

Exo-Earths from the Inside Out: Characterizing Earth-like Exoplanets with Future Space Telescopes

Samantha Rose Gilbert-Janizek

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2026

Reading Committee:

Rory K. Barnes, Chair

Thomas Quinn

Joshua Krissansen-Totton

Jacob Lustig-Yaeger

Peter Driscoll

Program Authorized to Offer Degree:
Astronomy

©Copyright 2026

Samantha Rose Gilbert-Janizek

University of Washington

Abstract

Exo-Earths from the Inside Out: Characterizing Earth-like Exoplanets with Future Space Telescopes

Samantha Rose Gilbert-Janizek

Chair of the Supervisory Committee:

Rory K. Barnes

Astronomy

Future space-based telescopes will seek extraterrestrial life on Earth-like planets via transmission and direct imaging spectroscopy. NASA has prioritized the development of the Habitable Worlds Observatory (HWO), a flagship ultraviolet/optical/infrared direct imaging telescope to study Earth-like planets around Sun-like (FGK) stars, and has recommended the development of the Origins Space Telescope (Origins), a probe-class mid-infrared telescope capable of transmission spectroscopy to study Earth-like planets around M-dwarf stars. To assess habitability and biosignatures with observations from these future missions, we require (1) coupled interior-atmosphere models to predict observables and discriminants to rule out abiotic planets, and (2) robust retrieval tools to analyze observed spectra and infer surface and atmospheric properties with their uncertainties.

In the first part of this thesis work, we develop a novel “whole-planet” model to simulate the abiotic Earth through time by enhancing a coupled stellar-core-mantle-lithosphere thermal evolution model to include the carbonate-silicate cycle, the deep water cycle, a global surface water cycle, and climate with radiative transfer calculations. The upgraded model predicts surface temperature, surface albedo, CO₂ and H₂O abundances in the atmosphere and oceans, and ocean pH, as well as mantle and core properties described in previous work. We validate our model on measurements of key properties of the pre-industrial Earth, including but not limited to surface temperature, surface albedo, atmospheric and oceanic abundances of CO₂ and H₂O, ocean pH, upper mantle temperature, mantle viscosity, and

the inner radius of the core.

Next, we extend our abiotic Earth model to explore the time-evolving boundaries of the habitable zone (HZ) around representative F-, G-, and K-type stars. We produce HZ phase diagrams by upgrading our whole-planet model with a self-consistent treatment of fractional ice coverage, exploring two end-member weathering prescriptions that bracket the poorly constrained efficiency of abiotic silicate weathering. We find that the width, stability, and temporal evolution of the HZ depends sensitively on both stellar type and weathering assumptions. In both models, the outer HZ boundary contracts over the star’s main sequence lifetime as volcanic degassing declines. Our nominal model predicts limit cycle regimes in the outer HZ for all three stellar types, while our suppressed weathering model predicts either stable habitability or permanent snowball states. The K-dwarf case is notably distinct: the runaway greenhouse boundary remains fixed throughout the evolution owing to the star’s slow luminosity evolution, and outer HZ planets show a pronounced tendency toward permanent glaciation in both models. This is a consequence of degassing declining faster than stellar luminosity increases. Our results are broadly consistent with previous work but predict substantially narrower HZ widths. In contrast to previous work, for G- and K-stars we find a narrow range of planets that remain habitable for the duration of the star’s main sequence lifetime.

In the latter portion of the thesis, we use a novel retrieval tool to assess accuracy in characterizing near-surface habitability and biosignatures via simulated transmission and direct imaging spectra, based on the Origins and LUVOIR (HWO precursor) mission concepts. In transmission, we assess our ability to discriminate between an Earth-like and a false-positive O_3 TRAPPIST-1 e, and in reflected light, we assess the robustness of retrieval results to un-modeled cloud extinction. We find that Origins-like transmission observations may detect the CO_2/CH_4 pair on M-dwarf planets and differentiate between biological and false-positive O_3 using H_2O and abundant CO . In contrast, direct-imaging observations with an HWO-like observatory are better suited to constraining O_2 and O_3 , and may be sensitive

to wavelength-dependent water cloud features, but will struggle to detect modern-Earth-like abundances of methane. For direct imaging, we weakly detect a stratospheric ozone bulge by fitting the near-UV wings of the Hartley band.

Finally, we assess the spectral resolving power required to detect and characterize biosignature gases across Earth-through-time atmospheres using HWO. Combining analytical detectability calculations spanning resolutions $R=20\text{--}5000$ with atmospheric retrievals, we find that the nominal visible resolution $R_{\text{vis}} = 140$ is sufficient for detecting O_2 in Phanerozoic-like atmospheres, while indirect inference via O_3 (detectable at $R_{\text{UV}}\sim 7$) may provide the most efficient path for characterizing low- O_2 Proterozoic atmospheres. In the near-IR, we find that $R_{\text{NIR}} \geq 40$ is necessary to avoid degeneracies between CO_2 and CO that could produce false positive detections of abundant CO , and the nominal $R_{\text{NIR}} = 70$ is sufficient for characterizing all Earth-through-time cases.

This thesis work contributes to HWO mission development by anticipating potential challenges in interpreting habitability and biosignatures due to abiotic planets, assessing the observational advantages of different stellar types for exoplanet surveys, and providing actionable guidance for finalizing spectrometer requirements while maintaining technological feasibility for the search for life on exoplanets.

DEDICATION

For my husband Joe, my friends, and my family.

ACKNOWLEDGEMENTS

Hope is a discipline.

– Mariame Kaba

Newly-inducted doctoral students are commonly advised, “A PhD is a marathon, not a sprint.” The advice is just as much a warning as it is an admonition: gird your loins, this will not be easy. Yet, not all PhDs are created equal – some PhDs are ultramarathons, and still others are Ironman triathlons. In the latter case, in my experience, it is particularly important to have a strong support network. I am fortunate to have a large village behind me, which spans family, friends, and collaborators from UW and beyond.

To my advisors and mentors, Rory, Jake, Avi, Josh, Peter, and more: Thank you for seeing and treating me as a scientist. Thank you for investing time, care, (and frequently, cold hard cash) into training me. From where I stand, together you are all responsible for **saving** my PhD.

To my fellow Raising eSTEAM volunteers: Thank you for sharing my vision of a better future. I know I am leaving eSTEAM in capable hands, and I cannot wait to see how you continue to improve our efforts and broaden our impact.

To the children of Echo Glen: Thank you for trusting me to teach you and share my love of science. Onward.

To my office-mates through the years, Nicole, Debby, and V: Thank you for wiping my tears, waking me up from midday naps on the office couch, and always being amenable to a study break. Thank you for never taking down my Pitbull flag – that is, thank you for accepting and loving me as I am. *Dale*.

To my PhD cohort/lab-mates, Héctor and Jess: Thank you for welcoming me into the group, answering my late-night messages, and commiserating in the darkest of times. And of course, thank you for helping me take the UW Astronomy PhD students on the high seas on a pirate ship.

To my brothers, Bryan, Brett, and Joseph: Thank you for showing me everything that is cool and interesting in the world. Thank you for always watching out for me. I’ll forgive you for all the teasing, too.

To my Mom and Dad, Karen and Irwin Gilbert: Thank you for always encouraging my curiosity. From the day I was born, you prioritized my education, but you never pressured me to achieve. When I pushed myself too hard, you were there to offer perspective. You gave me the foundation I needed to follow my dreams with confidence.

To my husband, Joe: Thank you for always supporting me, particularly throughout this journey. You are always there to listen and lend a hand, and you believed in me when all seemed lost. None of this would have happened without you.

CONTENTS

Dedication	i
Acknowledgements	ii
Contents	iii
List of Figures	vi
1 Introduction	1
1 Earth as an Exoplanet	2
2 The Search for Earth-like Exoplanets	4
3 Next-Generation Observatory Capabilities	6
4 What Makes a Planet Habitable?	10
5 Modeling the Whole Planet	12
6 Biosignatures and False Positives	14
7 Atmospheric Retrieval for Exoplanet Characterization	18
8 Thesis Overview	21
2 Modeling the Abiotic Earth Around the Sun	23
1 Properties of the Pre-Industrial Earth	25
1.1 Atmosphere	26
1.2 Surface Properties	29
1.3 Interior	30
2 Methods	34
2.1 Model Assumptions	36
2.2 Core, Mantle, and Basal Magma Ocean	37
2.3 Carbon Cycle	39
2.4 Water Cycle	43
2.5 Atmospheric Constituents	45
2.6 Ocean Chemistry	47
2.7 Surface Albedo	51
2.8 Climate	53
2.9 Reflected Light Spectra	55
2.10 Initial Conditions	56
3 Results	58
4 Discussion	68
4.1 Spectroscopic Predictions for HWO Observations	69
4.2 CH ₄ as a Greenhouse Gas	69
4.3 Abiotic Marine Calcite Budget	70
4.4 Implications for the Gaia Hypothesis	72
4.5 Comparison to Earth Evolution	76

4.6	Future Work	79
5	Conclusions	85
6	Constants	87
7	BMO Radiogenic Heating Calculation	90
7.1	Derivation of Effective Core Heating Rate	91
7.2	Required Radiogenic Enrichment	93
3	Modeling the Abiotic Earth around HWO Target Stars	95
1	Methods	98
1.1	Dynamic Model Initialization	98
1.2	Continental and Seafloor Weathering Assumptions	99
1.3	Fractional Ice Coverage	101
1.4	Stellar Cases and HZ Parameter Sweeps	105
2	Results	107
2.1	Representative Climate Evolutions	107
2.2	Model (a)	112
2.3	Model (b) (Nominal)	115
2.4	Model (a) and (b) Comparison	119
3	Discussion	121
3.1	Comparison to Previous Work	122
3.2	Implications for the Abiotic vs. Biotic HZ	126
3.3	Future Work	127
4	Conclusion	128
4	Retrievals for Terrestrial Atmospheres	131
1	Methods	133
1.1	Simulated Exoplanet Spectra	134
1.2	Retrieval Model	137
1.3	Model Comparison	145
2	Experiments	146
2.1	Transmission	148
2.2	Direct Imaging	151
3	Results	153
3.1	Transmission	154
3.2	Direct Imaging	160
4	Discussion	166
4.1	Comparison to Previous Work	170
4.2	Comparison of Transmission and Direct Imaging Capabilities	172
4.3	Model Limitations and Future Work	190
5	Conclusion	195
5	Assessing the Spectral Resolution of HWO	215
1	Methods	218
1.1	Atmospheric Cases	218
1.2	<code>rfast</code> Forward Model	218
1.3	<code>pyEDITH</code> Exposure Time Calculator	219

CONTENTS

1.4	Analytic Relation for Calculating Detectability as a Function of Resolution	222
1.5	<code>rfast</code> Retrieval Model, Setup, & Cases	223
2	Results	227
2.1	Detectability as a Function of Resolution in the Visible and the Near-IR	228
2.2	Retrieval Cases	234
3	Discussion	239
3.1	The Benefits of High Resolution in the Visible Require Lower Detector Noise	240
3.2	The Resolution in the Near-IR Should be Greater than 20 to Avoid Retrieval Ambiguities and Enable CO ₂ Detections	242
3.3	Comparison to Previous Work	244
3.4	Future Work	246
4	Conclusion	250
6	Conclusions and Future Work	259
1	Summary of Key Findings	260
1.1	Whole-Planet Modeling and the Abiotic Earth Baseline	260
1.2	Stellar Type and Habitable Zone Evolution	261
1.3	Transmission vs. Direct Imaging for Habitability and Biosignature Assessment	262
1.4	Spectral Resolution Requirements for HWO	263
2	Future Directions	264
2.1	Expanding the Abiotic Planet Parameter Space	264
2.2	Estimating the HZs of HWO FGK Stars	265
2.3	Retrieval Model Development	266
2.4	Mission Trade Studies and Target Prioritization	266
3	Concluding Remarks	267
	Bibliography	269

LIST OF FIGURES

1.1	Confirmed exoplanets plotted by mass and orbital period, courtesy of the NASA Exoplanet Archive.	5
1.2	“Tier A” HWO targets identified by the preliminary HWO target list (Mamajek & Stapelfeldt, 2023).	8
1.3	Earth through time atmospheres and spectra (Krissansen-Totton et al., 2025).	15
2.1	A schematic of the key processes included in our abiotic Earth model.	34
2.2	A schematic showing the atmospheric sources and sinks that define the new volatile cycling processes in VPLanet’s ThermInt module, inspired by Figure 1 in Foley (2015).	35
2.3	Incoming solar flux evolution with time, where the dot at 4.5 Gyr represents the present-day value.	60
2.4	Evolving bulk properties of the abiotic Earth.	61
2.5	Evolving ocean properties of the abiotic Earth.	62
2.6	Evolving mantle properties of the abiotic Earth.	66
2.7	Evolving core properties of the abiotic Earth.	67
2.8	A comparison of reflected light spectra for a realistic, cloudy Earth and the VPLanet abiotic Earth after 4.5 Gyr.	68
2.9	Comparison of the VPLanet abiotic Earth model evolution against paleo-climate proxy reconstructions from the literature (following Krissansen-Totton et al. (2018a).)	77
3.1	Nested loop structure of the dynamic initialization scheme.	100
3.2	A comparison of $p\text{CO}_2$ scale factors for the continental and seafloor weathering relations of Foley (2015) explored in this study.	102
3.3	A 2-dimensional grid of ice growth over land and ocean and global surface temperature as a function of luminosity and planet-star separation.	104
3.4	Ice fraction-surface temperature curve for both land and ocean surface-types.	105
3.5	A comparison of the luminosity evolutions of an F, G, and K-dwarf star.	108
3.6	Carbon cycle flux and surface temperature evolution of an abiotic Earth at 1.0 AU from the Sun under three model assumptions.	111
3.7	Climate phase diagram for an abiotic Earth orbiting a $1.2 M_\odot$ F-type star under a suppressed weathering rate.	113
3.8	Climate phase diagram for an abiotic Earth orbiting a $1.0 M_\odot$ G-type star under a suppressed weathering rate.	114
3.9	Climate phase diagram for an abiotic Earth orbiting a $0.8 M_\odot$ K-dwarf star under a suppressed weathering rate.	115

3.10	Climate phase diagram for an abiotic Earth orbiting a $1.2 M_{\odot}$ F-type star. .	117
3.11	Climate phase diagram for an abiotic Earth orbiting a $1.0 M_{\odot}$ G-type star. .	118
3.12	Climate phase diagram for an abiotic Earth orbiting a $0.8 M_{\odot}$ K-dwarf star. .	120
4.1	A comparison of the evenly-mixed ($\mathcal{M}_{iso}$), vertically-resolved with 3 – 5 fixed pressure points ($\mathcal{M}_{3P}$, $\mathcal{M}_{5P}$), and vertically-resolved with 3 free pressure points ($\mathcal{M}_{3F}$) forward models for retrieving O_3 from a simulated transmission spectrum of a clear-sky, Earth-like TRAPPIST-1 e.	141
4.2	A comparison of the vertically-resolved temperature structure and composition of each atmospheric case addressed in this study.	148
4.3	A comparison of simulated transmission spectra for TRAPPIST-1 e modeled with a self-consistent, Earth-like atmosphere.	150
4.4	A comparison of the posterior distributions for the Earth-like and abiotic false positive O_3 TRAPPIST-1 e using the evenly-mixed forward model ($\mathcal{M}_{iso}$). .	155
4.5	A comparison of the vertically-resolved O_3 profiles retrieved by the $\mathcal{M}_{3F}$ forward model for the Earth-like and abiotic TRAPPIST-1 e cases.	158
4.6	A comparison of O_3 total column abundance distributions calculated for retrieval models with and without vertical resolution for both transmission cases. .	159
4.7	A comparison of the median spectral fits with $\pm 1\text{-}\sigma$, $3\text{-}\sigma$ envelopes for the Earth-like and abiotic false positive O_3 TRAPPIST-1 e.	160
4.8	A comparison of the posterior distributions for the clear-sky and cloudy Earth using the evenly-mixed forward model ($\mathcal{M}_{iso}$).	161
4.9	A comparison of the vertically-resolved O_3 profiles retrieved by the $\mathcal{M}_{3P}$ and $\mathcal{M}_{3F}$ forward model for the clear-sky Earth.	164
4.10	A comparison of O_3 total column abundance distributions calculated for retrieval models with and without vertical resolution for both the clear-sky and cloudy direct imaging cases.	165
4.11	A comparison of the median spectral fits with 1- and 3- σ posterior distributions shown as shaded envelopes for the clear-sky and cloudy Earth.	166
4.12	The results of the retrieval on the Earth-like TRAPPIST-1 e transmission spectrum with the evenly-mixed forward model, $\mathcal{M}_{iso}$	199
4.13	The results of the retrieval on the Earth-like TRAPPIST-1 e transmission spectrum with the three-fixed points forward model, $\mathcal{M}_{3P}$	200
4.14	The results of the retrieval on the Earth-like TRAPPIST-1 e transmission spectrum with the five-fixed points forward model, $\mathcal{M}_{5P}$	201
4.15	The results of the retrieval on the Earth-like TRAPPIST-1 e transmission spectrum with the three-free points forward model, $\mathcal{M}_{3F}$	202
4.16	The results of the retrieval on the abiotic, false-positive O_3 TRAPPIST-1 e transmission spectrum with the evenly-mixed forward model, $\mathcal{M}_{iso}$	203
4.17	The results of the retrieval on the abiotic, false-positive O_3 TRAPPIST-1 e transmission spectrum with the three-fixed points forward model, $\mathcal{M}_{3P}$	204
4.18	The results of the retrieval on the abiotic, false-positive O_3 TRAPPIST-1 e transmission spectrum with the five-fixed points forward model, $\mathcal{M}_{5P}$	205
4.19	The results of the retrieval on the on the abiotic, false-positive O_3 TRAPPIST-1 e transmission spectrum with the three-free points forward model, $\mathcal{M}_{3F}$. .	206

4.20	The results of the retrieval on the clear-sky Earth direct imaging spectrum with the evenly-mixed forward model, $\mathcal{M}_{iso}$.	207
4.21	The results of the retrieval on the clear-sky Earth direct imaging spectrum with the three-fixed points forward model, $\mathcal{M}_{3P}$.	208
4.22	The results of the retrieval on the clear-sky Earth direct imaging spectrum with the five-fixed points forward model, $\mathcal{M}_{5P}$.	209
4.23	The results of the retrieval on the clear-sky Earth direct imaging spectrum with the three-free points forward model, $\mathcal{M}_{3F}$.	210
4.24	The results of the retrieval on the 50% cloudy (25% stratocumulus, 25% cirrus, 50% clear) Earth direct imaging spectrum with the evenly-mixed forward model, $\mathcal{M}_{iso}$.	211
4.25	The results of the retrieval on the 50% cloudy Earth direct imaging spectrum with the three-fixed points forward model, $\mathcal{M}_{3P}$.	212
4.26	The results of the retrieval on the 50% cloudy Earth direct imaging spectrum with the five-fixed points forward model, $\mathcal{M}_{5P}$.	213
4.27	The results of the retrieval on the 50% cloudy Earth direct imaging spectrum with the three-free points forward model, $\mathcal{M}_{3F}$.	214
5.1	SNR as a function of wavelength for the visible and near-infrared resolution retrieval investigations using the exposure time calculator.	227
5.2	Exposure time required for gas detection as a function of resolution for key gases in corresponding band-passes for all Earth through time cases explored in this study, assuming the nominal dark current (3×10^{-5} e-/pixel/s) throughout.	230
5.3	Contour plots showing the relationship between dark current, resolution in the visible band-pass, and exposure time required to detect H_2O , O_2 , and O_3 for the high CH_4 Proterozoic Earth.	233
5.4	Contour plots showing the relationship between dark current, resolution in the visible band-pass, and exposure time required to detect H_2O , O_2 , and O_3 for the Phanerozoic Earth.	234
5.5	Retrieval results for the Proterozoic Earth with varying resolutions in the visible band-pass (NIR resolution is held constant at 70).	235
5.6	Retrieval results for the Phanerozoic Earth with varying resolutions in the visible band-pass (NIR resolution is held constant at 70).	235
5.7	Retrieval results for the Archean Earth with varying resolutions in the near-IR (VIS resolution is held constant at 140).	235
5.8	Retrieval results for the Phanerozoic Earth with varying resolutions in the near-IR (VIS resolution is held constant at 140).	236
5.9	A comparison of the marginal and covariance plots of the retrieved CO_2 and CO abundances for the Phanerozoic Earth, with $R_{\text{NIR}} = 20$ and the nominal $R_{\text{NIR}} = 70$.	237
5.10	Full corner plot for the high- CH_4 Proterozoic case, including the nominal spectral resolution ($R_{\text{Vis}} = 140$) and higher resolutions ($R_{\text{Vis}} = 200, 500$, and 1000) in the visible (VIS) wavelengths.	252

5.11	Full corner plot for the Phanerozoic case, including the nominal spectral resolution ($R_{\text{Vis}} = 140$) and higher resolutions ($R_{\text{Vis}} = 200, 500, \text{ and } 1000$) in the visible (VIS) wavelengths.	253
5.12	Full corner plot for the Late Archean case, including the nominal spectral resolution ($R_{\text{NIR}} = 70$) and lower resolutions ($R_{\text{Vis}} = 60, 50, 40, \text{ and } 20$) in the near-infrared (NIR) wavelengths.	254
5.13	Full corner plot for the Phanerozoic case, including the nominal spectral resolution ($R_{\text{NIR}} = 70$) and lower resolutions ($R_{\text{Vis}} = 60, 50, 40, \text{ and } 20$) in the near-infrared (NIR) wavelengths.	255
5.14	Cloud parameter retrieval diagnostics for the near-IR channel as a function of spectral resolution for the Phanerozoic and Late Archean atmospheric cases.	256
5.15	Cloud-top pressure retrieval diagnostics for the visible channel as a function of spectral resolution for the Phanerozoic and High-CH ₄ Proterozoic atmospheric cases.	257

CHAPTER 1

INTRODUCTION

*Before long, in the history of science as well as in the popular imagination, a second scene of discovery is likely to become as famous as the one in which, during a chilly night in the late fall of 1609, Galileo raised his telescope from the Venice lagoon toward the Moon. It occurred to him in that moment, he said, that all planets are alike. Three centuries later, another discovery reversed the proposition: the Earth is a planet **like no other!** We have to acknowledge that the symmetry is really too perfect: whereas the first scientist discovered how to shift away from the narrow view of the Grand Canal he had from his window toward the infinite universe, the second discovered how to shift from the infinite universe back to the narrow limits of the blue planet.*

– Bruno Latour, *Gaia*, a (finally secular) figure for nature

1	Earth as an Exoplanet	2
2	The Search for Earth-like Exoplanets	4
3	Next-Generation Observatory Capabilities	6
4	What Makes a Planet Habitable?	10
5	Modeling the Whole Planet	12
6	Biosignatures and False Positives	14
7	Atmospheric Retrieval for Exoplanet Characterization	18
8	Thesis Overview	21

1 Earth as an Exoplanet

The epigraph for this introduction ([Latour, 2017](#)) originated in a series of 2013 lectures given by the late Bruno Latour, a philosopher and anthropologist who studied the practice of modern science and its role in society. The “second scientist” Latour refers to in the excerpt is chemist and inventor James Lovelock, who alongside microbiologist Lynn Margulis introduced “Gaia theory” in 1974 ([Lovelock & Margulis, 1974](#)). In its original form, Gaia theory argued that life is an integral piece of the self-regulating Earth system – life is not solely at the mercy of evolving externalities like a brightening star or global temperature fluctuations, but a forceful agent in shaping its own environment. The emergent concept of the “biosphere” thus reclassifies life as an architect of Earth’s geochemical and atmospheric evolution. To Latour, Gaia theory upends the Galilean conception of a universe where physical laws apply equally and unfeelingly to all of its component bodies. Galileo’s scientific revolution challenged the geocentric orthodoxy of his time by suggesting that Earth is no different than any other body in our Solar System. By contrast Lovelock and Margulis, Latour argues, show how Earth’s biosphere makes it a singular planet given the apparent lack of life elsewhere in the Solar System.

Yet Lovelock’s earlier work suggests a paradoxical takeaway: in conceptualizing terrestrial life as a generalizable system, Lovelock provided the quantitative framework needed to systematically search for life elsewhere. Nearly a decade before formalizing Gaia theory, Lovelock and philosopher Dian Hitchcock ([Hitchcock & Lovelock, 1967](#)) proposed that life could be detected remotely by identifying thermodynamic disequilibria in a planet’s atmosphere. Specifically, they proposed that the thermodynamically unfavorable coexistence of atmospheric O_2 and CH_4 is smoking gun evidence for the presence of life on the surface. This insight was geocentric only in the sense that Earth remains our sole example of a living world; the underlying principle is planetary-scale thermodynamics, physical laws as universal as the laws of motion. Life, Lovelock and Hitchcock argued, is a systemic process that

maintains atmospheric chemical disequilibria. This disequilibrium is a signature that can be observed across interstellar distances by studying how light interacts with gases in far-off planetary atmospheres. Arguably, Lovelock, Margulis, and Hitchcock did not break from the Galilean universe at all – rather, they developed the formalism needed to quantify the presence of life, thereby laying the foundations for astrobiological inquiry decades before the first exoplanet was ever discovered.

This dual perspective on Earth as the only (currently) known living planet forms the underlying principle for the astrobiology thesis presented here: our planet is seemingly unique in harboring the only known biosphere, but it is also representative of the physical, geological, and chemical principles that would apply to life anywhere in the universe. To effectively search for life beyond Earth, the Earth itself becomes our laboratory. We experiment by treating the Earth as an exoplanet, holding it at a distance and imagining how we might infer its properties from remote spectroscopic observations. Understanding life as a quantifiable phenomenon and building the telescopes we hope will eventually remotely detect similar worlds is an ongoing effort in the field of exoplanet astrobiology. In this thesis, we aim to advance these efforts by treating Earth as an exoplanet in two ways. First, we build a whole-planet model that reproduces our own present-day Earth from the core to the atmosphere, and then apply this model to Sun-like stars to quantify the likelihood of observing habitable planets as they evolve in time. Second, we deploy atmospheric retrieval models to quantify the capability of future space-based missions to correctly identify and characterize living planets from light-years away. Thus, the investigations presented here span "exo"-Earths from the inside out.

In the remainder of this introduction, we briefly review the state of the search for life beyond Earth, including the development of next-generation observatories, defining planetary habitability, modeling the “whole” planet, biosignatures and their false positive imitators, and atmospheric retrieval methods for characterizing exoplanet atmospheres. Finally, we provide an overview of the structure of this thesis.

2 The Search for Earth-like Exoplanets

Earth is currently the only known inhabited planet in the universe. In fact, geological evidence suggests that it was habitable with liquid water oceans for most of its 4.54 Gyr history (Cameron et al., 2024). Owing to its long-term habitability, Earth serves as the template that guides our search for biospheres elsewhere. Thus, the search for life begins with a search for Earth-like exoplanets, defined as rocky planets with approximately Earth’s mass and radius (Wolf et al., 2017) and active interior-atmosphere volatile exchange (Rodler & López-Morales, 2014; Fan et al., 2019). This broad definition likely includes planets with diverse compositions (Shim, 2014; Santerne et al., 2018), atmospheric structures (Vladilo et al., 2013), stellar hosts (Mamajek & Stapelfeldt, 2023), instellations (Kopparapu et al., 2013), and even habitable planets that never developed life (Cockell, 2014; Wordsworth & Pierrehumbert, 2014; López-Morales et al., 2019; Krissansen-Totton et al., 2021).

Our ability to detect Earth-like exoplanets around Sun-like (FGK) stars is limited by our current observational capabilities. Nonetheless, the discovery of a growing number of rocky planets (Gillon et al., 2016, 2017; Anglada-Escudé et al., 2016; Dittmann et al., 2017) has provided several plausible targets for atmospheric characterization. Figure 1.1 shows a scatter plot of the current population of confirmed exoplanets discovered via the transit, radial velocity, and direct imaging methods as a function of mass and orbital period (Christiansen et al., 2025). The transit detection method is most sensitive to larger planets orbiting closer to smaller, M-dwarf stars, resulting in the bulk of transiting planets having mass $M > M_{\oplus}$ (Kipping & Sanford, 2016) and short orbital periods (Beatty & Gaudi, 2008) (Figure 1.1). Of the detection methods shown in Figure 1.1, the transit method has thus far been most successful at identifying Earth-mass, rocky planets near our Solar neighborhood, such as the TRAPPIST-1 system (Gillon et al., 2016). The radial velocity method probes the gravitational effect of the planet on its host star, resulting in a similar bias towards more massive planets orbiting closer to the central star. The direct imaging detection technique blocks the

light of the star to observe the light reflecting off of its dimmer planets. As a result, current direct imaging efforts are limited by our ability to resolve low-contrast planets from their host stars – the current population of confirmed exoplanets detected by direct imaging are largely massive, outer gas giants that are hot enough to emit strongly in the infrared (e.g., [Marois et al., 2008, 2010; Macintosh et al., 2015](#)).

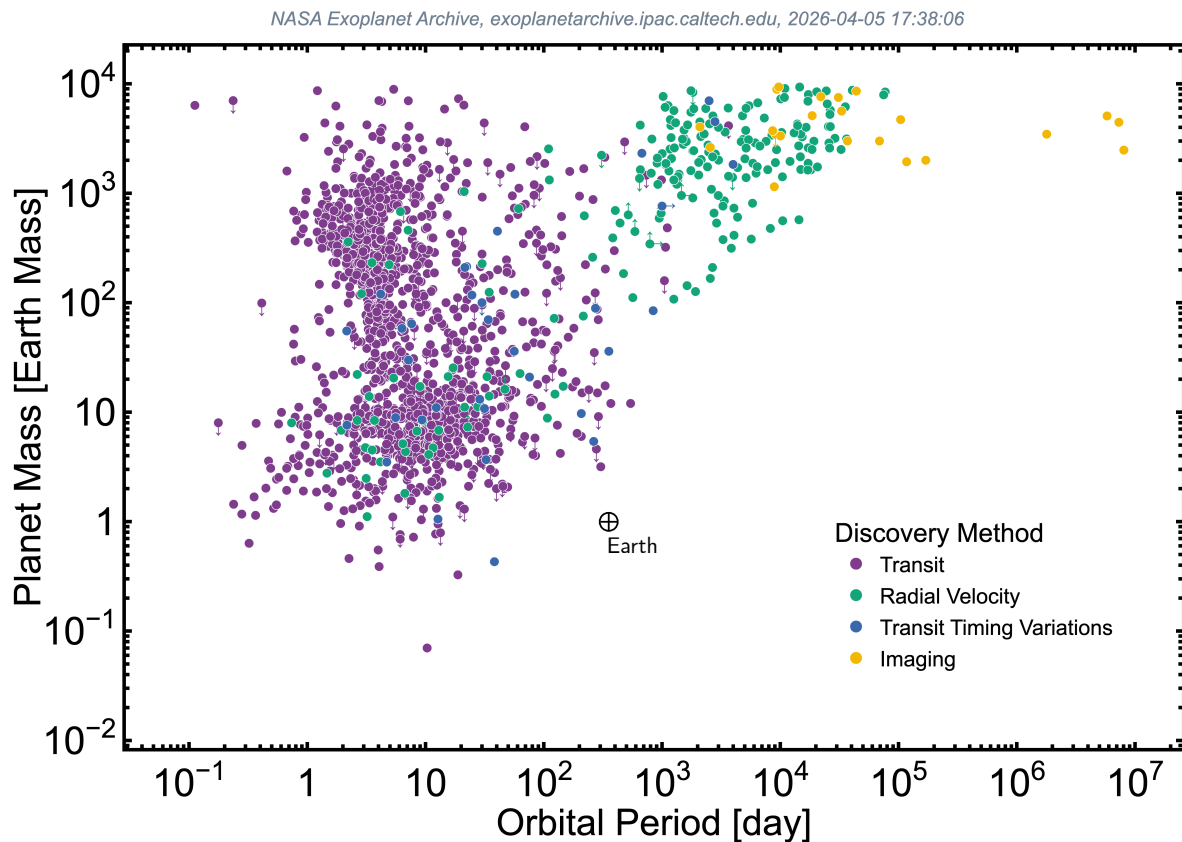


Figure 1.1: Confirmed exoplanets plotted by mass and orbital period, courtesy of the NASA Exoplanet Archive.¹ Colors denote the method of detection, and Earth ($1M_{\oplus}$, 365-day period) is marked with the circled plus sign.

To detect rocky, Earth-size planets around Sun-like stars and characterize their atmospheres to search for life, we require more powerful telescopes than are currently available. The next generation of space-based observatories, including NASA’s Habitable Worlds Observatory (HWO), are being designed precisely for this purpose. HWO will combine high-

¹Operated by the California Institute of Technology, under contract with the National Aeronautics and Space Administration under the Exoplanet Exploration Program.

contrast direct imaging with reflected light spectroscopy to probe the atmospheres of Earth-like planets around nearby FGK stars ([Feinberg et al., 2026](#)). However, interpreting observations from HWO and other future missions requires theoretical groundwork: self-consistent models of Earth-like planet evolution over billion-year (Gyr) timescales, an understanding of how HZ boundaries shift as stars age, and a clear picture of what a habitable-but-uninhabited planet looks like spectrally so that we can recognize the signatures of life when we find them. Beyond forward modeling, we must also ask what future observations can actually tell us: which atmospheric properties can be retrieved from reflected light or transmission spectra; how well biosignature abundances can be constrained given realistic noise and instrument performance; and how reliably biosignatures and habitability indicators can be detected with transmission versus direct imaging. The spectral resolution and wavelength coverage of future instruments will further determine which biosignatures are detectable and at what significance. This thesis develops a theoretical framework to help address each of these challenges. In the following section, we introduce next-generation observatories, including HWO, that will be capable of detecting and characterizing Earth-like exoplanets in the HZs of Sun-like stars.

3 Next-Generation Observatory Capabilities

Upcoming astrophysics missions will be designed to detect and characterize rocky, Earth-size planets and search for life, observing these planets in transmission, emission, or via reflected light spectroscopy (direct imaging) ([National Academies of Sciences, Engineering, and Medicine, 2021](#)). In the short-term, we will have access to transmission spectra of M-dwarf terrestrial atmospheres from the recently-launched JWST ([Batalha et al., 2018](#); [Wunderlich et al., 2019](#); [Lustig-Yaeger et al., 2019c](#)), and from the next generation of ground-based telescopes, which will see first light later this decade ([Guyon et al., 2012](#); [Snellen et al., 2013](#); [Rodler & López-Morales, 2014](#)).

In the coming decades, NASA has prioritized the development of HWO, its next flagship mission. HWO is planned to be a ~ 6 -meter space-based ultraviolet (UV), optical, and near-infrared (NIR) direct imaging telescope that will detect and characterize the atmospheres of nearby Earth-like exoplanets around FGK stars (Feinberg et al., 2026; National Academies of Sciences, Engineering, and Medicine, 2021; Mamajek & Stapelfeldt, 2023). A major priority of HWO will be to use reflected light spectroscopy to identify potentially habitable and inhabited planets. The Decadal Survey also identified a probe-class far-IR telescope (the Origins Space Telescope, *Origins*), which may be capable of transmission spectroscopy for characterizing Earth-like planets around M-dwarf stars (National Academies of Sciences, Engineering, and Medicine, 2021). A concept for a large space-based MIR interferometric mission to obtain spectral emission from habitable zone planets and search for signs of life is also being developed (Quanz et al., 2021).

Transmission and direct imaging are complementary, in that each technique is suited to a different class of host star. Transmission is primed to study the atmospheres of habitable zone planets in compact systems orbiting cooler M-dwarf stars, such as TRAPPIST-1 e, f, and g (Gillon et al., 2016), due to the high planet-star radius ratio and frequent transits of these planets (Suisa et al., 2020). Direct imaging will target systems with sufficient planet-star separation to study the atmospheres of Earth-like planets around more Sun-like, FGK stars (LUVOIR Team, 2019; Gaudi et al., 2020; Feinberg et al., 2026).

Nearby stars amenable to direct imaging with HWO have been recently identified, resulting in a tiered target list for community use (Mamajek & Stapelfeldt, 2023; Tuchow et al., 2025). For example, in one preliminary target list, “Tier A” stars are identified as having an angular separation that exceeds the putative inner working angle (IWA) of the telescope (83 milliarcseconds), planet-star brightness ratio above a minimum threshold (4×10^{-11}), and exoplanet apparent magnitude brighter than some threshold ($R_C < 31$) (Mamajek & Stapelfeldt, 2023). In Figure 1.2, we show designated “Tier A” FGK target stars from a preliminary HWO target list. Plotting these target stars in effective temperature and distance

Chapter 1. Introduction

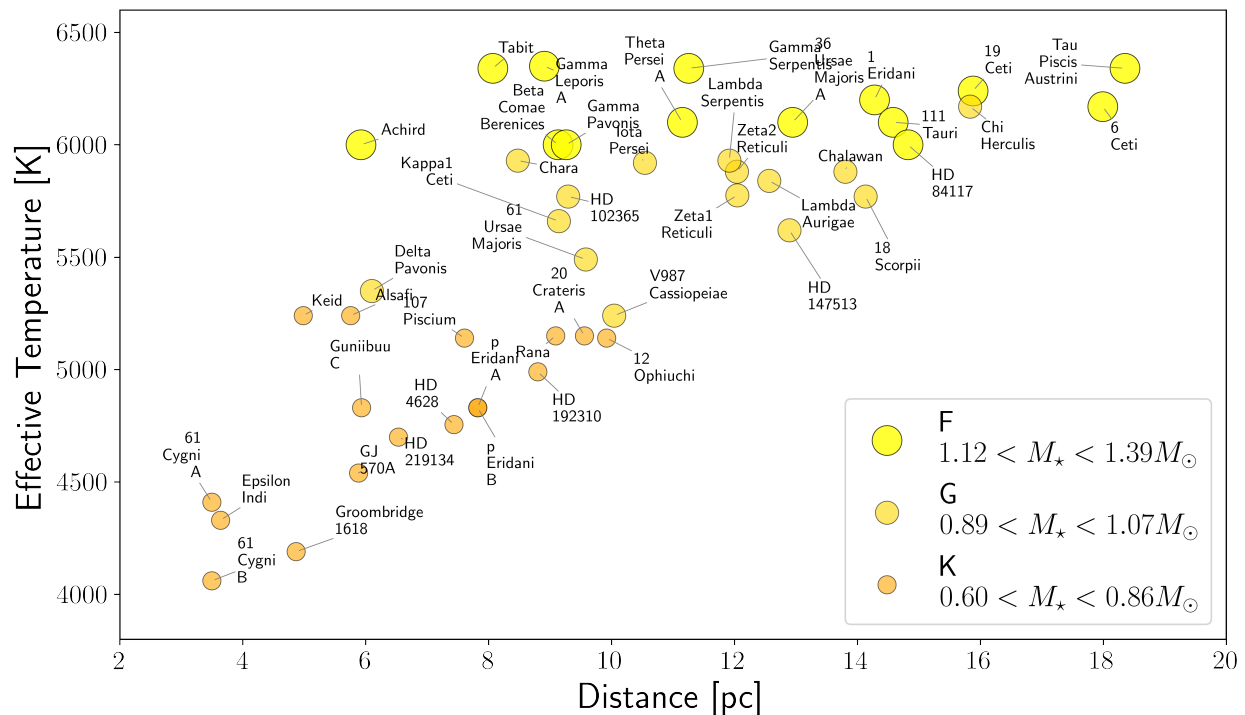


Figure 1.2: “Tier A” HWO targets identified by the preliminary HWO target list (Mamajek & Stapelfeldt, 2023) plotted in terms of their effective stellar temperatures and distances in parsecs. Most nearby stars (< 10 pc) are cooler K-dwarf stars.

Future missions will use transmission and direct imaging to probe rocky exoplanet atmospheres and look for direct and indirect signs of habitability. Transmission probes the terminator of the transiting planet, and is most sensitive to the upper troposphere and stratosphere, primarily due to the intrinsic molecular and scattering opacity of the atmosphere (Lincowski et al., 2018). In some cases, refraction can also significantly restrict access to the lower atmosphere in transmission observations (Muñoz et al., 2012; Bétrémieux & Kaltenegger, 2014; Misra et al., 2014a). Furthermore, cloud and aerosol extinction is known

to have a significant impact on transmission spectroscopy for larger planets (Fortney, 2005; Kreidberg et al., 2014), and this effect is likely highly wavelength dependent (Charnay et al., 2015). Modeling suggests that similar effects may occur on terrestrial worlds seen in transmission (e.g., Arney et al., 2017). In contrast to transmission, and by analogy with solar system observations, observations taken in reflected light could probe the planetary surface if there is broken cloud cover, and over a shorter atmospheric path, potentially making it less susceptible to obscuration by atmospheric gas and aerosol extinction (Fortney, 2005; Kreidberg et al., 2014). Key habitability indicators include the surface temperature, surface pressure, and surface liquid water content (Robinson, 2018; Meadows et al., 2018a). Habitability may be inferred by identifying and quantifying potential greenhouse gases, such as CO₂ and H₂O which may allow us to model the planet’s greenhouse warming and resultant surface temperature (Catling et al., 2018). Identifying the presence of an atmospheric “cold trap” may also aid in characterizing the habitability of the planetary surface by indicating warmer surface temperatures and volatile (e.g., water) condensation (Wordsworth & Pierrehumbert, 2014).

Transmission may be particularly sensitive to tropospheric temperature and pressure, since the atmospheric scale height ($\frac{RT}{\mu g}$) sets the size of the spectral features (Robinson, 2018). However, the continuum pressure need only supply a lower limit on the surface pressure, as sources of refraction and opacity may limit the transparency of the atmosphere to the surface (Misra et al., 2014b; Lincowski et al., 2018; Lustig-Yaeger et al., 2019a). In the absence of surface reflectivity variations and clouds, direct imaging could use the Rayleigh scattering slope and the width of molecular bands to help probe surface temperature and pressure, but this will likely be challenging due to the truncation of the visible path by even partial cloud cover, as well as degeneracies in pressure, gravity, and mean molecular weight, and poor prior constraints on related planet properties (Robinson, 2018). In the following section, we define planetary habitability and the HZ framework for target prioritization for future missions like HWO.

4 What Makes a Planet Habitable?

The defining trait of a habitable planet is its ability to sustain liquid water on its surface (e.g., [Dole & Asimov, 1964](#); [Robinson, 2017a](#)). As such, the HZ is defined as the range of orbital distances from a given host star at which a planet could plausibly maintain surface liquid water ([Kasting et al., 1993](#)). The HZ motivates a probabilistic framework for identifying priority targets for the search for life with future telescopic observations. For example, recent work argued that decision trees guiding future biosignature searches should focus HWO observations first on the detection of Earth-like abundances of water vapor as a key indicator of a potentially habitable surface below ([Young et al., 2024](#)).

The HZ is bounded by an inner and outer edge, referred to as the IHZ and OHZ, respectively. The inner edge may be defined by the runaway greenhouse effect, where surface oceans evaporate entirely ([Kopparapu et al., 2013](#)). [Kopparapu et al. \(2013\)](#) showed that this IHZ is not materially distinct from the IHZ set by atmospheric mass loss. Various definitions for the OHZ have been proposed, including the maximum greenhouse limit, where sufficiently high abundances of CO₂ result in increased atmospheric scattering and the tapering off of the greenhouse effect ([Kasting et al., 1993](#); [Kopparapu et al., 2013](#)). However, other work ([Haqq-Misra et al., 2016](#)) argued that this definition is too optimistic given CO₂ on Earth-like planets is cycled between the interior and the atmosphere over time ([Walker et al., 1981](#)). [Haqq-Misra et al. \(2016\)](#) argue that the OHZ should be set by the planet’s propensity to *maintain* dense CO₂ atmospheres given the carbonate-silicate cycle. [Kadoya & Tajika \(2019\)](#) build on the work of [Haqq-Misra et al. \(2016\)](#) by arguing that the OHZ should be set by the maximum greenhouse effect and the condensation of CO₂ given an ongoing carbonate-silicate cycle.

Beyond the OHZ, planets may experience global glaciations or “Snowball” events ([Haqq-Misra et al., 2016](#); [Kadoya & Tajika, 2019](#)). This follows studies of the evolution of Earth’s climate that have found evidence for intermittent Snowball periods in Earth’s history ([Hoff-](#)

man et al., 1998), possibly due to rapid oxygenation of the atmosphere by photosynthetic life (Pavlov et al., 2003) or limit cycles involving atmospheric CO₂ (Tajika, 2007; Mills et al., 2011). Limit cycles are oscillations between Snowball and ice-free states that occur when CO₂ degassing is not efficient enough to maintain a stable, temperate climate (Tajika, 2007; Abbot, 2016).

Previous work has used various modeling methods to estimate the evolving boundaries of the HZ around various types of stars, accounting for Snowball events and limit cycles. Kopparapu et al. (2013) used an updated 1-D radiative-convective climate model to estimate the static bounds of the HZ as a function of stellar mass. Haqq-Misra et al. (2016) used an updated energy balance model with a self-consistent carbonate-silicate cycle and seafloor weathering to revise previous estimates of HZ boundaries for abiotic planets. They also assessed the importance of limit cycles for determining the location of the OHZ. Kadoya & Tajika (2019) incorporated time-dependent planetary evolution by using a parameterized convection model to drive different CO₂ degassing histories for living planets. Their findings suggested narrower HZ regions than initially estimated by Kasting et al. (1993) and Kopparapu et al. (2013), and that living planets may be more susceptible to long-term glaciations. The reliance of Kadoya & Tajika (2019) on a biotic weathering reference state raises a deeper question: to what extent does the long-term habitability of an Earth-like planet depend on the presence of life in the first place?

Whether or not life is a requirement for habitability has vast implications for the search for life on exoplanets. If life is not required for habitability, then the detection of habitable conditions alone is insufficient evidence for life. Instead, distinguishing inhabited from habitable and uninhabited worlds will require identifying atmospheric gases or surface features that can only be produced by life.

Accurate estimation of the abiotic HZ requires a self-consistent model of planetary evolution that does not assume the presence of life. Previous work has made important progress, but each study leaves key physical processes unaccounted for. Kopparapu et al. (2013) and

[Haqq-Misra et al. \(2016\)](#) treat stellar properties as static snapshots rather than evolving quantities, and neither couples the carbon cycle to an explicit model of interior evolution. [Kadoya & Tajika \(2019\)](#) introduce time-dependent degassing driven by a parameterized convection model, but do not track ocean chemistry as an evolving carbon reservoir, do not allow land and ocean fractions to evolve as continental crust grows, and do not account for the role of the core.

5 Modeling the Whole Planet

The proper interpretation of photometry and spectroscopy of the “pale dots” detected in the HZ of nearby stars with HWO is predicated on a complete and self-consistent model of Earth-like planets. Many studies have simulated the co-evolution of the Earth and Sun based on geological evidence ([Barley et al., 2005](#); [Kharecha et al., 2005](#); [Zerkle et al., 2012](#); [Charney et al., 2013](#); [Arney et al., 2016](#); [Krissansen-Totton et al., 2018a](#); [Catling & Zahnle, 2020](#); [Wogan et al., 2022](#)) and extrapolated these results to predict the conditions of Earth-like planets around other stars ([Segura et al., 2005](#); [Kaltenegger et al., 2007](#); [Segura et al., 2010](#); [Frank et al., 2014](#); [Godolt et al., 2015, 2016](#); [Arney et al., 2017](#); [Gebauer et al., 2017](#); [Wolf et al., 2017](#); [Rugheimer & Kaltenegger, 2018](#); [Arney, 2019](#); [Young et al., 2023](#)). The existing work largely models the atmospheric evolution of these systems by coupling climate to photochemistry to achieve atmospheric compositions that are in radiative-convective equilibrium. However, atmospheric evolution cannot be fully understood in isolation from interior processes that regulate volatile cycling through the crust and mantle. While some studies have explored atmospheric evolution coupled to the mantle ([Krissansen-Totton et al., 2021](#)), to our knowledge no previous study has modeled the abiotic Earth by coupling a core-mantle-crust-climate model to an evolving Sun. To correctly interpret HWO spectra of Earth-like planets, we must close this gap by developing a “whole planet” model that co-evolves the core, mantle, crust, atmosphere, and star ([Foley & Driscoll, 2016](#)). Consequently, a validated

model of the entire planet is valuable for both understanding our own planet and optimizing the search for life on worlds outside the Solar System.

The emerging concept of “whole-planet” modeling was first introduced in a pair of papers (Driscoll & Bercovici, 2013; Foley & Driscoll, 2016), where it is argued that long-term habitability depends on evolving interior-atmosphere volatile exchange. Foley & Driscoll (2016) argue that climate influences whether or not plate tectonics develops, as low surface temperatures promote long-lived weak shear zones, or “damage”, in the lithosphere. In turn, plate tectonics facilitates volatile recycling at plate boundaries, which is particularly important for stabilizing the climate via the carbon cycle. Climate, therefore, can also affect the cooling rate of the interior, ultimately influencing the core cooling rate and thus magnetic field generation. Finally, the strength of the planetary magnetic field affects the rate of atmospheric escape, hypothesized to be a major factor in the desiccation of Venus and thus its distinct evolutionary path (Watson et al., 1981; Driscoll & Bercovici, 2013; Gillmann et al., 2022).

Despite its potential importance, many questions about whole planet modeling remain. Krissansen-Totton et al. (2021) developed a rigorously coupled mantle-surface-atmosphere model, which includes a magma ocean phase and a transition from a mobile lid to plate tectonics, to explore the long-term carbon and oxygen cycles. However, they did not focus on the coupled evolution of the mantle and core, instead assuming that the core cooling decays exponentially. Drawing on the ideas of Nimmo (2002), Armann & Tackley (2012) argue that the surface and core are interconnected, demonstrating that an episodic stagnant lid can suppress core cooling. Building on this work, Garcia et al. (2026) produced a coupled one-dimensional solar-atmosphere-lithosphere-mantle-core model of Venus that reproduces present-day atmospheric H_2O and CO_2 abundances and the lack of a core dynamo. Their ability to validate the model is limited by the dearth of observations of Venus’s surface and interior. Furthermore, as in Driscoll & Bercovici (2013), Garcia et al. (2026) assume a gray atmospheric model. To calculate the planet’s energy balance, Krissansen-Totton et al. (2021) use a pre-calculated radiative transfer grid to solve for the wavelength-dependent outgoing

long-wave radiation (OLR) from the planet, and use a simple parameterization of albedo, planet-star distance, and stellar luminosity to solve for the absorbed short-wave radiation (ASR). Thus, no whole-planet evolutionary model yet includes radiative transfer calculations to solve for the ASR and OLR balance of the changing planetary climate.

The atmospheres of “abiotic Earths” must be well-quantified to serve as null hypotheses for identifying life’s signatures and ruling out its imitators, or “false positives”, on exoplanets observed with HWO (Meadows et al., 2018b; Krissansen-Totton et al., 2021; Constantinou et al., 2025). In the next section, we introduce biosignatures, the molecular signs of life, and their possible false positive imitators.

6 Biosignatures and False Positives

Though living planets and their biosignatures may be incredibly diverse, life’s ~ 4 Gyr history on Earth acts as an informative baseline for mission design. Potentially inhabited planets may be identified by the detection of surface features and atmospheric biosignature gases, which may include biological pigments or the metabolic products of life, respectively (Des Marais et al., 2002; Seager et al., 2005; Meadows et al., 2018b; Schwieterman et al., 2018b; Schwieterman & Leung, 2024). Prior to the emergence of widespread photosynthetic life ~ 2.4 Ga, the anoxic Archean Earth atmosphere included abundant methane largely produced by life (Catling et al., 2001; Kasting, 2005; Thompson et al., 2022; Ebadi-rad et al., 2025). While atmospheric oxygen levels grew following the Great Oxidation Event ~ 2.4 Ga, the Proterozoic Earth atmosphere likely had O_2 and O_3 abundances orders of magnitude smaller than today (Kump, 2008; Lyons et al., 2014; Planavsky et al., 2014; Bellefroid et al., 2018). The Phanerozoic Earth atmosphere, with 21% O_2 , strong UV shielding due to stratospheric O_3 , and trace abundances of CH_4 represents the animal and plant-dominated planet we know today. Atmospheric water vapor and CO_2 , two possible indicators of habitability, have been present throughout Earth’s history (Catling et al., 2018). Based on Earth’s bio-

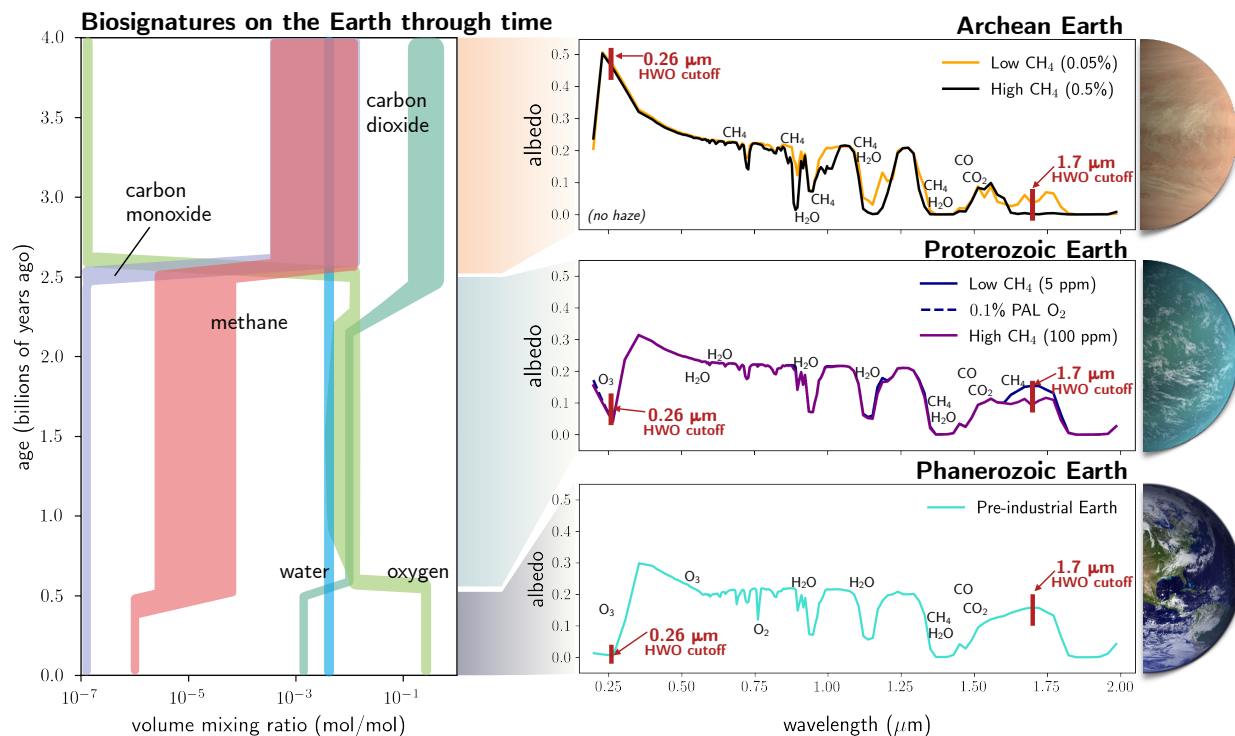


Figure 1.3: (*Left*) Atmospheric mixing ratios of key biosignature and habitability indicator gases across Earth’s geological history, spanning the Archean, Proterozoic, and Phanerozoic eras. (*Right*) Simulated HWO reflected light spectra for representative atmospheres in each era, computed at HWO’s baseline spectral resolutions ($R = 7$ UV, $R = 140$ optical, $R = 70$ NIR) with a 50% cloud fraction (Krissansen-Totton et al., 2025). The Archean spectra show prominent CH_4 absorption in the absence of O_2 or O_3 ; the Proterozoic spectra show weak O_3 features alongside CH_4 and H_2O ; and the Phanerozoic spectrum shows strong O_2 and O_3 features characteristic of the modern Earth. Red bars indicate HWO’s UV ($0.26 \mu\text{m}$) and NIR ($1.7 \mu\text{m}$) wavelength cutoffs. This figure originally appeared in Krissansen-Totton et al. (2025) (submitted to *Astrobiology*), to which the author contributed this analysis.

geochemical history, important molecules for detecting life on exoplanets thus include O_2 , O_3 , H_2O , CH_4 , and CO_2 . The evolution of the abundances of these gases in key eras are shown in Figure 1.3 (Krissansen-Totton et al., 2025).

