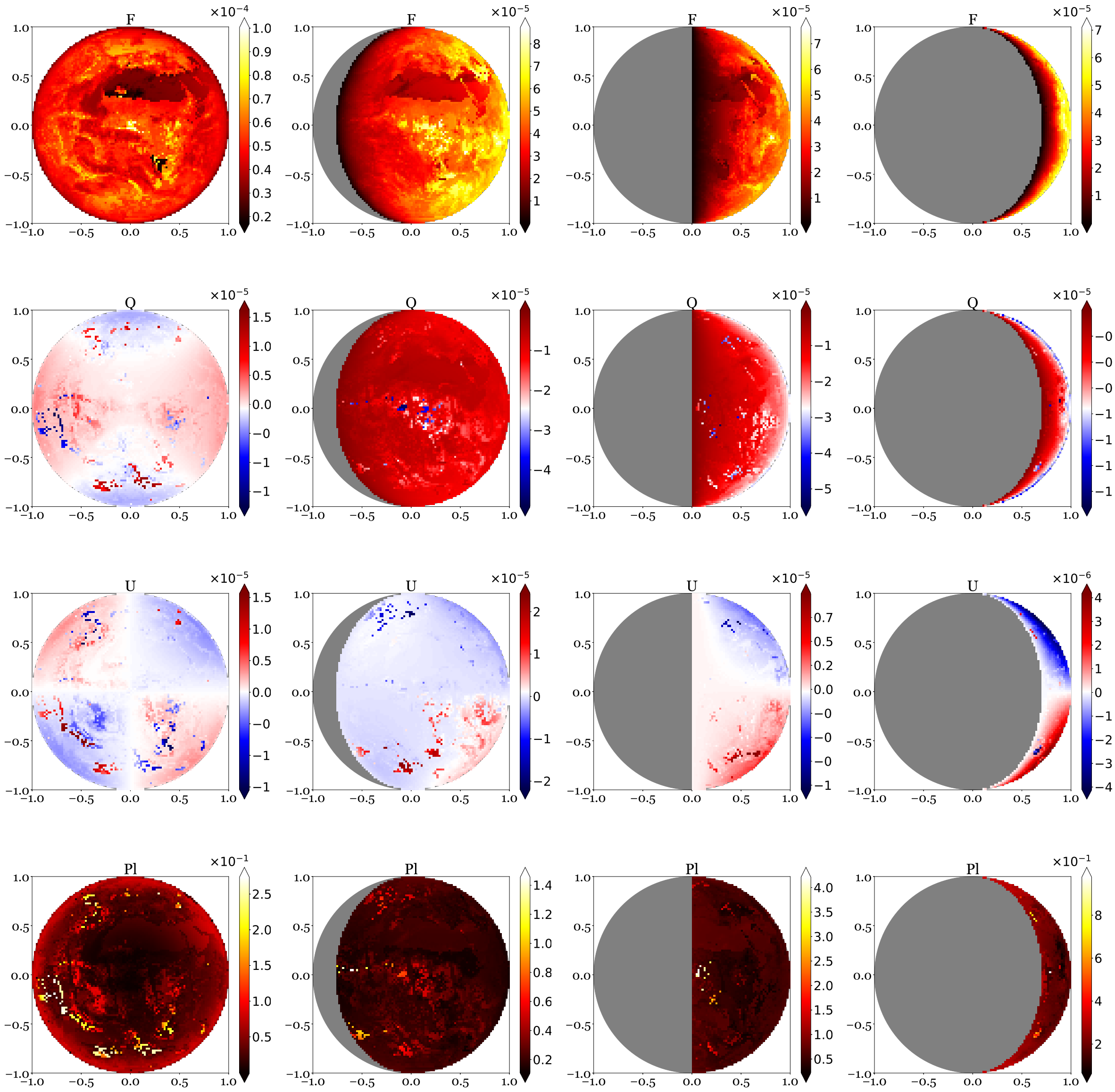


Figure 6.7: Resolved disk for $\lambda = 350\ nm$ at $\alpha = 0°, 40°, 90°$ and $135°$. In the upper four panels the stokes elements F, Q and U, and Pl are provided, respectively. The corresponding disk properties are provided in *Figure* 6.6. All disks are simulated with 100×100 pixels.

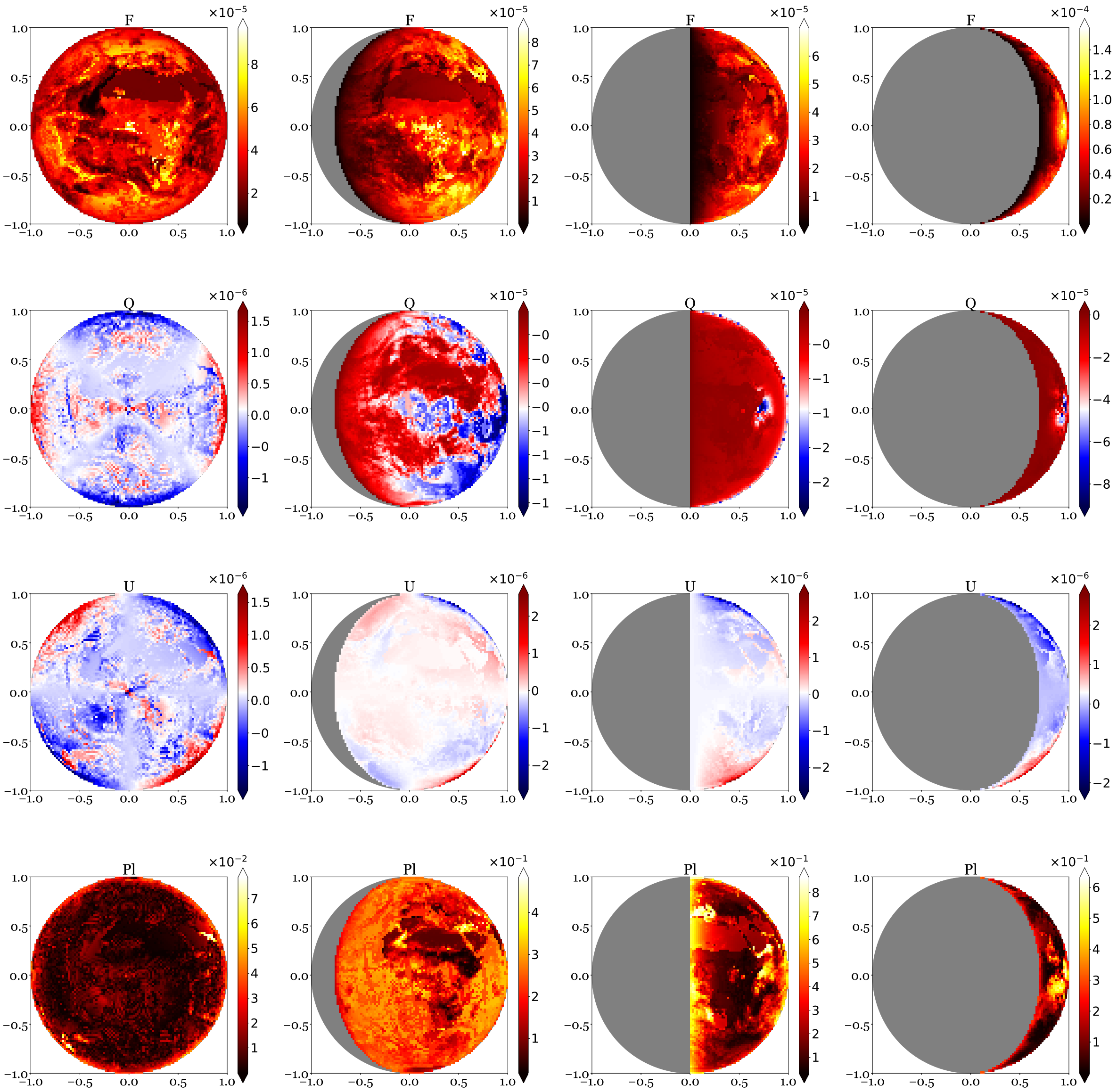


Figure 6.8: Similar to *Figure* 6.7, except for $\lambda = 550\ nm$.

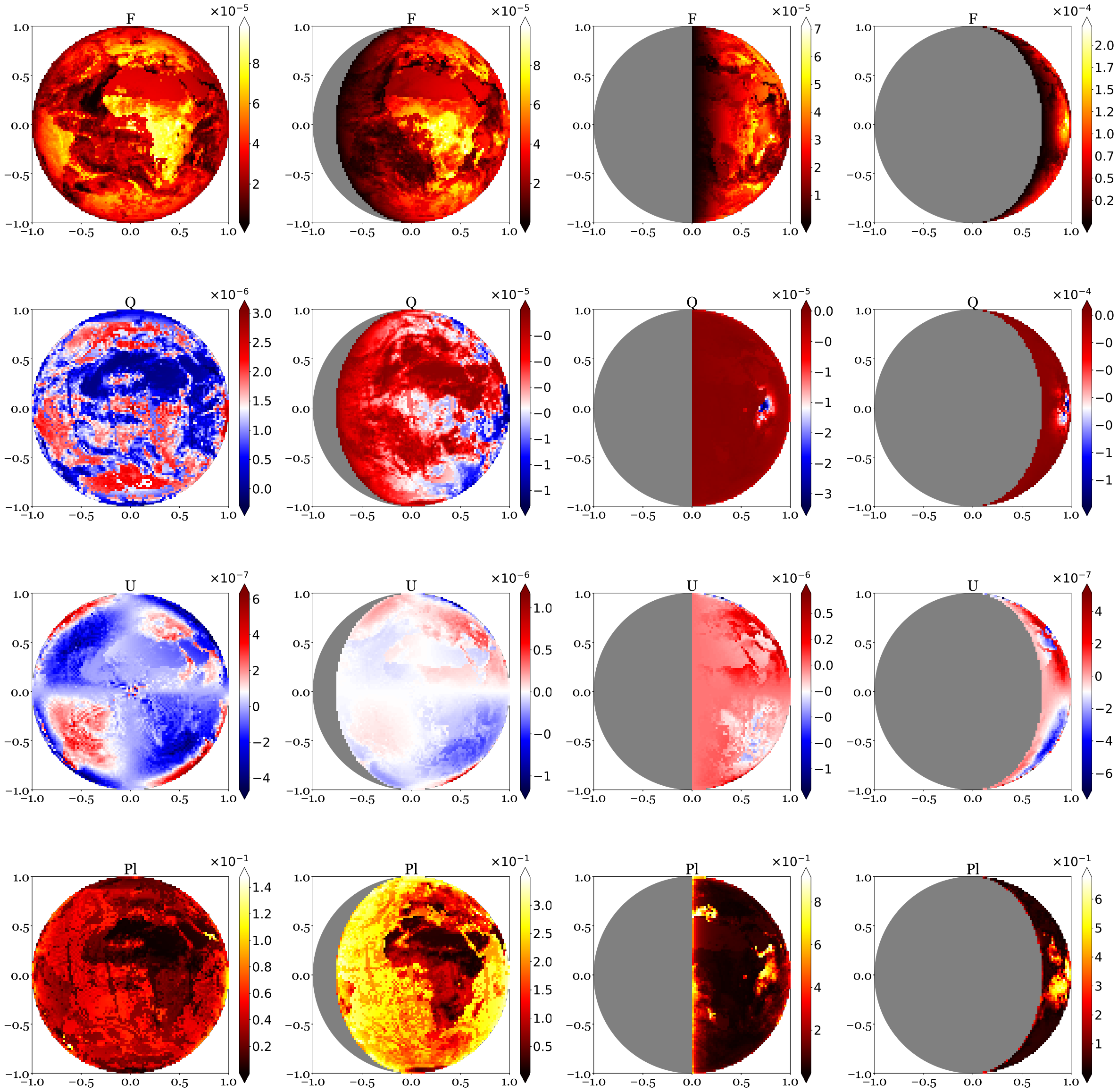


Figure 6.9: Similar to *Figure* 6.7, except for $\lambda = 865\ nm$.

- The primary rainbow at all λ is not affected. Near the secondary rainbow, at $\alpha \approx 56^\circ$, we observe a small dip in daily variability only at $\lambda = 350\ nm$.

For U we observe only a major increase at $350\ nm$, however, the absolute magnitude of the variations are still small. Also, we do not find a specific pattern that can potentially act as a signature of an Earth-like surface distribution or cover. We have retrieved some major changes in our reflected signals, but to which surface cover(s) can these be attributed? In the following figure we provide special end cases in which we provide the relative difference to the "complete" *Model 2* for every parameter.

In *Figure* 6.11 we replaced the vegetated surface covers with a black Lambertian surface cover to investigate its contribution to the model. For F we observe that the daily variability and absolute magnitude decreases significantly at $865\ nm$, which is attributed to the red-edge feature in the reflection spectrum of vegetation in general. The green bump, another enhancement in the reflection of vegetation, also causes a decrease in magnitude and daily variability. For spectral regions of low vegetation reflection, i.e. $\lambda = 350,\ 443$ and $670\ nm$, the absolute difference is least. The change is largest near full phase and decreases with increasing α. At crescent phases vegetated land covers have limited effect on the phase curves. For Q vegetation induces no significant change in the absolute amplitude and variability at $443,\ 670$ and $865\ nm$. For $\lambda = 550\ nm$ the absolute amplitude increases and the variability slightly decreases for a large region of phases near quadrature. An interesting feature near the secondary rainbow is observed at $350\ nm$, which shows an increase in the amplitude of a local daily variation. Pl shows a significant reduction of the daily variability induced on the primary rainbow, especially for λ at the red-edge feature. We observe the same at the glory. At the secondary rainbow we again see an increase in a local daily variation. At other α there occur slight increases in daily variations, which is largest for $\lambda = 550\ nm$. Lastly, for U at $350\ nm$, we also observe the increase in a local daily variation at the secondary rainbow. Furthermore, there occur changes in the variability at all other λ, but all are very small.

In *Figure* 6.12 we replaced the desert surface cover with a black Lambertian surface cover, to investigate its contribution to the model. For F we observe that the daily variability and absolute magnitude decrease and is strongest for longer wavelengths. As opposed to the reflection spectrum of vegetation, that of soil increases much more gradually, hence the gradual decrease in variability and absolute magnitude. The change is largest near full phase and decreases with increasing α. At crescent phases desert land cover have limited effect on the phase curves. For Q the desert model induces only changes in absolute amplitude and variability in a large region of phases around quadrature, being largest at $350\ nm$. An interesting feature near the secondary rainbow is also observed without the desert at $350\ nm$, but also at other λ and shows to decrease with increasing λ. At the primary rainbow the variability is affected, but not so much the amplitude of the variations. Pl essentially shows the same as in Q, but with a slight increase in daily variability at crescent phases for longer wavelengths. Lastly, for U at $350\ nm$ the change in daily variability is also apparent. Also, there occur changes in the variability at all other λ, but all are very small.

In *Figure* 6.13 we replaced the oceanic surface cover with a black Lambertian surface cover, to investigate its contribution to the model. For F we see that at crescent phases, at and after the intersection point, the magnitude and variability increases significantly, being strongest for longer wavelengths. In the absence of an ocean cover we do, however, still preserve a color alternation, which is attributed to the gas layers on top of the clouds as was described in *Section* 4.2.2. Also, at other phases we observe a decrease in absolute magnitude, which is strongest near full phase and for short wavelengths. This is caused by the fact that light, which is reflected by the ocean, is now completely absorbed. More specifically, from *Figure* 6.1a in Trees (2018) the albedo of the ocean just below the air-water interface is highest in the blue/ultra-violet wavelength region at which we also observe the largest decrease in F. For Q we completely lose the intersection point. Furthermore, the variability and amplitude near quadrature and at the primary rainbow decrease significantly. Again, for $\lambda = 350\ nm$ at the secondary rainbow, we observe an anomaly for ΔQ. Similarly, in Pl, the intersection point is lost, being of course a result of what we found for Q. In U we observe no interesting changes, but an overall decrease in daily variability, which is strongest for short λ.

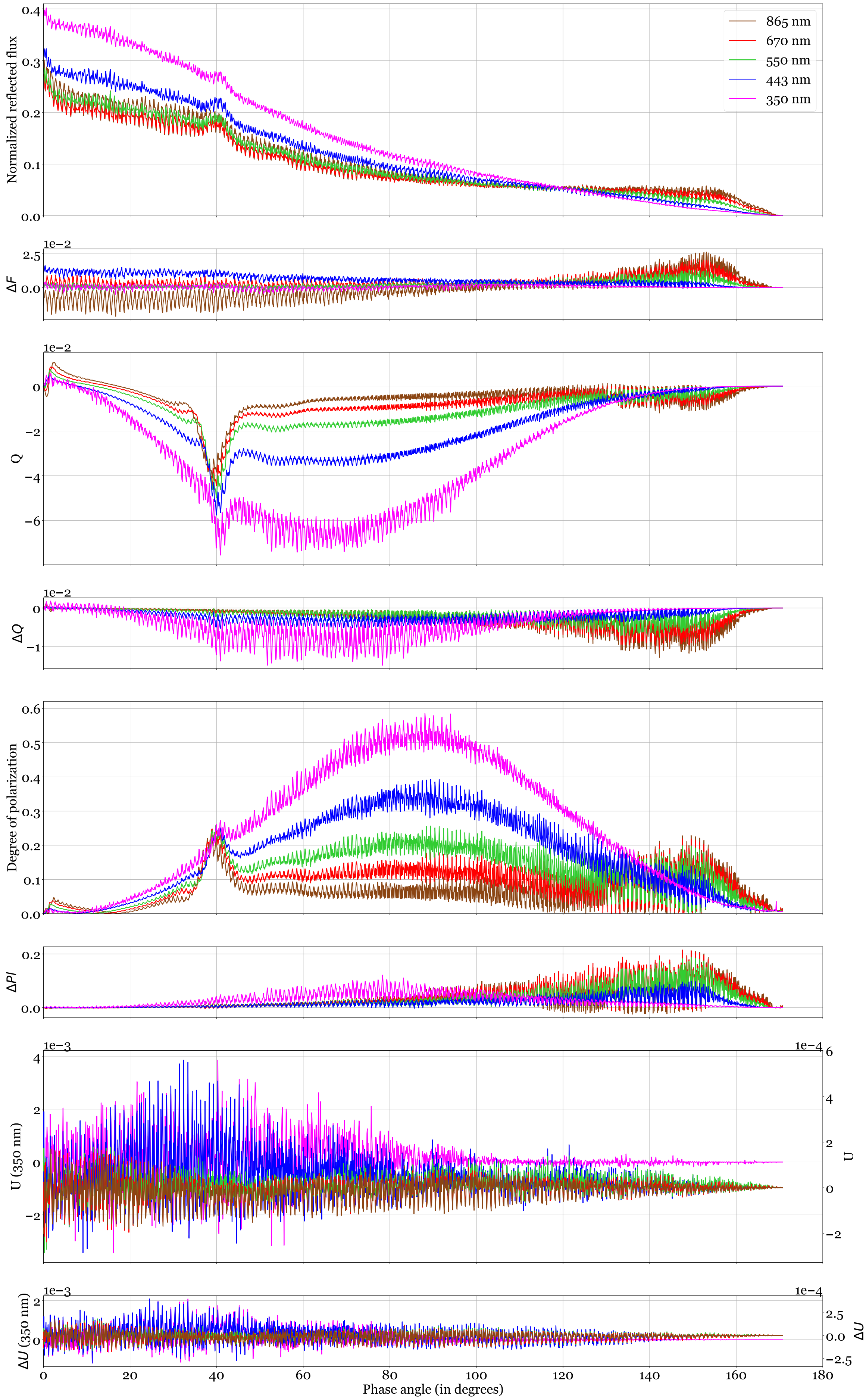


Figure 6.10: Semi-annual phase curves for F, Pl, Q and U. These phase curves are provided for five wavelength: $350, 443, 550, 670$ and 865 nm. The planetary model is that described in *Chapter* 3, except we replaced the Lambertian surface model of oceanic, vegetated and desert surface covers with anisotropic polarizing models. For every parameter we provide the absolute difference with respect to the Lambertian planetary model.

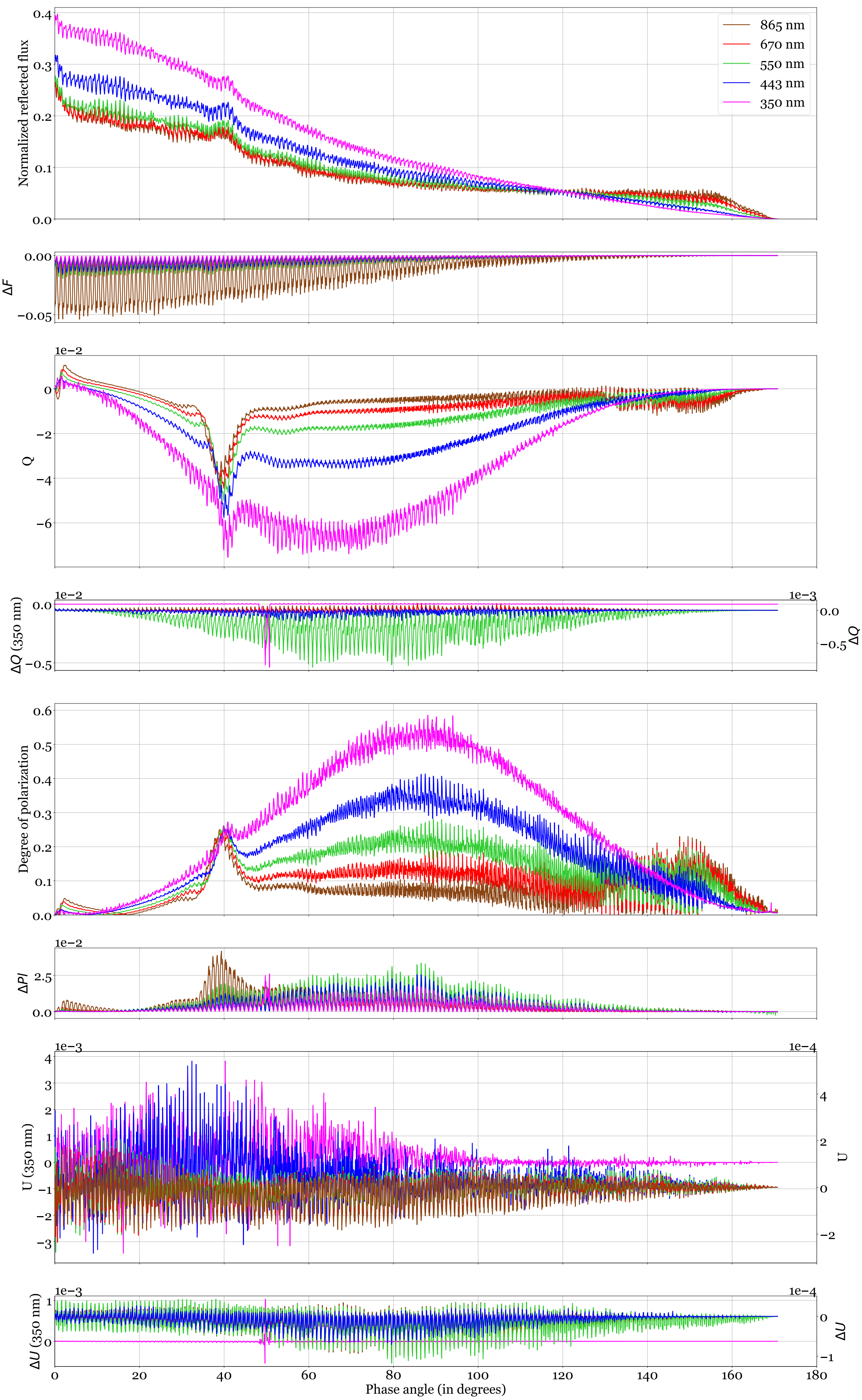


Figure 6.11: Similar to *Figure* 6.10, except without considering vegetated surface covers in the planetary model, i.e. the surface models are Lambertian with an albedo of 0. For every parameter we provide the absolute difference with respect to the "full" planetary model provided in *Figure* 6.10.

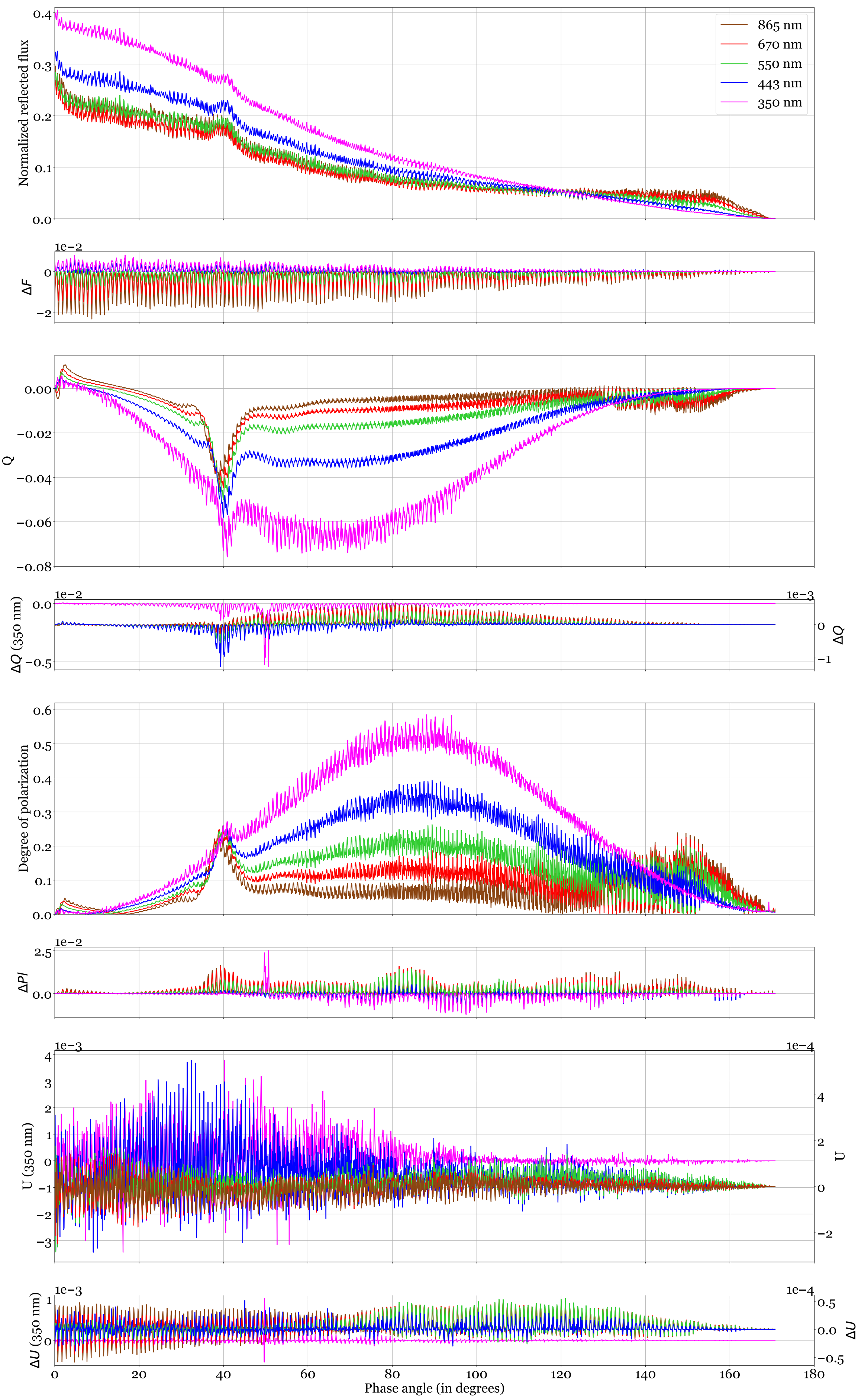


Figure 6.12: Similar to *Figure* 6.10, except without considering desert surface cover in the planetary model, i.e. the surface model are Lambertian with an albedo of 0. For every parameter we provide the absolute difference with respect to the "full" planetary model provided in *Figure* 6.10.

