Polarized signatures of a habitable world: comparing models of an exoplanet Earth with visible and near-infrared Earthshine spectra

Gordon, K.E.; Karalidi, T.; Bott, K.M.; Miles-Páez, P.A.; Mulder, W.; Stam, D.M.

Polarized Signatures of a Habitable World: Comparing Models of an Exoplanet Earth with Visible and Near-infrared Earthshine Spectra

Kenneth E. Gordon[1] , Theodora Karalidi[1] , Kimberly M. Bott[2,3] , Paulo A. Miles-Páez[4,7] , Willeke Mulder[5] , and Daphne M. Stam[6]
[1] Planetary Sciences Group, Department of Physics, University of Central Florida, 4111 Libra Drive, Orlando, FL 32816, USA; ke14gordon@knights.ucf.edu
[2] Department of Earth and Planetary Sciences, University of California, Riverside, Riverside, CA 92521, USA
[3] NASA Nexus for Exoplanet System Science, Virtual Planetary Laboratory Team, Box 351580, University of Washington, Seattle, WA 98195, USA
[4] European Southern Observatory, Karl-Schwarzschild-Straße 2, D-85748, Garching bei München, Germany
[5] Leiden Observatory, Leiden University, P.O. Box 9513, 2300 RA Leiden, The Netherlands
[6] Faculty of Aerospace Engineering, Astrodynamics and Space Missions, Delft University of Technology, Kluyverweg 1, 2629 HS, Delft, The Netherlands

Abstract

In the JWST, Extremely Large Telescopes, and LUVOIR era, we expect to characterize a number of potentially habitable Earth-like exoplanets. However, the characterization of these worlds depends crucially on the accuracy of theoretical models. Validating these models against observations of planets with known properties will be key for the future characterization of terrestrial exoplanets. Due to its sensitivity to the micro- and macro-physical properties of an atmosphere, polarimetry will be an important tool that, in tandem with traditional flux-only observations, will enhance the capabilities of characterizing Earth-like planets. In this paper we benchmark two different polarization-enabled radiative-transfer codes against each other and against unique linear spectropolarimetric observations of the earthshine that cover wavelengths from ~ 0.4 to $\sim 2.3\ \mu m$. We find that while the results from the two codes generally agree with each other, there is a phase dependency between the compared models. Additionally, with our current assumptions, the models from both codes underestimate the level of polarization of the earthshine. We also report an interesting discrepancy between our models and the observed 1.27 μm O_2 feature in the earthshine, and provide an analysis of potential methods for matching this feature. Our results suggest that only having access to the 1.27 μm O_2 feature coupled with a lack of observations of the O_2 A and B bands could result in a mischaracterization of an Earth-like atmosphere. Providing these assessments is vital to aid the community in the search for life beyond the solar system.

1. Introduction

Within the last decade, exoplanet missions including Kepler and the Transiting Exoplanet Survey Satellite have been aimed at determining $\eta_\oplus$, the occurrence rate of habitable-zone (HZ) rocky exoplanets around solar-type stars. Recently, Bryson et al. (2020) combined statistical data from the Kepler Data Release 25 planet candidate catalog along with stellar properties measured by the Gaia mission and found that $\eta_\oplus$ for the conservative HZ around main-sequence dwarf stars ranges from $0.37^{+0.48}_{-0.21}$ to $0.60^{+0.90}_{-0.36}$ planets per star, while for the optimistic HZ $\eta_\oplus$ ranges from $0.58^{+0.73}_{-0.33}$ to $0.88^{+1.28}_{-0.51}$ planets per star. Terrestrial exoplanets are therefore expected to be frequent, with more than 180 rocky planets confirmed thus far (as per the NASA Exoplanet Archive). These rocky exoplanets provide for intriguing studies as astronomers strive to discover life on distant worlds. However, to know whether a rocky exoplanet is habitable, we need to be able to characterize its atmosphere and surface.

Currently, transit spectroscopy and the direct detection of planetary flux are the most widely used techniques for performing these characterizations. While proven successful for the gaseous giant exoplanets at very short (transiting; e.g., Line et al. 2013; Sing et al. 2016) or very wide (imaging; e.g., Marois et al. 2008; Desgrange et al. 2022) orbital distances, Earth-sized exoplanets around Sun-like stars are more challenging and introduce a number of issues. In transit, the smaller planetary sizes and wider orbital distances of habitable terrestrial planets compared to gaseous giant planets mean that they only block a minuscule fraction of their host star upon transiting (e.g., Kaltenegger & Traub 2009; Pallé et al. 2011). Additionally, the longer orbital periods of terrestrial planets around solar-type stars require more time for follow-up studies to confirm any possible transits, and their thinner atmospheres complicate spectroscopic characterization (e.g., Bétrémieux & Kaltenegger 2014; Misra et al. 2014). For direct detection, the main issue lies in the minuscule amount of emitted and/or reflected flux from the small planet compared to its host star (e.g., Traub et al. 2010; Seager 2014). Therefore, full characterization of the atmospheres and surfaces of terrestrial exoplanets in the HZs of solar-type stars remains elusive using these traditional methods.

The recent launch of the JWST will help to alleviate some of these issues. JWST will allow for numerous follow-up characterizations of transiting planets across a range of masses and atmospheres (e.g., Greene et al. 2016). However, debate still exists on the feasibility of JWST to effectively characterize the atmospheres of potentially habitable Earth-like planets

---

[7] ESO Fellow.

orbiting cool dwarf stars, mainly due to the necessity for long observational timeframes (see, e.g., Robinson 2018 and Wunderlich et al. 2019, and references therein). For example, Barstow & Irwin (2016) analyzed the three potentially habitable Earth-like planets around TRAPPIST-1: planets 1b, 1c, and 1d. Using JWST's Near-InfraRed Spectrograph (NIRSpec) and Mid-InfraRed Instrument (MIRI), they found that $\gtrsim 60$ transits for 1b and $\gtrsim 30$ transits for 1c and 1d would be required to detect present-day Earth levels of ozone (O_3) on these planets.

Recently, Gialluca et al. (2021) showed that the detectability of biosignature gases at specific wavelength bins (i.e., only considering the central wavelength of strong absorption features) will be very challenging for JWST. However, integrating the spectral impact of a gas across the entire wavelength range of an instrument improves the observational time constraints and could allow for detections of CH_4, CO_2, O_2, and H_2O with a few tens of transits for cloud-free Earth analogs orbiting cool dwarf stars up to 15 pc away. Gialluca et al. (2021) stressed, though, that their results are highly dependent on the assumed planetary atmospheric chemistry and host-star environment of the models, so more in-depth analyses are required. Additionally, Gialluca et al. (2021) did not include clouds in their models, which can severely limit the detectability of key biosignatures such as H_2O (e.g., Suissa et al. 2020). Therefore, while JWST will provide some assistance, a number of degeneracies will still exist in the characterizations of habitable worlds by JWST, such as differentiating between the optical thicknesses and particle-size distributions of clouds.

Polarimetry is a powerful technique that has the ability to break these degeneracies, as it assesses physical aspects of light not measured in nonpolarimetric photometry or spectroscopy. Polarimetry measures light as a vector (by the orientation of the electric field oscillations) rather than only a scalar intensity, thus making it extremely sensitive to the physical and microphysical properties of an atmosphere (e.g., Hansen & Hovenier 1974; Hansen & Travis 1974). Due to the vector nature of polarimetry, it is also sensitive to the location of specific features on the disk of an object (see, e.g., Karalidi et al. 2013; Stolker et al. 2017). Polarimetry has helped to characterize bodies in the solar system, including the clouds of Venus (Hansen & Hovenier 1974) and the gas giants (Schmid et al. 2011) as well as the differing icy conditions of the Galilean moons (Dollfus 1975; Rosenbush 2002). Finally, polarimetry is especially useful in identifying whether reflection is coming from an absorptive surface or a cloud deck, making it highly applicable to studies of habitable worlds (e.g., Fauchez et al. 2017).

Various groups have modeled the flux (e.g., Lincowski et al. 2018; Meadows et al. 2018a; Lustig-Yaeger et al. 2019; Leung et al. 2020) and linear polarimetric signals (e.g., Stam 2008; Karalidi et al. 2011, 2012; Karalidi & Stam 2012; Bailey et al. 2018; Trees & Stam 2019; Groot et al. 2020; Trees & Stam 2022) of Earth-like planets as functions of orbital phase and wavelength. Models, however, need to be validated against observational data. To date, the Earth is the only known and observed habitable "Earth-like" planet, thus serving as a benchmark to infer the biosignatures of life as we know it today.

The best current method of studying the Earth-as-an-exoplanet is through observations of the earthshine: sunlight scattered or reflected by the dayside of the Earth and reflected back to the planet by the nightside of the Moon, where it can then be measured by ground-based facilities. Studies of the optical and near-infrared (NIR) earthshine flux spectra (e.g., Woolf et al. 2002; Turnbull et al. 2006; Pallé et al. 2009) reveal diagnostic biosignatures of the Earth, including the vegetation red edge (VRE), the ocean glint, and spectral features of atmospheric O_2 and H_2O. Studies on the spectropolarization of the earthshine (Dollfus 1957; Sterzik et al. 2012, 2019, 2020; Bazzon et al. 2013; Takahashi et al. 2013, 2021; Miles-Páez et al. 2014) detected the same biosignatures, and showed the sensitivity of polarization to features such as water clouds, varying surfaces, and ocean glint.

With the increased interest and expected near-future surplus of available data for terrestrial exoplanets, efforts have been made to benchmark and validate different codes focused on modeling these planets (e.g., Fauchez et al. 2021; Paradise et al. 2022). Here, we focus on the polarization of the reflected light from terrestrial planets and provide the first benchmarking of two independent polarized radiative-transfer codes against each other and against observations. In particular, we present the first comparisons of theoretical models from the Doubling Adding Program (DAP) code (de Haan et al. 1987; Stam 2008) and the Versatile Software for Transfer of Atmospheric Radiation (VSTAR; Spurr 2006; Bailey et al. 2018) against each other for a range of cloud and surface properties, and then compare these models against the visible to near-infrared (VNIR) observed earthshine spectrum presented in Miles-Páez et al. (2014). This spectrum provided the first extension of the wavelength range for linear spectropolarimetry of earthshine into the NIR regime, reporting the first detection of NIR H_2O around 0.93 and 1.12 μm and NIR O_2 around 1.27 μm in polarized earthshine data. To create our model exoplanet Earth we utilized cloud and land cover properties derived from observations by the Moderate Resolution Imaging Spectroradiometer (MODIS) instrument on board NASA's Terra and Aqua satellites (King et al. 2004), and corrected the observations for lunar depolarization.

The outline of this paper is as follows. Section 2 presents the observational data as well as the satellite data used to produce our models. In Section 3, we provide short descriptions of the two different codes that we compare in this paper. Then, in Section 4 we highlight the importance of clouds in modeling terrestrial exoplanets and the effects that changes in cloud parameters have on the resulting polarization signals. Section 5 compares the resulting degree of polarization from the two codes against each other and the earthshine observations. In Section 6 we provide further analysis on the 1.27 μm O_2 band in the earthshine spectrum. Finally, in Section 7, we summarize our results and describe possible improvements to the different numerical simulations, as well the implications of our models on future missions aimed at enhancing our understanding of Earth-like exoplanets.

2. The Data

2.1. Earthshine Observations

Miles-Páez et al. (2014) provided the first linear spectro-polarimetric measurements of the earthshine in the VNIR wavelengths. Miles-Páez et al. (2014) utilized the Andalucía Faint Object Spectrograph and Camera (ALFOSC) instrument of the 2.56 m Nordic Optical Telescope (NOT) as well as the
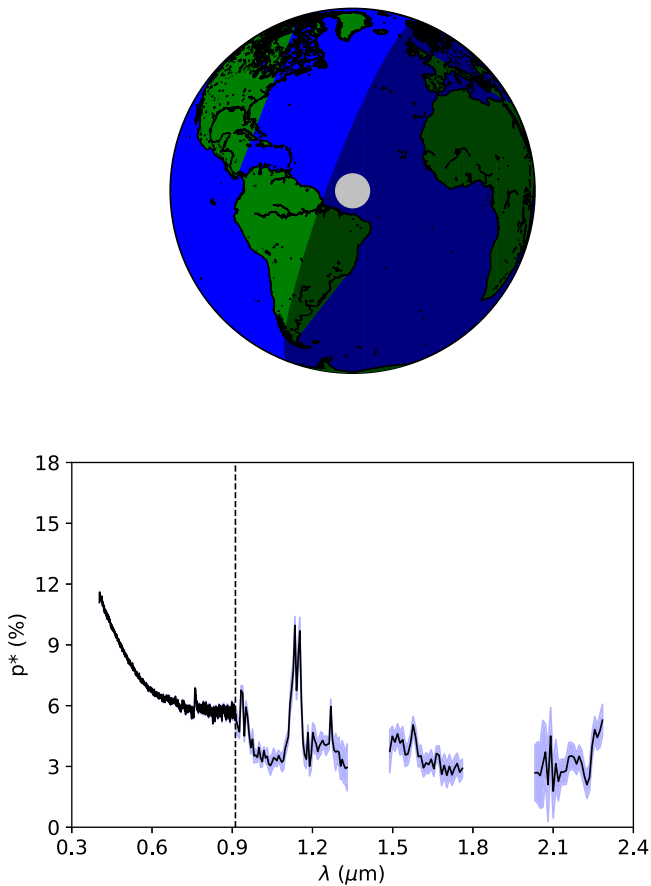
Figure 1. Top: the zenith location of the Moon on planet Earth at 21:20 UT on 2013 May 18: latitude 4° 36′ north, longitude 40° 36′ west. The lit portion of the globe is what was visible from the lunar surface. Bottom: the VNIR earthshine observations as published in Miles-Páez et al. (2014), showing the debiased degree of linear polarization p^* as a function of wavelength λ. Wavelengths of strong telluric absorption have been removed. The vertical dashed line around 0.9 μm separates the ALFOSC and LIRIS data. The uncertainty per wavelength is given by the blue shaded region.

Figure 2. Top: the planetary scattering plane defines the plane through the centers of the star, the planet (blue circle), and the observer (green circle). Here, the observed planet is assumed to be spherical with radius r and located at distance $d(\gg r)$ from the observer. The flux vector reflected by the planet toward the observer ($\pi \mathbf{F}$) depends on the incident flux vector from the star ($\pi \mathbf{F}_0$) and the planetary phase angle α. The black arrow represents the tilt of the reference plane for the earthshine observations with respect to the planetary scattering plane. Bottom: schematic diagram of how the phases of the Earth mirror the phases of the Moon as the latter orbits the former. The yellow arrows indicate the incident sunlight. The red (green) lines define the extent of the lunar (Earth) surface that can be seen from the Earth (Moon). In the leftmost (rightmost) panel, the Moon disk as seen from the Earth is illuminated 100% (0%), corresponding to a lunar phase $\alpha_M = 0°$ (180°), while the Earth disk as seen from the Moon is illuminated 0% (100%), corresponding to an Earth phase $\alpha_E = 180°$ (0°). The in-between phases are complementary to each other.