In particular, pairs of gases that are not in thermodynamic equilibrium could indicate active fluxes of gases at the surface that are produced by life. Three prominent pairs to consider for future telescopes are O_2/CH_4 (Hitchcock & Lovelock, 1967), O_3/CH_4 (Des Marais et al., 2002) and CO_2/CH_4 (Krissansen-Totton et al., 2016; Schwieterman et al., 2016). The canonical biosignature pair is O_2 and CH_4 – without life, these molecules should rapidly

destroy each other such that their simultaneous presence in an atmosphere strongly implies efficient production that is inconsistent with known abiotic sources (Hitchcock & Lovelock, 1967). O_3 is a photochemical byproduct of O_2 in our atmosphere, and may therefore act as a proxy for O_2 when O_2 is not observationally accessible (Des Marais et al., 2002). This substitution gives the O_3/CH_4 biosignature pair. However O_3 production is non-linear with O_2 abundance (Kasting & Donahue, 1980), especially for cooler M-dwarf stars (Segura et al., 2003; Kozakis et al., 2022), and inferring the O_2 abundance from a measurement of atmospheric O_3 may be challenging. Finally, both CH_4 and CO_2 have been present and potentially detectable (Meadows et al., 2023) throughout Earth’s Archean epoch despite drastic changes in the composition of the atmosphere, making it advantageous for applications on exoplanets that may be inhabited but lack a photosynthetic biosphere (Krissansen-Totton et al., 2016). Additionally, this biosignature pair may be particularly favorable for planets around M-dwarf (Segura et al., 2005) and K-dwarf (Arney, 2019) stars, which may experience a photochemical enhancement of CH_4 .

Photochemistry will complicate the task of interpreting biosignature pairs since incident stellar energy may enhance or suppress abiotic and biogenic molecules (Meadows et al., 2018b), leading to gradients in the chemical profile (Segura et al., 2003, 2005). Detecting these chemical gradients may provide insight into photochemical processes and thus a more robust interpretation of biosignature pairs in the context of their environment. However, the detection of these molecules alone does not constitute a confident life detection due to the possibility of “false positive atmospheres” (Domagal-Goldman et al., 2014; Meadows et al., 2018b; Krissansen-Totton et al., 2021; Thompson et al., 2022). False positive atmospheres mimic life’s presence via the abiotic generation of biosignature gases, and the effort to identify possible molecular discriminants to rule out such scenarios is ongoing (Harman et al., 2015; Schwieterman et al., 2016; Harman & Domagal-Goldman, 2018; Meadows et al., 2018b; Schwieterman et al., 2019b; Krissansen-Totton et al., 2021; Ulses et al., 2025). One advantage of biosignature pairs is that in addition to indicating a disequilibrium, they increase the

probability that at least one of the gases is due to a biological process by helping to rule out false positive mechanisms. For the O_2/CH_4 pair the CH_4 is not only an indication of a methanogenic biosphere, an additional biological process to the photosynthesis generating the O_2 , but it also acts as a false positive discriminant, making abiotic O_2 production a less likely explanation for observed O_2 and O_3 . However, false positive O_3 can potentially be generated by the photolysis of CO_2 in an environment that does not contain the catalysts needed for rapid recombination, such as H_2O (Gao et al., 2015). In this context, high abundances of free CO (Schwieterman et al., 2016), and vanishingly low abundances of H_2O may be useful discriminants to help identify a potential false positive (Gao et al., 2015).

For methanogenic life, recent work evaluating possible detections of CH_4 on Mars argued that abundant CO would imply an abiotic origin (Zahnle et al., 2011), since CO is a more energetically advantageous source of carbon than CO_2 (Ragsdale, 2004). Building on this argument, recent work proposed that the presence of abundant CO alongside CH_4 and H_2O is one possible “antibiosignature” for anoxygenic atmospheres (Wogan & Catling, 2020; Krissansen-Totton et al., 2022). Furthermore, in reducing biospheres, Schwieterman et al. (2019b) showed that biomass burning may only plausibly produce CO abundances as high as 10 ppm on planets around FGK stars. CO may also be relevant to the identification of oxygenic biospheres. Hu et al. (2020) showed that photolysis in habitable, CO_2 -rich atmospheres could lead to the abiotic accumulation of both CO and O_2 , potentially producing a false positive biosignature scenario where oxygen is present without life. However, in follow-up work Ranjan et al. (2023) demonstrated that this result was likely a model artifact related to the upper atmosphere boundary condition, finding that O_2 would remain at trace abundances rather than accumulating to modern-Earth-like levels, thereby preserving CO’s utility as an antibiosignature for Earth-through-time analogs.

7 Atmospheric Retrieval for Exoplanet Characterization

Spectroscopic retrieval is a powerful technique to infer planetary and atmospheric properties from remote spectroscopic data (Barstow et al., 2020) and to forecast observational needs and interpretation challenges prior to future data acquisition. Retrievals have been used extensively to analyze solar system bodies (e.g., Mahieux et al., 2010; Arney et al., 2014; Irwin et al., 2008; Spurr et al., 2001; Kleinböhl et al., 2009; Vinatier et al., 2007; Kim et al., 2011). More recently, pioneering retrieval work has inferred atmospheric properties of giant exoplanets using data from current observatories (e.g., Kreidberg et al., 2015; Wakeford et al., 2017; Benneke et al., 2019; Line et al., 2021). However, these applications have also identified critical challenges, such as degeneracies between clouds and mean molecular weight (Line & Parmentier, 2016) and biases due to non-uniform/multiple thermal profiles (Feng et al., 2016; Feng et al., 2020; Taylor et al., 2020), and others (e.g., Caldas et al., 2019; Pluriel et al., 2020; MacDonald & Lewis, 2021; Nixon & Madhusudhan, 2022), that motivate more investigations into how model over-simplifications can impact planetary scale interpretations.

The lessons learned from giant exoplanet work are only just beginning to be applied to rocky exoplanets, and many gaps still remain in identifying sources of potential biases and confounding factors when interpreting habitability and biosignatures from the simulated spectra of terrestrial exoplanet atmospheres. JWST has begun delivering the first transmission spectra of rocky exoplanet atmospheres, with early retrievals on TRAPPIST-1 planets and other M-dwarf terrestrials placing initial constraints on atmospheric composition and thickness (e.g., Lim et al., 2023; Lustig-Yaeger et al., 2023c; Zieba et al., 2023; Glidden et al., 2025; Radica et al., 2025), while direct imaging studies have looked ahead to the FGK regime accessible with future flagship missions. For future direct imaging missions, Feng et al. (2018) assessed the instrument requirements needed to make significant detections of water vapor, ozone, and oxygen on a directly-imaged Earth-twin with patchy water vapor clouds, but their investigation was limited to 0.4–1.0 μm , thereby excluding methane fea-

tures as well as the effects of clouds in the near-IR. By contrast, [Damiano & Hu \(2022\)](#) included methane features in their study and fit for the vertical structure of water vapor in the atmosphere to include the effects of clouds. Previous studies have also assessed our ability to spectrally retrieve the characteristics of potentially habitable planets around M-dwarfs in the near and mid-IR ([Tremblay et al., 2020](#)) and planets around white dwarfs ([Kaltenegger et al., 2020](#)) with transmission observations, but these studies were limited to non-self-consistent modern Earth atmospheres. Furthermore, [Feng et al. \(2018\)](#), [Tremblay et al. \(2020\)](#), and [Damiano & Hu \(2022\)](#) produced their synthetic data from atmospheres generated with unrealistic uniform vertical profiles for temperature and molecular abundances. Though [Lin et al. \(2021\)](#) investigated our ability to differentiate self-consistent prebiotic and modern Earth-like atmospheres on TRAPPIST-1 e with high resolution JWST transmission observations, they did not assess our ability to discriminate between an abiotic planet and an inhabited one. Finally, existing studies have yet to fully assess the comparative (and potentially complementary) strengths and weaknesses of transmission and direct imaging spectra for assessing habitability and biosignatures, and probing the surface environment, on rocky, habitable-zone planets.

The exploration of observatory trade space has benefited enormously from the application of spectral retrieval models because they help assess our ability to meaningfully constrain gas abundances ([Damiano & Hu, 2021, 2022](#); [Damiano et al., 2023](#); [Latouf et al., 2023a,b](#); [Gilbert-Janizek et al., 2024](#); [Tokadjian et al., 2024](#); [Latouf et al., 2025](#); [Young et al., 2025](#); [Krissansen-Totton et al., 2025](#); [Hagee et al., 2026](#)). Abundance constraints provide unique information about the character of planetary atmospheres that detections alone cannot. To ensure that HWO is capable of finding life on exoplanets, the community is considering how different observatory architectures and baseline instrument parameters affect our ability to detect and constrain these key gases. To observe and characterize comparatively dim exoplanet atmospheres around FGK stars, HWO will be equipped with a coronagraph and reflected light spectroscopy capable of UV/optical/NIR observations. In addition to spec-

troscopic wavelength range and band-pass signal-to-noise ratios (SNRs) (Krissansen-Totton et al., 2025), we must determine the spectral resolving power required to characterize the atmospheres of Earth-like exoplanets and ultimately identify habitable and/or living planets (LUVOIR Team, 2019; Gaudi et al., 2020).

Previous work argued that HWO’s spectral resolution ($\lambda/\Delta\lambda$) in the visible ($0.4 - 1.0 \mu\text{m}$) should be set by the detection of the narrow-band O_2 feature at $0.76 \mu\text{m}$ (Des Marais et al., 2002; Brandt & Spiegel, 2014; Feng et al., 2018; LUVOIR Team, 2019; Gaudi et al., 2020; Damiano & Hu, 2022), while the resolution in the near-IR ($1.0 - 1.7 \mu\text{m}$) should be set by the detection of CH_4 and CO_2 features (Des Marais et al., 2002; LUVOIR Team, 2019; Gaudi et al., 2020; Damiano & Hu, 2022). These recommendations have emerged from studies using a variety of analysis techniques, including measurements of band-pass absorption features (Des Marais et al., 2002), gas-specific detectability calculations (Brandt & Spiegel, 2014), and finally atmospheric retrieval simulations (Feng et al., 2018; Damiano & Hu, 2022). Recently, Feng et al. (2018) and Damiano & Hu (2022) used retrieval models to assess how two different spectral resolutions in the visible impact our ability to characterize Earth-like atmospheres. Both studies explored how $R = 70$ and $R = 140$ in the visible impact our ability to detect and constrain O_2 , O_3 , and H_2O . Additionally, Damiano & Hu (2021) showed that including $R = 40$ observations in the near-IR wavelengths provides important contextual information about atmospheric CO_2 . Lastly, the LUVOIR final report recommended $R_{\text{NIR}}=70$ for CO_2 detection with an integral field spectrograph and $R_{\text{NIR}}=200$ for single-point spectrographs (LUVOIR Team, 2019). However, no retrieval study to-date has fully and systematically explored the effect of spectral resolution in the visible and the near-IR on our ability to detect and constrain biosignatures for Earth-through-time atmospheres with HWO.

8 Thesis Overview

In this introduction, we have laid out the critical knowledge we have gained from previous work and identified potential gaps motivating further investigation. The rest of this thesis comprises four chapters describing original research and one chapter summarizing our broader conclusions and possible directions for future work. Here we briefly summarize the motivating gaps and the corresponding work that addresses them in the chapters ahead:

Chapter 2: Modeling the Abiotic Earth Around the Sun

No previous whole-planet model couples a core-mantle-crust thermal evolution model to a carbonate-silicate carbon cycle, deep and surface water cycles, and radiative transfer climate calculations for the abiotic Earth. We develop a coupled whole-planet model, validated on 19 pre-industrial Earth observations, and use it to generate a synthetic HWO reflected light spectrum as proof of concept for future detectability studies. We discuss our model’s implications for the Gaia hypothesis, as well as limitations of our model that should be addressed in future work.

Chapter 3: Modeling the Abiotic Earth Around HWO Target Stars

Previous HZ studies either treat stellar properties as static snapshots or lack explicit ocean carbon partitioning and evolving land/ocean fractions — none provides a self-consistent abiotic treatment of HZ evolution across stellar types. We extend the Chapter 2 model to FGK stellar hosts relevant to the HWO target list. In addition to exploring how stellar type and interior evolution shapes the habitability of Earth-like planets, we also investigate the sensitivity of our results to assumptions about the efficiency of continental and seafloor weathering.

Chapter 4: Retrievals for Terrestrial Atmospheres

No previous retrieval study has compared the complementary capabilities of transmission and direct imaging spectroscopy for habitability assessment and biosignature detection using self-consistent forward model atmospheres, or assessed our ability to discriminate

inhabited from abiotic planets. We use a Bayesian retrieval framework to transmission (Origins/TRAPPIST-1 e) and direct imaging (LUVOIR/Earth-Sun analog) by simulating observations of an Earth-like vs. abiotic false positive O_3 planet in transmission and a clear vs. cloudy Earth in direct imaging. We also perform vertically-resolved O_3 retrievals and Bayesian model comparison to assess the sensitivity of transmission and direct imaging to the stratospheric ozone bulge in Earth-like atmospheres. We discuss what each technique can and cannot constrain for habitability assessment and biosignature detection.

Chapter 5: Retrievals for Investigating Optimal Observatory Parameters

No retrieval study has systematically explored the effect of spectral resolution across both visible and near-IR wavelengths on biosignature detection and characterization for Earth-through-time atmospheres, leaving HWO spectrometer requirements insufficiently constrained. We perform analytic detectability calculations and targeted retrieval studies for assessing HWO spectral resolution needed for detecting biosignatures/habitability indicators on the Earth through time. We also assess the sensitivity of our results to the magnitude of the dark current. We use our results to identify the optimal spectral resolutions in the visible and near-IR wavelengths assuming a nominal dark current value for HWO's coronagraph.

The final chapter reviews our key findings and outlines areas for future work.

CHAPTER 2

MODELING THE ABIOTIC EARTH AROUND THE SUN

A version of this chapter has been submitted to The Planetary Science Journal and is currently under review, as: Gilbert-Janizek, S., et al., “A One-Dimensional, Whole-Planet Model of the Abiotic Earth through Time for Terrestrial Exoplanet Studies.” R.K. Barnes, P.E. Driscoll, and A.M. Mandell supervised this work; N.F. Wogan advised on the use of his climate model; J.L. Birky advised on the use of her maximum likelihood estimation code; L. Carone advised on the proper initial conditions for the transition from the magma ocean to the solid planet phase; R. Garcia advised on incorporating the deep water cycle into the model; all co-authors provided comments and edits on the manuscript; all other work is my own.

1	Properties of the Pre-Industrial Earth	25
1.1	Atmosphere	26
1.2	Surface Properties	29
1.3	Interior	30
2	Methods	34
2.1	Model Assumptions	36
2.2	Core, Mantle, and Basal Magma Ocean	37
2.3	Carbon Cycle	39

2.4	Water Cycle	43
2.5	Atmospheric Constituents	45
2.6	Ocean Chemistry	47
2.7	Surface Albedo	51
2.8	Climate	53
2.9	Reflected Light Spectra	55
2.10	Initial Conditions	56
3	Results	58
4	Discussion	68
4.1	Spectroscopic Predictions for HWO Observations	69
4.2	CH ₄ as a Greenhouse Gas	69
4.3	Abiotic Marine Calcite Budget	70
4.4	Implications for the Gaia Hypothesis	72
4.5	Comparison to Earth Evolution	76
4.6	Future Work	79
5	Conclusions	85
6	Constants	87
7	BMO Radiogenic Heating Calculation	90
7.1	Derivation of Effective Core Heating Rate	91
7.2	Required Radiogenic Enrichment	93

Here we present a one-dimensional, core-mantle-crust-climate model of the abiotic Earth that tracks the evolution of the core-mantle system, geochemistry, atmosphere, incoming solar flux, and solar effective temperature over the age of the Solar System. We use our model to reproduce 19 observed quantities of the pre-industrial Earth. We choose this calibration point as the modern Earth is not in radiative balance due to human-driven climate change. Due to large uncertainties in early Earth conditions, we do not attempt to validate the model on the conditions of the early Earth. Finally, as a proof of concept, we demonstrate that atmospheric states from our model can be used to generate realistic reflected light spectra comparable to observations HWO will obtain of Earth-like exoplanets.

In the next section, we describe the observed properties of the pre-industrial Earth and briefly describe relevant previous work modeling its evolution. We describe the whole planet abiotic Earth model in Section 2, and present its validated results in Section 3. We discuss future model upgrades as well as potential applications of our validated model for HWO observations of Earth-like exoplanets in Section 4. Finally, we summarize our results and conclusions in Section 5.

1 Properties of the Pre-Industrial Earth

The “pre-industrial” era defines the period of Earth’s history prior to the industrial revolution (c. 1800), when widespread changes to human land-use and atmospheric pollution first arose (Hawkins et al., 2017). From records, geological evidence, and modeling reconstructions, the PIE is known to have had lower atmospheric and ocean CO₂ concentrations and cooler global average surface temperatures than today. Measurements of these properties with their uncertainties and corresponding sources are summarized in Table 2.1. We report interior properties based on present-day measurements because (1) pre-industrial measurements of these properties are largely nonexistent, and (2) mantle and core properties evolve on a sufficiently slow timescale that any variation from their present-day values should be in-

significant. Similarly, we assume that the total mass of the ocean has not varied significantly in the last 300 years.

1.1 Atmosphere

Earth retains a 1-bar atmosphere consisting of 78% N_2 ([United States Committee on Extension to the Standard Atmosphere, 1976](#)). Oxygen, which comprises 21% of our atmosphere, became a major constituent after the proliferation of photosynthetic life ([Kump, 2008](#); [Farquhar et al., 2011](#); [Lyons et al., 2014](#)). The remaining atmosphere consists of “trace components”, including CO_2 and H_2O , which are effective greenhouse gases that help regulate planetary climate ([Tyndall, 1861](#)). Water vapor is the most efficient greenhouse gas ([Fleming, 1998](#)), and because warmer air can hold more water vapor, the atmospheric abundance of H_2O relates to global temperature in a positive feedback loop ([Held & Soden, 2000](#); [Patel & Kuttippurath, 2023](#)). However, water vapor’s short atmospheric residence time of ~ 10 days ([Trenberth, 1998](#)) – set by condensation and precipitation – means that H_2O cannot itself act as a long-term climate forcing. Instead, the non-condensing greenhouse gases such as CO_2 , with atmospheric lifetimes of centuries to millennia, set the temperature structure that determines the equilibrium water vapor abundance ([Lacis et al., 2010](#)). H_2O thus amplifies climate change driven by CO_2 rather than causing it. Below we detail measurements of Earth’s atmospheric CO_2 and H_2O abundances, as well as global average surface temperature.

A record of atmospheric CO_2 levels has been pieced together by analyzing air trapped in polar ice sheets. [Etheridge et al. \(1996\)](#) analyzed air bubbles in Antarctic ice cores dating back to 1000 C.E., finding that pre-industrial atmospheric CO_2 abundances were around 280 ppm with annual fluctuations of ~ 4 ppm. In Earth’s 1-bar atmosphere (101325 Pa), this abundance is equivalent to a partial pressure of 28.4 ± 0.4 Pa.

The abundance of H_2O in our atmosphere varies with latitude, with higher abundances in the humid tropics ([Mockler, 1995](#); [Allan et al., 2022](#)), as well as altitude, with much

Table 2.1: Selected properties of the pre-industrial Earth.

Property	Value	Reference(s)
<i>Atmosphere</i>		
Atmospheric CO ₂ ($p\text{CO}_2$)	28.4 ± 0.4 Pa	(1)
Atmospheric H ₂ O ($p\text{H}_2\text{O}$)	900 ± 700 Pa	(2, 3)
Global avg. surface temperature (T_{surf})	286.9 ± 0.1 K	(4, 5)
<i>Surface</i>		
Ground albedo (A_g)	0.14 ± 0.02	(6, 7, 8)
Ocean mass (w_{ocean})	1.4×10^{21} kg (1 T.O.)	(9)
Oceanic (CO ₂) _{aq}	8 ± 2 $\mu\text{mol/kg}$	(10)
Surface Ocean Total Dissolved Inorganic Carbon (DIC)	2000 ± 200 $\mu\text{mol/kg}$	(10, 11)
Ocean pH	8.18 ± 0.05	(12)
<i>Interior</i>		
Upper mantle temperature (T_{UM})	1587^{+164}_{-34} K	(13)
Core-mantle boundary temp. (T_{CMB})	4000 ± 200 K	(14)
Upper mantle heat flow (Q_{UM})	38 ± 3 TW	(13)
Core-mantle boundary heat flow (Q_{CMB})	11 ± 6 TW	(13)
Upper mantle viscosity (ν_{UM})	$10^{16.68 \pm 1} \text{ m}^2 \text{ s}^{-1}$ ($10^{20.2 \pm 1} \text{ Pa}\cdot\text{s}$)	(15)
Lower mantle viscosity (ν_{LM})	$10^{17.28 \pm 1} \text{ m}^2 \text{ s}^{-1}$ ($10^{21 \pm 1} \text{ Pa}\cdot\text{s}$)	(15)
Upper mantle melt fraction (f_{UM})	$11.5 \pm 3.5\%$	(16)
Mantle melt mass flux	$(1.3 \pm 0.8) \times 10^6$ kg/s	(17, 18)
Inner core radius (R_{IC})	1221 ± 1 km	(19)
Magnetic moment	$80 \pm 4 \text{ ZAm}^2$ ($1.00 \pm 0.05 \text{ E.U.}$)	(20, 21)
Magnetopause radius	$9.10 \pm 0.14 R_{\oplus}$ ($1.00 \pm 0.02 \text{ E.U.}$)	(22)

References: (1) [Etheridge et al. \(1996\)](#), (2) [Robinson et al. \(2011\)](#), (3) [Lustig-Yaeger et al. \(2023a\)](#), (4) [Hawkins & Sutton \(2016\)](#), (5) [Hawkins et al. \(2017\)](#), (6) [Solomon et al. \(2007\)](#), (7) [Kargel et al. \(2014\)](#), (8) [Genda \(2016\)](#), (9) [Charette & Smith \(2010\)](#), (10) [Feely et al. \(2001\)](#), (11) [Eide et al. \(2017\)](#), (12) [Jiang et al. \(2019\)](#), (13) [Jaupart et al. \(2015\)](#), (14) [Hirose et al. \(2013\)](#), (15) [Paulson et al. \(2005\)](#) with uncertainties from [Métivier et al. \(2016\)](#), (16) [Gale et al. \(2013\)](#), (17) [Cogné & Humler \(2004\)](#), (18) [Li et al. \(2015\)](#), (19) [Dziewonski & Anderson \(1981\)](#), (20) [de Pater & Lissauer \(2010\)](#), (21) [Kivelson & Bagenal \(2014\)](#), (22) calculated by [Driscoll & Bercovici \(2013\)](#) using equations from [Griekmeier et al. \(2004\)](#).

higher abundances below the atmospheric cold trap (Brewer, 1949; Randel & Park, 2019). The abundance and distribution of water vapor is also seasonally variable, but the resulting change to the global atmosphere is small, on the order of $\pm 1 \times 10^{15}$ kg H₂O, or less than 0.02% of the total atmospheric mass (Trenberth et al., 1987). Satellite and sounding measurements have provided anchor points for 1-D models to map the vertical distribution of water vapor throughout Earth’s atmosphere; for example, Robinson et al. (2011) analyzed archived satellite data corresponding to the dates of the EPOXI Earth-flyby mission in 2008, extracting water vapor abundance profiles at latitudes of 0°, 19°, 41°, and 66° (Lustig-Yaeger et al., 2023a, their Fig. 4). Each latitude profile represents the globally averaged vertical profile for all HEALPix bins with the corresponding north and south latitudes. The mean water vapor profile includes a near-surface water vapor volume mixing ratio of 0.0089, with near-surface abundances of 0.0158 and 0.00182 for the mean profiles acquired at 0° and 66°, respectively (Lustig-Yaeger et al., 2023a). Treating the equatorial and mid-latitude averages as upper and lower bounds, these measurements correspond to a present-day partial pressure of 900 ± 700 Pa, a large uncertainty that reflects the strong latitude-dependency of near-surface water vapor abundances.

Earth’s global average surface temperature can be traced through time directly via historical records, and indirectly by estimating evolving radiative forcings, using global climate model simulations, and finally by compiling proxy evidence from tree-rings, corals, and ice cores. Hawkins et al. (2017) combine all of these approaches to estimate that pre-industrial temperatures were 0.55–0.80°C cooler than the global average surface temperature from 1986 to 2005. Establishing the corresponding 1986-2005 baseline as 14.4°C (Hawkins & Sutton, 2016), we therefore assume the pre-industrial global average surface temperature to be 286.9 ± 0.1 K (Hawkins et al., 2017).

1.2 Surface Properties

To first order, Earth’s surface includes exposed land, oceans, and polar ice caps. These three surface types, along with atmospheric water vapor clouds, contribute to the overall brightness or albedo of the planet, regulating planetary absorption of stellar radiation. In addition to H_2O , the oceans contain dissolved inorganic carbon (DIC) in the form of aqueous CO_2 , carbonate (CO_3^{2-}), and bicarbonate (HCO_3^-), which together mediate the overall pH of the marine environment.

Earth sustains a surface liquid water ocean of mass 1.4×10^{21} kg, or by definition, 1 terrestrial ocean (T.O.) (Charette & Smith, 2010). Earth’s vast oceans are a major carbon sequestration reservoir. Atmospheric CO_2 dissolves into the oceans as aqueous CO_2 and reacts with water to form carbonic acid, which dissociates into H^+ and HCO_3^- , which itself dissolves into H^+ and CO_3^{2-} (Pilson, 2012).

Studies of ocean carbon isotopic fractionation reveal that the pre-industrial oceanic carbon budget is largely abiotic. Since photosynthetic respiration preferentially uses the ^{12}C isotope (O’Leary, 1981), ocean water enriched in heavier ^{13}C indicates abiotic dissolved inorganic carbon (Quay et al., 2003). Isotopic fractionation studies suggest non-human biological activity accounts for only $\sim 2\%$ of the total DIC in the oceans (Carroll et al., 2022). Though marine life pumps dissolved carbon from the near-surface water column to the deep ocean (Le Moigne, 2019), gravitational sinking dominates carbon export in modern oceans (Boyd et al., 2019; Nowicki et al., 2022; Siegel et al., 2023). Recent work using models to reconstruct pre-industrial ocean chemistry supports near-surface, pre-industrial DIC concentrations of 2000 ± 200 $\mu\text{mol/kg}$ (Feely et al., 2001; Eide et al., 2017), an aqueous CO_2 concentration of 8 ± 2 $\mu\text{mol/kg}$ (Feely et al., 2001), and a global average surface ocean pH of 8.18 ± 0.05 (Jiang et al., 2019).

Planetary surface albedo, a linear combination of albedo contributions from different surface types and clouds, affects climate by regulating the Earth’s absorbed shortwave radiation

from the Sun. Based on data from spectral libraries (Jet Propulsion Laboratory, 1999; Clark et al., 2003), previous planetary modeling work assumes that the albedos of seawater, rock, and ice are largely wavelength independent (Kaltenegger et al., 2007; Driscoll & Bercovici, 2013), where ocean albedo $A_{oc} = 0.1$, rock albedo $A_r = 0.17$, and ice albedo $A_i = 0.6$. The ocean covers 71% of Earth’s surface area (Genda, 2016), and the remainder comprises land with seasonal variations in ice coverage.

Satellite measurements have been used to quantify a perennial surface ice area of $16.01 - 16.04 \times 10^6 \text{ km}^2$ (Solomon et al., 2007; Kargel et al., 2014), or a perennial ice surface fraction of 3.4% that implies a remaining ice-free land surface fraction of 25.6%. On Earth, most of the ice-less land is covered by vegetation (Chu, 2019), with 54% grasslands ($A \sim 0.4$) and 30% forests ($A \sim 0.1$). Since we are investigating an abiotic planet without life of any kind, we will assume that the surface of the abiotic Earth would instead be dominated by rock. Calculating the total expected ground albedo (A_g) as a linear combination of surface fraction and albedo contributions, we thus calculate $A_g = 0.14$. Seasonal changes in Earth’s albedo imply variations of ± 0.02 (Bartman, 1980).

In reality, Earth’s *total* albedo also receives major contributions from atmospheric water vapor clouds, as satellite measurements suggest Earth’s total albedo is closer to $A_T = 0.29$ (Wielicki et al., 1996; Loeb et al., 2009). This total albedo is ~ 0.15 brighter than the composite ground albedo because clouds are highly reflective, and they partially block the path of stellar radiation to the ground (Ramanathan et al., 1989; Harrison et al., 1990). For an albedo representing an average of thick and thin clouds, de Pater & Lissauer (2010) and Driscoll & Bercovici (2013) use $A_{cl} = 0.89$.

1.3 Interior

Beneath a thin crustal layer, Earth has a convecting solid mantle, molten outer core, and solid iron inner core. Here we describe how various measurements and models have been used to establish values (with uncertainties) for key core and mantle properties of the Earth,

including temperature, heat flow, viscosity, upper mantle melt fraction, and magnetic moment.

The mantle geotherm describes how the laterally-averaged temperature of the mantle changes with depth. The temperature of the mantle at various depths may be estimated using petrological constraints on the temperature of mid-ocean ridge basalts (Kinzler & Grove, 1992) and the identification of several important phase changes: the 440 km deep olivine-wadsleyite phase change (Katsura et al., 2004), the 660 km deep post-spinel phase transition, (Katsura et al., 2003), and the inner core boundary (Alfè et al., 2002; Labrosse, 2003; Jaupart et al., 2015). Using the depth of these phase changes as anchor points, previous studies use isentropic profiles combined with thermal boundary layers to describe a continuous mantle geotherm (e.g., Katsura et al., 2004; Driscoll & Bercovici, 2014). Synthesizing these measurements and models has resulted in the following literature estimates for the temperature of the upper mantle (at the base of the lithosphere) and the core-mantle boundary: $T_{\text{UM}} = 1587^{+161}_{-34}$ K (Jaupart et al., 2015) and $T_{\text{CMB}} = 4000 \pm 200$ K (Hirose et al., 2013), respectively.

Over geological time the mantle undergoes solid state convection, transporting heat from the CMB to the surface. The viscosity of the mantle, an important quantity for determining the vigor of mantle convection and heat transfer, is expected to vary significantly between the upper and lower mantle due to variations in temperature, pressure, and composition (Jeanloz & Morris, 1986; Boehler, 1996; Jaupart et al., 2015). To estimate a 1D (radial) viscosity profile throughout the mantle Paulson et al. (2005) used satellite observations of secular changes in gravity near Hudson Bay to measure Post-Glacial Rebound, or the rate at which land masses gradually “bounce back” to their ice-free hydrostatic level after the melting of massive ice sheets.

Implementing 1-D averages of a 3-D viscosity model, Paulson et al. (2005, Fig. 3) report global-average dynamic viscosities of $\sim 10^{20.2}$ Pa·s (upper mantle, depth < 250 km) and $\sim 10^{21}$ Pa·s (lower mantle). Assuming $\rho_{\text{UM}} = 3300$ kg/m³ and $\rho_{\text{LM}} = 5200$ kg/m³, we

adopt nominal kinematic viscosities of $\nu_{\text{UM}} \approx 4.8 \times 10^{16} \text{ m}^2/\text{s}$ and $\nu_{\text{LM}} \approx 1.9 \times 10^{17} \text{ m}^2/\text{s}$, which we assume are uncertain by ± 1 order of magnitude. This envelope corresponds to approximately $\pm 3\sigma$ relative to the formal posterior uncertainty on 1-D globally-averaged upper-mantle viscosity reported by [Métivier et al. \(2016\)](#) ($10^{20.7 \pm 0.3} \text{ Pa}\cdot\text{s}$), and is comparable to the width of the bimodal lower-mantle posterior from the same study (their Fig. 6).

Although surface measurements constrain the upper mantle heat flow to $Q_{\text{UM}} = 38 \pm 3 \text{ TW}$ ([Jaupart et al., 2015](#)), the core-mantle boundary (CMB) heat flow can not be measured directly. However, the temperature jump across the lower mantle thermal boundary layer can be constrained by extrapolating the mantle geotherm down and the inner core boundary (ICB) temperature, i.e. the solidification temperature of the core, up to the CMB. This technique combined with the estimated viscosity of the lower mantle typically gives an estimate of Q_{CMB} to be $11 \pm 6 \text{ TW}$ ([Jaupart et al., 2015](#)). Thermal evolution models of the Earth, therefore, must be within these margins of uncertainty to be considered successful.

The subduction of Earth’s tectonic plates drives the upwelling of ambient mantle at mid-ocean ridges, facilitating the cycling of volatiles between the solid interior and atmosphere over geological timescales. The fraction of upwelling mantle that melts and its rate of flow determine the melt mass flux, which also plays a role in heat transport and can have a direct impact on planetary thermal evolution. The extent of melting of mid-ocean ridge basalts (MORB) has been estimated by compiling various measurements of basalt chemistry as a function of axial depth, and averaging the data over large distances to account for small-scale variability in basalt composition ([Klein & Langmuir, 1987](#)). This analysis suggests that the average extent of MORB melting is 8 – 15% ([Gale et al., 2013](#)).

Measurements of crustal production rates at plate boundaries probe mantle melting by providing estimates on the upward mass flux of mantle melt. [Cogné & Humler \(2004\)](#) use direct measurements of currently visible seafloor surfaces bounded by pairs of geophysical isochrones in the Atlantic, Indian, Antarctic, Pacific, Nazca, and Cocos ocean basins to estimate a seafloor production rate of 5 – 15 km^3/year . [Li et al. \(2015\)](#) use three-dimensional

seismic tomography to measure magmatic accretion at the Southwest Indian Ridge, measuring a crustal production rate of $\sim 20 \text{ km}^3/\text{year}$, providing an upper bound on mantle melt mass flux. Assuming an upper mantle density of $\sim 3300 \text{ kg/m}^3$, a crustal production rate range of $5 - 20 \text{ km}^3/\text{year}$ corresponds to a mantle melt mass flux of $(1.3 \pm 0.8) \times 10^6 \text{ kg/s}$ (values below zero not permitted). The amount of heat transported by the melt has been estimated to be $\sim 1.5 \text{ TW}$ (Nakagawa & Tackley, 2012).

Earth’s magnetic field is generated by convection in its liquid iron outer core. Core convection is primarily generated compositionally through the solidification of the solid iron inner core, where light elements are rejected into the liquid, thereby decreasing the local density of the fluid and driving fluid flow. The radius of the inner core has been measured to high precision by seismographs, yielding $R_{\text{IC}} = 1221 \pm 1 \text{ km}$ (Dziewonski & Anderson, 1981).

Key measurements of the terrestrial magnetic field include the dipolar magnetic moment, which quantifies the strength of the core field, and magnetopause radius, which is the boundary between the terrestrial magnetic field and that of the solar wind. Measurements of Earth’s magnetic field suggest a terrestrial magnetic moment of 80 ZAm^2 (de Pater & Lissauer, 2010; Kivelson & Bagenal, 2014), which we define as 1 Earth unit. Similarly, using the standoff distance equation from Grießmeier et al. (2004), Driscoll & Bercovici (2013) calculate a magnetopause radius of ~ 9.1 Earth radii (Driscoll & Bercovici, 2013), which we also define as 1 Earth unit. This calculated value is consistent with previous measurements from geosynchronous satellites which suggested a plausible magnetopause radius range between $9\text{--}10 R_{\oplus}$ (Shue et al., 1998). While the values for both the magnetic field moment and magnetopause radius lack formal uncertainty estimates, we adopt conservative validation tolerances of ± 0.015 Earth units for the magnetopause radius and ± 0.05 Earth units for magnetic moment. These $\leq 5\%$ tolerances effectively account for known variability in these quantities: the location of the magnetopause varies with solar wind conditions, while the magnetic dipole moment exhibits secular variation.

2 Methods

Our fully-coupled model builds on VPLanet (Barnes et al., 2020; Garcia et al., 2026) by incorporating published models of the carbon cycle (Foley, 2015), the deep water cycle (Seales & Lenardic, 2020), the surface hydrological cycle (Driscoll & Bercovici, 2013), and climate (Wogan et al., 2025b) for planets with plate tectonics. Our abiotic Earth model components are summarized in the schematic in Figure 2.1, with a more detailed schematic of CO_2 and H_2O cycling between the interior, surface, and atmosphere shown in Figure 2.2.

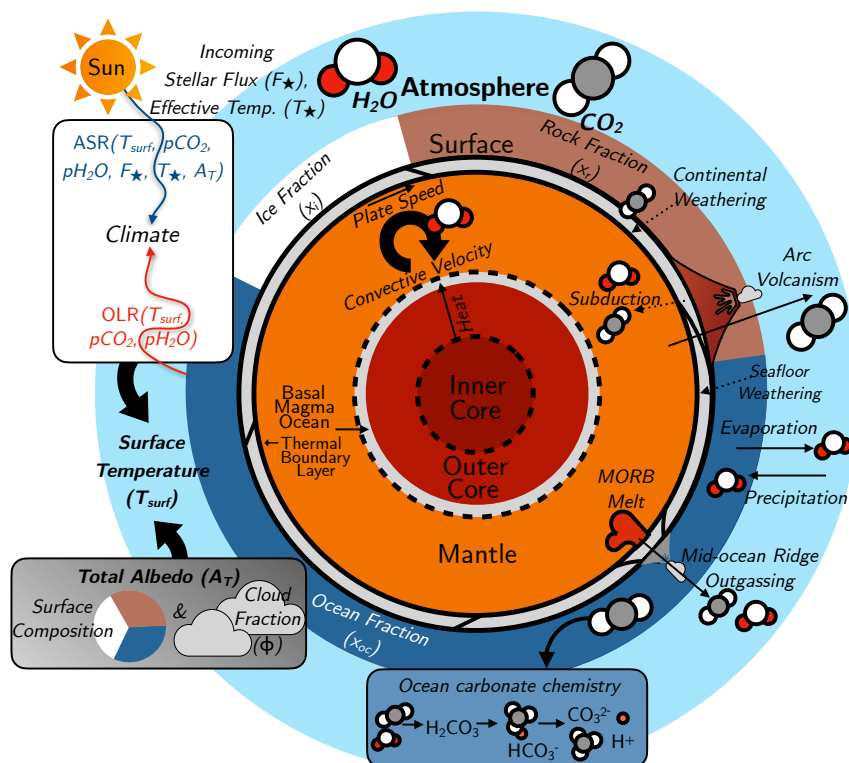


Figure 2.1: A schematic of the key processes included in our abiotic Earth model. Remotely observable quantities (surface temperature, albedo, and atmospheric H_2O and CO_2) are in bold. Volatile cycling is diagrammed in more detail in Figure 2.2.

In this section we describe the assumptions and governing equations used to upgrade VPLanet’s ThermInt module. Any equations and physical parameters not detailed here are implemented as described by Barnes et al. (2020, Appendix K). Model constants are defined in the Appendix.

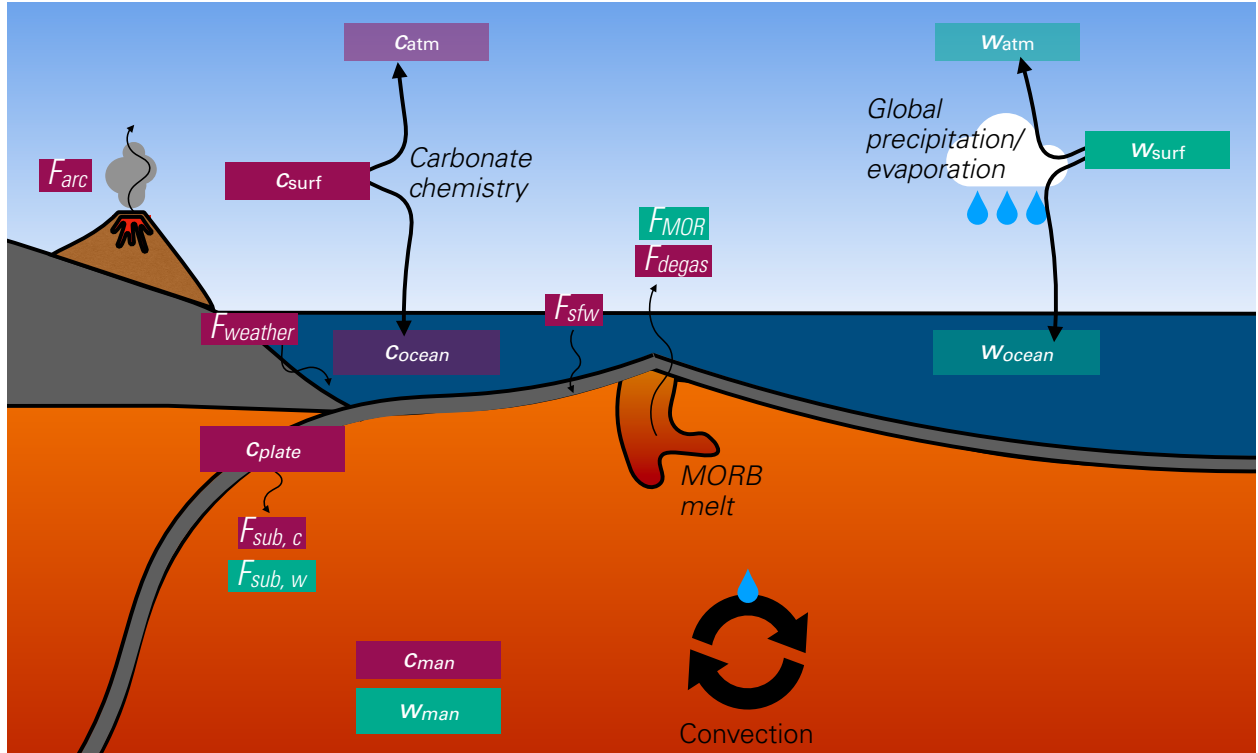


Figure 2.2: A schematic showing the atmospheric sources and sinks that define the new volatile cycling processes in VPLANET’s ThermInt module, inspired by Figure 1 in [Foley \(2015\)](#). Fluxes related to CO_2 are color-coded magenta, and fluxes related to H_2O are color-coded teal. We incorporate the climate-stabilizing carbon cycle from [Foley \(2015\)](#), tracking the mass of CO_2 across the mantle (c_{man}), plate (c_{plate}), and surface (c_{surf}) reservoirs. We include the deep water cycle from [Seales & Lenardic \(2020\)](#), and a precipitation/evaporation parameterization from [Driscoll & Bercovici \(2013\)](#), which allow us to track the mass of H_2O across the mantle (w_{man}) and surface (w_{surf}) reservoirs. At each timestep, surface carbon and water are partitioned into the atmosphere (c_{atm} , w_{atm}) and the ocean (c_{ocean} , w_{ocean}) by enforcing equilibrium at the air-water boundary. This schematic does not depict modeled ocean chemistry.

2.1 Model Assumptions

We make several simplifying assumptions about interior and atmosphere processes for model tractability. For the interior, we assume that the model is initialized immediately following the solidification of the magma ocean. Since this brief phase (1–100 Myr) likely dictated the initial composition of the atmosphere (Elkins-Tanton, 2008; Hamano et al., 2013; Lammer et al., 2018; Stueeken et al., 2020; Chao et al., 2021; Carone et al., 2025), we choose a set of appropriate initial conditions that apply after the bulk of the mantle has solidified.

Furthermore, we assume active plate tectonics throughout Earth’s history, consistent with Earth-moon geology that suggests plate tectonics began as early as the middle Hadean (Maruyama et al., 2018). A number of alternative hypotheses for Earth’s tectonic evolution exist. For example, Condie et al. (2016) invoke a transition from a stagnant lid mode to modern plate tectonics during the Archean to explain evidence that the mantle has not cooled constantly over time (Herzberg et al., 2010). Additionally, Al Asad & Lau (2024) suggest that a transition occurred from sluggish to active lid convection, and Korenaga (2006) argue that plate velocity slowed due to compositional buoyancy and lithospheric stiffening. However, given that we validate our model only against properties of the pre-industrial Earth rather than the early Earth, assessing the tectonic regime during the Hadean and Archean is currently outside the scope of this work.

For the atmosphere, since abiotic methane sources generate fluxes several orders of magnitude lower than biological production (Etiope & Sherwood Lollar, 2013; Wogan et al., 2020; Thompson et al., 2022), we focus our model on CO₂ and H₂O as the primary greenhouse gases regulating climate on an Earth without life, with a detailed discussion of more reduced species in Section 4.2. Additionally, given that the present-day rate of hydrogen escape from our atmosphere is slow (Tian et al., 2013; Gunell et al., 2018; Gronoff et al., 2020) and that most of our atmospheric water vapor inventory is kept close to the surface by the cold trap (Kasting et al., 1984; Gronoff et al., 2020), we neglect atmospheric escape in our model.

We also assume a constant 1 bar of N_2 throughout the Earth’s evolution. This assumption follows prior work (Krissansen-Totton et al., 2021) and is supported by isotopic analysis of 3.5-billion-year-old hydrothermal quartz suggesting that the atmospheric N_2 reservoir has changed little over Earth’s history (Marty et al., 2013).

Finally, we acknowledge that due to large uncertainties in early Earth conditions, we validate the model only against properties of the pre-industrial Earth. We therefore use observed pre-industrial continental weathering rates as our reference, acknowledging that these rates likely reflect some degree of biological enhancement by land vegetation and soil microbial communities. However, the quantitative contribution of biology to silicate weathering rates remains poorly constrained, with field estimates of the biotic enhancement factor spanning nearly two orders of magnitude (e.g., Dessert et al., 2003; Cochran & Berner, 1996), and no true abiotic control site exists on Earth today (Wild et al., 2022). We discuss addressing these uncertainties in Section 4.

2.2 Core, Mantle, and Basal Magma Ocean

We use the `VPLanet` core-mantle interior module, `ThermInt`, and the radiogenic heating module, `RadHeat`, as described by Barnes et al. (2020). We make two key modifications to the nominal model: (1) a basal magma ocean (BMO) layer, which treats radiogenic heating contributions from both the core and the BMO in one combined term prior to present-day solidification, and (2) a mantle viscosity influenced by the water content in the mantle as described in Garcia et al. (2026). The mantle water content is controlled by the deep water cycle and is described in Section 2.4.

Earth’s mantle may have solidified from the middle out (Boukaré et al., 2025), where a prolonged basal magma ocean (BMO) would have solidified slowly and become enriched in incompatible elements, including radiogenic species of K, Th, and U (Labrosse et al., 2007). With potential enrichment as high as ~ 200 times the bulk silicate Earth (Labrosse et al., 2007), the BMO may have provided a stable, thermally insulating layer that slowed core

cooling. However, the existence of the BMO is not yet widely accepted (Jackson et al., 2010; Campbell & O'Neill, 2012; Carlson et al., 2014). For example, Campbell & O'Neill (2012) note that the measured heat flux from upper mantle plumes is inconsistent with a radioactively-enriched layer in the lower mantle. To capture the thermal effect of an enriched BMO we define a new effective BMO + core radiogenic heating rate, $Q_{\text{rad,BMO}+C} = Q_{\text{rad,BMO}} + Q_{\text{rad,C}}$ that folds in the radiogenic heating of the BMO. With this modification the core energy balance in `ThermInt` becomes:

$$-\dot{T}_C = \frac{Q_{\text{rad,BMO}+C} - Q_{\text{BMO}}}{M_{CC}c_C + \frac{dM_{IC}}{dT_C}(L_H + E_G)}, \quad (2.1)$$

where M denotes mass, c_C denotes the core specific heat capacity, $\dot{T}$ denotes the evolving reservoir temperature, and Q denotes heat flow. The subscripts indicate the relevant reservoir, where C , CMB , and BMO represent the core, core-mantle-boundary, and basal magma ocean layer, respectively. Thus, $Q_{\text{rad,BMO}}$ represents the radiogenic heat production in the BMO, and $Q_{\text{rad,C}}$ represents the radiogenic heat production in the core. Additionally, $\frac{dM_{IC}}{dT_C}$ is the derivative of the inner core mass with respect to core temperature, and L_H and E_G give the latent and gravitational energy released at the inner core boundary per unit mass. For the full derivation, see the Appendix.

This approach approximates the effect of a BMO enriched in radiogenic elements without modeling its solidification. In reality, as the mantle gradually solidifies downward towards the core and the liquid BMO shrinks, it becomes further enriched in radiogenic elements (Labrosse et al., 2007). This effectively maintains a higher concentration of radiogenic elements outside the core and provides a stable thermally insulating layer.

To match estimates of upper mantle heat flow within uncertainties, we find that our model requires approximately 3.5 TW of combined core and lower mantle radiogenic power by 4.5 Gyr. Given that core radiogenic heating today is thought to be significantly lower than this amount (Watanabe et al., 2014; Blanchard et al., 2017; Faure et al., 2020), the bulk of the heating must come from the BMO. To assess whether this excess heating is physically

plausible we calculate the total present-day heat budget from a thin, enriched BMO layer.

Labrosse et al. (2007) estimate that if the early Earth's magma ocean gradually solidified downward toward the core, incompatible radiogenic elements would be expelled from solid mantle minerals and concentrated in the remaining liquid BMO. They find the primordial BMO could be enriched in the main radioactive elements (^{40}K , ^{232}Th , ^{235}U , and ^{238}U) by 200 times the concentrations in the bulk silicate Earth (see their Figure 4). Assuming the BMO is a spherical shell with thickness 4 km and that it is slightly denser than the solid mantle due to its relative enrichment in iron (6000 kg m^{-3}), we find a present-day combined heating budget of 3.67 TW due to ^{40}K , ^{232}Th , ^{235}U , and ^{238}U . This calculation demonstrates that attributing 3.5 TW of core heating to a radiogenically enriched, thin, long-lived BMO is physically plausible.

2.3 Carbon Cycle

The model conserves the mass of CO_2 over geological time as it cycles between three reservoirs: the mantle (c_{man}), plate (c_{plate}), and surface (c_{surf}), which includes the ocean (c_{ocean}), and atmosphere (c_{atm}) reservoirs. At each time step, we calculate the following fluxes as in Foley (2015): arc volcanic (F_{arc}), continental weathering (F_{weather}), seafloor weathering (F_{sfw}), degassing (F_{degas}), and subduction ($F_{\text{sub,c}}$) (see Figure 2.2). The fluxes are balanced between the planet's CO_2 reservoirs according to the following equations (Foley, 2015):

$$\dot{c}_{\text{plate}} = \frac{F_{\text{weather}}}{2} + F_{\text{sfw}} - F_{\text{sub,c}}, \quad (2.2)$$

$$\dot{c}_{\text{man}} = \chi_{\text{sub,c}} F_{\text{sub,c}} - F_{\text{degas}}, \quad (2.3)$$

where $\chi_{\text{sub,c}}$ is a factor that represents the subduction efficiency. The continental weathering flux is divided by a factor of 2 to account for the fact that half of the carbon initially sunk into the continental plate is returned to the atmosphere when carbonates form on the seafloor

(Berner et al., 1983; Foley, 2015). Finally, we have

$$\dot{c}_{\text{surf}} = F_{\text{arc}} + F_{\text{degas}} - \frac{F_{\text{weather}}}{2} - F_{\text{sfw}}, \quad (2.4)$$

where by mass conservation $c_{\text{surf}} = c_{\text{atm}} + c_{\text{ocean}}$ (see Section 2.6 for a detailed description of partitioning CO_2 into the atmosphere and ocean).

We define the fluxes according to Foley (2015), with some minor changes. The subduction flux is given by

$$F_{\text{sub}} = 2c_{\text{plate}}vL \times \frac{d_{\text{melt}}}{V_{\text{man}}}, \quad (2.5)$$

where v is the speed of the tectonic plates (see below), L is the present-day length of the ocean trenches, d_{melt} is the depth to the base of the upper mantle melt region (see Barnes et al. (2020, Appendix K)), and V_{man} is the volume of the mantle. The $d_{\text{melt}}/V_{\text{man}}$ term represents division by the effective area of the mantle melt. The arc volcanic flux is subsequently defined as

$$F_{\text{arc}} = (1 - \chi_{\text{sub,c}})F_{\text{sub}}, \quad (2.6)$$

where $1 - \chi_{\text{sub,c}}$ is the fraction of subducting carbon that degasses. The degassing flux is therefore given by

$$F_{\text{degas}} = 2f_{\text{UM}}\chi_{\text{degas,c}}c_{\text{man}}\frac{vLd_{\text{melt}}}{V_{\text{man}}}, \quad (2.7)$$

where f_{UM} is the upper mantle melt fraction, $\chi_{\text{degas,c}}$ is the fraction of mantle melt that degasses, and V_{man} is the volume of the mantle. We include the additional f_{UM} term to ensure that the degassing flux is tightly coupled to the size of the melt layer. As in Foley (2015), we assume that the length of the ocean ridges is equivalent to that of the trenches. We note that the upper mantle melt fraction describes partial melting of the upwelling mantle, and

not the whole-mantle volumetric melt fraction sometimes used in models that parameterize the solidification of the magma ocean (e.g., [Elkins-Tanton, 2008](#); [Krissansen-Totton et al., 2021](#)).

The arc and ridge degassing fluxes are the CO_2 sources in the surface (atmosphere + ocean). Both terms include efficiency factors $\chi_{\text{sub,c}}$ and $\chi_{\text{degas,c}}$. We calibrate these factors to give a total carbon outgassing flux similar to present-day observations of $\sim 1\text{--}5 \text{ Tmol/yr}$ ([Kerrick, 2001](#)), as well as PIE $p\text{CO}_2$ values.

Prior to subduction, seafloor and continental weathering deposit carbon into the oceanic plate. [Foley \(2015\)](#) gives the seafloor weathering flux as

$$F_{\text{sfw}} = F_{\text{sfw},0} \left(\frac{v}{v_0} \right) \left(\frac{x_{oc}}{x_{oc,0}} \right) \left(\frac{p\text{CO}_2}{p\text{CO}_{2,0}} \right)^{\alpha_1}, \quad (2.8)$$

where $F_{\text{sfw},0}$ is the present-day seafloor weathering flux, v_0 is the present-day speed of the tectonic plates, x_{oc} gives the ocean surface fraction, and $x_{oc,0}$ represents the present-day ocean surface fraction of 0.71. Lastly, $p\text{CO}_2$ is the partial pressure of CO_2 , $p\text{CO}_{2,0}$ is its present-day value, and α_1 quantifies the dependence of basalt carbonation on atmospheric CO_2 .