In conclusion, we saw that the polarizing desert cover mainly has an effect on the daily variations at phases near full phase until quadrature, but less so as compared to the vegetation model, which can be attributed to the fact that the desert is much less spread out over the Earth disk than for example vegetation or ocean. The intersection points induced on Q and Pl are also not caused by the polarizing desert and vegetation surface covers. By inspection of *Figures* 6.1 and 6.5 we do not observe an intersection point in the end cases of the vegetated and desert surface cover, whereas for the ocean planet end case (*Figure* 6.2) an intersection point was also found. We conclude the following:

- The polarizing vegetation has an effect on the daily variations in and strength of F, especially at $865\ nm$ for phases smaller than quadrature. For Q the vegetation model only induces significant variations for $\lambda = 550\ nm$.
- The polarizing desert induces additional variability on F that decrease with shorter λ and for larger α. For Q we observed an increase in daily variations at all λ for phase angles in a large region around quadrature.
- Trees (2018) showed that if an exoplanet is completely covered by a (frozen) ocean, then an intersection point in Q may be found, providing that the cloud fraction is less than 100% and/or the surface pressure is not well beyond that of Earth. Our model includes an Earth-like cloud cover that varies diurnally as well as an inhomogeneous surface cover. Under the influence of these inhomogeneities we are still able to retrieve a clear intersection point in Q, which showed to be solely caused by the presence of an ocean. In a similar fashion we also found the intersection point in Pl, again caused by presence of an ocean, which shows large correspondence at crescent phases to that of *Figure* 7.20 in Trees (2018). Lastly, Trees (2018) showed that without considering clouds an intersection point can also be found in F for a homogeneous ocean planet. In *Appendix* F, *Figure* F.2, we provide a simulation of *Model 2*, except that all cloud layers are considered as pure gas layers. More specifically, for our inhomogeneous surface cover we still obtain not only an intersection point in Q and Pl but also in F, in the absence of clouds. By inspection of the end cases in *Figure* 6.2 we find the same for an ocean planet. However, due to the highly reflecting clouds at crescent phases and for long λ (*Section* 4.2.2) this intersection point in F can not be unambiguously linked to the presence of the ocean. Also, Cowan et al. (2012) showed that highly reflecting poles can increase the flux at crescent phases, acting as a false positive for the detection of a glint. However, in Q the clouds did not cause such an intersection point and for the ice/snow covers it is reasonable to expect that the surface reflection is not highly polarizing (Peltoniemi et al. 2009). Hence, the intersection point in Q might provide us with an unambiguous signature of ocean on an Earth-like exoplanet.
- Other than retrieving the presence of a liquid water ocean on an Earth-like exoplanet, the position of the intersection point in F and Q might provide us with an estimation of the cloud fraction. Trees (2018) showed that for an idealized homogeneous ocean planet an increasing cloud fraction causes a shift of the intersection point to smaller and larger α for F and Q, respectively. By inspection of *Figure* 6.10 we retrieve: $\alpha_{intersect,F} \approx 120°$ and $\alpha_{intersect,Q} \approx 136°$. Assessing now *Figure* 7.28a and 7.29a in Trees (2018) allows us to retrieve a cloud fraction in both cases of $65 - 70\%$, agreeing well to the annual mean of MODIS observations that we used: 68% (see *Section* 3.3.2).

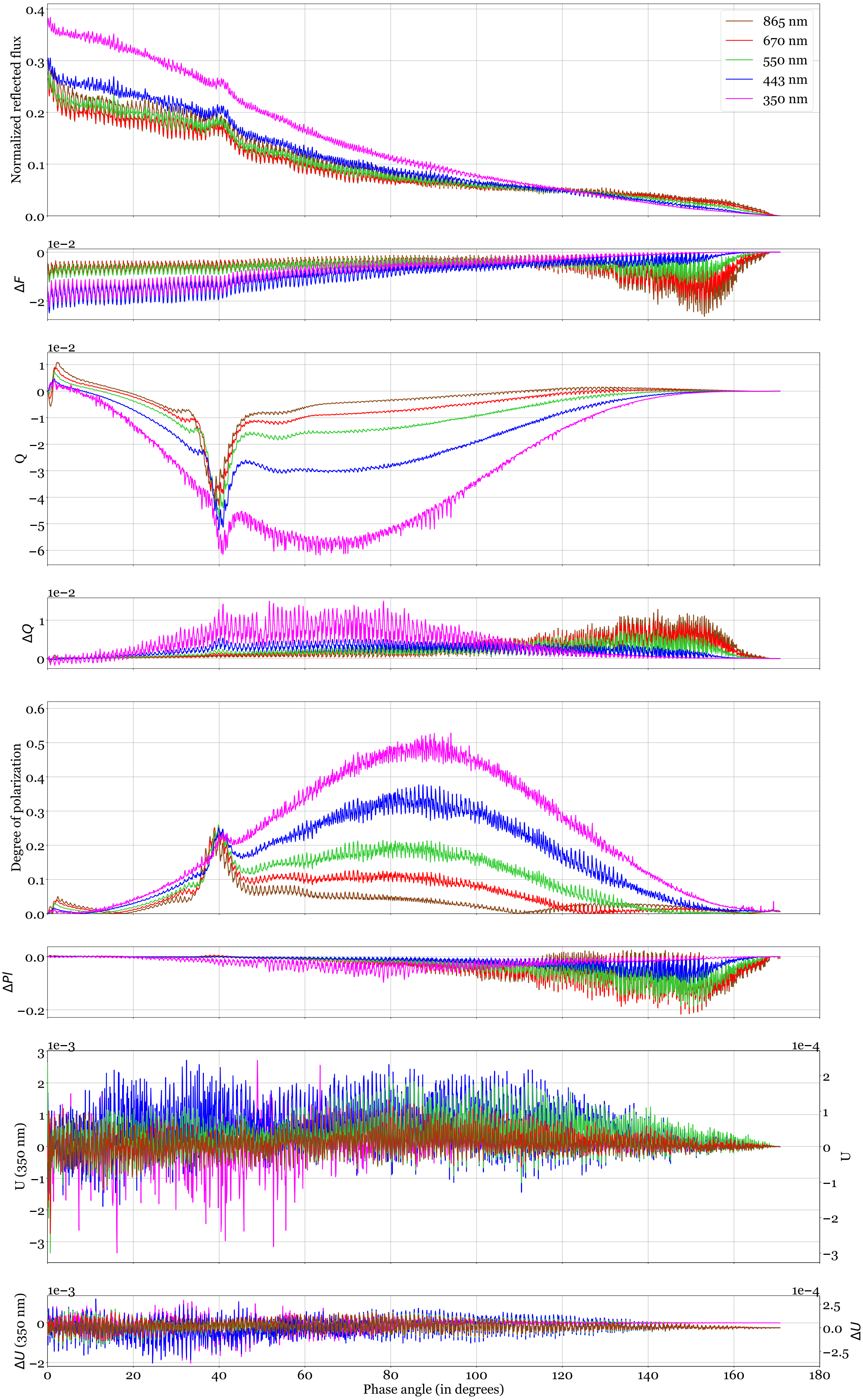


Figure 6.13: Similar to *Figure* 6.10, except without considering ocean surface cover in the planetary model, i.e. the surface model are Lambertian with an albedo of 0. For every parameter we provide the absolute difference with respect to the "full" planetary model provided in *Figure* 6.10.

7

Modeling Earthshine data

As we already know and have seen from our simulations is that polarimetry can be a strong asset to the characterization of terrestrial Earth-like exoplanets. We have seen that by including polarization in our measurements one can see strong indications of ocean bodies on Earth-like exoplanets as well as the presence of liquid water clouds. Up until know the only direct polarimetry observations from Earth have been made by the Polarization and Directionality of the Earth's Reflectances (POLDER) instrument (Deschamps et al. 1994). This instrument, however, is mounted on a satellite in low-Earth-orbit and can thus not provide representative results of a full disk-resolved or unresolved as observed from afar. Efforts are being made to design an instrument, called the Lunar Observatory for Unresolved Polarimetry of Earth (LOUPE), that will use the Moon as a platform to retrieve the spectropolarimetric signals of Earth (Karalidi, Stam, Snik, Bagnulo, Sparks and Keller 2012). In the mean time several observations of Earthshine have been made to retrieve photometric signals from Earth (Qiu et al. 2003). Spectropolarimetric signals of such kind have only recently been made by Sterzik et al. (2012) that show some interesting possible bio signatures from Earths spectrum. In this Chapter we will focus on the attempt to approach the observations made by Sterzik et al. (2012).
We start this chapter by a description of how Earthshine observations are retrieved (*Section* 7.1). Secondly, we will present the measurements that are retrieved by Sterzik et al. (2012) and earlier attempts to approximate this data (*Section* 7.2). In *Section* 7.3 we will present the model that will be used to simulate the Earthshine measurements, which will be presented in the last section, *Section* 7.4.

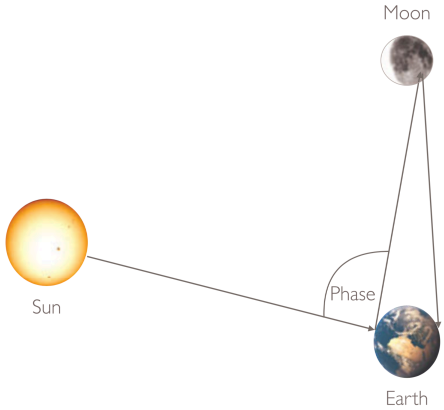
7.1. Observing Earth with Earthshine

Earthshine measurements are retrieved by observing the night side of the Moon from the night side of Earth. This side of the Moon is illuminated by the reflection of sunlight reflected on the day side of Earth, more clearly shown in the left panel of *Figure* 7.1. In the right panel one can clearly see the Earthshine visible on the "dark side" of the Moon. This method does have its limitations as one is not able to observe large phases Earth, because the night side of the Earth in view of the Moon is very small. Also, one needs a global system of observations telescopes which for the required precision is not achievable. Furthermore, we are dealing with a non isotropic Lunar surface, which one can easily conclude from the several craters on the Lunar surface and the darker surface features.

7.2. Polarization measurements from Earthshine

The spectropolarimetric results that are obtained with the Earthshine method are retrieved on the 25th of April 2011 on 09h00 UTC and the 10th of June 2011 on 01h00 UTC. These measurements are obtained with the Focal Reducer/Low-dispersion Spectrograph (FORS) that is mounted at the VLT based in Chile (Sterzik et al. 2012). The fraction of polarization P_Q obtained at these two epochs are provided in *Figure* 7.2. The red curves in these two panels represent the observed fraction of polarization P_Q, and the green curve represents the observed fraction of polarization P_U, which will not be assessed in this research. The inset in the two plots show the variations of P_Q[1] that are extracted from the

[1]In the remainder of this section we will refer to P_Q as Ps.



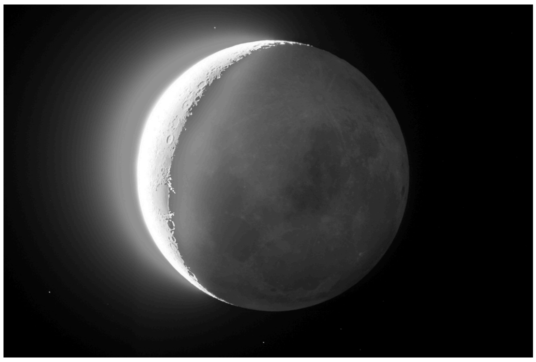


Figure 7.1: Left panel: Sun-Earth-Moon system with the illustration of Earthshine from the Moon (Sterzik et al. 2015). Right panel: Visible Earthshine on Moon at crescent phase.

continuum, i.e. the residual δP. This is achieved by fitting a fourth order polynomial to the data and subtracting this from the signal between 530 and 910 nm. The two major peaks in the residual show clearly the O_2 A-band at 760 nm. The horizontal black lines in this residual show two representative wavelength passband regions for which the NDVI is calculated. From the lower panel of *Figure* 7.2 one can see the land distribution for the two epochs. From these land distributions we can clearly see a large portion of vegetated (greenish) areas on April 25th as compared to June 10th. According to Sterzik et al. (2012) this major difference in the fraction of vegetated regions causes a dip in the observed signal, or a strengthening of the NDVI, which one can clearly see from the plots in *Figure* 7.2. For a more detailed discussion on the present features in the observations one can consult Sterzik et al. (2012). The black solid, dashed and dotted lines represent simulations that use the weighted quasi-homogeneous disk integrated approximation of Stam (2008*a*). The different weights are provided in *Table* 7.1. We can see that these simulations approximate the blueish/green part of the spectrum fairly well, but from 500 mu on wards the estimation becomes increasingly worse. The exact model parameters of these simulations can be retrieved from Sterzik et al. (2012). For April the simulation with the largest clear ocean portion provides the highest fraction of Pl on the entire spectrum. Simulating clouded vegetation and a higher fraction of clouded ocean provides the worst fit, especially for the blue region of the spectrum. When introducing some clear vegetation, the degree of polarization remains approximately the same as the first case, but agrees well to the observed data in the far blue. For June we see that for a high fraction of clouded ocean we see a large disagreement in the far blue/greenish spectrum. When introducing some clear vegetation the overall agreement with the data increases, but only by a small portion. The best fit is obtained by modeling a large fraction of the ocean as clear. As

Table 7.1: Model parameters of the simulations by Sterzik et al. (2012) used to approximate the observed Earthshine data. These simulations are based on the Quasi-homogeneous model by Stam (2008*a*). This table is fully taken from Sterzik et al. (2012).

Date/Model	Clouded ocean (%)	Clear ocean (%)	Clouded vegetation (%)	Clear vegetation (%)
April/Solid	48	40	0	12
April/Dashed	60	30	10	0
April/Dotted	44	56	0	0
June/Solid	40	60	0	0
June/Dashed	30	60	0	10
June/Dotted	27	73	0	0

mentioned in the previous section, the lunar surface depolarizes the signal. Although large ambiguities are presented in the spectral dependence and magnitude of this depolarization factor an estimation by Sterzik et al. (2012) is used, which is normalized at 550 nm and varies linearly according to

$$depol = 3.3\lambda/550 \tag{7.1}$$

So, to account for this depolarization and to be able to adequately compare our results with the observations and simulations we will have to divide our signal by this depolarization factor. The geometric

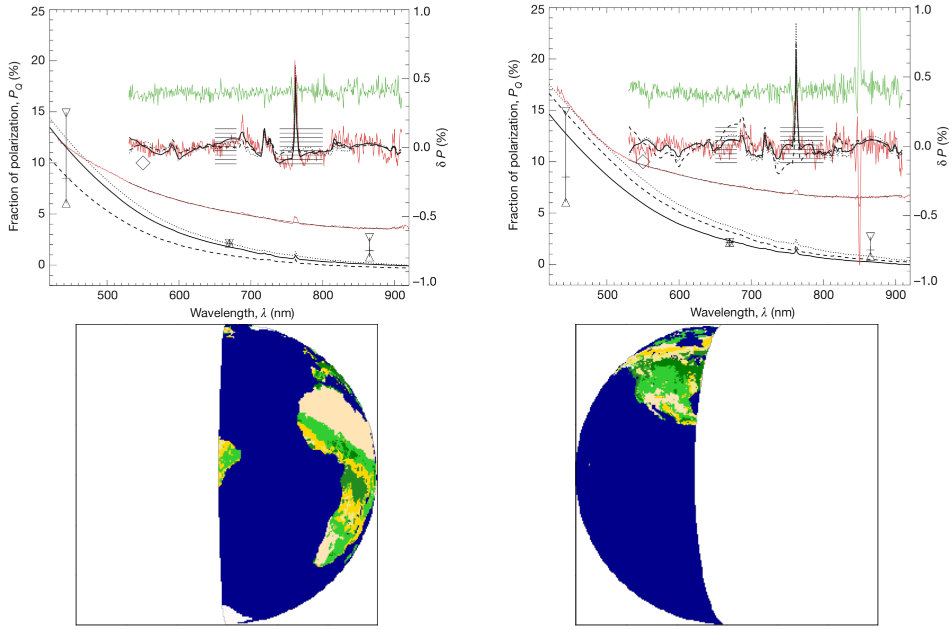


Figure 7.2: Upper left panel: Fraction of polarization P_Q measured from Earthshine on April 25th 2011 at 09h00 UTC, represented by the red curve. From these measurements the continuum is subtracted by fitting a fourth order polynomial, producing the inset in the plot. The green line represents P_U, which is on the same scale as the residuals. The triangles and diamond represent estimates based on POLDER and estimates from Dollfus, respectively. The black lines represent simulations with approximate Earth-like models. For more information about these estimates one can consult Sterzik et al. (2012). Upper right panel: Similar to the right panel, only for Earthshine retrieved on June 10th 2011 01h00 UTC. (Sterzik et al. 2012) Bottom left panel: simulated land cover on April 25th. Bottom right panel: simulated land cover on Jun 10th. Both land covers are discretized according to the IGBP classification (*Section* 3.2).

angles that we will be using are summarized in *Table* 7.2, which are based on the latitudinal and longitudinal position of the sub Lunar and sub Solar point on the Earth's surface. Because we will only have to compute one geometry per observations, we can use our full range of wavelengths: $350, 443, 500, 550, 670, 750$ and 865 nm.

Table 7.2: Summary of used geometrical angles for the two Earthshine epochs

	April 25th 2011	June 10th 2011
Sub Solar longitude (°)	44.517	-195.167
Sub Solar latitude (°)	13.133	22.967
Sub Lunar longitude (°)	-38.783	-93.783
Sub Lunar latitude (°)	-14.933	-5.383
Phase angle (°)	87	258
Input Obliquity (°)	-14.933	-5.383
Input Longitude (°)	-38.783	-93.783
Rotation (°)	14.427	20.720

7.3. The Earth model

The model that we will be using for the simulation of Earthshine data are very similar to that described in *Chapter* 6, but with a small addition. In summary, we will use the detailed cloud models that are described in *Chapter* 3 in combination with the polarized surfaces presented in *Chapter* 6. Additionally, we will use the scattering coefficients that we used to approximate a polarizing desert surface, to model dust aerosols in our atmosphere. These particles are modeled solely above clear sky pixels, so we will not model cloud layers in combination with dust aerosols layers. As one can read from *Chapter* 6, the polarizing desert surface were fitted to correspond to Aster spectral data for desert soil, by changing the single scattering albedo (SSA). These fitted SSA will also be used to approximate the wavelength dependent aerosol particles. In *Table* 7.3 the SSA's are listed for the corresponding wavelengths.

To be able to model these dust aerosols in a local atmosphere, we require the optical thickness

Table 7.3: Fitted Single Scattering Albedos (SSA's) to approximate dust aerosols with empirically obtained scattering coefficients by Moreno et al. (2006). The scattering coefficients were fitted to Lambertian Equivalent Reflectance (LER) albedos from Aster Spectral Database.

Wavelength $[nm]$	SSA [-]	LER [-]	Model albedo [-]
350	0.235	0.010604	0.010613
443	0.6103	0.05319	0.053186
500	0.7384	0.091416	0.091430
550	0.82353778	0.139387	0.139390
670	0.89436	0.215352	0.215391
750	0.91375	0.248771	0.248747
865	0.9279	0.279383	0.279351

of that layer and its vertical position. The optical thickness is obtained from MODIS data. The MODIS data that we will be using is the "AOD_550_Dark_Target_Deep_Blue_Combined_Mean_Mean" data set from the data product "MYD08-D", the same as that used in *Chapter* 3. In this data set the optical thickness values are retrieved at $\lambda = 550nm$, so we will scale the optical thickness similarly to the cloud optical thickness (*Equation* 3.4):

$$\tau^*(\lambda^*) = \frac{\sigma^*_{ext}}{\sigma_{ext}}\tau(\lambda) \tag{7.2}$$

where the asterisk denotes the desired optical thickness at a corresponding wavelength. The file that was fitted to aerosols observations did not include real valued scattering properties, so by use of Mie calculations extinction coefficients for $\lambda = 442, 443, 550, 663, 670, 850$ and $865\ nm$. The Mie calculations were performed by Dr. D. M. Stam. These Mie calculations are based on a fitted size distribution and measured refractive index by Moreno et al. (2006). *Figure* 1 in Moreno et al. (2006) shows the Log-Normal two parameter distribution with $R = 0.113\ \mu m$ (geometric mean radius) and $\sigma = 2.517$ (standard deviation) and the measured size distribution. Although there can be seen some small disagreement we find that the approximation is sufficient for the current discussion. For the refractive index, Moreno et al. (2006) uses a constant index for 442 and 633 nm: $n_r = 1.62 + 0.00001i$. To verify if this assumption is also valid for the ultra-violet and near-infrared domain, we assessed the *JPDOC Database of Optical Constants*. Sources like Pollack et al. (1994) confirm that the rock type "Olivine S" can be assumed constant for our entire wavelength domain. The last step now is to inter- and extrapolate the data to allow for calculations on all our wavelengths. We observed the data and found that it is certainly not a linear-like relation. In this regard, we fitted a 3rd order polynomial to the data, shown in *Figure* 7.3. In order to adequately model the aerosol optical depth with a reasonable

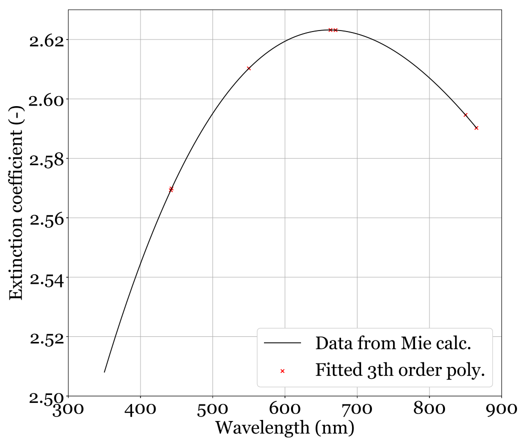


Figure 7.3: Red crosses indicates the data from Mie calculations by Dr. Daphne Stam. The black line is the result of a fitted 3rd order polynomial.

computation time and storage, we will again use a discretization. To provide such a discretization we will firstly look at the abundance distribution. To adequately represent the distribution we fitted 200 of the most commonly used continuous statistical distributions to the empirical data (Kotz 1994), from which we obtained a *Noncentral t-distribution*. From this distribution we can define confidence intervals, from which we strive to cover 95% of the data. The distribution and the confidence interval are shown in *Figure* 7.4. This distribution and histogram show that the highest abundance is located at ~ 0.1 optical thickness. 95% of this data lies between $0.05 - 0.53$ optical thickness. Secondly, we will look at the

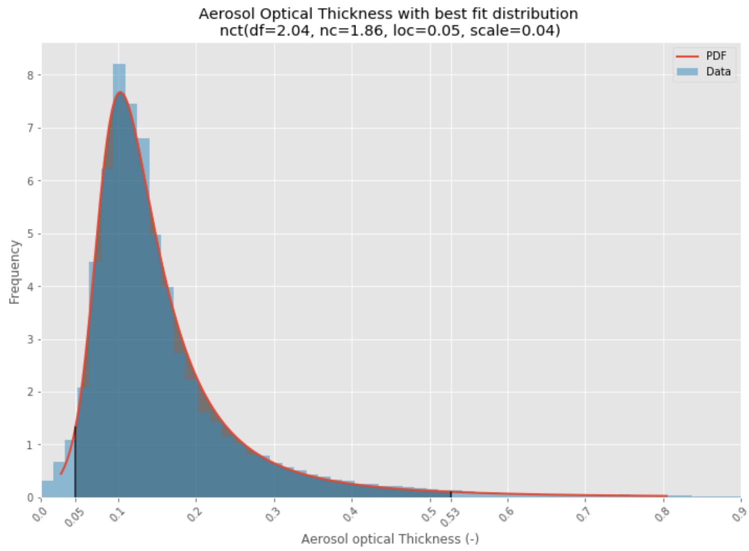


Figure 7.4: Histogram and fitted PDF to the aerosol optical thickness. A confidence interval of 95% is provided for the PDF.

distribution of the mean aerosol optical thickness as function of latitude. This is especially import as the dust aerosols are mainly located near the approximate latitude interval of $0 - 50$ degrees (Ginoux et al. 2001). Also, in this case we are mostly interested in the region surrounding the Saharan Desert as we will only model the dates on which the Earthshine measurements were retrieved of which we saw that April 25th 2011 fully includes the Saharan Desert, whereas July 10th 2011 only contains a large portion of the Pacific Ocean and the North and South American continents. If we look at *Figure* 7.5, one can see that for the region of $0° - 50°$ longitude the mean optical depth varies between $\sim 0.15 - 0.25$. Considering both the abundance of the full set of data and the latitudinal dependence we will use

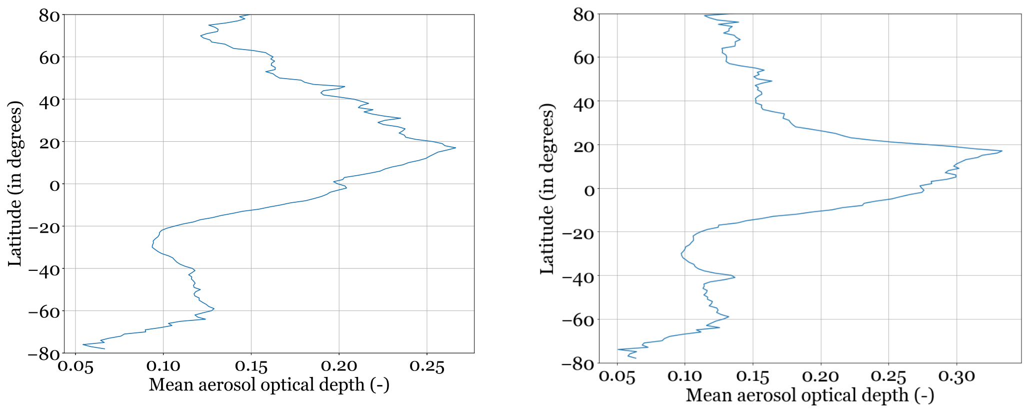


Figure 7.5: Left panel: Latitudinal dependence of mean aerosol optical depth, derived from monthly global data of 2011 entirely. Right panel: Latitudinal dependence of mean aerosol optical depth, derived from monthly data of 2011 entirely only for the region of $0° - 50°$ longitude.

the three bins provided in *Table* 7.4. The Aqua/MODIS Aerosol Optical Depth Daily L3 Global 1 Deg. CMG dataset was acquired from the Level-1 and Atmosphere Archive & Distribution System (LAADS) Distributed Active Archive Center (DAAC), located in the Goddard Space Flight Center in Greenbelt, Maryland (`https://ladsweb.nascom.nasa.gov/`). To include the observation of these aerosols

Table 7.4: Discretization bins of the aerosol optical thickness with associated bin values.

Bin number	1	2	3
Bin boundaries	0<AOT≦ 0.1875	0.1875<AOT≦0.2625	0.2625<AOT
Bin values	0.15	0.225	0.3

in our Earth-like model we need to make some assumptions. MODIS only computes aerosol properties for clear sky pixels, and thus not for cloudy pixels (Levy et al. 2013).