Long-slit Intermediate Resolution Infrared Spectrograph (LIRIS) of the 4.2 m William Herschel Telescope to capture simultaneous optical (ALFOSC; 0.4–0.9 μm, spectral resolution of 2.51 nm) and NIR (LIRIS; 0.9–1.4 μm, spectral resolution of 1.83 nm; 1.4–2.4 μm, spectral resolution of 2.91 nm) linear spectropolarimetric observations of the earthshine, thereby extending the knowledge of linear polarimetric earthshine data into redder wavelengths than previous studies and allowing for the detection of important biosignature features not found at visible wavelengths. To increase the signal-to-noise ratio (S/N) of the observations in the NIR, they applied a 10 pixel binning to the spectrum in this wavelength region. The final spectrum therefore had a variable resolving power R ($=\lambda/\Delta\lambda$) of $\sim$250 at $\lambda = 0.65$ μm and $R \sim 208$ at $\lambda = 1.27$ μm. For more information on the observations and calibrations of the data, we refer the reader to Miles-Páez et al. (2014).

Figure 1 shows the Earth as observed from the Moon during the Miles-Páez et al. (2014) observations (top) and the observed polarimetric earthshine spectrum (bottom). The reference plane of the observations is tilted by 29° from the planetary scattering plane, which is the plane through the centers of the star, planet, and observer (see top panel of Figure 2). The earthshine observations were obtained on a night
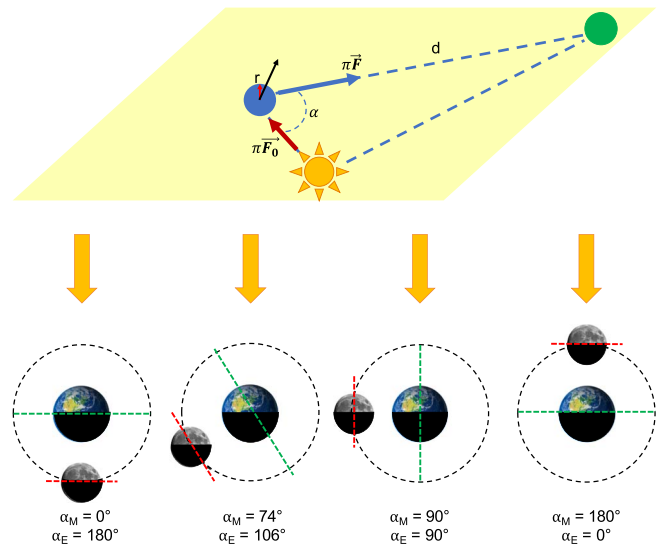
when the waxing Moon illuminated area was 59% (Miles-Páez et al. 2014). Due to the geometry of the Earth–Moon system, Earth phases are exactly opposite of lunar phases (see bottom panel of Figure 2). This means that the Earth would have been illuminated 41% on this night (as seen from the Moon), which corresponds to a phase angle of $\alpha = 106°$ during the observations. During the observations the largest contributors to the earthshine were the western Atlantic Ocean and the Amazon rainforest. In the wavelength range covered by these observations a number of strong absorption features exist, including those of atmospheric O_2 and H_2O, which are important biomarkers. Additionally, the presence of the Amazon rainforest in the field of view leaves a mark on the observations, with a small yet detectable VRE near $\lambda = 0.7$ μm.

2.2. Exoplanet-Earth Models

To model the horizontal inhomogeneity of the visible Earth disk, we divided our model planet into pixels of $2° \times 2°$ such that the local properties of the atmosphere and surface are plane-parallel and horizontally homogeneous. These pixels are large enough to be able to ignore effects from surrounding pixels (e.g., light that is reflected or scattered by clouds in one pixel toward the surface of an adjacent pixel). We then ran each pixel through a radiative-transfer code which produces the full spectropolarimetric signal of the pixel across all wavelengths, λ, in the VNIR (0.4–2.3 μm) and all phase angles, α, from 0° to 180° in steps of 2°. After running every pixel in the illuminated section of the Earth disk, we used a weighted averaging to combine the signals from all pixels and generate the

disk-integrated linear polarization spectrum for our model exoplanet Earth.

To test that our results are independent of the code we use, we compared the output of two separate codes (described in Section 3) for a range of key atmosphere and surface combinations (Section 5). In future work we will expand these comparisons to more surface–atmosphere combinations and to other planets of the solar system.

2.2.1. Model Atmospheres and Clouds

All model pixels have vertically heterogeneous atmospheres composed of stacks of horizontally homogeneous and locally plane-parallel layers, which each contain gas molecules and (if desired) cloud particles. The atmosphere is bounded below by a flat, homogeneous surface. Our model atmospheres have a temperature and pressure (T–P) profile representative of a midlatitude Earth atmosphere as defined by McClatchey et al. (1972), divided into 16 total layers (see Figure 15).

We modeled our clouds using MODIS Terra and Aqua satellite data (e.g., King et al. 2004). We utilized the Level-3 MODIS Atmosphere Daily Global (MADG) product for 2013 May 18, i.e., the day of observation for the earthshine measurements from Miles-Páez et al. (2014) for our model cloud properties. We describe a model cloud by three main properties: its cloud top pressure (p_c), its optical thickness (b_c), and its particle-size distribution. Our $2° \times 2°$ model pixels are much larger than typical horizontal variations seen in clouds on Earth, and they are larger than the MODIS pixel size used in the MADG product. We therefore calculated the average of the three cloud properties of all MODIS pixels within each model pixel and used these average values for the cloud properties of that model pixel. The MODIS database provides a p_c for all of its pixels, and after calculating the average p_c for our model pixel, we used our atmospheric T–P profile to determine in which layer the cloud should be placed. All clouds in our model pixels were limited in pressure and temperature so as to avoid mixed-phase clouds or ice clouds. We calculated the cloud b_c for each model pixel by averaging the cloud optical thicknesses for each of the MODIS pixels within our model pixel. The wavelength at which a b_c is specified in the MODIS data, λ_b, depends on the surface below the cloud, and we kept this dependency in our model runs. We capped b_c at 50 because values above this provided negligible impact on the resulting polarization (see Section 4.1).

We modeled our cloud particles using Mie theory (de Rooij & Van der Stap 1984). Following the MODIS data (Figure 3, blue dots), the model clouds use the two-parameter gamma distribution of Hansen & Travis (1974), which is governed by a particle effective radius, r_{eff}, in μm and a dimensionless effective variance, u_{eff}. Terrestrial liquid water clouds are composed of droplets with radii ranging from $\sim$4–5 μm up to $\sim$30 μm (e.g., Han et al. 1994). The MODIS cloud particle-size distribution measured on the day of the earthshine observations is indeed best matched by a two-parameter gamma distribution (Figure 3, solid black line) with radii ranging from $\sim$4 to $\sim$30 μm and a maximum in the distribution at $\sim$14 μm.

To analyze the full range of particle radii, we studied three different types of clouds: small droplets with $r_{\mathrm{eff}} = 6$ μm, $u_{\mathrm{eff}} = 0.4$ (type B), medium droplets with $r_{\mathrm{eff}} = 10$ μm, $u_{\mathrm{eff}} = 0.03$ (type C), and large droplets with $r_{\mathrm{eff}} = 14$ μm, $u_{\mathrm{eff}} = 0.04$ (type D). Clouds with larger water droplets require significantly more computing power and are more difficult for our
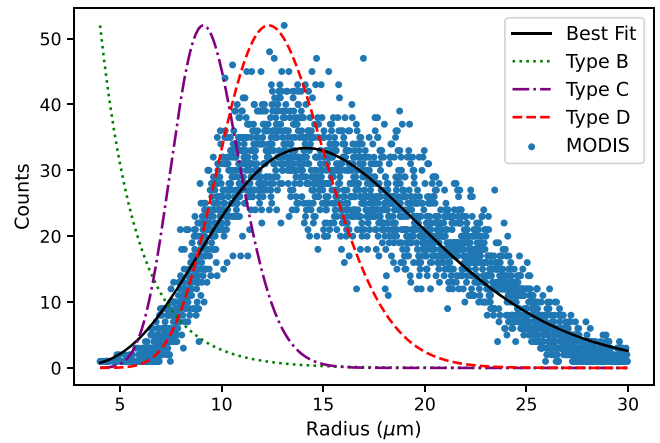

Figure 3. The MODIS liquid water cloud particle sizes (blue dots), as measured by the MADG product for 2013 May 18 (the day of the earthshine observations). Overplotted are the best-fit two-parameter gamma size distribution (solid black line) and the type B (dotted green line), type C (dashed–dotted purple line), and type D (dashed red line) particle-size distributions. The three cloud-type distributions have been normalized in this image to the maximum height of the MODIS data, and they span the majority of the particle sizes measured by MODIS.

codes to handle. Additionally, the type B clouds are similar to those used by previous theoretical models of the Earth (e.g., van Diedenhoven et al. 2007) as representative of average terrestrial liquid water clouds. Therefore, here we used type B cloud particles in our models. In Section 4.2 we discuss the effect that the different cloud particle sizes have on our model spectra.

The real part of the complex refractive index of water, n_w, is slightly wavelength dependent in the VIS and NIR wavelengths, ranging from 1.344 at $\lambda = 0.4$ μm to 1.320 at $\lambda = 1.0$ μm (Daimon & Masumura 2007; van Diedenhoven et al. 2007). The imaginary part of n_w (Im{n_w}) is small but varies greatly, from $\sim 10^{-8}$ at 0.3 μm to $\sim 10^{-3}$ at 3 μm, with a minimum of $\sim 8 \times 10^{-10}$ at 0.5 μm (Segelstein 1981; Pope & Fry 1997). In Section 4 our models have a wavelength-independent $n_w = 1.335 \pm 0.00001i$ (see Karalidi et al. 2011, and references therein) for simplicity. In Section 4.2 we discuss the impact that changing n_w has on our model spectra. Finally, for the comparisons between the two codes in Section 5.2, we used the wavelength-dependent n_w from Hale & Querry (1973).

2.2.2. Model Surfaces

We used the Level-3 MODIS Yearly Global Land Cover Types product for 2013 (i.e., the year of observations for the earthshine measurements) for our model surfaces. This product identifies 17 total MODIS surface classes, including 11 natural vegetation classes, three human-altered classes, and three nonvegetated classes. As we modeled the Earth-as-an-exoplanet, we condensed these 17 original MODIS surface classes down to five for simplicity: ocean, forest (a combination of deciduous and conifer), grass, sand, and snow and ice. When more than one MODIS surface type existed in our model pixel, we assigned the most abundant MODIS surface type as the surface of our model pixel. Our surfaces are modeled as ideal, depolarizing Lambertian surfaces with variable albedos as a function of wavelength. We used the NASA JPL EcoStress Spectral Library (Baldridge et al. 2009; Meerdink et al. 2019) to retrieve surface albedo spectra for each surface category.[8] In

---

[8] https://speclib.jpl.nasa.gov

a following paper we will study the effect of realistic surfaces by incorporating their bidirectional polarization distribution functions (BPDFs) so that their reflection is nonisotropic.

3. The Numerical Codes

3.1. Flux and Polarization Definitions

We describe the starlight that is reflected by a planet by the Stokes vector $\pi \mathbf{F}$ (see, e.g., Hansen & Travis 1974; Hovenier et al. 2004), as

$$\pi \mathbf{F} = \pi \begin{bmatrix} I \\ Q \\ U \\ V \end{bmatrix}, \qquad (1)$$

where parameter πI is the total flux, parameters πQ and πU are the linearly polarized fluxes, and parameter πV is the circularly polarized flux. All four parameters depend on the wavelength λ, and their units are in watt per square meter or W m^{-2} m^{-1} when defined per wavelength. Fluxes πQ and πU are defined with respect to a reference plane, and in this case we use the planetary scattering plane, which is the plane through the centers of the star, planet, and observer (see top panel of Figure 2).

Based on observations of the Sun (e.g., Kemp et al. 1987) and other active and inactive FGK stars (e.g., Cotton et al. 2017), the light of a solar-type star can be assumed to be unpolarized when integrated over the stellar disk. Starlight that has been reflected by a planet is usually polarized because it has been scattered by gases and aerosols or cloud particles in the planetary atmosphere and/or has been reflected by the surface. The total degree of polarization of this light is defined as the ratio of the polarized fluxes to the total flux, or $P_{\mathrm{tot}} = \sqrt{Q^2 + U^2 + V^2} / I$.

The parameter πV of sunlight that is reflected by an Earth-like planet is expected to be very small (e.g., Hansen & Travis 1974; Rossi & Stam 2018), and therefore we ignored it in our numerical simulations. Ignoring πV does not lead to any significant errors in the calculated total and polarized fluxes, as discussed by Stam & Hovenier (2005). P_{tot} can therefore usually be represented by the degree of linear polarization $P = \sqrt{Q^2 + U^2} / I$.

For a planet that is mirror-symmetric with respect to the planetary scattering plane (i.e., the planet is mostly homogeneous), parameter πU will be effectively zero (e.g., Hovenier 1970). In this case, we define the *signed* degree of linear polarization as

$$P_s = \frac{-Q}{I}. \qquad (2)$$

The signed degree of linear polarization also indicates the direction of the polarization: if $P_s > 0$, the light is polarized perpendicular to the plane containing the incident and scattered light, whereas if $P_s < 0$ the light is polarized parallel to the plane.

3.2. Doubling Adding Program Code

The radiative-transfer code DAP uses an efficient adding-doubling algorithm, first described by de Haan et al. (1987), which fully incorporates single and multiple scattering by atmospheric gases as well as aerosol and cloud particles. The adding method allows for the scattering properties of an atmosphere, composed of a stack of individual layers, to be calculated from the scattering properties of the individual layers, taking into account the repeated reflections at the interfaces between these layers. The doubling method is an extension of the adding method where, if the individual layers are identical, the scattering properties of the combined layer can be obtained through a rapid geometrical doubling manner (see, e.g., Hansen & Travis 1974, and references therein). The code then uses a fast, numerical disk-integration algorithm (Stam et al. 2006) to integrate the reflected light across the planet for all planetary phase angles, α. DAP has been used to calculate the flux and polarization signals of both terrestrial and gaseous exoplanets (Stam et al. 2006; Stam 2008; Karalidi et al. 2011, 2012; Karalidi & Stam 2012; Trees & Stam 2019; Groot et al. 2020).

DAP can model atmospheres of any composition, with as many layers as necessary to describe the full scattering properties of the atmosphere. However, increasing the number of layers results in an increase of the computation cost. Here we used 16 horizontally homogeneous and locally plane-parallel layers to make our calculations tractable, while also capturing the vertical variations of the atmosphere. Our models follow the T–P profile of a midlatitude Earth (McClatchey et al. 1972). Each layer contains H_2O, O_3, O_2, and N_2 molecules with the volume mixing ratio (VMR) profiles of McClatchey et al. (1972). To calculate the absorption properties of these gases, DAP uses the HITRAN 2020 molecular line lists (Gordon et al. 2022), and the k-coefficient method. Depending on the MODIS cloud p_c data, atmospheric layers in our models can also contain water cloud droplets of varying b_c.