Finally, [Foley \(2015\)](#) derives a continental weathering flux that accounts for supply-limited behavior at high $p\text{CO}_2$ such that

$$\begin{aligned} F_{\text{weather}} = F_{\text{weather},s} \times & \left\{ 1 - \exp \left[- \frac{F_{\text{weather},0}}{F_{\text{weather},s}} \frac{x_r}{x_{l,0}} \left(\frac{v}{v_0} \right)^{\alpha_2} \right. \right. \\ & \times \left(\frac{p\text{CO}_2}{p\text{CO}_{2,0}} \right)^{\beta_1} \left(\frac{P_{\text{sat}}}{P_{\text{sat},0}} \right)^{a_1} \\ & \left. \left. \times \exp \left(\frac{E_a}{R_g} \left(\frac{1}{T_{\text{surf},0}} - \frac{1}{T_{\text{surf}}} \right) \right) \right] \right\}, \end{aligned} \quad (2.9)$$

where $F_{\text{weather},s}$ is the supply limit to weathering, $F_{\text{weather},0}$ is the present-day silicate weathering rate, β_1 is the partial pressure of CO_2 ($p\text{CO}_2$) scaling parameter for silicate weathering, P_{sat} is the saturation vapor pressure, $P_{\text{sat},0}$ is the present-day saturation vapor pressure, a_1 is the saturation vapor pressure scaling parameter for silicate weathering, E_a is the activation

energy of the weathering reaction, R_g is the universal gas constant, T_{surf} is the surface temperature (see Section 2.8), and $T_{\text{surf},0}$ is the present-day surface temperature. The equation also includes a term for scaling the evolving exposed rock fraction, x_r , at a given time step by the present-day land fraction, $x_{l,0}$. Similar to the seafloor weathering flux equation, we incorporate a scaling factor related to the present-day plate speed raised to the power of α_2 to account for the strength of the dependence of weathering on plate speed. The treatment of evolving surface composition is described in further detail in Section 2.7.

The seafloor and continental weathering fluxes strongly influence the $p\text{CO}_2$ evolution in our model. Though there is not an explicit dependence on surface temperature in the seafloor weathering parameterization, it is implicitly dependent on surface temperature via the $p\text{CO}_2$ and plate speed terms.

The supply limit to weathering represents the maximum achievable flux given the available weatherable rock on the surface of the planet. We calculate the supply limit using the relation $F_{\text{weather},s} = (A_{\text{Earth}} x_r E_{\text{max}} f_{cc} \rho_r) / \bar{m}_{cc}$, where E_{max} is the maximum erosion rate, f_{cc} is the fraction of Mg, Ca, K, and Na in the continental crust, ρ_r is the density of the regolith, and $\bar{m}_{cc}$ is the average molar mass of Mg, Ca, K, and Na (Foley, 2015). Mg, Ca, K, and Na are relevant as these are the weatherable components of felsic continental crust (West, 2012). This relation gives the supply-limit in units of mol/s, which we convert to kg/s by multiplying by the molar mass of CO_2 , $\bar{m}_c$. The saturation vapor pressure is related to its present-day value and the surface temperature of the planet according to the relation

$$P_{\text{sat}} = P_{\text{sat,ref}} \exp \left\{ \left[-\frac{\bar{m}_w L_w}{R_g} \left(\frac{1}{T_{\text{surf}}} - \frac{1}{T_{\text{sat,ref}}} \right) \right] \right\}, \quad (2.10)$$

where $\bar{m}_w$ is the molar mass of water, L_w is the latent heat of water, and $T_{\text{sat},0}$ is the reference saturation vapor temperature (Driscoll & Bercovici, 2013).

The fluxes F_{arc} , F_{sfw} , F_{degas} , and $F_{\text{sub,c}}$ are related to the speed of the tectonic plates, v . We implement the Seales & Lenardic (2020) parameterization, which relates v to the mantle Rayleigh number:

$$v = \frac{a_2 \kappa}{2(R_{\text{man}} - R_c)} \left(\frac{Ra}{Ra_{\text{crit}}} \right)^{2\beta_2}, \quad (2.11)$$

where a_2 and β_2 are scaling terms, κ is the thermal diffusivity, R_{man} is the radius of the mantle, R_c is the radius of the core, Ra is the mantle Rayleigh number (see [Barnes et al. \(2020\)](#), Appendix K), and $Ra_{\text{crit}} = 660$ is the critical Rayleigh number.

2.4 Water Cycle

We track the mass of water as it cycles between the mantle (w_{man}) and the surface (w_{surf}), which includes the ocean (w_{ocean}), and the atmosphere (w_{atm}), via a subduction flux ($F_{\text{sub,w}}$) and a mid-ocean ridge degassing flux (F_{MOR}). The fluxes balance according to the following system of differential equations:

$$\dot{w}_{\text{man}} = F_{\text{sub,w}} - F_{\text{MOR}}, \quad (2.12)$$

and

$$\dot{w}_{\text{surf}} = F_{\text{MOR}} - F_{\text{sub,w}}, \quad (2.13)$$

where mass conservation requires that $w_{\text{surf}} = w_{\text{ocean}} + w_{\text{atm}}$.

The abundance of water in the mantle impacts the viscosity of the upper mantle and thus the efficiency of heat transfer via the Arrhenius viscosity. [Garcia et al. \(2026\)](#) give the Arrhenius component of the viscosity as

$$\nu_{\text{ref}} \times \exp \left\{ \left[\frac{E_{a,\text{man}}}{T_{\text{UM}} R_g} - \frac{\Delta E_{\text{H}_2\text{O}}(w_{\text{man}}/M_{\text{man}})}{T_{\text{UM}} R_g} \right] \right\}, \quad \nu_{\text{Arr}} = \quad (2.14)$$

where ν_{ref} is the reference viscosity of the upper mantle, $E_{a,\text{man}}$ is the viscosity activation energy of the upper mantle, $\Delta E_{\text{H}_2\text{O}}$ is the depression of mantle viscosity activation energy due

to water, M_{man} is the mass of the mantle, and T_{UM} is the temperature of the upper mantle. The viscosity of the upper mantle is then calculated by dividing the Arrhenius viscosity term by the upper mantle viscosity melt factor (Garcia et al., 2026). The melt viscosity factor is calculated following Driscoll & Barnes (2015). First, we define the normalized melt fraction:

$$\Phi = \frac{f_{\text{UM}}}{\phi_s}, \quad (2.15)$$

where f_{UM} is the upper mantle melt fraction and $\phi_s = 0.8$ is the reference melt fraction parameter. Next, we calculate the melt geometry factor:

$$F = (1 - \xi) \cdot \text{erf} \left[\frac{\sqrt{\pi}}{2(1 - \xi)} \Phi (1 + \Phi^\gamma) \right], \quad (2.16)$$

where the viscosity-melt reduction coefficient $\xi = 5 \times 10^{-4}$, the viscosity-melt reduction exponent $\gamma = 6.0$, and erf is the error function. Finally, the melt viscosity reduction factor is given by:

$$\eta_{\text{melt}} = \frac{1 + \Phi^\delta}{(1 - F)^{B\phi_s}}, \quad (2.17)$$

where the viscosity-melt reduction exponent $\delta = 6.0$ and the viscosity-melt reduction parameter $B = 2.5$ (Costa et al., 2009).

To calculate the fluxes of water into and out of the mantle, we assume the definitions of $F_{\text{sub,w}}$ and F_{MOR} given by Seales & Lenardic (2020) so that

$$F_{\text{sub,w}} = f_h \rho_{\text{man}} D_{\text{hydr}} S \chi_{\text{regas,w}}, \quad (2.18)$$

where f_h is the mass fraction of water in the serpentinized layer, ρ_{man} is the density of the mantle, D_{hydr} is the thickness of the serpentinized layer, S is the spreading rate of the tectonic plates, and $\chi_{\text{regas,w}}$ is the regassing efficiency factor. The thickness of the serpentinized layer is related to surface temperature T_{surf} according to the relation $D_{\text{hydr}} = (k |T_{\text{surf}} - 973|) / q_m$,

where k is the thermal conductivity and q_m is the heat flux through the upper mantle (see [Barnes et al. \(2020\)](#), Appendix K). Note that this parameterization of the subduction flux assumes that there is always a sufficient supply of water to drive the serpentinizing reaction. The spreading rate is related to the speed of the plates as $2Lv$, where the factor of $2L$ represents bidirectional spreading.

The mid-ocean ridge degassing flux is thus given by

$$F_{\text{MOR}} = \rho_{\text{man}} f_{\text{UM}} X_{\text{melt}} d_{\text{melt}} S \chi_{\text{degas,w}}, \quad (2.19)$$

where X_{melt} is the abundance of water in the erupting magma and $\chi_{\text{degas,w}}$ is the degassing efficiency factor. The parameters f_{UM} and d_{melt} are calculated as described by [Barnes et al. \(2020, Appendix K\)](#), while the fraction of water in the melt can be found by taking $X_{\text{melt}} = \chi_m / [D_{\text{H}_2\text{O}} + f_{\text{UM}}(1 - D_{\text{H}_2\text{O}})]$, where χ_m is the mass fraction of water in the mantle ($w_{\text{man}} / M_{\text{man}}$) and $D_{\text{H}_2\text{O}}$ is the bulk distribution coefficient ([Seales & Lenardic, 2020](#)).

The mass of H_2O on the surface (atmosphere + ocean) results from the balance of the mid-ocean ridge degassing sources and the subduction sinks at the plate boundaries. In our model, both terms include degassing and regassing efficiency factors, $\chi_{\text{degas,w}}$ and $\chi_{\text{regas,w}}$, respectively. We tune these factors to give the PIE mass of water in the oceans and atmosphere after 4.5 Gyr, and to give a degassing flux on the order of present-day observations, $\sim 11 \text{ Tmol/yr}$ ([Fischer, 2008](#)).

2.5 Atmospheric Constituents

Our Earth-like atmospheres consist of N_2 , H_2O , and CO_2 . As in previous work ([Krissansen-Totton et al., 2021](#)), we assume that the atmosphere maintains a constant 1 bar of N_2 , since our 1-bar, abiotic atmospheres lack photosynthetic oxygen. Atmospheric H_2O and CO_2 are allowed to vary based on the balance of sources and sinks from volatile cycling. We do not include atmospheric escape in our model given we aim only to validate the model on the

PIE, when escape rates were likely insignificant.

We neglect CH_4 in our model, as abiotic methane fluxes are insufficient to contribute meaningfully to Earth’s greenhouse warming. Modern terrestrial observations show that abiotic CH_4 sources – including volcanic outgassing, serpentinization, and hydrothermal processes – generate approximately 1.6 Tg/yr (0.1 Tmol/yr) (Etiope & Sherwood Lollar, 2013), while biological sources produce 500–600 Tg/yr (31–37 Tmol/yr) (Saunio et al., 2019), a factor of approximately 300 times larger. Laboratory experiments and thermodynamic modeling support these observational constraints: McCollom (2003, 2016) found that kinetic barriers limit abiotic methane production rates even under thermodynamically favorable conditions, and Wogan et al. (2020) demonstrated that volcanic CH_4 fluxes remain orders of magnitude below biological rates even under highly reducing mantle conditions.

Even on the early, more volcanically active Earth, abiotic methane production would have been substantially lower than microbial methanogenesis. Modern atmospheric methane reaches approximately 1.7 ppmv (Dlugokencky et al., 2011), requiring biological fluxes of 500–600 Tg CH_4 yr^{-1} to persist against photochemical destruction. With abiotic production rates 300 times lower, steady-state methane mixing ratios would remain negligible even with extended atmospheric lifetimes in an anoxic atmosphere (Pavlov et al., 2001), and are thus unlikely to contribute significantly to greenhouse warming compared to CO_2 and H_2O .

In the model framework, we treat both H_2O and CO_2 as combined surface reservoirs. At each time step, we partition each gas between the atmosphere and the ocean by enforcing instantaneous equilibrium. Thus, by mass conservation we require that $c_{\text{surf}} = c_{\text{ocean}} + c_{\text{atm}}$, and $w_{\text{surf}} = w_{\text{ocean}} + w_{\text{atm}}$. Here, we describe the partitioning of water, while the partitioning of CO_2 is described in the following subsection.

To maintain hydrostatic equilibrium at the air-water boundary, at each time step we partition water between the atmosphere and the ocean by initially calculating the mass of water in the atmosphere (w_{atm}) in excess of the amount of moisture the air can hold at a given time step (Driscoll & Bercovici, 2013). Since the Earth’s oceans and atmosphere achieve

equilibrium on the order of days, whereas `ThermInt` processes evolve on geological timescales, we assume the flow of water between the atmosphere and the ocean occurs instantaneously in the model. At each time step, we calculate the saturation vapor pressure (Foley, 2015) as a function of T_{surf} , and we then multiply by a factor of $RH \times (M_{\text{atm,T}}/P_{\text{T}}) \times (\bar{m}_w/w_{\text{atm}}^-)$ to obtain m_{moist} , the maximum mass of water that the air can hold, where RH represents the tuneable global average relative humidity.

When $w_{\text{atm}} > m_{\text{moist}}$, the excess water mass is subtracted from w_{atm} and added to w_{ocean} . In the inverse case when $w_{\text{atm}} < m_{\text{moist}}$, to maintain equilibrium between the air and the ocean, we subtract the mass difference from w_{ocean} and add it to w_{atm} . To obtain $p\text{H}_2\text{O}$ in Pa, we convert w_{atm} to moles of H_2O , divide by the total moles in the atmosphere, and multiply by P_{T} . This simplified 0-D approximation provides the initial atmospheric H_2O partial pressure needed as input to the climate calculation (Section 2.8), where it is subsequently refined using the vertically-resolved atmospheric structure.

2.6 Ocean Chemistry

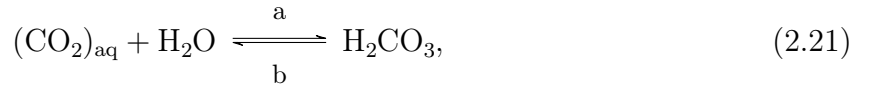
The total reservoir of carbon in the ocean (total dissolved inorganic carbon, or DIC) is the sum of the aqueous $[\text{CO}_2]$, $[\text{CO}_3^{2-}]$, and $[\text{HCO}_3^-]$. Since we enforce equilibrium at the air-ocean boundary, we assume that all ocean chemistry values are representative of surface waters. Initially, atmospheric CO_2 dissociates in ocean water following Henry's law (Pilson, 2012), which quantifies the solubility of a gas according to its partial pressure above a liquid so that,

$$(\text{CO}_2)_{\text{aq}} = H \times p\text{CO}_2, \quad (2.20)$$

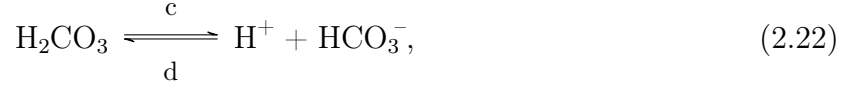
where $(\text{CO}_2)_{\text{aq}}$ gives the aqueous concentration of CO_2 in units of mol/kg H_2O , H is Henry's law coefficient in units of mol/kg/atm, and $p\text{CO}_2$ is the partial pressure of atmospheric CO_2 . We obtain c_{ocean} by multiplying $(\text{CO}_2)_{\text{aq}}$ by the total mass of water in the ocean, w_{ocean} , and

the molar mass of CO_2 . By definition, $c_{\text{atm}} = (p\text{CO}_2/P_{\text{T}}) \times (\bar{m}_c/\bar{m}_{\text{atm}}) \times M_{\text{atm, T}}$, where P_{T} is the total pressure, $\bar{m}_{\text{atm}}$ is the average molar mass of the atmosphere, and $M_{\text{atm, T}}$ is the total mass of the atmosphere. Substituting $p\text{CO}_2 = \text{CO}_{2\text{aq}}/H$ into the mass conservation constraint, we solve a quadratic equation to obtain $(\text{CO}_2)_{\text{aq}}$ based on the integration of Equation (2.4) at each time step. Finally, to obtain $p\text{CO}_2$ in Pa, we take the $(\text{CO}_2)_{\text{aq}}$ solution and divide by the Henry's law coefficient, and convert from atm to Pa.

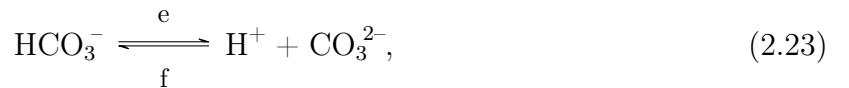
Next, aqueous CO_2 and H_2O combine to make carbonic acid in the following equilibrium reaction:



where the forward reaction (a) is governed by the rate constant k_a , and the reverse reaction (b) is governed by the rate constant k_b . Depending on the ocean pH, carbonic acid readily dissociates into H^+ and bicarbonate in the following equilibrium reaction



with k_c and k_d for the forward and reverse reactions, respectively. Finally, bicarbonate may further dissolve into H^+ and a carbonate ion



where k_e and k_f give the dissolution and combination rate constants, respectively.

As in [Schwieterman et al. \(2019b\)](#) and [Krissansen-Totton et al. \(2021\)](#), we assume that calcite is saturated so that

$$[\text{CO}_3^{2-}] = \frac{\Omega_{\text{cal}} \times K_{\text{sp}}(T_{\text{surf}})}{[\text{Ca}^{2+}]}, \quad (2.24)$$

where Ω_{cal} is the unit-less global ocean calcite saturation factor, K_{sp} is the calcite solubility product, and $[\text{Ca}^{2+}]$ is the concentration of calcite in the oceans in mol/kg. To calculate

solubility self-consistently with surface temperature, we use the temperature and salinity-dependent solubility product formula given by [Millero \(2005\)](#). As in [Schwieterman et al. \(2019a\)](#), for all solubility parameterizations we assume that ocean salinity remains constant at 35 parts per thousand. We also assume that the temperature of the ocean is equivalent to the global average surface temperature, as we assume that the ocean and atmosphere equilibrate faster than our model’s geological timescales.

Using Equation (2.24), a saturation factor, and a calcite concentration, we can calculate the carbonate concentration in the ocean. [Schwieterman et al. \(2019a\)](#) assumed $\Omega_{\text{cal}} = 1$, but calcite may need to be supersaturated ($\Omega_{\text{cal}} \sim 10\text{--}20$) for abiotic precipitation to occur. Furthermore, over Earth’s history, the calcite concentration has varied from $[\text{Ca}^{2+}] = 10^{-2}$ to 3×10^{-1} mol/kg ([Halevy & Bachan, 2017](#)). We therefore tune $\Omega_{\text{cal}}/[\text{Ca}^{2+}]$ to match the observed total dissolved inorganic carbon in the surface oceans for pre-industrial conditions (see Table 2.4).

This tuning approach requires justification, as our model represents an Earth without oxygenic photosynthesis or complex skeletal organisms, analogous to conditions during the Archean and early Proterozoic eons. In the modern ocean, carbonate precipitation is predominantly biogenic ([Chave & Suess, 1970](#); [Bialik et al., 2022](#)), though [Bialik et al. \(2022\)](#) estimate that abiotic aragonite precipitation accounts for approximately 15% of CO_2 efflux in oligotrophic regions where surface waters remain supersaturated with respect to calcium carbonate. However, evidence supports that abiotic carbonate precipitation processes, while kinetically slower than biogenic precipitation, can achieve similar steady-state DIC concentrations over geological timescales. Laboratory experiments show that abiotic calcite precipitates at rates two to five orders of magnitude slower than biotic calcite ([Carpenter & Lohmann, 1992](#)), but given billion-year timescales, kinetic limitations become less constraining.

Furthermore, geological evidence supports widespread abiotic carbonate precipitation in Earth’s early history. The Precambrian geological record shows extensive carbonate platform

deposition (Grotzinger, 1989; Grotzinger & Kasting, 1993; Grotzinger & James, 2000), including meter-thick beds of fibrous calcite and aragonite precipitated directly on the Archean seafloor (Sumner & Grotzinger, 1996). Sumner & Grotzinger (1996) argued that high rates of abiotic precipitation in Archean oceans were facilitated by elevated Fe^{2+} concentrations in anoxic seawater, which paradoxically promoted aragonite precipitation while inhibiting calcite nucleation. This geological evidence demonstrates that abiotic processes can generate substantial carbonate deposits, supporting our assumption that calcite saturation is achievable without biology.

We find that the optimal $\Omega_{\text{cal}}/[\text{Ca}^{2+}]$ to match pre-industrial DIC is 408.1 kg/mol. Recent spatial mapping of global calcite saturation of present-day oceans shows an average $\Omega_{\text{cal}} = 4.3$ (Shaik et al., 2025), and the present-day $[\text{Ca}^{2+}]$ is 10.28 mmol/kg (Halevy & Bachan, 2017), corresponding to a present-day ratio of $\Omega_{\text{cal}}/[\text{Ca}^{2+}] = 398$. Conservatively estimating that a supersaturation factor of $\Omega_{\text{cal}} = 20$ is required for abiotic carbonate precipitation, our calibrated $\Omega_{\text{cal}}/[\text{Ca}^{2+}]$ corresponds to a calcite concentration of 49 mmol/kg. We discuss the possible limitations of this tuning approach in Section 4.3.

We model this chemical system using the following equations adopted from Pilson (2012):

$$k_e = \frac{[\text{H}^+][\text{HCO}_3^-]}{(\text{CO}_2)_{\text{aq}}}, \quad (2.25)$$

and

$$k_f = \frac{[\text{H}^+][\text{CO}_3^{2-}]}{[\text{HCO}_3^-]}. \quad (2.26)$$

Since $k_e/k_f = [\text{HCO}_3^-]^2/([\text{CO}_3^{2-}][\text{CO}_2])$, we can combine these rate constant equations to solve for the concentration of bicarbonate:

$$[\text{HCO}_3^-] = \sqrt{(\text{CO}_2)_{\text{aq}}[\text{CO}_3^{2-}] \times \frac{k_e}{k_f}}. \quad (2.27)$$

Finally, we can solve for the concentration of $[\text{H}^+]$ by taking

$$[\text{H}^+] = \frac{K_{sp}[\text{HCO}_3^-]}{[\text{CO}_3^{2-}]}, \quad (2.28)$$

where $\text{pH} = -\log_{10} [\text{H}^+]$.

2.7 Surface Albedo

The surface albedo of the planet in our model consists of contributions from the ground and atmospheric water vapor clouds. Based on previous work ([Driscoll & Bercovici, 2013](#)), we simulate the evolving total planetary surface albedo A_T by allowing the ground albedo A_g and cloud fraction ϕ to vary with time so that

$$A_T = A_g(1 - \phi) + A_{cl}\phi. \quad (2.29)$$

We assume that our Earth-like water vapor clouds have an albedo of $A_{cl} = 0.89$, which represents an average albedo of thick and thin clouds ([de Pater & Lissauer, 2010](#); [Driscoll & Bercovici, 2013](#)). We assume a constant fraction of the atmospheric H_2O reservoir condenses into clouds, and tune the wavelength-independent cloud opacity pressure to ensure we match the PIE surface temperature at 4.5 Gyr. As in [Driscoll & Bercovici \(2013\)](#), we then calculate the cloud reflectivity fraction using the two-stream Eddington approximation,

$$\phi = \frac{\gamma(1 - e^{-2\beta\tau_{cl}})}{1 - \gamma^2 e^{-2\beta\tau_{cl}}}, \quad (2.30)$$

with $\tau_{cl} = 4800$ Pa as the constant cloud opacity pressure and scattering constants

$$\gamma = \frac{\beta - 2(1 - \bar{w})}{\beta + 2(1 - \bar{w})} \quad (2.31)$$

and

$$\beta = \sqrt{3(1 - \bar{w})(1 - \bar{w}g_w)}, \quad (2.32)$$

where $\bar{w}$ is the single-scattering albedo $1 - \bar{w} = 1 \times 10^{-7}$ and $g_w = 0.74$ is the asymmetry factor (Hashimoto & Abe, 2001). Given this simplified cloud model cannot account for more complex effects like radiative heating due to clouds and wavelength-dependent reflectivity, we tune the cloud opacity pressure to match pre-industrial Earth temperature. Thus, we will *not* attempt to match the pre-industrial Earth cloud fraction with the model.

The ground albedo A_g consists of contributions from surface rock (x_r), ocean (x_{oc}), and ice (x_i) fractions, where

$$A_g = x_r A_r + x_{oc} A_{oc} + x_i A_i. \quad (2.33)$$

We take the albedo of rock to be $A_r = 0.17$ (Kaltenegger et al., 2007; Driscoll & Bercovici, 2013), the albedo of ocean water to be $A_{oc} = 0.1$ (de Pater & Lissauer, 2010; Driscoll & Bercovici, 2013), and the albedo of ice to be $A_i = 0.6$ (Kaltenegger et al., 2007; Driscoll & Bercovici, 2013). We assume a conical ocean basin as in Driscoll & Bercovici (2013), with a slope angle of $\theta_{oc} = 0.06188^\circ$ calibrated to give $x_{oc} = 0.71$ when $w_{\text{ocean}} = 1.4 \times 10^{21}$ kg. To model ice growth, we include the simple relation from Driscoll & Bercovici (2013), which gives $x_i = x_{i,\text{max}} \times (1 - x_{oc}) \times (x_{oc}/x_{oc,0})$, where $x_{i,\text{max}} = 0.12$ is the maximum ice fraction, $(1 - x_{oc})$ is the fraction of land, and $x_{oc,0}$ is the present-day ocean fraction of 0.71.

Previous work has established that habitability requires only some amount of liquid water on the planetary surface (Kasting et al., 1993), and that global average surface temperatures can be meaningfully compared to spatially-resolved three-dimensional climate states when assessing habitability thresholds (Charnay et al., 2013; Arney et al., 2016). Charnay et al. (2013) demonstrated through 3D general circulation modeling that complete glaciation (e.g., a “snowball” Earth) likely does not occur until global average surface temperatures fall below 240 K, owing to spatial heterogeneity in ice coverage and meridional heat transport that

maintains liquid water reservoirs at the equator. We therefore allow $x_i = 1.0$ when the global average surface temperature is ≤ 240 K, though like [Driscoll & Bercovici \(2013\)](#) our model does not currently handle transitions between global glaciation phases and $x_{i,\max}$.

2.8 Climate

Planetary surface temperature is related to surface and atmospheric characteristics as well as the radiation received from the star. We determine surface temperature by calculating thermal equilibrium between the absorbed shortwave radiation (ASR) and the outgoing long-wave radiation (OLR). OLR is a function of surface temperature and the atmospheric abundances of water vapor and CO_2 , while ASR is additionally a function of the total albedo, incoming stellar flux, and effective stellar temperature.

To calculate ASR and OLR, we use the `clima` module of the `Photochem` software package [Wogan et al. \(2025b\)](#). `clima` uses standard two-stream methods ([Toon et al., 1989](#)) to solve the radiative transfer equations, incorporating opacity tables to simulate continuum UV absorption, Rayleigh scattering, and collision-induced absorption, while accounting for line absorption with the correlated-k method. This approach explicitly captures the wavelength-dependent absorption and scattering processes that gray atmosphere models cannot. `clima` has been benchmarked against other community radiative transfer codes including SOCRATES ([Wolf et al., 2022](#)), ExoRT ([Wolf et al., 2022](#)), SMART ([Meadows & Crisp, 1996](#)), and the radiative transfer code used by [Kopparapu et al. \(2013\)](#), demonstrating good agreement across a range of atmospheric compositions and stellar spectra. Furthermore, the code reproduces the observed pressure-temperature profiles of Venus, Earth, Mars, Jupiter and Titan ([Wogan et al., 2025b](#)).

To resolve 1-D atmospheric structure, we use the mode of `clima` that assumes an isothermal stratosphere above a convective troposphere with a pseudo-moist adiabat. The stratospheric temperature is estimated via the skin temperature for an assumed bond albedo, A_b . As in any 1-D climate model, `clima` has a limited ability to self-consistently calculate the radiative

impact of clouds. As a result, we follow previous work (e.g., [Kopparapu et al., 2013](#)) and “paint” them on the surface by incorporating them into the total surface albedo, which is an input to the radiative transfer calculation (see Section 2.7).

Using the stellar evolution module in `VPLanet` ([Baraffe et al., 2015](#)), we calculate the evolving incoming stellar flux ($F_\star$) and temperature ($T_\star$) as a function of time. To determine the top-of-atmosphere solar flux, `clima` calculates a black-body distribution corresponding to $T_\star$. This ensures that ASR accounts for how the evolving Sun impacts wavelength-dependent atmospheric opacity. Rather than calling the full radiative transfer calculation at every timestep, which would be computationally prohibitive for billion-year integrations, we pre-compute a grid of OLR and ASR values as functions of surface temperature, atmospheric CO_2 partial pressure, the total H_2O surface inventory, stellar flux, stellar effective temperature, and surface albedo.

The grid spans surface temperatures from 100 K to 500 K (41 points, spaced at 10 K intervals), CO_2 partial pressures from 10^{-6} to 10^4 bar (21 points, logarithmically spaced at 0.5 dex intervals), H_2O total surface pressures from 10^{-3} to 10^4 bars (15 points, logarithmically spaced at 0.5 dex intervals), stellar flux from 800 to 2000 W/m^2 (13 points, spaced at 100 W/m^2 intervals), surface albedoes of 0.0, 0.1, 0.3, 0.5, 0.7, and 0.9 (6 discrete values), and stellar effective temperatures from 5000 to 6000 K (5 points, spaced at 250 K intervals). At each model timestep, we pass the current values of T_{surf} , $p\text{CO}_2$, total surface water inventory (in bars), stellar flux ($F_\star$), A_{T} , and stellar effective temperature ($T_\star$) to a multi-dimensional linear interpolation routine that returns the corresponding OLR and ASR from the pre-computed grid. The complete grid contains approximately 9.7 million pre-computed atmospheric states.

We then use Brent’s method ([Brent, 2013](#)) to iteratively solve for the equilibrium surface temperature where ASR equals OLR, using a bracket of [100, 600] K and a tolerance of 2×10^{-12} . Once convergence is achieved, we pass the solved T_{surf} back to `VPLanet` to update all temperature-dependent parameters. This grid-based approach maintains the approximate

accuracy of full radiative transfer calculations ($< \sim 1$ K error) while reducing computational cost by approximately three orders of magnitude, enabling efficient exploration of long-term planetary evolution.

In `clima`, we parameterize the H_2O inventory as a pressure column (in bars) representing the total surface water reservoir (w_{surf} converted to an equivalent pressure). `clima` internally partitions the combined surface reservoir between atmospheric vapor (limited by saturation vapor pressure at each vertical level) and condensed phases, capturing the coupled radiative-thermodynamic effects of water vapor feedback. Once a steady-state climate is computed, we extract the atmospheric composition using `clima`'s vertically-resolved solution. We extract the surface water vapor mixing ratio and the total surface pressure from the interpolated grid output, which represent `clima`'s self-consistent partitioning of water between the atmosphere and ocean when accounting for temperature stratification throughout the atmospheric column. This more physically accurate partitioning refines our initial surface-only estimates described in Sections 2.5. We recalculate w_{atm} , w_{ocean} , c_{atm} and c_{ocean} to match `clima`'s atmospheric state, ensuring that the water distribution used in subsequent carbon cycle and thermal evolution calculations remains consistent with the radiative transfer solution.

2.9 Reflected Light Spectra

We use the `PICASSO` code (Batalha et al., 2019) to compute synthetic HWO reflected light spectra ($0.2\text{-}2.0\ \mu\text{m}$) for the true Earth and the abiotic Earth's atmospheric state after 4.5 Gyr. We use a set of $R = 15,000$ resampled opacities archived on Zenodo (Wogan et al., 2025a), derived from a similar set of HITEMP (Rothman et al., 2010) and HITRAN (Gordon et al., 2017) opacities used by `clima` for climate calculations (Section 2.8). The opacity database accounts for line absorption from the following relevant species: CO_2 , H_2O , O_2 , O_3 , and CH_4 . Relevant CIA partners include $\text{CO}_2\text{-CO}_2$, $\text{H}_2\text{O-H}_2\text{O}$, $\text{H}_2\text{O-N}_2$, $\text{N}_2\text{-N}_2$, $\text{N}_2\text{-O}_2$, and $\text{O}_2\text{-O}_2$. `PICASSO` also includes Rayleigh scattering for all specified molecules.

We compute two reflected light spectra: (1) a realistic modern Earth spectrum using

temperature and mixing ratio profiles from the Intercomparison of Radiation Codes in Climate Models (ICRCCM) case 62, which represents the averaged mid-latitude Earth during the summer months (Lincowski et al., 2018), and (2) an abiotic Earth spectrum using the `vPLanet` atmospheric state after 4.5 Gyr, which includes only CO_2 and H_2O as greenhouse gases along with 1 bar of N_2 . For both spectra, we input an expected surface albedo of 0.14 (without clouds) and include a cloud deck. We model clouds as a single gray deck extending from 0.6 to 0.7 bar with an optical depth of 10, a single-scattering albedo of 0.99, and an asymmetry parameter of 0.85. We compute both cloudy and cloud-free spectra and weight them by a cloud fraction of 50% to approximate Earth’s observed cloud coverage. Both calculations assume solar illumination at quadrature (90° phase angle) and are binned to $R = 140$.

The ICRCCM Earth reference profile has been re-interpolated onto the `vPLanet` pressure grid to enable direct comparison. The `vPLanet` grid has a surface pressure of 1.016 bar compared to the ICRCCM value of 0.989 bar, representing a difference of approximately 0.03 bar, or 3% higher surface pressure. The top-of-atmosphere (TOA) extends to 1.1×10^{-5} bar in the `vPLanet` grid versus 1×10^{-7} bar in the original ICRCCM profile. The number of atmospheric layers differs as well, with 51 layers in the `vPLanet` grid compared to 63 layers in the original ICRCCM profile. These differences are minor – the 3% surface pressure difference causes negligible changes in atmospheric column density, and the higher TOA cutoff removes only the uppermost stratosphere and mesosphere, which are optically thin and contribute minimally to the spectrum. The temperature-pressure profile, water vapor distribution, and other trace gas abundances have been linearly interpolated in log-pressure space to preserve the atmospheric structure of the original ICRCCM profile.

2.10 Initial Conditions

Our model requires the initialization of key parameters, including those describing the initial conditions of the interior and the surface volatile inventories. Here we briefly describe how

we initialized our model at 50 Myr; i.e., following the magma ocean phase. Most other initial values are calibrated to match the relevant measurements of the PIE. Though the model evolution is largely insensitive to the initial CO₂ and H₂O inventories, we base the total initial masses of volatiles on previous work.

To compute a plausible evolutionary model for Earth, we optimized 16 initial conditions and constants of the interior (given in Table 2.2) to reproduce Earth’s present-day interior properties (given in Table 2.1). To perform the optimization, we assumed a Gaussian likelihood function based on the mean and uncertainty of each output parameter, and minimized the negative log-likelihood using the Nelder-Mead simplex algorithm (Nelder & Mead, 1965) implemented in SciPy (Jones et al., 2001). Due to the multi-modal nature of likelihood space, we initialized the optimization at 200 different starting locations uniformly sampled over a wide parameter range, and selected the maximum likelihood fit out of all runs that converged.

As in Seales & Lenardic (2020) we initialize our model with 2 terrestrial oceans (TOs) of water. From Sleep & Zahnle (2001), we assume that Earth’s total initial CO₂ inventory is 1.1×10^{21} kg. Initially, to partition the volatiles between the interior and surface reservoirs, we follow estimations of the primordial Earth atmosphere immediately following the solidification of the magma ocean. Elkins-Tanton (2008) predicts primordial Earth CO₂–H₂O atmospheres ranging from 90–3350 bars. Consistent with the lower end of these model estimates, we initialize the surface with 5 bars of CO₂ and 530 bars of H₂O (2 T.O.) for a combined primordial atmosphere of 535 bars.

Though the magma ocean phase may conclude with a higher surface inventory of CO₂, through initial model testing we found that the $p\text{CO}_2$ evolution is largely independent of the initial atmospheric carbon inventory. This model behavior occurs because our seafloor weathering flux (Equation [2.8]) is assumed to be directly dependent on plate speed and $p\text{CO}_2$ relative to the present-day, which may be very large immediately following the magma ocean phase, though primordial Earth conditions remain uncertain. As a result, basaltic

seafloor weathering is found to be efficient enough to draw nearly all of the CO_2 out of the atmosphere. Therefore, we initialize the model with 5 bars to maximize model stability – all CO_2 reservoirs must initially contain non-zero mass, and the remaining CO_2 is split between the plate and the mantle.

In summary, our model couples interior thermal evolution (`ThermInt`, `RadHeat`), volatile cycling (carbon and water), ocean chemistry, and climate (`Clima`) into a unified framework. We track CO_2 and H_2O masses across mantle, ocean, and atmospheric reservoirs with fluxes controlled by plate tectonics and weathering. Key model features include a radiogenically enriched BMO, water-dependent mantle viscosity, calcite-saturated ocean chemistry, and evolving planetary surface albedo from ground and cloud contributions. We initialize the model as described above, then evolve the system forward 4.5 Gyr, optimizing 17 interior parameters to match 19 pre-industrial Earth properties.

3 Results

In this section, we show that the whole-planet model successfully reproduces the pre-industrial Earth after 4.5 Gyr of model evolution. We also generate a corresponding reflected light spectrum to compare against that of a realistic, cloudy Earth. In Table 2.2 we report the calibrated interior parameter values following our optimization procedure described in Section 2.10, and model results and a comparison to measured values within uncertainties are summarized in Table 2.3.

In Figure 2.3, we show the incoming solar flux at the planet as a function of time. Using the Baraffe et al. (2015) model, the evolving solar flux starts at 943 W/m^2 ($\sim 0.7L_\odot$) and reaches 1370 W/m^2 ($1.0L_\odot$) by 4.5 Gyr. The final incoming stellar flux is thus consistent with its present-day value. Furthermore, the early stellar flux is consistent with predictions of other stellar evolution models used to model the faint young Sun (Sagan & Mullen, 1972; Newman & Rood, 1977; Gough, 1981).

Parameter	Value
Initial ^{40}K power in the mantle	$5.14 \times 10^{13} \text{ W}$
Initial ^{40}K power in the core	$4.05 \times 10^{13} \text{ W}$
Initial ^{232}Th power in the mantle	$7.01 \times 10^{12} \text{ W}$
Initial ^{232}Th power in the core	$1.30 \times 10^{11} \text{ W}$
Initial ^{238}U power in the mantle	$1.04 \times 10^{13} \text{ W}$
Initial ^{238}U power in the core	$1.0 \times 10^{11} \text{ W}$
Initial ^{235}U power in the mantle	$1.66 \times 10^{13} \text{ W}$
Initial ^{235}U power in the core	$4.68 \times 10^{11} \text{ W}$
Initial mantle temperature	2700 K
Initial core temperature	5700 K
Melt eruption efficiency	0.01 [n.d.]
Core liquidus depression	$4.04 \times 10^{-4} \text{ K}$
Lower/upper mantle viscosity ratio	1.5 [n.d.]
Depression of mantle viscosity activation energy due to water	$9.642 \times 10^6 \text{ K}$
Viscosity activation energy	$2.64 \times 10^5 \text{ J mol}^{-1}$
Upper mantle thermal conductivity	4.24 W/m/K

Table 2.2: Initial conditions and constants of the interior Earth model. Core radiogenic power includes BMO contribution (see Section 2.2).

Figure 2.4 shows the evolution of bulk planetary parameters including surface temperature, total albedo (ground and clouds), atmospheric $p\text{H}_2\text{O}$, and atmospheric $p\text{CO}_2$. By 4.5 Gyr, the planet surface temperature reaches 286.9 K (a), the ground albedo (solid line) is 0.14 while the total albedo including clouds (dashed line) is 0.198 (b), the atmospheric $p\text{H}_2\text{O}$ is 1632 Pa with an assumed relative humidity of 100% (c), and the atmospheric $p\text{CO}_2$ is 28.2 Pa (d). The atmospheric water vapor abundance (c) shows a strong relationship with planetary temperature (a) due to the efficiency of OLR absorption and strong temperature-dependence of saturation vapor pressure. After 4.5 Gyr, our model suggests that the total degassing flux of CO_2 is 1.2 Tmol/yr and that of H_2O is 10 Tmol/yr – both are reasonably close to the present-day observations of 1–5 and 11 Tmol/yr, respectively.

At the start of the evolution, following a brief (~ 100 Myr) equilibration period, the planet is relatively cold, with $T_{\text{surf}} = 272 \text{ K}$, which is still well above the 240 K threshold for complete glaciation (Section 2.7). In response, the weathering is low while the degassing flux is high, allowing the atmospheric CO_2 abundance to spike to ~ 1 bar 100,000 years

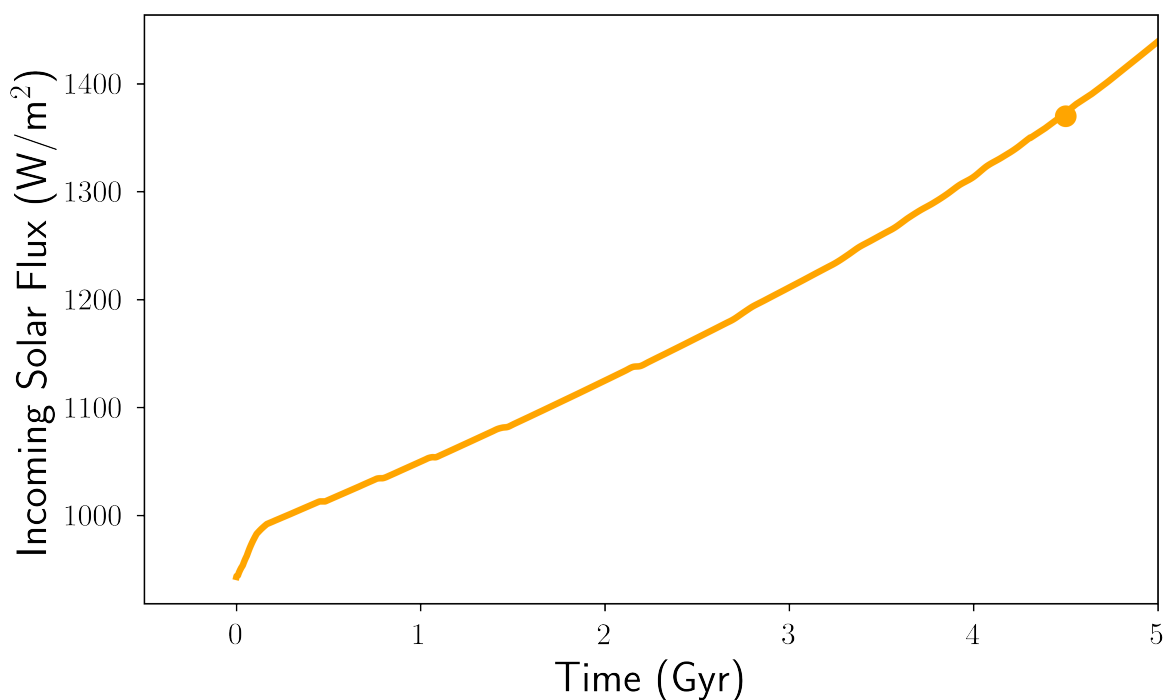


Figure 2.3: Incoming solar flux evolution with time, where the dot at 4.5 Gyr represents the present-day value.

after model initialization. As the planet warms, weathering is enhanced and the carbon sinks begin to overwhelm the sources, resulting in a gradual decrease in $p\text{CO}_2$. Initial oscillations in CO_2 (<5 Myr) occur due to a disequilibrium between CO_2 partitioning and the rapidly cooling mantle and surface temperatures. The final state of the model is ultimately insensitive to these initial model conditions. Furthermore, the transition between magma ocean solidification and the onset of plate tectonics is poorly understood, so modeling this transition falls outside the scope of this work.

The total albedo includes contributions from the ground and clouds, and is therefore higher. Clouds are tuned to match the PIE temperature. From tuning the constant cloud opacity pressure, we find a total albedo of 0.198 corresponding to a cloud fraction of 8%. We note that this is likely an unrealistically small cloud fraction for the PIE, most likely due to our simplified cloud implementation (see Section 4.6).

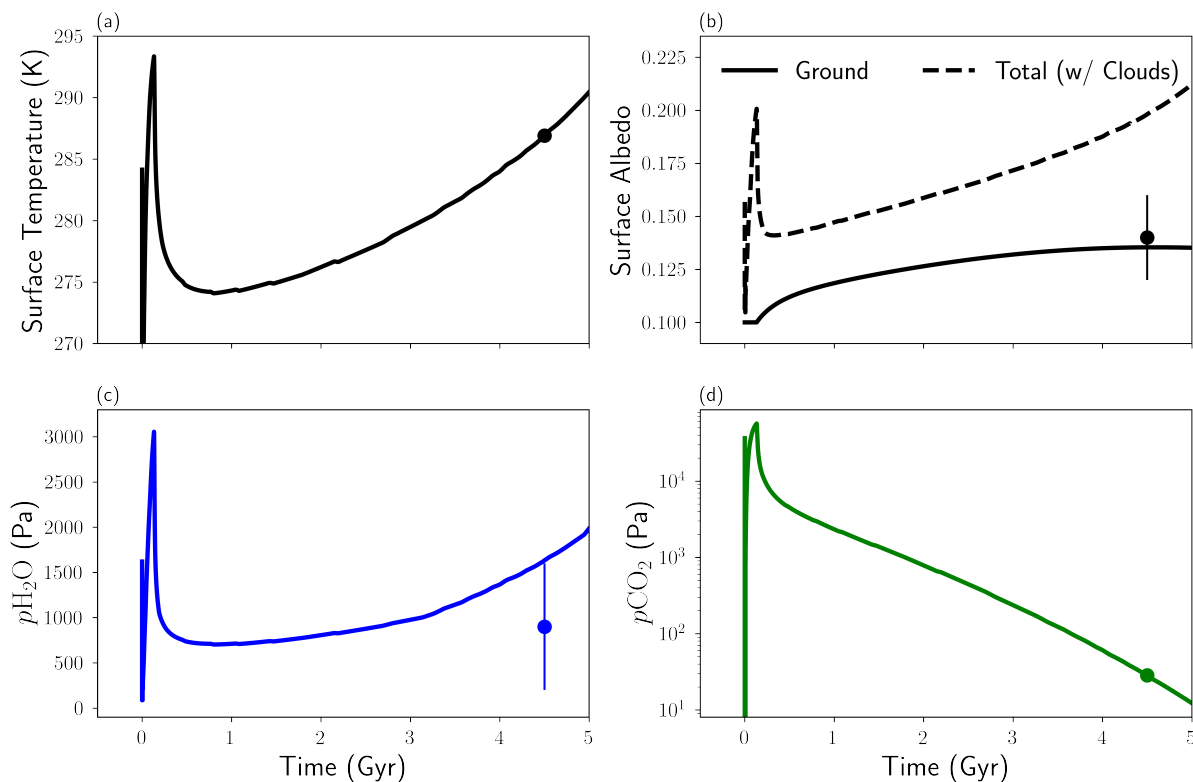


Figure 2.4: Some bulk properties of the abiotic Earth as a function of time. (a) Global average surface temperature warms as the star gradually brightens. (b) The ground albedo increases as dark ocean water subducts into the mantle and comparatively brighter land becomes exposed. Total albedo also increases as the warming planet’s atmosphere can hold more water vapor, and more clouds form. (c) As the planet warms, the atmosphere can hold more water and the H_2O partial pressure increases. (d) Finally, atmospheric CO_2 spikes early on when the planet is cold, and gradually decreases as weathering becomes more efficient on the warming planet. On all plots, dots represent measurements of PIE (4.5 Gyr) properties with uncertainties.

Figure 2.5 shows the evolution of the surface and mantle water budgets (a), the aqueous CO_2 concentration (b), ocean pH (c), and the surface ocean’s total dissolved inorganic carbon (DIC) concentration (d). We initialize the surface with 2 TOs, resulting in 1 TO in the mantle after 4.5 Gyr. Using Henry’s law, we calculate an aqueous CO_2 concentration of $10.8 \mu\text{mol/kg}$ by 4.5 Gyr (b). Assuming a constant carbonate concentration based on the abiotic calcite budget, we obtain a pH of 8.28 after 4.5 Gyr (c). Finally, the total dissolved inorganic carbon (DIC) at the ocean’s surface is $1919 \mu\text{mol/kg}$ after 4.5 Gyr (d). Ocean

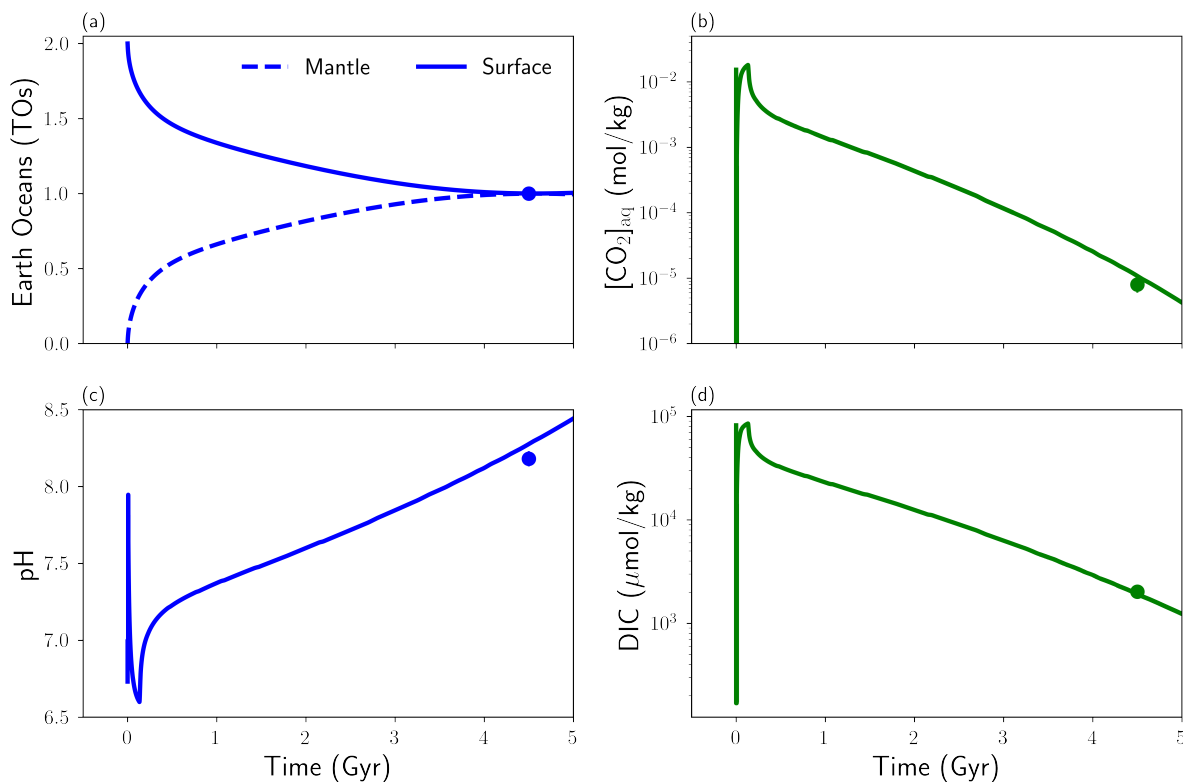


Figure 2.5: Properties of the ocean as a function of time. (a) After 4.5 Gyr, 1 TO has entered the mantle, ultimately leaving 1 TO on the surface. (b) The concentration of CO_2 in the oceans tracks closely with the atmospheric CO_2 inventory, following Henry’s law. (c) Prior to equilibration, the pH initially spikes because of the initial concentration of carbonate and bicarbonate in the calcite-saturated oceans is also very high. Following equilibration, the pH gradually becomes more alkaline as the abundance of CO_2 in the atmosphere and ocean gradually decreases. (d) The total dissolved inorganic carbon at the ocean’s surface, which includes aqueous CO_2 , carbonate, and bicarbonate, tracks closely with the atmospheric CO_2 budget. On all plots, dots represent measurements of PIE (4.5 Gyr) properties with uncertainties.

pH exhibits strong temperature-dependent behavior, while the DIC evolution tracks closely with the atmospheric CO_2 evolution. When the atmospheric CO_2 abundance is initially high (Figure 2.4, d), the aqueous CO_2 concentration is also high (Figure 2.4, b). Following model equilibration, as atmospheric CO_2 abundances sharply increase prior to 1 Gyr, the aqueous CO_2 concentration rises to 2×10^{-2} mol/kg, and the ocean pH is ~ 6.6 , or $30\times$ more acidic than present-day Earth oceans. As the atmospheric CO_2 content gradually wanes to PIE values, the aqueous CO_2 concentration and DIC follow suit, resulting in increasingly alkaline

oceans.

Figure 2.6 shows the evolution of the mantle. In Figure 2.6(a) we show the evolution of the average temperature of the mantle (T_M), the evolving upper mantle temperature T_{UM} , and the evolving lower mantle temperature T_{LM} . In panel (b) of the same figure, we show the evolution of heat flow in the upper mantle (Q_{UM}), and the radiogenic heat flow in the mantle ($Q_{rad,M}$). Panel (c) shows the changing boundary layer depths of the upper (δ_{UM}) and lower mantle (δ_{LM}). Note that the thickness of the assumed BMO always remains smaller than the thickness of the lower boundary layer ($< 60\text{--}100$ km), preventing the radiogenically enriched BMO from mixing into the convecting mantle. Panel (d) shows the viscosity of the upper (ν_{UM}) and lower mantle (ν_{LM}) with time. Finally, panel (e) shows the upper mantle melt fraction, and panel (f) shows the melt mass flux over geological time. By 4.5 Gyr, we confirm that T_{UM} , Q_{UM} , ν_{UM} , ν_{LM} , F_{melt} , and mantle melt mass flux all match measurements of Earth properties within measurement uncertainties.

In Figure 2.7, we show the evolution of key core properties over time. Panel (a) shows the evolution of the core-mantle boundary (T_{CMB}) and core (T_C) temperatures, while Panel (b) shows the evolution of the heat flow at the core-mantle boundary and the heat flow in the core due to radiogenics ($Q_{rad,C}$). Panel (c) shows the radius of the inner core with time, and panel (d) shows the core buoyancy flux. Finally, panel (e) shows the magnetic moment in Earth units, and panel (f) shows the magnetopause radius in Earth units.

For all properties shown in this figure, we note a discontinuity that occurs ~ 4 Gyr. This feature represents inner core nucleation in the model, when light elements begin to be rejected from the solidifying inner core (Driscoll & Bercovici, 2014). The thermal buoyancy flux of the core gradually decreases as the core cools (b). Once the inner core begins to solidify light elements are injected into the base of the liquid outer core, generating compositional buoyancy and boosting the total core buoyancy flux (b). This buoyancy injection into the outer core boosts the magnetic moment and magnetopause radius (c and d). All four core related properties in Figure 2.7 match the final expected values within measurement

uncertainties.

Table 2.3: Abiotic Earth model values compared to pre-industrial Earth measurements. Model values are reported at 4.5 Gyr.