1. L3 aerosols are only considered for clear sky pixels.
2. For a L3 pixel with AOT>0 the full clear sky L3 pixel is replaced by one with an aerosol layer.
3. AOT values are averaged by an arithmetic mean over the pixels that fall into our disk pixels.
4. The total disk pixel is approximated by a weighted summation, including clear sky, aerosol and cloud reflection matrices.

The total disk pixel TOA reflection is computed with (analogously to *Equation* 3.3)

$$\mathbf{R}_{tot} = \mathbf{R}(\tau_{cld})\frac{\sum_{i=1}^{N} Fcld_i}{N} + \mathbf{R}(\tau_0)(1 - \frac{\sum_{i=1}^{N} Fcld_i}{N} - \frac{N_{aero}}{N_{clear}}) + \mathbf{R}(\tau_{aero})(\frac{N_{aero}}{N_{clear}}) \tag{7.3}$$

where $\mathbf{R}_{tot}$ is the total reflection of the pixel, $\mathbf{R}(\tau_T^{cld})$ is the reflection of the cloudy part of the pixel, $\mathbf{R}(\tau_0)$ the reflection of the clear part of the pixel, $\mathbf{R}(\tau_0^{aero})$ the reflection of the aerosol part of the pixel, $\sum_{i=1}^{N} Fcld_i$ the total fraction of only cloud pixels and N is the total number of pixels (cloudy and clear), N_{aero} the number of **L3** pixels that have AOT>0 and N_{clear} the number of clear **L3** pixels. N_{aero} can thus never exceed N_{clear}.

The last parameter, the aerosol top pressure, is not a dataset in MODIS. To keep the computations simple we will use one top pressure. According to Ginoux et al. (2001), decreases the amount of dust aerosols rapidly with height. They found in their simulations that the highest abundance of dust is located near $\sim 800\ mb$. We will also use an aerosol top pressure (ATP) of $800\ mb$ with a vertical extend similar to our clouds of $100\ mb$.

7.4. Simulated Earthshine with the Earth model

In the attempt to approach the Earthshine data we will perform different simulations that consist of various combinations of land cover models and atmospheric models. We will primarily distinguish two types of simulations, one with a nominal cloud cover that we have used so far in all of the presented simulations, but also a reduced cloud cover data set, i.e. the Retrieval Cloud Fraction (for more information about these data sets consult *Section* 3.3). We will first perform a simulation that utilizes only Lambertian reflecting surfaces. In the subsequent three sections, we utilize the polarizing ocean model, the polarizing ocean model plus polarizing desert model and lastly the polarizing ocean, desert and vegetation models. For more information about these models one can consulted *Chapter* 6. In the fifth section, we will include dust aerosols in our model in combination with the polarizing surfaces. In the last section we will conclude on the findings in the previous sections and discuss possible shortcomings.

Lambertian surfaces

As a first order approximation we modeled our Earth-like exoplanet model with Lambertian surfaces, as we have already done in *Chapter* 4 and 5. In *Figure* 7.6 one can see the simulations with nominal cloud fraction in the upper two panels and that with reduced cloud fraction in the lower panels. In the two left panels the simulation for April is provided. For the nominal cloud fraction these show to correspond closely to two quasi-homogeneous simulations by Sterzik et al. (2012), but provides a slightly higher fraction op polarization in the red. The approximation in this part of the spectral domain is, however, still far to low. With reduced cloud cover we see a major increase in the blue part of the spectrum and a slight increase in the red part. We have seen that clouds are generally highly depolarizing and have a relatively high reflection. When smaller cloud fractions are used the degree of polarization will thus

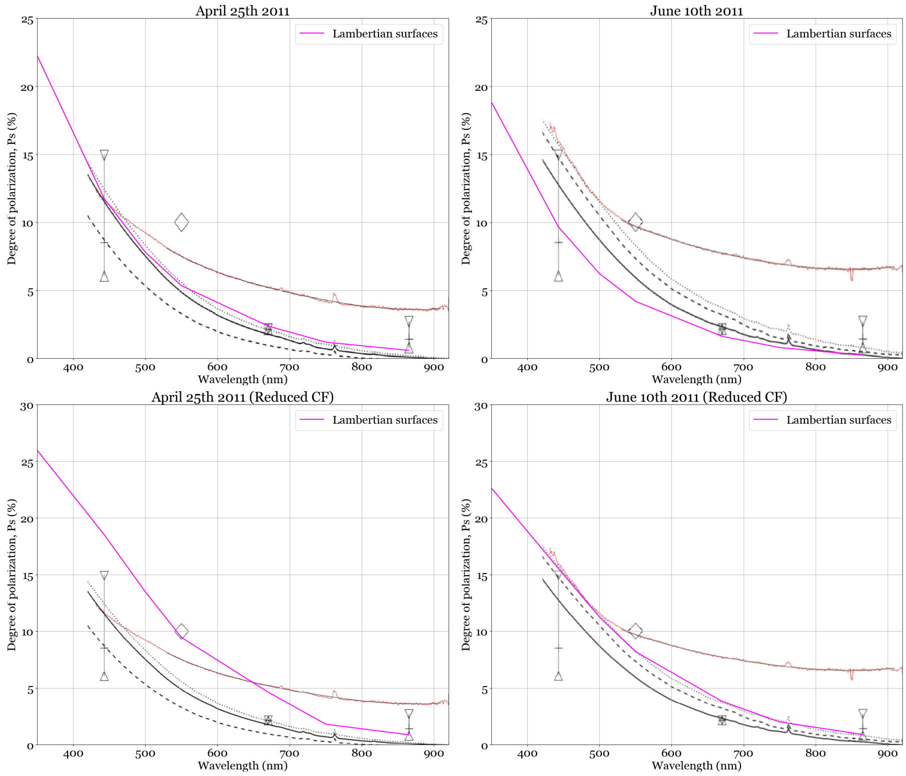


Figure 7.6: In this figure one can find the simulations of Earthshine for April 25th 2011 (left panels) and June 10th 2011 (right panels). The simulations were performed with a nominal cloud cover (upper panels) and a reduced cloud cover (bottom panels) (*Section* 3.3). The surfaces are approximated by Lambertian reflection. The red line represents the observed Earthshine data, whereas the solid, dotted and dashed lines represent Quasi-homogeneous simulations.

increase. The behaviour is largest at blue wavelengths as Rayleigh scattering becomes very effective. The simulations of June are provided in the right two panels. For the nominal cloud fractions we see a large underestimation of Ps in the blue and the red altogether, where even the quasi-homogeneous simulations provide a better fit in the blue and green parts of the spectrum. For the reduced cloud fractions we see a much better fit and approximate the blue/greenish part of the Earthshine well. In the red part we see a major underestimation, but a slightly better fit than the quasi-homogeneous simulations.

Lambertian+polarizing ocean surfaces

By adding a polarizing ocean surface to our model we already saw that the effect of the glint has major influence on F and in particular Q (*Chapter* 6). For both epochs, we expect a major effect for June. In that configuration the fictive position of the glint is right on top of the ocean cover, whereas depending on the wind speed this might only be partly true for April. From our simulations, presented in *Figure* 7.7, one can see that for our nominal cloud cover we see only small improvements for both epochs, where the effect is largest for June. The simulations with reduced cloud fractions show a much larger increase in Ps, especially for June. This difference between these simulations is devoted to the fact that much less clouds cover the surface, and potentially the position of the glint; allowing for a major increase in polarized flux Q. For April we see that this causes an even bigger overestimation in the blue, but a fairly good approximation in the red part of the spectrum. For June we see a similar effect but with much better agreement to the Earthshine.

Lambertian+polarizing ocean and desert surfaces

In *Chapter* 6 we constructed a polarizing desert surface. The effect of this polarizing desert surface as compared to a Lambertian equivalent is shown in *Figure* 7.8. In the right two panels we see that this addition provides no significant improvement to our model in June. If one looks back to *Figure* 7.2 the cause can directly be observed. In June there is barely any land cover visible that we considered as desert. Furthermore, the visible desert regions are located far from the center of the disk and

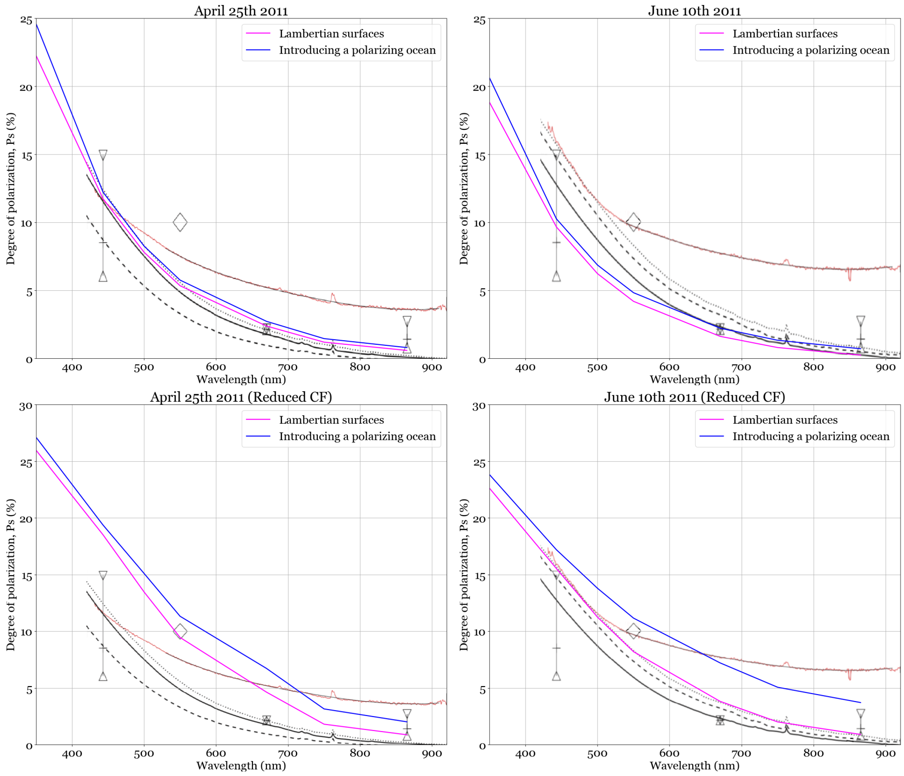


Figure 7.7: Similar to *Figure* 7.6, but now we added a polarizing ocean surface to the clear and "thin"-clouded (COT=5) pixels.

contribute little to the spectropolarimetric signal. For April we see a slightly higher contribution, but again very small, even in the case of reduced cloud cover. This is mostly attributed to the fact that in April the cloud cover above the Sahara Desert in both cases is very low to even completely clear.

All polarizing surfaces

From the disk-resolved cases in *Chapter* 4 and 6 we saw that vegetation has quite some influence on the normalized reflected flux in the red wavelength region. In the Earthshine simulations, presented in *Figure* 7.9, we see a significant contribution of this polarizing model for April, especially for the reduced cloud cover run. We essentially see a tilting effect that we require to approximate the Earthshine measurements better, where Ps in the red part increases and in the blue spectral regions decrease. For the nominal cloud cover we, however, do not see a significant improvement in the red part of the spectrum. For the reduced cloud cover we now see a fairly accurate approximation in the red but a majorly overestimated signal in the blue/green. For June the contributions are far less if one compares them to April. This is again caused by the low amount of vegetated surface pixels and the fact that they all lie far from the center of the disk. We see a similar rotation as in April, but with nominal cloud cover the estimation maintains rather worse. For the reduced cloud cover, which showed already good agreement, we keep an overestimation in the blue and a larger underestimation in the red.

All polarizing surface+dust aerosols

The effect of the rather experimental addition of dust aerosols in our model atmosphere are presented in *Figure* 7.10. For both April and June with nominal cloud fraction we barely see the influence of these aerosols. If we closely look at the greenish/blue part of the spectrum we see for both epochs a decrease in Ps. For April we see a very good agreement in the blue of $< 1\%$, but a $\sim 3-4\%$ disagreement in the red. The worst approximation is found for the simulation of June, where we find a disagreement of $\sim 5-6\%$ in the blue and $\sim 6-7\%$ in the red. If we now look at the reduced cloud fraction cases, an enhancement of the latter behaviour can be observed. In these cases we also see a slight decrease in the red part of the spectrum. For June this provides for a good agreement in the blue, but a rather weak agreement in the red were we are still off by $\sim 3-4\%$. For April we see an overestimation of $\sim 6-7\%$ at the blue and only $\sim 1\%$ at the red part of the spectrum.

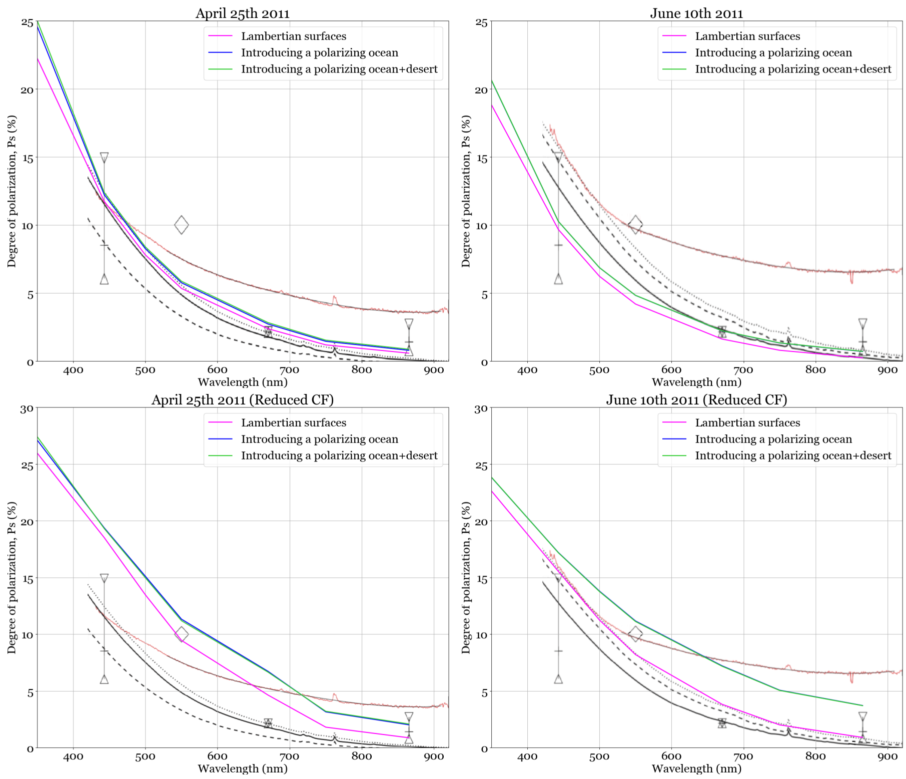


Figure 7.8: Similar to *Figure* 7.7, but now we also added a polarizing desert surface to the clear and "thin"-clouded (COT=5) pixels.

Concluding remarks on the simulated Earthshine data

In the previous sections we have seen that we were not able to accurately agree with the retrieved Earthshine data, despite the fact that we constructed a very comprehensive model. In conclusion, we have seen that by adding the polarized surfaces for the ocean, desert and vegetation the overall agreement with the Earthshine data increased. By adding dust aerosols, we observed a slight tilt, decreasing the degree of polarization Ps at the red wavelengths and increasing Ps in the blue. Hence, the addition of aerosols induces a slight decrease in the overall agreement. To further investigate the cause of this disagreement, we will present some simulations for which we customized the cloud distribution for both dates. All of the following simulations are performed without considering dust aerosols.

For comparison, we have plotted all cases in a single panel. In the left panel of *Figure* 7.11 we simulated April 25th of 2011 and in the right panel June 10th of 2011. The simulation is made using all polarizing surfaces without aerosols is labeled with "Full model".
Throughout this thesis we have showed that clouds generally depolarize reflected light. To get a feel for what range of Ps values we can physically model, an end case that simulates our land cover distribution without clouds is computed. To clarify, we have considered clear sky models for the ocean, vegetation, desert and ice cover pixels. From this simulation one can see that we basically overestimated the Earthshine data for the entire spectral domain, which basically tells us that it is possible to approximate the red part of the spectrum. Albeit, the overestimation for April is only minor, which tells us that under real cloudy conditions full agreement with the Earthshine data might seem unrealistic. Contrarily, for June we see a large overestimation, essentially telling us that by altering the cloud distribution even more an overall agreement to the data might be possible. It has to be noted that in the previous sections we already saw a much better agreement for June than for April at least when we used the reduced cloud cover.
In the previous chapter we saw that the most dominant polarizing land cover type for all wavelengths considered is the ocean, provided that no land cover masks the position of the glint. For the next simulation we do not consider clouds for the ocean pixels only. This shows an overall decrease of Ps and a close agreement in the red for April, but still a large disagreement at the blue. For June we still

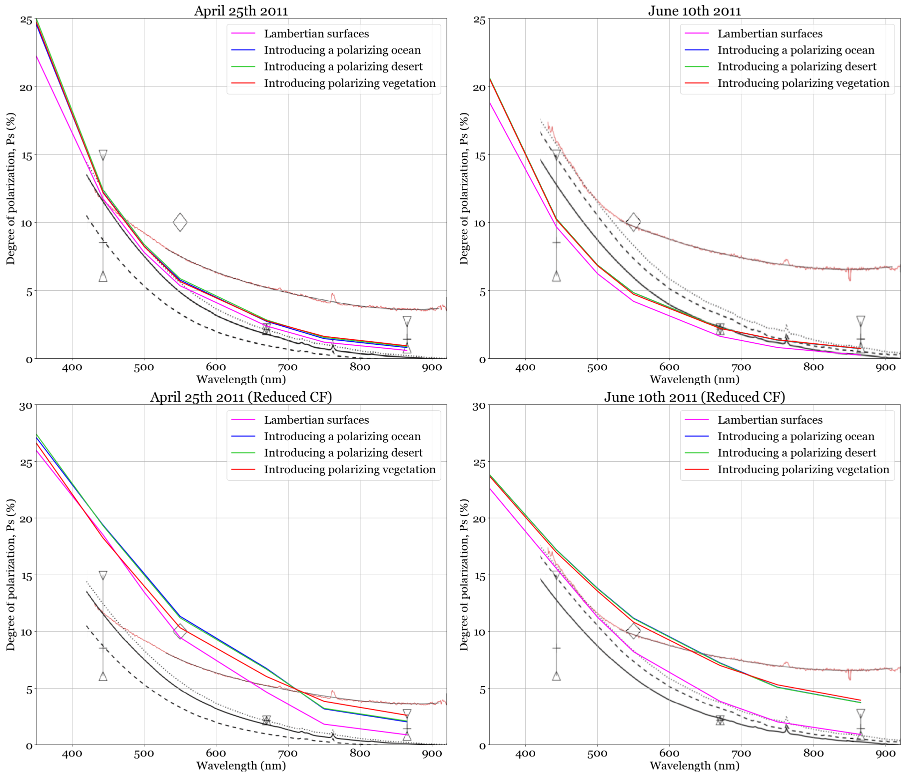


Figure 7.9: Similar to *Figure* 7.8, but now we also added a polarizing vegetation surface to the clear pixels.

see a large overestimation for the entire spectrum.
To be somewhat more realistic we will again include cloud over the ocean, except for a small portion of pixels that form the highly reflecting part of the glint. We want to see the effect of the glint alone, because we know that the dominant polarizing part of the ocean is the glint. The simulations show that overestimation is completely gone for both epochs and that we again underestimate the Earthshine data with this configuration. On the one hand this might seem unexpected as we allowed the extremely polarizing glint to radiate relatively easy through the exoplanet atmosphere, but the addition of the highly reflecting cloud on top of the other ocean pixels has a dominant effect on Ps.
modeling an even larger clear sky region at the ocean glint shows to have little effect for April, because the glint is partly "masked" by the African continent. For June one can see a rotation towards the Earthshine data, but again to small for a significant agreement.
The last case we will model is that of the extended clear region over the glint in combination with the reduced cloud fraction data set. For both observations this shows significant agreement at the red part of the spectrum, but especially a major overestimation at the blue region for April.

We have seen that we were able to approximate the Earthshine data at the near-infrared wavelengths quite well by altering the distribution of the clouds. This shows that the exact distribution of the clouds is quite dominant. The MODIS observations that we use are not retrieved exactly at the time of observation and thus might be the source of the initial disagreement in the previous sections. Although, we have seen that for our most accurate simulations there was either an overestimation in the blue, an underestimation in the red wavelengths or both. What we essentially want is a tilt towards higher Ps in the red and towards lower Ps in the blue wavelength region. This is both the case for the observed data in April and June and might suggest that the approximate depolarizing factor of the Lunar surface does not allow full agreement at all.

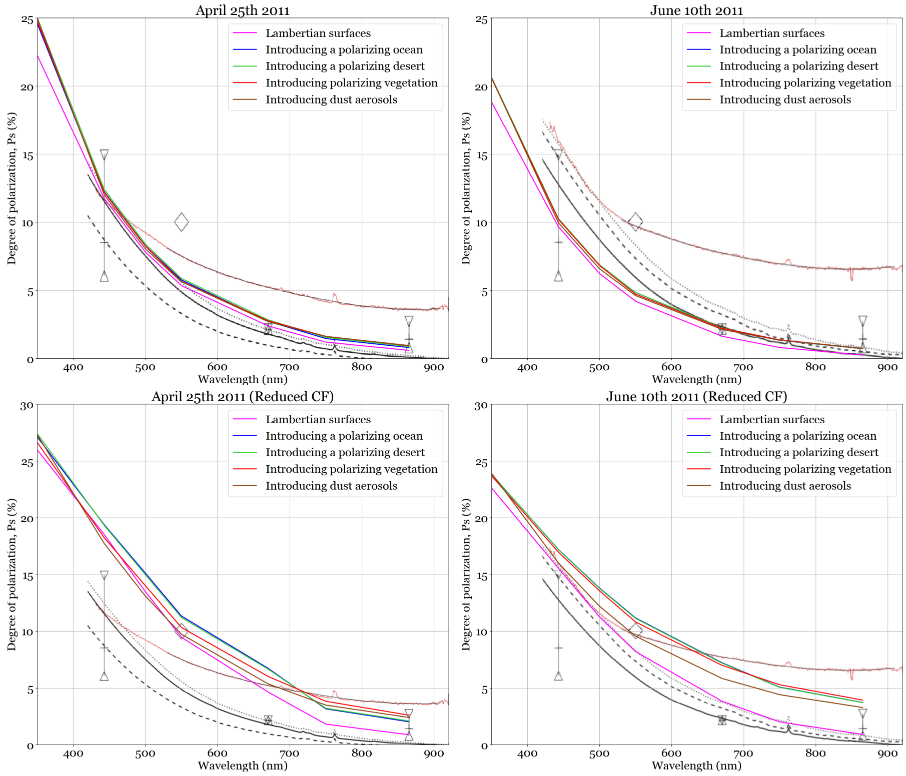


Figure 7.10: Similar to *Figure* 7.9, but now we added aerosols to clear pixels according to MODIS data.

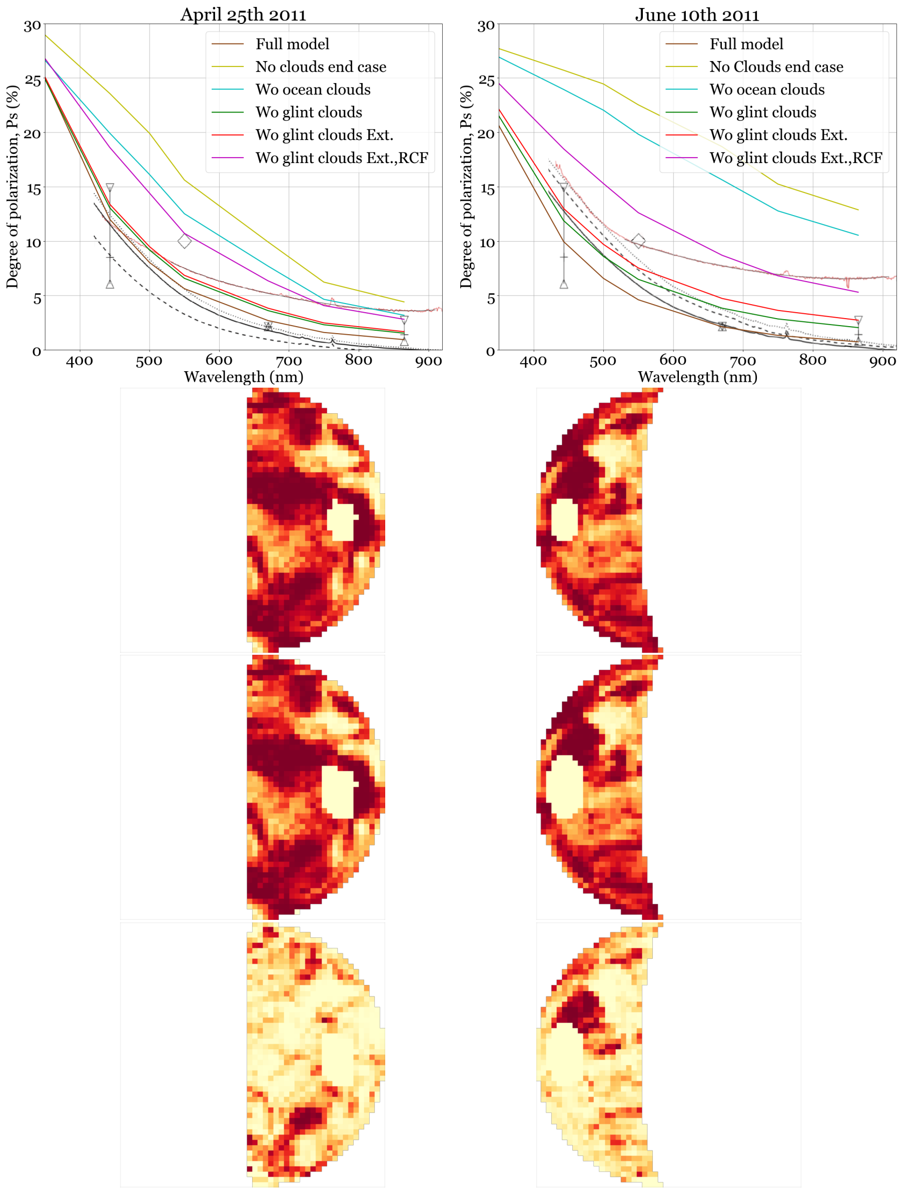


Figure 7.11: Earthshine simulations with customized cloud covers. In the upper left and right panels we provide the different cases for April 25th 2011 and June 10th 2011, respectively. In the lower panels we provided the customized cloud fraction distributions for both epochs, from top to bottom for the glint, the extended glint and the extended glint in combination with the reduced cloud fraction data set. For the lower panels yellow corresponds to a zero cloud fraction and red to a cloud fraction of one.

8

Discussion and Conclusions

The research objective was stated as follows:

> *The research objective is to retrieve spectropolarimetric signals from an Earth-like exoplanet model in an edge-on configuration to be able to rationalize future disk integrated observations, by use of a radiative transfer algorithm in combination with Earth observations.*

In this research we provided simulations of (polarized) flux reflected from an edge-on Earth-like exoplanet. The model that we implemented is based on daily varying MODIS data. More specifically, the cloud layer in our plane parallel vertically inhomogenous local surface-atmosphere system varies spatially and temporally. The invariant cloud layer is modeled by its cloud top pressure, cloud optical thickness and cloud particle effective radius. The surface discretization is based on the most dominant surface types: oceanic, vegetated, desert and ice/snow surfaces. The gaseous part of the model atmosphere is invariant and assumed to be in hydrostatic equilibrium. A discretization of the cloud data sets allows us to model a horizontally inhomogeneous planetary disk that utilizes 36 different cloudy models and one model with a pure gaseous atmosphere. Additionally, we utilized the ability to include anisotropic polarizing surfaces, which model vegetated and oceanic land covers[1]. To account for the desert surface, we constructed an anisotropic polarizing desert model from empirical data of "Olivine-S" dust particles. For the snow/ice surface pixels we continued to use a Lambertian reflecting surface. The main conclusions are provided next.