DAP defines the flux vector of stellar light that has been reflected by a spherical planet with radius, r, at a distance, d, from the observer (where $d \gg r$) as

$$\pi \mathbf{F}(\lambda, \alpha) = \frac{1}{4} \frac{r^2}{d^2} \mathbf{S}(\lambda, \alpha) \pi \mathbf{F}_0(\lambda), \qquad (3)$$

where λ is the wavelength of the light and α is the planetary phase angle, i.e., the angle formed by the star–planet–observer (see top panel of Figure 2). $\pi \mathbf{F}_0$ is the flux vector of the incident starlight and $\mathbf{S}$ is the 4×4 planetary scattering matrix with elements a_{ij}, which is calculated by DAP (for more information, see Stam et al. 2006). For our calculations, we normalized Equation (3) assuming $r = 1$ and $d = 1$, and assumed unpolarized incident starlight (e.g., Kemp et al. 1987; Cotton et al. 2017) so that $\pi F_0 = 1$. The total flux reflected by the planet is thus simplified to

$$\pi F_n(\lambda, \alpha) = \frac{1}{4} a_{11}(\lambda, \alpha), \qquad (4)$$

where a_{11} is the (1, 1) element of the $\mathbf{S}$ matrix (see Stam 2008), and the subscript n on the flux indicates that it is now normalized. The normalized fluxes πF_n can be straightforwardly scaled for any planetary system using Equation (3), and inserting the correct values for r, d, and πF_0. P_s ($=-Q/I$) is independent of these values and thus does not require any scaling.

3.3. Versatile Software for Transfer of Atmospheric Radiation

VSTAR uses the discrete-ordinate method for the radiative-transfer calculations, and fully incorporates single and multiple scattering by atmospheric gases as well as aerosol and cloud particles. Similar to DAP, VSTAR models the reflected light and polarization phase curves of a planet, in addition to its emission and transmission spectra, for given atmosphere and surface properties. VSTAR utilizes the widely used and validated (see, e.g., Kopparla et al. 2018) vectorized linear discrete ordinate radiative-transfer (VLIDORT) polarized light solution (Spurr 2006) to solve the polarized, or vector, radiative-transfer equation:

$$\mu \frac{d\boldsymbol{I}_\nu(\tau, \mu, \phi)}{d\tau} = \boldsymbol{I}_\nu(\tau, \mu, \phi) - \boldsymbol{S}_\nu(\tau, \mu, \phi), \quad (5)$$

where $\boldsymbol{I}_\nu$ is the Stokes vector describing the polarized light at frequency ν (similar to the Stokes vector of Equation (1)), τ is the optical depth of the atmospheric layer, and (μ, ϕ) describe the direction of the light, where μ is the cosine of the zenith angle and ϕ the azimuthal angle. The source function $\boldsymbol{S}_\nu$ captures scattering of outside radiation into the beam, thermal emission from the planet, and direct illumination of the atmosphere by an external source. Similar to DAP, VSTAR assumes that this external source is also unpolarized.

VLIDORT replaces the integral that appears in the scattering term of $\boldsymbol{S}_\nu$ with a sum using Gaussian quadrature. VSTAR uses a double Gauss scheme in which a separate set of quadrature angles is used for the light beam traveling down through the atmosphere and for the light beam traveling back up through the atmosphere. Following the discrete-ordinate method, VSTAR limits the angular distribution of light beams to a discrete number and then sums up these different angular solutions to get the final solution. Increasing the number of quadrature angles (or streams) thus increases the accuracy of the angular representation of the models. However, more streams leads to longer computing time (for more detailed descriptions and discussions, see Bailey et al. 2018). In this study we used 32 streams for our VSTAR models. We confirmed through preliminary tests that no major improvements to the models were acquired with any higher number of streams.

Similar to DAP, VSTAR can model atmospheres of any composition with any number of atmospheric layers. Different atmospheric layers of the VSTAR models, like DAP, can contain water cloud droplets depending on the MODIS cloud data and different VMRs of the H_2O, O_3, O_2, and N_2 molecules. However, while DAP uses k-coefficients, VSTAR uses a line-by-line approach to calculate the absorption properties of these atmospheric gases. As we benchmark VSTAR and DAP against each other (see Section 5.2), all of the VSTAR models copy the same atmospheric structure and surface properties as those used in the DAP models. Finally, both DAP and VSTAR used the 2020 HITRAN molecular line lists for the atmospheric gas absorptions.

While DAP and VSTAR follow different approaches to solving the radiative-transfer equation, both codes calculate the resulting reflected light from a planet, and both approaches are historically well validated for vectorized treatments of light (e.g., Spurr 2008). Here we perform the first benchmark of these two approaches for exoplanet modeling against each other and against earthshine data.
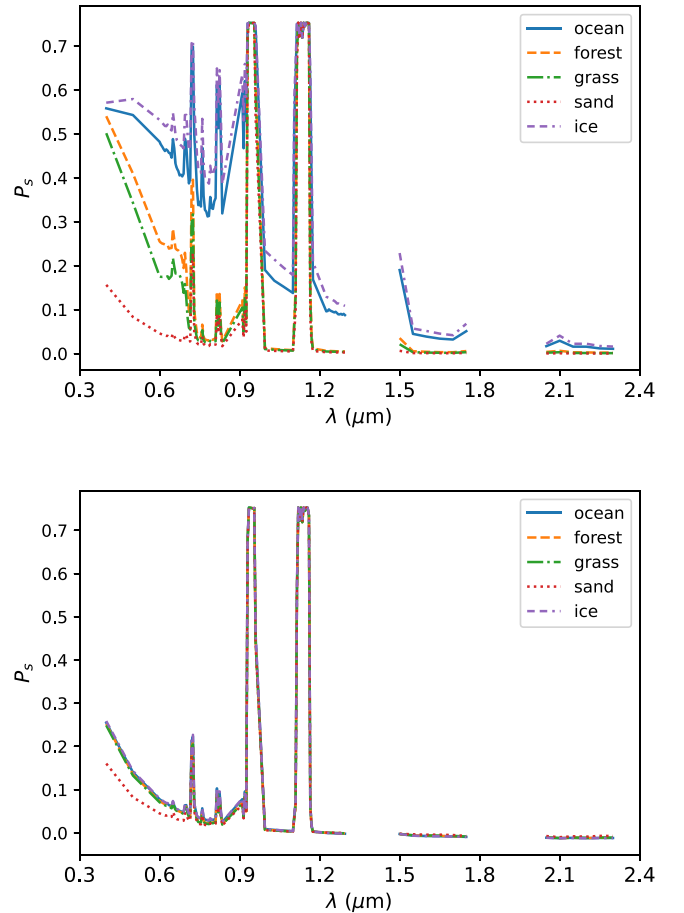

Figure 4. P_s as a function of λ for $\alpha = 70°$ of a planet with an ocean (solid blue line), forest (dashed orange line), grass (dashed–dotted green line), sand (dotted red line), or ice (dashed–dashed–dotted purple line) surface. Our model planets have clear atmospheres (top panel) or atmospheres with one water cloud with $b_c = 8$ and $p_c = 0.710$ bar, consisting of type B particles (bottom panel). The addition of clouds suppresses the surface effects on the resulting spectra.

4. Impact of Clouds

Clouds on Earth play a significant role in determining the overall polarization state of the planet. Interactions of the incident radiation with cloud particles depolarizes light due to multiple scattering within the clouds, thereby adding more unpolarized background light to the signal and lowering the resulting P_s. In this section we use DAP to assess the effects that different properties of water clouds play on the resulting P_s of the reflected light from an Earth-like planet. Unless otherwise stated, the models were generated at a variable resolving power, R, ranging from ~60 to ~260 in the 0.4–1.3 μm range and ~27 to ~46 in the 1.35–2.3 μm range, with $R \sim 250$ around 1.27 μm, similar to the resolving power of the binned earthshine data from Miles-Páez et al. (2014).

4.1. Effects of Cloud Optical Thickness

Figure 4 shows our model P_s as a function of λ for the five different surfaces we consider in our models, and for clear atmospheres (top panel) and cloudy atmospheres (bottom panel). Our cloudy models contain one liquid water cloud layer comprised of type B particles, with $b_c = 8$ and $p_c = 0.710$ bar. All atmospheres contain O_2 and H_2O vapor, which lead to the strong O_2 A band at 0.76 μm and the NIR H_2O bands near 0.9 and 1.1 μm. Our cloudless model spectra show clear features of
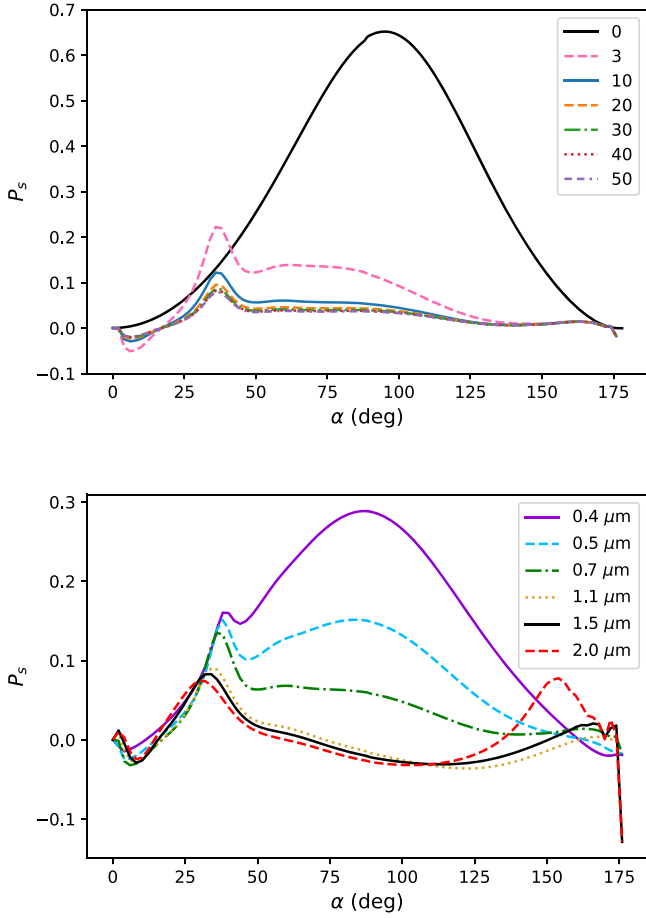
Figure 6. $P_s(\alpha)$ at $\lambda = 0.7\,\mu$m for the three different types of liquid water clouds tested with DAP, for a planet with a forest surface and a water cloud with $b_c = 10$ and $p_c = 0.710$ bar. The lines represent type B (dashed–dotted green line), type C (solid black line), and type D particles (dashed blue line). Larger particles result in stronger angular features and higher P_s across most phase angles.

layer to act like a completely depolarizing solid surface and make P_s insensitive to further increases in b_c. Any features detected in P_s for $b_c \gtrsim 30$ are due to single, or a few times, scattered light on the upper parts of our model clouds. Finally, the bottom panel of Figure 5 shows the wavelength dependence of the rainbow feature for the water clouds, with the location of the rainbow peak shifting to lower α with increasing λ (see, e.g., Bailey 2007; Karalidi et al. 2012).

4.2. Effects of Particle Size and Cloud Absorption

In Figure 6 we show P_s as a function of α, at $\lambda = 0.7\,\mu$m, for the three different particle-size distributions we explored. These models have a forest surface and a single cloud layer with $b_c = 10$ and $p_c = 0.710$ bar. Our model clouds are simulated using a two-parameter gamma distribution with $r_{\rm eff} = 6\,\mu$m, $u_{\rm eff} = 0.4$ (type B); $r_{\rm eff} = 10\,\mu$m, $u_{\rm eff} = 0.03$ (type C); and $r_{\rm eff} = 14\,\mu$m, $u_{\rm eff} = 0.04$ (type D) (see Section 2.2.1). Increasing $r_{\rm eff}$ results in higher $|P_s|$, with the larger particles producing stronger angular features than the smaller particles. In particular, the larger particles (types C and D) lead to a stronger and more pronounced primary rainbow than the smaller particles (type B). Additionally, C and D show a pronounced secondary rainbow feature around $\alpha = 60°$. The differences between the P_s of the three models at the largest phase angles are relatively small because there most of the reflected starlight has been scattered in the atmospheric layers above the clouds, and therefore the influence from the clouds themselves on the phase curves decreases.

The imaginary part of the complex refractive index of water n_w varies by ~ 5 orders of magnitude in the VNIR (see Section 2.2.1). To study the effect of our choice of Im$\{n_w\}$ on our model exoplanet Earth, we also modeled liquid water clouds with a complex refractive index of $1.335 \pm 0.0001i$ so that the imaginary part is one order of magnitude larger than our chosen value. The models in Figure 7 incorporate all pixels of our exo-Earth data grid, therefore including clouds with varying b_c and p_c above all five surfaces. These horizontally inhomogeneous spectra were calculated using the weighted-sum approximation (see, e.g., Stam 2008; Karalidi & Stam 2012). The flux vector of a planet covered by M different types of pixels is calculated as $\pi \boldsymbol{F}(\alpha) = \sum_{m=1}^{M} w_m \pi \boldsymbol{F}_m(\alpha)$,

Figure 5. Top panel: the phase dependence of P_s for different values of b_c at $\lambda = 0.7\,\mu$m for planets with an ocean surface and one water cloud layer composed of type B particles with $p_c = 0.710$ bar. The more optically thick the cloud layer, the lower the overall P_s. Bottom panel: the wavelength dependence of $P_s(\alpha)$ for liquid water clouds, for a planet with an ocean surface and one cloud layer with $b_c = 8$ and $p_c = 0.710$ bar. As expected, the peak of the rainbow feature ($\alpha \sim 25°$–$40°$) shifts to smaller phase angles at larger wavelengths while the strength of the Rayleigh peak ($\alpha \sim 90°$) diminishes.

the surfaces, such as the VRE for the forest and grass surfaces around $0.7\,\mu$m. As expected, the introduction of clouds in our model atmosphere reduces, or even erases, the impact of the surfaces on our spectra. The VRE that is easily detectable in the clear atmosphere models (top panel) is reduced in the cloudy models (bottom panel). Additionally, the surface albedos have much smaller effect in the cloudy models, as the spectra for all five surfaces have roughly the same P_s across all wavelengths. These changes are due to the clouds blocking the incoming light and increasing the amount of multiple scattering, thus lowering P_s.