Properties	Measured Value	Model Value
Atmospheric CO ₂ ($p\text{CO}_2$)	28.4 ± 0.4 Pa	28.2 Pa
Atmospheric H ₂ O ($p\text{H}_2\text{O}$)	900 ± 700 Pa	1632 Pa
Surface temperature (T_{surf})	286.9 ± 0.1 K	286.9 K
Ground albedo (A_g)	0.14	0.14
Ocean mass (w_{ocean})	1.4×10^{21} kg	1.4×10^{21} kg
Oceanic (CO ₂) _{aq}	8 ± 2 $\mu\text{mol/kg}$	10.8 $\mu\text{mol/kg}$
Total Dissolved Inorganic Carbon (DIC)	2000 ± 200 $\mu\text{mol/kg}$	1919 $\mu\text{mol/kg}$
pH	8.18 ± 0.05	8.28
Upper mantle temperature (T_{UM})	1587^{+164}_{-34} K	1562 K
Core-mantle boundary temperature (T_{CMB})	4000 ± 200 K	3889 K
Upper mantle heat flow (Q_{UM})	38 ± 3 TW	39 TW
Core-mantle boundary heat flow (Q_{CMB})	11 ± 6 TW	16 TW
Upper mantle viscosity (ν_{UM})	$10^{16.68 \pm 1}$ m ² s ⁻¹	$10^{17.46}$ m ² s ⁻¹
Lower mantle viscosity (ν_{LM})	$10^{17.28 \pm 1}$ m ² s ⁻¹	$10^{17.73}$ m ² s ⁻¹
Upper mantle melt fraction (F_{UM})	$11.5 \pm 3.5\%$	8.5%
Melt flux	$(1.3 \pm 0.8) \times 10^6$ kg/s	2.0×10^6 kg/s
Inner core radius (R_{IC})	1221 ± 1 km	1221 km
Magnetic moment	1.00 ± 0.05 E. Unit	1.01 E. Unit
Magnetopause radius	1.00 ± 0.02 E. Unit	1.00 E. Unit

Note. All measured values and uncertainties are detailed in Section 1, and summarized in Table 2.1.

Table 2.3 summarizes our final model values for 19 pre-industrial characteristics. The two parameters that fall slightly outside measurement uncertainty are oceanic (CO₂)_{aq} and ocean pH. Our modeled (CO₂)_{aq} of 10.8 $\mu\text{mol/kg}$ slightly exceeds the measured range of 8 ± 2 $\mu\text{mol/kg}$, while our modeled ocean pH of 8.28 is marginally more alkaline than the measured value of 8.18 ± 0.05 (an offset of 0.10 pH units, or roughly twice the reported measurement uncertainty). Assuming the measurement and uncertainties associated with these two parameters represent normal underlying distributions, our model results are no

more than $2\text{-}\sigma$ off from the mean. We interpret this as sufficient agreement within the estimated measurement uncertainties. Nonetheless, since ocean pH is determined by the carbonate equilibrium system, these modest deviations from expectation for both ocean pH and $(\text{CO}_2)_{\text{aq}}$ can be attributed to our simplifying assumptions regarding calcite concentration. We discuss the implications of this assumption in Section 4.3.

Finally, as a proof of concept, in Figure 2.8, we compare reflected light spectra of a realistic cloudy Earth and a snapshot of the abiotic Earth after 4.5 Gyr (a), as well as their corresponding abundance profiles (b). Volume-mixing ratio profiles are shown in log-log space as a function of decreasing pressure in bars on the y-axis. While both atmospheres include CO_2 , H_2O , and N_2 , the abiotic atmosphere lacks O_2 (and therefore O_3) and CH_4 . The lack of O_2 and O_3 contributes to the most significant differences in absorption between the two spectra. For example, in the near-UV $\sim 0.25\text{ }\mu\text{m}$, the true Earth spectrum shows absorption due to the O_3 Hartley band. In the visible, the Earth spectrum has absorption features at $\sim 0.6\text{ }\mu\text{m}$ due to the O_3 Chappuis band, and a prominent narrow-band feature at $\sim 0.76\text{ }\mu\text{m}$ due to the O_2 A-band. In the near-infrared, the Earth spectrum shows absorption at $\sim 1.27\text{ }\mu\text{m}$ due to the weak O_2 feature. Finally, the true Earth spectrum has $\sim 2\text{ ppm}$ of CH_4 , which produces an absorption feature at $\sim 1.69\text{ }\mu\text{m}$ in the near-infrared. This low abundance results in a weak absorption feature that ultimately represents only a minor difference between the true and abiotic Earth spectra.

Figure 2.8b demonstrates some minor differences between the N_2 , CO_2 , and H_2O profiles. While the ICRCCM atmosphere is 79% N_2 , the abiotic atmosphere has 1 bar to compensate for the lack of bulk O_2 . Furthermore, the ICRCCM atmosphere was benchmarked to more recent terrestrial atmospheric abundances. As a result, the ICRCCM atmosphere has a higher CO_2 abundance of 330 ppm versus our 280 ppm. Lastly, though both water vapor profiles have similar near-surface abundances, our tropopause appears at a slightly lower pressure (and higher altitude) than the ICRCCM profile. This difference occurs for two reasons. First, the ICRCCM profile represents an average mid-latitude climate during the

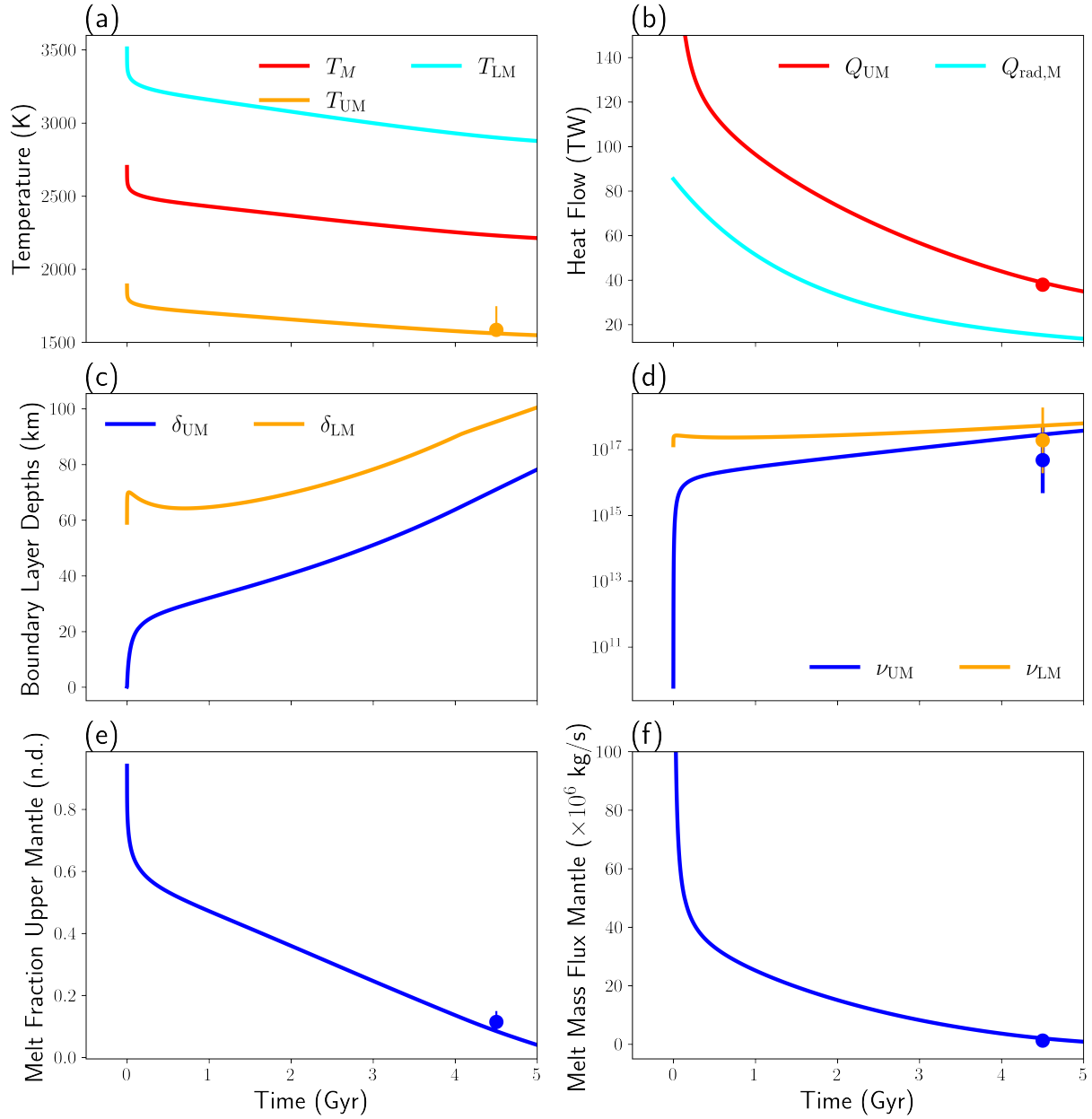


Figure 2.6: Key properties of the Earth’s mantle through time. (a) The mantle and core gradually cool over time. (b) As the core and mantle cool, the heat flow across the upper mantle and the core-mantle boundary and the heat flow due to radiogenics in the core and mantle decrease. (c) Both the upper and lower mantle boundaries grow deeper as the core-mantle system cools and the mantle solidifies downwards. (d) The lower mantle is initially less viscous than the upper mantle. As water subducts into the cooling mantle, both the upper and lower mantle viscosities gradually increase. (e) As the interior cools, the upper mantle melt fraction rapidly decreases. (f) Similarly, the melt flux from the mantle also decreases as the mantle cools. On all plots where applicable, dots represent measurements of PIE (4.5 Gyr) properties with uncertainties.

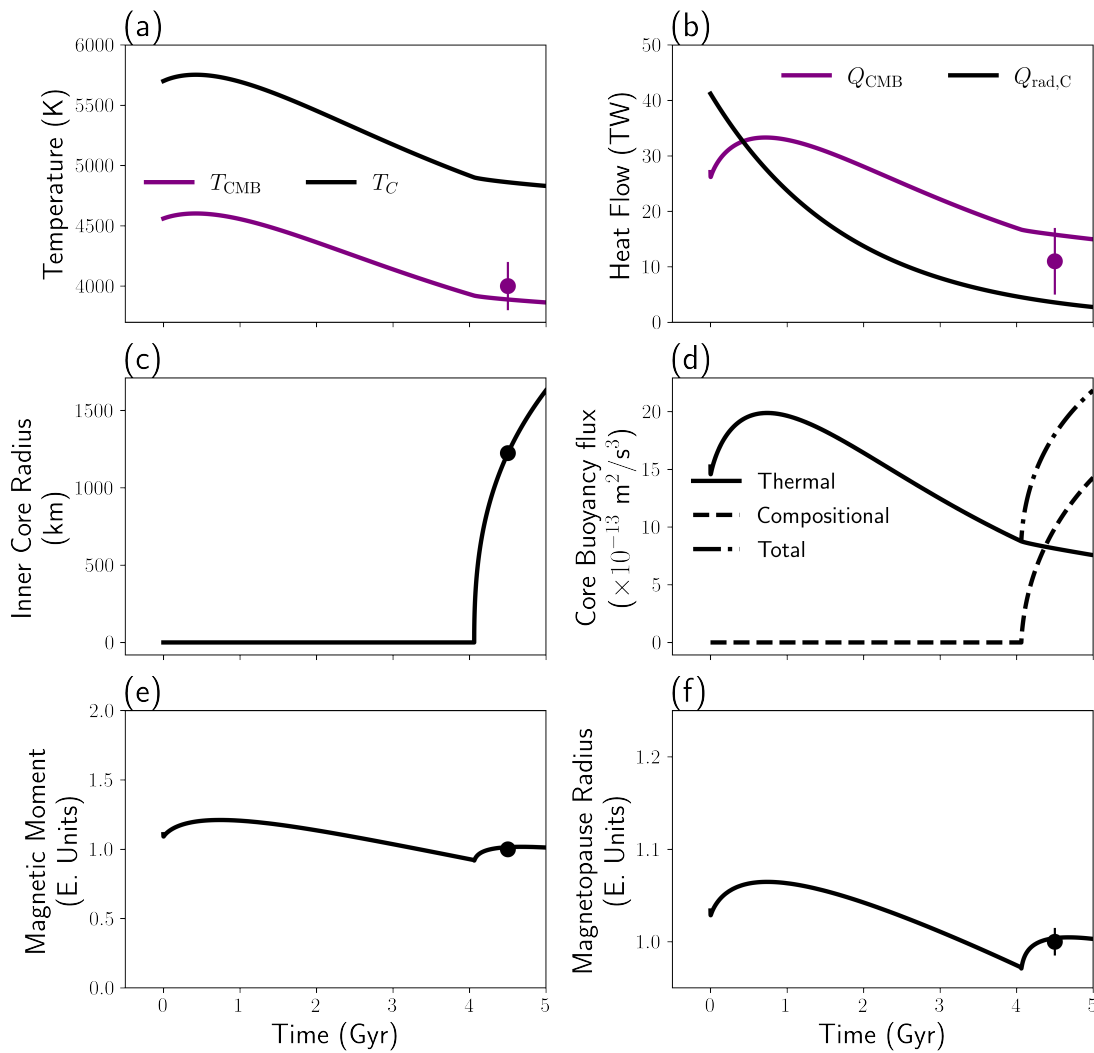


Figure 2.7: Key properties of the Earth’s core through time. (a and b) The temperature and heat flows of the core and core-mantle boundary secularly decrease with time. (c) Once the interior has sufficiently cooled, the inner core rapidly solidifies at ~ 4 Gyr. (d) The thermal buoyancy flux of the core gradually decreases as the core cools. Once the inner core solidifies and exsolves light elements, the compositional changes inject energy into the outer core and boost the overall core buoyancy flux. (e and f) The late energy increase in the outer core due to inner core solidification causes coinciding discontinuities in the magnetic moment and the magnetopause radius. On all plots where applicable, dots represent measurements of PIE (4.5 Gyr) properties with uncertainties.

summer months whereas our profile represents a globally averaged profile. Second, `clima` estimates the location of the tropopause by calculating the atmospheric skin temperature rather than explicitly calculating the radiative-convective boundary. Nonetheless, these dif-

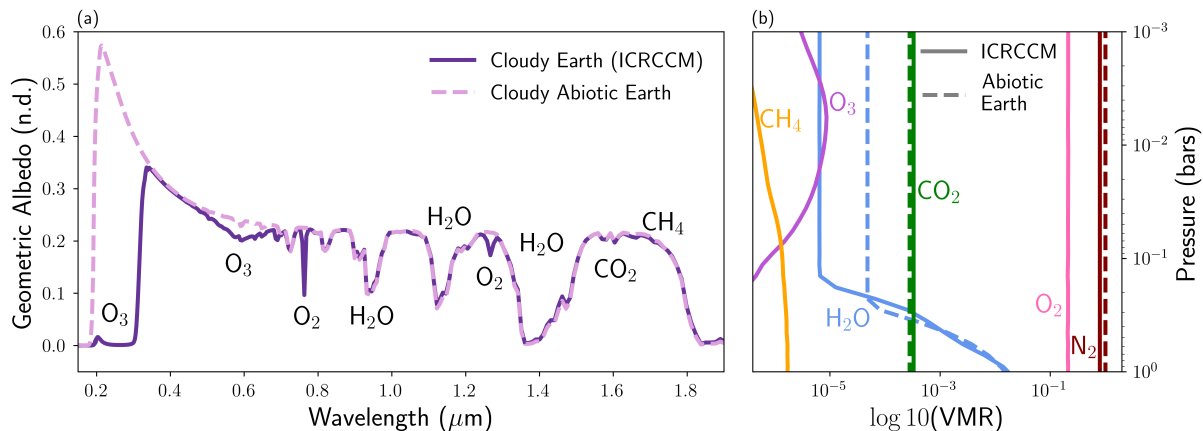


Figure 2.8: (a) A comparison of reflected light spectra for a realistic, cloudy Earth and the abiotic atmospheric state after 4.5 Gyr. The abiotic atmosphere lacks O_2 (and therefore O_3) and CH_4 , contributing to differences in absorption. (b) A comparison of the input molecular abundance profiles used to generate each reflected light spectrum. We address minor differences between the CO_2 , H_2O , and N_2 profiles in the accompanying text. Though the true profiles extend to a top-of-atmosphere pressure of 1×10^{-5} bars, we only show the surface to the upper stratosphere to better emphasize the differences in the water vapor profiles given the logarithmic scale.

ferences in the atmospheric profiles do not contribute to major differences in the resulting spectra – the weak CO_2 bands are insensitive to small differences in abundance, and most spectral water vapor absorption occurs below the tropopause.

4 Discussion

We have presented a coupled core-mantle-crust-climate evolution model that reproduces 19 properties of the pre-industrial Earth largely within measurement uncertainties, with two ocean chemistry parameters showing modest offsets attributable to our simplified treatment of calcite saturation. We have also shown that our model can be used to generate synthetic HWO-like reflected light spectra. Our findings suggest that the abiotic Earth could evolve to a similar state (in terms of global geophysics and climate) to present-day Earth without life, but without abundant atmospheric oxygen. In this section, we discuss the implications of key assumptions in our model, such as neglecting CH_4 and abiotic calcite saturation in the

oceans. Finally, we discuss our findings in the context of the Gaia hypothesis and forecast future research directions.

4.1 Spectroscopic Predictions for HWO Observations

The atmospheric states produced by our coupled evolution model can generate realistic reflected light spectra suitable for comparison with future HWO observations (see Figure 2.8). The most significant spectral differences arise from the absence of biologically-mediated gases: the modern Earth spectrum exhibits strong absorption features from O_2 , O_3 , and CH_4 , none of which appear in our abiotic atmosphere.

This comparison validates that our model produces realistic atmospheric states that can be propagated through radiative transfer codes to generate synthetic HWO observations. The abiotic spectrum represents the expected appearance of a habitable but lifeless Earth-like planet, establishing a null hypothesis against which to test biosignature detections. While both spectra show prominent H_2O and CO_2 features, only the biological Earth exhibits the O_2 (or O_3)/ CH_4 disequilibrium signature that would constitute compelling evidence for modern Earth-like life (Meadows et al., 2018b). Future parameter space explorations varying planetary properties and host star types will reveal the diversity of reflected light spectra representing abiotic habitable worlds, establishing the range of null hypothesis scenarios that HWO must discriminate from life.

4.2 CH_4 as a Greenhouse Gas

Our model establishes that the carbonate-silicate cycle can operate effectively to maintain habitable surface temperatures over geological timescales using only CO_2 and H_2O as greenhouse agents, without requiring additional radiative forcing from trace gases like CH_4 . This finding has important implications for understanding both Earth’s climate history and the potential habitability of rocky exoplanets.

Abundant atmospheric CH_4 due to microbial life on the early Earth has been proposed

as a solution to the faint young Sun paradox (Pavlov et al., 2000; Kharecha et al., 2005). Pavlov et al. (2000) showed that atmospheric CH_4 mixing ratios of 100–1000 ppmv, combined with elevated CO_2 levels, could maintain above-freezing temperatures despite 20–30% lower solar luminosity. While biological fluxes of CH_4 may have warmed Earth’s climate in certain epochs – with some models suggesting temperature increases of 5–10°C from elevated methane during the Archean (Roberson et al., 2011) – our results demonstrate that such biological enhancement is not necessary for maintaining habitability. An abiotic Earth may remain habitable through the CO_2 - H_2O greenhouse alone, provided the carbonate-silicate cycle operates.

Future work should nevertheless explore the sensitivity of climate evolution to low levels of abiotic CH_4 . Even modest trace gas abundances could influence atmospheric photochemistry and oxidation state. The atmospheric lifetime of methane depends strongly on photochemically generated hydroxyl radicals (OH^-) produced through water vapor photolysis (Pavlov et al., 2001). In an anoxic atmosphere without oxygenic photosynthesis, extended methane lifetimes could potentially allow greater accumulation than our simplified analysis suggests given sufficiently available OH^- . Additionally, planets orbiting M- and K-dwarf stars may experience different photochemical regimes due to their host stars’ UV spectra, potentially affecting methane lifetimes and the viability of methane as a warming agent on abiotic worlds (Segura et al., 2005; Arney et al., 2016). Incorporating volcanic CH_4 outgassing coupled to mantle redox state evolution would establish a more complete abiotic baseline for interpreting future observations of exoplanets orbiting different stars.

4.3 Abiotic Marine Calcite Budget

Tuning the calcite saturation and concentration to achieve the observed pre-industrial ocean DIC and approximate pH represents an oversimplification of the terrestrial carbon cycle, as demonstrated by the larger model disagreement for ocean pH and aqueous CO_2 concentration. We tune the calcite saturation and concentration to match the observed total DIC, but

this choice then constrains the carbonate ion concentration via the calcite saturation equation (Equation [2.24]), which in turn determines the partitioning among carbonate species through equilibrium constants (Equations [2.25], [2.26], and [2.27]), and ultimately fixes the pH. This cascade of constraints means that solely tuning calcite concentration to match total DIC does not guarantee agreement with other observables that depend on the same carbonate chemistry, including pH and $(\text{CO}_2)_{\text{aq}}$.

The larger pH offset highlights a fundamental limitation of our simplified ocean chemistry model: we assume calcite saturation and do not explicitly track total alkalinity as an independent variable, such as in [Krissansen-Totton et al. \(2021\)](#). In a more complete carbonate system, total alkalinity and DIC jointly determine pH and carbonate speciation. By fixing calcite concentration to match DIC, we implicitly constrain alkalinity through the carbonate equilibrium, removing a degree of freedom that would otherwise allow independent adjustment of pH.

The primary distinction between our abiotic model and a biologically mediated carbon cycle lies in how carbon is packaged for gravitational transport from the surface to the seafloor. The biological pump creates heavier, rapidly-sinking aggregates (marine snow, fecal pellets, shells) from organic and inorganic carbon ([Passow & Carlson, 2012](#)), allowing more carbon to reach the deep ocean before re-mineralization. Life effectively enhances gravitational export of carbon and influences the depth at which re-mineralization occurs. Organisms migrating vertically through the water column also transport metabolized carbon ([Le Moigne, 2019](#)), but gravitational sinking dominates in modern oceans ([Boyd et al., 2019](#); [Nowicki et al., 2022](#); [Siegel et al., 2023](#)). In an abiotic ocean, gravitational settling of inorganic precipitates would still occur, but the size distribution, sinking velocities, and spatial patterns of carbonate formation would likely differ from the biologically-mediated system.

Furthermore, it is unclear whether abiotic calcite is less thermodynamically stable than its biogenic counterpart. [Zhuang et al. \(2018\)](#) suggest that abiotic calcite is less stable

than biogenic calcite due to lower crystallinity and activation energy, but [Stalport et al. \(2005\)](#) found that biologically precipitated calcite thermally degrades at temperatures approximately 40°C cooler than abiotic calcite, suggesting greater thermal stability of abiotic crystals at elevated temperatures. Pessimistically, these differing material properties combined with the orders-of-magnitude slower precipitation kinetics ([Carpenter & Lohmann, 1992](#)) imply that achieving high calcite concentrations in an abiotic ocean may require either longer equilibration times or higher degrees of supersaturation than in living systems.

The extent to which an uninhabited planet could accumulate and maintain the calcite concentrations necessary to operate an effective carbonate-silicate thermostat remains an open question with implications for the long-term habitability of abiotic, Earth-like exoplanets. The calcite concentration directly impacts carbonate chemistry and thus the efficiency of oceanic carbon sequestration and weathering feedbacks. Future work should incorporate a more complete carbonate system with total alkalinity as a prognostic variable, following approaches like [Krissansen-Totton et al. \(2021\)](#). Additionally, systematic parameter space explorations should assess the sensitivity of ocean chemistry, atmospheric CO₂ regulation, and surface temperature to variations in calcite concentration and alkalinity. Such studies would help establish the range of abiotic carbonate cycling rates compatible with maintaining temperate surface conditions on terrestrial planets over Gyr timescales.

4.4 Implications for the Gaia Hypothesis

Since its introduction by [Lovelock & Margulis \(1974\)](#), the Gaia hypothesis has spurred long-running debates over life’s role in long-term climate stability (i.e., habitability). In evaluating Lovelock’s original theory, [Kirchner \(1989\)](#) argues for re-categorizing Gaia into “weak” and “strong” variants – weak Gaia argues that interactions between living and abiotic processes have some stabilizing effect on the Earth; strong Gaia asserts that such interactions are the *most significant* process stabilizing Earth’s climate. Strong Gaia proponents thus argue that life stabilizes the planetary thermostat by regulating temperature, atmospheric compo-

sition, and ocean chemistry to maintain conditions favorable to its survival and propagation (Lovelock & Margulis, 1974; Lovelock, 1983, 2000, 2016).

Generally, criticisms of the Gaia hypothesis acknowledge that life alters its environment, but point out that, despite some compelling examples, life’s activities are not always favorable to its long-term survival. For instance, on the modern Earth, marine plankton “plunder” surface ocean nutrients, creating widespread biological deserts (Volk, 2002; Kirchner, 2003). Furthermore, anthropogenic climate change is a salient contemporary example of how biological activity can *disrupt* rather than stabilize the climate system (Kirchner, 2002, 2003). These examples demonstrate that biological feedbacks are not intrinsically homeostatic.

More recently, modeling studies have investigated whether life can actively contribute to maintaining long-term planetary habitability. Nicholson et al. (2018) developed the ExoGaia model, which simulates how abiotic geochemistry and microbial metabolisms interact to influence surface temperature. They found that under stable conditions, microbial metabolisms can prevent planets from reaching inhospitable temperatures that would otherwise occur on their lifeless counterparts. Alcabes et al. (2020) extended this work by subjecting planets to three types of global climate perturbations, finding that planets with Gaian feedbacks are most resilient to global climate change, especially when changes are gradual. Together, these ExoGaia studies suggest that if life does emerge on a planet, it may under certain conditions evolve stabilizing feedbacks that enhance the long-term habitability of the host planet.

However, these models investigate whether life *can* stabilize planetary climate under a weak Gaia hypothesis, not whether life is *necessary* for climate stability as in the strong Gaia hypothesis. As further evidence against a strong Gaia mechanism, Kirchner (1989) identifies purely geophysical models that produce long-term climate stability without invoking life at all (Berner et al., 1983; Walker et al., 1981). Our results are in agreement with these models. We demonstrate that an Earth-like planet can maintain habitable surface temperatures over 4.5 Gyr of evolution (and longer) through abiotic processes alone, without requiring biological feedbacks for climate regulation.

The carbonate-silicate cycle, driven by the interplay between silicate weathering, volcanic outgassing, and ocean chemistry, provides a robust negative feedback that keeps atmospheric CO_2 and surface temperature within habitable bounds despite the Sun’s luminosity increasing by approximately 30% over 4.5 Gyr. Critically, our model demonstrates this climate stability while self-consistently evolving all major planetary subsystems: the interior cools and outgassing rates decline; the magnetic dynamo weakens and the magnetopause contracts; ocean chemistry adjusts to changing atmospheric CO_2 ; and the incoming stellar flux steadily increases. Despite these dramatic changes, the coupled system maintains surface temperatures between 272–290 K throughout the evolution. This represents a more complete test of climate stability than previous models that considered the carbon cycle in isolation, as it demonstrates that the carbonate-silicate thermostat functions robustly even when embedded within a fully evolving planetary system with realistic feedbacks between interior, surface, atmosphere, and external forcing from the brightening Sun.

As discussed above, our model does simplify ocean chemistry, which biology enhances on the modern Earth (Section 4.3). However, the fundamental mechanism of climate stabilization – the temperature and CO_2 dependence of silicate weathering coupled to carbonate precipitation and subduction – operates independently of life. While life may affect the *spatial distribution* and *kinetics* of carbonate precipitation, these biological enhancements may not be *necessary* for long-term climate stability.

Our finding that life may not be a requirement of a habitable planet has important implications for understanding Earth’s habitability and for predicting conditions on rocky exoplanets. The persistence of habitable conditions over geological time may not require life to actively regulate the climate, rather potentially emerging from volatile cycling between planetary reservoirs. Habitable conditions may have been created on the early Earth by the abiotic carbonate-silicate cycle. Once life emerged, it may have developed stabilizing feedbacks as suggested by ExoGaia-type models (Nicholson et al., 2018; Alcabes et al., 2020), potentially altering weathering rates, ocean chemistry, and atmospheric composition. How-

ever, our results suggest that the underlying abiotic stabilization mechanism provided by the carbonate-silicate cycle can maintain habitability independently. While we cannot rule out that life enhances climate stability (weak Gaia), our results suggest that life may not be a requirement of a habitable planet.

The implication that life may not be a requirement of a habitable planet raises a complementary question: given a habitable planet, how long does it take for life to emerge? On Earth, mineralogical evidence confirms that liquid water oceans existed as early as 4.4 Ga (Cameron et al., 2024), while the oldest widely accepted evidence for life – stromatolites from the Pilbara Craton – dates to 3.5 Ga (Walter et al., 1980). Though more ancient biogenic signatures have been proposed (and continue to be highly disputed), the stromatolite evidence provides a conservative estimate that abiogenesis occurred within approximately 900 Myr of ocean formation. It has been suggested that this relatively rapid emergence of life on Earth implies that abiogenesis may also be rapid on habitable exoplanets (Kipping, 2025). Our results support this conjecture by showing how Earth could have been habitable but abiotic for billions of years, yet life still appeared relatively shortly after the appearance of habitable conditions. On the other hand, our model suggests that habitable surface conditions can persist for billions of years through purely abiotic mechanisms, potentially providing a prolonged “prebiotic window” during which the chemical building blocks of life could accumulate and undergo the reactions necessary for creating life.

Potentially habitable exoplanets observed by future missions may be captured at various stages of their prebiotic evolution: some may harbor early life similar to Earth at 3.5 Ga, while others could still be accumulating prebiotic chemistry despite having maintained habitable conditions for comparable or even longer durations. Indeed, our results suggest that rocky planets with plate tectonics, liquid water oceans, and silicate weathering can maintain habitable surface conditions over billion-year timescales without requiring biology. Thus, the mere presence of habitable conditions is not sufficient evidence that a planet is inhabited. Some worlds may remain prebiotic for prolonged periods, or never experience abiogenesis at

all. Distinguishing inhabited from uninhabited-but-habitable worlds will therefore require (1) detecting disequilibrium biosignature gases (Lovelock & Margulis, 1974; Meadows et al., 2018b) or surface features directly attributable to biological activity (Schwieterman et al., 2018b), but also (2) a well-quantified abiotic baseline for the potential mimics of biosignature gases and surface biosignatures (e.g., Wogan et al., 2020; Krissansen-Totton et al., 2021).

4.5 Comparison to Earth Evolution

While we did not aim to validate our results on the early Earth we can broadly compare our model predictions to measurements of global average paleo-climate conditions in the literature. Following Krissansen-Totton et al. (2018a), we compare the evolution of surface temperature, atmospheric $p\text{CO}_2$, and ocean pH against paleo-climate proxy data in Figure 2.9a, b, and c, respectively. We also qualitatively compare our interior thermal evolution to previous work. Overall, we find that our abiotic model largely reproduces paleo $p\text{CO}_2$ and ocean pH estimates across much of Earth’s history and yields a mantle thermal history that is consistent with most previous models, while predicting systematically cooler surface temperatures than biotic proxy reconstructions suggest.

Across the Archean (~ 0.5 – 2.0 Gyr) and Proterozoic (~ 2.0 – 4.0 Gyr) eons, we find that our predicted $p\text{CO}_2$ remains broadly consistent with measurements established via mass balance models of weathering (Driese et al., 2011; Sheldon, 2006) and a Precambrian upper bound estimated from calcification of cyanobacterial sheaths (Kah & Riding, 2007). Near the neo-Archean/paleo-Proterozoic boundary (~ 2 Gyr), our $p\text{CO}_2$ is nearly an order of magnitude smaller than the abundances predicted by Kanzaki & Murakami (2015) based on modeling of aqueous soil chemistry. By contrast, our model predicts global surface temperatures that are consistently several K cooler than measurements of glacial deposits (Geboy et al., 2013; Kuipers et al., 2013; Williams, 2005; Ojakangas et al., 2014; Young et al., 1998; De Wit & Furnes, 2016), $\delta^{18}\text{O}$ and deuterium isotopes in 3.42 Ga cherts (Hren et al., 2009), and $\delta^{18}\text{O}$ and organic phosphate (Blake et al., 2010), but consistent with lower-end estimates derived

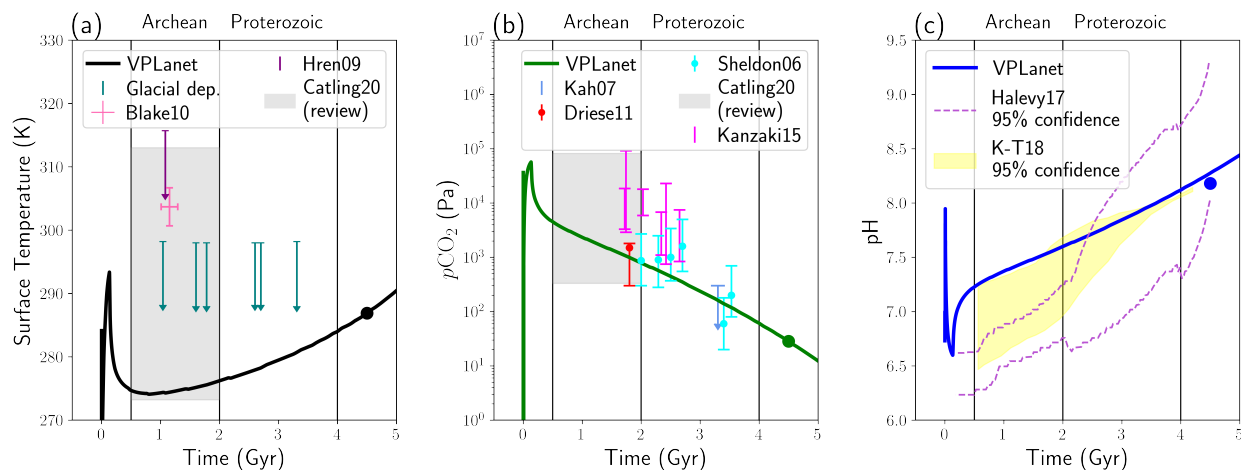


Figure 2.9: Comparison of the `vPLANet` abiotic Earth model evolution (solid curves) against paleo-climate proxy reconstructions from the literature, following the framework of Krissansen-Totton et al. (2018a). Vertical black lines demarcate the Archean (~0.5–2.0 Gyr) and Proterozoic (~2.0–4.0 Gyr) eons. Filled black, green, and blue circles mark pre-industrial Earth values. **(a)** Surface temperature evolution compared with glacial deposit occurrences (teal arrows; Geboy et al., 2013; Kuipers et al., 2013; Williams, 2005; Ojakangas et al., 2014; Young et al., 1998; De Wit & Furnes, 2016), $\delta^{18}\text{O}$ and organic phosphate constraints (Blake10; Blake et al., 2010), $\delta^{18}\text{O}$ and deuterium isotopes in 3.42 Ga cherts (Hren09; Hren et al., 2009), and the habitable temperature range reviewed in Catling & Zahnle (Catling20; 2020). Downward arrows denote upper limits. **(b)** Atmospheric $p\text{CO}_2$ evolution compared with cyanobacterial sheath calcification upper bounds (Kah07; Kah & Riding, 2007), paleosol-based estimates from Driese et al. (Driese11; 2011) and Sheldon (Sheldon06; 2006), aqueous soil chemistry models (Kanzaki15; Kanzaki & Murakami, 2015), and the habitable range reviewed in Catling20. **(c)** Ocean pH evolution compared with 95% confidence intervals from the carbonate-alkalinity model of Halevy & Bachan (Halevy17; 2017) and the coupled carbon-cycle model of Krissansen-Totton et al. (K-T18; 2018a).

from models fitted to these constraints (Krissansen-Totton et al., 2018a). Similarly, our ocean pH throughout the evolution is consistent with the higher end of estimates from Krissansen-Totton et al. (2018a), but is consistently more alkaline than estimates from Halevy & Bachan (2017). The offset in pH likely reflects our simplified treatment of ocean alkalinity (see Section 4.6), while our consistently cooler surface temperatures are in line with prior work identifying biogenic CH_4 from methanogens as an important greenhouse contributor on the early Earth (Pavlov et al., 2000; Kharecha et al., 2005). An abiotic world would lack such elevated CH_4 abundances and, under a faint young Sun, may evolve along a colder trajectory.

While proxy evidence gives a validation baseline for paleo-Earth climate, Earth’s early interior evolution is harder to constrain. Herzberg et al. (2010) compiled petrologic estimates of mantle temperature over Earth history, which reveal more or less monotonic mantle cooling since the Archean. However, there is significant uncertainty in mantle temperature and cooling rates in the Archean and Hadean, with models predicting both early heating periods (Korenaga, 2008a,b) or monotonic cooling (Driscoll & Bercovici, 2014).

A subtler tension in this comparison is worth acknowledging: the present-day Earth system parameters against which we tune and validate our model are themselves the product of ~ 4 Gyr of coupled biological and geological evolution. In calibrating an abiotic model to reproduce these endpoint values, we are implicitly asking whether an abiotic Earth *could* have arrived at a state quantitatively similar to the present-day. Our results indicate that such a solution is achievable in principle, but the broader question of what global properties depend on biotic versus abiotic processes requires future work.

Nonetheless, our model demonstrates that many of the present-day global geophysical conditions remain achievable along an abiotic pathway. We emphasize that this solution is not necessarily unique: other plausible evolutionary trajectories could also reproduce modern conditions, and our results likely represent one of a suite of Earth-like solutions. More broadly, our findings suggest that the presence or absence of life may meaningfully contribute to the thermal trajectory of an Earth-like planet — an abiotic Earth plausibly

evolves through a substantially cooler climate history than the biotic Earth inferred from proxy data, even while converging on the same present-day state. This work is a step toward systematically mapping the broader space of both biotic and abiotic evolutionary histories consistent with present-day Earth.

4.6 Future Work

Our core-mantle-crust-climate model reproduces the conditions of the pre-industrial Earth without life, and provides a foundation for future studies aiming to simulate the evolving habitability of rocky exoplanets as well as the observable signatures of these states. However, we found that our simple cloud parameterization is likely insufficient to model realistic radiative forcing from clouds. We tuned a wavelength-independent cloud opacity pressure to match pre-industrial Earth surface temperature, resulting in an unrealistically low cloud fraction of 8% compared to Earth’s observed 67% average cloud coverage (Wielicki et al., 1996). This discrepancy arises because we “paint” clouds onto the surface rather than self-consistently calculating cloud formation and radiative effects within the atmospheric column (Fauchez et al., 2018). Future work should reconcile this inconsistency.

Our treatment of atmospheric processes spans a spectrum of physical fidelity. For surface temperature, atmospheric water vapor partitioning, and radiative transfer, we employ first-principles physics: `clima` solves the radiative transfer equations with wavelength-dependent opacities, while water vapor is constrained by the Clausius-Clapeyron relation at each atmospheric level. By contrast, clouds are parameterized – the cloud opacity scale and reflectance parameters are calibrated to reproduce the PIE’s cloud radiative effect, and are not derived from first-principles cloud microphysics.

While more sophisticated 1-D cloud treatments exist (Ackerman & Marley, 2001; Zsom et al., 2012; Fauchez et al., 2018; Helling, 2019; Windsor et al., 2023), implementing them would largely trade our current free parameters for others (e.g., sedimentation efficiency in microphysical models), without fundamentally resolving the uncertainties inherent to repre-

senting 3-D cloud processes in 1-D globally-averaged models. Clouds remain the dominant source of uncertainty in 3-D general circulation models due to the complex interplay of microphysics, planet-wide dynamics, and radiative transfer (Mauritsen et al., 2012; Vial et al., 2013; Way et al., 2018). For the purposes of investigating long-term planetary evolution across diverse parameter spaces, a more pragmatic approach is to acknowledge this uncertainty explicitly and marginalize over plausible cloud parameter ranges when conducting broad explorations of Earth-like exoplanets. By varying our cloud opacity pressure and cloud albedo parameters across physically motivated ranges in future studies, we can assess how cloud radiative forcing uncertainties affect predicted habitability outcomes without requiring self-consistent calculations of cloud microphysics.

Further improvements to the model would ensure that it can capture additional feedback mechanisms of planets on diverse evolutionary paths. These improvements include incorporating additional volatile cycles and early Earth processes that affect the planet’s initial conditions. Below we discuss potential model improvements and describe applications to exoplanet studies, which may include coupling our model outputs to 3-D general circulation models (GCMs) to better capture the complexity of surface processes like clouds and ice coverage.

Our model currently only handles CO₂ and H₂O cycling, which is a subset of Earth’s geologically active volatiles. Following Krissansen-Totton et al. (2021), we assume that N₂ remains a bulk gas by setting it to a constant of 1 bar throughout the evolution. Future work should incorporate abiotic nitrogen cycling into our whole-planet model, as the abundance of atmospheric N₂ affects surface pressure and impacts the surface temperature of the planet via collision-induced absorption (Schwieterman et al., 2015) and pressure broadening of CO₂ and H₂O (Goldblatt et al., 2009). Furthermore, Laneuville et al. (2018) showed that, in an abiotic context, the abundance of atmospheric nitrogen is highly dependent on surface and interior processes, suggesting this form of volatile cycling is ripe for inclusion in a whole-planet modeling framework.

Future work should explore the sensitivity of long-term climate evolution to different assumptions about the efficiency of continental weathering on an abiotic planet without land vegetation. Previous habitability studies have suggested that land plants enhance continental weathering rates beyond purely abiotic levels, potentially driving abiotic planets to equilibrate at a higher $p\text{CO}_2$ (Haqq-Misra et al., 2016) and enhancing the likelihood of living planets experiencing global glaciation events (Kadoya & Tajika, 2019). However, the quantitative contribution of biology to silicate weathering rates remains poorly constrained. Field estimates of the biotic enhancement factor span nearly two orders of magnitude (e.g., Dessert et al., 2003; Cochran & Berner, 1996), and no true abiotic control site exists on Earth today (Wild et al., 2022). Marginalizing over a range of weathering assumptions would help quantify the effect of this uncertainty and bracket the range of plausible $p\text{CO}_2$ trajectories and surface temperatures for abiotic Earth-like planets, establishing more robust predictions for HWO observations.

While low abiotic CH_4 abundances likely do not directly contribute to greenhouse warming, methane nevertheless plays an important role in atmospheric chemistry and redox balance (Pavlov et al., 2001). The atmospheric abundance and lifetime of CH_4 depends on the availability of photochemically generated oxidants, which in turn depends on the planet's oxidation state, which is itself influenced by hydrogen escape to space (Catling et al., 2001; Zahnle et al., 2019). Atmospheric escape is therefore critical for determining both the feasibility of detecting abiotic methane and for interpreting CH_4 observations as potential biosignatures.

More broadly, atmospheric escape processes – particularly XUV-driven hydrodynamic and thermal escape – can substantially alter volatile inventories, atmospheric composition, and planetary oxidation state over Gyr timescales (Catling et al., 2001; Luger & Barnes, 2015; Gialluca et al., 2024). This possibility is especially important for planets orbiting lower mass stars like M- and K-dwarfs, which experience extended pre-main-sequence phases compared to G- and F-dwarfs, with high XUV fluxes that can drive significant atmospheric loss

(Chadney et al., 2015; Atri & Mogan, 2021). For example, on M-dwarf planets, preferential hydrogen escape oxidizes the atmosphere and potentially the mantle, affecting volcanic outgassing composition (the CH_4/CO_2 ratio), ocean chemistry, and the operation of the carbonate-silicate cycle (Gaillard et al., 2022; Schaefer et al., 2016).

Giant impacts during planetary formation and early bombardment can also drive atmospheric loss while simultaneously creating transient reduced atmospheres enriched in H_2 and potentially CH_4 (Zahnle et al., 2020). Future work should couple atmospheric escape and impact-driven atmospheric evolution to our model to track how early volatile loss shapes long-term habitability, particularly for planets around active K-dwarfs, which will be prime targets for atmospheric characterization with HWO (Arney, 2019; Mamajek & Stapelfeldt, 2023). Incorporating volcanic CH_4 outgassing, escape-driven oxidation, and impact-driven reducing conditions will establish a complete abiotic baseline against which to interpret methane detections in exoplanet atmospheres.

Our model is initialized after the magma ocean has crystallized. However, the magma ocean phase shapes the atmospheric composition and surface temperature of the early Earth. Volatiles incompatible with the magma ocean are released into the atmosphere (Elkins-Tanton, 2012), and the rate at which the magma ocean cools may also contribute to whether or not a planet develops plate tectonics (Schaefer & Elkins-Tanton, 2018). Thus, our model potentially underestimates the initial surface temperature and atmospheric volatile inventory of the planet, and does not account for how this early evolutionary process ultimately shapes Earth’s habitability.

Though this work does not couple to the magma ocean, VPLanet already includes a magma ocean model. Carone et al. (2025) updated MagmOc, VPLanet’s magma ocean module, demonstrating good agreement with previous magma ocean models (Elkins-Tanton, 2008; Nikolaou et al., 2019; Lichtenberg et al., 2021). Future work could integrate the MagmOc module into this whole-planet modeling framework, and continuously model the planet from the magma ocean phase into the solid mantle phase (e.g., Krissansen-Totton et al., 2021).

Finally, our 1-D model is not designed to account for complicated 3-D planetary phenomena like transitions to global glaciation states. Modeling a realistic ice-albedo feedback and snowball Earth transitions likely requires latitude-dependent insolation and heat transport that 1-D globally-averaged models cannot wholly capture (Hoffman et al., 2017; Abbot et al., 2013; Rose et al., 2017; Wilhelm et al., 2022). Similarly, substantial discrepancies exist between 1-D and 3-D models regarding the height of cloud formation, cloud optical properties, and the climate impacts of clouds (Yang et al., 2016; Way et al., 2018; Komacek & Abbot, 2019; Helling, 2019).

Rather than attempting to fully incorporate these 3-D phenomena into our 1-D framework, we envision a complementary modeling approach. Our evolutionary model provides realistic time-dependent boundary conditions for GCM simulations: atmospheric composition, volatile inventories, interior heat flux, and incoming stellar flux at snapshots throughout planetary evolution. GCMs can then simulate the spatial climate patterns, ice coverage, cloud distributions, and circulation for these evolved atmospheric states (Wolf et al., 2022). Conversely, GCM-derived quantities – such as a spatially-averaged albedo that includes realistic cloud distributions, ice surface fractions, and meridional heat transport efficiencies – can inform improved parameterizations in 1-D evolutionary models without sacrificing the computational speed required to cover large parameter sweeps over Gyr-timescales. This iterative dialogue between complementary modeling approaches leverages the strengths of each: 1-D models are appropriate for long-term interior-atmosphere coupling, while 3-D models are useful for simulating spatially resolved climate states with realistic cloud and ice feedbacks. The integration of different modeling techniques will likely be essential for interpreting future observations of exoplanet atmospheres, and for predicting which planets are most likely to maintain habitable conditions over geological time.

With the additional model developments described above, a whole-planet framework could be used to assess abiotic mimics of life on Earth-like exoplanets. For example, an improved version of the model could build on the work of Krissansen-Totton et al. (2021) by assessing

the evolution of abiotic O_2 in the atmospheres of exoplanets around Sun-like stars with a fully coupled core model. Our model would include the effect of an evolving core on the planet’s magnetosphere and the efficiency of hydrogen escape, which regulates the accumulation of abiotic O_2 . The improved model could also be used to investigate the abiotic generation of other biosignature gases, like CH_4 . This would expand on previous investigations of abiotic mimics by considering the evolution of the core-mantle system in addition to the star and planetary atmosphere.

In its current form, our model is well-poised for large parameter sweep explorations of Earth-like exoplanets throughout the HZs of FGK stars, both for the exploration of habitable planet properties as well as the simulation of variations in observable signatures of habitability. For this Earth validation study, we solely show our model coupled to solar evolution, but the `VPLanet` stellar model can simulate main sequence stars with masses up to $1.4M_\odot$. This mass range encompasses the Sun-like FGK stars that HWO will eventually target for observation.

Using preliminary HWO stellar target lists from [Mamajek & Stapelfeldt \(2023\)](#) and [Tuchow et al. \(2025\)](#), future work could simulate planets with the mass, radius, and tectonic regime of the Earth at various separations from FGK host stars. Since our coupled climate model performs full ASR and OLR calculations, it can be used to simulate the conditions at both the inner and outer edges of the HZ. At the inner edge, planets cease to be habitable when they enter the runaway greenhouse phase and at the outer edge, CO_2 scattering sets the maximum greenhouse limit ([Kopparapu et al., 2013](#)). To explore a wide range of planets, we can vary the initial CO_2 and H_2O inventories, as well as interior parameters like the radiogenics budget and the initial mantle and core temperatures. The unique luminosity evolution of each target star combined with various model parameters will shape the climate of a given planet over time, ultimately determining its potential long-term habitability.

Importantly, our evolutionary framework allows us to generate atmospheric states and synthetic HWO observations not just at a single endpoint, but at multiple snapshots throughout

a planet’s history. A single Earth-like planet may exhibit dramatically different spectroscopic signatures at 1 Gyr, 3 Gyr, and 5 Gyr as its atmosphere, interior, and surface evolve in response to changing stellar forcing and volatile cycling. This application of our model could reveal how the stability of the Earth system is affected by different stellar types. For example, the comparatively slow luminosity evolution of K-dwarfs may promote long-term climate stability over the extended main sequence lifetime of these stars, while the relatively rapid brightening of F-dwarfs may challenge the climate stability of Earth-like exoplanets. The time-resolved atmospheric states across different evolutionary epochs can then be used to produce simulated HWO observations, providing insight on the range of spectral characteristics consistent with purely abiotic planetary conditions. This temporal dimension is crucial: the diversity of abiotic habitable worlds HWO will sample reflects not only variations in initial conditions and host star properties, but also the natural evolution of planetary systems captured at different ages.

5 Conclusions

Here we have presented a coupled core-mantle-crust-climate model that reproduces geophysical, climate, and ocean properties of the pre-industrial Earth. Our model predicts that life may not be required to maintain a habitable planet resembling Earth-like conditions. This work thus contributes to the ongoing debate regarding the Gaia hypothesis by suggesting that life and habitability may be somewhat decoupled.

In terms of future applications, our model is intended to be applied to the evolution of rocky exoplanets. Our model shows that the evolution of the planetary interior plays an important role in determining the long-term habitability of Earth-like exoplanets. Thus, including realistically evolving planetary interiors is a crucial aspect of whole planet modeling and should account for potentially non-Earth-like compositions in order to illuminate the potential equilibrium states of rocky planets.

Though exoplanet interiors themselves will likely never be directly observable, the effects of interior evolution are apparent in the composition of the atmosphere and the albedo of the surface environment, both of which HWO will observe on Earth-like planets around Sun-like stars via reflected light spectroscopy. We have demonstrated that atmospheric states from our model can generate realistic reflected light spectra that capture the key differences between inhabited and uninhabited worlds, establishing an abiotic baseline for biosignature interpretation. The model presented here is compatible with modeling the long-term habitability of theoretical exoplanets around key HWO targets. For quantifying the null hypothesis and establishing an abiotic baseline, additional model development is required. Nevertheless, our core-mantle-crust-ocean-atmosphere-stellar model represents significant progress towards the ambitious goal of whole-planet modeling, which itself is a key step towards advancing our search for life in the universe.

Appendix

6 Constants

Table 2.4: All constants appear here in alphabetical order, with Greek letters appearing last.

Values without units are denoted as non-dimensional (n.d.).

Symbol	Parameter	Value
a_1	saturation vapor pressure scaling parameter for silicate weathering	0.3 [n.d.]
a_2	plate velocity scaling parameter	5.38 [n.d.]
A_b	bond albedo	0.29 [n.d.]
A_i	ice albedo	0.6 [n.d.]
A_{oc}	ocean albedo	0.1 [n.d.]
A_r	rock albedo	0.17 [n.d.]
A_{Earth}	surface area of the Earth	$5.1 \times 10^{14} \text{ m}^2$
B	viscosity-melt reduction coefficient	2.5 [n.d.]
$D_{\text{H}_2\text{O}}$	bulk H ₂ O distribution coefficient	0.01 [n.d.]
E_a	silicate weathering activation energy	$42 \times 10^3 \text{ J/mol}$
$E_{a,\text{man}}$	mantle viscosity activation energy	$3 \times 10^5 \text{ J/mol}$
E_{max}	maximum erosion rate	$3.2 \times 10^{-10} \text{ m/s}$
f_{cc}	fraction of Mg, Ca, K, and Na in the continental crust	0.08 [n.d.]
f_h	mass fraction of water in the serpentinized layer	0.03 [n.d.]

Continued on next page

Table 2.4: All constants appear here in alphabetical order, with Greek letters appearing last.

Values without units are denoted as non-dimensional (n.d.). (Continued)

Symbol	Parameter	Value
$F_{\text{sfw},0}$	present-day seafloor weathering flux	55,454 mol/s
$F_{\text{weather},0}$	present-day silicate weathering flux	380,257 mol/s
g	surface gravity of the Earth	9.8 m/s ²
k	thermal conductivity	4.2 W/m/K
L	present-day length of ocean trenches	6×10^7 m
L_w	latent heat of water	2.469×10^6 J/kg
$\bar{m}_c$	molar mass of CO ₂	44×10^{-3} kg/mol
$\bar{m}_{cc}$	average molar mass of Mg, Ca, K, and Na	32×10^{-3} kg/mol
$\bar{m}_w$	molar mass of water	18×10^{-3} kg/mol
M_{man}	mass of the mantle	4×10^{24} kg
$p\text{CO}_{2,0}$	present-day partial pressure of atmospheric CO ₂	33 Pa
$P_{\text{sat},0}$	present-day saturation vapor pressure	1391 Pa
$P_{\text{sat,ref}}$	reference saturation vapor pressure	610 Pa
R_c	radius of Earth's core	3481×10^3 m
R_g	universal gas constant	8.314 J/K/mol
RH	relative humidity	100%
R_{man}	radius of Earth's mantle	6371×10^3 m
Ra_{crit}	critical Rayleigh number	660 [n.d.]
S	salinity	35 ppt

Continued on next page

Table 2.4: All constants appear here in alphabetical order, with Greek letters appearing last.

Values without units are denoted as non-dimensional (n.d.). (Continued)

Symbol	Parameter	Value
S_0	solar constant	1369 W/m ²
$T_{\text{sat,ref}}$	reference saturation vapor temperature	273 K
$T_{\text{surf},0}$	present-day surface temperature	285 K
v_0	present-day plate speed	1.58×10^{-9} m/s
V_{man}	volume of Earth's mantle	9.1×10^{20} m ³
ν_{ref}	reference kinematic mantle viscosity	4.578×10^9 m ² /s
$x_{i,\text{max}}$	maximum ice surface fraction	0.12 [n.d.]
$x_{l,0}$	present-day land fraction	0.29 [n.d.]
$x_{\text{oc},0}$	present-day ocean fraction	0.71 [n.d.]
α_1	strength of dependence of basalt carbonation on atmospheric CO ₂	0.25 [n.d.]
α_2	strength of the dependence of continental weathering flux on plate speed	1.5 [n.d.]
β_1	partial pressure of CO ₂ scaling parameter for silicate weathering	0.55 [n.d.]
β_2	scaling exponent of Rayleigh number for plate speed	1/3 [n.d.]
γ	viscosity-melt reduction coefficient	6.0 [n.d.]
δ	viscosity-melt reduction exponent	6.0 [n.d.]