The Earth-like model

By simulating our Earth-like model, with Lambertian surfaces for all land cover types, as resolved disks we were able to retrieve the direct contribution of the surface and atmosphere to a disk-resolved signal. These disks are computed at $\lambda = 350,\ 550$ and $865\ nm$ at the same sub-observer longitude. For longer wavelengths, the contribution of surface reflection and reflection from the cloud layers become more apparent in F (total flux). The surface reflection does not contribute to the polarized fluxes Q and U for all λ. The sensitivity of Pl (degree of polarization) to the surface reflection is thus induced solely by F. The clouds only show strong affects on the disks of Q near full phase and $\alpha = 40^{\circ}$.

The sensitivity of polarized flux Q on scattering from clouds is caused by the micro-physical properties of the cloud particles considered. We model spherical liquid water particles to form our cloud layer. Light beams which scatter once on these particles exhibit a high degree of reflection and polarization at certain scattering angles. Depending on the optical depth, vertical position and particle effective radii, enhancements in the planetary phase curves appear, generally known as the glory near full phase, the primary rainbow near $\alpha \approx 40^{\circ}$, and the secondary rainbow near $\alpha \approx 56^{\circ}$. By inspection of the planetary phase curves at $\lambda = 350,\ 443,\ 550,\ 670$ and $865\ nm$ (*Figure* 4.12) we know that these signatures decrease in strength and appearance for shorter λ. Q provides a valuable tool to retrieve the primary rainbow as it remains visible for all wavelengths considered. For F and Pl we compared the strength of the primary rainbow to Karalidi, Stam and Hovenier (2012), showing that the primary rainbow is much more apparent in our simulations, especially for F, because our cloud

[1]See Cheung (2018) and Trees (2018) for a full description of the vegetation and ocean-atmosphere model, respectively.

layers are optically thicker and consist of particles with larger effective radii. On all phase curves, daily variations are induced, because (1) we simulate the horizontally inhomogeneity of cloud and surface covers and (2) the planet rotates around its own axis. F, Pl and U are most sensitive to these spatial and temporal inhomogeneities, whereas in Q these are apparent at $350\ nm$ only. In F we retrieved an intersection point around $\alpha = 120°$, where the phase curves at different λ alternate. This is caused by the decreasing dominance of the gaseous atmosphere on top of the clouds at long λ, allowing more light to be reflected on the cloud layers. For Pl, Q and U no such reversal is retrieved.
The seasonality in cloud observations of 2011 induce a region of variability that attains the same order of relative magnitude, on the continuum of the planetary phase curves, as the daily variations. Hence, any retrieval of seasonality in an exoplanetary atmosphere from these phase curves is ambiguous.
For all simulated planetary phase curves we retrieve very small values of U. There withal we were not able to retrieve any dominant signature other than some unsubstantiated suggestions.

Retrieval of cloud variability
The daily variability induced on the planetary phase curves allowed us to apply the *Discrete Fourier Technique* and the *autocorrelation method* to retrieve Earth's rotation period. We showed that this retrieval greatly depends on the level of noise in the photometric and polarimetric signal, the temporal sampling and observation interval, i.e. the range of phases on which the methods are applied. Due to time constraints we only considered the phase curve at $550\ nm$. Exploiting the ability to retrieve the correlation at multiple consecutive rotation periods with the autocorrelation method, we found that the presence of a temporally invariant cloud cover can be retrieved by comparing the autocorrelation at multiple consecutive rotation periods from Q to that of F and/or Pl.

The changing cloud cover provides different mean cloud parameters[2] for any specific phase angle and cloud observation. The question arose whether these changes are correlated to variability in F, Pl, Q and/or U. To investigate this, we simulated a large set of data points at $\alpha = 40°$ and $\alpha = 90°$. First of all, we found that U does not provide any significant correlation with any of the mean cloud parameters. Also, surprisingly, we found that the cloud optical thickness has no correlation with any of the Stokes parameters or Pl, although we found that in the resolved disks this parameter showed significant dominance. For a set of λ on both phases we found significant correlations for the mean cloud fraction, top pressure and particle effective radius. Using both phases simultaneously, we defined a shape parameter. With this parameter we were able to reproduce the more idealized results from Rossi and Stam (2017). Furthermore, we found the following pronounced correlations:

1. Cloud fraction is correlated to F at $\lambda = 350\ nm$;
2. Cloud fraction is correlated to Pl at $\lambda = 550\ nm$;
3. Cloud fraction is correlated to Q at $\lambda = 550$ and $865\ nm$;
4. cloud particle effective radius is correlated to F at $\lambda = 865$;
5. cloud particle effective radius is correlated to Pl at $\lambda = 550$ and $865\ nm$;
6. Cloud top pressure is correlated to Pl at $\lambda = 550\ nm$;
7. Cloud top pressure is correlated to Q at $\lambda = 350\ nm$.

Extended Earth-like model
The Lambertian surface approximation completely depolarizes any reflected light. By implementation of realistic anisotropic polarizing surface models we were able to simulate more realistic spectropolarimetric signals. We incorporate vegetation models for steppe and deciduous forests, a wind speed dependent ocean model, for which we provide observations with a temporal resolution of 6 hours, and a desert model fitted to the Entisol specimen.

From the disk-resolved simulations we retrieve a stronger distinction in F between vegetated and desert surface covers for all λ and phases. The polarized reflection of these two covers cause apparent features of their spatial distribution in both Q and U. A major appearance of the specular reflection

[2]The mean is computed from of all visible and illuminated cloudy pixels.

from the ocean cover, i.e. the glint, is retrieved for F, Q and Pl, being most apparent for Q.

The dominance of this glint in the spectropolarimetric signals is also well retrieved in the planetary phase curves. From the phase curves at different λ (*Figure* 6.10), we are able to unambiguously retrieve the presence of the ocean under the influence of (1) realistic ocean surface wind observations, (2) the presence of an Earth-like cloud cover, and (3) the presence of continents. More specifically, the specular reflection from the ocean causes an intersection point of the phase curves at different wavelengths in F, Pl and Q. In particular, the intersection point in Q is found to be solely caused by the presence of a ocean on an Earth-like exoplanet. Also, the locations of the intersection point in F and Q allowed us to retrieve an estimate of the mean cloud fraction of $0.65 - 0.7$, agreeing well to the mean cloud fraction from MODIS data of 0.68.

Earthshine simulations

Using the Earth-like model with anisotropic polarizing surface, we have attempted to approximate the observed Earthshine data on April 25th 2011 and June 10th 2011. In earlier attempts by Emde et al. (2017); Sterzik et al. (2012), full agreement with the data was not found. In our model we also incorporate dust aerosols at a fixed vertical position and with an optical thickness according to MODIS data. These aerosols are modeled with the scattering matrices from the anisotropic polarizing desert model. Our simulations show moderate agreement for all λ, which can be caused by (1) neglecting the presence of other aerosols, such as maritime aerosols, or (2) the approximation of the correction for the depolarizing behaviour of the Lunar surface.

Final remarks

Before we started this study the following research questions were set:

1. What is the spectropolarimetric signal for a resolved and unresolved Earth-like exoplanet?
   (a) How does light reflect from an Earth-like exoplanet?
   (b) How can an Earth-like exoplanet be modeled?
   (c) Which features characterize Earth and how can these features be used in future exoplanet characterization?
2. Which signatures from spectropolarimetric signals can be identified such that Earth-like exoplanets can be characterized?
   (a) Can Earth biomarkers be characterized in spectropolarimetric signals?
   (b) Can spectropolarimetric signatures characterize exoplanet surfaces?
   (c) Can spectropolarimetric signatures characterize exoplanet atmospheres?
   (d) How can spectropolarimetry be used to identify planetary and orbital elements?

In *Section* 2.1 we provided a basic understanding of how light reflects from an Earth-like exoplanet (1a). In this research we used the radiative transfer code PyMieDAP in combination with Earth observations to construct and model an Earth-like exoplanet. The steps that were taken to introduce MODIS data as input for PyMieDAP in order to construct a horizontally inhomogeneous planetary disk is provided in *Section* 2.3 (1b). The construction of the actual planetary model from these observations is provided in *Chapter* 3 (1c).
In *Chapter* 4 we showed that the red edge feature in vegetation is clearly visible in the disk-resolved cases. From the phase curves, we could not retrieve any unambiguous signature related to this feature. Also, we found that the primary rainbow can be retrieved confidently for the full spectral domain from which we can potentially characterize the clouds in an exoplanet atmosphere. We were not able to retrieve an unambiguous signature of the gaseous atmosphere, because (1) it is spatially and temporally invariant and (2) we did not consider absorption of the gaseous constituents. Induced by the partially ocean covered exoplanet surface, an unambiguous intersection point is retrieved in the spectropolarimetric signal Q, whose position also provides an indication of the cloud cover (2a,b,c). In *Chapter* 5 we provided estimates on the retrieval of the rotation rate in photometric as well as polarimetric signals (1d).

We have discussed the presence of spectropolarimetric signatures, which can potentially be directly retrieved from, or aid in, the interpretation of future exoplanet observations. Moreover, in the design of future telescopes the characteristics of these signatures may be considered. These signatures are retrieved from photometric and polarimetric signals, which are created using a horizontally inhomogeneous model. This model allowed us to include the spatially variability in cloud and surface cover, for example the appearance of the glint through the patchy cloud cover or its absence when continents are in sight. Conclusively, utilizing a set of wavelengths could potentially allow one to retrieve information about the presence, abundance and micro-physical properties of clouds in the gas atmosphere of, and also the presence of an ocean cover on, an Earth-like exoplanet.

9
Recommendations

The recommendations for future followup studies are itemized as follows:

- General optimization of the PyMieDAP code to allow one to increase the number of cloud types, i.e. the discretization of the cloud observations. This will also allow the user to define more types of surface cover, but we expect that this smaller discretization has little effect. Also, it would be interesting to model the spatial and temporal variability of specific constituents in the gaseous atmosphere, e.g. O_3. Such a study would be most effective if absorption of these constituents is also considered.
- Recompute the Stokes vectors to include the circularly polarized fluxes. Muñoz (2015); Rossi and Stam (2018) already provided simulations of circularly polarized fluxes, but no such analysis has been provided for a large set of cloud types that spatially vary in time according to Earth observations. For example, it might be possible to retrieve some effect of seasonality as this was not unambiguously retrieved in our simulations.
- By modeling aerosols in our atmosphere, we slightly altered the spectral dependence of the Earthshine simulation. It is interesting to construct more realistic models not only for Saharan dust, but also for e.g. maritime aerosols, ice clouds, biomass burning, etc. and assessing their effect on the spectropolarimetric signals. For ice clouds, however, a thorough analysis is already provided by Karalidi, Stam and Hovenier (2012).
- Analyze the relations of Stokes vectors with mean cloud parameters by weighting them with the position of the disk. By inspection of *Equation* 2.39 we know that the reflection of each pixels is weighted by the incident and emission zenith angles. By accounting for the position of the pixels, a better fit may be achieved. Furthermore, in the correlations found we did not take into account the polarizing surfaces. By inspection of *Figures* 4.12 and 6.10 we expect that especially for Q at $\alpha = 90°$ the surfaces will attain a dominant role. Hence, it is also recommended to include the fractions of surface cover into this analysis.
- The planetary model that is constructed bases the surface cover distribution on annual MODIS observations. To be able to retrieve the seasonal changes in this cover it is advised to include observations of ice/snow and vegetation cover that vary on a shorter temporal scale.
- The retrieval of dynamic weather by use of the autocorrelation method showed some promising results. However, due to time limitation we were only able to provided the results at $\lambda = 550\ nm$. In a further study it is advised to also take into account the ultra-violet and near-infrared spectral region as the reflection from clouds and surface covers are influenced by the strong wavelength dependence of the gaseous atmosphere. Furthermore, we provided that the temporal resolution of, and noise levels in, the reflected signals have major effect on the positive retrieval of the rotation period at $550\ nm$, but how are the retrievals at other wavelength affected? Moreover, telescopes in the near and far future require certain integration times, limiting the temporal resolution. In addition, a study into the expected rotation periods of exoplanets is required to assess the applicability of such integration times.

- The unambiguous signature of an ocean-atmosphere system in Q was first retrieved by Trees (2018). By incorporating their ocean model in our planetary model we were able to retrieve that signature under the influence of realistic ocean surface wind observations, the presence of an Earth-like cloud cover and the presence of continents. Furthermore, they showed that this signature is suppressed when the cloud cover is completely overcast and the surface pressure attains $\sim 10\ bar$. From our point of view it is interesting to investigate to what extend the ocean surface fraction can be lowered to still retrieve this signature, where not only the fraction of ocean cover plays a role but also the spatial distribution as the ocean glint is per definition located on the planetary scattering plane.

- A code to asses the feasibility of quasi homogeneous approximations (see e.g. Stam 2008*a*) is already available. In a future study it would be interesting to investigate whether the large inhomogeneities in the cloud and surface cover can be approximated well with a weighted sum of horizontally homogeneous planet end cases. This quasi homogeneous approximation would allow the user to significantly lower the computational effect.

- In our analysis the signatures of vegetated and desert land covers were not retrieved, which is most definitely due to the less distinct anisotropy (polarized) reflection of these covers as compared to for example an ocean cover. For vegetation we know that the red-edge enhancement in the spectrum of total reflection is a very important biosignature (Berdyugina et al. 2016; Horler et al. 1983). In previous studies this signature was found for total flux as well as for the degree of polarization (Hamdani et al. 2006; Montañés-Rodríguez et al. 2006; Sterzik et al. 2012; Tinetti et al. 2006). In a followup study it would be interesting to increase the spectral resolution, at least around the red-edge, to assess whether this biosignature can also be retrieved from spectropolarimetric signal of an Earth-like exoplanet.

Bibliography

Ackerman, S. A., Strabala, K. I., Menzel, W. P., Frey, R. A., Moeller, C. C. and Gumley, L. E. (1998), 'Discriminating clear sky from clouds with modis', *Journal of Geophysical Research: Atmospheres* **103**(D24), 32141–32157.

Ackerman, S., Holz, R., Frey, R., Eloranta, E., Maddux, B. and McGill, M. (2008), 'Cloud detection with modis. part ii: validation', *Journal of Atmospheric and Oceanic Technology* **25**(7), 1073–1086.

Bailey, J. (2007), 'Rainbows, Polarization, and the Search for Habitable Planets', *Astrobiology* **7**, 320–332.

Berdyugina, S. V., Kuhn, J. R., Harrington, D. M., Šantl-Temkiv, T. and Messersmith, E. J. (2016), 'Remote sensing of life: polarimetric signatures of photosynthetic pigments as sensitive biomarkers', *International Journal of Astrobiology* **15**, 45–56.

Bogumil, K., Orphal, J., Homann, T., Voigt, S., Spietz, P., Fleischmann, O., Vogel, A., Hartmann, M., Kromminga, H., Bovensmann, H. et al. (2003), 'Measurements of molecular absorption spectra with the sciamachy pre-flight model: instrument characterization and reference data for atmospheric remote-sensing in the 230–2380 nm region', *Journal of Photochemistry and Photobiology A: Chemistry* **157**(2-3), 167–184.

Cheung, J. C. (2018), The Flux and Linear Polarization of Vegetated Earth-like Exoplanets, Master's thesis, Delft University of Technology.

Chowdhary, J., Cairns, B. and Travis, L. D. (2006), 'Contribution of water-leaving radiances to multiangle, multispectral polarimetric observations over the open ocean: bio-optical model results for case 1 waters', *Applied Optics* **45**, 5542–5567.

Ciddor, P. E. (1996), 'Refractive index of air: new equations for the visible and near infrared', *Applied Optics* **35**, 1566.

Coakley, J. (2003), 'Reflectance and albedo, surface'.

Cowan, N. B., Abbot, D. S. and Voigt, A. (2012), 'A False Positive for Ocean Glint on Exoplanets: The Latitude-Albedo Effect', *The Astrophysical Journal* **752**, L3.

Cowan, N. B., Agol, E., Meadows, V. S., Robinson, T., Livengood, T. A., Deming, D., Lisse, C. M., A'Hearn, M. F., Wellnitz, D. D., Seager, S. and Charbonneau, D. (2009), 'Alien Maps of an Ocean-bearing World', *The Astrophysical Journal* **700**, 915–923.

Cox, C. and Munk, W. (1954), 'Measurement of the roughness of the sea surface from photographs of the sun's glitter', *Journal of the Optical Society of America (1917-1983)* **44**, 838–850.

Daimon, M. and Masumura, A. (2007), 'Measurement of the refractive index of distilled water from the near-infrared region to the ultraviolet region', *Applied optics* **46**(18), 3811–3820.

de Haan, J. F., Bosma, P. B. and Hovenier, J. W. (1987*a*), 'The adding method for multiple scattering calculations of polarized light', *Astronomy & Astrophysics* **183**, 371–391.

de Haan, J. F., Bosma, P. B. and Hovenier, J. W. (1987*b*), 'The adding method for multiple scattering calculations of polarized light', *Astronomy & Astrophysics* **183**, 371–391.

de Rooij, W. A. and van der Stap, C. C. A. H. (1984), 'Expansion of Mie scattering matrices in generalized spherical functions', *Astronomy & Astrophysics* **131**, 237–248.

Deschamps, P.-Y., Breon, F.-M., Leroy, M., Podaire, A., Bricaud, A., Buriez, J.-C. and Seze, G. (1994), 'The POLDER mission: instrument characteristics and scientific objectives', *IEEE Transactions on Geoscience and Remote Sensing* **32**, 598–615.

Diem, M. (1948), 'Messungen der grosse von wolkenelementen ii', *Meteor. Rundschau* **1**, 261–273.

Domke, H. (1974), 'The expansion of scattering matrices for an isotropic medium in generalized spherical functions', *Astrophysics & Space Science* **29**, 379–386.

Emde, C., Buras-Schnell, R., Sterzik, M. and Bagnulo, S. (2017), 'Influence of aerosols, clouds, and sunglint on polarization spectra of earthshine', *Astronomy & Astrophysics* **605**, A2.

EOS Data Products Handbook Volume 1 (2000) (n.d.), https://eospso.nasa.gov/sites/default/files/publications/data_products_1.pdf. Accessed: 2018-03-20.

EOS Data Products Handbook Volume 2 (2000) (n.d.), https://eospso.gsfc.nasa.gov/sites/default/files/publications/data_products_vol2_0.pdf. Accessed: 2018-03-21.

Ford, E. B., Seager, S. and Turner, E. L. (2001), 'Characterization of extrasolar terrestrial planets from diurnal photometric variability', *Nature* **412**, 885–887.

Fouquart, Y., Buriez, J., Herman, M. and Kandel, R. (1990), 'The influence of clouds on radiation: A climate-modeling perspective', *Reviews of Geophysics* **28**(2), 145–166.

Friedl, M. A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A. and Huang, X. (2010), 'MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets', *Remote Sensing of Environment* **114**, 168–182.

Fujii, Y., Kawahara, H., Suto, Y., Fukuda, S., Nakajima, T., Livengood, T. A. and Turner, E. L. (2011), 'Colors of a Second Earth. II. Effects of Clouds on Photometric Characterization of Earth-like Exoplanets', *The Astrophysical Journal* **738**, 184.

Fujii, Y., Kawahara, H., Suto, Y., Taruya, A., Fukuda, S., Nakajima, T. and Turner, E. L. (2010), 'Colors of a Second Earth: Estimating the Fractional Areas of Ocean, Land, and Vegetation of Earth-like Exoplanets', *The Astrophysical Journal* **715**, 866–880.

Ginoux, P., Chin, M., Tegen, I., Prospero, J. M., Holben, B., Dubovik, O. and Lin, S.-J. (2001), 'Sources and distributions of dust aerosols simulated with the gocart model', *Journal of Geophysical Research: Atmospheres* **106**(D17), 20255–20273.

Gisler, D., Schmid, H. M., Thalmann, C., Povel, H. P., Stenflo, J. O., Joos, F., Feldt, M., Lenzen, R., Tinbergen, J., Gratton, R. et al. (2004), Cheops/zimpol: a vlt instrument study for the polarimetric search of scattered light from extrasolar planets, *in* 'Ground-based instrumentation for astronomy', Vol. 5492, International Society for Optics and Photonics, pp. 463–475.

Hahn, C. J., Rossow, W. B. and Warren, S. G. (2001), 'Isccp cloud properties associated with standard cloud types identified in individual surface observations', *Journal of Climate* **14**(1), 11–28.

Hale, G. M. and Querry, M. R. (1973), 'Optical constants of water in the 200-nm to 200-μm wavelength region', *Appl. Opt.* **12**(3), 555–563.
URL: *http://ao.osa.org/abstract.cfm?URI=ao-12-3-555*

Hamdani, S., Arnold, L., Foellmi, C., Berthier, J., Billeres, M., Briot, D., François, P., Riaud, P. and Schneider, J. (2006), 'Biomarkers in disk-averaged near-uv to near-ir earth spectra using earthshine observations', *Astronomy & Astrophysics* **460**(2), 617–624.

Han, Q., Rossow, W. B. and Lacis, A. A. (1994*a*), 'Near-global survey of effective droplet radii in liquid water clouds using isccp data', *Journal of Climate* **7**(4), 465–497.

Han, Q., Rossow, W. B. and Lacis, A. A. (1994*b*), 'Near-global survey of effective droplet radii in liquid water clouds using isccp data', *J. Climate* **7**, 465–497.

Hansen, J. E. (1971), 'Multiple Scattering of Polarized Light in Planetary Atmospheres Part II. Sunlight Reflected by Terrestrial Water Clouds.', *Journal of Atmospheric Sciences* **28**, 1400–1426.

Hansen, J. E. and Travis, L. D. (1974*a*), 'Light scattering in planetary atmospheres', *Space Science Reviews* **16**, 527–610.

Hansen, J. E. and Travis, L. D. (1974*b*), 'Light scattering in planetary atmospheres', *Space Science Reviews* **16**, 527–610.

Hoeijmakers, H. J., Arts, M. L. J., Snik, F., Keller, C. U., Stam, D. M. and Kuiper, J. M. (2016), 'Design trade-off and proof of concept for LOUPE, the Lunar Observatory for Unresolved Polarimetry of Earth', *Optics Express* **24**, 21435.

Horler, D. N. H., Dockray, M. and Barber, J. (1983), 'The red edge of plant leaf reflectance', *International Journal of Remote Sensing* **4**, 273–288.

Hough, J. H., Lucas, P. W., Bailey, J. A. and Tamura, M. (2003), A high-sensitivity polarimeter for the direct detection and characterization of extra-solar planets, *in* 'Polarimetry in Astronomy', Vol. 4843, International Society for Optics and Photonics, pp. 517–524.

Hough, J., Lucas, P., Bailey, J., Tamura, M., Hirst, E., Harrison, D. and Bartholomew-Biggs, M. (2006), 'Planetpol: a very high sensitivity polarimeter', *Publications of the Astronomical Society of the Pacific* **118**(847), 1302.

Hovenier, J. W., Van Der Mee, C. and Domke, H., eds (2004), *Transfer of polarized light in planetary atmospheres : basic concepts and practical methods*, Vol. 318 of *Astrophysics and Space Science Library*.

Hubanks, P., Platnick, S., King, M. and Ridgway, B. (2015), 'Modis atmosphere l3 gridded product algorithm theoretical basis document (atbd) & users guide', *ATBD reference number ATBD-MOD-30, NASA* **125**.

Jiang, J. H., Zhai, A. J., Herman, J., Zhai, C., Hu, R., Su, H., Natraj, V., Li, J., Xu, F. and Yung, Y. L. (2018), 'Using deep space climate observatory measurements to study the earth as an exoplanet', *arXiv preprint arXiv:1805.05834* .

Kaltenegger, L. and Traub, W. A. (2009), 'Transits of Earth-like Planets', *The Astrophysical Journal* **698**, 519–527.

Karalidi, T., Stam, D. and Hovenier, J. (2012), 'Looking for the rainbow on exoplanets covered by liquid and icy water clouds', *Astronomy & Astrophysics* **548**, A90.

Karalidi, T., Stam, D. M. and Hovenier, J. W. (2011), 'Flux and polarisation spectra of water clouds on exoplanets', *Astronomy & Astrophysics* **530**, A69.

Karalidi, T., Stam, D., Snik, F., Bagnulo, S., Sparks, W. and Keller, C. (2012), 'Observing the earth as an exoplanet with loupe, the lunar observatory for unresolved polarimetry of earth', *Planetary and Space Science* **74**(1), 202–207.

Kawahara, H. and Fujii, Y. (2010), 'Global Mapping of Earth-like Exoplanets From Scattered Light Curves', *The Astrophysical Journal* **720**, 1333–1350.

Kawahara, H. and Fujii, Y. (2011), 'Mapping Clouds and Terrain of Earth-like Planets from Photometric Variability: Demonstration with Planets in Face-on Orbits', *The Astrophysical Journal* **739**, L62.

Kemp, J. C., Henson, G. D., Steiner, C. T., Beardsley, I. S. and Powell, E. R. (1987), 'The optical polarization of the Sun, measured at a sensitivity of parts in ten million', *Nature* **328**, 92.

King, M. D., Radke, L. F. and Hobbs, P. V. (1990), 'Determination of the spectral absorption of solar radiation by marine stratocumulus clouds from airborne measurements within clouds', *Journal of the Atmospheric Sciences* **47**(7), 894–908.

Koepke, P. (1984), 'Effective reflectance of oceanic whitecps', *Applied Optics* **23**, 1816–1824.

Kokhanovsky, A. (2004), 'Optical properties of terrestrial clouds', *Earth-Science Reviews* **64**(3-4), 189–241.

Kotz, J. (1994), 'Balakrishan, continuous univariate distributions, volumes i and ii'.

Lacis, A. A. and Hansen, J. (1974), 'A Parameterization for the Absorption of Solar Radiation in the Earth's Atmosphere.', *Journal of Atmospheric Sciences* **31**, 118–133.

Lammer, H., Bredehöft, J. H., Coustenis, A., Khodachenko, M. L., Kaltenegger, L., Grasset, O., Prieur, D., Raulin, F., Ehrenfreund, P., Yamauchi, M., Wahlund, J. E., Grießmeier, J. M., Stangl, G., Cockell, C. S., Kulikov, Y. N., Grenfell, J. L. and Rauer, H. (2009), 'What makes a planet habitable?', *Astronomy and Astrophysics Review* **17**, 181–249.

Levy, R., Mattoo, S., Munchak, L., Remer, L., Sayer, A., Patadia, F. and Hsu, N. (2013), 'The collection 6 modis aerosol products over land and ocean', *Atmospheric Measurement Techniques* **6**(11), 2989.

Li, Z., Cribb, M., Chang, F. and Trishchenko, A. (2004), Validation of modis-retrieved cloud fractions using whole sky imager measurements at the three arm sites, *in* 'Proc. 14th ARM Science Team Meeting', pp. 1–6.

Liou, K. N. (1980), *An introduction to atmospheric radiation.*

Liou, K.-N. and Yang, P. (2016), *Light scattering by ice crystals: fundamentals and applications*, Cambridge University Press.