For completeness, Figure 5 displays P_s as a function of α for cloudy models with $p_c = 0.710$ bar, b_c ranging from 0 to 50, and at $\lambda = 0.7\,\mu$m (top panel); and with $p_c = 0.710$ bar, $b_c = 8$, and wavelengths ranging from 0.4 to $2.0\,\mu$m (bottom panel). As expected, the clear atmosphere shows the characteristic bell-shaped curve due to Rayleigh scattering (top panel of Figure 5; solid black line). The addition of clouds decreases P_s across most phase angles and for most b_c values. The only exception is for the $b_c = 3$ model, where the primary rainbow of our liquid water clouds (for $\alpha \sim 30°$–$40°$) shows a stronger P_s than the Rayleigh scattering. The P_s of our models rapidly converges for $b_c \gtrsim 30$, since b_c larger than this cause the cloud
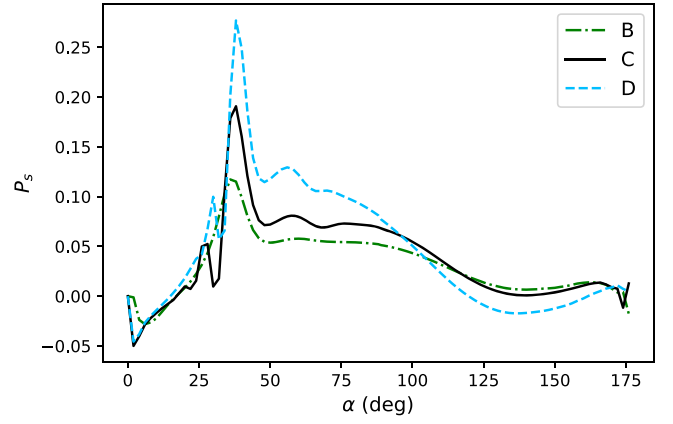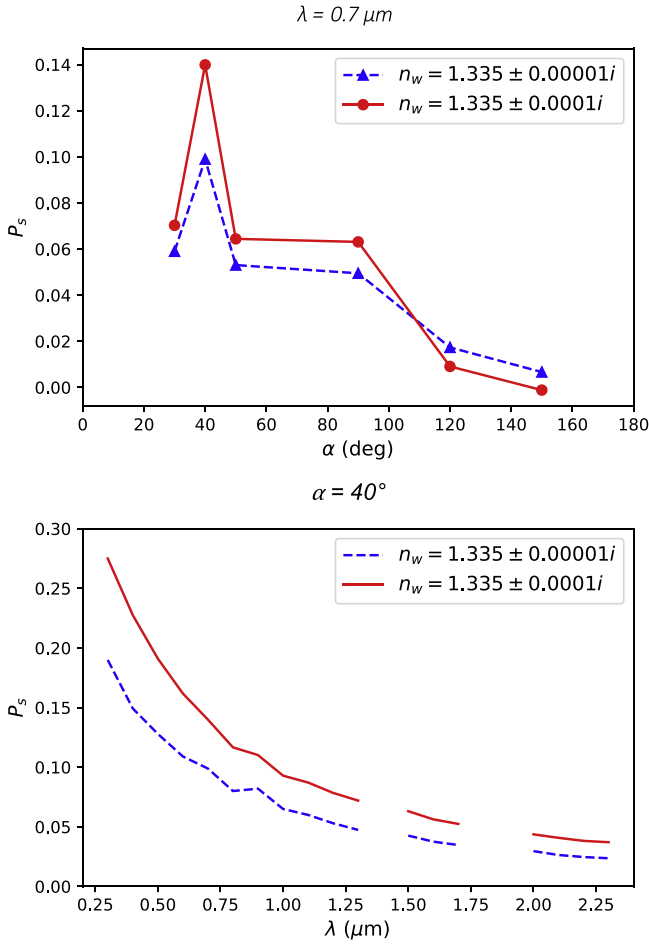
Figure 7. P_s as a function of α (top panel) and λ (bottom panel) for our model exoplanet Earth with water clouds having a constant, wavelength-independent n_w. The clouds have either a smaller ($\text{Im}\{n_w\} = 0.00001$; dashed blue lines) or a larger absorption ($\text{Im}\{n_w\} = 0.0001$; solid red lines). As expected, changing $\text{Im}\{n_w\}$ affects the total exoplanet $P_s(\alpha)$ and $P_s(\lambda)$.

where πF_m is the flux vector from a single m pixel and w_m is the fraction of type m pixels on the inhomogeneous planet, so that $\sum_{m=1}^{M} w_m = 1$.

Figure 7 (top panel) shows the P_s of our model exoplanet Earth as a function of α at $\lambda = 0.7\,\mu m$, for the two different values of n_w. To reduce computation time we only ran these models at six key phase angles. The angles were chosen to highlight the primary rainbow feature around $\alpha = 40°$, the planet at its furthest distance from its star at $\alpha = 90°$, and a few larger phase angles in order to approximate the full shape of the phase dependence. For most of the phase angles, the model with the larger imaginary part of n_w (red line) produced higher levels of polarization than the model with the smaller imaginary part of n_w (blue line), due to more absorption of the light through the water particles. This higher absorption leads to less multiple scattering of the light by the clouds and therefore less depolarization of the light. As we approach limb-viewing geometry at larger α, most of the reflected starlight gets scattered by the atmospheric gases in the layers above the clouds, and light that does end up reaching the cloud layers is more likely to be absorbed rather than scattered by the water droplets with the greater imaginary part of n_w. This leads to an overall lower level of polarization than the particles with smaller imaginary part of n_w.

Finally, the bottom panel of Figure 7 shows the $P_s(\lambda)$ for the two different n_w at $\alpha = 40°$, where the polarization of the liquid water clouds is strongest due to the primary rainbow feature. Here we are not interested in the spectral features of the atmosphere, rather just the change in the polarization continua due to the different $\text{Im}\{n_w\}$. Thus, these spectra are broad band, with a constant spectral resolution of $0.1\,\mu m$. As expected, the model with the larger $\text{Im}\{n_w\}$ (red line) shows higher P_s across the entire spectrum.

5. Model Comparisons

5.1. Correcting for Lunar Depolarization

To use earthshine observations as a benchmark for theoretical models we need to correct our observations for the lunar depolarization. Reflection off the Moon results in a depolarization of the polarized earthshine. This is due to the rough particulate regolith of the Moon, which induces high levels of back-scattering. Dollfus (1957) showed that the bright regions on the Moon depolarize the continuum of the incident light more strongly than the dark regions, due to the different porosity levels and particulate radii of the lunar soil in mare as compared to highlands and craters. Dollfus (1957) also noted a wavelength dependence of the depolarization for visible wavelengths, at a factor of approximately $3.3\lambda/550$ (λ, in nanometers).

In general terms, the lunar depolarization factor is defined as $\epsilon = P^{\text{out}}/P^{\text{in}}$, where P^{out} is the P of the light reflected by the Moon (i.e., the earthshine) and P^{in} is the P that is incident on the Moon (i.e., the true polarization of the Earth). To date no direct measurements of ϵ exist. Therefore, here we follow the discussion of Bazzon et al. (2013) and assume that

$$\log \epsilon(\lambda, a_{603}) = -0.61 \log a_{603} - 0.291 \log \lambda - 0.955, \quad (6)$$

where $\log a_{603}$ is the lunar albedo at $603\,\text{nm}$, and the wavelength $\log \lambda$ is measured in micrometers. This method was validated in previous studies of earthshine polarimetry by Sterzik et al. (2019, 2020). We acknowledge that this equation from Bazzon et al. (2013) was determined based on their observations at visible wavelengths only, but at this time we are not aware of any study providing a depolarization factor for the NIR.

Due to the small length of the slit at the NOT used for the earthshine observations, Miles-Páez et al. (2014) had to alternate exposures of the Moon and sky positions throughout their observing night. The optical part of their measured spectrum is thus an average from multiple regions of the lunar surface. Therefore, for our calculations of ϵ, we use an average lunar albedo of $a_{603} = 0.1359$ (based on measurements by Velikodsky et al. 2011). After calculating ϵ through Equation (6), the corrected, true polarization of the Earth is then calculated through $P^{\text{in}} = P^{\text{out}}/\epsilon$.

5.2. DAP versus VSTAR

This paper is part of a larger project aimed at comparing DAP and VSTAR. Here we compared the two codes to confirm that the results for our exoplanet-Earth model are independent of the code we used and thus the two different solutions to the radiative-transfer equation. In future work we will provide more in-depth comparisons, incorporating more surface–
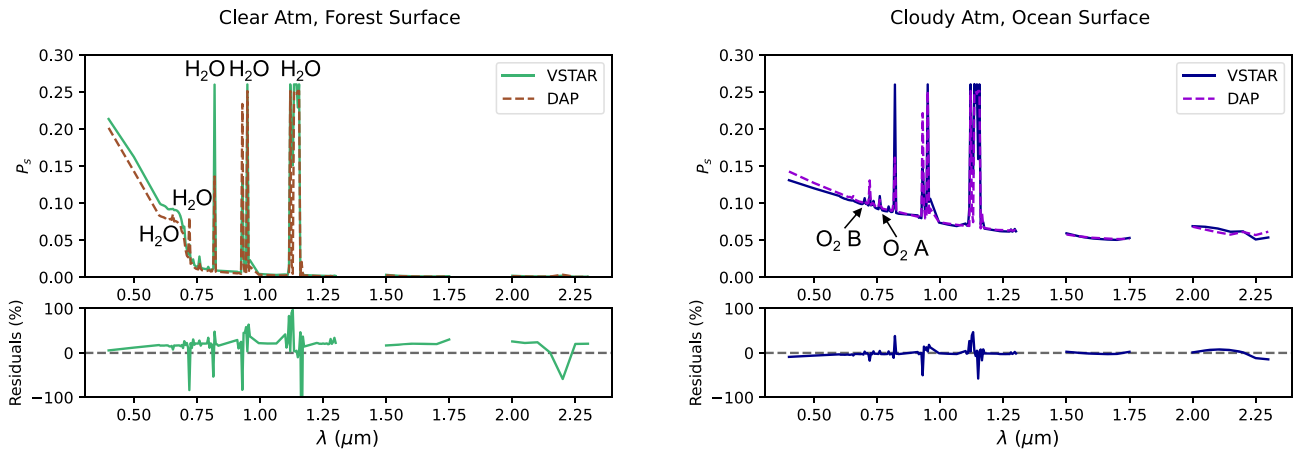
Figure 8. $P_s(\lambda)$ for VSTAR (solid lines) and DAP (dashed lines) models at $\alpha = 40°$. Our models have a forest surface and clear atmosphere (left panel) or an ocean surface and one water cloud layer with $b_c = 10$ and $p_c = 0.710$ bar, consisting of type B particles with a wavelength-dependent n_w (right panel). The major H_2O spectral features are labeled in the clear forest P_s plot (left panel), while the major O_2 spectral features are labeled in the cloudy ocean P_s plot (right plot). The differences seen in the absorption lines is due to the different approaches DAP and VSTAR use to model absorption.

atmosphere interactions and comparing to other planets in the solar system.

In Figure 8 we compare $P_s(\lambda)$ for DAP (dashed lines) and VSTAR (solid lines), at a phase angle of $\alpha = 40°$. The models were generated at constant spectral resolutions of 5 nm for $\lambda = 0.4$–$1.3\,\mu$m ($R \sim 250$ at $\lambda = 1.27\,\mu$m) and 50 nm for $\lambda = 1.3$–$2.3\,\mu$m ($R \sim 32$ at $\lambda = 1.6\,\mu$m), since we are only interested in the level of the continua and not the spectral features at the longest wavelengths. The left panel shows models for clear atmospheres above a forest surface. The right panel shows models with an ocean surface and atmospheres that contain one liquid water cloud layer with $b_c = 10$ and $p_c = 0.710$ bar, consisting of type B particles with the wavelength-dependent n_w of Hale & Querry (1973). Also shown are the residuals between the two codes, calculated as the relative percent change in the VSTAR spectra with respect to the DAP spectra.

The continua of the models for both cases match well across the entire spectral range, with the majority of differences appearing around the O_2 and H_2O absorption bands (see residuals). Both codes use the HITRAN 2020 molecular line lists (Gordon et al. 2022) to calculate the O_2 and H_2O absorption, but VSTAR uses line-by-line and DAP k-coefficients for these calculations. The small discrepancies in the peaks and widths of the absorption features between the models are due to the wavelength grid that the k-coefficients were made with, which lost some of the finer detail of the absorption lines compared to VSTAR's line-by-line approach. The larger residuals of the NIR continuum for the clear forest comparisons is due to the low polarization of these models at longer wavelengths, where slight differences in P_s leads to large relative percent change. However, the two models still match well across this continuum ($P_s = 0.38\%$ (0.32%) for VSTAR (DAP) at 1.03 μm; $P_s = 0.15\%$ (0.13%) for VSTAR (DAP) at 1.6 μm).

Figure 9 shows the residuals between the VSTAR and DAP models for four different α's ranging from 20° to 160°. All models now have an ocean surface and atmospheres that are either clear (left panel) or with one liquid water cloud layer, similar to Figure 8 (right panel). The VSTAR-DAP residuals exhibit a phase angle dependence. For the cloudy models (right panel), the minimum residuals at $\alpha = 40°$ (i.e., near the primary

rainbow feature) are due to the strong rainbow polarization of the liquid water clouds dominating P_s. The maximum residuals at $\alpha = 160°$ were also observed in Karalidi & Stam (2012) and, similar to that study, are most likely caused by the different disk-integration methods used in VSTAR and DAP. Finally, in both the clear and cloudy models, the greatest residuals occur in the NIR H_2O absorption bands (49% (clear) and 83% (cloudy) at 0.93 μm; 74% (clear) and −24% (cloudy) at 1.12 μm; versus 19% (clear) and −1.6% (cloudy) at 1.6 μm (the continuum) for the $\alpha = 80°$ comparisons) and are due to the wavelength grid with which the k-coefficients were created.

All VSTAR and DAP models compared here used the same (Lambertian) surface treatment and had the same atmospheric structure, including number of atmospheric layers as well as T–P, composition, and (when applicable) cloud profiles. Both codes followed the derivations of Peck & Reeder (1972) for Rayleigh scattering in dry air and modeled clouds using Mie theory. Finally, both codes utilized the HITRAN 2020 database for the molecular absorption, with VSTAR using line-by-line and DAP k-coefficient calculations. While some differences between the VSTAR and DAP models can be attributed to the different absorption calculation methods and disk-integration methods, a number of discrepancies exist that we cannot currently identify the source of. We will address these discrepancies in future work.

5.3. Earthshine versus Models

In this section we compare different models of our exoplanet Earth from both DAP and VSTAR, at a phase angle of $\alpha = 106°$, to each other and to both the original earthshine observations as well as the observations corrected for lunar depolarization.

In Figure 10 we compare the P_s of the exoplanet-Earth models generated by DAP and VSTAR. Due to computational time restrictions, the exoplanet-Earth VSTAR models (solid lines) represent a crude blending of four models: the two models shown in Figure 8; a model with an ocean surface and a clear atmosphere; and a model with a forest surface and the same cloud as in Figure 8. For pixels in the MODIS data with ocean or ice surfaces we used the VSTAR ocean models, while the VSTAR forest models were used for all other surface types. The cloudy VSTAR models were used for pixels containing
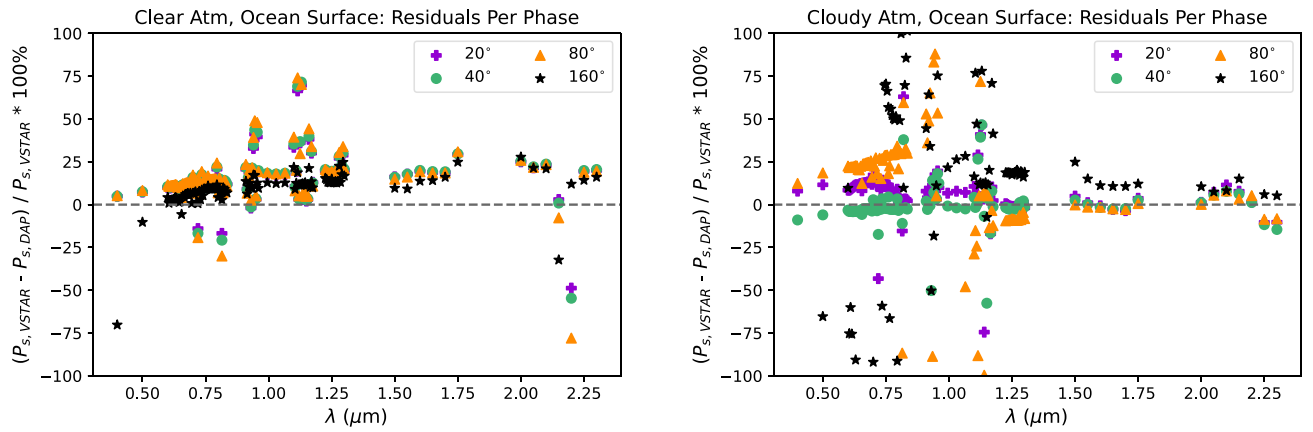
Figure 9. Residuals between the VSTAR and DAP models as a function of λ for α's of 20° (purple pluses), 40° (green circles), 80° (orange triangles), and 160° (black stars). All models have an ocean surface and either clear atmospheres (left panel) or atmospheres containing one liquid water cloud layer, similar to Figure 8 (right panel). The residuals show an α dependence, with the lower residuals noted at the primary rainbow feature (~40°).

clouds with $b_c \geqslant 3$ (and any p_c), and the clear VSTAR models were used for all other pixels. The DAP model was made by blending just the four models used in the VSTAR model (left panel, dashed red line), or all the surface and cloud models included in the MODIS data, but now for water cloud particles with a wavelength-independent n_w of $1.335 \pm 0.00001i$, due to computational time restrictions (right panel, dashed green line). This full exoplanet-Earth DAP model in the right panel was generated using the same weighted-sum approximation as described in Section 4.2. All models had spectral resolutions of 5 nm for $\lambda = 0.4$–1.3 μm and 50 nm for $\lambda = 1.3$–2.3 μm, resulting in $R \sim 250$ at $\lambda = 1.27$ μm. The residuals of the two comparisons are shown in the bottom of each panel, calculated as the relative percent change in the VSTAR spectra with respect to the DAP spectra.