Continued on next page

Table 2.4: All constants appear here in alphabetical order, with Greek letters appearing last.

Values without units are denoted as non-dimensional (n.d.). (Continued)

Symbol	Parameter	Value
$\Delta E_{\text{H}_2\text{O}}$	depression of mantle viscosity activation energy due to water	$2.22 \times 10^6 \text{ K/wt. frac.}$
κ	thermal diffusivity	$1 \times 10^{-6} \text{ m}^2/\text{s}$
ξ	viscosity-melt reduction coefficient	$5 \times 10^{-4} \text{ [n.d.]}$
ρ_{man}	density of the mantle ($M_{\text{man}}/V_{\text{man}}$)	4395 kg/m^3
ρ_r	density of regolith	2500 kg/m^3
σ_{SB}	Stefan-Boltzmann constant	$5.67 \times 10^{-8} \text{ W/m}^2/\text{K}$
τ_{cl}	cloud opacity pressure	4050 Pa
$\chi_{\text{degas,c}}$	CO ₂ degassing efficiency	0.365 [n.d.]
$\chi_{\text{sub,c}}$	CO ₂ subduction efficiency	0.28798 [n.d.]
$\chi_{\text{degas,w}}$	water degassing efficiency	0.7156 [n.d.]
$\chi_{\text{regas,w}}$	water regassing efficiency	0.1248 [n.d.]
$\Omega_{\text{cal}}/[\text{Ca}^{2+}]$	marine calcite concentration	408.1 kg/mol

7 BMO Radiogenic Heating Calculation

Here we provide a detailed derivation of the effective core radiogenic heating rate, which includes contributions from an enriched basal magma ocean (BMO) layer. We calculate the ^{40}K enrichment needed to generate up to 3.5 TW of heat in the core, which our model requires to match observed upper mantle heat flow.

7.1 Derivation of Effective Core Heating Rate

The core and BMO energy balances may be written as

$$-M_C c_C \dot{T}_C = Q_{\text{rad},C} + Q_{IC} - Q_{\text{CMB}},$$

and

$$-M_{\text{BMO}} c_{\text{BMO}} \dot{T}_B = Q_{\text{CMB}} + Q_{\text{rad},\text{BMO}} - Q_{\text{BMO}},$$

respectively, where M denotes mass, c denotes specific heat, $\dot{T}$ denotes the time derivative of reservoir temperature, and Q denotes heat flow. The subscripts indicate the relevant reservoir: C and IC denotes the core and inner core respectively, CMB gives the core-mantle boundary, and BMO is the basal magma ocean layer. Thus, $Q_{\text{rad},\text{BMO}}$ represents the radiogenic heat flow in the BMO, and $Q_{\text{rad},C}$ represents the radiogenic heat flow in the core. We neglect tidal heating, as it is not a significant source of internal heating for Earth ([Ray et al., 1996](#)). The negative sign on the left-hand side represents secular cooling. Adding these equations yields

$$-M_C c_C \dot{T}_C - M_{\text{BMO}} c_{\text{BMO}} \dot{T}_{\text{BMO}} = Q_{\text{rad},\text{BMO}} + Q_{\text{rad},C} + Q_{IC} - Q_{\text{BMO}}.$$

We assume that the core and BMO cool at the same rate such that $\dot{T}_C \approx \dot{T}_{\text{BMO}}$. This assumption is justified if the more viscous solid mantle regulates the cooling rate of both reservoirs. This allows us to write

$$-\dot{T}_C (M_C c_C + M_{\text{BMO}} c_{\text{BMO}}) = Q_{\text{rad},\text{BMO}} + Q_{\text{rad},C} + Q_{IC} - Q_{\text{BMO}}. \quad (2.34)$$

The heat released at the inner core boundary is given by

$$Q_{IC} = \dot{T}_C \frac{dM_{IC}}{dT_C} (L_H + E_G), \quad (2.35)$$

where dM_{IC}/dT_C represents the change in inner core mass with temperature, and L_H and E_G are the latent and gravitational energy released per unit mass at the inner core boundary. Substituting into Equation (2.34) yields

$$-\dot{T}_C = \frac{(Q_{rad,BMO} + Q_{rad,C}) - Q_{BMO}}{(M_{BMO}c_{BMO} + M_Cc_C) + \frac{dM_{IC}}{dT_C}(L_H + E_G)}. \quad (2.36)$$

The mass of the BMO is likely much smaller than that of the core. To show this, we can compare the density of each, finding that

$$\frac{M_{BMO}}{M_C} = \left[\left(\frac{R_{BMO}}{R_C} \right)^3 - 1 \right] \frac{\rho_{BMO}}{\rho_C}, \quad (2.37)$$

where $R_{BMO} = R_C + L_{BMO}$ with $R_C = R_{CMB} = 3481$ km the radius of the core-mantle boundary (i.e., the outer surface of the core), and L_{BMO} is the thickness of the BMO layer. Consistent with seismic observations of the density jump at the CMB ([Dziewonski & Anderson, 1981](#)) and with high-pressure constraints showing that iron-enriched silicate melts are less dense than the metallic core ([Labrosse et al., 2007](#)), we assume that the BMO is half as dense as the core, $\rho_{BMO}/\rho_C \approx 0.5$. Using this conservative estimate of the density ratio, we find that M_{BMO} is maximally 20% that of the core for an initial BMO thickness of 400 km, as in [Labrosse et al. \(2007\)](#). The rapidly-thinning BMO is even less massive than the core – for example, $M_{BMO}/M_C \approx 0.044$ for a 100 km-thick BMO, and $M_{BMO}/M_C \approx 0.009$ for a 20 km-thick BMO layer. Thus, for the majority of the evolution $M_{BMO} \ll M_C$, and we neglect the $M_{BMO}c_{BMO}$ term in Equation (2.36). Combining the radiogenic heat from the BMO and core into one term, we find the final expression given in Equation (2.1) in Section 2.2:

$$-\dot{T}_C = \frac{Q_{rad,BMO+C} - Q_{BMO}}{M_Cc_C + \frac{dM_{IC}}{dT_C}(L_H + E_G)}. \quad (2.38)$$

7.2 Required Radiogenic Enrichment

To estimate the possible heating power generated by a radiogenically enriched BMO, we account for heating from the four radioactive isotopes included in our model: ^{40}K , ^{232}Th , ^{235}U , and ^{238}U . We define the present-day heating contribution from each isotope as,

$$Q_{\text{rad,BMO},i,0} = M_{\text{BMO}} \times x_{\text{BMO}} \times c_i \times h_i, \quad (2.39)$$

where M_{BMO} is the mass of the BMO, x_{BMO} is the concentration of the parent species in the BMO, c_i is the relative abundance of each isotope, and h_i is the total radiogenic heating power per unit mass of the isotope. Table 2.5 provides the values of x_{BMO} , c_i , and h_i for each species. To solve for the mass of the BMO, we can write

$$M_{\text{BMO}} = \rho_{\text{BMO}} \frac{4\pi}{3} ((R_{\text{CMB}} + L_{\text{BMO}})^3 - (R_{\text{CMB}})^3) \quad (2.40)$$

where ρ_{BMO} is the density of the BMO, and we solve for the volume of a L_{BMO} -thick BMO using the radius of the core mantle boundary, $R_{\text{CMB}} \approx 3481$ km.

Labrosse et al. (2007) modeled a primordial BMO of $L_{\text{BMO}} \sim 400$ km that fractionally crystallizes to ~ 8 km over 4.5 Gyr. However, as the BMO solidifies and thins, incompatible elements are continuously expelled into the residual liquid, such that a thinner present-day BMO is expected to be more concentrated than a thicker earlier one. We assume a present-day BMO thickness of $L_{\text{BMO}} = 4$ km, which we interpret as the highly enriched residual of a larger primordial BMO. Finally, we assume a BMO density of $\rho_{\text{BMO}} = 6000 \text{ kg m}^{-3}$, which is slightly denser than the solid mantle due to higher pressures. Substituting into Equation 2.39 to solve for the individual radiogenic heat contributions for each of the 4 species, we find that the present-day total $Q_{\text{rad,BMO},0} = 3.67$ TW for a 4 km-thick BMO. This calculation demonstrates that attributing 3.5 TW of core heating to a radiogenically enriched, thin present-day BMO is physically plausible.

Parameter	Species	Value	Reference
Bulk Silicate Earth (BSE) concentration	K	260 ppmw	McDonough (2025)
	Th	77.6 ppbw	
	U	20.6 ppbw	
BMO enrichment factor (x_{BMO})	all	$200\times$ BSE	Labrosse et al. (2007, their Figure 4)
Isotopic abundance (c_i)	^{40}K	1.28×10^{-4}	Korenaga (2006)
	^{232}Th	1	
	^{235}U	7.2×10^{-3}	
	^{238}U	0.9927	
Radiogenic heat per unit mass (h_i , $\mu\text{W}/\text{kg}$)	^{40}K	27.9	Turcotte & Schubert (2002)
	^{232}Th	26.9	
	^{235}U	569	
	^{238}U	93.7	

Table 2.5: Physical and Earth properties used for radiogenic heating calculations in the basal magma ocean (BMO).

CHAPTER 3

MODELING THE ABIOTIC EARTH AROUND HWO TARGET STARS

All models are wrong, but some are useful.

– George Box

1	Methods	98
1.1	Dynamic Model Initialization	98
1.2	Continental and Seafloor Weathering Assumptions	99
1.3	Fractional Ice Coverage	101
1.4	Stellar Cases and HZ Parameter Sweeps	105
2	Results	107
2.1	Representative Climate Evolutions	107
2.2	Model (a)	112
2.3	Model (b) (Nominal)	115
2.4	Model (a) and (b) Comparison	119
3	Discussion	121
3.1	Comparison to Previous Work	122
3.2	Implications for the Abiotic vs. Biotic HZ	126

3.3	Future Work	127
4	Conclusion	128

In this work, we apply the self-consistent whole-planet abiotic Earth model from Chapter 2 to explore the time-evolving boundaries of the HZ around F, G, and K stars. Characterizing these boundaries is directly relevant to the Habitable Worlds Observatory (HWO), which will search for biosignatures in the reflected-light spectra of rocky planets around nearby FGK stars and whose target list is drawn largely from stars expected to host worlds within their HZs. Crucially, the HZ is not static: its edges migrate as a star brightens over its main sequence lifetime and as a planet’s interior evolution modulates the volcanic degassing that sustains its atmosphere, so whether a given HWO target is likely to host a habitable and *observable* planet depends on coupled stellar and geophysical processes that few habitability models evolve together fully self-consistently. By coupling an explicit interior model of mantle thermal evolution and volatile outgassing to a carbon cycle and radiative-convective climate model, we produce HZ phase diagrams that capture the coupled evolution of stellar luminosity, volcanic degassing, and weathering feedbacks across the full main sequence lifetimes of selected FGK stars.

As we will show, the outer HZ proves far more sensitive to poorly constrained carbon-cycle assumptions—particularly the behavior of seafloor weathering during glaciations—than the inner edge, which is governed by the comparatively robust runaway greenhouse limit; this asymmetry, together with a pronounced tendency of K-dwarf planets toward permanent glaciation, has direct consequences for HWO target prioritization that we develop in [section 3](#). In [section 1](#), we describe various improvements made to the base model, including an improved initialization scheme and a treatment for fractional ice coverage. Furthermore, we identify model sensitivities to seafloor and continental weathering, which we briefly explore in [section 2](#). We then describe the selected FGK stars and the setup for our HZ parameter sweeps. In [section 2](#), we show representative climate evolutions and HZ phase diagrams for selected FGK stars. We compare our results to previous works and discuss future model improvements in [section 3](#), and offer conclusions based on these preliminary results in [section 4](#).

1 Methods

The underlying `vPlanet` model used in this work is described in the previous chapter. Here, we briefly describe model improvements made to enable the application of the abiotic Earth model to a broad range of star-planet separations. These improvements include (1) a dynamic initialization scheme to find stable warm climates by iteratively re-partitioning CO_2 between the surface and interior, (2) adjustments to the parameterization of continental and seafloor weathering fluxes, and (3) a self-consistent treatment of fractional ice coverage for both rock and ocean surface types. Lastly, we describe the experimental setup for the HZ parameter sweep of high-priority HWO target stars. We additionally note that `clima` determines the fraction of atmospheric CO_2 in the gas phase based on the total pressure and surface temperature; condensed CO_2 , whether solid or liquid, is excluded from the greenhouse calculation, which places a physical upper limit on the greenhouse warming achievable at the outer edge of the HZ (Turbet et al., 2017).

1.1 Dynamic Model Initialization

The initial thermochemical state of each terrestrial planet is determined through a nested iterative procedure that enforces self-consistency among the surface temperature, atmospheric composition, and carbon-water partitioning prior to the onset of coupled evolution. The thermal interior state is first established by calling the mantle and core verification routines, which set the upper and lower mantle temperatures, core-mantle boundary temperature, and all derived viscosity and melt fraction quantities through a two-iteration melt loop. The carbon reservoir is then partitioned between the mantle, tectonic plate, and surface (atmosphere plus ocean) by bisection on the surface fraction of the total planetary carbon budget, with the remainder split equally between the mantle and plate reservoirs.

The following initialization scheme is summarized in Figure 3.1. For each candidate CO_2 partition, an inner fixed-point iteration converges the surface temperature by alternately

computing the atmospheric CO₂ partial pressure from the current carbon and water state, updating the surface albedo assuming ice-free conditions and a fixed ocean area fraction, and calling the radiative-convective climate model to obtain a new surface temperature. The inner iteration uses a damping factor of 0.5 to suppress oscillations arising from the bistable nature of the ice-albedo feedback. The outer bisection targets a surface temperature in the range 295—320 K, with the bisection direction determined by both the converged temperature and whether the atmospheric CO₂ partial pressure exceeds the saturation vapor pressure of CO₂ at the converged temperature, computed from a Clausius-Clapeyron fit to the CO₂ phase diagram. This prevents the bisection from accumulating atmospheric CO₂ beyond the condensation threshold at cold temperatures. If no warm solution is found within 30 outer iterations, the partition yielding the highest surface temperature is retained. Following convergence, a single equilibration pass with the full liquid ocean present redistributes carbon between the atmosphere and ocean via Henry’s law dissolution before the evolution loop begins, ensuring continuity between the initialized state and the first timestep.

1.2 Continental and Seafloor Weathering Assumptions

The model described in the previous chapter adopted the continental and seafloor weathering equations of [Foley \(2015\)](#) (see [Equation 2.8](#) and [Equation 2.9](#)). In initial testing of this model, we made two key findings: (1) the nominal model’s weathering feedbacks vastly exceed initial CO₂ degassing rates such that planets at separations >1 AU cannot maintain habitable temperatures, and (2) evolutionary tracks are highly sensitive to the strength of the weathering feedbacks especially whether or not seafloor weathering proceeds during a planetary snowball. Regarding (1), at the beginning of the evolution, when the interior is hot and the upper melt fraction high, the nominal model predicts initial seafloor weathering rate $\sim 10^8$ kg/s, or 50,000 times the present-day seafloor weathering rate. To mitigate this potential model defect, we implement a hard ceiling on the seafloor weathering of 10^7 kg/s. Regarding (2), we found that the strong dependence on $p\text{CO}_2$ relative to present-day values

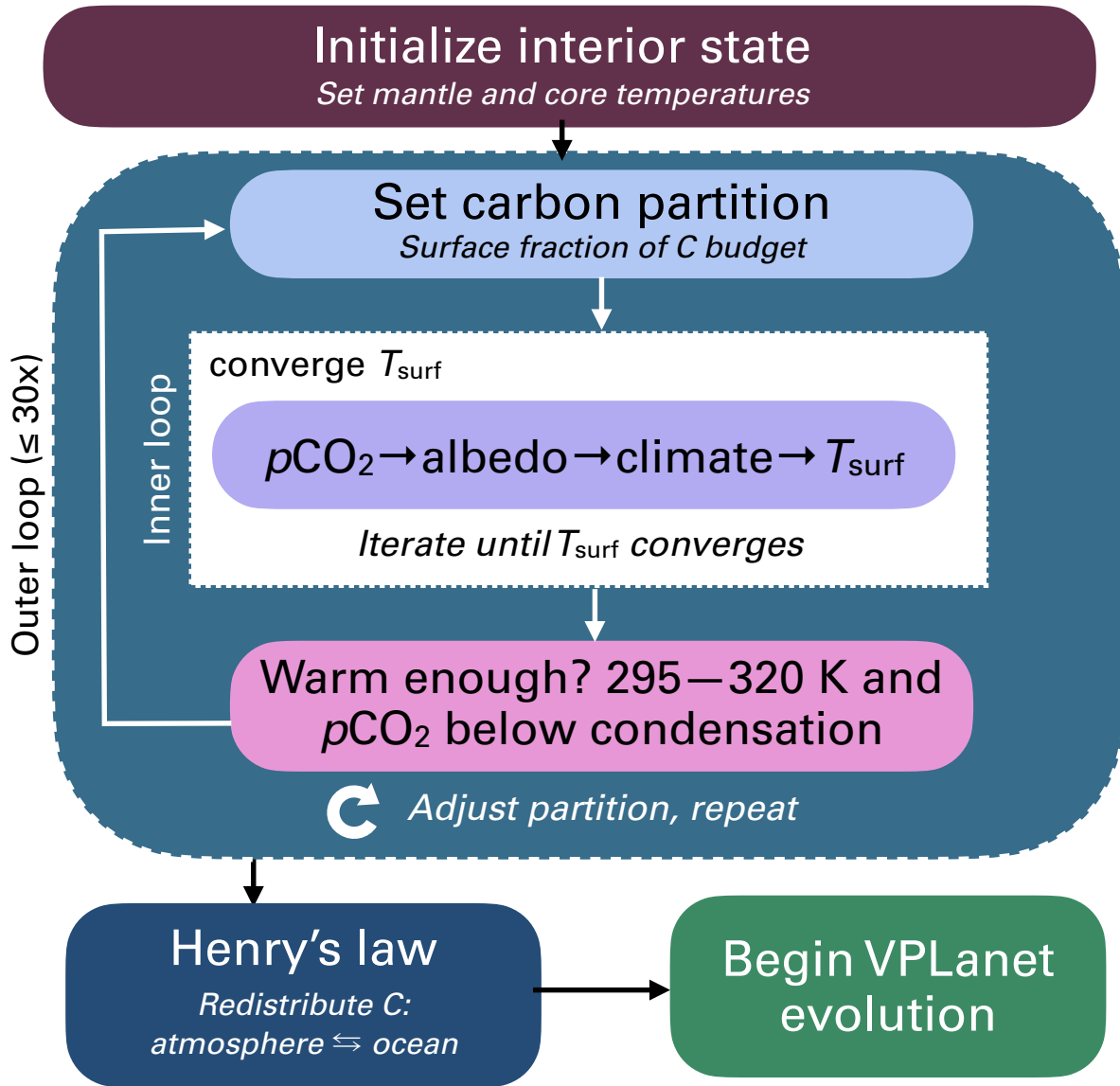


Figure 3.1: Nested loop structure of the dynamic initialization scheme. After the thermal interior state is set once, an outer bisection loop partitions the planetary carbon budget between the surface and interior reservoirs, while an inner fixed-point loop converges the surface temperature for each candidate partition by alternately computing the atmospheric CO_2 partial pressure, updating the ice-free surface albedo, and calling the radiative-convective climate model, with a damping factor of 0.5 applied to suppress oscillations from the ice-albedo feedback. The outer loop accepts a partition only when the converged surface temperature falls within 295–320 K and the atmospheric $p\text{CO}_2$ lies below the CO_2 condensation threshold, retaining the warmest solution if no warm state is found within 30 iterations. A single Henry’s-law equilibration pass then redistributes carbon between the atmosphere and ocean before the coupled evolution begins.

prevented the accumulation of atmospheric CO_2 on planets in the outer HZ. We therefore explored two modifications to the nominal model. To isolate the effect of the ocean fraction dependence in the seafloor weathering relation (Equation 2.8), we ran a case using the nominal Foley (2015) weathering exponents but with the $x_{oc}/x_{oc,0}$ term removed. Since seafloor weathering no longer scales with ocean coverage, this ensures seafloor weathering continues on a globally glaciated planet, although continental weathering ceases independently because its exposed rock fraction term $x_r/x_{l,0}$ goes to zero when the surface is ice-covered. Second, we explored a case with reduced $p\text{CO}_2$ scaling exponents to weaken the overall strength of the weathering feedbacks. In Figure 3.2, we compare the $p\text{CO}_2$ scaling factors of the nominal Foley (2015) model to this reduced case. The nominal exponents are 0.25 for seafloor weathering and 0.55 for continental weathering; the reduced case uses exponents of 0.18 and 0.1 respectively. We apply both modified models alongside the nominal model to explore the habitable zones of F, G, and K stars and compare the resulting climate evolutions.

1.3 Fractional Ice Coverage

Realistically capturing the ice-albedo feedback, a key control on snowball transitions and HZ stability, requires relating surface ice coverage to climate, yet our radiative-convective model resolves only a single global-mean surface temperature. Rather than couple a latitudinally-resolved EBM or GCM at every timestep, we use a seasonal EBM to derive a parameterization mapping global-mean temperature to ice fraction – separately for land and ocean, which differ in heat capacity and hence in their freeze and melt thresholds– bridging the spatially-resolved physics into our 1D model at negligible cost. To quantify the relationship between global surface temperature and ice fraction, we use POISE (“Planetary Orbit-Influenced Simple EBM”) (Deitrick et al., 2018; Wilhelm et al., 2022), a one-dimensional seasonal EBM in VPLanet (Barnes et al., 2020). POISE has been used to reproduce Earth’s annual climate and its Milankovitch cycles (see North & Coakley Jr, 1979; Huybers & Tziperman, 2008). The model calculates the flow of heat between neighboring latitude bins, each of which are

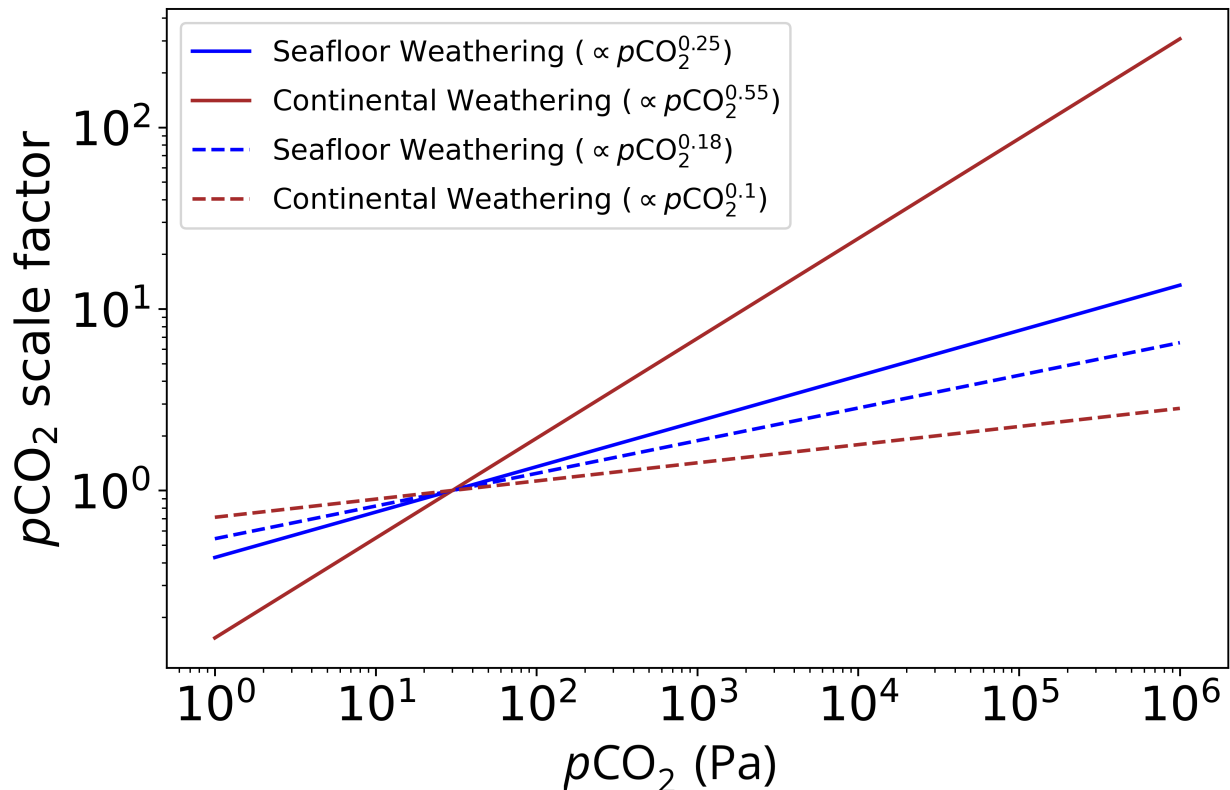


Figure 3.2: A comparison of $p\text{CO}_2$ scale factors for the continental and seafloor weathering relations of [Foley \(2015\)](#) explored in this study. The choice of exponent effects the strength of the weathering dependency on the partial pressure atmospheric CO_2 .

partitioned into land and ocean with differing heat capacities ([Wilhelm et al., 2022](#)). Using relations from [Huybers & Tziperman \(2008\)](#), `POISE` self-consistently calculates global surface temperature and ocean and land ice coverage based on bedrock depression, lithospheric rebound, and ice sheet height. Following previous work, the model assumes a constant albedo of 0.6 for ice ([Shields et al., 2013](#); [Deitrick et al., 2018](#)). For a full description of the model used here, please see [Wilhelm et al. \(2022\)](#).

Since `POISE` determines global surface temperature and ice fraction for static inputs, we sample 100 solar luminosity values ranging from 0.14–1.5 $L_\odot$ and 80 semi-major axes from 0.8–1.5 AU for a total of 8,000 simulations. The ranges explored for both luminosity and semi-major axis are selected to ensure sufficient coverage of parameter space for our HZ investigations, which include evolving stellar luminosity. For each simulation, we assume the

baseline land and ocean fractions (0.34 and 0.66, respectively) in each latitude bin remain near their present-day values. Each `VPLanet` run in the parameter sweep produces one equilibrium record with a global average surface temperature (T_{global}), the area of land covered by ice ($f_{\text{ice,land}}$), and the area of ocean covered by ice ($f_{\text{ice,ocean}}$).

We show the outputs of these parameter sweeps in the contour plots in [Figure 3.3](#). These plots show that, for ice on both land and ocean, smaller planet-star separations and higher luminosities are associated with lower ice surface fractions, while larger planet-star separations and lower luminosities are associated with higher ice surface fractions. Given its higher heat capacity relative to land, sea ice has a lower temperature threshold for total melting and total freezing. As a result, fractional sea ice coverage extends to longer planet-star separations and lower stellar luminosities. The final panel of the subplot shows global mean temperature as a function of semi-major axis and stellar luminosity. Higher surface temperatures are associated with planets closer to higher-luminosity stars, while colder temperatures are associated with planets further from lower-luminosity stars. The temperature contour maps cleanly onto the land and ocean ice fraction contours.

Ice fraction lookup tables for land and ocean surfaces were derived from the `POISE` parameter sweep. Equilibrium ice fractions were normalized by the assumed land and ocean area fractions, binned to a 0.5 K temperature grid, and summarized by the median within each bin. The resulting binned relationship was enforced to be monotonically decreasing in temperature using isotonic regression, which adjusts only those points that violate monotonicity rather than modifying the full sequence. This guarantees that each global-mean surface temperature maps to a single ice fraction, so the parameterization returns a unique, well-defined albedo at every timestep. Residual flat segments introduced by isotonic regression, where consecutive knots share identical values, were eliminated by replacing each such run with a linear ramp between its bounding values, ensuring strict monotonicity. A Piecewise Cubic Hermite Interpolating Polynomial (PCHIP) was then fit to the resulting knot sequence. PCHIP was selected over standard cubic splines because it preserves local monotonicity and

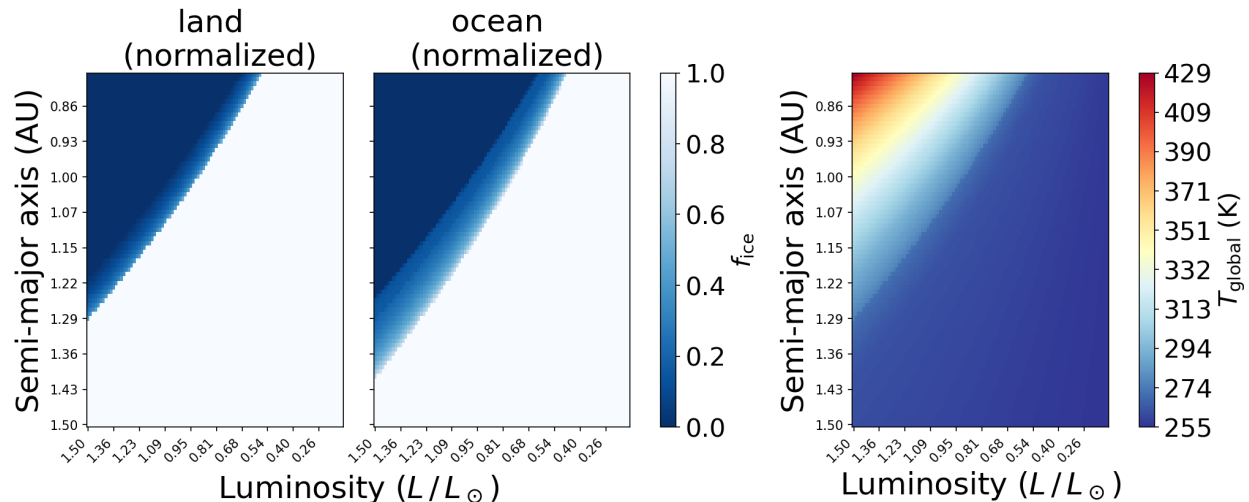


Figure 3.3: A POISE parameter sweep over luminosity and planet-star separation gives a 2-dimensional grid of land (a) and ocean (b) ice fractions and global average surface temperature (c). As expected, planets closer to higher-luminosity stars are characterized by higher global average surface temperatures and lower ice fractions. Similarly, planets farther from lower-luminosity stars have lower global average surface temperatures and higher ice fractions.

does not overshoot between knots, producing a smooth curve. The knot arrays are stored directly in the model source and evaluated at runtime via binary-search linear interpolation on the 0.5 K grid, for which the approximation error relative to the full PCHIP curve is negligible. The resulting relationships between land and ocean ice surface fraction and global average surface temperature are shown in Figure 3.4. Below distinct temperature thresholds (271 K for $f_{ice,land}$, 263 K for $f_{ice,ocean}$), land and ocean surfaces become completely covered in ice. Similarly, above surface temperature thresholds of 293 K and 285 for land and ocean respectively, ice completely melts. Between these temperature boundaries, fractional ice coverage monotonically decreases as a function of increasing temperature. The derivatives of the fractional ice coverage-temperature curves ($\frac{df}{dT}$) for both land and ocean remain negative, demonstrating this monotonic behavior.

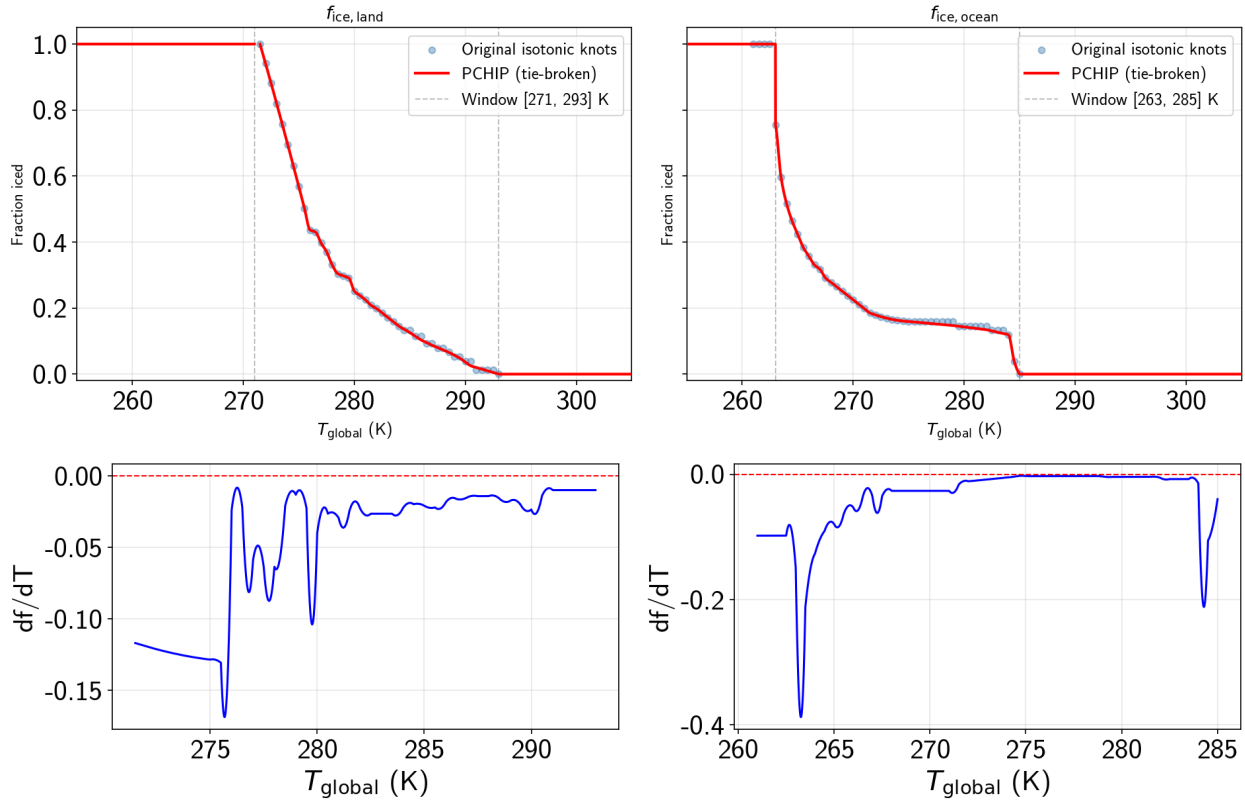


Figure 3.4: Flattening the 2-dimensional grid gives an ice fraction-surface temperature curve for both land and ocean surface-types. In the bottom 2 panels, we plot the corresponding derivatives of the fractional ice coverage-temperature curves ($\frac{df}{dT}$). Isotonic regression is used to enforce monotonicity and ensure smooth model behavior, and a Piecewise Cubic Hermite Interpolating Polynomial is fit to the data.

1.4 Stellar Cases and HZ Parameter Sweeps

We next apply our abiotic Earth models to one F, one G, and one K star. Our G star is based on the Sun, with $M_{\star} = M_{\odot}$. For the F star, we choose a star with $M_{\star} = 1.2M_{\odot}$ and for the K star, we choose a star with $M_{\star} = 0.8M_{\odot}$. These stars roughly correspond to “Tier A” HWO target stars, such as 19 Ceti (F7V) and 107 Piscium (K1V). For each spectral type, we calculate a unique grid to interpolate and solve for climate following the same procedure described in the previous chapter. We maintain the same grid points as the Earth-model for the surface temperature, partial pressures of CO_2 and H_2O , surface albedo, and instellation at the planet. We recalculate the grids for a range of stellar effective temperatures appropriate

to each spectral type. For the F star, the grid spans stellar effective temperatures from 6000 to 7500 K (6 points, spaced at 250 K intervals). For the G star, the grid spans stellar effective temperatures from 3500 to 5500 K (8 points, spaced at 250 K intervals).

To determine the range of planet-star separations for the habitable zone parameter sweeps, we computed the inner and outer edges of the habitable zone for representative stars of each spectral class using the “optimistic” empirical flux limits of [Kopparapu et al. \(2013\)](#). The inner edge corresponds to the runaway greenhouse limit and the outer edge to the maximum greenhouse limit, with effective stellar flux scaled by stellar luminosity and temperature following [Kopparapu et al. \(2013\)](#). For each spectral class, we selected a representative target from the HWO Tier A target list ([Mamajek & Stapelfeldt, 2023](#)) and computed the corresponding orbital separations for the star at the beginning and end of its main sequence evolution ([Table 3.1](#)). The minimal inner edge corresponds to the star at the beginning of its main sequence lifetime, when it is dimmest, while the maximal outer edge corresponds to the star at the conclusion of its main sequence lifetime. For each spectral class case, we ran 120 simulations uniformly sampling the separation range shown in [Table 3.1](#), for a total of 360 simulations across all three stellar types. The parameter sweep for each grid was extended slightly beyond the nominal HZ edges to capture limit cycle and snowball states near the boundaries, as climate transitions are expected to occur within this margin.

Table 3.1: Stellar properties and habitable zone separations for HWO Tier A targets.

Spectral Class	Example	$M_{\star} [M_{\odot}]$	Separation [AU]
F	19 Ceti	1.2	1.2 – 2.9
G	Sun	1.0	0.8 – 1.5
K	107 Piscium	0.8	0.5 – 1.1

Each simulation timestep was independently classified into one of four climate regimes based on instantaneous surface temperature. A timestep was classified as *habitable* if the surface temperature exceeded 263 K, which is the threshold for freezing the entire planet including surface oceans, and as *snowball* if it fell below this threshold. A *limit cycle* was identified when a simulation exhibited at least two crossings of the 263 K freezing threshold

after an initial settling period of 10% of the total run length; all time steps spanning from the first to last crossing were assigned to this regime. A *runaway greenhouse* state was assigned to simulations that terminated before the star exited the main sequence while still in a warm state. Early termination occurs when the surface temperature exceeds the upper boundary of the climate lookup table (500 K) within a single timestep, causing the integration to halt. Because the model timestep is 1000 years, the onset time of the runaway greenhouse is uncertain by at most this amount, which is negligible relative to the Gyr timescales of interest.

2 Results

In Figure 3.5, we compare the luminosity evolution of the F, G, and K-dwarf stars explored in this work. The more massive F star begins its main sequence lifetime more than 1.5 times as bright as the Sun. After only ~ 4.5 Gyr, it has moved off the main sequence with a final luminosity $> 2.5L_{\odot}$. By contrast, the `VPLanet` stellar model predicts that the Sun is 70% as bright as present-day at the start of its evolution, and after 8 Gyr, it has a final luminosity of $\sim 1.3L_{\odot}$. Finally, the cooler and smaller K star has a luminosity of $\sim 0.24L_{\odot}$ following an initial spike. Owing to its slow evolution on the main sequence, its final luminosity after 10 Gyr is only $\sim 10\%$ brighter.

2.1 Representative Climate Evolutions

We first show representative climate and carbon cycle flux evolutions for the Earth at 1 AU from the Sun. Each figure contains three subplots: Earth’s climate evolution under the assumptions of (a) the model with reduced weathering and continuous seafloor weathering during snowball events, (b) the nominal model that ceases seafloor weathering during snowball events, and (c) the nominal model that allows seafloor weathering to proceed during snowballs. Examining these evolutions is instructive for comparing the different models and

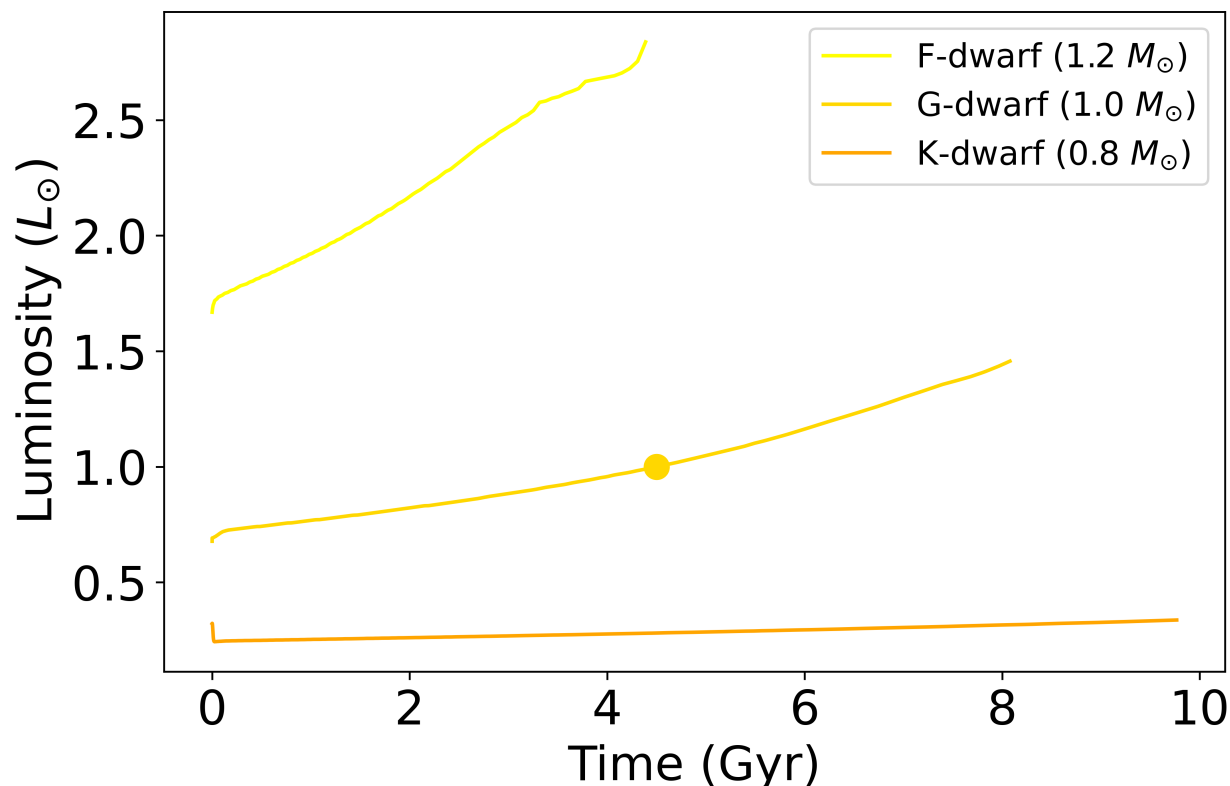


Figure 3.5: A comparison of the luminosity evolutions of an F, G, and K-dwarf star show that more massive F-dwarfs burn bright and move off the main sequence relatively quickly. Conversely, K-dwarfs are less luminous and have longer main sequence lifetimes compared to the Sun. The dot on the G-dwarf track represents present-day (4.5 Gyr) solar luminosity.

decoding the underlying model behavior of the HZ phase diagrams, which comprise the main result of this study.

In [Figure 3.6](#), we show representative climate evolutions for the abiotic Earth at a distance of 1.0 AU from the Sun. Each subplot shows both flux (left y-axis) and surface temperature (right y-axis) as a function of time in Gyr. In [Figure 3.6a](#), we see that under the assumptions of the reduced weathering model with continuous seafloor weathering, seafloor weathering remains lower than the arc volcanic flux and the planet remains habitable for the entirety of its evolution. Seafloor and continental weathering cease after ~ 7 Gyr of evolution, after which all CO_2 has been scrubbed from the atmosphere in response to rising temperatures induced by the brightening Sun.

By contrast, [Figure 3.6b](#) shows that the nominal model without continuous seafloor weath-

ering predicts intermittent snowball events from 1.5–4 Gyr. The inset plot in [Figure 3.6b](#) zooms in on one of these limit cycles ~ 1.58 Gyr, and shows that a temperature increase coincides with a rapid increase in the continental and seafloor weathering fluxes from zero. This occurs because sufficient $p\text{CO}_2$ has accumulated in the atmosphere to melt the prior snowball state, exposing rock and ocean and reigniting both weathering processes. The sudden onset of seafloor and continental weathering reduces the atmospheric CO_2 , leading to a decrease in temperature. Since seafloor weathering and continental weathering are both dependent on a $p\text{CO}_2$ scaling factor, they also decrease. However, the carbon sinks are not suppressed enough to prevent the planet from crashing back to a snowball state. This cycle repeats several times until the stellar flux is high enough to sustain a stable, habitable climate state ~ 4 Gyr.

Lastly, [Figure 3.6c](#) shows the evolution of the same planet under the assumptions of the nominal model with continuous seafloor weathering. Despite starting as habitable, the planet crashes to a snowball state after ~ 1.2 Gyr. Thereafter, this model predicts that the planet remains in a snowball state, until it enters a runaway greenhouse at the end of the star’s main sequence lifetime. Zooming in on the temperature crash ~ 1.3 Gyr, the inset plot in [Figure 3.6c](#) shows that temperature increases following an increase in $p\text{CO}_2$. The strong dependence of seafloor and continental weathering on $p\text{CO}_2$ causes both fluxes to increase in kind, scrubbing CO_2 out of the atmosphere and triggering the crash to the snowball state. Thereafter, continuous seafloor weathering prevents CO_2 from accumulating and melting the ice-covered planet.

Across all three models, the degassing flux secularly decreases as the planet cools. Once the interior has cooled enough that the upper mantle melt fraction goes to zero and the mantle fully solidifies, degassing of CO_2 ceases (though arc volcanic flux continues). We note that the timing of degassing cessation differs for each model, though the interior equations remain unchanged, because the temperature gradient of the interior is dependent on the surface temperature of the planet. For the model (c) case, the cold surface during the terminal

snowball event cools the interior more rapidly, leading to the loss of degassing ~ 5.175 Gyr. By contrast, the same condition is met in model (a) after 5.5 Gyr, and in model (b) after 5.45 Gyr, as the surface temperature of the planets at this time differ only modestly. Thus, surface temperature differences across all cases explored here will lead to relatively minor variations in the timing of mantle solidification. By the same token, mantle solidification limits a source of atmospheric CO_2 , which will affect the remainder of the planet’s climate evolution.

The three model configurations produce qualitatively distinct climate behaviors for the present-day Earth. Model (a), with suppressed weathering and continuous seafloor weathering, predicts a stably habitable planet throughout the Sun’s main sequence lifetime, with weathering fluxes that remain below the arc volcanic outgassing rate. Model (b), the nominal model without seafloor weathering during snowball events, predicts intermittent limit cycle behavior from ~ 1.5 –4 Gyr before settling into a stable habitable state. Critically, the cessation of seafloor weathering during snowball events allows atmospheric CO_2 to accumulate via arc volcanism, enabling recovery from glaciated states. This suggests that the limit cycle behavior of the nominal model is itself a consequence of the overpowered weathering feedbacks: the strong $p\text{CO}_2$ dependence drives the planet into a snowball, but the absence of seafloor weathering during the snowball permits eventual recovery. Model (c), the nominal model with continuous seafloor weathering, predicts that the Earth enters an unrecoverable snowball state after ~ 1.2 Gyr. Since this is clearly inconsistent with Earth’s observed climate history, we do not use model (c) to explore HZ phase space in what follows. However, it serves as a useful end-member case for illustrating the sensitivity of model outcomes to assumptions about seafloor weathering during snowball events. Hereafter, we present results from models (a) and (b) only.

We compile the models (a) and (b) planet evolutions into HZ phase diagrams which represent the climate state of a planet in semi-major axis bins as a function of time. Thus, each pixel in the phase diagram represents a snapshot in time of the evolution of an abiotic

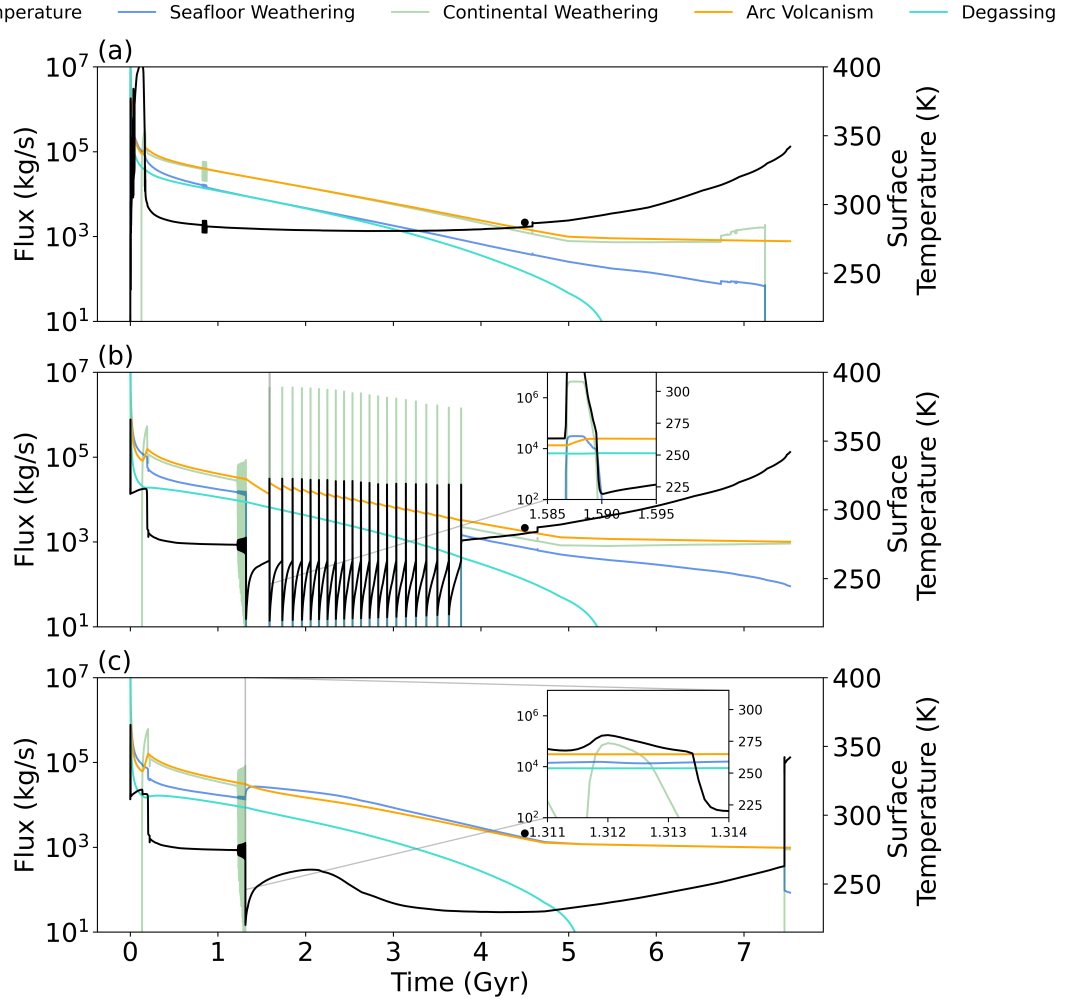


Figure 3.6: Carbon cycle flux and surface temperature evolution of an abiotic Earth at 1.0 AU from the Sun under three model assumptions: (a) reduced weathering with continuous seafloor weathering, (b) nominal weathering without seafloor weathering during snowball events, and (c) nominal weathering with continuous seafloor weathering. Insets zoom in on individual limit cycle and transient deglaciation events. Black dots show the surface temperature of the pre-industrial Earth at 4.5 Gyr. While models (a) and (b) demonstrate feasibly Earth-like climate evolutions, model (c) predicts an unrealistic terminal snowball.

Earth at a given distance from its star. In the next section, we describe the HZ phase diagrams predicted by model (a).

2.2 Model (a)

In Figures 3.7, 3.8, and 3.9, we show the HZ phase diagrams for F, G, and K stars respectively using our model with a suppressed weathering response to $p\text{CO}_2$ and continuous seafloor weathering during snowballs. Across all three plots, we note that no planets enter a limit cycle regime under the assumptions of this model. We describe the behavior of each phase diagram and compare it to its nominal model counterpart.

Figure 3.7 shows the climate regimes of an abiotic Earth at various separations from the F-type star using our suppressed weathering model. Under these model assumptions, the outer edge of the HZ initially moves inward as the star brightens. After this initial brightening, the HZ boundary lies at 1.7 AU. Identical to the nominal model, planets at the inner edge (1.2 AU) enter a runaway greenhouse after 1.5 Gyr, and the inner edge of the HZ moves outward. After 2.7 Gyr, the outer edge gradually moves outward and formerly snowball planets experience late deglaciations before the F star evolves off of the main sequence at 4.5 Gyr. We also note that planets at 1.9 AU appear to re-enter snowball states after 4.2 Gyr.

Figure 3.8 shows the climate regimes of an abiotic Earth planet at various separations from a $1.0M_\odot$ Sun-like G-type star. Our suppressed weathering model initially predicts a hard boundary between habitable and snowball planets at ~ 1.2 AU, or 0.02 AU farther than the nominal model. Identical to the nominal model, after 3.5 Gyr, the abiotic Earth at 0.8 AU enters a runaway greenhouse regime, and the inner edge moves outward as the Sun brightens. On this plot, we again note the location of the present-day Earth with a circled plus. We note that this model predicts that the present-day Earth would remain habitable over the majority of the Sun’s lifetime, entering a runaway greenhouse regime after about 7.5 Gyr. After 4.5 Gyr, the outer edge of the HZ moves outward, and snowball planets at a distance > 1.2 AU gradually become habitable following late deglaciations. Though this model does not predict limit cycles, it does predict that HZ planets > 1.05 AU become

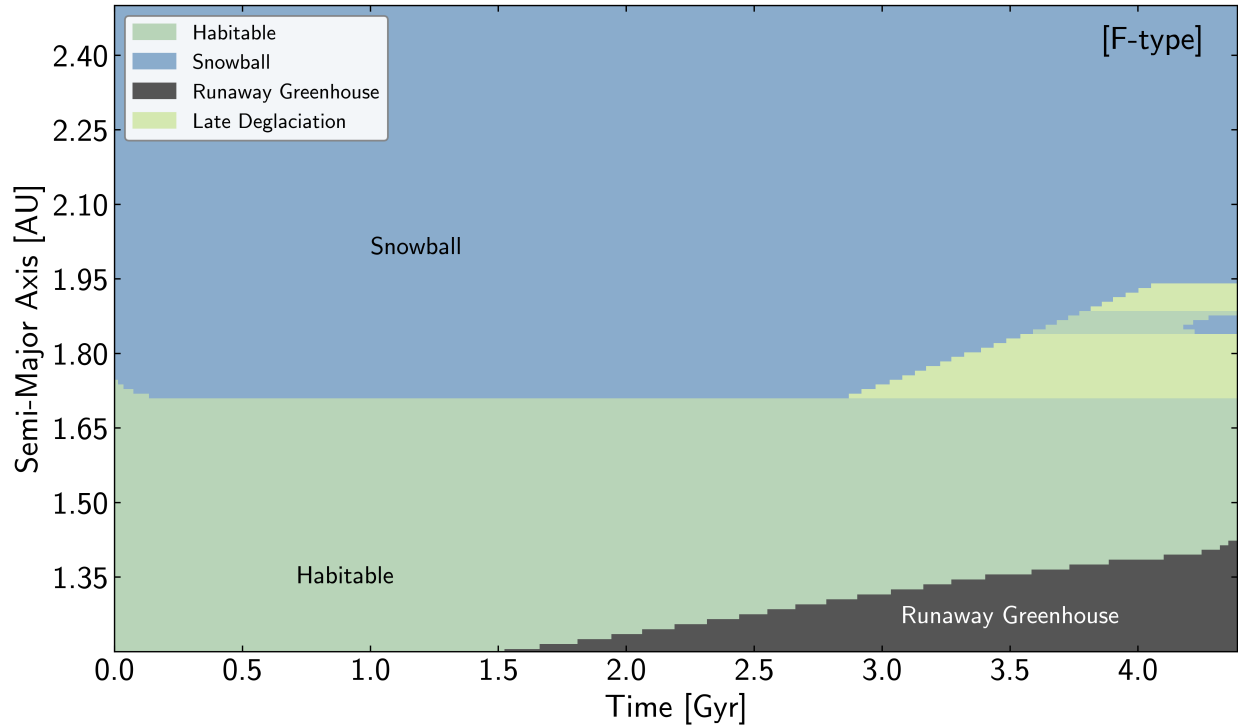


Figure 3.7: Climate phase diagram for an abiotic Earth orbiting a $1.2 M_{\odot}$ F-type star under a suppressed weathering rate. The outer HZ boundary settles at ~ 1.7 AU, and planets at ~ 1.2 AU enter a runaway greenhouse after ~ 1.5 Gyr, consistent with the nominal model. Outer HZ planets that enter limit cycles in the nominal case (Figure 3.10) remain stably habitable here. The star exits the main sequence at 4.5 Gyr.

snowballs after ~ 3.5 Gyr. All planets ultimately recover habitable states towards the end of the evolution, but at differing rates. The suppressed weathering model predicts that the final outer edge lies ~ 1.45 AU. We note the occurrence of perplexing features in this phase diagram, including the crash to and recovery from snowball states of mid-HZ planets, as well as the snowball/habitable “stripes” appearing at the end of the evolution for outer HZ planets.