Litvinov, P., Hasekamp, O. and Cairns, B. (2011), 'Models for surface reflection of radiance and polarized radiance: Comparison with airborne multi-angle photopolarimetric measurements and implications for modeling top-of-atmosphere measurements', *Remote Sensing of Environment* **115**(2), 781–792.
URL: *http://dx.doi.org/10.1016/j.rse.2010.11.005*

Lucas, P., Hough, J. and Bailey, J. (2006), Planetpol: polarimetry of hot jupiters at the parts per million level, *in* 'Tenth Anniversary of 51 Peg-b: Status of and prospects for hot Jupiter studies', pp. 334–341.

Maignan, F., Bréon, F. M., Fédèle, E. and Bouvier, M. (2009), 'Polarized reflectances of natural surfaces: Spaceborne measurements and analytical modeling', *Remote Sensing of Environment* **113**(12), 2642–2650.
URL: *http://dx.doi.org/10.1016/j.rse.2009.07.022*

Marchand, R., Ackerman, T., Smyth, M. and Rossow, W. B. (2010), 'A review of cloud top height and optical depth histograms from misr, isccp, and modis', *Journal of Geophysical Research: Atmospheres* **115**(D16).

Mayor, M. and Queloz, D. (1995), 'A Jupiter-mass companion to a solar-type star', *Nature* **378**, 355–359.

Mishchenko, M. I. and Travis, L. D. (1997), 'Satellite retrieval of aerosol properties over the ocean using polarization as well as intensity of reflected sunlight', *Journal of Geophysical Research* **102**, No.D14:16989–17013.

MODIS Land Validation Web Site (n.d.), https://landval.gsfc.nasa.gov/ProductStatus.php?ProductID=MOD12&_ga=2.166296701.297105975.1532679131-655636017.1490210080. Accessed: 2018-03-26.

Monahan, E. C. and Ó Muircheartaigh, I. (1980), 'Optimal Power-Law Description of Oceanic Whitecap Coverage Dependence on Wind Speed', *Journal of Physical Oceanography* **10**, 2094–2099.

Montañés-Rodríguez, P., Pallé, E., Goode, P. R. and Martín-Torres, F. J. (2006), 'Vegetation Signature in the Observed Globally Integrated Spectrum of Earth Considering Simultaneous Cloud Data: Applications for Extrasolar Planets', *The Astrophysical Journal* **651**, 544–552.

Morel, A. (1974), 'Optical properties of pure water and pure sea water', *Optical Aspects of Oceanography* **1**, 1–24.

Morel, A. and Maritorena, S. (2001), 'Bio-optical properties of oceanic waters: A reappraisal', *Journal of Geophysics Research* **106**, 7163–7180.

Moreno, F., Vilaplana, R., Munoz, O., Molina, A. and Guirado, D. (2006), 'The scattering matrix for size distributions of irregular particles: An application to an olivine sample', *Journal of Quantitative Spectroscopy and Radiative Transfer* **100**(1-3), 277–287.

Muñoz, A. G. (2015), 'Towards a comprehensive model of earth's disk-integrated stokes vector', *International Journal of Astrobiology* **14**(3), 379–390.

Munoz, O., Moreno, R., Guirado, D., Dabowska, D. D., Volten, H. and Hovenier, J. W. (2012), The Amsterdam-Granada Light Scattering Database., *in* 'Digital Handbook. The 2012 European Aerosol Conference', p. 53.

Nadal, F. and Breon, F.-M. (1999), 'Parameterization of surface polarized reflectance derived from POLDER spaceborne measurements', *IEEE Transactions on Geoscience and Remote Sensing* **37**, 1709–1718.

Nakajima, T. (1983), 'Effect of wind-generated waves on the transfer of solar radiation in the atmosphere-ocean system', *Journal of Quantitative Spectroscopy and Radiative Transfer* **29**, 521–537.

Nakajima, T. and King, M. D. (1990), 'Determination of the optical thickness and effective particle radius of clouds from reflected solar radiation measurements. part i: Theory', *Journal of the atmospheric sciences* **47**(15), 1878–1893.

NASA Releases Kepler Survey Catalog with Hundreds of New Planet Candidates (n.d.), https://www.nasa.gov/press-release/nasa-releases-kepler-survey-catalog-with-hundreds-of-new-planet-candidates. Accessed: 2018-08-22.

Oakley, P. H. H. and Cash, W. (2009), 'Construction of an Earth Model: Analysis of Exoplanet Light Curves and Mapping the Next Earth with the New Worlds Observer', *The Astrophysical Journal* **700**, 1428–1439.

Oreopoulos, L., Cahalan, R. F. and Platnick, S. (2007), 'The plane-parallel albedo bias of liquid clouds from modis observations', *Journal of Climate* **20**(20), 5114–5125.

Pallé, E., Ford, E. B., Seager, S., Montañés-Rodríguez, P. and Vazquez, M. (2008), 'Identifying the Rotation Rate and the Presence of Dynamic Weather on Extrasolar Earth-like Planets from Photometric Observations', *The Astrophysical Journal* **676**, 1319–1329.

Peck, E. R. and Reeder, K. (1972), 'Dispersion of Air', *Journal of the Optical Society of America (1917-1983)* **62**, 958.

Peltoniemi, J., Hakala, T., Suomalainen, J. and Puttonen, E. (2009), 'Polarised bidirectional reflectance factor measurements from soil, stones, and snow', *Journal of Quantitative Spectroscopy and Radiative Transfer* **110**, 1940–1953.

Piirola, V., Berdyugin, A. and Berdyugina, S. (2014), Dipol-2: A double image high precision polarimeter, *in* 'Ground-based and Airborne Instrumentation for Astronomy V', Vol. 9147, International Society for Optics and Photonics, p. 91478I.

Platnick, S., King, M. D., Meyer, K. G., Wind, G., Amarasinghe, N., Marchant, B., Arnold, G. T., Zhang, Z., Hubanks, P. A., Ridgway, B. et al. (2015), 'Modis cloud optical properties: User guide for the collection 6 level-2 mod06/myd06 product and associated level-3 datasets', *Version* **1**, 145.

Pollack, J. B., Hollenbach, D., Beckwith, S., Simonelli, D. P., Roush, T. and Fong, W. (1994), 'Composition and radiative properties of grains in molecular clouds and accretion disks', *The Astrophysical Journal* **421**, 615–639.

Pope, R. M. and Fry, E. S. (1997), 'Absorption spectrum (380 -700 nm) of pure water. II. Integrating cavity measurements', *Applied Optics* **36**, 8710–8723.

Qiu, J., Goode, P., Pallé, E., Yurchyshyn, V., Hickey, J., Montañés Rodriguez, P., Chu, M.-C., Kolbe, E., Brown, C. and Koonin, S. (2003), 'Earthshine and the earth's albedo: 1. earthshine observations and measurements of the lunar phase function for accurate measurements of the earth's bond albedo', *Journal of Geophysical Research: Atmospheres* **108**(D22).

Rayleigh, L. (1871), 'Scientific papers 8 and 9', *Philos Mag* **41**(107), 274.

Rossi, L., Berzosa-Molina, J. and Stam, D. M. (2018), 'PYMIEDAP: a Python-Fortran tool for computing fluxes and polarization signals of (exo)planets', *Astronomy & Astrophysics* **616**, A147.

Rossi, L. and Stam, D. M. (2017), 'Using polarimetry to retrieve the cloud coverage of Earth-like exoplanets', *Astronomy & Astrophysics* **607**, A57.

Rossi, L. and Stam, D. M. (2018), 'Circular polarization signals of cloudy (exo)planets', *Astronomy & Astrophysics* **616**, A117.

Rossow, W. B. and Schiffer, R. A. (1999), 'Advances in Understanding Clouds from ISCCP.', *Bulletin of the American Meteorological Society* **80**, 2261–2288.

Roujean, J.-L., Leroy, M. and Deschamps, P.-Y. (1992), 'A Bidirectional Reflectance Model of the Earth's Surface for the Correction of Remote Sensing Data', *Journal of Geophysical Research* **97**(92).

Saar, S. and Seager, S. (2003), 'Uses of linear polarization as a probe of extrasolar planet atmospheres', *arXiv preprint astro-ph/0305436* .

Sancer, M. (1969), 'Shadow-corrected electromagnetic scattering from a randomly rough surface', *IEEE Transactions on Antennas and Propagation* **17**, 577–585.

Schaepman-Strub, G., Schaepman, M. E., Painter, T. H., Dangel, S. and Martonchik, J. V. (2006), 'Reflectance quantities in optical remote sensing-definitions and case studies', *Remote Sensing of Environment* **103**(1), 27–42.

Schmid, H. M., Beuzit, J.-L., Feldt, M., Gisler, D., Gratton, R., Henning, T., Joos, F., Kasper, M., Lenzen, R., Mouillet, D., Moutou, C., Quirrenbach, A., Stam, D. M., Thalmann, C., Tinbergen, J., Verinaud, C., Waters, R. and Wolstencroft, R. (2006), Search and investigation of extra-solar planets with polarimetry, *in* C. Aime and F. Vakili, eds, 'IAU Colloq. 200: Direct Imaging of Exoplanets: Science & Techniques', pp. 165–170.

Seager, S. and Sasselov, D. D. (2000), 'Theoretical transmission spectra during extrasolar giant planet transits', *The Astrophysical Journal* **537**(2), 916.

Slingo, A. (1989), 'A gcm parameterization for the shortwave radiative properties of water clouds', *Journal of the Atmospheric Sciences* **46**(10), 1419–1427.

Smith, B. (1967), 'Geometrical shadowing of a random rough surface', *IEEE Transactions on Antennas and Propagation* **15**, 668–671.

Smith, R. C. and Baker, K. S. (1981), 'Optical properties of the clearest natural waters (200-800 nm)', *Applied Optics* **20**, 177–184.

Snik, F., Karalidi, T. and Keller, C. U. (2009), 'Spectral modulation for full linear polarimetry', *Applied Optics* **48**, 1337–1346.

Snyder, D. L., Helstrom, C. W., Lanterman, A. D., Faisal, M. and White, R. L. (1995), 'Compensation for readout noise in ccd images', *JOSA A* **12**(2), 272–283.

Snyder, J. P. (1987), *Map projections–A working manual*, Vol. 1395, US Government Printing Office.

Sogandares, F. M. and Fry, E. S. (1997), 'Absorption spectrum (340 -640 nm) of pure water. I. Photothermal measurements', *Applied Optics* **36**, 8699–8709.

Sparks, W., Hough, J., Kolokolova, L., Germer, T., Chen, F., DasSarma, S., DasSarma, P., Robb, F., Manset, N., Reid, I. et al. (2009), 'Circular polarization in scattered light as a possible biomarker', *Journal of Quantitative Spectroscopy and Radiative Transfer* **110**(14-16), 1771–1779.

Stam, D. and Hovenier, J. (2005), 'Errors in calculated planetary phase functions and albedos due to neglecting polarization', *Astronomy & Astrophysics* **444**(1), 275–286.

Stam, D., Hovenier, J. and Waters, L. (2004*a*), 'Using polarimetry to detect and characterize jupiter-like extrasolar planets', *Astronomy & Astrophysics* **428**(2), 663–672.

Stam, D. M. (2008*a*), 'Spectropolarimetric signatures of Earth-like extrasolar planets', *Astronomy & Astrophysics* **482**, 989–1007.

Stam, D. M. (2008*b*), 'Spectropolarimetric signatures of Earth-like extrasolar planets', *Astronomy & Astrophysics* **482**, 989–1007.

Stam, D. M., Aben, I. and Helderman, F. (2002), 'Skylight polarization spectra: Numerical simulation of the Ring effect', *Journal of Geophysical Research (Atmospheres)* **107**, 4419.

Stam, D. M., de Rooij, W. A., Cornet, G. and Hovenier, J. W. (2006), 'Integrating polarized light over a planetary disk applied to starlight reflected by extrasolar planets', *Astronomy & Astrophysics* **452**, 669–683.

Stam, D. M., Hovenier, J. W. and Waters, L. B. F. M. (2004*b*), 'Using polarimetry to detect and characterize Jupiter-like extrasolar planets', *Astronomy & Astrophysics* **428**, 663–672.

Stephens, G. L. and Tsay, S.-C. (1990), 'On the cloud absorption anomaly', *Quarterly Journal of the Royal Meteorological Society* **116**, 671–704.

Sterzik, M. F., Bagnulo, S. and Emde, C. (2015), Bio-signatures of Planet Earth from Spectropolarimetry, *in* K. N. Nagendra, S. Bagnulo, R. Centeno and M. Jesús Martínez González, eds, 'Polarimetry', Vol. 305, pp. 305–312.

Sterzik, M. F., Bagnulo, S. and Palle, E. (2012), 'Biosignatures as revealed by spectropolarimetry of Earthshine', *Nature* **483**, 64–66.

Swain, M. R., Tinetti, G., Vasisht, G., Deroo, P., Griffith, C., Bouwman, J., Chen, P., Yung, Y., Burrows, A., Brown, L. R., Matthews, J., Rowe, J. F., Kuschnig, R. and Angerhausen, D. (2009), 'Water, Methane, and Carbon Dioxide Present in the Dayside Spectrum of the Exoplanet HD 209458b', *The Astrophysical Journal* **704**, 1616–1621.

Swain, M. R., Vasisht, G. and Tinetti, G. (2008), 'The presence of methane in the atmosphere of an extrasolar planet', *Nature* **452**, 329–331.

Tinetti, G., Meadows, V. S., Crisp, D., Kiang, N. Y., Kahn, B. H., Fishbein, E., Velusamy, T. and Turnbull, M. (2006), 'Detectability of Planetary Characteristics in Disk-Averaged Spectra II: Synthetic Spectra and Light-Curves of Earth', *Astrobiology* **6**, 881–900.

Tinetti, G., Vidal-Madjar, A., Liang, M.-C., Beaulieu, J.-P., Yung, Y., Carey, S., Barber, R. J., Tennyson, J., Ribas, I., Allard, N., Ballester, G. E., Sing, D. K. and Selsis, F. (2007), 'Water vapour in the atmosphere of a transiting extrasolar planet', *Nature* **448**, 169–171.

Todorov, K. (2008), 'Determining the temperature of exoplanet hat-p-1b', *Physics, Astronomy and Geophysics Honors Papers. Paper* **1**.

Traub, W. A., Levine, M., Shaklan, S., Kasting, J., Angel, J. R., Brown, M. E., Brown, R. A., Burrows, C., Clampin, M., Dressler, A., Ferguson, H. C., Hammel, H. B., Heap, S. R., Horner, S. D., Illingworth, G. D., Kasdin, N. J., Kuchner, M. J., Lin, D., Marley, M. S., Meadows, V., Noecker, C., Oppenheimer, B. R., Seager, S., Shao, M., Stapelfeldt, K. R. and Trauger, J. T. (2006), TPF-C: status and recent progress, *in* 'Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series', Vol. 6268, p. 62680T.

Trees, V. (2018), Transfer of Polarized Light in an (Exo)planetary Atmosphere-Ocean System, Master's thesis, Delft University of Technology.

Tsang, L., Kong, J. A. and Shin, R. T. (1985), *Theory of microwave remote sensing*, Wiley Interscience, New York.

User Guide to Collection 6 MODIS Land Cover (MCD12Q1 and MCD12C1) Product (n.d.*a*), https://lpdaac.usgs.gov/sites/default/files/public/product_documentation/mcd12_user_guide_v6.pdf. Accessed: 2018-05-30.

User Guide to Collection 6 MODIS Land Cover (MCD12Q1 and MCD12C1) Product (n.d.*b*), https://speclib.jpl.nasa.gov. Accessed: 2018-04-02.

van de Hulst, H. C. (1957), *Light Scattering by Small Particles*.

Vlachos, M., Yu, P. and Castelli, V. (2005), On periodicity detection and structural periodic similarity, *in* 'Proceedings of the 2005 SIAM international conference on data mining', SIAM, pp. 449–460.

Waterman, P. C. (1971), 'Symmetry, Unitarity, and Geometry in Electromagnetic Scattering', *Physical Review D: Particles, Fields, Gravitation & Cosmology* **3**, 825–839.

Williams, D. M. and Gaidos, E. (2008), 'Detecting the glint of starlight on the oceans of distant planets', *Icarus* **195**, 927–937.

Williams, D. M. and Pollard, D. (2003), 'Extraordinary climates of Earth-like planets: three-dimensional climate simulations at extreme obliquity', *International Journal of Astrobiology* **2**, 1–19.

Zhai, P.-W., Hu, Y., Chowdhary, J., Trepte, C. R., Lucker, P. L. and Josset, D. B. (2010), 'A vector radiative transfer model for coupled atmosphere and ocean systems with a rough interface', *Journal of Quantitiative Spectroscopy and Radiative Transfer* **111**, 1025–1040.

Zugger, M. E., Kasting, J. F., Williams, D. M., Kane, T. J. and Philbrick, C. R. (2010*a*), 'Light Scattering from Exoplanet Oceans and Atmospheres', *The Astrophysical Journal* **723**, 1168–1179.

Zugger, M. E., Kasting, J. F., Williams, D. M., Kane, T. J. and Philbrick, C. R. (2010*b*), 'Light Scattering from Exoplanet Oceans and Atmospheres', *The Astrophysical Journal* **723**, 1168–1179.

Zugger, M. E., Kasting, J. F., Williams, D. M., Kane, T. J. and Philbrick, C. R. (2011*a*), 'Erratum: 'Light Scattering from Exoplanet Oceans and Atmospheres'', *The Astrophysical Journal* **739**, 55(5pp).

Zugger, M. E., Kasting, J. F., Williams, D. M., Kane, T. J. and Philbrick, C. R. (2011*b*), 'Searching for Water Earths in the Near-infrared', *The Astrophysical Journal* **739**, 12.

A

Modified PyMieDAP routine for modeling Earth-like inhomogeneous planetary disks

Figure A.1 shows a flowchart of the developed PyMieDAP Planet_pixels function. The principal input objects are the Earth Observations, Input parameters and a Model Object. The Earth Observations comprise of any observations that the user favors in the format that will subsequently be described. The input parameters consist of those that define the geometry and type of exoplanet one wants to model, but also what temporal resolution and interval one wants to observe or simply a single date at a modified geometry. The core input parameters for the doubling-adding routine and Fourier file construction can also be applied in this first instance. Before one can start computing disk resolved or integrated Stokes elements a *Model Object* needs to be defined. In this *Model Object* one needs to specify all parameters that will be constant in the Earth-like model, such as for example the mean molecular mass or in case of a locally homogeneous surface, the surface reflection (a more detailed description is provided in *Section* 2.2.1. This *Model Object* will at the same time be used to store all data throughout the computation, including for example the Stokes elements, phase angles used, disk geometrical properties, etc.

As part of the *Planet_pixels* function the following operations will be executed. Using the specified planetary geometry the unmodified *getgeos* function is called to calculate all the relevant disk properties such as the number of visible and illuminated pixels, the pixel areas, the solar and emission zenith angles, the azimuthal difference angles, the rotation β for each pixel, and the coordinates for each pixel. Using these parameters the extended *Mask_Planet* function can now calculate a specific mask for each pixel that is based on Earth observations. In general the pixel models are distinct by surface type and by cloud type. It thus depends on the users preference how many possible pixel models there can be called by the *Mask_planet* function. The next step is to calculate the radiative transfer model for each unique combination of observations. This can be computed beforehand or during a run, because the Fourier files are stored and assigned an unique label. If a pixel model has not been calculated yet, this needs to be handled by a new function in PyMieDAP. The *Model_generator* function does exactly this by using the *mie_code* and *compute_model* functions that were already developed in PyMieDAP, after which the generated *Fourier file* is read using the *read_dap* function. In case the *Fourier file* exists the Fourier coefficients are immediately read by the *read_dap* function.

With the extended *Planet_pixel* function we can now obtain the resolved Stokes vector for an arbitrarily amount of data points. Because the user also may want to analyze the disk integrated spectropolarimetric signal, the *plot_pixel* PyMieDAP function is extended to do just that.

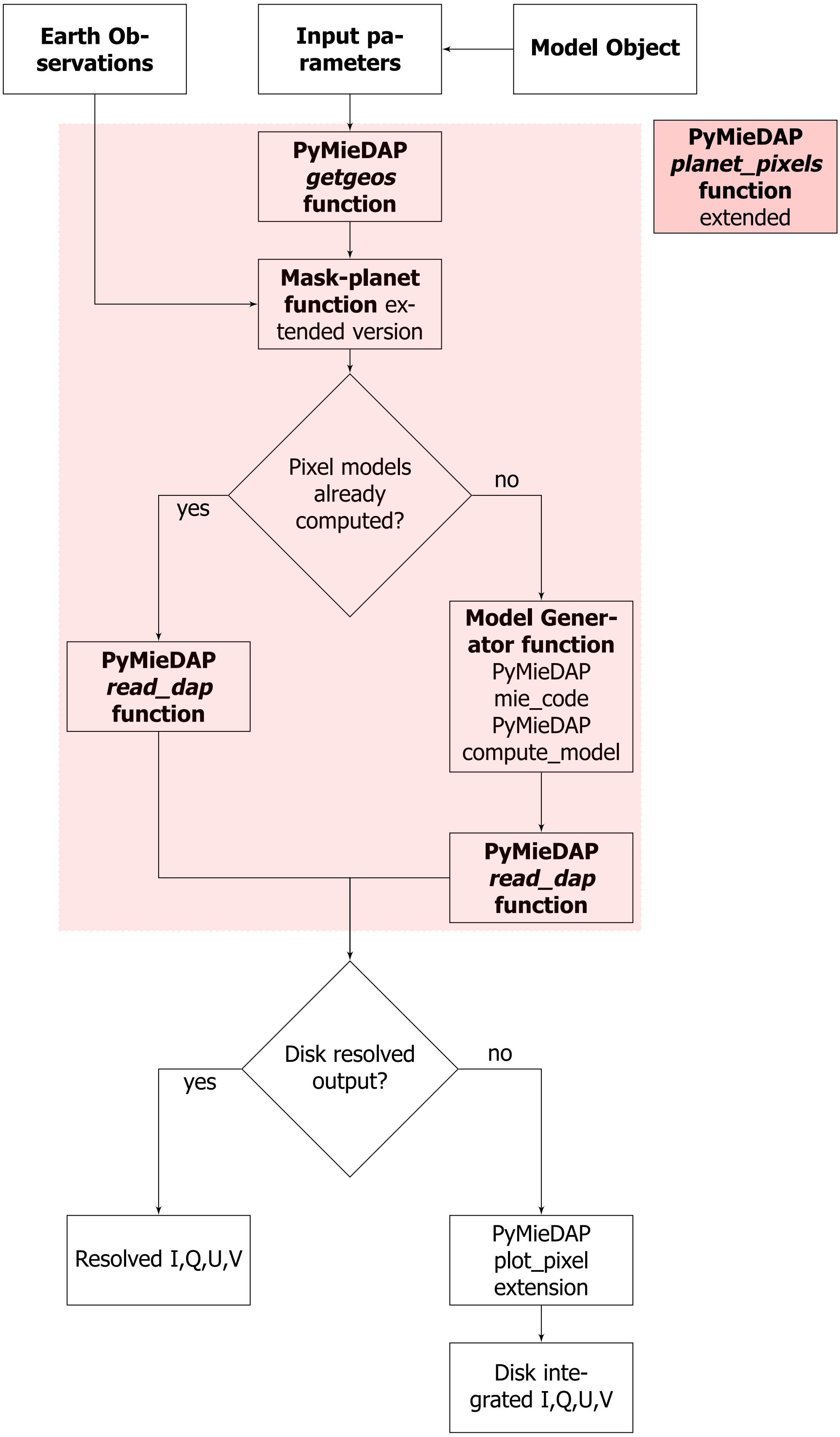


Figure A.1: Flowchart of the general process with which an user of PyMieDAP can model an Earth-like exoplanet, based on Earth Observations. Red boxes denote newly developed or extended functions. Blue boxes denote existing functions. The remaining boxes denote decisions or input parameters.

B

Verification of coordinate conversion

The coordinate transformation used in the *mask_planet* function is verified according to two verification cases provided by Snyder (1987). In *Table* B.1 an initial obliquity and longitudinal position of $(\Phi_0, \lambda_0) = 0°$ is used. In *Tables* B.2 and B.3 an initial obliquity and longitudinal position of $(\Phi_0, \lambda_0) = (40°, 0°)$ is used. For both cases: Radius of sphere = 1.0 .

Table B.1: Origin: (x,y)=0 at $(\Phi_0, \lambda_0) = 0°$.

	Theoretical values from Snyder (1987)										
Long.		$0°$	$10°$	$20°$	$30°$	$40°$	$50°$	$60°$	$70°$	$80°$	$90°$
Lat.	y	*x*									
$90°$	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
$80°$	0.9848	0.0000	0.0302	0.0594	0.0868	0.1116	0.1330	0.1504	0.1632	0.1710	0.1736
$70°$	0.9397	0.0000	0.0594	0.1170	0.1710	0.2198	0.2620	0.2962	0.3214	0.3368	0.3420
$60°$	0.8660	0.0000	0.0868	0.1710	0.2500	0.3214	0.3830	0.4330	0.4698	0.4924	0.5000
$50°$	0.7660	0.0000	0.1116	0.2198	0.3214	0.4132	0.4924	0.5567	0.6040	0.6330	0.6428
$40°$	0.6248	0.0000	0.1330	0.2620	0.3830	0.4924	0.5868	0.6634	0.7198	0.7544	0.7660
$30°$	0.5000	0.0000	0.1504	0.2962	0.4330	0.5567	0.6634	0.7500	0.8138	0.8529	0.8660
$20°$	0.3420	0.0000	0.1632	0.3214	0.4698	0.6040	0.7198	0.8138	0.8830	0.9254	0.9397
$10°$	0.1736	0.0000	0.1710	0.3368	0.4924	0.6330	0.7544	0.8529	0.9254	0.9698	0.9848
$0°$	0.0000	0.0000	0.1736	0.3420	0.5000	0.6428	0.7660	0.8660	0.9397	0.9848	1.0000
	Results from mask_planet function										
$90°$	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
$80°$	0.9848	0.0000	0.0302	0.0594	0.0868	0.1116	0.1330	0.1504	0.1632	0.1710	0.1736
$70°$	0.9397	0.0000	0.0594	0.1170	0.1710	0.2198	0.2620	0.2962	0.3214	0.3368	0.3420
$60°$	0.8660	0.0000	0.0868	0.1710	0.2500	0.3214	0.3830	0.4330	0.4698	0.4924	0.5000
$50°$	0.7660	0.0000	0.1116	0.2198	0.3214	0.4132	0.4924	0.5567	0.6040	0.6330	0.6428
$40°$	0.6248	0.0000	0.1330	0.2620	0.3830	0.4924	0.5868	0.6634	0.7198	0.7544	0.7660
$30°$	0.5000	0.0000	0.1504	0.2962	0.4330	0.5567	0.6634	0.7500	0.8138	0.8529	0.8660
$20°$	0.3420	0.0000	0.1632	0.3214	0.4698	0.6040	0.7198	0.8138	0.8830	0.9254	0.9397
$10°$	0.1736	0.0000	0.1710	0.3368	0.4924	0.6330	0.7544	0.8529	0.9254	0.9698	0.9848
$0°$	0.0000	0.0000	0.1736	0.3420	0.5000	0.6428	0.7660	0.8660	0.9397	0.9848	1.0000

Table B.2: Origin: (x,y)=0 at $(\Phi_0, \lambda_0) = (40°, 0°)$.