The general shape and slope of $P_s(\lambda)$ match very well between the DAP and VSTAR models. The continua of the spectra overlap across most wavelengths, and the differences in the large H_2O spectral features are expected due to the different methods of handling absorptions between the two codes (line-by-line for VSTAR versus k-coefficients for DAP). The inclusion of extra surfaces in the full-blown DAP exoplanet-Earth model (right panel, dashed green line) leads to more differences between the DAP and the VSTAR models. In particular, the largest variations occur around the VRE ($\lambda \sim 0.7$ μm), with δP_s ($= P_{s,VSTAR} - P_{s,DAP}$) ~ 0.009 for the crude exoplanet-Earth model comparison (left panel, inset plot) and $\delta P_s \sim 0.012$ for the crude VSTAR model compared to the full DAP model (right panel, inset plot). This is because the crude VSTAR exoplanet-Earth model only includes forest surfaces for its land pixels, while the full-blown DAP exoplanet-Earth model replaces some of these vegetated surfaces with sand or ice surfaces depending on the specific pixel in the data grid, thereby slightly decreasing the strength of the VRE. The small differences at the longer wavelengths in the right panel can be attributed to the different n_w used for the liquid water clouds, which can lead to slight differences in the absorptive properties of the cloud particles and therefore alter the resulting P_s (see Section 4.2).

It is important to note that these spectra show P_s ($= -Q/I$), so they hold information about the direction of P of the reflected light (Section 3.1). For the given $\alpha_\oplus$ at the time of the observations, $P_{s,DAP}$ is low, with the continuum becoming negative at $\lambda \gtrsim 0.75$ μm, indicating that the direction of the

polarization has flipped to being parallel rather than perpendicular to the scattering plane. This result is in line with previous studies and is due to a combination of the physical and microphysical properties of the clouds used in our models (e.g., Stam 2008). Analyzing the angles of the polarization of the crude heterogeneous VSTAR model, we find a small flip in the angles at the same wavelength as the DAP model, with the VSTAR angles tending toward 0°. Finally, we note that for the VSTAR spectrum in Figure 10 we plot the ratio of Stokes Q to Stokes I ($-Q/I$) in order to provide a more direct comparison between the VSTAR and DAP models. This is due to the weighted-sum approximation used for our full exoplanet-Earth DAP model, which results in the Stokes U being zero. On the other hand, the output from VSTAR can be used to produce a nonzero Stokes U. In future papers we will use the method of Karalidi et al. (2012) to produce realistic heterogeneous models ($U \neq 0$) for DAP and compare against the full heterogeneous VSTAR models (taking into account all surfaces and cloud variations).

The effect of planet heterogeneity on Stokes U can be seen in Figure 11, where we plot the ratio of linearly polarized flux πU to total flux πI for different models generated by VSTAR. These include a homogeneous ocean planet with a cloud-free atmosphere (black line) along with four crude exoplanet-Earth models generated using different combinations of the simple VSTAR models from Figure 10: a mixed surface planet with forest and ocean surfaces and a cloud-free atmosphere (red line); a combination of clear and cloudy forest models (green line); a combination of clear and cloudy ocean models (blue line); and the crude VSTAR exoplanet-Earth model of Figure 10 that incorporates both surfaces and clear and cloudy models (gold line). All four heterogeneous cases are mapped to the full exoplanet-Earth data grid (Section 2.2) using a similar method as the crude VSTAR exoplanet-Earth model of Figure 10. As expected, the homogeneous case has $U/I \sim 0$ at all wavelengths due to the disk symmetry. The introduction of heterogeneity breaks this symmetry and results in $U/I \neq 0$, with the cloud heterogeneity (green and blue lines) affecting U/I more than the surface heterogeneity (red line). Note that the spectrum of the homogeneous ocean planet with the clear and cloudy atmospheres (blue line) is almost completely covered by the spectrum of the full-blown crude VSTAR exoplanet-Earth model (gold line). This is due to the strong
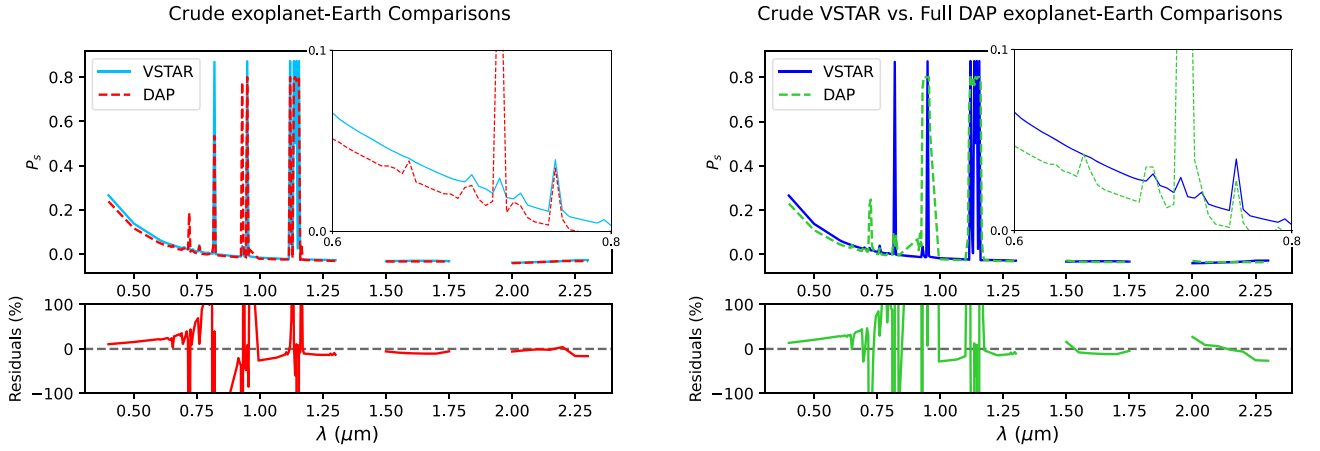
Figure 10. Comparison between blended models of DAP (dashed lines) and VSTAR (solid lines) for P_s as a function of λ. The left panel shows crude exoplanet-Earth models for both codes generated from blending four simple cloudy and clear ocean and forest models. The right panel displays the crude VSTAR exoplanet-Earth model compared against the full-blown DAP exoplanet-Earth model, which takes into account all five surface types and all clouds with type B particles (with a wavelength-independent n_w) and varying b_c and p_c. These models were created at $\alpha = 106°$ to match the geometry of the Earth during the earthshine observations. Residuals are shown for both comparisons. The inset plot of each panel shows a zoomed-in view of the spectra to highlight the main differences around the VRE (centered at $\lambda \sim 0.7$ μm).
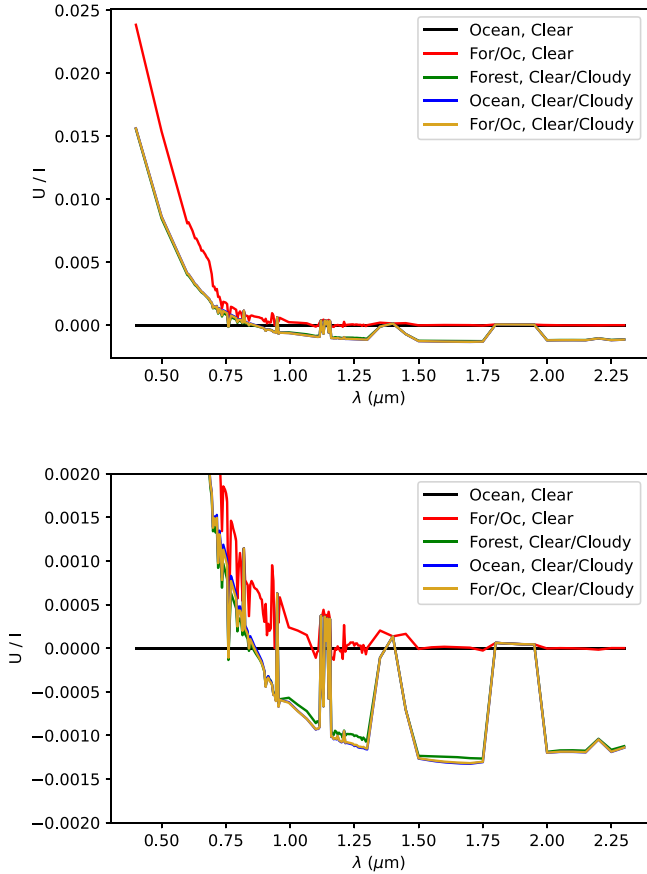

Figure 11. Comparisons between the ratios of Stokes U to Stokes I for a fully homogeneous planet (black line) and planets with different heterogeneities generated by VSTAR. All models are for $\alpha = 106°$. The homogeneous model is for a clear atmosphere above an ocean surface, whereas the heterogeneous models show the resulting U/I for crude exoplanet-Earth models created from different combinations of the four simple VSTAR models from Figure 10. The bottom panel shows the spectra with a zoomed-in view on the y-axis to better visualize their structures at longer wavelengths.

contribution of the Atlantic Ocean on our exoplanet-Earth data grid (see Figure 1), which has a stronger effect on the resulting polarization than the forest surface pixels.

Figure 12 displays comparisons between the exoplanet-Earth models generated by DAP (solid blue line) and VSTAR (dashed gold line) against both the original earthshine observations (dotted black line) and the observations corrected for lunar depolarization (dashed–dotted red line). The VSTAR model was generated from the same four simple models used to generate the crude VSTAR exoplanet-Earth model (Figure 10), but now incorporates Stokes U as well. Since $U = 0$ for the DAP data, the DAP model is simply taken to be the absolute value ($P_{DAP} = \sqrt{Q^2/I}$) of the DAP spectrum from the right panel of Figure 10 for the full-blown DAP exoplanet Earth. Again, the models had spectral resolutions of 5 nm for $\lambda = 0.4$–1.3 μm and 50 nm for $\lambda = 1.3$–2.3 μm, resulting in $R \sim 250$ at $\lambda = 1.27$ μm. The earthshine observations were originally measured at $R \sim 690$ in the NIR but the spectrum was binned down to increase the S/N, resulting in $R \sim 208$ at $\lambda = 1.27$ μm.

Comparing the DAP and VSTAR models in Figure 12, we see similar trends as those discussed in previous sections. The shape and slope of the models match each other well, with only slight differences due to the different treatments of the atmospheric absorptions between the two codes, as well as the inclusion of more surfaces in the DAP model. Additionally, the inclusion of Stokes U in the VSTAR spectrum may add to the slight increase in P in the shorter wavelengths compared to the DAP spectrum (see Figure 11). As expected, correcting the earthshine for lunar depolarization (Section 5.1) resulted in an increase of P by $\sim$4%–12% (depending on λ), with δP reaching $\sim$18% in the 1.12 μm NIR H_2O band.

Interestingly, the spectral slope of the earthshine observations differs from that of the models. While both models agree with one another, they both underestimate the polarization of the earthshine, especially after correcting for lunar depolarization. In an attempt to find a realistic value for the lunar albedo that would improve the fit between our models and the corrected earthshine, we tested different values for a_{603} (see Equation (6)). However, while lower values of a_{603} decreased the resulting P of the corrected earthshine, no value provided a match between the models and the observed spectra. We therefore opted to stick with the average value of
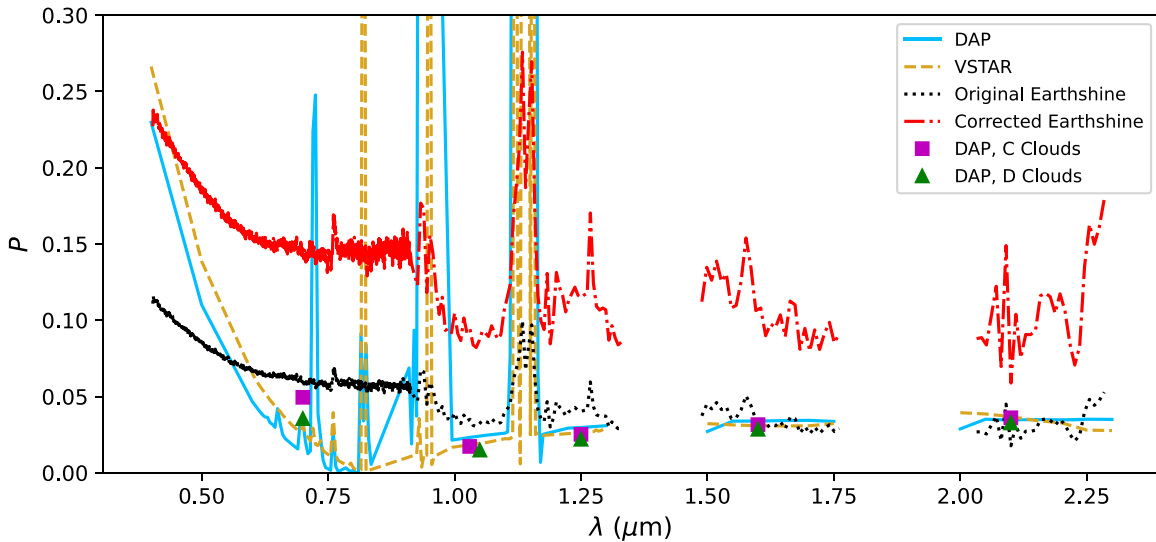
Figure 12. Comparison between the full-blown exoplanet-Earth model generated by DAP (solid blue line) and a crude exoplanet-Earth model created from the four simple VSTAR models from Figure 10 (dashed gold line) with the original earthshine observations (dotted black line) and the earthshine observations corrected for lunar depolarization (dashed–dotted red line). The DAP and VSTAR models are at $\alpha = 106°$ to match the observations. The DAP model includes all five surface types and clouds with varying b_c and p_c, while the VSTAR model includes only forest and ocean surfaces and clouds with a set b_c and p_c. The DAP and VSTAR model clouds are composed of type B particles. The VSTAR cloud particles used the wavelength-dependent n_w of Hale & Querry (1973), while the DAP cloud particles used a wavelength-independent n_w due to computational time restrictions. Also included in this plot are data points from two additional crude exoplanet-Earth models created by DAP, but now using the type C (purple squares) and type D (green triangles) cloud particles with the same wavelength-independent n_w as the type B clouds of the DAP model.