Figure 3.9 shows the climate regimes of an abiotic Earth planet at various separations from the K-dwarf star using the suppressed weathering model. At the start of the evolution, we find that the hard boundary between habitable and snowball planets lies at ~ 0.78 AU, 0.03 AU further than the nominal model predicts. After 4 Gyr, abiotic Earths from 0.65–0.75 AU enter a snowball state. By contrast, planets closer to the inner edge remain habitable.

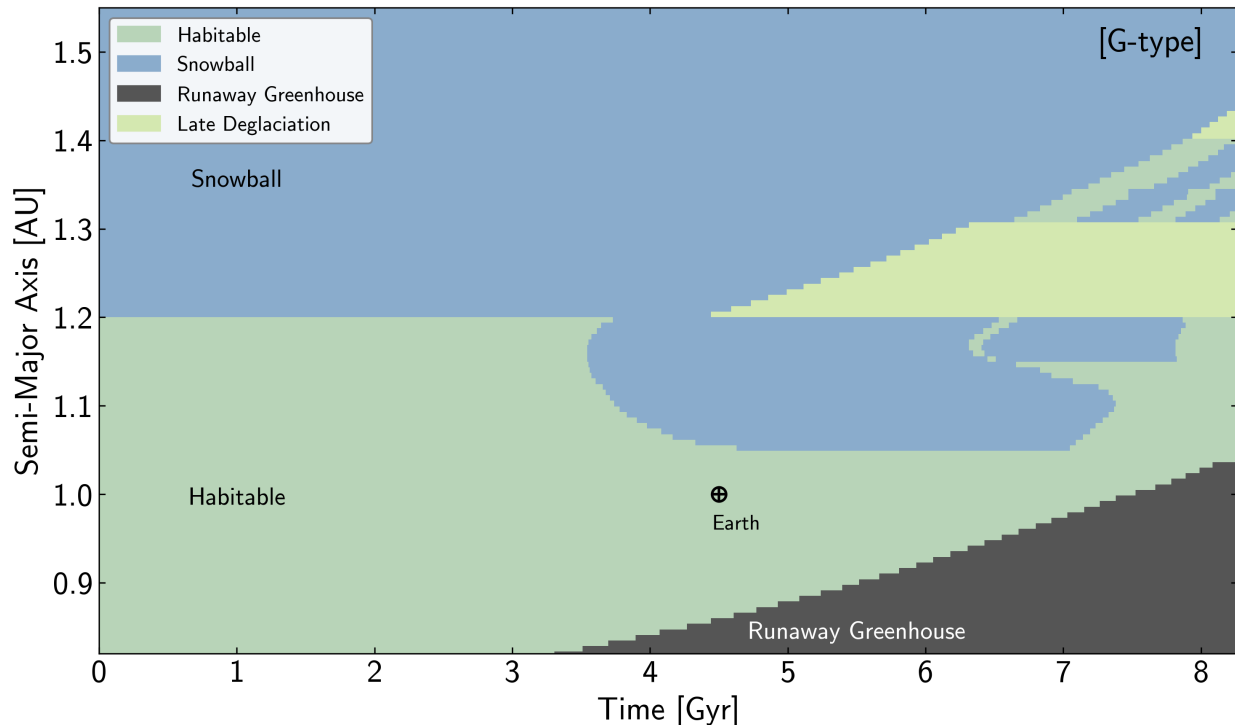


Figure 3.8: Climate phase diagram for an abiotic Earth orbiting a $1.0 M_{\odot}$ G-type star under a suppressed weathering rate. The initial habitable–snowball boundary lies at ~ 1.2 AU, 0.02 AU farther than the nominal model. The present-day Earth (marked with a cross) remains habitable for the majority of the Sun’s lifetime, entering a runaway greenhouse after ~ 7.5 Gyr. Planets beyond ~ 1.05 AU enter snowball states after ~ 3.5 Gyr before recovering habitable conditions, replacing the limit cycles predicted by the nominal model. The final outer HZ edge (~ 1.45 AU) is slightly narrower than in the nominal case (~ 1.5 AU). The star exits the main sequence at 8.1 Gyr.

We note that the HZ outer edge is non-linear after 4 Gyr, with planets < 0.65 AU and > 0.7 AU becoming snowballs later than planets lying between 0.65 and 0.7 AU. We also note that planets at orbital separations of ~ 0.7 AU appear to enter snowball states faster than the planets that are comparatively closer to and farther from the star. Similarly, planets at orbital separations of ~ 0.75 AU appear to remain habitable for billions of years longer than their closer-in counterparts. Lastly, we note that this model predicts that the runaway greenhouse region remains flat with time, consistent with the relatively flat evolution of the K-star’s luminosity (Figure 3.5).

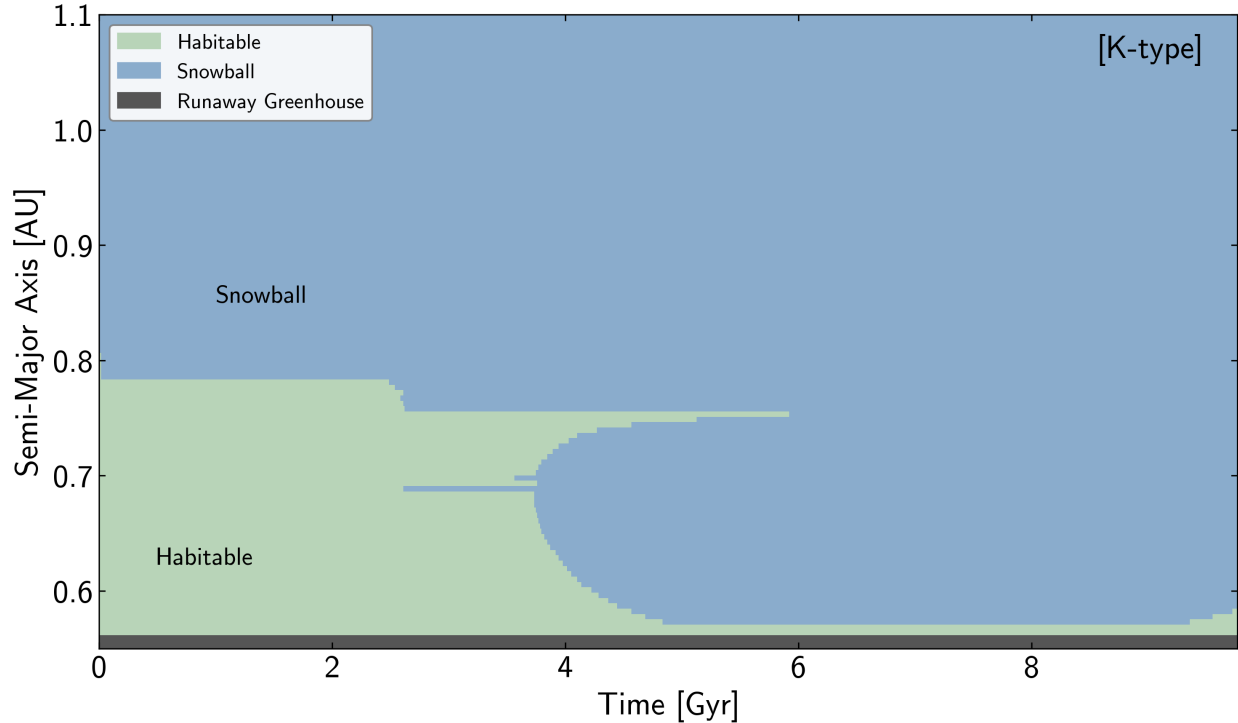


Figure 3.9: Climate phase diagram for an abiotic Earth orbiting a $0.8 M_{\odot}$ K-dwarf star under a suppressed weathering rate. The initial habitable–snowball boundary lies at ~ 0.78 AU, 0.03 AU farther than the nominal model. After ~ 4 Gyr, planets at 0.65–0.75 AU enter a snowball state from which none recover, in contrast to the G-dwarf suppressed weathering case. The outer HZ boundary is non-linear, mirroring the limit cycle structure of the nominal model. As in the nominal case, the runaway greenhouse boundary remains flat with time. The star exits the main sequence at 9.8 Gyr.

2.3 Model (b) (Nominal)

In Figures 3.10, 3.11, and 3.12, we show the resulting HZ phase diagrams for F, G, and K stars, respectively, using our nominal seafloor and continental weathering model, which assumes seafloor weathering ceases during global glaciation events. Across all phase diagrams, green indicates a habitable planet, blue indicates a snowball planet, white indicates a limit cycle, and black indicates a runaway greenhouse. In some plots, a light green region appears that indicates a late deglaciation to habitable conditions.

Figure 3.10 shows the climate regimes of an abiotic Earth planet at various separations from a $1.2M_{\odot}$ F-type star as it evolves on the main sequence. Our nominal model predicts

that the F star evolves off of the main sequence at 4.5 Gyr. At the start of the F-star's evolution, we find a hard boundary between habitable and snowball planets at ~ 1.75 AU. After 0.5 Gyr, abiotic Earths in the outer HZ (~ 1.58 – 1.75 AU) enter a limit cycle regime. We note that the Habitable-Limit Cycle edge is non-linear, as planets at ~ 1.58 and ~ 1.75 AU begin to limit cycle later than planets at ~ 1.65 AU. After 1.5 Gyr, the abiotic Earth at 1.2 AU enters a runaway greenhouse regime. As the star continues to brighten, planets at successive separations also enter a runaway greenhouse, establishing a roughly linear slope to the runaway greenhouse region. After 3.0 Gyr, the limit cycle region begins to move outward as a function of time.

Both the nominal model (b) and the suppressed weathering model (a) predict comparable inner and outer edge locations. However, while the nominal weathering model predicts that outer HZ planets (~ 1.58 – 1.75 AU) enter unstable limit cycle regimes after ~ 0.5 Gyr, the suppressed weathering model predicts that these planets remain stably habitable throughout the star's main sequence lifetime.

Figure 3.11 shows the climate regimes of an abiotic Earth planet at various separations from a $1.0M_{\odot}$ Sun-like G-type star. Our model predicts that the Sun will evolve off of the main sequence at about 8.1 Gyr. At the start of the Sun's evolution, we find a hard boundary between habitable and snowball planets at ~ 1.18 AU. By 1.0 Gyr, abiotic Earths in the outer HZ (~ 1.58 – 1.75 AU) enter a limit cycle regime. Similar to that of the F-star case, we note that the Habitable-Limit Cycle edge is non-linear. After 3.5 Gyr, the abiotic Earth at 0.8 AU enters a runaway greenhouse regime. As the star continues to brighten, planets at successive separations also enter a runaway greenhouse, establishing a roughly linear slope to the runaway greenhouse region. After 2.5 Gyr, the limit cycle region begins to move outward as a function of time. On this plot, we mark the location of the present-day Earth with a circled plus. We note that this model predicts that the present-day Earth would be in a limit cycle regime today, though it starts and concludes its evolution as a habitable planet. We also note that after ~ 1 Gyr a discontinuous snowball region occurs at the

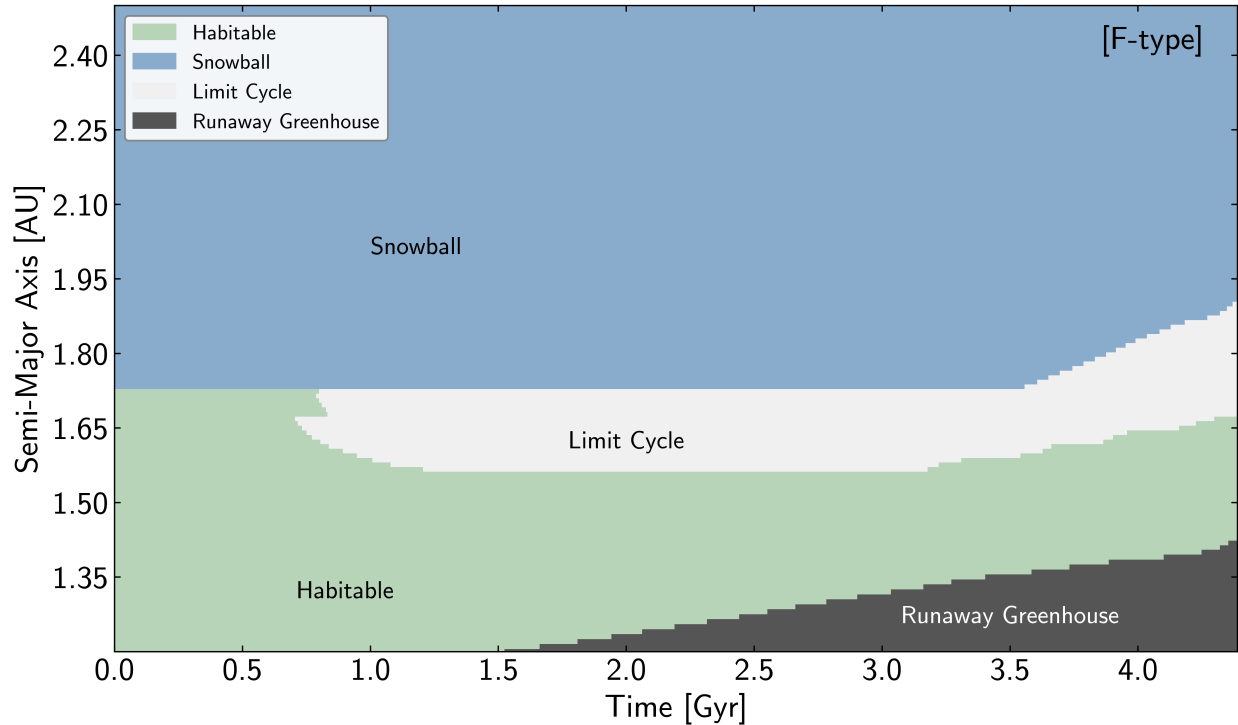


Figure 3.10: Climate phase diagram for an abiotic Earth orbiting a $1.2 M_{\odot}$ F-type star. The habitable–snowball boundary is initially located at ~ 1.75 AU, with planets in the outer HZ (~ 1.58 – 1.75 AU) entering a limit cycle regime after ~ 0.5 Gyr. Planets at progressively larger separations enter a runaway greenhouse state as the star brightens, establishing a linear inner boundary. The limit cycle region migrates outward after ~ 3.0 Gyr. The star exits the main sequence at 4.5 Gyr.

boundary of the habitable and limit cycle regions for planets at orbital separations of ~ 1.1 AU. Finally, after 7.8 Gyr, the planets on the outer edge of the limit cycle region experience a late deglaciation, briefly becoming habitable before the Sun exits the main sequence.

The non-linear boundaries of the limit cycle regions predicted by the nominal model resemble the non-linear Snowball-Habitable boundaries of the late-snowball planets predicted by model (b), but the nominal model predicts that outer HZ planets enter limit cycles by ~ 1.0 Gyr, while the suppressed weathering model instead drives planets beyond ~ 1.05 AU into snowball states after ~ 3.5 Gyr, from which they eventually recover habitable conditions via late deglaciation. Lastly, though the nominal model predicts that the outer edge of the HZ migrates to 1.5 AU by the conclusion of the Sun’s evolution, the suppressed weathering model predicts that the final outer edge lies closer to 1.45 AU.

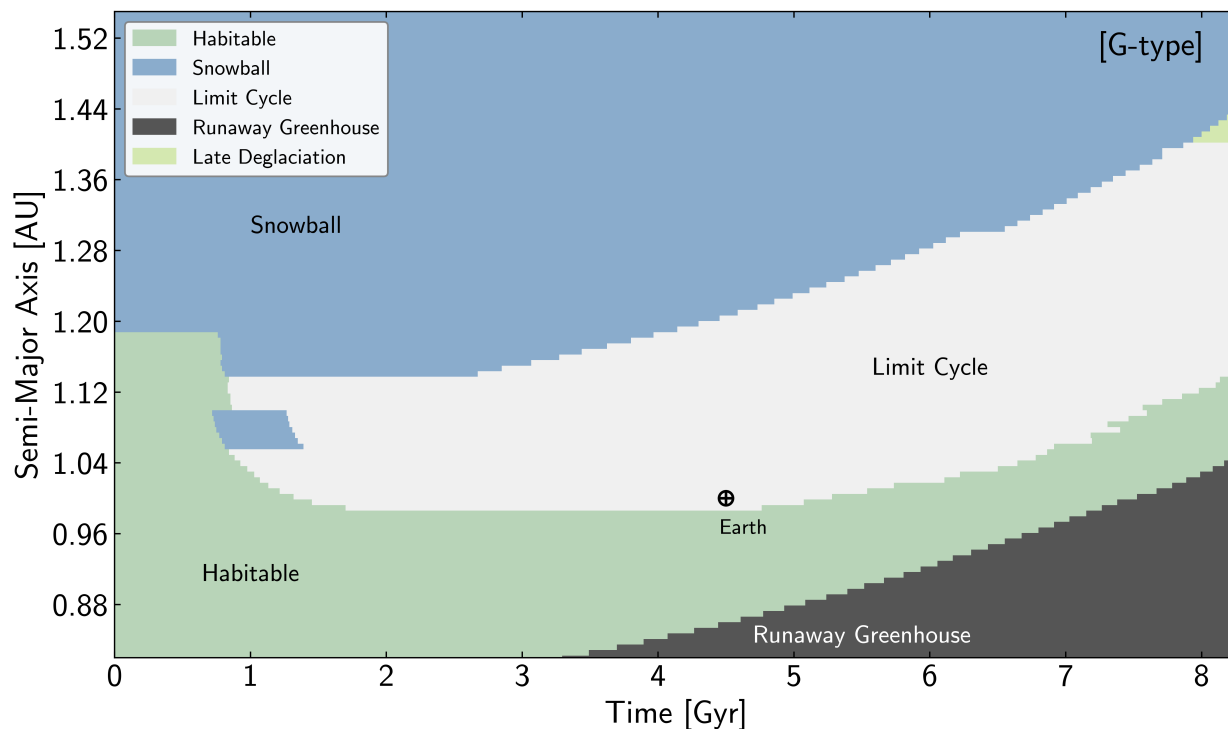


Figure 3.11: Climate phase diagram for an abiotic Earth orbiting a $1.0 M_{\odot}$ G-type star. The habitable–snowball boundary is initially located at ~ 1.18 AU, with outer HZ planets entering a limit cycle regime by ~ 1.0 Gyr. The habitable–limit cycle boundary is non-linear, mirroring the F-star case. Planets at progressively larger separations enter a runaway greenhouse state beginning at ~ 3.5 Gyr, and the limit cycle region migrates outward after ~ 2.5 Gyr. The present-day Earth (marked with a circled plus) falls within the predicted limit cycle regime, though the planet begins and ends its evolution in a habitable state. After ~ 7.8 Gyr, planets at the outer limit cycle boundary undergo a late deglaciation, briefly becoming habitable before the star exits the main sequence at 8.1 Gyr.

Figure 3.12 shows the climate regimes of an abiotic Earth planet at various separations from a $0.8 M_{\odot}$ K-dwarf star. Our model predicts that the K-dwarf evolves off of the main sequence at 9.8 Gyr. At the start of the evolution, we find that the hard boundary between habitable and snowball planets lies at ~ 0.75 AU. By 1.0 Gyr, abiotic Earths in the outer HZ (~ 0.65 – 0.75 AU) enter a snowball state prior to the limit cycle regime. By contrast, planets closer to the inner edge directly enter a limit cycle regime. Similar to that of the F and G-dwarf cases, we note that the Habitable-Limit Cycle edge is non-linear. After 2.0 Gyr, the limit cycle region begins to move outward as a function of time. Finally, after 7.5 Gyr, the planets closer to the inner edge (~ 0.56 – 0.58 AU) briefly become habitable before

the star exits the main sequence. Unlike the F and G-stars, the runaway greenhouse region does not move outward as a function of time, but remains flat. Lastly, we note several gaps in the continuity of the limit cycle region containing significant “snowball” patches, mostly affecting planets at orbital separations of 0.60–0.78.

The non-linearity of the limit cycle region predicted by the nominal model is similar to the non-linearity of the HZ outer edge predicted by the suppressed weathering model. While both models predict that outer HZ planets (~ 0.65 – 0.75 AU) enter snowball or limit cycle states, the nominal model produces persistent limit cycles with intermittent snowball patches, whereas the suppressed weathering model drives these planets into permanent snowball states after ~ 4 Gyr with no subsequent deglaciation. Lastly, we note that similar to the nominal model, this model predicts that the runaway greenhouse region remains flat with time, consistent with the relatively flat evolution of the K-star’s luminosity (Figure 3.5).

2.4 Model (a) and (b) Comparison

Across all three spectral types, the nominal and suppressed weathering models differ substantially in their predictions of climate stability for outer HZ planets. The nominal model produces limit cycle regimes in the outer HZ for all three stellar types, driven by the interplay between ice-albedo feedback and carbonate-silicate weathering. The suppressed weathering model eliminates limit cycles entirely, instead producing either stably habitable planets (F-type) or permanent snowball states from which planets do not recover (K-dwarf). The G-dwarf case represents an intermediate behavior: the suppressed weathering model drives outer HZ planets into snowball states analogous to the limit cycle region of the nominal model, but these planets eventually recover habitable conditions via late deglaciation rather than remaining permanently frozen. In both models, the runaway greenhouse inner boundary evolves similarly for F and G stars, migrating outward as the star brightens, while the K-dwarf runaway greenhouse boundary remains flat with time regardless of weathering prescription owing to its slow main sequence evolution. The suppressed weathering model

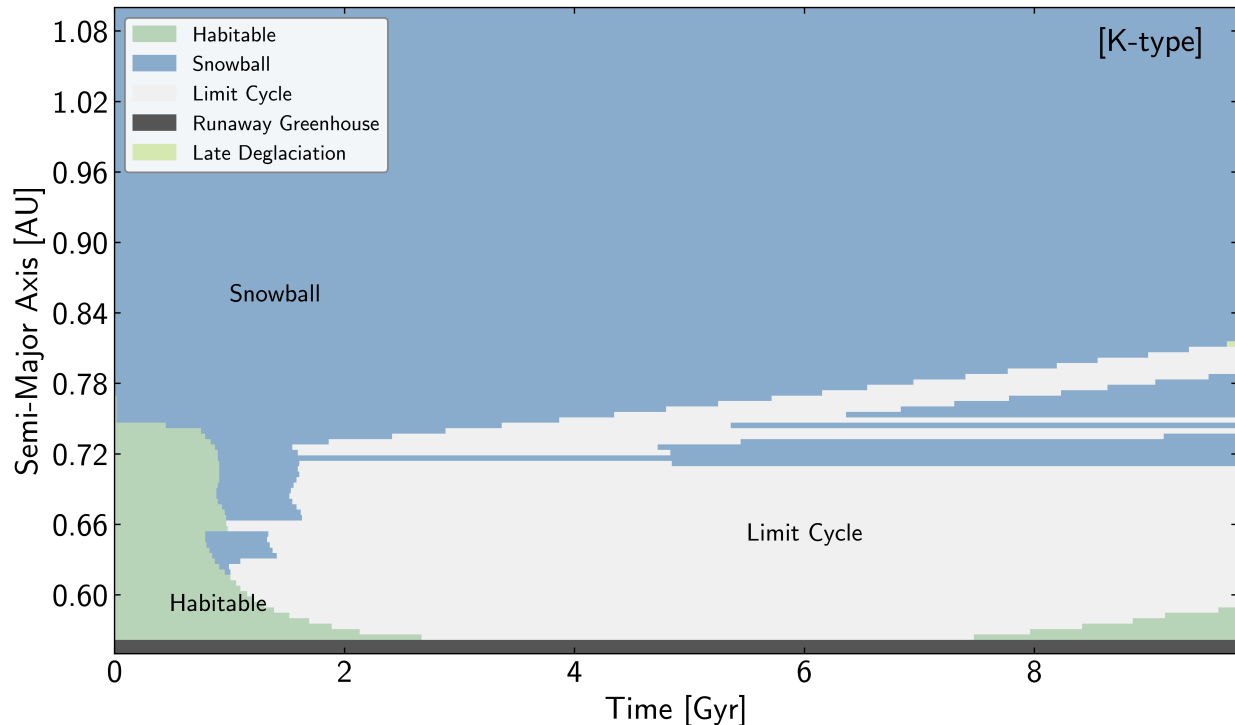


Figure 3.12: Climate phase diagram for an abiotic Earth orbiting a $0.8 M_{\odot}$ K-dwarf star. The habitable–snowball boundary is initially located at ~ 0.75 AU. By 1.0 Gyr, outer HZ planets (~ 0.65 – 0.75 AU) transition through a snowball state before entering a limit cycle regime, while planets closer to the inner edge enter the limit cycle directly. The habitable–limit cycle boundary is non-linear, consistent with the F and G-dwarf cases. The limit cycle region migrates outward after ~ 2.0 Gyr. After ~ 7.5 Gyr, planets near the inner edge (~ 0.56 – 0.58 AU) briefly become habitable before the star exits the main sequence at 9.8 Gyr. Unlike the F and G-dwarf cases, the runaway greenhouse boundary remains at a fixed separation throughout the evolution. Notable snowball patches are present within the limit cycle region, reflecting intermittent glaciation episodes.

predicts outer HZ boundaries that are ~ 0.02 – 0.03 AU farther from the star compared to the nominal model, and slightly narrower final HZ widths at the end of each star’s main sequence lifetime. We summarize the initial and final orbital ranges for the HZs of each stellar type using models (a) and (b) in [Table 3.2](#).

In summary, our parameter sweeps reveal that the width and stability of the habitable zone depends sensitively on both stellar type and the assumed strength of the carbonate–silicate weathering feedback. Across all three stellar types, the nominal model predicts that outer HZ planets enter limit cycle regimes driven by the interplay between ice–albedo feed-

Table 3.2: Habitable zone orbital ranges by stellar type for models (a) and (b). In model (b), limit cycling occurs at the outer portions of the listed ranges; for G-type stars, late deglaciation additionally occurs at 1.4–1.45 AU.

Stellar Type	Initial HZ (AU)	Final HZ (AU)	
		Model (a) No Limit Cycles	Model (b) Limit Cycles
F-type	1.2–1.7	1.4–1.9	1.4–1.6 (limit cycling: 1.6–1.9)
G-type	0.8–1.2	1.05–1.45	1.05–1.1 (limit cycling: 1.1–1.45)
K-type	0.54–0.78	0.54–0.58	0.54–0.58 (limit cycling: 0.58–0.7)

back and $p\text{CO}_2$ -dependent weathering, while the suppressed weathering model predicts either stable habitability or permanent snowball states. The inner HZ boundary migrates outward over time for F and G stars as the star brightens, while the K-dwarf runaway greenhouse boundary remains fixed. In both models, the outer HZ boundary is non-linear and time-varying, reflecting the coupled evolution of stellar luminosity and interior carbon outgassing. The K-dwarf case is notably distinct from the F and G-dwarf cases, exhibiting a flat runaway greenhouse boundary and a greater tendency toward permanent glaciation throughout the HZ. We discuss the physical interpretation of these results and their implications for habitability around FGK stars in the following section.

3 Discussion

Our results demonstrate that the width, stability, and temporal evolution of the HZ around FGK stars depends sensitively on both stellar type and the strength of the carbonate-silicate weathering feedback. In both models, the outer HZ boundary is time-varying and non-linear, contracting over the star’s main sequence lifetime as volcanic degassing declines. The nominal model predicts that outer HZ planets enter limit cycle regimes driven by the interplay between ice-albedo feedback and $p\text{CO}_2$ -dependent weathering, while the suppressed weathering model eliminates limit cycling entirely, replacing it with either stable habitability or permanent snowball states depending on stellar type.

The K-dwarf case is notably distinct: unlike F and G stars, whose inner HZ boundaries

migrate outward as the star brightens, the K-dwarf runaway greenhouse boundary remains fixed throughout the evolution, and outer HZ planets show a pronounced tendency toward permanent glaciation in both models. For K-dwarfs, stellar luminosity increases only 10% over 10 Gyr, so the star cannot compensate for declining degassing by providing more insolation. Degassing shuts off early as the interior cools, removing the primary CO_2 source. Without sufficient atmospheric CO_2 , seafloor weathering out-competes arc volcanic outgassing and drives the planet into a permanent snowball. By contrast, for F and G stars, the star brightens rapidly enough that the planet eventually receives sufficient insolation to sustain habitability even as degassing declines. This asymmetry explains why our results suggest that the outer HZ contracts dramatically for K-dwarfs but not for F and G stars.

Across all stellar types, the dominant uncertainty in our predictions traces to the treatment of seafloor weathering during snowball events, which determines whether CO_2 can accumulate sufficiently to drive deglaciation. We discuss the physical origins of these results, their dependence on model assumptions, and briefly cover their implications for the prioritization of habitable zone targets around FGK stars.

3.1 Comparison to Previous Work

[Kopparapu et al. \(2013\)](#), [Haqq-Misra et al. \(2016\)](#), and [Kadoya & Tajika \(2019\)](#) each explored the bounds of the HZ based on stellar type, enabling direct comparison between our results and prior works. Here we compare our results, and discuss relevant differences in model assumptions. Ultimately, [Kadoya & Tajika \(2019\)](#) provides the best opportunity for an apples-to-apples comparison, given their approach most closely mirrors ours in treating the HZ as a time-evolving quantity governed by coupled stellar and planetary evolution.

[Kopparapu et al. \(2013\)](#) found static HZ limits for snapshots of stars in time, with the solar HZ lying between 0.99 and 1.7 AU, that of a $1.2M_\odot$ F-star between 1.5–2.5 AU, and that of a $0.8M_\odot$ K-star between 0.5–1.0 AU. [Haqq-Misra et al. \(2016\)](#) found that the static solar HZ lies between 0.95–1.58 AU. Within these bounds, the authors found a limit cycling

region that exists between 0.97–1.12 AU. [Kadoya & Tajika \(2019\)](#) explicitly addressed the boundaries of the inner and outer HZ for an evolving star. For a slightly more massive F-star than considered here ($1.25M_{\odot}$), the authors found that the HZ evolves from 1.4–2.25 AU to 1.8–1.9 AU over the star’s 5 Gyr main sequence lifetime, with significant limit cycling from 1.9–2.5 AU at the end of the evolution. For a Sun-like G-star, they found an initial HZ range from 0.85–1.45 AU, and after 10 Gyr only a limit cycling region from 1.3–1.7 AU. Finally, for a slightly less massive K-star ($0.75M_{\odot}$), they found that the initial HZ lies between 0.5–0.8 AU. After 10 Gyr on the main sequence, they found only a limit cycle region from 0.55–0.8 AU.

For Sun-like, G-type stars, we find that the inner edge of the HZ is closer to the star than all three previous studies. The final ranges are broadly similar to that of [Kadoya & Tajika \(2019\)](#), though their final limit cycling region extends to farther orbital distances. [Haqq-Misra et al. \(2016\)](#) predicts a limit cycling region that lies closer to the star than in this work. For F-type stars, we find HZ ranges that are consistently narrower and closer to the star than both [Kopparapu et al. \(2013\)](#) and [Kadoya & Tajika \(2019\)](#). In particular, the HZ of [Kadoya & Tajika \(2019\)](#) extends much further out (to 2.5 AU), roughly 0.6 AU beyond this study’s outer limit, though this can be partially attributed to their more massive F-star being more luminous. Finally, for K-type stars, the initial HZ ranges are broadly similar across all studies. The biggest divergence is in the location of the HZ limits at the conclusion of the star’s main sequence lifetime: this study’s habitable region (model (b)) collapses dramatically to a narrow band near 0.54–0.7 AU, while the final limit cycling region of [Kadoya & Tajika \(2019\)](#) extends to 0.8 AU and [Kopparapu et al. \(2013\)](#)’s static estimate reaches 1.0 AU.

Both [Kopparapu et al. \(2013\)](#) and [Haqq-Misra et al. \(2008\)](#) lack the time-evolving properties that drive the most distinctive features of our results. While both consider HZ boundaries across a range of stellar types, neither evolves stellar luminosity over the main sequence lifetime – thus, their boundaries are effectively snapshots in time for fixed stellar properties.

Kopparapu et al. (2013) establishes a purely radiative outer boundary with no carbon cycle at all, while Haqq-Misra et al. (2016) introduces weathering feedbacks and limit cycling but treats outgassing as a fixed parameter. Neither can capture the secular decline in volcanic degassing that, in our model, progressively contracts the outer HZ boundary over the star’s main sequence lifetime, particularly for K-dwarfs. Our results are therefore broadly consistent with both studies at early times, but diverge increasingly as the interior evolves, with the final HZ in our model being substantially narrower than either prior estimate.

Of the studies considered here, the comparison with Kadoya & Tajika (2019) is most directly informative. The primary distinctions are that we drive volatile fluxes with an explicit interior model rather than parameterized degassing histories, and that we allow for continental growth over time, both of which introduce additional degrees of freedom in the carbon cycle evolution. Though both our model and that of Kadoya & Tajika (2019) (for full description of convection parameterization, see Kadoya & Tajika, 2015) produce comparable average mantle temperatures, mantle heat fluxes, and spreading rates for the present-day Earth, the *evolution* of these values likely differs. Furthermore, following Haqq-Misra et al. (2016), Kadoya & Tajika (2019) account for water ice’s lower albedo in the near-IR, likely allowing for more generous outer HZ edges than in this work, where gray albedos are assumed for all surface types. Similarly, our different cloud parameterizations result in our finding shallower inner edges across all stellar types, as well as stable habitable planets for G- and K-type stars at the end of their main sequence lifetimes. Finally, we have shorter main sequence lifetimes than Kadoya & Tajika (2019) for the F and solar G-type stars, owing to our use of different stellar evolution models.

A further distinction concerns the weathering reference state and the absolute degassing flux. Kadoya & Tajika (2019) implicitly assume a biotic planet, in which soil $p\text{CO}_2$ is enhanced by a factor of 10–100 relative to the atmosphere by microbial and plant respiration, strengthening the weathering feedback relative to the abiotic formulation (Haqq-Misra et al., 2016). All else being equal, an abiotic weathering formulation would tend to make our

planets less prone to snowball states and limit cycling relative to [Kadoya & Tajika \(2019\)](#). However, our interior-derived CO_2 degassing rate ($\sim 1 \text{ Tmol yr}^{-1}$) is substantially lower than the present-day Earth value assumed by both [Haqq-Misra et al. \(2016\)](#) and [Kadoya & Tajika \(2019\)](#) ($\sim 7.5\text{--}8 \text{ Tmol yr}^{-1}$). Furthermore, unlike previous approaches, our model explicitly partitions carbon between the atmosphere and ocean via Henry’s law dissolution, introducing an additional carbon sink that suppresses atmospheric $p\text{CO}_2$ relative to models that treat the surface carbon reservoir as purely atmospheric. Both effects push atmospheric $p\text{CO}_2$ lower than in previous work for a given degassing rate. Combined, lower degassing rates and lower atmospheric $p\text{CO}_2$ tends to push planets toward the snowball and limit cycle boundary ([Kadoya & Tajika, 2019](#)). The net result is HZ ranges that are narrower than [Kadoya & Tajika \(2019\)](#) despite our less aggressive weathering parameterization.

Notably, the prevalence of limit cycling in our results depends sensitively on the assumed behavior of seafloor weathering during snowball states: when seafloor weathering continues during glaciation, CO_2 is unable to accumulate sufficiently to trigger deglaciation, and limit cycles are replaced by permanent snowball states. This is consistent with the broader point made by [Haqq-Misra et al. \(2016\)](#) and [Abbot \(2016\)](#), who find that limit cycling ultimately reflects the balance between outgassing and weathering fluxes rather than either in isolation. In our model, where outgassing is determined by the interior evolution, this balance hinges entirely on the assumed weathering behavior during glaciated states – whether seafloor weathering continues to draw down CO_2 during a snowball episode or is suppressed. Apparently it is this single assumption, more than any other, that determines whether our model predicts either limit cycling or long-term glaciations.

More broadly, the fundamental uncertainties in any HZ model of this kind reduce to uncertainties in the time-evolving balance of carbon cycle fluxes and long-term stellar evolution. Parameterized approaches like that of [Haqq-Misra et al. \(2016\)](#) make these uncertainties explicit by directly varying the degassing flux histories. Our interior model enforces self-consistency between the fluxes – degassing, continental growth, and weathering are not

varied independently but emerge from a common physical framework – yet this comes at the cost of propagating uncertainty through additional processes: mantle thermal evolution, melt generation, continental crust production, and their coupled effects on the carbon cycle. In a sense, we are marginalizing over the same underlying flux uncertainty, but distributing it across a richer model structure. Whether this approach is an advantage depends on how well the interior model captures the true diversity of planetary evolutionary pathways. This open question in exoplanet climate evolution studies motivates treating our two weathering scenarios not as bracketing cases, but as illustrative end-members of a broader continuum of possible flux evolutions.

3.2 Implications for the Abiotic vs. Biotic HZ

Kadoya & Tajika (2019) argue, following Haqq-Misra et al. (2016), that the CO₂ degassing rate at the boundary between warm and limit cycling states is lower for an abiotic planet than for a biotic one, leading them to conclude that land plants may increase the susceptibility of a planet to snowball limit cycles, and conversely, that abiotic planets may enjoy longer durations of stable warm climates. However, this conclusion rests on the assumption that abiotic weathering is substantially weaker than biotic weathering by a well-defined factor. Field estimates of the biotic enhancement to silicate weathering span from essentially 1 (no enhancement) to ~100 depending on organism, mineral substrate, and environmental context, and no true abiotic control site exists on Earth today against which to calibrate these estimates (Wild et al., 2022). If the lower end of this range is representative, the distinction between biotic and abiotic weathering efficiency largely collapses. Furthermore, in supply-limited weathering regimes (Foley, 2015), where fresh rock availability rather than atmospheric CO₂ or temperature governs the weathering rate, the biological enhancement of soil $p\text{CO}_2$ becomes largely irrelevant regardless of its magnitude. Our results suggest that abiotic planets are not necessarily less susceptible to snowball states. Furthermore, our results, which treat weathering strength somewhat as a free parameter, suggest that the

susceptibility of a planet to long-term glaciation depends more sensitively on the behavior of seafloor weathering during snowball episodes and the absolute degassing rate than on whether the planet hosts land life.

3.3 Future Work

Across nearly all of our phase diagrams, we have noted puzzling features that require further investigation. These features include transient cold excursions and late-onset deglaciations (e.g., [Figure 3.8](#)), which appear in our HZ phase diagrams as “glitches” in otherwise smooth model behavior. At this time, we have not ruled out the possibility that these features appear due to modeling artifacts including numerical instabilities in the model, which may cause un-physical oscillations in $p\text{CO}_2$. Such instabilities are resolved by reducing the size of the model time step in the vicinity of the oscillation. Alternatively, one possible physical cause of discontinuous model behavior is climate bistability, or the co-existence of a warm, habitable solution and a frozen, snowball solution which both satisfy planetary thermal equilibrium (e.g., [Boschi et al., 2013](#); [Murante et al., 2020](#)). Future work should carefully examine individual climate evolutions and assess the balance of fluxes in the epochs of concern to determine whether the underlying model behavior is physical or not, and whether bistable climatic solutions can explain the observed “glitches.”

We also identify several natural extensions of this work. First, the weathering model could be constrained by vetting it against early Earth climate evolution, using the range of plausible evolutions consistent with the geological record to bound the interior and carbon cycle uncertainties explored here. Second, the relative contributions of visible and near-IR albedo differences for ice-covered surfaces – which are not accounted for in our gray albedo treatment – could be incorporated, as water ice absorbs more in the near-IR than the visible, reducing the ice-albedo contrast for cooler stars and potentially affecting the limit cycle boundaries for K-dwarfs in particular. Third, the fate of CO_2 ice at the outer HZ boundaries warrants explicit treatment; our model currently excludes condensed CO_2 from the planetary

greenhouse but does not track whether it remains available for weathering or degassing, which could affect the recovery from glaciated states. Finally, numerical instabilities in our model could likely be smoothed by adopting an implicit solver in the integrator.

More generally, parameterized approaches similar to [Haqq-Misra et al. \(2016\)](#) address fundamental carbon cycle uncertainties by directly varying carbon cycle fluxes. Our interior model offers a complementary and potentially more physically grounded avenue: rather than varying the fluxes themselves, one could vary the underlying planetary parameters that govern them – such as the initial abundance of radiogenic elements in the core and mantle, initial mantle temperature, or core mass fraction – and propagate these through the interior model to produce self-consistent flux histories. This would allow the degassing and weathering evolution to co-vary in physically motivated ways rather than being varied independently, potentially providing tighter and more interpretable constraints on the range of plausible HZ evolutions for Earth-like exoplanets. Conversely, an insensitivity to interior evolution parameters would suggest that evolution uncertainties can be sufficiently captured by marginalizing over carbon cycle fluxes, or by using convective parameterizations of the interior as in [Kadoya & Tajika \(2019\)](#).

4 Conclusion

We have presented an improved version of our whole-planet abiotic Earth model, including a treatment for fractional ice coverage, and assessed the sensitivity of our model to assumptions about seafloor and continental weathering. We then applied two versions of the model – (1) with suppressed weathering that continues during snowball events, and (2) with nominal weathering that ceases during snowballs – to exploring the HZ boundaries of selected F-, G-, and K-type stars.

In agreement with previous results ([Haqq-Misra et al., 2016](#); [Abbot, 2016](#)), we find that the occurrence of limit cycles strongly depends on our assumptions about the evolving car-

bon cycle. In particular, we find that limit cycles occur only if we assume that seafloor weathering cannot continue during global glaciations. While previous work has explored the dependence on the assumed CO_2 degassing rate, we have shown that these climate instabilities are also susceptible to assumptions about planetary weathering. We recommend re-tuning the weathering model by identifying those which reproduce early Earth history within measurement uncertainties. Additionally, to further explore the range of plausible evolutions of Earth-like planets around FGK stars, we recommend a parameter sweep over key properties of the interior (e.g., budget of radionuclides, energy from early impact events). This will build on previous work that directly varied the CO_2 degassing rate by allowing the degassing rate to scale self-consistently with realistic interior dynamics.

Our evolving HZ boundaries for a $1.2M_\odot$ F-type star, solar-type G-dwarf, and $0.8M_\odot$ star are broadly consistent with those found by [Kadoya & Tajika \(2019\)](#). However, we find slightly more generous inner edges across all stellar types; less generous outer edges; and slightly broader HZs for G- and K-dwarfs at the conclusion of their main sequence lifetimes. For the F- and K-type stars, some of these differences may be attributed to differences in stellar mass, but we also identified key differences in our wavelength-independent treatment of surface albedo and our cloud parameterizations. Future work can further explore the sensitivity of model results to these parameterization choices.

These results have direct implications for target prioritization with HWO. The strong dependence of outer HZ climate stability on poorly constrained geophysical processes – particularly the behavior of seafloor weathering during glaciated states – introduces a fundamental uncertainty in predicting which planets in the outer HZ are likely to be habitable at any given point in their evolution. This suggests that planets near the inner HZ edge, where habitability is governed primarily by atmospheric water loss or the runaway greenhouse limit rather than the carbon cycle, may represent higher priority targets for habitability and biosignature searches. Among stellar types, K-dwarfs present a nuanced case: while their slow luminosity evolution and long main sequence lifetimes are observationally advantageous, our results

suggest that outer HZ K-dwarf planets are particularly susceptible to permanent glaciation. Assuming Earth-like interior heating and degassing histories, planets near the inner K-dwarf HZ edge may therefore offer the best combination of long-term habitability and observational accessibility. Similarly, our results suggest that demonstrably habitable planets found farther from K stars may indicate young systems (<1 Gyr), but it is unclear if this is sufficient time for life to emerge. Alternatively, K-stars may require larger radiogenic inventories or tidal heating to maintain higher rates of internal heating to maintain higher degassing rates and sustain longer-term habitability. More broadly, the abiotic baseline established here provides a physically motivated null hypothesis for interpreting reflected light spectra of HWO targets: atmospheric and surface states consistent with our model predictions need not invoke life.

CHAPTER 4

RETRIEVED ATMOSPHERES AND INFERRED SURFACE PROPERTIES FOR TERRESTRIAL EXOPLANETS USING TRANSMISSION AND REFLECTED LIGHT SPECTROSCOPY

A version of this chapter was published as: Gilbert-Janizek, S., Meadows, V. S., & Lustig-Yaeger, J. 2024, “Retrieved Atmospheres and Inferred Surface Properties for Terrestrial Exoplanets Using Transmission and Reflected-light Spectroscopy,” The Planetary Science Journal, 5, 148, doi: [10.3847/PSJ/ad381e](https://doi.org/10.3847/PSJ/ad381e). It is reproduced here under the terms of the Creative Commons Attribution 4.0 license. V. S. Meadows supervised the research; J. Lustig-Yaeger provided technical mentorship on the use of his retrieval model; both co-authors provided comments and edits on the manuscript; all other work is my own.

1	Methods	133
1.1	Simulated Exoplanet Spectra	134

1.2	Retrieval Model	137
1.3	Model Comparison	145
2	Experiments	146
2.1	Transmission	148
2.2	Direct Imaging	151
3	Results	153
3.1	Transmission	154
3.2	Direct Imaging	160
4	Discussion	166
4.1	Comparison to Previous Work	170
4.2	Comparison of Transmission and Direct Imaging Capabilities	172
4.3	Model Limitations and Future Work	190
5	Conclusion	195

Here we present a study to understand the relative effectiveness of transmission observations and direct imaging for characterizing the near-surface environment of exoplanets to assess habitability and biosignatures. We investigate performing spectroscopic retrievals in the presence of clouds for both transmission and direct imaging, as well as a false positive biosignature assessment for transmission observations. Despite the compelling scientific prospects of emission spectroscopy, habitable zone planets may be too cool to achieve comparably high signal-to-noise ratios in thermally emitted light. Therefore, in this study, we exclusively focus on transmission spectroscopy for transiting planets orbiting M-dwarf stars due to its advantageously high signal-to-noise ratio. For both transmission and direct imaging, we assess our ability to probe the photochemical productivity of the atmosphere via the vertical structure of O_3 . In the following sections, we describe our methods for comparing the capabilities of transmission and direct imaging spectroscopy (Section 1) and the self-consistent planet/atmospheric cases we will explore (Section 2). We provide a summary of our results in Section 3, and discuss the impacts they have for assessing habitability and biosignatures on terrestrial, habitable zone planets in Section 4. Finally, we provide a summary of these takeaways and identify areas for future exploration in Section 5.

1 Methods

We use a retrieval model with and without vertically-resolved molecular abundance profiles to infer the composition and physical properties of the atmospheres of terrestrial exoplanets from simulated observations, and thereby characterize the surface environment and habitability of these worlds. Part of assessing the composition of the atmosphere is understanding the vertical resolution of some species that have distinctive structure with altitude, such as photochemically-generated or destroyed molecules (e.g. O_3) and water vapor, whose abundance is strongly tied to the vertical temperature structure. By probing the vertical structure we can better understand these important gases in the full context of the planetary envi-

ronment. Our novel retrieval model solves for the distribution of atmospheric states that best fits the simulated spectrum, assuming either evenly-mixed or vertically-resolved molecular profiles. This state-of-the-art terrestrial retrieval model generates spectra and fits them to the simulated data to infer atmospheric parameters and likely abundance profiles. We are capable of running the model for vertically-resolved molecular profiles, where the algorithm is either constrained to pre-selected pressure points or allowed to freely fit for the pressure points. For each observatory, we determine whether the vertically-resolved models appreciably improved the fit compared to the simpler evenly-mixed case. Finally, to assess the interpretability of the results, we compare the retrieved abundance profiles to the true profiles used to generate the simulated data. The models and our inputs are detailed below.

1.1 Simulated Exoplanet Spectra

To produce synthetic transmission and reflectance observations of exoplanets, we use a coupled 1-D climate-photochemistry model to produce self-consistent atmospheres, a line-by-line radiative transfer model to generate high-resolution spectra, and an instrument simulator to degrade the resolution of the spectra to that of the relevant instruments and add astrophysical, telescope, and instrumental noise. These synthetic spectra therefore incorporate the full atmospheric complexity rendered by our models, including a non-isothermal temperature structure, heterogeneously mixed gas abundances, and pressure-broadened absorption lines. As a result, our subsequent retrieval experiments using these simulated observations reflect the ability of the model to accurately infer realistic atmospheric states subject to the necessary simplifying retrieval model assumptions.

1.1.1 Simulated Atmospheres & Synthetic Spectra

To generate thermally and photochemically self-consistent atmospheres, we couple the VPL Climate model to *atmos*, a photochemistry model. VPL Climate (Lincowski et al., 2018) is a 1-D radiative-convective equilibrium model that uses the Spectral Mapping and Radiative

Transfer (SMART) code to internally calculate radiative fluxes and heating rates. SMART is a line-by-line, multi-stream, multi-scattering radiative transfer code, which includes layer-dependent gaseous and aerosol absorption, emission, and scattering (Meadows & Crisp, 1996; Stamnes et al., 1988, originally developed by D. Crisp). Finally, *atmos* is a 1-D atmospheric chemistry model that has been used to model the photochemistry of various terrestrial exoplanets in a range of studies (Segura et al., 2005; Arney et al., 2017; Meadows et al., 2018b). Here, we use an updated version of the publicly-available *atmos* model as described in Lincowski et al. (2018).

Once we have generated self-consistent atmospheres, we use the Line By Line ABSorption Coefficients (LBLABC) model and SMART to generate high-resolution spectra. LBLABC calculates rotational-vibrational line absorption coefficients for the gases in the atmosphere under consideration (Meadows & Crisp, 1996) using the HITEMP2010 and HITRAN2012 line lists (Rothman et al., 2010, 2013). LBLABC assumes air ($\text{N}_2 + \text{O}_2 + \text{Ar}$) line broadening with a Lorentzian wing cut-off of 1000 cm^{-1} for CO_2 and H_2O , as they are known to have broad wings for Earth and Venus. For all other gases, we use a wing cut-off of 50 cm^{-1} . For a complete discussion of how LBLABC calculates the line absorption coefficients, please see Lustig-Yaeger et al. (2023b). Line absorption coefficients are calculated by LBLABC for each molecular species at high spectral resolution to fully resolve individual lines and only convolved to the lower resolution of the data after the radiative transfer calculation has been computed via SMART. SMART uses the absorption coefficients output by LBLABC to solve the radiative transfer equations and produce a high-resolution spectrum. SMART is capable of generating atmospheric spectra in transmitted, emitted, and reflected light, and has been rigorously validated on observations of the Earth (Robinson et al., 2011) and Venus (Meadows & Crisp, 1996; Arney et al., 2014). SMART has also been used to model simulated spectra for theoretical habitable M dwarf planets (Segura et al., 2003, 2005; Meadows et al., 2018b) and uninhabitable planets (Lincowski et al., 2018, 2019). SMART calculates cloud opacities independently from rotational-vibrational line coefficient opacities.

SMART is also capable of computing layer-dependent flux Jacobians and top-of-atmosphere radiance Jacobians. Radiance Jacobians are matrices consisting of the partial derivatives of the top-of-atmosphere radiance field with respect to each user-specified property of the atmosphere at each atmospheric layer. These matrices thus relate the spectrum to small perturbations to the atmospheric state, allowing us to quickly determine which vertical sections of the atmosphere the spectrum is sensitive to. SMART can compute radiance Jacobians with respect to temperature, gas abundances, and surface albedo. While SMART does not currently support the calculation of transit depth Jacobians, this is being considered for future work.

1.1.2 Instrument Simulator

To simulate the *Origins* and LUVOIR instrument data, we use the open-source Python `coronagraph` model to degrade the high resolution synthetic spectra and add noise (Robin et al., 2016; Lustig-Yaeger et al., 2019c). `coronagraph` was originally developed to simulate data from telescopes equipped with a coronagraph and has since been modified to generate simulated data (signal and noise) for a range of exoplanet observing methods using space-based (Bolcar et al., 2016; Mennesson et al., 2016) and ground-based telescope architectures (Meadows et al., 2018a). The user provides a high-resolution exoplanet spectrum and the model calculates the data points by degrading the spectrum to the instrument resolution using a boxcar convolution. Though a triangle convolution function is a more realistic simulator of a slit spectrometer, we did not find significant differences when implementing the boxcar filter. `coronagraph` is then used to calculate the observed photon count rates for a range of observing scenarios based on inputs specified by the user, including telescope diameter and temperature, instrument throughput, and wavelength-dependent spectral resolution. `coronagraph` calculates noise contributions from zodiacal and exozodiacal dust, telescope thermal emission, coronagraph speckles, dark current, read noise, and clock-induced charge. The model then provides the wavelength-dependent signal-to-noise (S/N) ratios from which

the error bars are generated on the synthetic data. Previous applications of this model are detailed in [Robinson et al. \(2016\)](#), and the inputs to the noise model are described in Table 4.1. We note that the model can be used to simulate data points with randomized noise instances, but this step is unnecessary for this application as shown by [Feng et al. \(2018\)](#).

Table 4.1: Summary of the mission concepts used as a template for comparing transmission and direct imaging in this report. We adapt instrument specifications from the *Origins* and LUVOIR final reports, respectively. In the noise parameters column, R stands for resolution and t stands for throughput, with coronagraph throughput values calculated by [Kopparapu et al. \(2021\)](#) and optical throughput values calculated by [Stark et al. \(2019\)](#) for direct imaging, and throughput values taken from the *Origins* final report for transmission ([Meixner et al., 2019](#)).