Theoretical values from Snyder (1987)										
Long. Lat.	0°	10°	20°	30°	40°	50°	60°	70°	80°	90°
90°	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	(0.7660)	(0.7660)	(0.7660)	(0.7660)	(0.7660)	(0.7660)	(0.7660)	(0.7660)	(0.7660)	(0.7660)
80°	0.0000	0.0302	0.0594	0.0868	0.1116	0.1330	0.1504	0.1632	0.1710	0.1736
	(0.6428)	(0.6445)	(0.6495)	(0.6577)	(0.6689)	(0.6827)	(0.6986)	(0.7162)	(0.7350)	(0.7544)
70°	0.0000	0.0594	0.1170	0.1710	0.2198	0.2620	0.2962	0.3214	0.3368	0.3420
	(0.5000)	(0.5033)	(0.5133)	(0.5295)	(0.5514)	(0.5785)	(0.6099)	(0.6447)	(0.6817)	(0.7198)
60°	0.0000	0.0868	0.1710	0.2500	0.3214	0.3830	0.4330	0.4698	0.4924	0.5000
	(0.3420)	(0.3469)	(0.3614)	(0.3851)	(0.4172)	(0.4568)	(0.5027)	(0.5535)	(0.6076)	(0.6634)
50°	0.0000	0.1116	0.2198	0.3214	0.4132	0.4924	0.5567	0.6040	0.6330	0.6428
	(0.1736)	(0.1799)	(0.1986)	(0.2290)	(0.2703)	(0.3212)	(0.3802)	(0.4455)	(0.5151)	(0.5868)
40°	0.0000	0.1330	0.2620	0.3830	0.4924	0.5868	0.6634	0.7198	0.7544	0.7660
	(0.0000)	(0.0075)	(0.0297)	(0.0660)	(0.1152)	(0.0.1759)	(0.2462)	(0.3240)	(0.4069)	(0.4924)
30°	0.0000	0.1504	0.2962	0.4330	0.5567	0.6634	0.7500	0.8138	0.8529	0.8660
	(-0.1736)	(-0.1652)	(-0.1401)	(-0.0991)	(-0.0434)	(0.0252)	(0.1047)	(0.1926)	(0.2864)	(0.3830)
20°	0.0000	0.1632	0.3214	0.4698	0.6040	0.7198	0.8138	0.8830	0.9254	0.9397
	(-0.3420)	(-0.3328)	(-0.3056)	(-0.2611)	(-0.2007)	(-0.1263)	(-0.0400)	(0.0554)	(0.1571)	(0.2620)
10°	0.0000	0.1710	0.3368	0.4924	0.6330	0.7544	0.8529	0.9254	0.9698	0.9848
	(-0.5000)	(-0.4904)	(-0.4618)	(-0.4152)	(-0.3519)	(-0.2739)	(-0.1835)	(-0.0835)	(0.0231)	(0.1330)
0°	0.0000	0.1736	0.3420	0.5000	0.6428	0.7660	0.8660	0.9397	0.9848	1.0000
	(-0.6428)	(-0.6330)	(-0.6040)	(-0.5567)	(-0.4924)	(-0.4132)	(-0.3214)	(-0.2198)	(-0.1116)	(0.0000)
−10°	0.0000	0.1710	0.3368	0.4924	0.6330	0.7544	0.8529	0.9254	0.9698	–
	(-0.7660)	(-0.7564)	(-0.7279)	(-0.6812)	(-0.6179)	(-0.5399)	(-0.4495)	(-0.3495)	(-0.2429)	–
−20°	0.0000	0.1632	0.3214	0.4698	0.6040	0.7198	0.8138	0.8830	–	–
	(-0.8660)	(-0.8568)	(-0.8296)	(-0.7851)	(-0.7247)	(-0.6503)	(-0.5640)	(-0.4686)	–	–
−30°	0.0000	0.1504	0.2962	0.4330	0.5567	0.6634	0.7500	–	–	–
	(-0.9397)	(-0.9312)	(-0.9061)	(-0.8651)	(-0.8095)	(-0.7408)	(-0.6614)	–	–	–
−40°	0.0000	0.1330	0.2620	0.3830	0.4924	–	–	–	–	–
	(-0.9848)	(-0.9773)	(-0.9551)	(-0.9188)	(-0.8696)	–	–	–	–	–
−50°	0.0000	–	–	–	–	–	–	–	–	–
	(-1.0000)	–	–	–	–	–	–	–	–	–
Results from mask_planet function										
90°	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	(0.7660)	(0.7660)	(0.7660)	(0.7660)	(0.7660)	(0.7660)	(0.7660)	(0.7660)	(0.7660)	(0.7660)
80°	0.0000	0.0302	0.0594	0.0868	0.1116	0.1330	0.1504	0.1632	0.1710	0.1736
	(0.6428)	(0.6445)	(0.6495)	(0.6577)	(0.6689)	(0.6827)	(0.6986)	(0.7162)	(0.7350)	(0.7544)
70°	0.0000	0.0594	0.1170	0.1710	0.2198	0.2620	0.2962	0.3214	0.3368	0.3420
	(0.5000)	(0.5033)	(0.5133)	(0.5295)	(0.5514)	(0.5785)	(0.6099)	(0.6447)	(0.6817)	(0.7198)
60°	0.0000	0.0868	0.1710	0.2500	0.3214	0.3830	0.4330	0.4698	0.4924	0.5000
	(0.3420)	(0.3469)	(0.3614)	(0.3851)	(0.4172)	(0.4568)	(0.5027)	(0.5535)	(0.6076)	(0.6634)
50°	0.0000	0.1116	0.2198	0.3214	0.4132	0.4924	0.5567	0.6040	0.6330	0.6428
	(0.1736)	(0.1799)	(0.1986)	(0.2290)	(0.2703)	(0.3212)	(0.3802)	(0.4455)	(0.5151)	(0.5868)
40°	0.0000	0.1330	0.2620	0.3830	0.4924	0.5868	0.6634	0.7198	0.7544	0.7660
	(0.0000)	(0.0075)	(0.0297)	(0.0660)	(0.1152)	(0.0.1759)	(0.2462)	(0.3240)	(0.4069)	(0.4924)
30°	0.0000	0.1504	0.2962	0.4330	0.5567	0.6634	0.7500	0.8138	0.8529	0.8660
	(-0.1736)	(-0.1652)	(-0.1401)	(-0.0991)	(-0.0434)	(0.0252)	(0.1047)	(0.1926)	(0.2864)	(0.3830)
20°	0.0000	0.1632	0.3214	0.4698	0.6040	0.7198	0.8138	0.8830	0.9254	0.9397
	(-0.3420)	(-0.3328)	(-0.3056)	(-0.2611)	(-0.2007)	(-0.1263)	(-0.0400)	(0.0554)	(0.1571)	(0.2620)
10°	0.0000	0.1710	0.3368	0.4924	0.6330	0.7544	0.8529	0.9254	0.9698	0.9848
	(-0.5000)	(-0.4904)	(-0.4618)	(-0.4152)	(-0.3519)	(-0.2739)	(-0.1835)	(-0.0835)	(0.0231)	(0.1330)
0°	0.0000	0.1736	0.3420	0.5000	0.6428	0.7660	0.8660	0.9397	0.9848	1.0000
	(-0.6428)	(-0.6330)	(-0.6040)	(-0.5567)	(-0.4924)	(-0.4132)	(-0.3214)	(-0.2198)	(-0.1116)	(0.0000)
−10°	0.0000	0.1710	0.3368	0.4924	0.6330	0.7544	0.8529	0.9254	0.9698	–
	(-0.7660)	(-0.7564)	(-0.7279)	(-0.6812)	(-0.6179)	(-0.5399)	(-0.4495)	(-0.3495)	(-0.2429)	–
−20°	0.0000	0.1632	0.3214	0.4698	0.6040	0.7198	0.8138	0.8830	–	–
	(-0.8660)	(-0.8568)	(-0.8296)	(-0.7851)	(-0.7247)	(-0.6503)	(-0.5640)	(-0.4686)	–	–
−30°	0.0000	0.1504	0.2962	0.4330	0.5567	0.6634	0.7500	–	–	–
	(-0.9397)	(-0.9312)	(-0.9061)	(-0.8651)	(-0.8095)	(-0.7408)	(-0.6614)	–	–	–
−40°	0.0000	0.1330	0.2620	0.3830	0.4924	–	–	–	–	–
	(-0.9848)	(-0.9773)	(-0.9551)	(-0.9188)	(-0.8696)	–	–	–	–	–
−50°	0.0000	–	–	–	–	–	–	–	–	–
	(-1.0000)	–	–	–	–	–	–	–	–	–

Table B.3: Continued: Origin: (x,y)=0 at $(\Phi_0, \lambda_0) = (40^{\circ}, 0^{\circ})$.

Long. / Lat.	100°	110°	120°	130°	140°	150°	160°	170°	180°	
	Theoretical values from Snyder (1987)									
90°	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	–
	(0.7660)	(0.7660)	(0.7660)	(0.7660)	(0.7660)	(0.7660)	(0.7660)	(0.7660)	(0.7660)	–
80°	0.1710	0.1632	0.1504	0.1330	0.1116	0.0868	0.0594	0.0302	0.0000	–
	(0.7738)	(0.7926)	(0.8102)	(0.8262)	(0.8399)	(0.8511)	(0.8593)	(0.8643)	(0.8660)	–
70°	0.3368	0.3214	0.2962	0.2620	0.2198	0.1710	0.1170	0.0594	0.0000	–
	(0.7580)	(0.7950)	(0.8298)	(0.8612)	(0.8883)	(0.9102)	(0.9264)	(0.9364)	(0.9397)	–
60°	0.4924	0.4698	0.4330	0.3830	0.3214	0.2500	0.1710	0.0868	0.0000	–
	(0.7192)	(0.7733)	(0.8241)	(0.8700)	(0.9096)	(0.9417)	(0.9654)	(0.9799)	(0.9848)	–
50°	0.6330	0.6040	0.5567	0.4924	0.4132	0.3214	0.2198	0.1116	0.0000	–
	(0.6586)	(0.7281)	(0.7934)	(0.8524)	(0.9033)	(0.9446)	(0.9751)	(0.9937)	(1.0000)	–
40°	0.7544	0.7198	0.6634	0.5868	–	–	–	–	–	–
	(0.5779)	(0.6608)	(0.7386)	(0.8089)	–	–	–	–	–	–
30°	0.8529	0.8138	–	–	–	–	–	–	–	–
	(0.4797)	(0.5734)	–	–	–	–	–	–	–	–
20°	0.9254	–	–	–	–	–	–	–	–	–
	(0.3669)	–	–	–	–	–	–	–	–	–
	Results from mask_planet function									
90°	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	–
	(0.7660)	(0.7660)	(0.7660)	(0.7660)	(0.7660)	(0.7660)	(0.7660)	(0.7660)	(0.7660)	–
80°	0.1710	0.1632	0.1504	0.1330	0.1116	0.0868	0.0594	0.0302	0.0000	–
	(0.7738)	(0.7926)	(0.8102)	(0.8262)	(0.8399)	(0.8511)	(0.8593)	(0.8643)	(0.8660)	–
70°	0.3368	0.3214	0.2962	0.2620	0.2198	0.1710	0.1170	0.0594	0.0000	–
	(0.7580)	(0.7950)	(0.8298)	(0.8612)	(0.8883)	(0.9102)	(0.9264)	(0.9364)	(0.9397)	–
60°	0.4924	0.4698	0.4330	0.3830	0.3214	0.2500	0.1710	0.0868	0.0000	–
	(0.7192)	(0.7733)	(0.8241)	(0.8700)	(0.9096)	(0.9417)	(0.9654)	(0.9799)	(0.9848)	–
50°	0.6330	0.6040	0.5567	0.4924	0.4132	0.3214	0.2198	0.1116	0.0000	–
	(0.6586)	(0.7281)	(0.7934)	(0.8524)	(0.9033)	(0.9446)	(0.9751)	(0.9937)	(1.0000)	–
40°	0.7544	0.7198	0.6634	0.5868	–	–	–	–	–	–
	(0.5779)	(0.6608)	(0.7386)	(0.8089)	–	–	–	–	–	–
30°	0.8529	0.8138	–	–	–	–	–	–	–	–
	(0.4797)	(0.5734)	–	–	–	–	–	–	–	–
20°	0.9254	–	–	–	–	–	–	–	–	–
	(0.3669)	–	–	–	–	–	–	–	–	–

C

Reflected stokes parameters as function of opacity and particle effective radius

In *Figure* C.1 we have provided an illustration of the geometries that are used to calculate the top of atmosphere (TOA) stokes elements. In *Figure* C.2-C.13 the TOA stokes elements are provided. In the left panels these elements are plotted as function of opacity for four different particle effective radii. In the right panel we provide the relations as function of the logarithmic of the opacity.

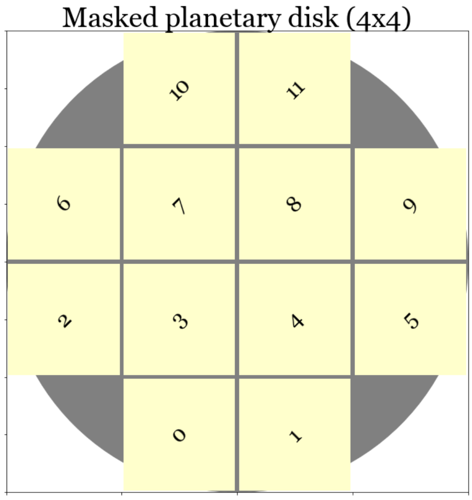


Figure C.1: Position of the pixels that are used to computed the top of atmosphere reflection for the stokes parameters F, Q and U as function of opacity and particle effective radius

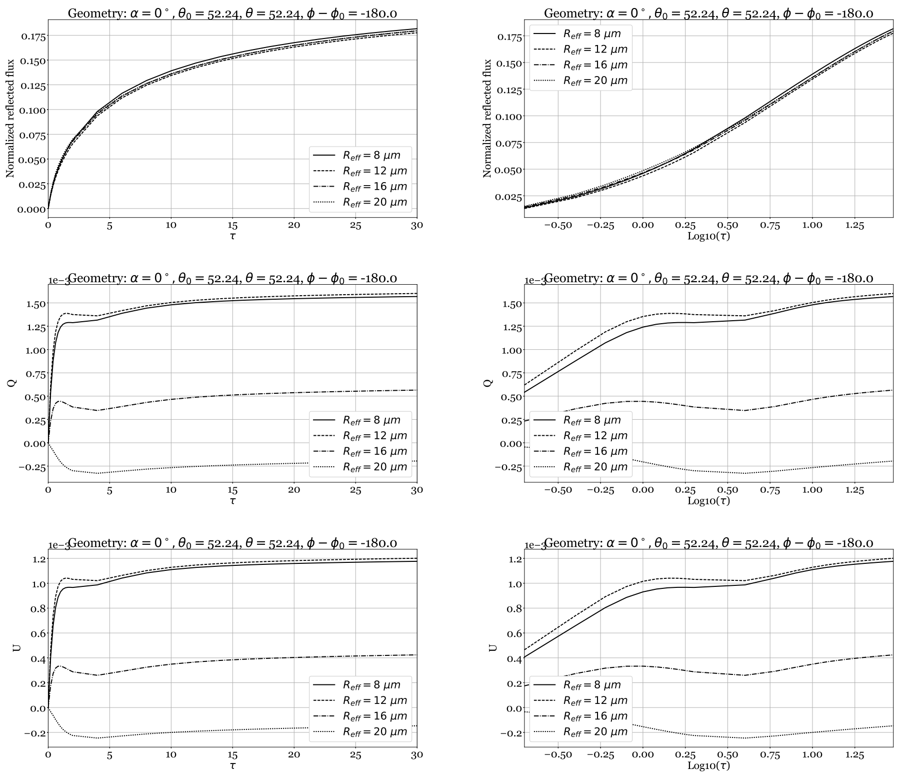


Figure C.2: Normalized reflected flux (upper panel) and polarized fluxes Q (middle panel) and U (lower panel) as function of cloud optical thickness for four values of cloud particle effective radius. The computations are made for a pixel 0 as referred to *Figure* C.1. The left panels show the cloud optical thickness on the x-axis and the right panels the logarithmic value of cloud optical thickness.

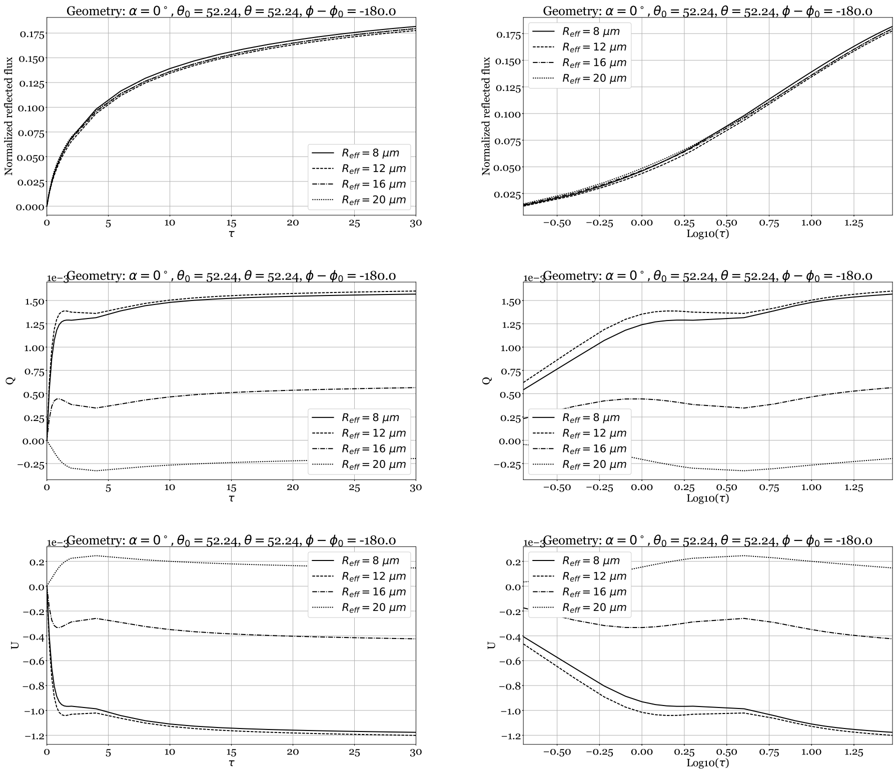


Figure C.3: Similar to *Figure* C.2, but for pixel 1.

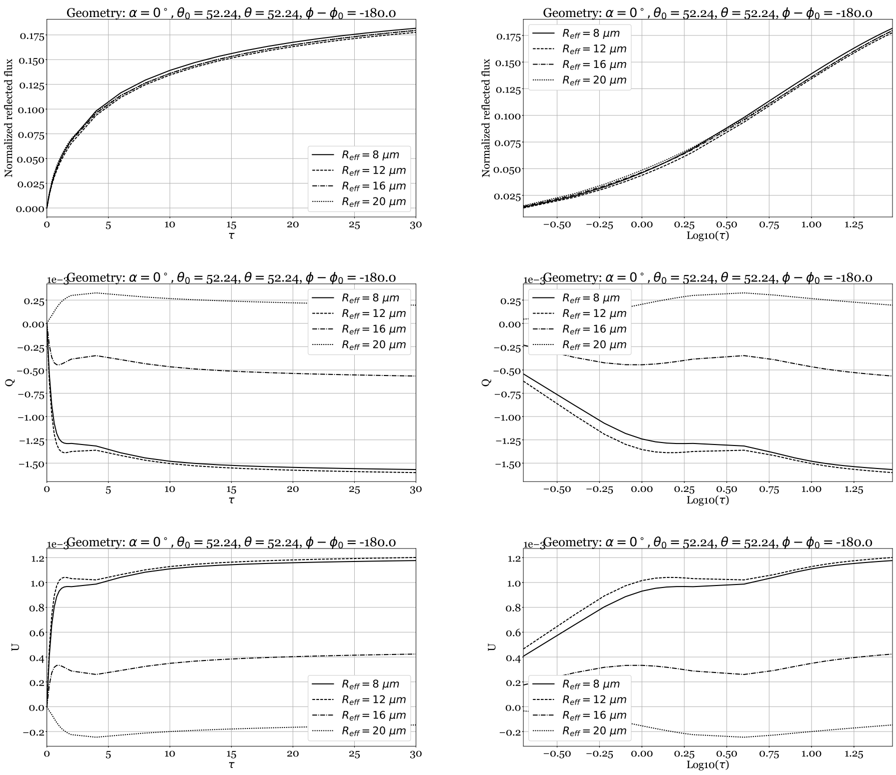


Figure C.4: Similar to *Figure* C.2, but for pixel 2.

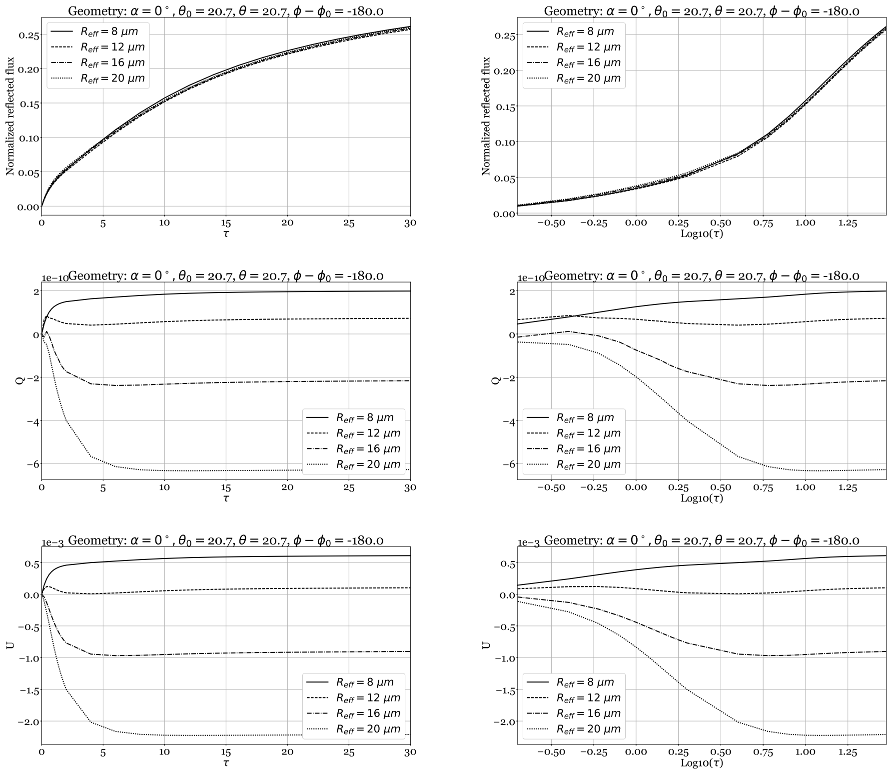


Figure C.5: Similar to *Figure* C.2, but for pixel 3.

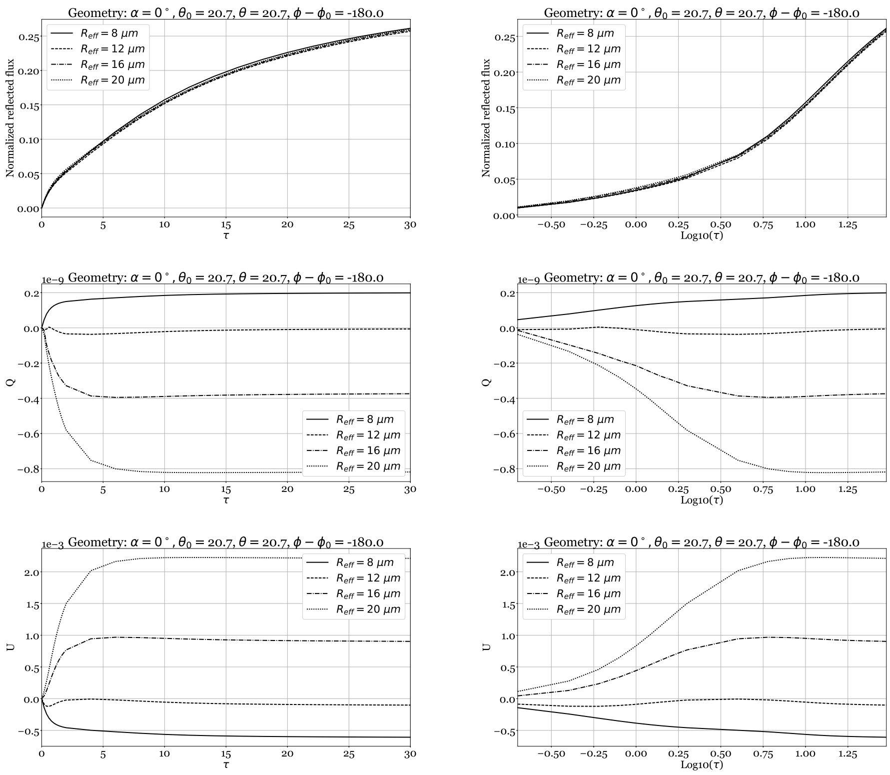


Figure C.6: Similar to *Figure* C.2, but for pixel 4.

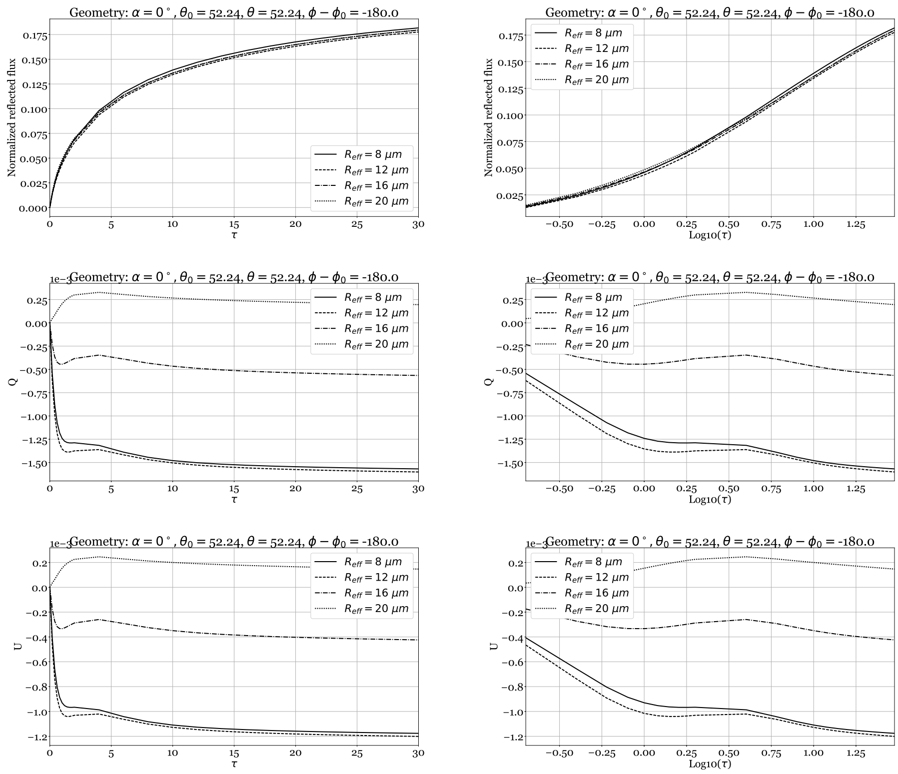


Figure C.7: Similar to *Figure* C.2, but for pixel 5.