$a_{603} = 0.1359$. Both of the models and the observations show a Rayleigh-like slope at the shorter wavelengths, and the P of the models match the P of the corrected earthshine at ~ 0.4 μm. However, the earthshine observations are much flatter than the model spectra across all wavelengths.

We note that in both the original and corrected earthshine observations there is an offset in the level of P between the reddest wavelengths of the optical spectrum ($\lambda \gtrsim 0.8$ μm) and the NIR continuum. During their observations, Miles-Páez et al. (2014) collected the optical and NIR data simultaneously, but were unable to point the two telescopes at exactly the same region of the Moon. The offset is therefore not expected to be real, but instead a product of observing different lunar areas with different properties.

Changes in the spectral slope at shorter wavelengths are connected with the existence and properties of clouds in an atmosphere. Additionally, as discussed in Section 4.2, changes in the micro- and macro-physical properties of the clouds affect the resulting P in the NIR. The difference between our models and the observations indicate that our approach of using just a single cloud layer for a given pixel, with one type of model cloud (type B) across all pixels, is oversimplified and we would need a broader range of cloud properties to match the observations. Addressing this issue will be part of future work. Additionally, the flatter spectral slope of the earthshine at shorter wavelengths ($\sim \lambda^{-1.25}$) could possibly be attributed to haze, soot, and/or other aerosols in the Earth's atmosphere during the time of observation, which were not simulated in our models.

To test the effect of the adopted cloud particle size on the resulting spectra, we tested two additional crude exoplanet-Earth DAP models, similar to those in Figure 10, but now with type C (purple squares) or type D (green triangles) cloud particles (see Figure 12). Similar to the type B clouds used in the DAP model (solid blue line), the type C and type D cloud particles used a wavelength-independent n_w of

$1.335 \pm 0.00001i$ due to computational time restrictions. As discussed in Section 4.2, larger cloud particles lead to higher P at most phase angles. However, while we see some changes in the P between the three DAP models, at $\alpha = 106°$ neither of the models with larger cloud particles produced P close to that of the corrected earthshine. The differences between the models and observed spectra are therefore not solely due to the cloud particle-size distribution we adopted. Instead, these differences could be attributed to the fact that we assumed a single kind of liquid water clouds and did not take into account ice or mixed-phase clouds.

Water clouds on Earth show a variety of cloud particle sizes and their particle-size distribution, or even phase (liquid or ice), can vary with altitude (e.g., in cumulonimbus clouds). Additionally, liquid water clouds can be covered by water ice clouds (e.g., cirrus clouds), and the global coverage of these ice clouds varies by season and latitude. The MODIS data measured on the day of the earthshine observations show a global mean coverage of $\sim 43\%$ for ice clouds and at least $\sim 26\%$ for mixed-phase clouds. Even though ice clouds have been shown to not be important for the global P at some phase angles (e.g., Karalidi et al. 2012 showed that they cannot mask the liquid water rainbow feature at $\alpha = 40°$), at $\alpha = 106°$ their single scattering P is much larger than at $\alpha = 40°$ (see Karalidi et al. 2012, their Figure 3). This suggests that the effect of ice clouds on the global P is stronger at the observed α of the earthshine and could attribute to the discrepancy between the earthshine and our models.

Finally, the observed spectrum and the two model spectra differ considerably in the appearance of the 1.27 μm O_2 feature. While apparent in the observations, this feature does not appear in either model. The O_2 feature peaks at a wavelength of $\lambda \sim 1.2685$ μm, with an increase in the polarization of $\sim 1.8\%$ above the NIR continuum in the original observations and $\sim 5.2\%$ above the NIR continuum after correcting for lunar depolarization. We do not reproduce this O_2 feature at the R of

our models ($R \sim 250$ at $\lambda = 1.27\,\mu$m), which are slightly higher than the R of the binned earthshine ($R \sim 208$ at $\lambda = 1.27\,\mu$m). This presents an intriguing discrepancy between the models and observations, which we investigate further in the following section.

6. The 1.27 μm O$_2$ Feature

The recent launch of the JWST and the large number of terrestrial exoplanets we will be able to characterize in the coming years means that being able to characterize biosignatures in the atmospheres of these planets is becoming increasingly important. O$_2$ is among the best biosignatures due to its photosynthetic origins and relatively higher abundance in planetary atmospheres compared to other biosignatures such as H$_2$O (Meadows 2017). O$_2$ produces readily detectable spectral features across the VNIR wavelengths, including the 0.69 μm O$_2$ B band and the strong 0.76 μm O$_2$ A band, as well as the 1.06 and 1.27 μm NIR bands, which are caused by a combination of molecular O$_2$ absorption as well as O$_2$–O$_2$ collision-induced absorption. However, O$_2$ can also be formed abiotically and could therefore constitute a false-positive biosignature, making it highly important to understand the planetary environment when searching for signs of life (Meadows et al. 2018b). For example, Leung et al. (2020) showed that the detection of a strong 0.69 μm O$_2$ band complemented with a weaker or undetected 1.27 μm band could indicate the presence of an ocean-loss, high-O$_2$ atmosphere, suggesting that the observed planet is no longer habitable. Nevertheless, O$_2$ remains an important biosignature as it is difficult for terrestrial planets in the HZ to maintain O$_2$-rich atmospheres without the aid of life (Meadows et al. 2018b).

Here we analyze the 1.27 μm O$_2$ feature that was detected in the earthshine observations but was not reproduced in any of our heterogeneous Earth models (Section 5.3), and the implications this has for future observations. In particular, we test the effect that different resolving powers have on our models and do a parametric exploration of model parameters (T–P profiles and O$_2$ content of the atmosphere), in an attempt to match the observed O$_2$ feature in the earthshine spectrum from Miles-Páez et al. (2014).

6.1. The Effect of the Model R

Each spectral line has a finite width and depth that are dependent on the levels of line-broadening mechanisms in the planetary atmosphere (Petty 2006). To be able to distinguish a specific feature, our instruments and models need to have larger R than the feature itself, otherwise we risk not capturing it in observed or model spectra. Here we used DAP to determine at which R we lose any signs of the 1.27 μm O$_2$ feature in our model spectra. All models were run for a planet with a clear atmosphere above an ocean surface and for a phase angle of $\alpha = 106°$, to match the geometry of the Earth–Moon system on the night of the earthshine observations. The models cover wavelengths from 1.26 μm to 1.28 μm and have $R \sim 250$, 500, 690, and 1000 (at $\lambda = 1.27\,\mu$m). The $R \sim 250$ model is at the R of the models used in previous sections, whereas the $R \sim 690$ model represents the R of the original (unbinned) earthshine observations.

Figure 13 shows the πF (top panel) and P_s (bottom panel) of our model planets at the different R's. The P_s plot also includes
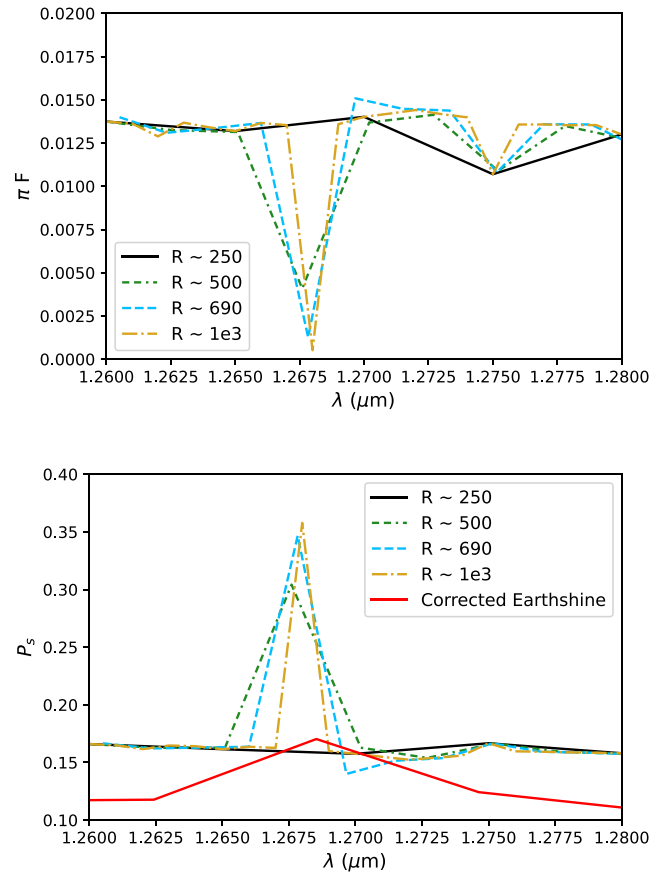

Figure 13. The wavelength dependence of model planets generated by DAP for varying R, centered around the 1.27 μm O$_2$ feature. All models are for a cloudless atmosphere above a homogeneous ocean surface, for $\alpha = 106°$. Top panel: πF as a function of λ. Bottom panel: P_s as a function of λ. The O$_2$ feature disappears in both πF and P_s at $R \sim 250$. Also included in the bottom panel is the 1.27 μm O$_2$ feature as seen in the binned earthshine observations ($R \sim 208$, binned down from $R \sim 690$) after correcting for lunar depolarization.

the 1.27 μm O$_2$ feature from the binned earthshine observations, corrected for the lunar depolarization (solid red line). Two distinct spectral lines of O$_2$ are visible in πF: one at $\lambda \sim 1.268\,\mu$m and one at $\lambda \sim 1.275\,\mu$m. As expected, decreasing the R of our models results in decreased intensity and broader widths in our model lines. In P_s, a detectable feature is seen near $\lambda = 1.27\,\mu$m only for the models with larger R. For the model with R comparable to the binned earthshine data ($R \sim 250$, solid black line), though, the feature is lost entirely due to the low number of sampled wavelengths, which causes the feature to blend with the surrounding continuum. To test that the chosen wavelength bins of the low-R model did not miss the core of the 1.27 μm O$_2$ feature, we ran additional low-R models in which we shifted the wavelength grid used to calculate the k-coefficient tables for our O$_2$ absorptions. We found that no tested wavelength grid produced a detectable 1.27 μm O$_2$ feature at $R \sim 250$. The results here are in line with previous studies analyzing low-R (e.g., Brandt & Spiegel 2014; Tremblay et al. 2020) and high-R (e.g., Lopez-Morales et al. 2019) simulated spectra of terrestrial exoplanets in flux. Our results show that the same holds true for polarization as well.

To test that our results are independent of our code and the inclusion of polarization in the radiative-transfer calculations, Figure 14 shows part of a simulated high-R reflectance spectrum of an Earth-like terrestrial planet orbiting Proxima
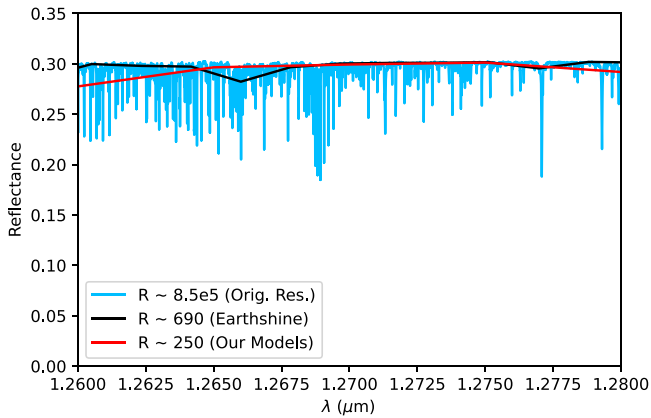
Figure 14. Simulated Earth-like atmospheric reflectance spectrum (blue line) for a planet orbiting Proxima Centauri, as generated by Leung et al. (2020), centered around the 1.27 μm O_2 feature. The original resolution of this spectrum had $R \sim 850,000$ at $\lambda \sim 1.27$ μm. We binned the model down to $R \sim 690$ (black line) to match the R of the original earthshine observations from Miles-Páez et al. (2014) and $R \sim 250$ (red line) to match the R of our models used throughout this paper. The O_2 spectral lines are greatly diminished in the $R \sim 690$ spectrum and disappear in the $R \sim 250$ spectrum.

Centauri from Leung et al. (2020). The original spectrum has $R \sim 850,000$ at $\lambda = 1.27$ μm. We binned the spectrum down to $R \sim 690$ and $R \sim 250$ to match the R of the unbinned earthshine spectrum as well as the R of our models from previous sections. While the 1.27 μm O_2 feature is detectable at $R \sim 850,000$, in the lower-R models we see only a very shallow and wide feature ($R \sim 690$) or no feature at all ($R \sim 250$), in agreement with our previous discussion.

6.2. Retrieving O_2 from Exoplanet Atmospheres

Even the highest-R models that produced an observable 1.27 μm O_2 feature did not match that of the earthshine observations corrected for lunar depolarization (see Figure 13). In the coming years telescopes like JWST and the Roman Space Telescope (RST) will provide us with NIR/IR spectra of multiple terrestrial exoplanets (note that JWST will only observe unpolarized light, but the RST will include a polarimeter). Recently, Fauchez et al. (2020) simulated the number of transits needed for a 5σ detection with JWST of the O_2 A band (at $R = 100$) for a modern Earth-like cloudy atmosphere on TRAPPIST-1 e orbiting a TRAPPIST-1-like star at several distances from Earth. They found that the strong cloud opacity of their model at short wavelengths hindered the signal from this band and required upwards of 1000 transits for an accurate detection with JWST. Therefore, unless complementary ground-based observations are taken we will not have full access to the O_2 A and B bands to accurately constrain the O_2 content of an atmosphere. With this in mind, we proceeded with scanning the parameter space of different O_2 abundances, surface conditions, cloud parameters, and T–P profiles to test what our retrieved planetary properties could be if we only had access to the 1.27 μm O_2 feature in polarization.

6.2.1. Scanning T–P Profiles and O_2 Atmospheric Contents

In the Earth atmosphere O_2 can be approximated as being evenly mixed, with a VMR of $\sim 21\%$. Our models so far used this constant VMR along with a midlatitude Earth T–P profile (McClatchey et al. 1972). However, the diversity of exoplanets observed to date suggests that different VMR profiles might
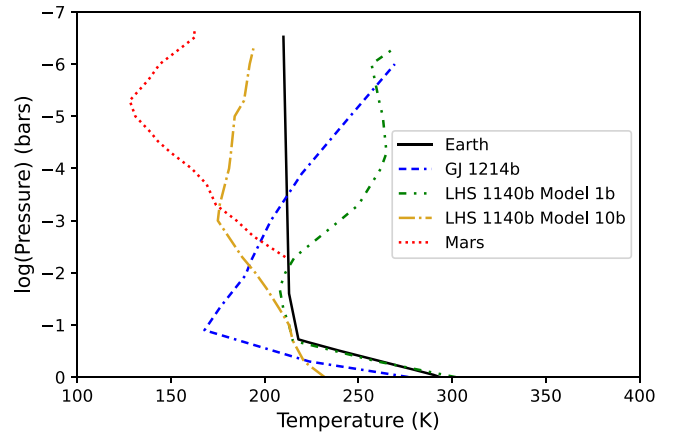

Figure 15. The T–P profiles of the five different planets modeled in this analysis, including the original T–P profile used in our exoplanet-Earth models from previous sections (solid black line).

exist in habitable planets. Planetary gravity g, T–P profiles, and VMR profiles are codependent in atmospheres. Therefore, we tested how using different values of g, T–P profiles, and VMRs in our models affect the fit of the modeled 1.27 μm O_2 feature with the polarized earthshine observations.