Observing Technique	Mission/ Instrument Template	Wavelength [μm]	Noise Parameters	Inscribed Diameter [m]	Temperature [K]
Transmission	<i>Origins</i> MISC-T	2.8–20.0	$R = 100, t = 0.34$ ($2.8 < \lambda < 10.5 \mu\text{m}$) $R = 200, t = 0.26$ ($10.5 < \lambda < 20.0 \mu\text{m}$)	5.9	4.5
Direct Imaging	LUVOIR-A (B) ECLIPS	0.2–2.0	$R = 7, t = 0.05 \times 0.27$ ($0.2 < \lambda < 0.38 \mu\text{m}$) $R = 140, t = 0.2 \times 0.27$ ($0.38 < \lambda < 0.74 \mu\text{m}$) $R = 100, t = 0.3 \times 0.27$ ($0.74 < \lambda < 2.0 \mu\text{m}$)	15 (6.7)	150

1.2 Retrieval Model

To solve for the atmospheric parameters of the simulated exoplanets, we use the SMART Exoplanet Retrieval (SMARTER) model ([Lustig-Yaeger, 2020](#); [Lustig-Yaeger et al., 2022](#), [Lustig-Yaeger et al. 2023, submitted](#)). This model uses SMART to generate synthetic spectra that are then fit to the observed spectrum to infer posterior probability distributions for a given set of atmospheric parameters, e.g., temperature, surface pressure, gas abundances, and albedo. As in most atmospheric retrievals, these posteriors for the atmospheric characteristics are generated by combining a forward model and an inverse model to solve a Bayesian inference problem.

To solve for the atmospheric parameters that produced an observed or simulated spectrum, we implement Bayesian inference to solve for the posterior probability distributions for atmospheric parameters given the observed spectrum, or the data, $\mathcal{D}$. In this case, the posterior probability refers to the probability that a given atmospheric composition would produce the observed spectrum. The posterior distributions are given by

$$\mathcal{P}(\theta|\mathcal{D}, \mathcal{M}) = \frac{\mathcal{P}(\theta|\mathcal{M})\mathcal{L}(\mathcal{D}|\theta, \mathcal{M})}{\mathcal{Z}(\mathcal{D}|\mathcal{M})}. \quad (4.1)$$

where $\mathcal{P}(\theta|\mathcal{M})$ is the *prior* probability of the model parameters, $\mathcal{L}(\mathcal{D}|\theta, \mathcal{M})$ is the *likelihood* of the observed spectral data $\mathcal{D}$ being produced by a given model at a particular set of parameters $\mathcal{M}(\theta)$, and $\mathcal{Z}(\mathcal{D}, \mathcal{M})$ is the marginal likelihood or *evidence* normalization factor. In this case, the prior refers to the probability of the model parameters at a particular atmospheric state given previous measurements or observations. Since we currently have no prior observations of terrestrial exoplanet atmospheres relevant to this study, we use uninformative priors for all parameters, where the probability of any value within a finite logarithmic range is given an equal weight. As is common in Bayesian inference problems, we implement the logarithm of Equation 4.1. For data with uncorrelated errors, we define the log-likelihood function in relation to the traditional χ^2 goodness of fit metric such that $\log \mathcal{L} = -\frac{1}{2}\chi^2$, where this metric describes the goodness of fit that a given atmospheric state, and its corresponding spectrum given by the forward model, provide to the observed spectrum. The log-likelihood function is a sum over m wavelengths in the spectrum such that,

$$\mathcal{L} = -\frac{1}{2} \sum_{l=1}^m \frac{[\mathcal{M}(\theta_N)_l - \mathcal{D}_l]^2}{\sigma_l^2}. \quad (4.2)$$

Evaluating the likelihood at an atmospheric state θ_N initiates a run of the forward model $\mathcal{M}$ at this state θ_N , and allows Bayes theorem to be evaluated at this point. The marginal likelihood is defined as

$$\mathcal{Z}(\mathcal{D}, \mathcal{M}) = \int \mathcal{P}(\theta, \mathcal{M}) \mathcal{L}(\mathcal{D}|\theta, \mathcal{M}) d\theta. \quad (4.3)$$

This multi-dimensional integral accounts for the likelihood of all atmospheric parameters as well as the number of degrees of freedom allowed in the radiative transfer forward model, where degrees of freedom in this case equals the number of free parameters. While this multi-dimensional integral is generally difficult to calculate, techniques like nested sampling (see below) can more tractably estimate this term and provide a metric for comparing models $\mathcal{M}$ (Skilling, 2006) and determining which model best describes the observed spectrum. Next, we describe our nominal forward model as well as vertically-resolved variants that we compare against.

1.2.1 SMART (Forward Model)

SMART is the core of the retrieval forward models, $\mathcal{M}$. As described above, SMART calculates transmission, emission, and reflectance spectra for exoplanets using stellar and planetary input parameters combined with absorption coefficients for the given atmospheric gases. We use `coronagraph` to degrade this high-resolution spectrum to the resolution of the observing instrument. In general, the forward model takes a set of N atmospheric parameters θ_N to generate a model spectrum, $\mathcal{M}(\theta_N)$, at the same resolution as the data. We now describe in detail the sequence of events that occur inside the forward models.

We make assumptions in LBLABC and SMART to reduce computation time in the retrieval model. We use the previously computed rotational-vibrational line absorption coefficients for each gas produced for the synthetic data, as described in Section 1.1.1, within the retrieval. Although we do not regenerate new absorption coefficient files in successive iterations as the atmospheric state changes due to computational expense, this simplifying assumption does not produce errors in the vicinity of the solution. This necessary assumption is discussed further in Section 4. Additionally, solar zenith angle and the spectral mapping binning criteria are kept fixed for all SMART runs in a given retrieval.

We enable the forward model to construct an atmosphere based on the state vector in one of three modes: evenly-mixed, vertically-resolved with **fixed** pressure points, and vertically-resolved with **free** pressure points. In the evenly-mixed mode, each gas profile is fixed throughout the vertical extent of the atmosphere at the abundance sampled from the prior. In the vertically-resolved mode with fixed pressure points, the user specifies a number of pressure points for a selected gas. Due to the additional computational expense of each fixed pressure point, which corresponds to an additional free parameter in gas abundance space, in this study we limit our exploration to 3 or 5 fixed pressure points per gas. Similarly, although multiple gases could be vertically resolved in a given retrieval, in this study we explore one gas to both reduce computational expense and isolate the effect on the fit. All other gases are assumed to have evenly-mixed profiles. Once the gas abundances have been drawn from the prior for each pressure point in the vertically-resolved profile, the forward model performs a linear interpolation to construct a gas abundance profile.

In the vertically-resolved mode with free pressure points, instead of specifying the pressure point values, the user includes three pressure points *within* the state vector, along with three abundance points. This allows the model to freely select the pressures and abundances that provide the best fit to the observed spectrum, and then perform a linear interpolation to construct the vertical profile. We impose a simple prior that the pressures maintain order from smallest to largest to preserve the uniqueness of the points, thereby breaking a labelling degeneracy. As in the fixed points version of the model, all other gases besides the user-selected, vertically-resolved gas are given evenly-mixed profiles. We show a comparison of these four forward model modes with their corresponding spectra to examine the effect on O_3 , a vertically inhomogeneous gas, in the spectrum of a clear-sky, Earth-like TRAPPIST-1 e in Figure 4.1. We note that the example O_3 abundance profile fits generate visible differences in the depth of the O_3 spectral absorption features, with changes to the the $9.6 \mu\text{m}$ band being more pronounced, relative to the $4.7 \mu\text{m}$ O_3 band where the difference is much smaller.

For all three modeling modes, the full atmosphere is constructed by assuming an isother-

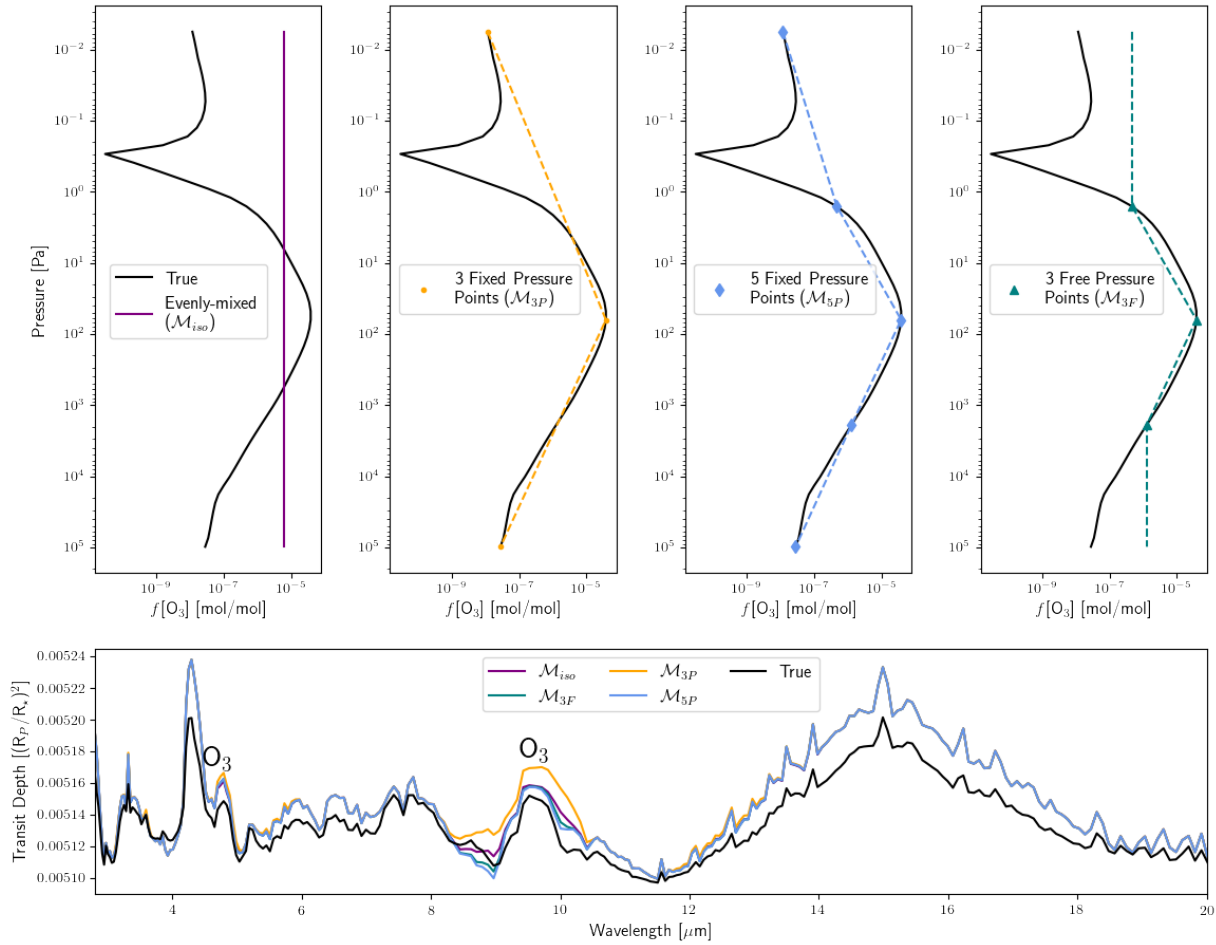


Figure 4.1: A comparison of the evenly-mixed ($\mathcal{M}_{iso}$), vertically-resolved with 3 – 5 fixed pressure points ($\mathcal{M}_{3P}$, $\mathcal{M}_{5P}$), and vertically-resolved with 3 free pressure points ($\mathcal{M}_{3F}$) forward models for retrieving O₃ from a simulated transmission spectrum of a clear-sky, Earth-like TRAPPIST-1 e. The four top panels in the figure show an example of an O₃ profile that the forward model may try to produce to generate a fit to the spectral data. The bottom panel compares the corresponding spectra generated by each forward model to the true simulated spectrum. The assumption of isothermal pressure-temperature profiles in each of the forward models also modifies the continuum and the depth of other spectral features relative to the true spectrum.

mal temperature-pressure (T-P) profile, which is computationally straightforward for the retrieval. Although very few terrestrial planets exhibit isothermal T-P profiles, especially near the surface, isothermal stratospheres are possible, especially for planets orbiting M dwarfs (Lincowski et al., 2018). Since both transmission and reflectance spectra are only weakly sensitive to atmospheric thermal structure Robinson (2017a); Feng et al. (2018), the

isothermal assumption is common in retrieval modeling work. The assumption of an isothermal temperature structure inside the retrieval is also implemented by [Feng et al. \(2018\)](#). We discuss this assumption further in Section 4. We also backfill the atmosphere with N_2 in all three modes to ensure the total volume mixing ratios of each atmospheric state add to unity so that,

$$f[N_2] = 1 - \sum_{j=1}^{n_{gas}} f_j, \quad (4.4)$$

where f_j is the volume mixing ratio of the j -th gas for a total of n_{gas} user-specified gases. N_2 is a fill gas in Venus, Mars, and Earth, the terrestrial planets with significant atmospheres in our Solar System, and similar studies also backfill terrestrial exoplanet atmospheres with N_2 (e.g., [Barstow & Irwin, 2016](#); [Krissansen-Totton et al., 2018b](#); [Changeat et al., 2019](#); [Barstow et al., 2020](#)). However, other fill gases could be used if scientifically warranted, like Ar ([Lustig-Yaeger, 2020](#)), or a mixture of H_2/He (e.g., [Tsiaras et al., 2016](#); [Mollière et al., 2019](#)).

Once the volume mixing ratio of the filler gas has been calculated, the forward model calculates the mean molecular weight μ of the atmosphere using the relation

$$\mu = \frac{\sum_k^{n_P} P_k (\sum_j^{n_{gas}} m_j f_{j,k})_k}{\sum_k^{n_P} P_k}, \quad (4.5)$$

where P_k is the k -th of n_P total pressure layers in the atmosphere, m_j is the molecular mass of the j -th gas out of n_{gas} total gases, and $f_{j,k}$ is the volume mixing ratio of the j -th gas at the k -th pressure layer. Accordingly, backfilling the atmosphere creates a bias for the filling gas that in turn propagates a bias in mean molecular weight. Since observable molecular spectral features are proportional to the scale height of the atmosphere—which is governed by atmospheric mean molecular weight and temperature, as well as gravity (planetary mass)—then a bias in the atmospheric mean molecular weight may lead to the propagation of biases in retrievals of the other properties, and any results that depend on those properties. In this

sense, backfilling with N_2 can create a false constraint that tends not to bias the retrieval of any of the other gases in the transmission spectrum, but does bias other parameters like the surface temperature and also the bulk gas volume mixing ratio if that bulk gas is in fact not N_2 . To avoid this bias caused by using filler gases, other retrieval models use centered-log ratio priors (Aitchison, 1982) for the gas abundances (Benneke & Seager, 2012; Damiano & Hu, 2021). Instead of requiring that all volume mixing ratios sum to one, centered-log ratio priors require that the log of all volume mixing ratios sum to zero. Thus, instead of assuming a bulk gas and a flat prior for all trace gases, centered-log ratio priors allow any retrieval gas to be the bulk gas as well without favoring or disfavoring any species. The implementation of centered-log ratio priors in SMARTER is outside the scope of this study, but will be explored in future work.

Once we have constructed an atmosphere, we pass this atmosphere to SMART to generate either a transmission or reflectance spectrum. In transmitted light, SMART calculates the change in flux ΔF during transit as the fractional transit depth,

$$\Delta F = \left(\frac{R_P}{R_\star} \right)^2, \quad (4.6)$$

where R_P is the planet radius and $R_\star$ is the stellar radius. For these transmission spectrum calculations, we use the SMART ray-tracing model described in Robinson (2017b) with refraction turned on. To calculate the transmission spectrum, SMART may either include the solid body or assume a solid body and compute the differential due to the atmosphere, where this differential constitutes the opacities along a line of sight as well as the refraction, which is a function of the viewing geometry of the star-planet system. Consequently, SMART can calculate a number of different transit products, including the total transit depth (atmosphere and solid body), the relative transit depth (atmosphere *above* the solid body), as well as the effective transit height relative to the surface in kilometers. In the transmission retrieval mode, the user may therefore include the solid-body planet radius R_P and calculate the

total transit depth of the atmosphere in the transmission forward model, as we do here in this study. In the absence of transit Jacobians, we use the effective transit height to indicate where in the atmosphere the true spectrum is sensitive to a given gas. The retrieval user may also include continuum pressure P_0 , and temperature T as free parameters. Using the assumption of an isothermal temperature-pressure profile, the retrieved temperature may also effectively refer to the surface temperature, T_0 .

In reflected light, SMART gives the unit-less planet-to-star flux ratio as,

$$\frac{F_P}{F_\star} = \left(\frac{R_P}{a} \right)^2 \frac{F_{P,TOA}}{F_{\star,TOA}}, \quad (4.7)$$

where a is the semi-major axis of the planet's orbit, $F_{P,TOA}$ is the top of atmosphere planet flux, and $F_{\star,TOA}$ is the top of atmosphere stellar flux. In the reflectance retrieval mode, the user may include surface albedo A_s and surface temperature T_0 as free parameters. SMARTER produces the reflected light spectrum assuming a uniform or gray surface albedo, where we take the average reflectivity of the Earth as $\bar{A}_e = 0.2$ to generate the synthetic spectral data.

Once we have generated the high-resolution spectrum, we use **coronagraph** to produce the degraded spectrum, $\mathcal{M}(\theta_N)$. **coronagraph** convolves the high resolution spectrum with a boxcar filter to down-bin the spectrum to the resolution of the data. With more realistic instruments, this down-binning process can be more complicated if there are known instrument characteristics that need to be accounted for, such as non-linearities in pixel resolution. Since we are simulating hypothetical instruments that have not yet been designed, built, or characterized, we implement this simplification in place of a more defined instrument response function.

Once we have generated the degraded spectrum $\mathcal{M}(\theta_N)$ from the atmospheric state θ_N , we pass the spectral inputs to the likelihood function, $\mathcal{L}$. This in turn triggers the logic of the inverse model. The inverse model maximizes the posterior probability and minimizes the residuals in the fit to the observed spectrum. After assessing the fit to the data $\mathcal{D}$, the

inverse model selects the next set of parameters randomly drawn from the priors. Next, we describe the inverse model component of the retrieval framework.

1.2.2 PyMultiNest (Inverse Model)

For this work, we employ a nested sampling inverse model to numerically solve Bayes theorem called `PyMultiNest` (Buchner et al., 2014). `PyMultiNest` is a Python wrapper for `MultiNest`, an open-source Fortran program (Feroz & Hobson, 2008; Feroz et al., 2009, 2013). Introduced by Skilling (2004), nested sampling is commonly used in atmospheric retrievals as a computationally efficient method for solving the Bayesian inference problem (Benneke & Seager, 2013; Waldmann et al., 2015; Lavie et al., 2017; Gandhi & Madhusudhan, 2018). Nested sampling algorithms maintain a user-specified set of live points, each containing a sample of the parameters θ from the prior distributions $\mathcal{P}(\theta)$. At each successive iteration, nested sampling implements a cost function by removing the live point from the set with the lowest likelihood, and replaces it with a new live point with parameter samples randomly drawn from the prior distributions such that the new live point samples have a higher likelihood than the last. Also at each iteration, `PyMultiNest` calculates the log evidence, and then compares the change in evidence between successive iterations using the live points. Once the change in evidence is equal to or smaller than a tolerance value, the process is complete. The evidence tolerance is thus the model convergence criterion, where we implement `PyMultiNest` with 1000 live points and an evidence tolerance of ≤ 0.5 .

1.3 Model Comparison

To compare our models and assess their accuracy, we analyze the model evidence. Nested sampling has the added benefit of concurrently calculating the evidence term by reducing the multidimensional integral $\mathcal{Z}(\mathcal{D})$ described above.

The evidence metric rewards the model with the best fit to the data while penalizing excess parameter dimensions, so the model with the higher evidence is typically taken to be

the preferred model given fits to the same data. Thus, to compare two models $\mathcal{M}_1$ and $\mathcal{M}_2$, we merely calculate the difference in their log-evidence terms so that,

$$B_{1,2} = \ln Z_1 - \ln Z_2, \quad (4.8)$$

where B is known as the Bayes factor. In this formulation, a Bayes factor with $B_{1,2} > 0$ indicates that $\mathcal{M}_1$ is the favored model, whereas a Bayes factor with $B_{1,2} < 0$ indicates that $\mathcal{M}_2$ is preferred instead. In this study, we interpret the significance of our Bayes factors based on the empirically-derived odds ratios proposed by [Jeffreys \(1998\)](#).

To assess the quality of the fits provided to the spectra by our models, we compare the χ^2 for the median spectral solutions produced by the retrieval. To calculate χ^2 , we use the fact that $\mathcal{L} = -\frac{1}{2}\chi^2$. The Bayes factor analysis allows us to determine whether additional model complexity is justified by the data, while separately analyzing the χ^2 values allows us to determine which model provided the best fit to the data regardless of model complexity. In combination, these analyses help determine whether improved data resolution or precision may increase the evidence in favor of the more complex model.

2 Experiments

In this study, we apply both our evenly-mixed and vertically-resolved retrieval models to a set of atmospheric cases in transmitted and reflected light, which we describe in detail here. In transmission, we simulate an Earth-like, inhabited TRAPPIST-1 e and an abiotic, uninhabited TRAPPIST-1 e. Our uninhabited TRAPPIST-1 e represents a possible O_3 false-positive case, with a 1-bar, false positive O_3 , desiccated atmosphere based on the most extreme case described in [Gao et al. \(2015\)](#). In reflected light, we simulate an inhabited, clear-sky Pre-industrial Earth at 10 parsecs and the same atmosphere with thermally self-consistent H_2O clouds. We describe these cases in detail below, and summarize them in Table 4.2 along with our chosen exposure times for both methods. We compare two distinct

telescope architectures using two different observation methods that will be optimized for different wavelength ranges and planetary systems. To make the investigation as fair as possible given these intrinsic differences, we run our retrieval model for similar *Origins* and LUVOIR-B exposure times ($t = 40$ hours). For the larger aperture LUVOIR-A, less exposure time is required to achieve the same quality of data. Choosing similar *Origins* and LUVOIR-B exposure times allows us to draw conclusions about the comparative value of observing optimal transmission and direct imaging targets. Overheads due to, for example, slew, settle, and out-of-transit baseline, are neglected from these exposure calculations. We show a comparison of the structure and composition of each atmospheric case addressed in Figure 4.2. We note that in the case of the clear-sky, false positive O_3 , H_2O -poor TRAPPIST-1 e, the abiotically-generated O_3 abundances are comparable to that of the Earth-like TRAPPIST-1 e with an oxygenic biosphere. Both profiles show a stratospheric O_3 bulge, though the desiccated planet has a much shallower bulge.

Table 4.2: A summary of the observing cases explored in this work. While *Origins* is also capable of observing some exoplanets in emission and LUVOIR is also capable of transmission observations for a limited number of targets, we focus on each of the observatory’s main observing modes to provide instrument templates for comparing transmission and direct imaging.

Mission	Observing Method	Planet/Star	Biosphere	Comparison Case
<i>Origins</i>	Transmission $t_{exp} = 40$ hours	TRAPPIST-1 e	Clear-Sky Earth-like, 1-bar, N_2 dominated	Modern Abiotic 1-bar N_2 CO ₂ dominated, false positive O_3 , desiccated (Gao et al., 2015)
LUVOIR-A (B ^c)	Direct Imaging $t_{exp} = 10$ ($\sim 40^c$) hours	Earth-Sun (10 pc)	Clear-Sky Earth, 1-bar, N_2 dominated	Modern Earth with H ₂ O Clouds, 1-bar, N_2 dominated

^cSeparate noise models and retrievals are not run for LUVOIR-B, and differences between architecture mirror diameter and throughputs complicate scaling between A and B exposure times.

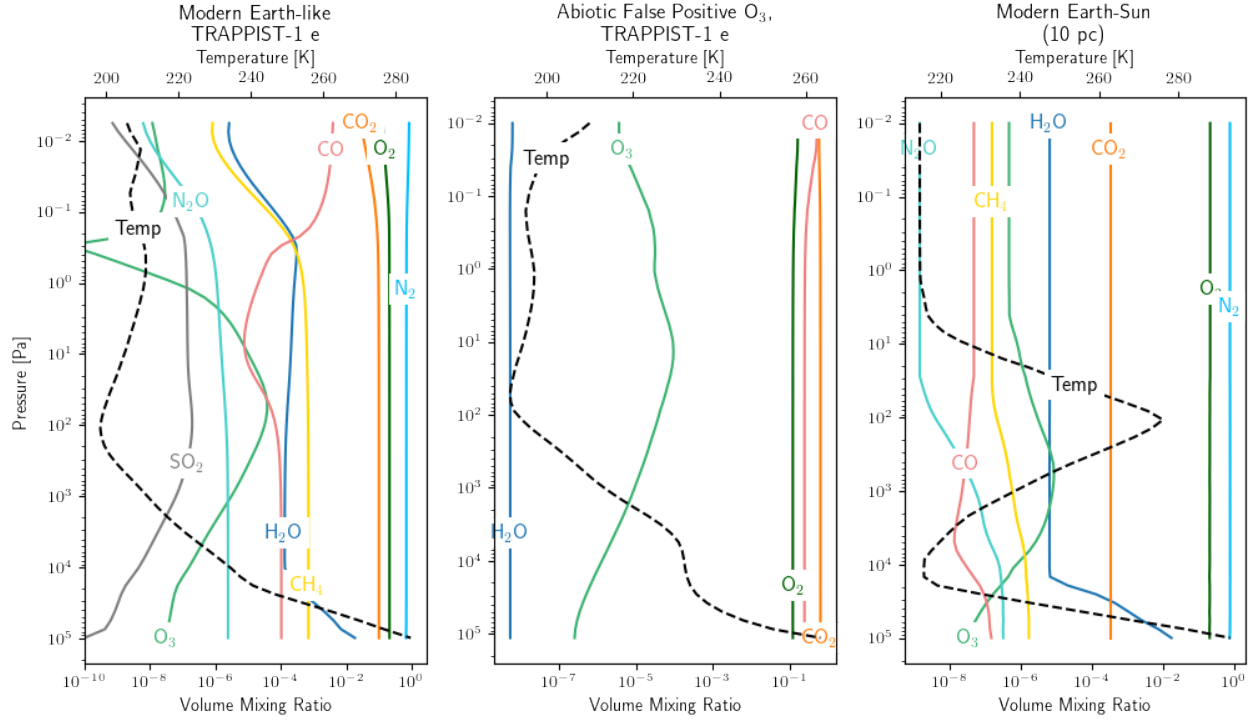


Figure 4.2: A comparison of the vertically-resolved temperature structure and composition of each atmospheric case addressed in this study. The Earth-like and abiotic, ozone false-positive planets (first and second panels) have similar ozone inventories with visible stratospheric bulges.

2.1 Transmission

For the transmission experiments, we use spectra generated for an Earth-like TRAPPIST-1 e with a photosynthetic biosphere and an abiotic, desiccated TRAPPIST-1 e with false positive O_3 as the synthetic data inputted to the retrieval. Our Earth-like TRAPPIST-1 e transmission spectrum is generated based on temperature and mixing ratio profiles self-consistently calculated for a habitable zone planet around an M8V star (Davis et al., *in prep*). The abiotic case represents a possible O_3 false-positive generated by CO_2 photolysis in a dry, 1-bar atmosphere (Gao et al., 2015). In this false positive scenario, the extremely low abundance of H_2O prevents CO_2 from recombining after photodissociation, thereby allowing enough O_3 to accumulate to mimic the O_3 levels we would expect on an inhabited planet. This is an important false positive mechanism, as an *Origins*-like mission would observe

transiting planets in the mid-IR where O_3 must be used as a proxy for O_2 .

The Earth-like TRAPPIST-1 e atmosphere is also modeled with thermally self-consistent water (stratocumulus) and ice (cirrus) clouds. The vertical cloud distributions and ice cloud optical properties are self-consistently calculated with a single scattering model for a variety of particle types as described in [Meadows et al. \(2023\)](#). However, when we simulate the spectroscopic effects of these clouds on the planet’s transmission spectrum, we see no significant changes to the spectrum compared to the clear-sky case, except for minimal changes in transit height probed by the $8 - 13 \mu\text{m}$ window region, as shown in Figure 4.3, which effectively raises the minimum altitude probed from 10km to 12km. Therefore, we conclude that Earth-like water and ice clouds are unlikely to have a significant impact on our transmission retrieval experiments and we omit them here. These results are consistent with [Meadows et al. \(2023\)](#), who show that clouds do not significantly impact the transit depths of spectroscopic features for an Earth-like TRAPPIST-1 e for wavelengths greater than $2.8 \mu\text{m}$. As a caveat, we note that 3-D, General Circulation Models (GCMs) include more sophisticated dynamics than 1-D models, and can predict different cloud altitudes as well as a distinct day-side cloud accumulation pattern for tidally locked planets like those in the TRAPPIST-1 system ([Komacek et al., 2020](#)), which may or may not be synchronously rotating. However, the 1-D results used here predict clouds that are at altitudes comparable to, if not slightly higher than those predicted for TRAPPIST-1 e by 3-D models ([Fauchez et al., 2019](#); [Meadows et al., 2023](#)). Though 3-D models are currently computationally expensive and thus difficult to implement within a spectroscopic retrieval framework, current work is exploring how to make the implementation of 3-D models in atmospheric retrievals more tractable by speeding up radiative transfer calculations ([MacDonald & Lewis, 2021](#)).

For the 1-bar false positive O_3 planet, we do not consider or model aerosols, because we concluded that neither clouds, hazes, or dust would likely affect the observations. Since the abundance of H_2O is well below the saturation vapor pressure required for water to condense, we conclude that water clouds may not form. Additionally, since this is a methane-free,

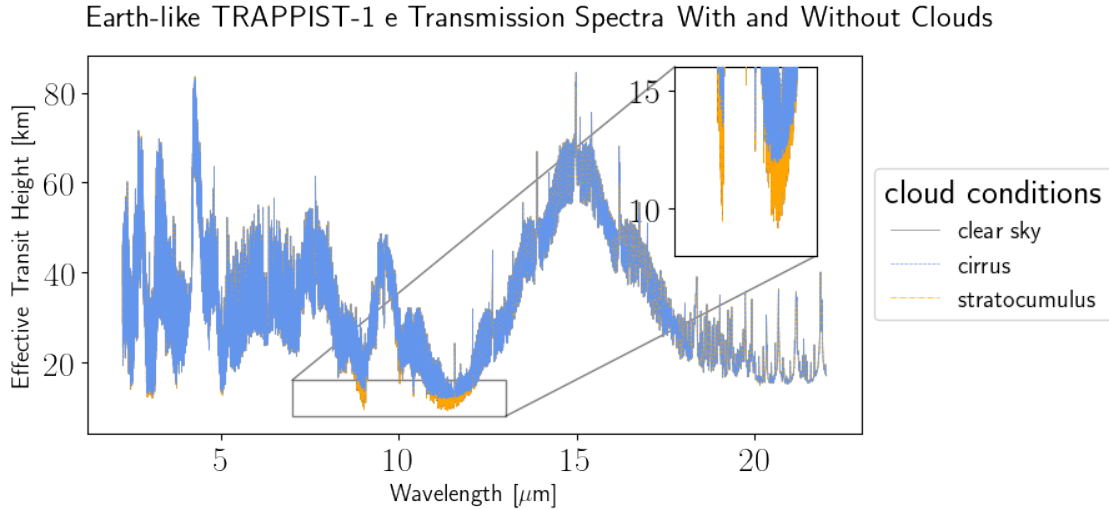


Figure 4.3: A comparison of simulated transmission spectra for TRAPPIST-1 e modeled with a self-consistent, Earth-like atmosphere. When self-consistent ice and water clouds are included, no significant spectroscopic changes occur to the spectrum. We see small impacts due to clouds in the relatively transparent portion of the atmosphere between 8 – 13 μm , where the clouds reduce the minimum altitude probed from 10 to 12 km.

highly-oxidizing Mars-like atmosphere, we also rule out the presence of organic hazes since [Arney et al. \(2016\)](#) showed that a CH_4/CO_2 ratio of ≥ 0.2 is required for haze to form (and become spectrally significant) in an Earth-like atmosphere. However, given a desiccated Mars-like atmospheric composition, it is reasonable that global dust storms could occur on such a planet, and that these dust particles may impact the opacity of the atmosphere. Global dust storms on Mars are reasonably well-studied, with observations reporting “rocket storms” that may loft dust as high as 70 km ([Wang et al., 2018](#)). However, the Martian atmosphere is considerably thinner than the 1-bar atmosphere simulated here, so dust storms on Earth represent a better analogue to our Earth-like planet. On Earth, the atmospheric transport of Saharan dust is a well-studied phenomenon. Thus, we can look to these studies as an analogue for dust transport within a 1-bar atmosphere to understand how dust particles may impact the transmission spectra of the 1-bar, H_2O -desiccated planet in our study. Lidar observations of the Earth taken from space over a 5 year period have shown that the peak vertical distribution of the dust does not exceed 5 km ([Tsamalis et al., 2013](#)). We therefore find that dust particles in this atmosphere are unlikely to impact observations

in transmission.

In transmitted light, we attempt to retrieve the planet radius, continuum pressure, surface temperature, and gas abundances. A summary of the parameters and their associated priors are described in Table 4.3.

Table 4.3: A summary of the parameters and priors explored in the two transmission cases of this study (TRAPPIST-1 e Earth-like and false positive O₃ planet). In the non-vertically-resolved retrievals, all gas profiles from the forward model are generated assuming evenly-mixed abundances. R_p is planet radius, P_0 is continuum pressure, and T_0 is surface temperature. Finally, f indicates the gas mixing ratio of a given molecule (e.g., H₂O, CH₄, CO₂, O₂, O₃, CO).

Parameter θ		Prior $\mathcal{P}(\theta)$	Lower Bound	Upper Bound
Planetary radius	R_p	Uniform	$0.85R_{\oplus}$	$0.95R_{\oplus}$
Continuum pressure	P_0	Uniform in log-space	10^2 Pa	10^6 Pa
Surface temperature	T_0	Uniform	100 K	400 K
Gas abundance	$f[x]$	Uniform in log-space	10^{-10} mol/mol	1 mol/mol
^{a,b} Vertically-resolved abundance	O ₃ $f[\text{O}_3]_n$ ($n = 3, 5$)	Uniform in log-space	10^{-10} mol/mol	1 mol/mol
^b Vertically-resolved pressure	O ₃ $P(\text{O}_3)_n$ ($n = 3, 5$)	Uniform in log-space	10^{-2} Pa	10^5 Pa

^aParameters considered in the fixed points models ($\mathcal{M}_{3P}$, $\mathcal{M}_{5P}$) and ^bthose considered in the free points model ($\mathcal{M}_{3F}$).

2.2 Direct Imaging

For the two direct imaging experiments, we use spectra generated for an Earth-like environment with and without clouds as the synthetic data input to the retrieval. Our Earth spectrum is generated based on temperature and mixing ratio profiles from case 62 of the Intercomparison of Radiation Codes in Climate Models (ICRCCM), representing the averaged midlatitude Earth during the summer months, as shown in [Lincowski et al. \(2018\)](#). Comparison of spectra generated using ICRCCM and full-disk-averaged spectra of the Northern hemisphere in the spring ([Robinson et al., 2011](#)) showed only minor differences in the depths of water vapor bands, which are known to be highly spatially and temporally variable on

the Earth. Unlike in the M dwarf transmission cases, the known potential false positives for abundant O_2 for planets orbiting G dwarfs are difficult to generate and difficult to distinguish from true biospheres (Wordsworth & Pierrehumbert, 2014; Meadows, 2017). In particular, G-dwarf planets with low abundances of non-condensable gases (such as N_2) may produce inefficient cold traps that allow water into higher levels of the atmosphere, where it is more readily photolyzed to produce O_2 (Wordsworth & Pierrehumbert, 2014). Therefore, for our comparison case we instead consider the impact of realistic clouds on our ability to retrieve planetary properties in direct imaging. For noise calculations, we assume the observed planets are orbiting a G dwarf, at a distance of 10 pc.

While cloud decks have been shown to effectively truncate the atmospheric scale height and thus reduce the size of spectral features in transmission (Lincowski et al., 2018; Fauchez et al., 2019; Lustig-Yaeger et al., 2019c; Komacek et al., 2020), the behavior of atmospheric clouds in reflected light is more complex and may actually enhance spectral features in some cases (e.g., Rugheimer et al., 2013). For our synthetic spectral observations, the vertical cloud distributions and ice cloud optical properties are again self-consistently calculated with a single scattering model for a variety of particle types as described in Meadows et al. (2023). We simulate Earth with realistic patchy clouds by linearly combining the water (25%), ice (25%), and clear-sky (50%) spectra to produce the cloudy spectrum. Finally, we assume a gray (constant with wavelength) surface albedo of 0.2 as in Feng et al. (2018) and Damiano & Hu (2022). Wavelength dependent changes in surface albedo expected from a realistic planetary surface, which likely contribute significantly to an observed spectrum, are beyond the scope of this study and will be the subject of future work. A summary of the parameters and their associated priors are described in Table 4.4.

Table 4.4: A summary of the parameters and priors explored in the direct imaging (clear and cloudy Earth) cases of this study. In the non-vertically-resolved retrievals, all gases profiles from the forward model are generated assuming evenly-mixed abundances. A_s is a uniform surface albedo of 0.2 reported as a unitless ratio and T_0 is surface temperature. Finally, f indicates the gas mixing ratio of a given molecule (e.g., H_2O , CH_4 , CO_2 , O_2 , O_3 , CO).

Parameter θ		Prior $\mathcal{P}(\theta)$	Lower Bound	Upper Bound
Surface albedo	A_s	Uniform in log-space	10^{-2}	1
Surface temperature	T_0	Uniform	100 K	400 K
Gas abundance	$f[x]$	Uniform in log-space	10^{-12} mol/mol	1 mol/mol
^{a,b} Vertically-resolved abundance	O_3 $f[\text{O}_3]_n$ ($n = 3, 5$)	Uniform in log-space	10^{-12} mol/mol	1 mol/mol
^b Vertically-resolved pressure	O_3 $P(\text{O}_3)_n$ ($n = 3, 5$)	Uniform in log-space	10^{-2} Pa	10^5 Pa

^aParameters considered in the fixed points models ($\mathcal{M}_{3P}$, $\mathcal{M}_{5P}$) and ^bthose considered in the free points model ($\mathcal{M}_{3F}$).

3 Results

The results of our study consist of two main components: posterior distributions generated by the retrieval, which provide the distribution of atmospheric properties that best match the spectrum, and the fits to the spectrum generated by running SMART for atmospheric states sampled from these posteriors. We compare the posterior distributions to the true profiles to assess the accuracy of our retrieval model and our ability to constrain parameters from the data. We supply *all* posterior distributions with covariances in Appendix 5. In the main text of this paper, we show the marginal posterior distributions for the evenly-mixed forward model cases. All non- O_3 posteriors for the vertically-resolved cases are within $1\text{-}\sigma$ of their evenly-mixed counterparts, so we omit them in the main text. We show the posterior distributions for the significant vertically-resolved O_3 profiles in the main text. The calculation of the posteriors includes the calculation of the evidence, $\log \mathcal{Z}$. We use the evidence term as a metric for determining the relative likelihood of a given model, and we compare model evidences via Bayes factors to perform model comparisons. Finally, we use

the χ^2 values for the median spectral fit to assess the goodness-of-fit of each model. We first individually compare our pairs of transmission and direct imaging cases, and then compare the results of our transmission cases to the results of our direct imaging cases.

3.1 Transmission

We first compare our retrieval cases in transmitted light. We compare the posterior distributions from the evenly-mixed cases for the Earth-like and abiotic false positive O₃ TRAPPIST-1 e in Figure 4.4 below. To indicate the atmospheric layers to which the true, noiseless spectrum is most sensitive, we rerun the input spectrum without a given gas and determine the wavelengths over which the gas is spectrally active. Then, we calculate the size of these absorption features in the true spectrum in terms of the effective transit height, and convert this to the corresponding range of pressures in the atmosphere that are displayed in Figure 4.4.

For the Earth-like TRAPPIST-1 e case, we first assess the retrieval of the surface pressure and the (nominally) evenly-mixed gases in the true atmosphere, which include CH₄, CO₂, O₂, and CO. We retrieve a median continuum pressure of 8×10^4 Pa, which is lower by a factor of 80% from the true surface pressure (± 1 dex). This median retrieved pressure is consistent with the pressure found at an altitude of ~ 2 km in the true atmosphere. However, the broad $1\text{-}\sigma$ distribution of pressures we retrieve span $10^4 - 4.1 \times 10^5$ Pa, or 0.1 – 4.1 bars. The lower bound of this range corresponds to an altitude of 23 km in the true atmosphere, while the 4.1-bar upper bound exceeds the true surface pressure of 1 bar and therefore has no altitude analog in the true atmosphere. For CH₄ and CO₂, we retrieve median abundances that are within a factor of 2 of the true mixing ratios. We note that at $1\text{-}\sigma$, the retrieved CH₄ abundance may be as low as 43 ppm, but at 2- and $3\text{-}\sigma$ the abundance may be as low as 8 ppm and 2 ppm, respectively. The retrieved O₂ volume mixing ratio is not constrained at all, with a posterior that provides no more information than the prior. Lastly, the CO volume mixing ratio is not precisely constrained, with a distribution peak that aligns with

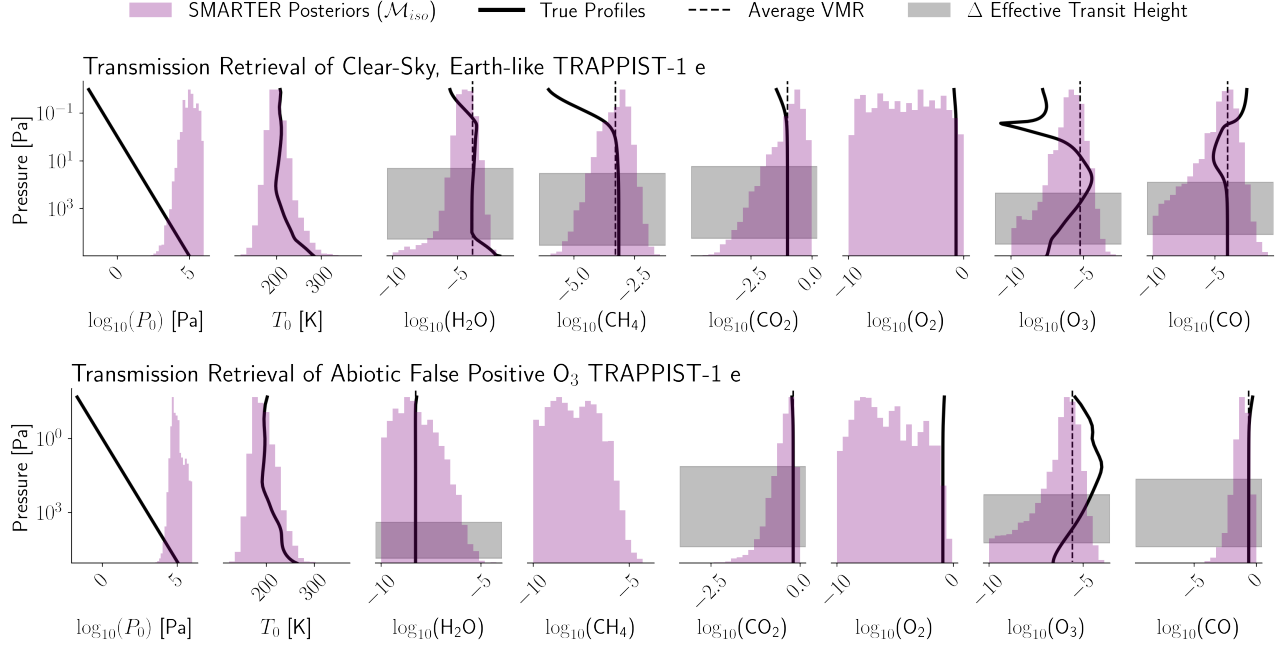


Figure 4.4: A comparison of the posterior distributions for the Earth-like and abiotic false positive O₃ TRAPPIST-1 e using the evenly-mixed forward model ($\mathcal{M}_{iso}$). The black lines represent the true profiles used to generate the observed spectrum, and the dotted black lines indicate the corresponding average volume mixing ratios (VMR) of the true gas abundance profiles. We calculate the change in effective transit height for the spectral features of each gas in the noiseless input spectrum, to indicate the atmospheric layers to which the spectrum is sensitive (gray shading). For both atmospheres, there are no O₂ absorption features in the 2.8 – 20.0 μm mid-IR wavelength range we considered, so we are unable to perform this analysis. For the abiotic planet, there is no CH₄ in the atmosphere, so we do not have true values to compare to.

the evenly-mixed portion of the CO profile to which the spectrum is sensitive, but a 1- σ interval spanning 5 orders of magnitude in volume mixing ratio.

Assessing the retrieval of parameters that change with pressure in the true atmosphere is more complex, and requires us to consider a number of reference points along the true profile to which the retrieval may be sensitive. We note that the retrieved isothermal temperature distribution is consistent with the temperature in the upper atmosphere $\sim 50 - 70$ km ($\sim 10^2$ Pa), but biased low by 30% relative to the true surface temperature, T_0 . This suggests that the transmission spectrum is not probing the surface environment despite the lack of clouds in this atmosphere. Similarly, the retrieved H₂O volume mixing ratio is underestimated by 80%

relative to the true surface water vapor abundance, but the change in effective transit height for the H_2O features imply that the true spectrum should be sensitive primarily to the lower stratospheric abundance as shown in Figure 4.4, which is evenly mixed. However, even if we adjust our comparison to account for this spectral sensitivity, the retrieved H_2O volume mixing ratio distribution is within a factor of 5 lower than this isothermal, evenly-mixed region of the true water profile. Finally, O_3 is not precisely constrained with the evenly-mixed model, with a broad posterior distribution that spans the entire width of the prior. Furthermore, the retrieved O_3 abundance is underestimated relative to the bulge abundance, but the upper bound of the distribution coincides with this maximum O_3 abundance. The retrieved O_3 abundance is consistent with both the true abundance at $\sim 10^3$ Pa, which is within the atmospheric column probed by the spectrum, and the average volume mixing ratio of the true ozone profile within $1\text{-}\sigma$ (albeit underestimated within a factor of 7).

For the abiotic false positive O_3 TRAPPIST-1 e, the evenly-mixed gases now include H_2O , as the planet is sufficiently desiccated so as to inhibit condensation. We retrieve a median continuum pressure of 1.0×10^5 Pa, which is within a factor of 0.9 of the true surface pressure (± 1 dex). This median retrieved pressure is consistent with the pressure found at an altitude of ~ 1 km in the true atmosphere. However, similar to the Earth-like TRAPPIST-1 e, we find that the $1\text{-}\sigma$ pressure distribution we retrieve spans a broad range of pressures from $3 \times 10^4 - 4 \times 10^5$ Pa (0.3 – 4 bars), where the lower bound corresponds to an altitude of 11 km in the true atmosphere, and the upper bound does not correspond to any altitude in the true atmosphere. We retrieve an upper limit on the H_2O volume mixing ratio, which is within a factor of 33 of the true volume mixing ratio, and the spectrum appears to be sensitive to the optically thin H_2O features closer to the surface than in the case of the Earth-like atmosphere. At $1\text{-}\sigma$, the upper limit on the water vapor abundance is 167 ppb, while at $2\text{-}\sigma$ it is 5 ppm and at $3\text{-}\sigma$ it is 13 ppm. Similarly, we retrieve a $1\text{-}\sigma$ upper limit of ~ 70 ppb on CH_4 , a very low abundance that is consistent with the absence of CH_4 in this atmosphere. However, the 2- and $3\text{-}\sigma$ upper limits are much higher at 5 ppm and 27

ppm, respectively. For CO_2 , we retrieve a median volume mixing ratio of 40%, which is underestimated by a factor of 1.5 relative to the truth for this CO_2 -dominated atmosphere. Similar to the Earth-like case, the O_2 volume mixing ratio is not constrained. Lastly, the CO volume mixing ratio is underestimated by a factor of 4.

For this desiccated planet, the pressure-dependent parameters include temperature and O_3 , which shows a stratospheric bulge akin to the Earth-like atmosphere. The retrieved isothermal temperature distribution is biased low by 3% from the coolest point (~ 200 K) in the true temperature profile. Finally, the broader peak of the retrieved O_3 distribution for the evenly-mixed forward model ($\mathcal{M}_{iso}$) more closely aligns with the average volume mixing ratio than in the case of the Earth-like TRAPPIST-1 e, but it is still slightly underestimated. As in the case of the Earth-like atmosphere, the retrieved O_3 abundance is consistent with the true abundance in one of the atmospheric layers probed by the spectrum $\sim 10^4$ Pa, but the retrieved distribution is very broad and spans the entire width of the prior. We summarize and compare all retrieved values from both of our transmission cases to their corresponding truths in Table 4.5.

The vertically-resolved forward models ($\mathcal{M}_{3P}$, $\mathcal{M}_{5P}$, $\mathcal{M}_{3F}$) are unable to retrieve an O_3 abundance profile with the characteristic stratospheric bulge for both the Earth-like and abiotic TRAPPIST-1 e, from the transmission spectra. In fact, for both atmospheric cases, all three vertically-resolved models effectively reproduce an evenly-mixed O_3 volume mixing ratio profile due to the lack of constraints in abundance space (abundance and pressure space) for $\mathcal{M}_{3P}$ and $\mathcal{M}_{5P}$ ($\mathcal{M}_{3F}$). We show a comparison of the vertically-resolved O_3 profiles for the Earth-like and abiotic atmospheric cases using the forward model with 3 free pressure points ($\mathcal{M}_{3F}$) in Figure 4.5. In Figure 4.6, we show secondary total column abundance posteriors calculated by sampling the pressure, temperature, and O_3 abundance profiles from the retrieval posteriors, generating the putative atmosphere using the relevant forward model (evenly-mixed or vertically-resolved), and solving for the column abundance using the ideal gas and hydrostatic equilibrium equations. We compare the size and location

of these distributions to the total column abundances calculated for the “true” atmospheres used to generate the synthetic data. For both the Earth-like and abiotic cases, and across all models, we find that the retrieval of the total column abundance is highly unconstrained from these transmission spectra. In the case of the Earth-like planet, no total column abundance distribution is biased higher from the truth by more than $1.1\text{-}\sigma$, but the peak column abundances from these distributions vary from being a factor of 1.5 to 200 times larger than truth. For the abiotic planet, these peaks vary from being a factor of 0.4 to 5 times larger than truth. In both cases, the 5-fixed points model ($\mathcal{M}_{5P}$) shows the largest bias in total column abundance higher than the truth. For the Earth-like planet, the 3-free points model produces the total column abundance that is least biased relative to the true value, while for the abiotic planet no model cases are particularly biased, but the evenly-mixed model produces the least accurate distribution.

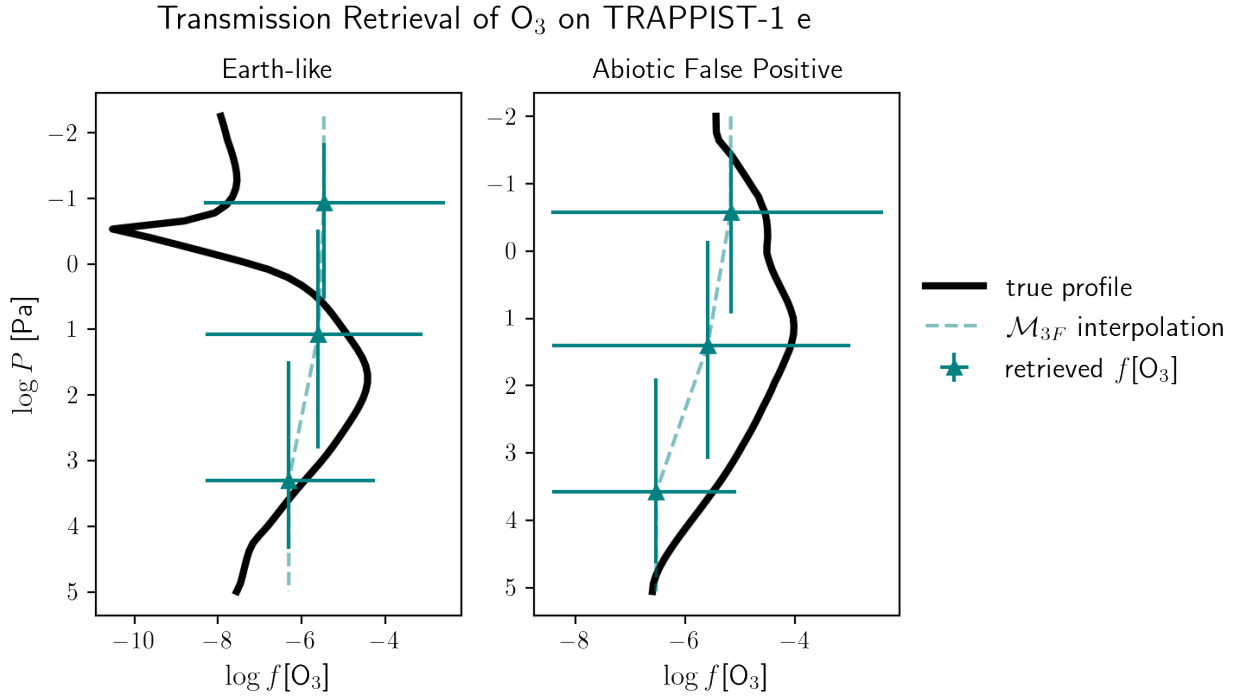


Figure 4.5: A comparison of the vertically-resolved O₃ profiles retrieved by the $\mathcal{M}_{3F}$ forward model for the Earth-like and abiotic TRAPPIST-1 e cases. We note that, for both atmospheres, the retrieval effectively reproduced an evenly-mixed O₃ profile despite being allowed to move freely in pressure and abundance space for 3 points along the profile.

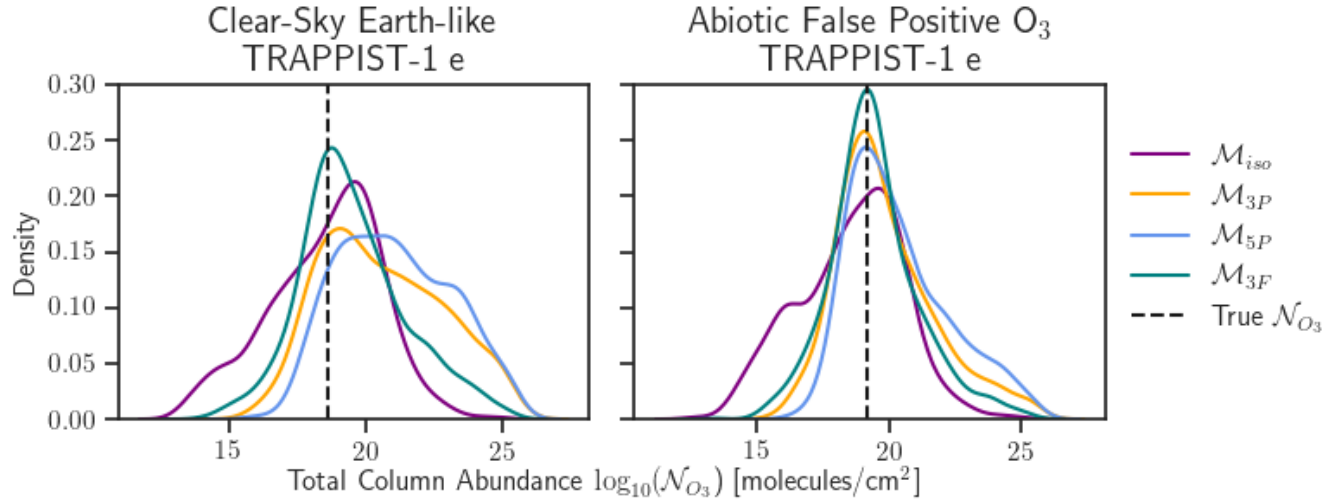


Figure 4.6: A comparison of O_3 total column abundance distributions calculated for retrieval models with and without vertical resolution for both transmission cases. For both cases and across all models, we find that the total column abundance of O_3 is broadly distributed and unconstrained, suggesting that the transmission spectra are relatively insensitive to the column even in the absence of aerosols.