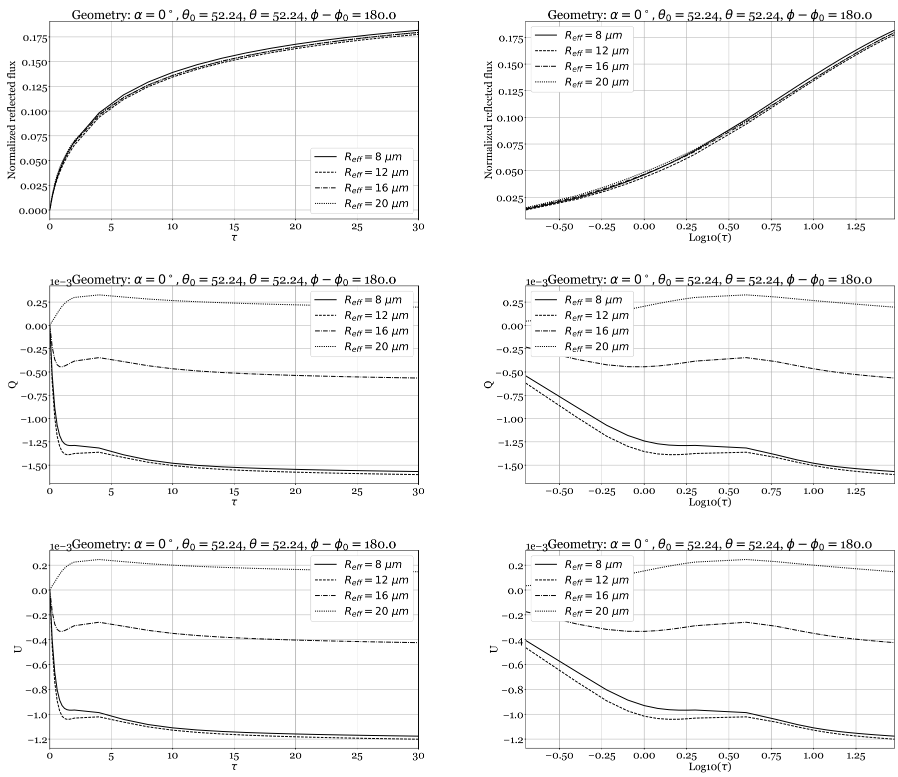


Figure C.8: Similar to *Figure* C.2, but for pixel 6.

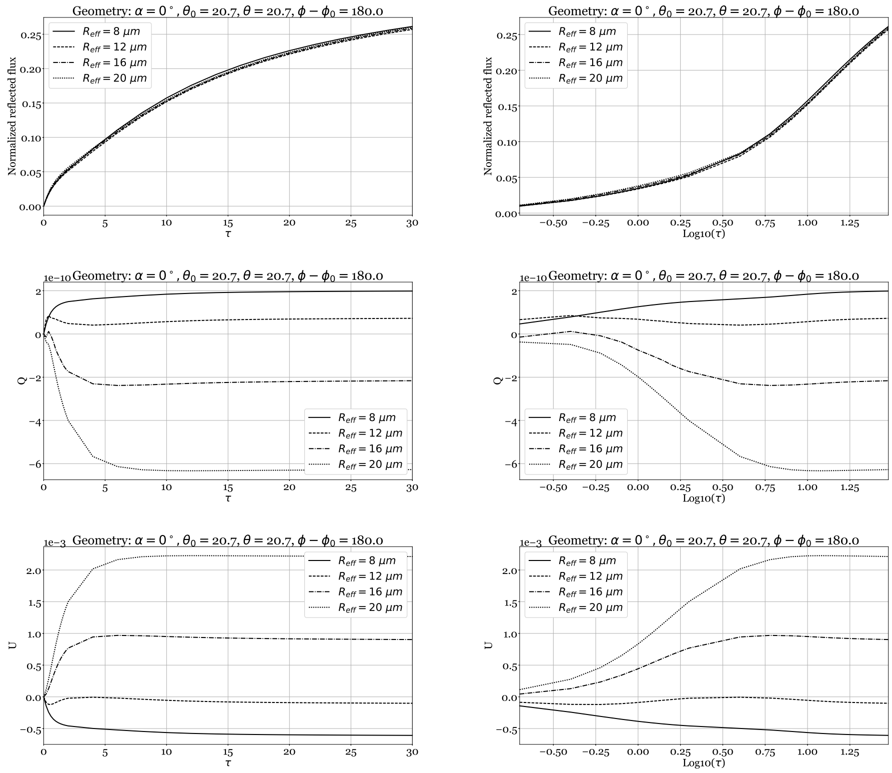


Figure C.9: Similar to *Figure* C.2, but for pixel 7.

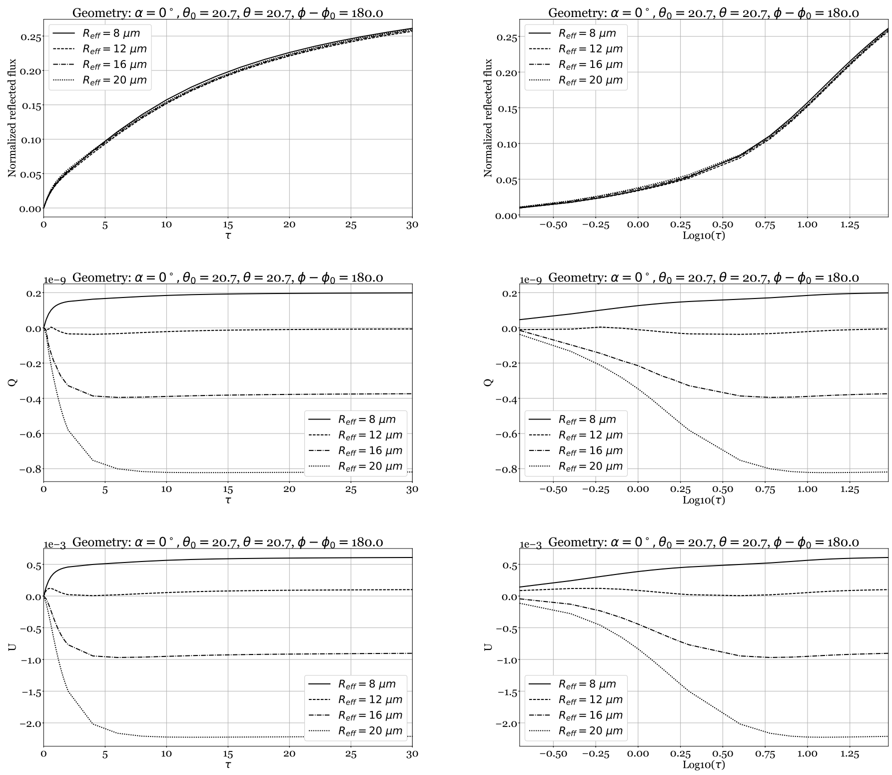


Figure C.10: Similar to *Figure* C.2, but for pixel 8.

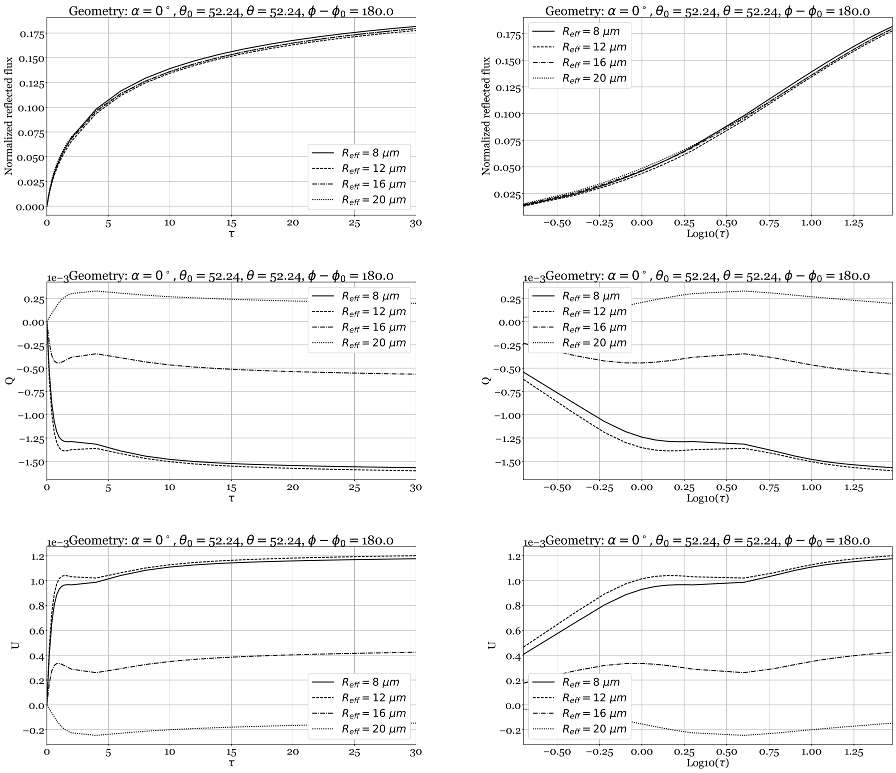


Figure C.11: Similar to *Figure* C.2, but for pixel 9.

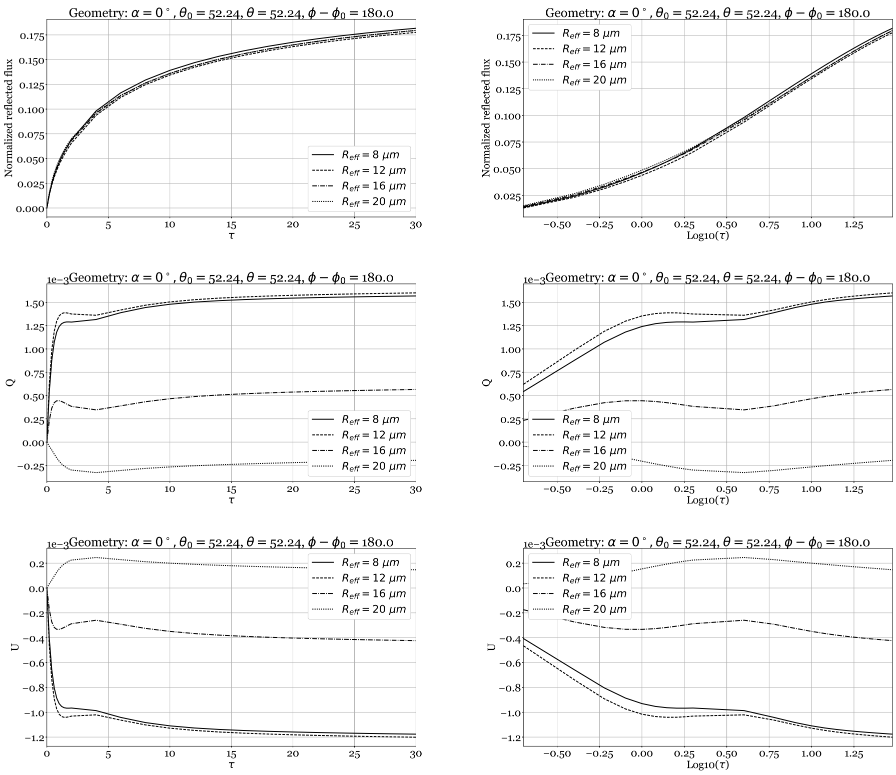


Figure C.12: Similar to *Figure* C.2, but for pixel 10.

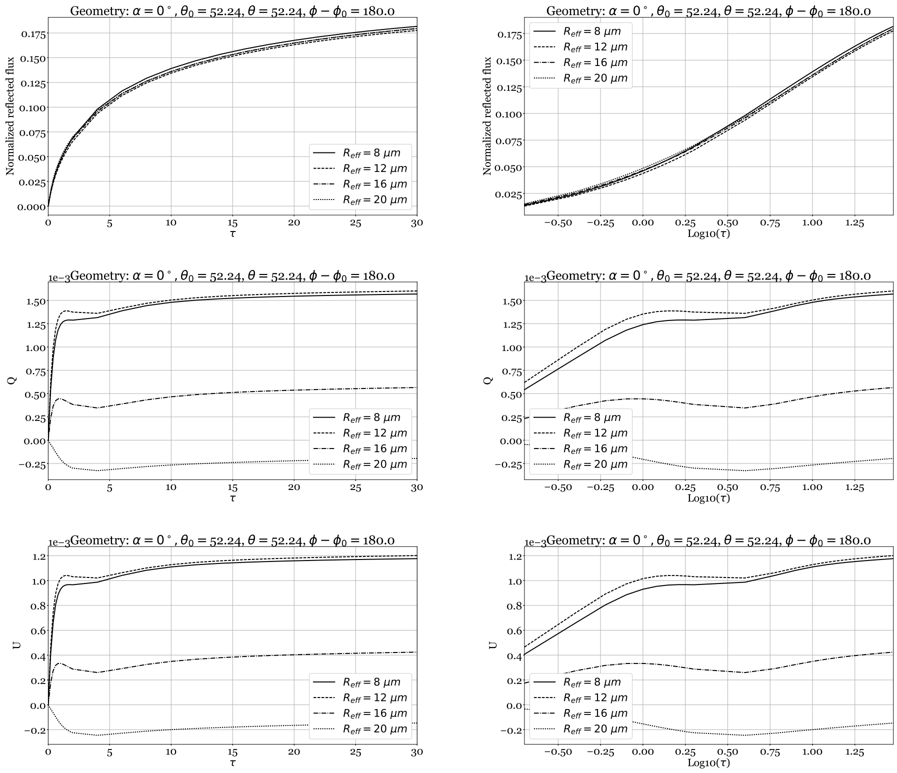


Figure C.13: Similar to *Figure* C.2, but for pixel 11.

D

Polar plots cloudy pixel models

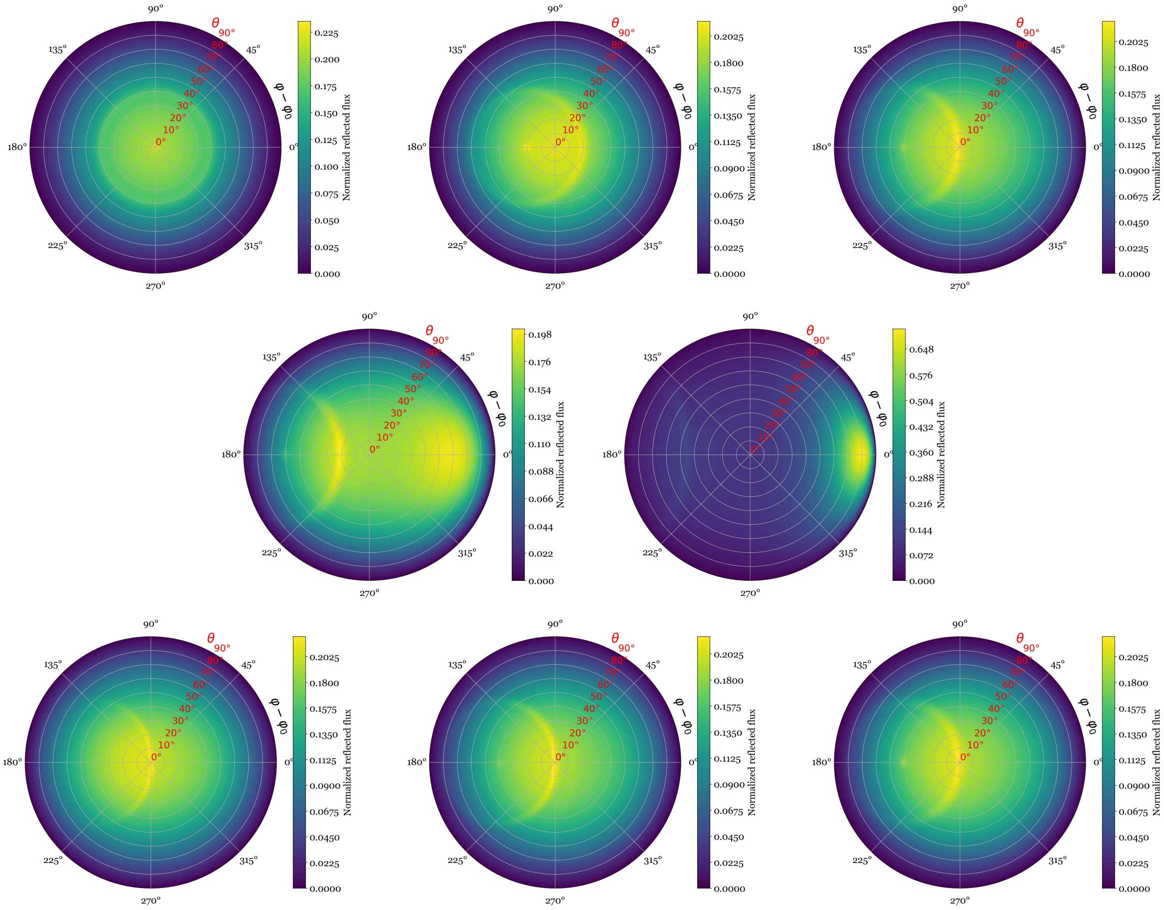


Figure D.1: Polar plots representing the TOA reflection F of a locally horizontally homogeneous, but vertically inhomogeneous atmosphere. The atmosphere includes a layer of clouds. The upper left to the middle right panels correspond to solar zenith angles of $\theta_0 = 0, 20, 40, 60$ and $85°$,respectively. In the bottom panel the polar plots for $\lambda = 350,\ 550$ and $865\ nm$ are provided.

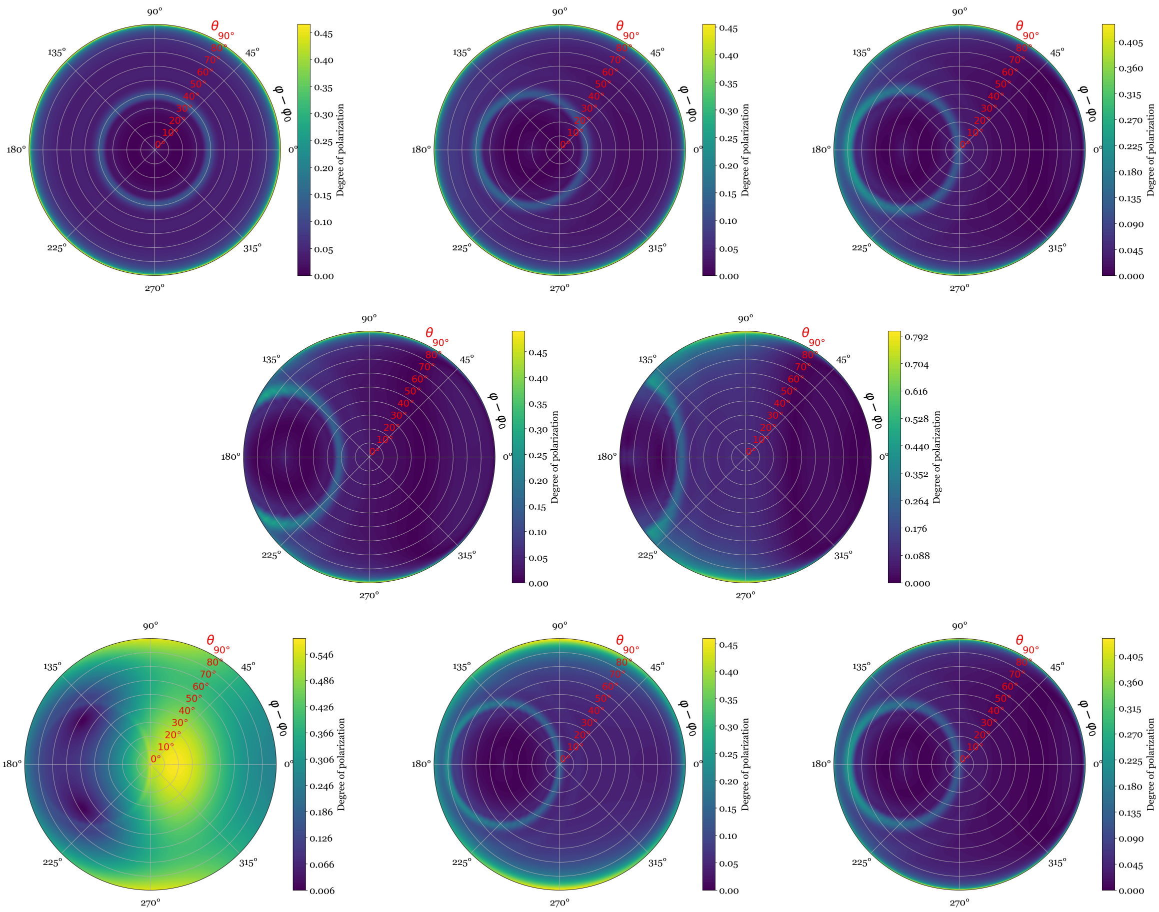


Figure D.2: Similar as *Figure* D.1, except for Pl.

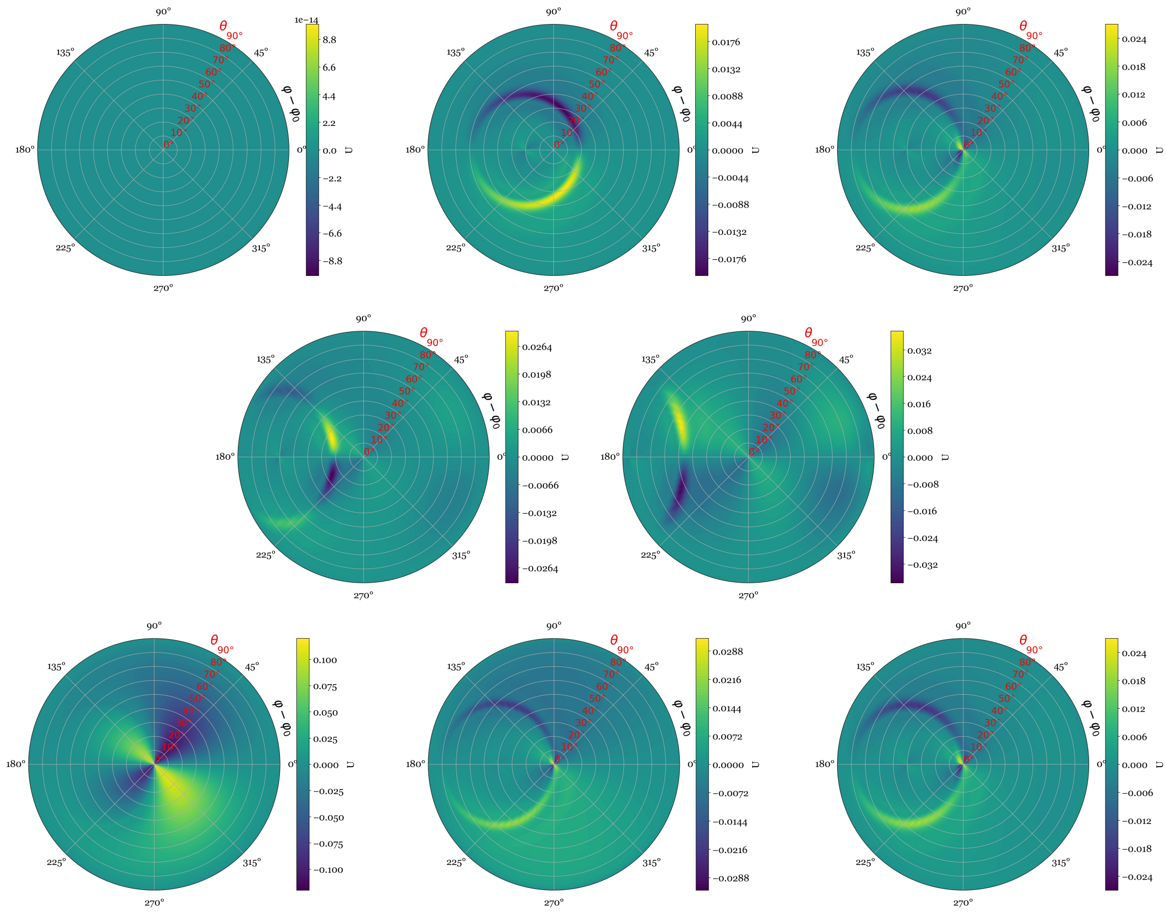


Figure D.3: Similar as *Figure* D.1, except for U.

E

Mutual effect of two cloud parameters on reflected light

In *Section* 5.2 we saw multiple apparent correlations between cloud parameters and the different Stokes elements. Some of these results appeared to be majorly influenced by other parameters, for example in the relations between Pl and CER at $865\ nm$ near the rainbow peak in *Figure* 5.7. In this plot we saw a clear correlations between the mentioned parameters, but also regions with high abundances of data points. In these cases we expect a major dependency on another cloud parameter. Similarly to *Section* 5.2 we will not include any dependency on the cloud optical thickness as we did not retrieve any apparent relationships. The colors that indicate the influence of a second cloud parameter in the scatter plots are defined as: yellowish correspond to high values and blueish to low values. Also, the interfering cloud parameter is provided on the y axis.

The relation between F and CF showed to decrease with increasing wavelength at both phase angles. In the corresponding panels in *Figure* E.1 we can see that especially at $\alpha = 90°$ the interference of CER is dominantly weakening the linear like relation. At the near-infrared CER becomes so dominant that we see some division into three regions, corresponding to high, average and low effective sizes of liquid water particles. For values near the rainbow peak we retrieve the same effect but much less distinctively. The influence of CTP shows to be scattered in a non correlated way. Similarly, the correlation between F and CTP for every wavelength and both phases did not show any strong correlation at all. For CER we saw that a strong correlation occurred only for long wavelength. In *Figure* E.1 one retrieve that at short wavelength CF has large influence on the width of the distribution, essentially disturbing any possible correlation. The same can be found at $550\ nm$, whereas for long wavelength this influence is more scattered over all data points. CTP at both phases, for which we found no strong correlations, shows to be clearly dominated by CF at short wavelength and by CER at long wavelength.

As mentioned in the introduction, we retrieved some correlations that clearly show the presence of another dominating cloud parameter in Pl. The effect of CF was most pronounced for long wavelengths and the rainbow phase. At $865\ nm$ we also saw again some division into regions of high abundant data points. From *Figure* E.2 we can see that this is clearly caused by a strong relation of CER with Pl, which strengthens for longer wavelengths. Also, similarly as for F, CTP shows no clear dominance in CF induced correlations. The retrieved correlations for CER with Pl show to be less affected by any other cloud parameter, albeit showing a regions of high abundant data points. This is in return caused by the variability in cloud fraction, but far less pronounced than for the correlations of Pl with CF. Again, CTP shows to have no dominance in these results and as we observe the scatter plots of Pl versus CTP we see that both CER and CF are clearly dominating the reflecting behaviour, whereas the former shows to be again the strongest degenerescence.