In particular, we used the observed profiles of Mars and model profiles from two of the best-studied super-Earths to date: GJ 1214b (Charbonneau et al. 2009) and the nearby temperate super-Earth LHS 1140b (Dittmann et al. 2017). Both of these exoplanets were discovered by the MEarth project survey, and due to their close distances to Earth serve as exceptional candidates for future follow-up observations and characterizations by both ground- and space-based telescopes. For GJ 1214b we utilized the T–P profile, VMRs, and gravity of Morley et al. (2015) for the model with $300\times$ solar metallicity and $0.01\times$ the planet's incident flux, as this model was most similar to that of Earth and allowed for the condensation of water clouds in the atmosphere. For LHS 1140b we utilized the T–P profiles, VMRs, and gravities for two of the models of Wunderlich et al. (2021), their Model 1b and their Model 10b. These models were chosen because they both have a constant CH_4 VMR of 1×10^{-3}, which studies have shown is expected for terrestrial planets in the HZ around mid-M dwarf stars (e.g., Wunderlich et al. 2019). Additionally, Model 1b has an H_2-dominated atmospheric composition similar to that of Neptune, while Model 10b has a CO_2-dominated atmospheric composition similar to that of Mars and Venus, thereby giving us a wide range of planetary atmospheric compositions to model (see Table 3 in Wunderlich et al. 2021, and references therein). The T–P profiles for our five scenarios are shown in Figure 15. For computational efficiency we limited the number of species used in our model super-Earth atmospheres compared to the original models. In particular, we used only N_2, O_3, CO_2, CH_4, He, H_2, O_2, and H_2O.

Previous observations and analyses of these super-Earths show that they could contain thick cloud decks, hazes, and/or potentially large levels of greenhouse effects (e.g., Kreidberg et al. 2014; Dittmann et al. 2017). However, clouds and hazes significantly lower the resulting polarization from a planet and require much longer computing times to model. For computational efficiency, and in order to acquire higher signals that are comparable to the earthshine observations, we ran all models
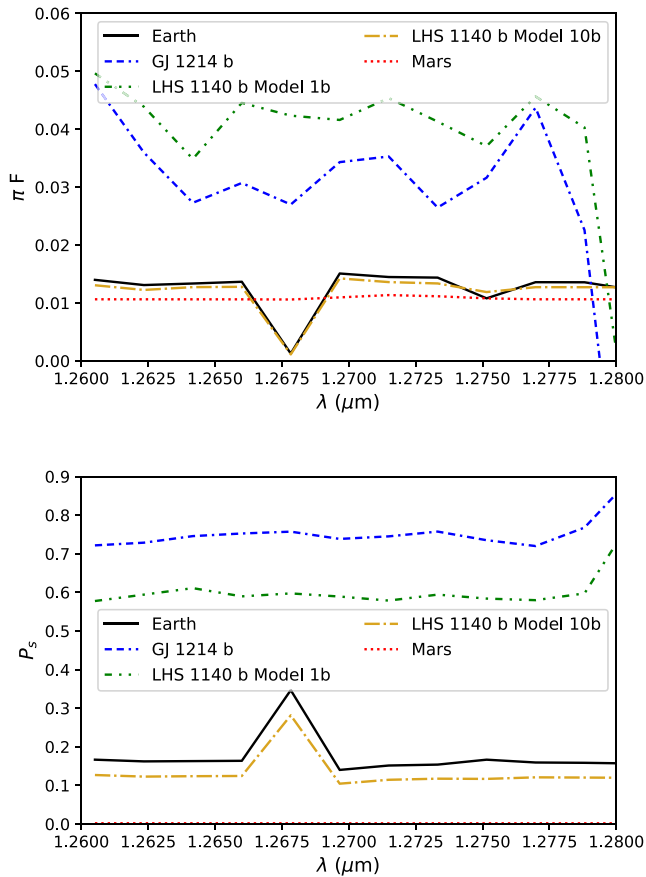
Figure 16. πF (top panel) and P_s (bottom panel) of models generated by DAP for the five different planetary T–P profiles, centered around the 1.27 μm O_2 feature. All models are for planets with a clear atmosphere above a homogeneous ocean surface and for $\alpha = 106°$. The models were generated at $R \sim 690$ at $\lambda = 1.27$ μm. Only the original Earth (solid black line) and the LHS 1140b Model 10b (dashed–dotted gold line) spectra show signs of the 1.27 μm O_2 feature in P_s.

Figure 17. πF (top panel) and P_s (bottom panel) of Earth models with varying levels of atmospheric O_2 content, centered around the 1.27 μm O_2 feature. All models have a clear atmosphere and a homogeneous ocean surface, and were ran at $\alpha = 106°$. The models have $R \sim 690$ at $\lambda = 1.27$ μm. As expected, increasing the amount of O_2 in the atmosphere leads to higher P_s in the 1.27 μm O_2 band.

for cloud-free atmospheres above an ocean surface, and for $\alpha = 106°$ to match the earthshine observations. The models cover wavelengths ranging from 1.26 to 1.28 μm, and were all ran at $R \sim 690$ at $\lambda = 1.27$ μm.

Figure 16 shows the πF (top panel) and P_s (bottom panel) of the reflected light off our model planets. The πF of the two models with H_2-dominated atmospheres, GJ 1214b (blue dashed–dotted lines) and LHS 1140b Model 1b (green dashed–dotted–dotted lines), are higher than those of the other three models. However, as these super-Earths have atmospheres more akin to Neptune than to Earth, their models have very small O_2 VMRs, resulting in spectra that are featureless in both πF and P_s. For the Mars model (red dotted lines), the resulting πF and P_s are featureless, as well, which is due to the small VMRs of O_2 and H_2O used in our Martian model. The low P_s ($P_s \sim 0$) is due to the fact that the Martian atmosphere is thin and our surface was treated as Lambertian. We note that the inclusion of a realistic surface as well as dust in the atmosphere would increase P_s to levels comparable to the observed Martian polarization (e.g., Stam et al. 2008). The LHS 1140b Model 10b (gold dashed–dotted lines) produces πF and P_s that are both similar in shape to those of the Earth model (black solid lines), but at slightly lower values. The appearance of a strong O_2 feature for these two models is most likely due to the higher O_2 VMRs used in these models when compared to
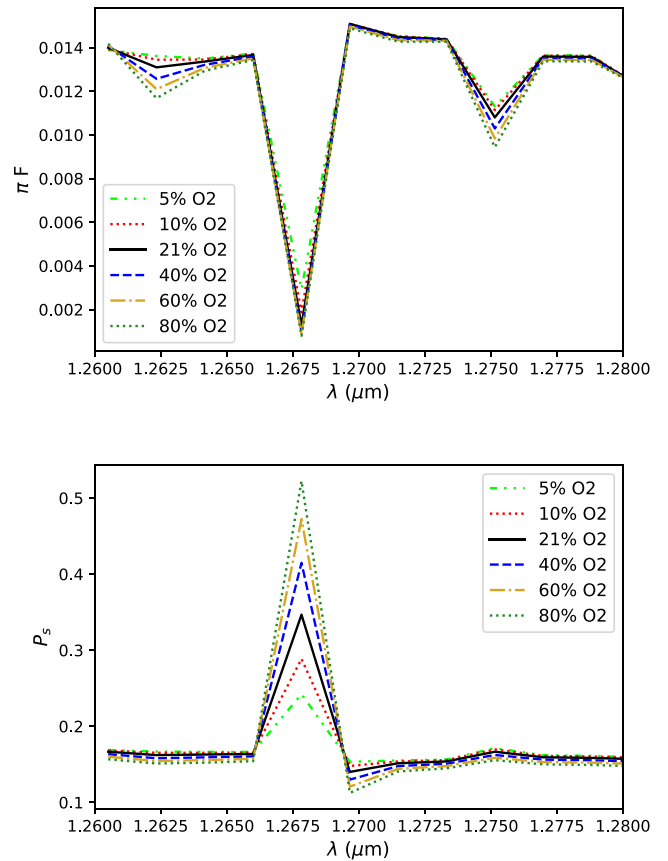
the other three: 21% for the Earth case and 30% for the LHS 1140b Model 10b case.

To verify this assumption, we ran additional Earth models with a range of O_2 VMRs from 5% to 80%. To ensure that the mass of the model Earth atmosphere remained the same across all models, any increase (decrease) in the O_2 VMR was balanced by a corresponding decrease (increase) in the N_2 VMR. These models have a clear atmosphere above a homogeneous ocean surface and cover wavelengths ranging from 1.26 to 1.28 μm, with $R \sim 690$ at $\lambda = 1.27$ μm. Figure 17 shows the πF (top panel) and P_s (bottom panel) of the reflected light for these models. For $\text{VMR}_{O_2} \gtrsim 10\%$ the center of the O_2 feature becomes saturated in πF, so any additional O_2 in the atmosphere results in a negligible increase in the depth of the line. On the other hand, there is a strong correlation between P_s and the O_2 VMR with P_s increasing by $\sim 28\%$ (absolute δP_s) between the 5% and the 80% models. This confirms our assumption that higher atmospheric O_2 concentrations lead to a larger and more pronounced 1.27 μm O_2 feature in the resulting P_s of our model planets.

6.2.2. Matching the Observed O_2 Feature

We performed a parametric scan using the different T–P profiles and varying the O_2 VMR between 5% and 80% for all model atmospheres in an attempt to match the 1.27 μm O_2 feature seen in the earthshine observations of Miles-Páez et al. (2014),
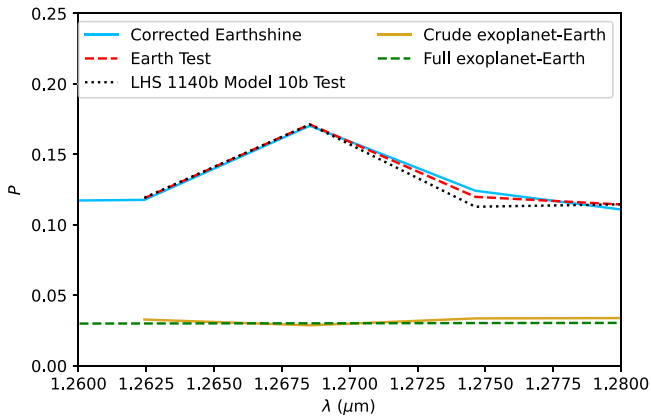
Figure 18. Comparison of the two best-fit models using the average Earth (dashed red line) and LHS 1140b Model 10b (dotted black line) *T*–*P* profiles to the earthshine observations corrected for lunar depolarization (solid blue line). The Earth model had a clear atmosphere with 10% O_2 and a heterogeneous surface covered by 70% ocean and 30% forest. The LHS 1140b Model 10b model had a clear atmosphere with 13% O_2 above a homogeneous ocean surface. The two models were both generated at $R \sim 690$ and then binned down to $R \sim 208$. Also included are the full-blown exoplanet-Earth DAP spectrum from Figure 12 (dashed green line) at $R \sim 250$ as well as a crude exoplanet-Earth DAP spectrum (solid gold line) generated from four high-*R* models. Both of these exoplanet-Earth models were also binned down to $R \sim 208$. All model spectra are for $\alpha = 106°$.

corrected for lunar depolarization. Figure 18 shows the two best-fit polarized spectra compared to the corrected earthshine observations (blue solid line). These best-fit spectra were generated at $R \sim 690$ and then binned down to $R \sim 208$ at $\lambda = 1.27 \mu m$. Figure 18 also shows the full DAP exoplanet-Earth spectrum from Figure 12 ($R \sim 250$ at $\lambda = 1.27 \mu m$, dashed green line), and a crude exoplanet-Earth spectrum (solid gold line) created from four high-resolution ($R \sim 1000$) DAP models and generated in the same way as the crude DAP model of Figure 10 (left panel).

The inclusion of any clouds in our models lowered the NIR continua, and misaligned them with respect to the corrected earthshine observations (see the exoplanet-Earth models in Figure 18). Additionally, as expected (e.g., Fauchez et al. 2017), the clouds significantly reduced the intensity of the O_2 feature. The best-fit model ($\chi^2 = 0.000365$ versus 0.818948 for the crude exoplanet-Earth model and 0.792116 for the full exoplanet-Earth model) that matched both the peak of the O_2 feature and the NIR continuum of the earthshine observations was of a planet with a clear, cloud-free atmosphere and an Earth-like *T*–*P* profile with 10% O_2 (red dashed line). The model surface was heterogeneous, with 70% ocean and 30% forest pixels. The second best-fit model ($\chi^2 = 0.001362$) was of a planet with a clear atmosphere and the *T*–*P* profile of LHS 1140b Model 10b, with 13% O_2 (black dotted line). The model surface was a homogeneous ocean. Both models have lower O_2 VMR than their respective physically consistent models. Higher O_2 abundances in the Earth and LHS 1140b Model 10b models caused the peak of the O_2 feature to rise further and mismatched the peak with respect to the corrected earthshine feature.

While we did not perform a full retrieval, rather only a limited parameter scan, these results are a cautionary tale for future observations of terrestrial exoplanets aimed at searching for biosignatures. We found that both an Earth-like planet and a super-Earth-like planet with low O_2 contents could produce

statistically similar polarized signals for the 1.27 μm O_2 feature as measured for the Earth. It will therefore be important that future studies make sure to analyze multiple bands of O_2 at the same time in order to truly characterize this biosignature, which could be achieved by combining ground- and space-based observations to cover the full VNIR wavelength range at high resolution.

7. Discussion and Conclusions

Polarimetry of exoplanets is an underutilized tool that will help the community to characterize exoplanets and break degeneracies that unpolarized light observations have. To prepare for the characterization of future polarimetric observations it is important that we benchmark our polarized radiative-transfer codes against each other and against observations of planets with known properties. Here we presented the first benchmark comparison of two polarized radiative-transfer codes, DAP and VSTAR, against each other and against polarimetric earthshine observations. Future work will build upon these comparisons by incorporating more surface and atmospheric parameters, including BPDFs, additional surface types, different cloud compositions, and additional cloud particle-size distributions.

The DAP and VSTAR models tested here showed good agreement with each other, both in the continua and in major biosignature features (O_2 and H_2O). However, a few discrepancies were noted between the models, which can mostly be attributed to the different methods of calculating the absorptions between the two codes (*k*-coefficient method for DAP versus line-by-line method for VSTAR). Additionally, we noted an interesting phase dependence in the VSTAR-DAP residuals, occurring in both clear atmosphere and cloudy atmosphere model comparisons. As all of the compared models had the same surface, atmosphere, Rayleigh scattering, and (when applicable) cloud treatments, we are unsure at this time as to the cause of these discrepancies. Addressing these differences is part of ongoing work.