The behaviour of Q showed to provide clear correlations for CF at the green and red wavelengths, whereas we saw a clear behaviour for CTP in the blue. The correlations that were found did not

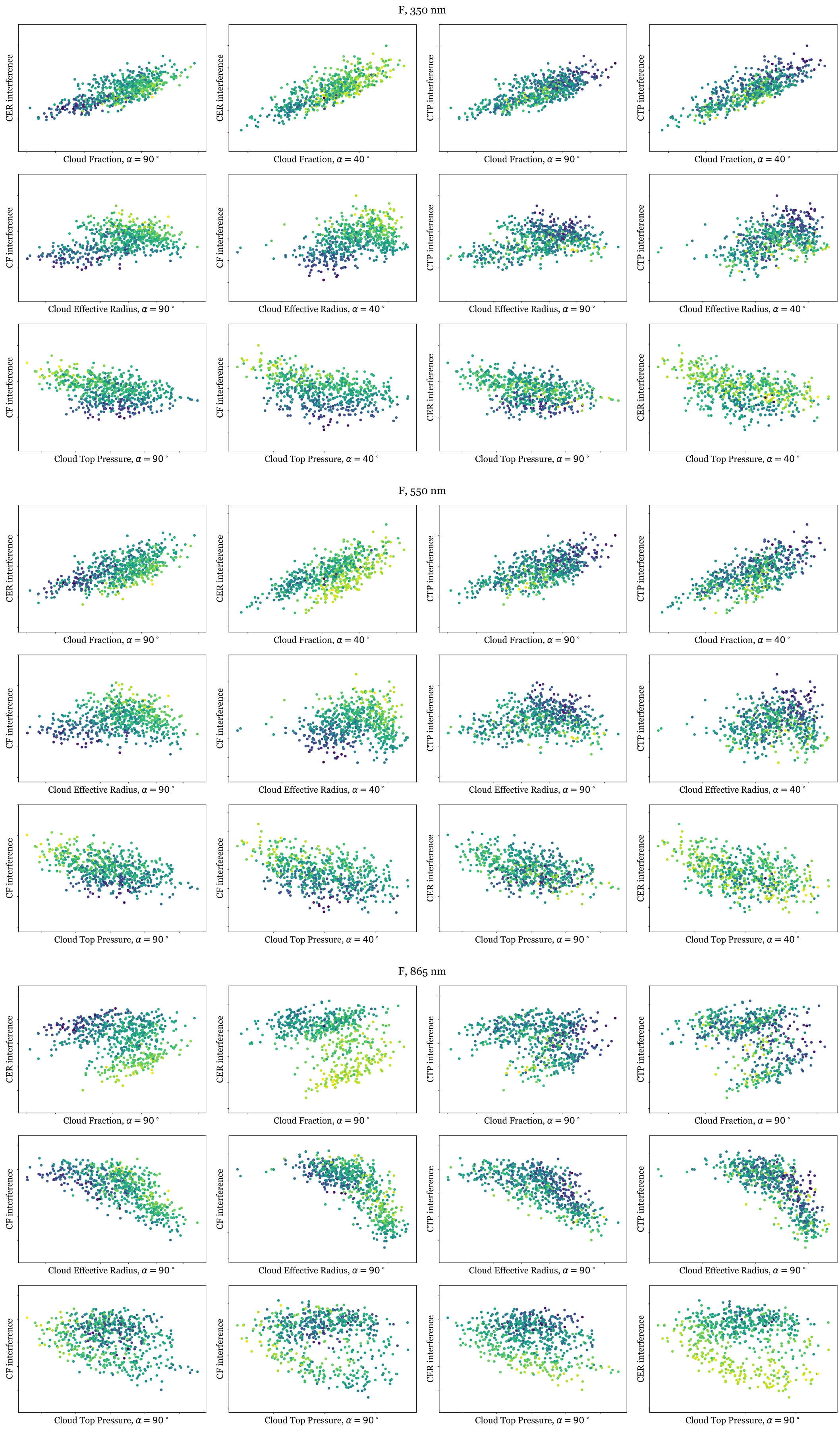


Figure E.1: The influence of a second cloud parameter on the correlations of a single cloud parameter with F. From top to bottom we have provided several plots for 350, 550 and 865 nm. In each sub figure we provided the relation of cloud fraction, cloud particle effective radius and cloud top pressure from top to bottom, respectively. The interference of the cloud fraction, cloud particle effective radius and cloud top pressure are provided for each row. The colors that indicate the influence of a second particular cloud parameter in the scatter plots are defined as: yellowish to high values and blueish to low values. Also, the interfering cloud parameter is provided on the y axis.

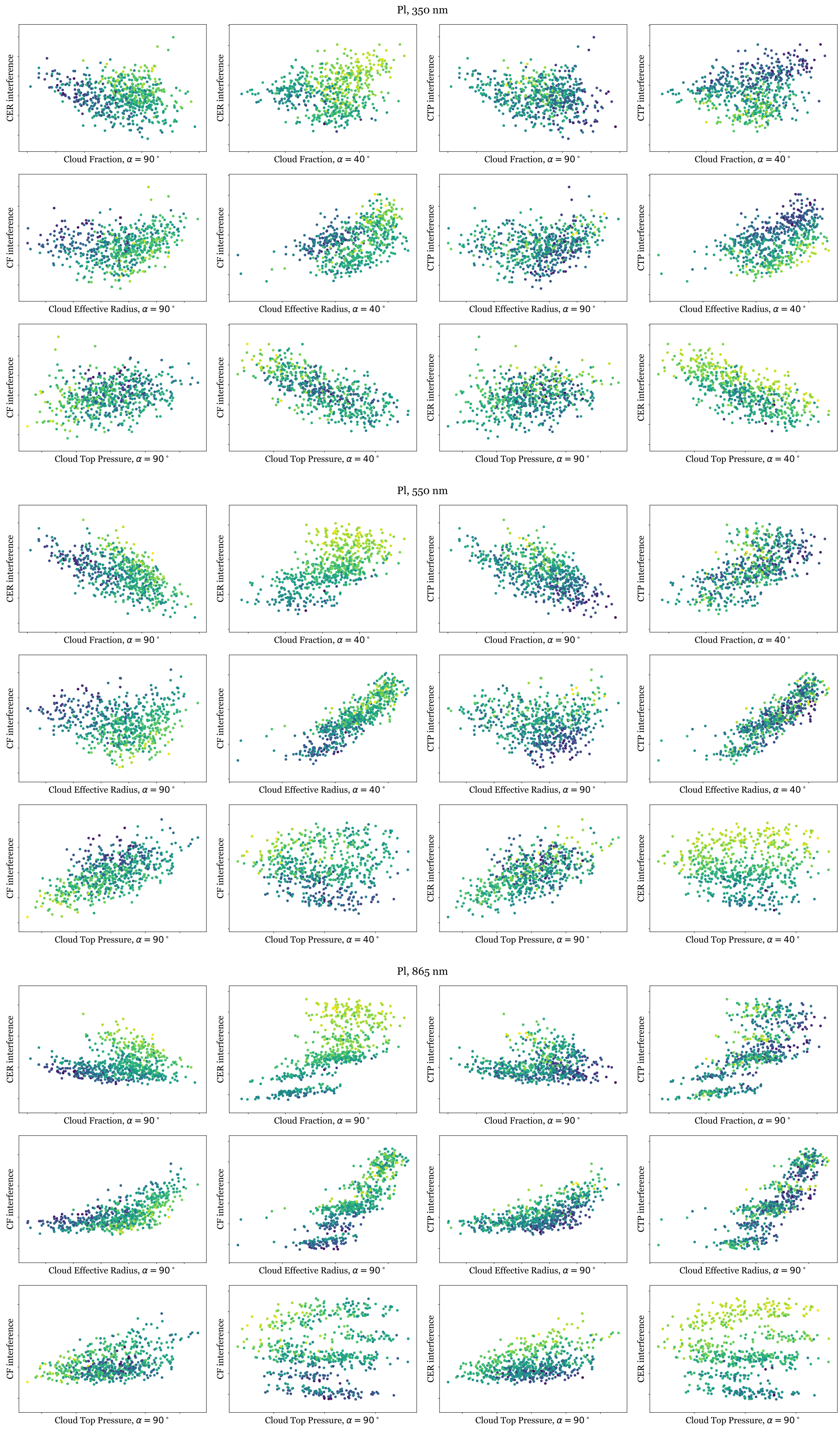


Figure E.2: Similar to *Figure* E.1, except for Pl.

show any major presence of a dominating second cloud parameter, but we will show the dependencies anyway. In *Figure* E.3 the degenerescence for all plots are provided. The correlations of Q with CF show to be barely affected by CER or CTP at long wavelengths. We do, however, observe some agreement with CF and CER in terms of decreasing Q with increasing parameter values. At short wavelength some dominance from CTP is retrieved, not surprising as we already saw that CTP is strongly correlated at this wavelength. For CER we see the same behaviour at the blue wavelength, but with a much more degenerescence from CF for longer wavelengths. For CTP both CF and CER are dominantly present.

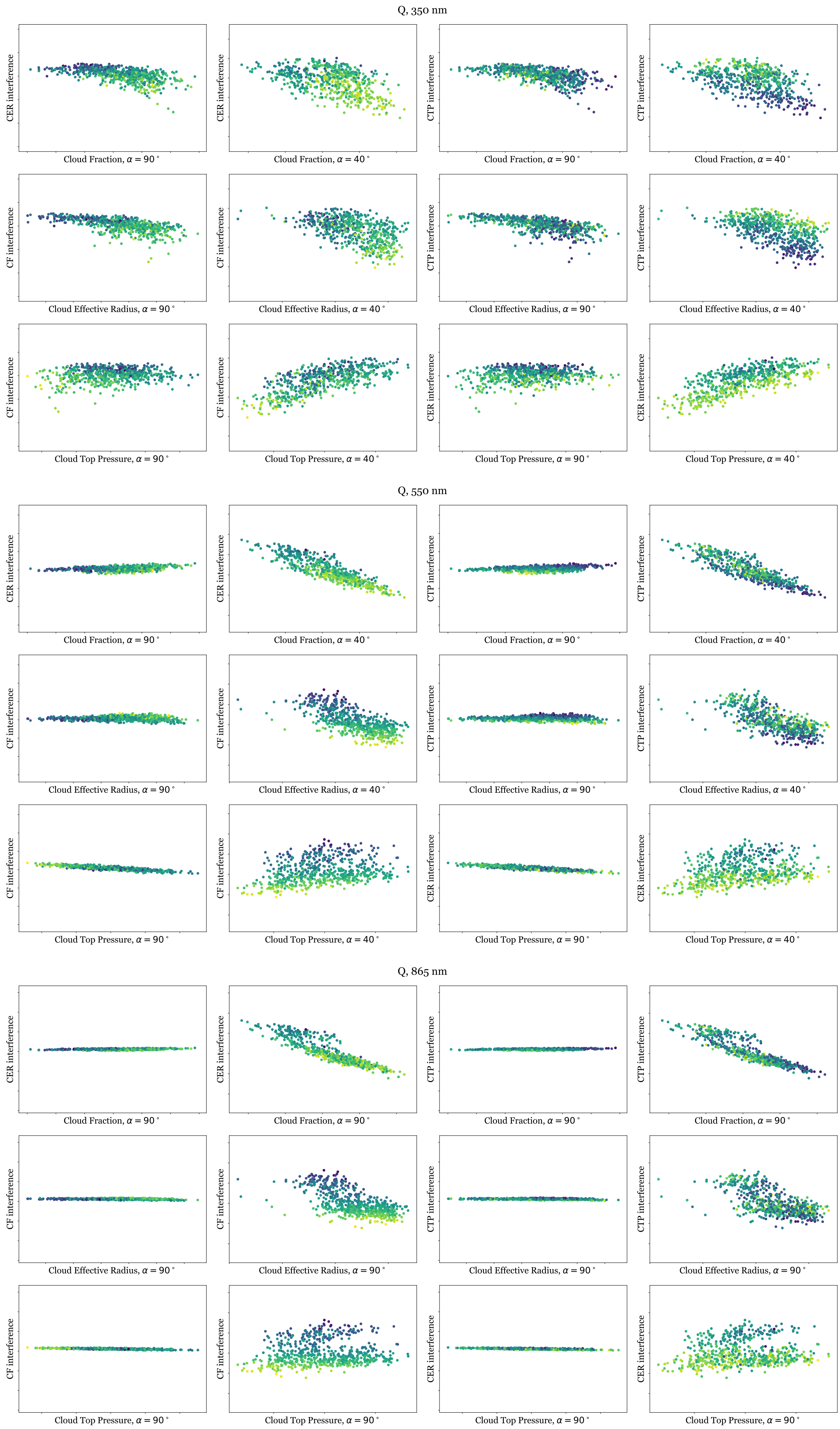


Figure E.3: Similar to *Figure* E.1, except for Q.

F

Additional phase curves

Figure F.1 shows plots that are equivalent to that in *Figure* 4.6, except that we introduced anisotropic polarizing surfaces to the planetary model. In *Figure* F.2 we used *Model 2*, except that all cloud layers are considered as pure gas layers. We also provide the absolute difference with respect to our "complete" *Model 2*. We show this figure to illustrate the effect of the polarizing surface models without the (daily) variability of the clouds.

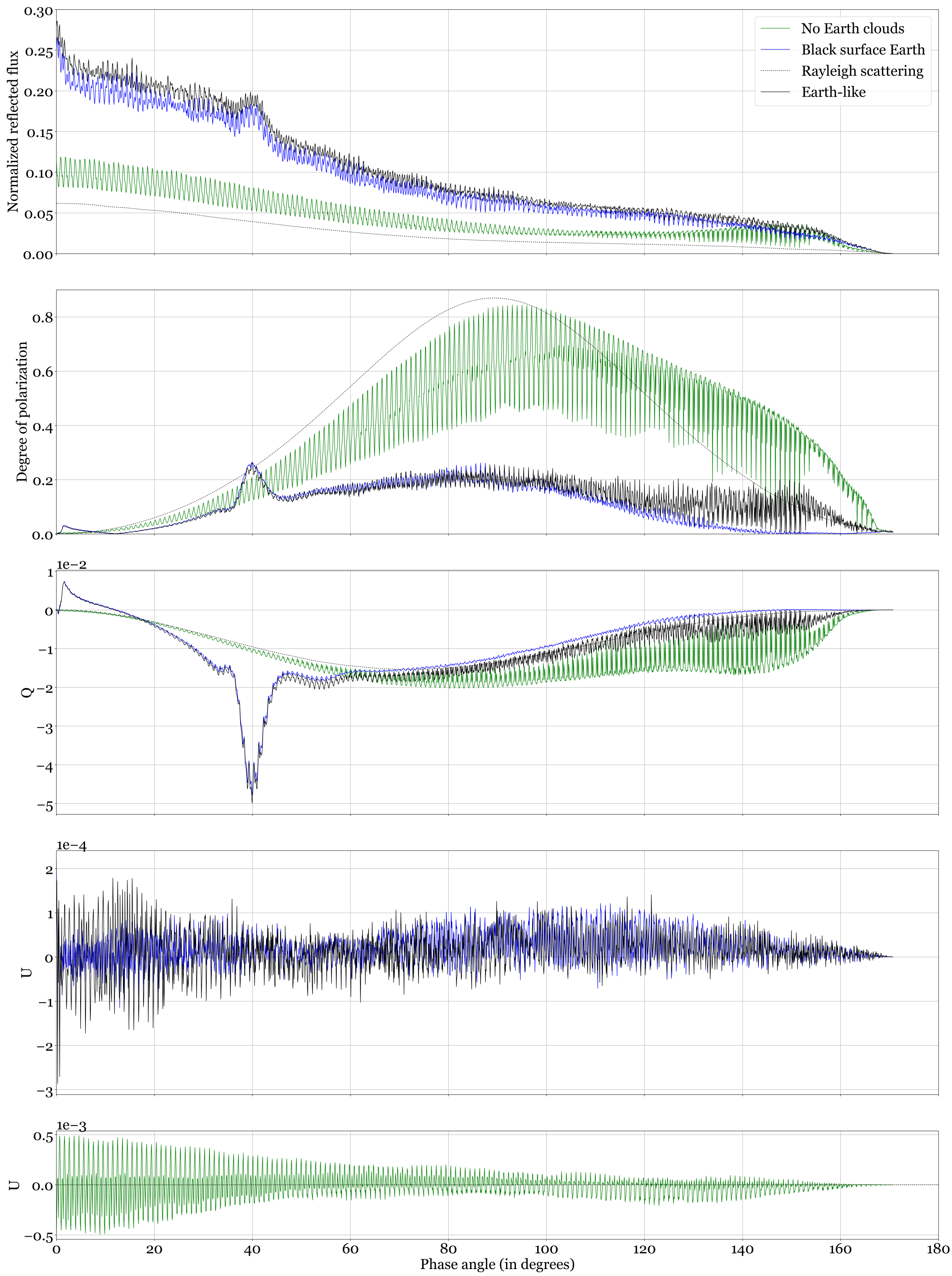


Figure F.1: Phase curves computed at $\lambda = 550\ nm$. From top to bottom we provide F, Q, U and Pl. All subplots show four different cases: the planetary model (Earth-like), the planetary model without cloud layers and a homogeneous black surface (Rayleigh scattering), the planetary model with a black homogeneous surface (Black surface Earth), and the planetary model without cloud layers (No Earth clouds).

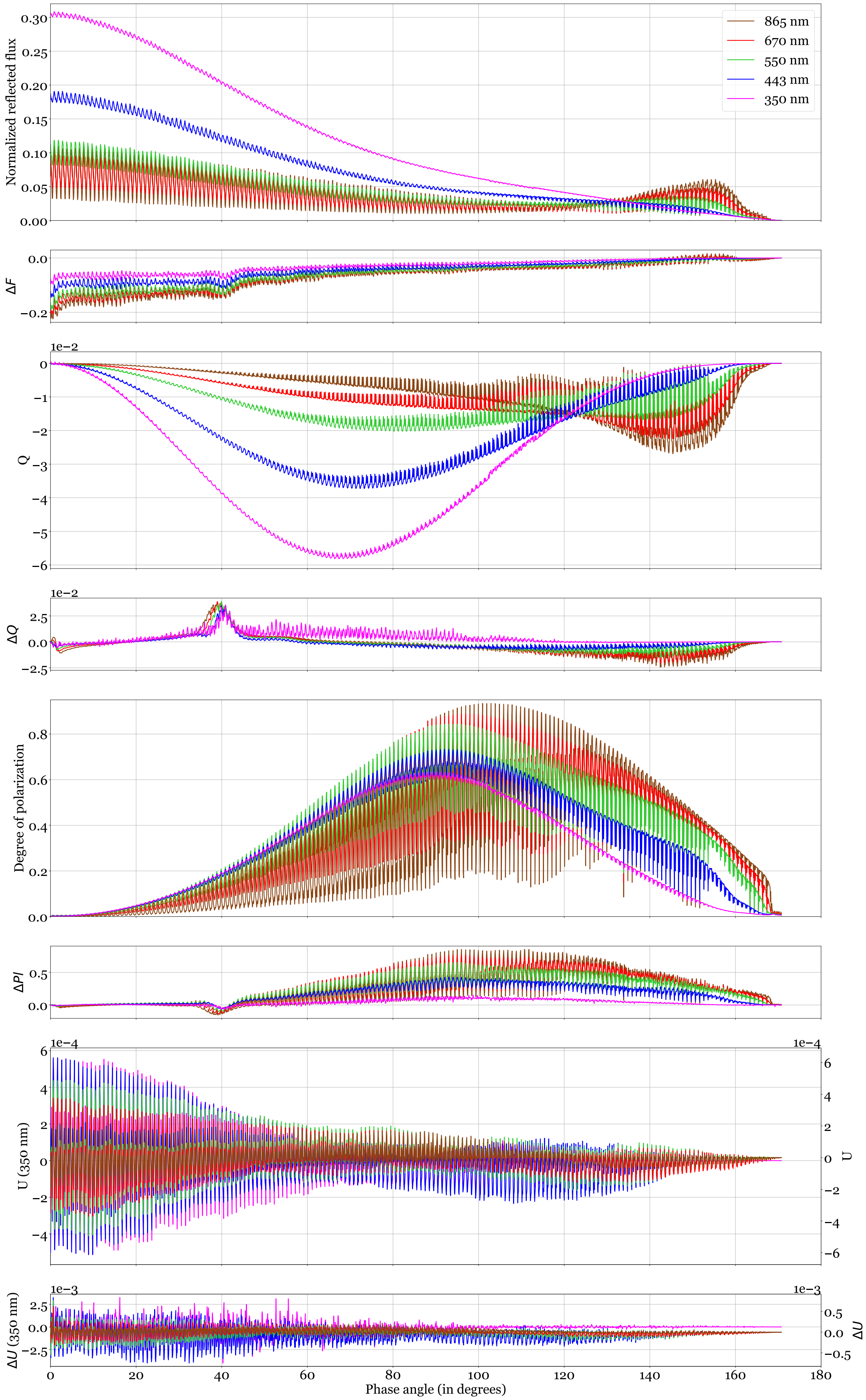


Figure F.2: Similar to *Figure* 6.10, except without considering clouds in the planetary model. For every parameter we provide the absolute difference with respect to the "full" planetary model provided in *Figure* 6.10.

G

Position of glint on a homogeneous ocean planet

Figures G.1 and G.2 show the vertical extend and position of the glint for $\alpha = 90^{\circ}$ and 135°, respectively.

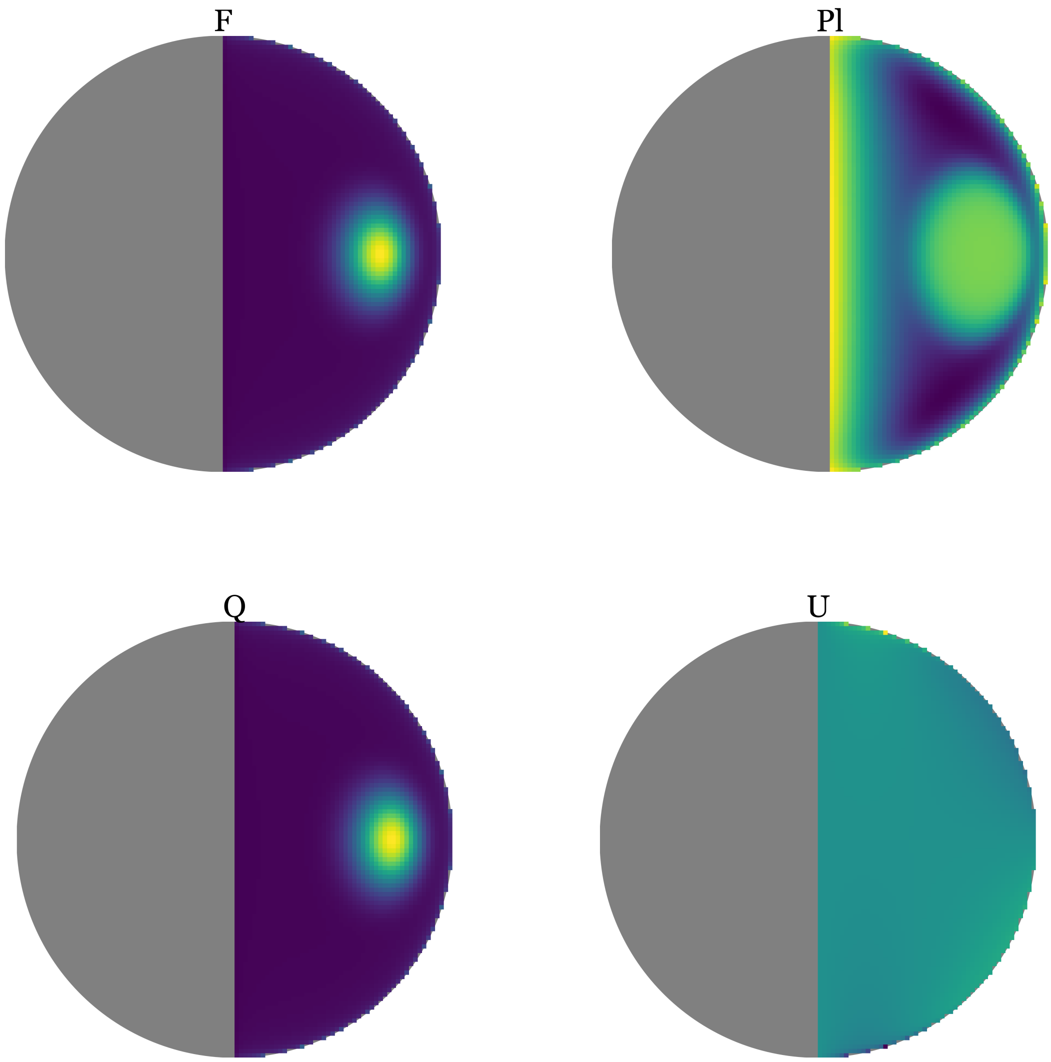


Figure G.1: Spatial extend of the glint for $\alpha = 90°$.

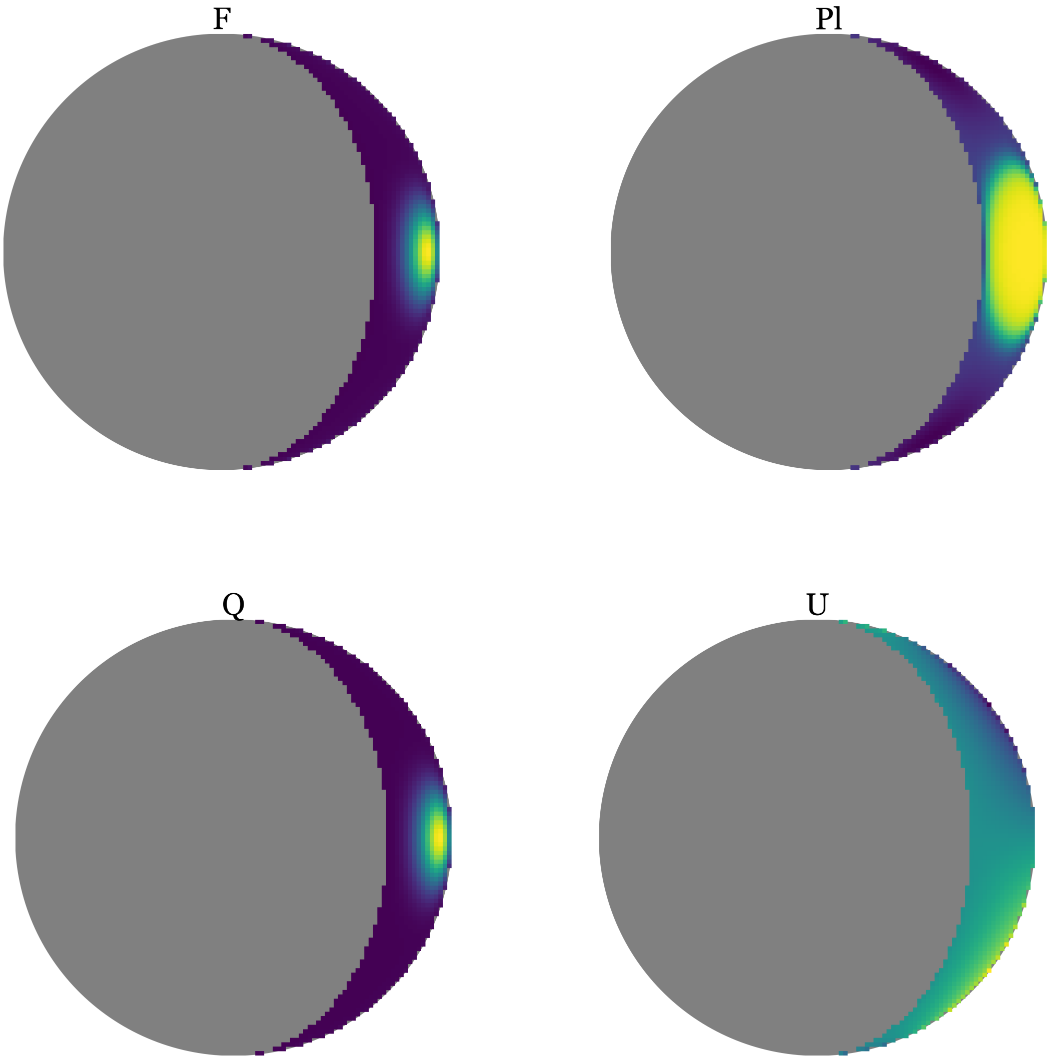


Figure G.2: Spatial extend of the glint for $\alpha = 135°$.