Comparing our models against the earthshine observations of Miles-Páez et al. (2014), corrected for lunar depolarization, we noted that the models underestimate the earthshine *P*, while the earthshine also has a flatter spectral slope than the models predict. This suggests that the simplifications we used affected our resulting model spectra. In particular, our assumptions of a single cloud layer per pixel and the adoption of a single cloud parameterization across the planetary disk affected the overall shape of our spectra. Our model simplifications were based on common simplifications done for modeling exoplanets. These simplifications are necessitated from the fact that an observed exoplanet will be occupying a pixel or less in our images and any rotationally asymmetric features will only be resolved through time-resolved observations. The small signal of an exoplanet Earth would require long integration times, so that any information about small-scale variation across the disk would be lost in the observations. However, our results suggest that such simplifications can lead to large discrepancies between the data and our models. These discrepancies would have affected the retrieved properties of the exoplanet Earth if we were using our models to retrieve the properties of the Earth using the earthshine observations.

The effect of simplifying cloud properties on exoplanet spectra has already been noted for unpolarized observations of

giant planets and brown dwarfs (Luna & Morley 2021). For terrestrial planets, Feng et al. (2018) showed that incorporating the full cloud three-dimensional structure and variability is important for the proper constraint of the planetary parameters. Here we showed that the simplification of the cloud variability in an atmosphere also affects the polarimetric models of terrestrial planets. The effect of the inclusion of realistic variance in cloud parameters, such as varying optical thickness with height, inclusion of ice crystals or dust particulates, variable refractive index of water (n_w), and varying cloud particle effective radii across the Earth's disk and with altitude, will be part of future work. We note, however, that increasing the complexity of models will greatly increase the computing power and time needed to run models and perform retrievals (see also Feng et al. 2018), so understanding which assumptions can be made to the models and in which scenarios will be important.

In Section 6.2 we showed that our model R affected the fit of our models to the observed 1.27 μm O_2 feature. Models with $R \sim 250$ at $\lambda = 1.27$ μm, similar to the R of the binned earthshine observations, could not reproduce the O_2 feature at all, and only models with $R \gtrsim 500$ produced a detectable O_2 feature. Interestingly, even at higher R no realistic Earth-like model could reproduce the observed O_2 feature at 1.27 μm after correcting for lunar depolarization. We performed a parametric scan of a range of planetary properties (T–P profiles, O_2 VMR) to constrain the nature of the exoplanet Earth assuming that we knew nothing about the planet and only had access to the 1.27 μm O_2 feature.

We found that two models could fit the feature: an Earth-like model and a super-Earth-like model with clear atmospheres and lower O_2 VMRs than the Earth. Our results highlight the importance that the synergy between ground-based extremely large telescopes and space-based observatories (such as JWST and RST) will have for the full characterization of terrestrial exoplanets, even when polarimetric observations are available. Specifically, complementary observations of the O_2 A and B bands would hint at the nature of the 1.27 μm feature and aid in the confirmation that the observed O_2 features originate from an Earth-like planet.

While the 1.27 μm O_2 feature is an important biosignature and is easily detectable in Earth spectra, the large strength of this feature in the relatively low-R earthshine spectrum of Miles-Páez et al. (2014) is intriguing. The fact that this feature could not be matched by any of our Earth-like models suggests that the depth of this observed band could be anomalous. Observations from NASA's Solar Terrestrial Relations Observatory and Solar and Heliospheric Observatory detected an Earth-directed coronal mass ejection (CME) on the morning of 2013 May 17, a day and a half before the observations by Miles-Páez et al. (2014). CMEs are known to cause geomagnetic storms, which can deplete O_3 levels in the upper Earth atmosphere for up to several days after the event (e.g., Jackman et al. 2001). When O_3 is broken down by this UV radiation, it splits into molecular and atomic oxygen through photolysis. The CME event during the time of the earthshine observations could have caused rapid depletion of O_3, thereby increasing the amount of excited O_2 in the atmosphere and leading to a surplus of airglow on Earth. One of the strongest spectral signals of the Earth's airglow is the 1.27 μm O_2 band (Wayne 1994). This could have been picked up in the observations of the earthshine by Miles-Páez et al. (2014)

and caused the large 1.27 μm O_2 feature seen in their spectrum. If their observed 1.27 μm O_2 feature turns out to be an abnormal detection related to the CME event, this would have great implications for the detection of biomarkers in polarization in the presence of an active parent star.

A vital path forward for understanding the polarimetric spectra of Earth-like planets would be to acquire more observations of the earthshine across the VNIR wavelengths. To date, the observations of Miles-Páez et al. (2014) are the only earthshine measurements to fully extend polarized observations into the NIR. (Takahashi et al. 2021 also studied the NIR polarimetry of the earthshine but only for broadband measurements.) The observations of Miles-Páez et al. (2014), however, only covered one phase angle of the Earth and included a large O_2 feature at 1.27 μm that was unmatched by our models. While we couldn't reproduce the observed earthshine polarization from Miles-Páez et al. (2014), our models were able to reproduce the flux signal of the 1.27 μm O_2 feature from EPOXI observational data described in Livengood et al. (2011). This hints further to an abnormal 1.27 μm O_2 strength on the day of the earthshine observations. Capturing the earthshine at more phase angles across the full VNIR wavelengths would help to build upon the results of Miles-Páez et al. (2014) and produce the full polarized phase curve of the Earth. These additional observations would greatly improve the benchmarking of theoretical models against observational data, and allow us to test whether the strength of their measured 1.27 μm O_2 feature is a commonality or an abnormal detection (e.g., due to increased solar activity). This would help pave the way for future characterizations of the biosignatures of Earth-like exoplanets.

ORCID iDs

Kenneth E. Gordon ⬤ https://orcid.org/0000-0002-4258-6703
Theodora Karalidi ⬤ https://orcid.org/0000-0001-7356-6652
Kimberly M. Bott ⬤ https://orcid.org/0000-0002-4420-0560
Paulo A. Miles-Páez ⬤ https://orcid.org/0000-0003-2446-8882
Willeke Mulder ⬤ https://orcid.org/0000-0001-5762-9385

Daphne M. Stam ⬤ https://orcid.org/0000-0003-3697-2971

References

Bailey, J. 2007, AsBio, 7, 320
Bailey, J., Kedziora-Chudczer, L., & Bott, K. 2018, MNRAS, 480, 1613
Baldridge, A. M., Hook, S. J., Grove, C. I., & Rivera, G. 2009, RSEnv, 113, 711
Barstow, J. K., & Irwin, P. G. J. 2016, MNRAS: Lett., 461, L92
Bazzon, A., Schmid, H. M., & Gisler, D. 2013, A&A, 556, A117
Bétrémieux, Y., & Kaltenegger, L. 2014, ApJ, 791, 7
Brandt, T. D., & Spiegel, D. S. 2014, PNAS, 111, 13278
Bryson, S., Kunimoto, M., Kopparapu, R. K., et al. 2020, AJ, 161, 36
Charbonneau, D., Berta, Z. K., Irwin, J., et al. 2009, Natur, 462, 891
Cotton, D. V., Marshall, J. P., Bailey, J., et al. 2017, MNRAS, 467, 873
Daimon, M., & Masumura, A. 2007, ApOpt, 46, 3811
de Haan, J. F., Bosma, P., & Hovenier, J. 1987, A&A, 183, 371
de Rooij, W., & Van der Stap, C. 1984, A&A, 131, 237
Desgrange, C., Chauvin, G., Christiaens, V., et al. 2022, A&A, 664, A139
Dittmann, J. A., Irwin, J. M., Charbonneau, D., et al. 2017, Natur, 544, 333
Dollfus, A. 1957, SAnAp, 4, 3
Dollfus, A. 1975, Icar, 25, 416
Fauchez, T., Rossi, L., & Stam, D. M. 2017, ApJ, 842, 41
Fauchez, T. J., Turbet, M., Sergeev, D. E., et al. 2021, PSJ, 2, 106
Fauchez, T. J., Villanueva, G. L., Schwieterman, E. W., et al. 2020, NatAs, 4, 372
Feng, Y. K., Robinson, T. D., Fortney, J. J., et al. 2018, AJ, 155, 200
Gialluca, M. T., Robinson, T. D., Rugheimer, S., & Wunderlich, F. 2021, PASP, 133, 054401
Gordon, I. E., Rothman, L. S., Hargreaves, R., et al. 2022, JQSRT, 277, 107949
Greene, T. P., Line, M. R., Montero, C., et al. 2016, ApJ, 817, 17
Groot, A., Rossi, L., Trees, V., Cheung, J., & Stam, D. 2020, A&A, 640, A121
Hale, G. M., & Querry, M. R. 1973, ApOpt, 12, 555
Han, Q., Rossow, W. B., & Lacis, A. A. 1994, JCli, 7, 465
Hansen, J. E., & Hovenier, J. W. 1974, JAtS, 31, 1137
Hansen, J. E., & Travis, L. D. 1974, SSRv, 16, 527
Hovenier, J. 1970, A&A, 7, 86
Hovenier, J. W., Van der Mee, C. V., & Domke, H. 2004, Transfer of Polarized Light in Planetary Atmospheres: Basic Concepts and Practical Methods (Astrophysics and Space Science Library), Vol. 318 (Dordrecht: Kluwer)
Jackman, C. H., McPeters, R. D., Labow, G. J., et al. 2001, GeoRL, 28, 2883
Kaltenegger, L., & Traub, W. A. 2009, ApJ, 698, 519
Karalidi, T., & Stam, D. 2012, A&A, 546, A56
Karalidi, T., Stam, D., & Guirado, D. 2013, A&A, 555, A127
Karalidi, T., Stam, D., & Hovenier, J. 2011, A&A, 530, A69
Karalidi, T., Stam, D., & Hovenier, J. 2012, A&A, 548, A90
Kemp, J. C., Henson, G. D., Steiner, C. T., & Powell, E. R. 1987, Natur, 326, 270
King, M., Closs, J., Spangler, S., et al. 2004, EOS Data Products Handbook, Volume 1 (Greenbelt, MD: NASA Goddard Space Flight Center)
Kopparla, P., Natraj, V., Crisp, D., et al. 2018, AJ, 156, 143
Kreidberg, L., Bean, J. L., Désert, J.-M., et al. 2014, Natur, 505, 69
Leung, M., Meadows, V. S., & Lustig-Yaeger, J. 2020, AJ, 160, 11
Lincowski, A. P., Meadows, V. S., Crisp, D., et al. 2018, ApJ, 867, 76
Line, M. R., Knutson, H., Deming, D., Wilkins, A., & Desert, J.-M. 2013, ApJ, 778, 183
Livengood, T. A., Deming, L. D., A'Hearn, M. F., et al. 2011, AsBio, 11, 907
Lopez-Morales, M., Ben-Ami, S., Gonzalez-Abad, G., et al. 2019, AJ, 158, 24
Luna, J. L., & Morley, C. V. 2021, ApJ, 920, 146

Lustig-Yaeger, J., Meadows, V. S., & Lincowski, A. P. 2019, AJ, 158, 27
Marois, C., Macintosh, B., Barman, T., et al. 2008, Sci, 322, 1348
McClatchey, R. A., Fenn, R., Selby, J. E. A., et al. 1972, Optical Properties of the Atmosphere (Third Edition), AFCRL-72.0497 (Hanscom AFB: United States Air Force Cambridge Research Laboratories)
Meadows, V. S. 2017, AsBio, 17, 1022
Meadows, V. S., Arney, G. N., Schwieterman, E. W., et al. 2018a, AsBio, 18, 133
Meadows, V. S., Reinhard, C. T., Arney, G. N., et al. 2018b, AsBio, 18, 630
Meerdink, S. K., Hook, S. J., Roberts, D. A., & Abbott, E. A. 2019, RSEnv, 230, 111196
Miles-Páez, P., Pallé, E., & Osorio, M. Z. 2014, A&A, 562, L5
Misra, A., Meadows, V., & Crisp, D. 2014, ApJ, 792, 61
Morley, C. V., Fortney, J. J., Marley, M. S., et al. 2015, ApJ, 815, 110
Pallé, E., Osorio, M. R. Z., Barrena, R., Montañés-Rodríguez, P., & Martín, E. L. 2009, Natur, 459, 814
Pallé, E., Osorio, M. R. Z., & Muñoz, A. G. 2011, ApJ, 728, 19
Paradise, A., Macdonald, E., Menou, K., Lee, C., & Fan, B. L. 2022, MNRAS, 511, 3272
Peck, E. R., & Reeder, K. 1972, JOSA, 62, 958
Petty, G. W. 2006, A First Course in Atmospheric Radiation (Madison, WI: Sundog Publishing)
Pope, R. M., & Fry, E. S. 1997, ApOpt, 36, 8710
Robinson, T. 2018, in Handbook of Exoplanets, ed. H. J. Deeg & J. A. Belmonte (Berlin: Springer), 67
Rosenbush, V. K. 2002, Icar, 159, 145
Rossi, L., & Stam, D. M. 2018, A&A, 616, A117
Schmid, H. M., Joos, F., Buenzli, E., & Gisler, D. 2011, Icar, 212, 701
Seager, S. 2014, PNAS, 111, 12634
Segelstein, D. J. 1981, PhD thesis, Univ. Missouri-Kansas City
Sing, D. K., Fortney, J. J., Nikolov, N., et al. 2016, Natur, 529, 59
Spurr, R. 2008, Light Scattering Reviews 3 (Berlin: Springer), 229
Spurr, R. J. D. 2006, JQSRT, 102, 316
Stam, D. 2008, A&A, 482, 989
Stam, D., De Rooij, W., Cornet, G., & Hovenier, J. 2006, A&A, 452, 669
Stam, D., & Hovenier, J. 2005, A&A, 444, 275
Stam, D. M., Laan, E., Snik, F., et al. 2008, Third International Workshop on The Mars Atmosphere: Modeling and Observations, Vol. 1447 (Williamsburg, VA: LPI Contributions), 9078
Sterzik, M. F., Bagnulo, S., Emde, C., & Manev, M. 2020, A&A, 639, A89
Sterzik, M. F., Bagnulo, S., & Palle, E. 2012, Natur, 483, 64
Sterzik, M. F., Bagnulo, S., Stam, D. M., Emde, C., & Manev, M. 2019, A&A, 622, A41
Stolker, T., Min, M., Stam, D. M., et al. 2017, A&A, 607, A42
Suissa, G., Mandell, A. M., Wolf, E. T., et al. 2020, ApJ, 891, 58
Takahashi, J., Itoh, Y., Akitaya, H., et al. 2013, PASJ, 65, 38
Takahashi, J., Itoh, Y., Matsuo, T., et al. 2021, A&A, 653, A99
Traub, W. A., Oppenheimer, B. R., & Seager, S. 2010, Direct Imaging of Exoplanets (Tucson, AZ: Univ. Arizona Press)
Trees, V. J., & Stam, D. M. 2019, A&A, 626, A129
Trees, V. J. H., & Stam, D. M. 2022, A&A, 664, A172
Tremblay, L., Line, M. R., Stevenson, K., et al. 2020, AJ, 159, 117
Turnbull, M. C., Traub, W. A., Jucks, K. W., et al. 2006, ApJ, 644, 551
van Diedenhoven, B., Hasekamp, O. P., & Landgraf, J. 2007, JGRD, 112, D15208
Velikodsky, Y., Opanasenko, N. V., Akimov, L. A., et al. 2011, Icar, 214, 30
Wayne, R. P. 1994, Res. Chem. Intermed., 20, 395
Woolf, N. J., Smith, P. S., Traub, W. A., & Jucks, K. W. 2002, ApJ, 574, 430
Wunderlich, F., Godolt, M., Grenfell, J. L., et al. 2019, A&A, 624, A49
Wunderlich, F., Scheucher, M., Grenfell, J. L., et al. 2021, A&A, 647, A48