For both the Earth-like and abiotic TRAPPIST-1 e cases, we find that the evenly-mixed model and the vertically-resolved models provide plausible fits to the data, so the evidence values effectively favor the least complex, evenly-mixed model. All χ^2 , evidence, and Bayes factor values are reported in Table 4.7. Though the more flexible models produce smaller χ^2 values, the model evidence demonstrates that these more complex models are not statistically favored at the signal-to-noise level explored here. In fact, in the case of the 3-free points model $\mathcal{M}_{3F}$, the Bayes factor $B = -1.8$ suggests weak evidence in favor of the least flexible, evenly-mixed model $\mathcal{M}_{iso}$. This is because the model evidence penalizes models with additional parameters that do not sufficiently improve the fit to the data. This indicates that we find no statistical evidence in favor of vertically parameterized ozone profiles for the transmission cases.

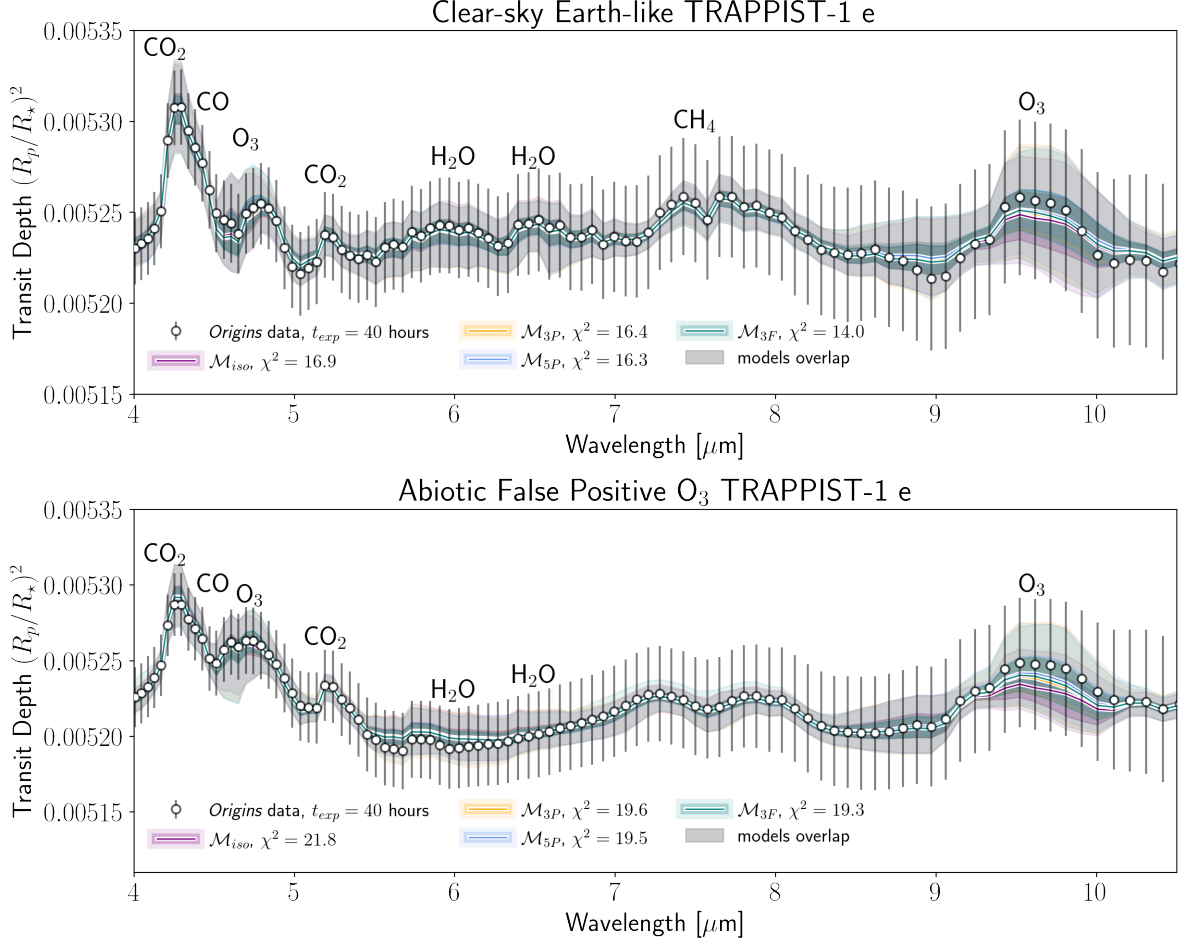


Figure 4.7: A comparison of the median spectral fits with $\pm 1\sigma$, 3σ envelopes for the Earth-like and abiotic false positive O_3 TRAPPIST-1 e. On each plot, we show the fit to the spectrum supplied by each version of the forward model and report its associated χ^2 value. Note that the lines and envelopes overlap significantly and appear gray when they do so. All models provide plausible fits to all spectral features for both transmission cases. We zoom in on the 4-12 μm range of the spectrum, as this is where the signal-to-noise is largest.

3.2 Direct Imaging

We now compare our retrieval cases in reflected light, beginning with the posterior distributions from the evenly-mixed cases for the clear-sky and cloudy Earth in Figure 4.8. To show the atmospheric layers to which the true spectrum is sensitive, we compute the radiance Jacobian as a function of gas abundance in each atmospheric layer, to calculate wavelength dependent spectral sensitivity to each gas, vertically-resolved throughout the atmospheric column. We then integrate the Jacobians over all wavelengths in the simulated data range

for each gas, and report the result as a pressure-dependent Jacobian contribution.

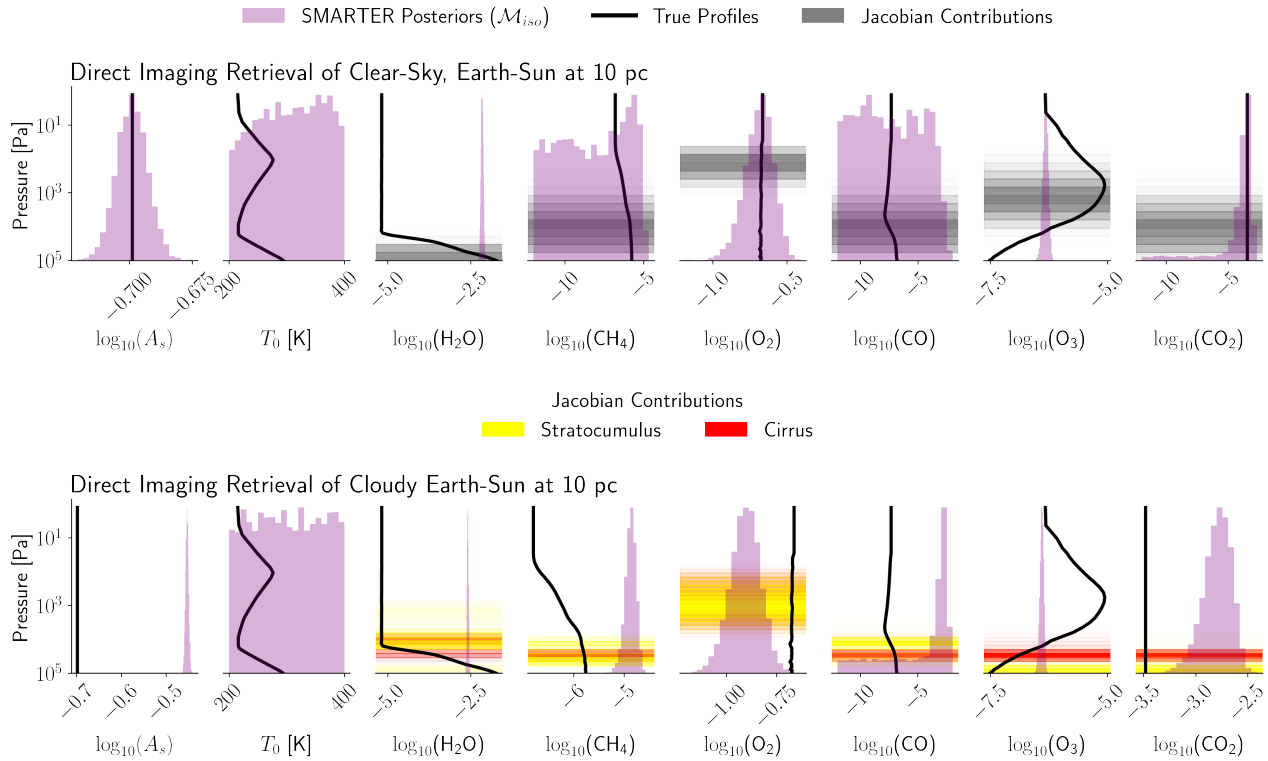


Figure 4.8: A comparison of the posterior distributions for the clear-sky and cloudy Earth using the evenly-mixed forward model ($\mathcal{M}_{iso}$). Note that the y-axis on the albedo subplot provides a reference axis for interpreting the pressure-dependent profiles of all other parameters in the plot; we do not retrieve albedo as a function of pressure. The black lines represent the true profiles used to generate the observed spectrum. For the clear-sky case, the gray shaded regions indicate the Jacobian contributions to the true spectrum, where the opacity of these regions scales with their relative weight. Thus, a darker gray region indicates the atmospheric layer(s) to which each spectral feature is most sensitive. For the cloudy case, we show the Jacobian contributions to the true spectrum for an atmosphere with either 100% stratocumulus (yellow) or 100% cirrus clouds (red). The opacity of these regions also scales with their relative weight, and regions with significant overlap between stratocumulus and cirrus clouds Jacobian contributions take on an orange hue.

For the clear-sky Earth case, beginning with the largely pressure-independent parameters, we observe that the average surface albedo is accurately retrieved and well constrained. In contrast, both the retrieved CH_4 and CO mixing ratio distributions are not constrained, though we gain valuable upper limits on the abundances of both gases. For CH_4 , we retrieve a $1\text{-}\sigma$ upper limit of 1.3 ppm, and a $3\text{-}\sigma$ upper limit of 1.86% VMR. The retrieved O_2

volume mixing ratio is confidently constrained at 21%, which matches the true Earth oxygen abundance of 21%. Lastly, the CO₂ volume mixing ratio appears to be accurately retrieved within a factor of 2 ($<1\text{-}\sigma$) of the true value, but the distribution contains a tail trailing towards lower abundances. The $1\text{-}\sigma$ upper bound of the retrieved CO₂ distribution is 489 ppm, and the $3\text{-}\sigma$ upper bound is 91% VMR.

For the pressure-dependent parameters, we observe that the retrieved isothermal temperature distribution is completely unconstrained, but the retrieved H₂O volume mixing ratio distribution is confidently constrained within 10% at $1\text{-}\sigma$ near the true surface value, to which the spectrum is sensitive. Specifically, our median retrieved H₂O volume mixing ratio is within a factor of 3 of the true surface value, and is consistent with the tropospheric abundance at 0.7 bars (7×10^4 Pa, ~ 2.5 km). Finally, the retrieved O₃ volume mixing ratio for the evenly-mixed forward model ($\mathcal{M}_{iso}$) is underestimated relative to both the average volume mixing ratio of the true profile ($<2\text{-}\sigma$) and the bulge abundance ($<3\text{-}\sigma$). The spectrum is most sensitive to O₃ in the stratospheric bulge layers, but the retrieved abundance does not overlap with the true profile in these layers of the atmosphere.

For the cloudy Earth case (50% clear, 25% stratocumulus, 25% cirrus), we show the Jacobian contributions to the true spectrum in Figure 4.8 for each gas for an atmosphere with 100% stratocumulus clouds and an atmosphere with 100% cirrus clouds to visualize the impact of water and ice clouds on the spectral features. The water Jacobians for each cloud type show the altitudes at which these clouds are distributed in our model (~ 18 km, $\sim 10^4$ Pa). Importantly, we restate that clouds are *not* considered in the retrieval model, but are included in the observed spectrum. For the pressure-independent parameters, we note that the retrieved average surface albedo is twice as bright as the truth ($>20\text{-}\sigma$) due to the presence of clouds in the data and the omission of clouds in the retrieval model. Interestingly, the retrieved CH₄, CO, and CO₂ abundances are all confidently overestimated, and CO has a long tail trailing towards lower abundances. CH₄ is overestimated by a factor of 10, while CO₂ is overestimated by a factor of 50. For CO, the distribution is non-normal and

nearly bimodal, with a well-defined peak and a long tail trailing towards smaller abundances. Though this peak in the CO posterior gives a much larger abundance than the true profile, approximately 40% of the distribution lies *below* abundances of 10^{-5} , which suggests that the retrieval has not confidently ruled out these lower abundances. For CH₄ and CO, cirrus clouds appear to increase the sensitivity of the spectral features to lower altitudes compared to the clear-sky case. For CO₂, both cloud types appear to increase the sensitivity of the spectrum to lower atmospheric layers in the true spectrum. We also note that the tail of the CO₂ posterior distribution has been reduced compared to its clear-sky counterpart. Lastly, the spectrum's sensitivity to O₂ has not changed with altitude due to the addition of clouds, but the median retrieved O₂ abundance (13% VMR) is underestimated by a factor of 2 compared to the true profile.

For the pressure-dependent parameters, the retrieved isothermal temperature distribution is completely unconstrained as in the clear-sky case. Also compared to the clear-sky case, the retrieved H₂O mixing ratio is offset lower from the true surface abundance within a factor of 7, and would be consistent with the true profile in the atmosphere around 0.55 bars (5.5×10^4 Pa, ~ 5 km). The cirrus clouds appear to make the H₂O spectral features sensitive closer to the tropopause, helping to explain the lower retrieved H₂O abundance relative to the clear-sky case.

Finally, the retrieved O₃ abundance is now underestimated by a more significant margin relative to both the true average volume mixing ratio ($<2.5\sigma$) and the true bulge abundance ($<5\sigma$). However, the clouds appear to increase the number of atmospheric layers to which the spectrum is sensitive, with stratocumulus clouds pushing the spectral sensitivity to higher altitudes. We summarize and compare all retrieved values from both our clear-sky and cloudy direct imaging cases to their corresponding truths in Table 4.6.

The $\mathcal{M}_{3P}$ and $\mathcal{M}_{3F}$ vertically-resolved forward models retrieve an O₃ abundance profile with a characteristic stratospheric bulge for both the clear-sky and cloudy Earth. The $\mathcal{M}_{5P}$ forward model is not able to constrain the vertical structure of the profile, effectively

returning an evenly-mixed profile. We show a comparison of the vertically-resolved O_3 profiles for the clear-sky Earth using the $\mathcal{M}_{3P}$ and $\mathcal{M}_{3F}$ forward models in Figure 4.9. In Figure 4.10, we show secondary total column abundance posteriors calculated using the same method described in section 3.1. We find that, in both the clear-sky and cloudy cases, the evenly-mixed model ($\mathcal{M}_{iso}$) shows the largest bias to higher column abundances than the truth when compared to any of the vertically-resolved models ($\mathcal{M}_{3P}$, $\mathcal{M}_{5P}$, $\mathcal{M}_{3F}$). The peaks of the evenly-mixed total column abundance distributions for both the clear-sky and cloudy cases are overestimated by 40% relative to the truth. The vertically-resolved models perform better, producing total column abundance distributions with peaks that slightly underestimate the true value by only 5% in the clear-sky case and 15% in the cloudy case.

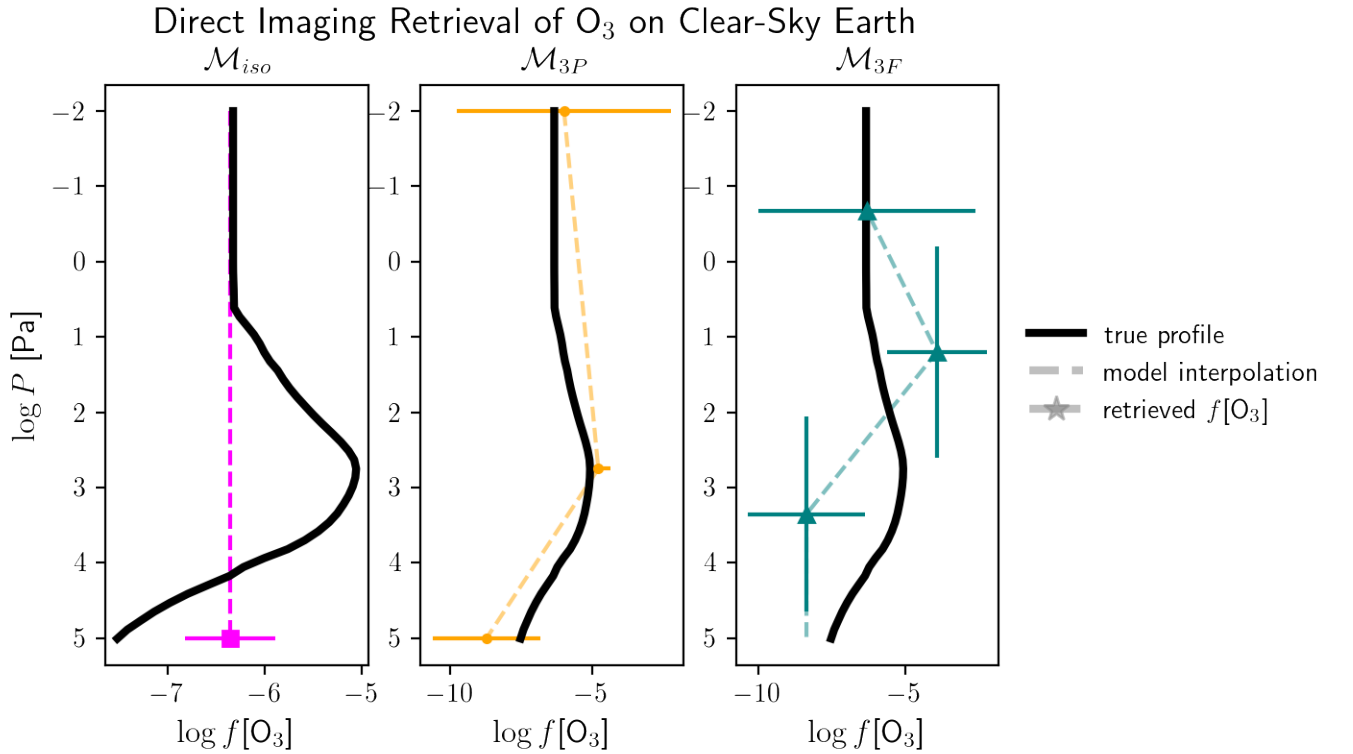


Figure 4.9: A comparison of the vertically-resolved O_3 profiles retrieved by the $\mathcal{M}_{3P}$ and $\mathcal{M}_{3F}$ forward model for the clear-sky Earth. The Bayes factors weakly favor the 3-fixed points model ($\mathcal{M}_{3P}$) over the evenly-mixed case ($\mathcal{M}_{iso}$) and the 3-free points model ($\mathcal{M}_{3F}$), which has more free parameters.

Comparing the evidence for the models in both the clear-sky and cloudy cases, we find

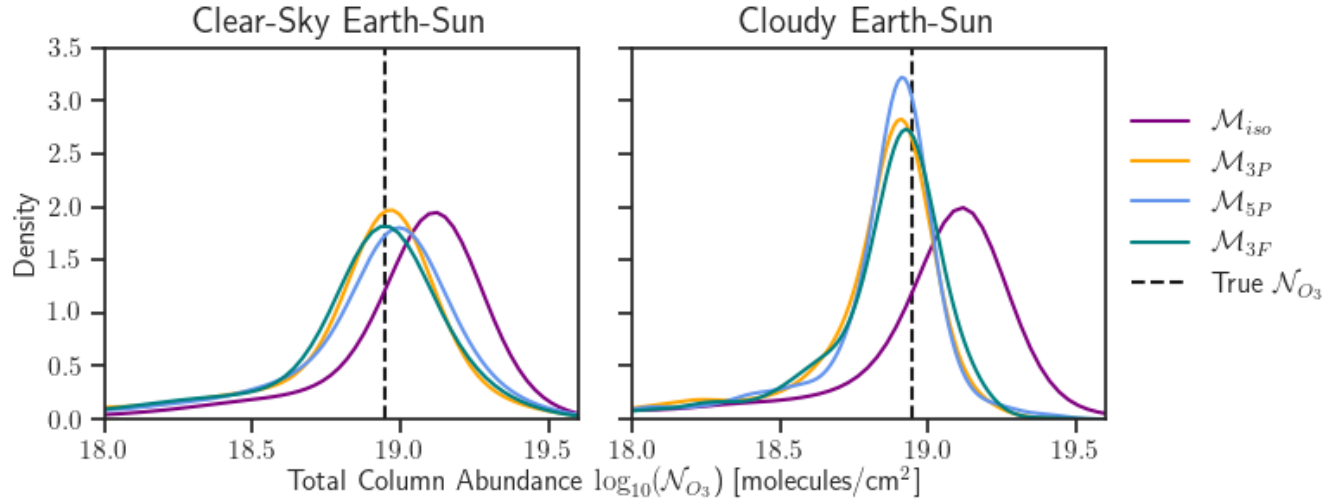


Figure 4.10: A comparison of O_3 total column abundance distributions calculated for retrieval models with and without vertical resolution for both the clear-sky and cloudy direct imaging cases. For both cases, we find that the total column abundance of O_3 is well-constrained using any of the vertically-resolved models, while the evenly-mixed model shows a small bias towards higher column abundances that worsens in the presence of clouds.

that there is weak evidence in favor of the $\mathcal{M}_{3P}$ forward model when compared to $\mathcal{M}_{iso}$ and $\mathcal{M}_{3F}$ ($0.6 < B < 1.7$). All model evidences ($\log \mathcal{Z}$) are reported in Table 4.7. We also compare the forward model fits to the spectrum for both the clear-sky and cloudy Earth cases in Figure 4.11. For the fits to the simulated spectral data of the clear-sky Earth, all four versions of the forward model provide a plausible fit to the spectrum with the given error bars, but the most flexible model with 3 free points ($\mathcal{M}_{3F}$) has the smallest χ^2 of 26.9. In particular, we observe that the $\mathcal{M}_{3F}$ model produces the best fit to the $0.30 \mu\text{m}$ O_3 feature, and produces a better fit to the true O_3 profile than the evenly-mixed model. Like the clear-sky case, the cloudy direct imaging case also shows weak evidence in favor of the 3-fixed points ($\mathcal{M}_{3P}$) model as well as the 5-fixed points ($\mathcal{M}_{5P}$) model, but we note that the fit to the spectrum is visibly poor in the bottoms of the water vapor bands and at the $1.6 \mu\text{m}$ CO/CO_2 feature. All χ^2 values are reported in Table 4.7.

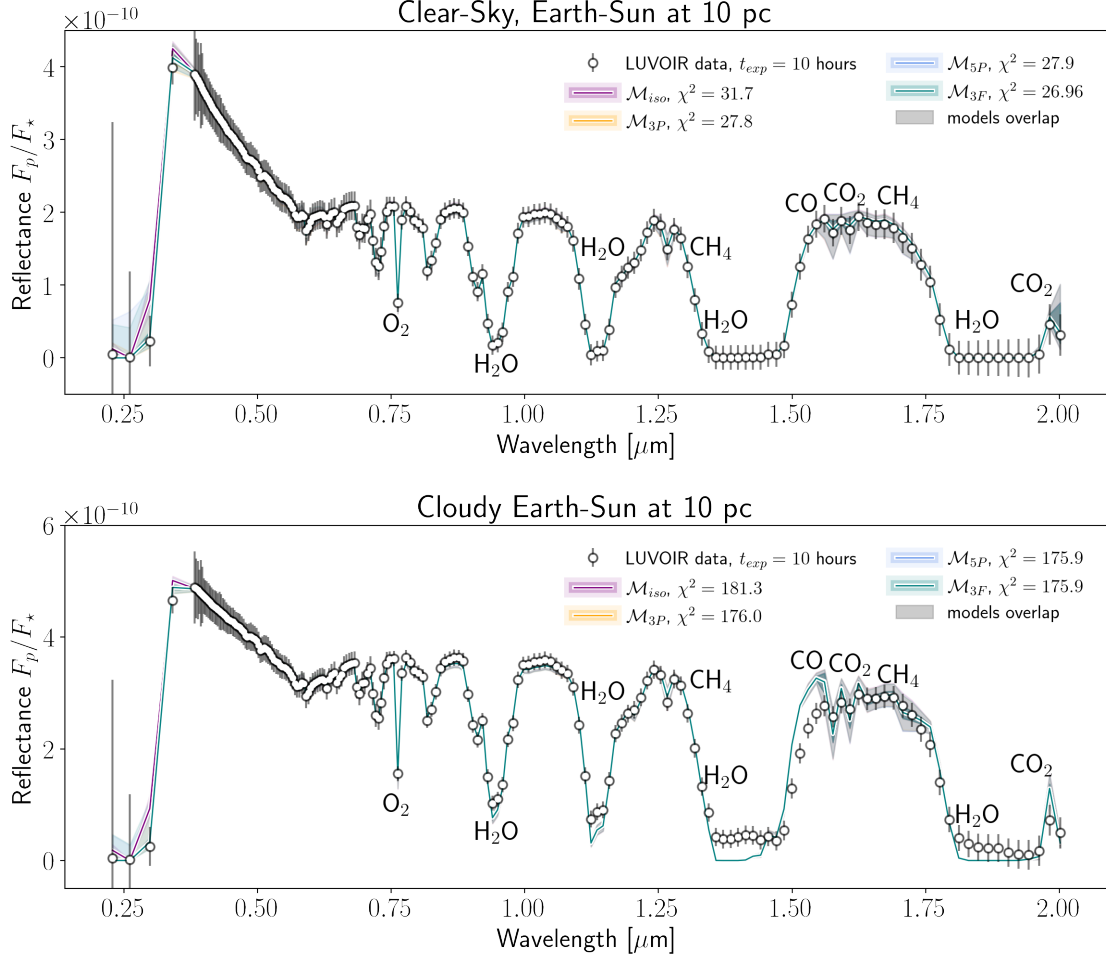


Figure 4.11: A comparison of the median spectral fits with 1- and 3- σ posterior distributions shown as shaded envelopes (dark and pale, respectively) for the clear-sky and cloudy Earth. Note that the lines and envelopes corresponding to the different models overlap significantly and appear gray when they do so. On each plot, we show the fit to the spectrum supplied by each version of the forward model and report its associated χ^2 value. For the clear-sky Earth, the evenly-mixed model visibly provides a better fit to the 0.3 μm feature. In the upper panel, note the larger distribution on the CO posteriors, and that vertically resolving the O_3 profile results in the most significant changes to the fit to the spectrum in the region of the UV Hartley band. In the lower panel, note the large residuals in the water bands and in the CO/ CO_2 features $\sim 1.6 \mu\text{m}$ when wavelength-dependent clouds are present in the data but not included in the retrieval.

4 Discussion

We used a novel terrestrial exoplanet retrieval code to compare the accuracy of retrievals for transmission and direct imaging spectroscopy, to assess both retrieved abundances and our

Table 4.5: Retrieved constraints on Earth-like and abiotic TRAPPIST-1 e transmission spectra generated with the evenly-mixed forward model. As noted in the text, the vertically-resolved models only significantly impacted the retrieved distributions of O₃. The results of the vertically-resolved comparisons are summarized in Table 4.7. For O₃, and CO, which are not evenly-mixed gases in the Earth-like atmosphere, we use the average volume mixing ratio of the true profiles as our true values. Similarly, for H₂O, we designate the isothermal region of the true profile to be our comparison value given the apparent sensitivity of the water absorption features in the transmission spectrum to this region of the atmosphere. Finally, “UL” denotes parameters for which we obtained only an upper limit, and parameters that are not constrained are marked “Unconstrained”.

Parameters	Earth-like T-1 e			Abiotic O ₃ T-1 e		
	Truths	Retrieved	Ratio	Truths	Retrieved	Ratio
$R_p [R_\oplus]$	0.91	$0.920^{+0.002}_{-0.002}$	$1.010^{+0.002}_{-0.002}$	0.91	$0.9200^{+0.0009}_{-0.0020}$	$1.0000^{+0.0009}_{-0.0020}$
$P_0 [\text{Pa}]$	9.8×10^4	$8^{+33}_{-7} \times 10^4$	$0.8^{+3.4}_{-0.7}$	1.2×10^5	$1.0^{+3.0}_{-0.7} \times 10^5$	$0.9^{+3.0}_{-0.6}$
$T_0 [\text{K}]$	284	202^{+34}_{-30}	$0.7^{+0.1}_{-0.1}$	263	194^{+30}_{-25}	$0.74^{+0.01}_{-0.09}$
$f[\text{H}_2\text{O}]$	158 ppm	32^{+330}_{-30} ppm	$0.2^{+2.0}_{-0.2}$	5 ppb	7^{+160} ppb (UL)	$1.3^{+31.0}$ (UL)
$f[\text{CH}_4]$	696 ppm	403^{+1900}_{-360} ppm	$0.6^{+3.0}_{-0.5}$	0	16^{+55} ppb (UL)	$160000^{+600000}$ (UL)
$f[\text{CO}_2]$	0.10	$0.07^{+0.30}_{-0.06}$	$0.7^{+3.0}_{-0.6}$	0.63	$0.4^{+0.3}_{-0.2}$	$0.6^{+0.5}_{-0.3}$
$f[\text{O}_2]$	0.20	Unconstrained	-	0.12	Unconstrained	-
$f[\text{O}_3]$	6 ppm	$0.84^{+13.00}_{-0.80}$ ppm	$0.2^{+2.0}_{-0.2}$	7.3 ppm	$0.73^{+5.00}_{-0.72}$ ppm	$0.100^{+0.700}_{-0.098}$
$f[\text{CO}]$	606 ppm	$12.4^{+650.0}_{-12.0}$ ppm	$0.02^{+1.00}_{-0.02}$	0.24	$0.064^{+0.200}_{-0.054}$	$0.3^{+0.8}_{-0.2}$

Table 4.6: Retrieved constraints on clear-sky and cloudy Earth-twin reflected light spectra generated with the evenly-mixed forward model. As noted in the text, the vertically-resolved models only significantly impacted the retrieved distributions of O₃. The results of the vertically-resolved comparisons are summarized in Table 4.7. For H₂O, we use the surface abundance of water in the true atmosphere as our comparison value. For O₃, we use the bulge abundance in the true atmosphere as our comparison value given the apparent sensitivity of the ozone absorption features in the spectrum to this region of the atmosphere. As in Table 4.5, “UL” denotes parameters for which we obtained only an upper limit, and parameters that are not constrained are marked “Unconstrained”.

Parameters	Clear-sky Earth			Cloudy Earth		
	Truths	Retrieved	Ratio	Truths	Retrieved	Ratio
A_s	0.2	$0.200^{+0.003}_{-0.003}$	$1.0^{+0.01}_{-0.01}$	0.2	$0.350^{+0.003}_{-0.003}$	$1.75^{+0.02}_{-0.02}$
$T_0 [\text{K}]$	294	Unconstrained	-	294	Unconstrained	-
$f[\text{H}_2\text{O}]$	0.02	$0.0070^{+0.0006}_{-0.0005}$	$0.95^{+0.08}_{-0.08}$	0.02	$0.0030^{+0.0002}_{-0.0002}$	$0.35^{+0.02}_{-0.02}$
$f[\text{CH}_4]$	1.50 ppm	6^{+1300} ppb (UL)	0.004^{+900}	1.50 ppm	14^{+4}_{-3} ppm	9^{+2}_{-2}
$f[\text{O}_2]$	0.21	$0.21^{+0.05}_{-0.05}$	$0.99^{+0.2}_{-0.2}$	0.21	$0.13^{+0.02}_{-0.02}$	$0.6^{+0.1}_{-0.1}$
$f[\text{CO}]$	90 ppb	30^{+60000} ppb (UL)	0.32^{+609}	90 ppb	528^{+2000}_{-528} ppm	5600^{+22000}_{-5600}
$f[\text{O}_3]$	4.2 ppm	$0.50^{+0.07}_{-0.06}$ ppm	$0.11^{+0.02}_{-0.02}$	4.2 ppm	$0.40^{+0.04}_{-0.04}$ ppm	$0.094^{+0.009}_{-0.009}$
$f[\text{CO}_2]$	330 ppm	187^{+302}_{-170} ppm	$0.57^{+0.92}_{-0.52}$	330 ppm	2000^{+500}_{-500} ppm	$5.0^{+1.6}_{-1.4}$

ability to infer the characteristics of an exoplanet’s surface environment. In transmitted light, we performed this comparison for both a modern Earth-like and an abiotic TRAPPIST-1 e

Table 4.7: Summary of the results of the retrieval tests for our observing cases and their atmospheric subcases with evenly-mixed ($\mathcal{M}_{iso}$) and vertically-resolved ($\mathcal{M}_{3P}$, $\mathcal{M}_{5P}$, $\mathcal{M}_{3F}$) forward models. The model that minimizes χ^2 provides the best fit to the data. The model with the largest evidence ($\log \mathcal{Z}$) is the preferred model, but our comparison of the effectiveness of the models depends on the significance (Jeffreys, 1998) of the difference in their evidences compared to $\mathcal{M}_{iso}$, shown in the Bayes Factor column of the table below. Here, Bayes factors $|B| < 1$ are considered to provide inconclusive evidence, while negative Bayes factors $-2 < B \leq -1$ suggest weak evidence in favor of $\mathcal{M}_{iso}$, and positive Bayes factors $1 \leq B < 2$ suggest weak evidence in favor of a given vertically-resolved forward model. Dashes are reported in this column for $\mathcal{M}_{iso}$ in each case given this is the model used for comparison. For both the Earth-like TRAPPIST-1 e transmission case and the false-positive O_3 transmission case, we find that there is either no significant difference in the model evidence between $\mathcal{M}_{iso}$ and $\mathcal{M}_{3P}$ and $\mathcal{M}_{5P}$, and that $\mathcal{M}_{iso}$ is weakly preferred when compared to $\mathcal{M}_{3F}$. For the clear-sky Earth-twin direct imaging case, we find weak evidence in favor of the 3-fixed points model ($\mathcal{M}_{3P}$). Like the clear-sky case, the cloudy direct imaging case also shows weak evidence in favor of the 3-fixed points ($\mathcal{M}_{3P}$) as well as the 5-fixed points ($\mathcal{M}_{5P}$) models, but we note that the fit to the spectrum is visibly poor in the bottoms of the water vapor bands and at $1.6 \mu\text{m}$.

Observing Method	Atmosphere	Model ($\mathcal{M}$)	# Params. (N)	Goodness of Fit (χ^2)	Evidence ($\log \mathcal{Z}$)	Bayes Factor (B)
Transmission	Earth-like T-1 e	$\mathcal{M}_{iso}$	9	16.9	-15.9 ± 0.1	-
		$\mathcal{M}_{3P}$	11	16.4	-15.7 ± 0.3	0.2
		$\mathcal{M}_{5P}$	13	16.3	-16.3 ± 0.1	-0.4
		$\mathcal{M}_{3F}$	14	14.0	-17.7 ± 0.0	-1.8
	Abiotic O_3 T-1 e	$\mathcal{M}_{iso}$	9	21.8	-16.5 ± 0.0	-
		$\mathcal{M}_{3P}$	11	19.6	-15.9 ± 0.1	0.6
		$\mathcal{M}_{5P}$	13	19.5	-17.0 ± 0.1	-0.5
		$\mathcal{M}_{3F}$	14	19.3	-18.0 ± 0.0	-1.5
Direct Imaging	Clear-sky Earth	$\mathcal{M}_{iso}$	8	31.7	-25.5 ± 0.1	-
		$\mathcal{M}_{3P}$	10	27.8	-24.2 ± 0.1	1.3
		$\mathcal{M}_{5P}$	12	27.9	-24.8 ± 0.1	0.7
		$\mathcal{M}_{3F}$	13	26.9	-25.9 ± 0.1	-0.4
	Cloudy Earth	$\mathcal{M}_{iso}$	8	181.3	-120.2 ± 0.0	-
		$\mathcal{M}_{3P}$	10	176.0	-118.3 ± 0.2	1.9
		$\mathcal{M}_{5P}$	12	175.9	-118.4 ± 0.1	1.8
		$\mathcal{M}_{3F}$	13	175.9	-119.8 ± 0.1	0.4

with false positive O_3 . In reflected light, we compared a clear-sky and cloudy Earth-twin at 10 pc.

For the cases considered, we found that we were able to detect and constrain the abundances of numerous key species. For transmission observations with exposure time $t = 40$ hours for an *Origins*-like mid-IR probe, which is assumed to be nearly twice as sensitive as

JWST according to the instrument parameters described by [Meixner et al. \(2019\)](#) over the wavelength range $2.8 - 20.0 \mu\text{m}$, our model detects and retrieves 20% of the true H_2O , 60% of the true CH_4 , and 20% of the true O_3 abundances for the modern Earth-like atmosphere. For the abiotic false positive planet, we detect and retrieve 130% of the true H_2O , 10% of the true O_3 , and a useful upper limit on the CH_4 . We also accurately constrain CO_2 abundances for the inhabited planet to 70% of the true profile, and retrieve 60% of the true CO_2 abundance for the abiotic planet. For the modern Earth-like TRAPPIST-1 e observed in transmission, we obtained good fits to the simulated data, but found that our retrieved water abundance coincides with the true stratospheric abundance of this planet rather than the surface abundance. For this case, we also retrieved a median temperature of 202 K, which, given our calculations of transmission’s sensitivity to that region, is likely to be that of stratospheric altitudes between 50 and 70 km.

For our direct imaging cases, we found that the retrieval produced relatively high accuracy constraints on O_3 with inferred abundances within 20% of the true value, and a near-surface abundance of H_2O in a clear-sky, Earth-twin atmosphere retrieved to within 10%. Similarly, O_2 is constrained with $1-\sigma$ errors of $\pm 5\%$ VMR for the clear-sky case, and within a factor of 2 with clouds. Additionally, when we implement a vertically-resolved forward model in these cases, we obtain weak sensitivity to the presence of a stratospheric O_3 bulge for both the clear-sky and cloudy Earth based on improved fits to the Hartley band in the ultraviolet. In addition, our vertically resolved model more accurately retrieves the true averaged column abundance compared to the evenly mixed assumption. The evenly-mixed model retrieves the true value within 40% errors, while the vertically resolved models retrieve the true value within $\sim 10\%$ errors in the clear-sky case and within $\sim 20\%$ errors in the cloudy case ([4.10](#)). For the clear-sky Earth, our inferred abundances fit the reflected light observations over the entire wavelength range within $1-\sigma$ error bars.

4.1 Comparison to Previous Work

There are benchmark retrieval studies in the literature that we can compare to, although subtle differences in the synthetic observations and retrieval forward models limit direct comparisons. The retrieval results we present for the photochemically self-consistent clear-sky, Earth-like TRAPPIST-1 e in transmitted light may be compared to the transmission retrieval of Tremblay et al. (2020), for their non-photochemically-self-consistent Earth orbiting TRAPPIST-1, which has the gas mixing ratios and atmospheric structure of the true Earth. Consequently, although similar, the atmospheric compositions and vertical distributions for the two cases are not identical. Similarly, the results we present for the cloudy Earth-twin in reflected light may be compared to the Earth twin planets modeled by Feng et al. (2018) and Damiano & Hu (2022). In all three of these studies, the authors use a forward model to produce synthetic data by assuming evenly-mixed gas and isothermal temperature profiles (except in the case of H_2O , which is the condensing gas, in Damiano & Hu [2022]). All three studies use the same forward model to generate their data and perform the retrieval, in which cloud properties are also included as free parameters. In particular, Damiano & Hu (2022) implement a cloud parameterization scheme that vertically resolves the profile of the condensing gas at the top/bottom of the cloud deck and at the planet surface, allowing exploration of cloud behavior and cloud-gas absorption degeneracies. All three studies show potential degeneracies between cloud parameters, surface albedo, and gas absorption.

In comparison, our study used the vertically-resolved atmospheric structure of a 1-D-modeled planet to generate the synthetic data, allowing us to explore the biases introduced by using a simplified forward model to fit spectra of complex planetary environments. Our study lacks a cloud parameterization for the retrieval model, but instead includes additional free parameters to vertically resolve non-condensing, photochemically-active gases. Our study shows the biases that are introduced when a cloud-free model is used to fit a cloudy

atmosphere, and how vertical parameterization impacts the retrieval and interpretation of non-condensing, photochemically-active gases.

Comparing our results to the retrieval of synthetic transmission spectra at $R=100$ and $\lambda = 3.0 - 30.0 \text{ } [\mu\text{m}]$ for Earth-like planets orbiting TRAPPIST-1 from Tremblay et al. (2020), we more precisely constrain H_2O and CH_4 , while Tremblay et al. (2020) put more precise constraints on CO_2 and O_3 . However, this difference in retrieval sensitivity is most likely due to compositional differences in our model atmospheres. Tremblay et al. (2020) did not generate synthetic data for a climatically and photochemically self-consistent Earth-like planet around an M dwarf, and our photochemical models (Figure 4.2) show that the spectrum of the M dwarf star would have suppressed O_3 and enhanced CH_4 abundance (Segura et al., 2005) compared to modern Earth. Consequently, Tremblay et al. (2020) find an O_3 column abundance that is enhanced compared to ours, while their CH_4 is significantly reduced. Furthermore, their paper includes data generated with more transits than this study, and they scale TRAPPIST-1’s stellar spectrum to a more optimistic K -band magnitude of 8 (100 times brighter than TRAPPIST-1’s observed brightness) to mimic the average brightness of nearby M dwarf stars. This also serves to increase the SNR of the spectrum particularly at the longer wavelengths where prominent O_3 and CO_2 absorption features appear.

In contrast, the synthetic observations in reflected light all assume a planetary atmospheric composition close to true Earth and so provide a better comparison with our results. In particular, the $S/N=15$, $R=140$ simulations of Feng et al. (2018) and the $S/N=10$, $R=140$ simulations of Damiano & Hu (2022) make a better comparison with our $S/N=15$, $R=100$ cloudy Earth-twin retrieval results. For H_2O in the cloudy case (50% clear/50% cloudy), we are able to more precisely constrain the near-surface water abundance (7% errors) than both Feng et al. (2018) (100% – 200% errors) and Damiano & Hu (2022) (200 – 700% errors), and we appear to retrieve a near-surface tropospheric water abundance consistent with the true profile at 5 km ($5.6 \times 10^4 \text{ Pa}$), between the surface and cloud deck. Similar to Feng et al. (2018) and Damiano & Hu (2022), O_2 and O_3 are well constrained, as we

retrieve 60% of the true O_2 and 10% of the O_3 bulge abundance, and we are able to rule out abundances lower than 9% for O_2 and abundances lower than 0.36 ppm for O_3 . However, unmodeled, wavelength-dependent cloud absorption (Robinson et al., 2011) causes our retrieval to overestimate the abundance of CH_4 and CO_2 as it increases their VMRs in an attempt to match the additional cloud absorption. By contrast, Damiano & Hu (2022) obtain more accurate but highly imprecise retrieved abundances and Feng et al. (2018) did not model or retrieve these gases in their study. We note that this inaccuracy of our retrieval model will likely be remedied by the inclusion of cloud parameters in the retrieval forward model, which we intend to explore in future studies.

4.2 Comparison of Transmission and Direct Imaging Capabilities

Our results have broader implications for the differing capabilities of transmission and direct imaging spectroscopy to extract accurate information regarding a terrestrial planet’s environmental properties. For the two observing modes below we compare the sensitivity to cloud absorption, the ability to assess the surface habitability of an exoplanet, and the interpretation of biosignature pairs and their discriminants in the context of the planetary environment.

4.2.1 Sensitivity to Clouds

Clouds are likely to be a significant feature of the planetary atmospheres explored here, whether they are impediments to observables (e.g., Lincowski et al., 2018; Fauchez et al., 2019; Suissa et al., 2020; Komacek et al., 2020) or observables themselves (Kreidberg et al., 2014; Wakeford & Sing, 2015; Arney et al., 2018; Benneke et al., 2019). While the SMART forward model is capable of modeling aerosols, and we use this capability in creating the simulated observations, the aerosols must be rigorously defined in terms of particle size distributions and wavelength-dependent optical properties (Robinson et al., 2011), which are typically much more complex than the simpler parameterizations used in existing retrieval

models (e.g., [Feng et al., 2018](#); [Tremblay et al., 2020](#); [Damiano & Hu, 2022](#)), and therefore computationally expensive to use in a retrieval model. As a result, we focus our study on the outcomes we obtain when realistic aerosols are present in exoplanet measurements but not fitted for and retrieved on.

We saw relatively little impact between clear and cloudy spectra in our *Origins*-like mid-IR transmission simulations of an Earth-like TRAPPIST-1 e with an exposure time of 40 hours. Similar results were found in the retrieval study of [May et al. \(2021\)](#), who showed that water ice cloud variability did not affect retrieved abundances at detectable levels for synthetic observations of TRAPPIST-1 e with JWST’s NIRSpec PRISM. This is despite previous evidence that clouds can potentially truncate transmission spectra in observations of hotter exoplanets ([Kreidberg et al., 2014](#)) and models of Earth-like habitable-zone terrestrial planets ([Fauchez et al., 2019](#); [Suisa et al., 2020](#); [Komacek et al., 2020](#); [Lincowski et al., 2018](#)), and can also affect the retrieval of atmospheric properties ([Line & Parmentier, 2016](#)).

Our lack of cloud impact in the transmission case compared to previous results is due to two factors: 1) primarily that our photochemically-self-consistent TRAPPIST-1 e atmosphere contains higher abundances of both CH₄ and CO₂ than the true Earth cases explored previously (e.g., [Fauchez et al., 2019](#)), thereby providing greater atmospheric opacity with a pseudo-continuum of absorption and 2) to a lesser extent, that the two studies take different approaches to modeling the ice and water clouds included in the modern Earth-like atmospheres. Our inferred lack of spectral truncation by clouds is consistent with the cloud altitudes—ice clouds at 300 mbar (3×10^4 Pa) and water clouds at 900 mbar (9×10^4 Pa)—which are below the regions of the atmosphere where the retrieval is sensitive (c.f. Figure 4.4). Although it has been suggested that 3-D GCMs may form clouds higher than predicted by 1-D climate models ([Fauchez et al., 2019](#)), clouds in our model are in fact located a few kilometers higher at the terminator, where transmission probes, compared to [Fauchez et al. \(2019\)](#). Instead, truncation of the spectrum occurs because our higher trace gas abundances produce stronger bands with extended wings (especially CO₂ and CH₄) that form a

pseudo-continuum of opacity for most wavelengths $>2\ \mu\text{m}$. This can be seen in Figure 4.3 in this work and Figure 11 in Meadows et al. (2023), when compared to the atmospheric spectra with lower CH_4 and CO_2 abundances in Figure 11 in Lincowski et al. (2018), and Figure 5a in Fauchez et al. (2019). The additional spectrally-broad opacity from these higher gas abundances raises the minimum atmospheric level probed for even the clear-sky case, thereby minimizing the difference between the atmospheric layers probed in the clear and cloudy cases, and effectively reducing the impact of the addition of clouds. The effect of this additional opacity can explain much, but not all, of the reduced spectral truncation due to clouds found in this study relative to Fauchez et al. (2019); cloud modeling differences are therefore also needed to fully explain this spectral truncation gap. Even though our cloud altitudes are similar, clouds may produce a second order effect on the minimum transmission altitude between these two sets of results, likely due to differences in cloud optical depths and particle distributions in the two models.

For the clear and cloudy direct imaging comparison, clouds do truncate the probed atmospheric region, and we retrieve 40% less O_2 , 20% less O_3 , and 60% less near-surface water abundance in the cloudy case when compared to the clear sky case. We also find that the presence of wavelength-dependent cloud absorption in the data (especially near $1.6\ \mu\text{m}$) causes us to overestimate the abundances of CO_2 , CH_4 , and CO , which all overlap these cloud absorption features. This overestimate occurs because the model currently cannot fit for cloud-induced, wavelength-dependent reflectivity changes, and so attempts to compensate by enhancing the wavelength-independent surface albedo as well as increasing or decreasing the inferred abundances of gases with overlapping absorption features to better fit the spectrum.

However, the surface albedo can only affect the spectral continuum, and not the bottoms of saturated absorption bands where photons do not reach the surface, while true clouds, which are partially gray scatterers and higher in the atmosphere, may increase the reflectivity in the bottoms of the bands. While the near-surface water abundance we retrieve assuming

evenly-mixed vertical profiles provides a reasonable fit to the shape of the water bands, it will overestimate the water abundance in the cloud-truncated regions of the atmosphere, and this manifests as a much deeper core of the band than seen in the synthetic data. We find that, when assuming evenly-mixed profiles, the near-surface abundance we retrieve provides a good fit to the overall shapes of the water bands, but that only the cloud-top abundance can provide a satisfactory fit to the core of the bands. One such data-model mismatch is particularly noticeable in the poor fit to the H_2O bands $>1\ \mu\text{m}$ (c.f. Figure 4.11). We can quantify the size of the large residuals in the bottoms of these near-IR H_2O bands by calculating the quadrature sum of the S/N at each wavelength of a given absorption feature. Using this method, we find that the residuals in the near-IR H_2O bands have S/N as large as 6.5. Another visible data-model mismatch occurs in the CO band at $\sim 1.55 - 1.6\ \mu\text{m}$ (c.f. Figure 4.11). The CO, CO_2 , and CH_4 absorption features near $1.6\ \mu\text{m}$ overlap with a large ice cloud absorption feature from $1.5 - 1.6\ \mu\text{m}$, and since the ice cloud feature is not present in the model, this results in an overestimate of the inferred abundances of these gases.

4.2.2 Habitability Assessment

Constraining surface temperature, surface pressure, the presence and abundance of water vapor (Robinson, 2018; Meadows et al., 2018a), and the abundance of atmospheric CO_2 (Catling et al., 2018) are key environmental parameters needed to constrain a climate model and thereby assess the likelihood that a terrestrial planet is habitable. Atmospheric temperature and pressure are fundamental properties that can reveal the structure of the atmosphere, and their surface values can help determine whether liquid water can be sustained on the planet’s surface. The abundance of greenhouse gases, such as water vapor and CO_2 , especially when combined with any knowledge of atmospheric pressure and temperature, provide further constraints for climate modeling of the surface temperature. The detection of atmospheric water vapor near the surface of the planet can also help assess climate, and may also increase the likelihood that surface liquid water is present. Below we describe and compare

the strengths and weaknesses of the transmission or direct imaging technique for obtaining the information needed to constrain habitability assessment via a climate model.

Characterizing habitability using transmission observations—Transmission observations are not sensitive to the planetary surface (Lincowski et al., 2018), and our results did not provide informative constraints on planetary surface temperature, or near-surface water abundance for both the Earth-like planet and the abiotic planet studied here. Furthermore, given our isothermal temperature model for the Earth-like and abiotic planets, our $1\text{-}\sigma$ retrieved pressure ranges of $0.1 - 4.1$ bars ($1 \times 10^4 - 4.1 \times 10^5$ Pa) and $0.3 - 4$ bars ($3 \times 10^4 - 4 \times 10^5$ Pa) translate to altitudes at or below ~ 23 km and ~ 11 km, respectively. The high-pressure ~ 4 bar upper bound for both planets indicates an insensitivity to how much additional atmosphere may lie below the region probed by the transmission spectrum. The retrieved median isothermal temperature agrees well with the stratosphere ($\sim 10^2$ Pa, $50 - 70$ km) of the Earth-like planet within 10 K, and that of the abiotic planet within several K, although both temperatures are below the freezing point of water. We were also not able to retrieve precise constraints on the atmospheric CO_2 abundances for either planet, making it challenging to model the planetary climate.

We retrieve median pressures for both the living and abiotic TRAPPIST-1 e that are near that of the true surface pressure, but in an agnostic setting it will be difficult to know whether we were truly seeing the surface of the planet or the top of a cloud deck given the nature of transmission observations (Benneke & Seager, 2012; Lustig-Yaeger et al., 2019b). In transmitted light, the retrieved maximum pressure corresponds to the spectral continuum, which probes to lower altitudes than molecular absorption. However, for a habitable planet this level is still likely to be above the surface, due to the intrinsic opacity of molecules in the atmosphere (Meadows et al. (2018a) and Figure 4.4) as well as clouds and hazes (Fortney, 2005; Kreidberg et al., 2014; Arney et al., 2017). Depending on the host star and planetary distance, refraction could also limit access to the near-surface atmosphere (Misra et al., 2014b; B  tr  mieux & Kaltenegger, 2014; Lincowski et al., 2018; Lustig-Yaeger et al.,

2019a).

For the spectrum of the clear-sky Earth-like TRAPPIST-1 e (which is indistinguishable from the cloudy case at almost all wavelengths, c.f., Figure 4.3) we retrieved a median pressure from the continuum that is consistent with the true atmosphere’s ~ 2 km altitude pressure—which is 80% of the true surface pressure (± 1 dex) for the Earth-like planet. For the cloud-free abiotic planet, we retrieve a median continuum pressure consistent with 1 km altitude, or 90% of the true surface pressure. These differences in median retrieved pressure reflect differences in scale height between the two atmospheres. Although these median retrieved pressures are close to the true surface values, the retrieved distribution suggests a broader range of pressures, and without knowledge of the ground truth, we would not be able to rule out that we had probed to a cloud-top, rather than close to the surface, without additional information. However, the relative sensitivity of transmission spectra to stratospheric temperature in combination with tighter constraints on the near-surface pressure may potentially be used with a climate model to infer the atmospheric temperature structure if a cold trap is present (Robinson & Catling, 2014).

CO₂ is a key greenhouse gas, and while the peaks of our retrieved distributions for stratospheric CO₂ are relatively accurate (70% and 60% of the true value for the Earth-like and abiotic planets, respectively), we retrieve imprecise distributions spanning ~ 3 orders of magnitude for both cases. Since CO₂ is commonly evenly mixed in terrestrial planetary atmospheres (Hu et al., 2012), the stratospheric abundance we retrieve is likely a good proxy for the abundance throughout the atmosphere, which can help constrain surface temperature. However, the broad distributions we retrieve correspond to a $1\text{-}\sigma$ range of CO₂ abundances of approximately 1%-40% for the Earth-like planet and 1%-70% of the abiotic planet. Given that TRAPPIST-1 e has $\sim 60\%$ Earth’s insolation (Gillon et al., 2017), this broad range of CO₂ abundances could give both habitable and uninhabitable scenarios assuming Earth-like fluxes (Turbet et al., 2020; Meadows et al., 2023). This insensitivity to CO₂ abundance is in part due to the strong absorption of CO₂, which produces spectral features that saturate at

low abundances. Additionally, CO₂ absorption in the MIR is a product of both abundance and atmospheric thermal structure, and an assumed isothermal temperature structure could compromise the ability to fit CO₂ features in the spectrum. The correlation between temperature and thermally-dependent CO₂ absorption features (e.g., the 5.3 μm CO₂ feature) could be used in future retrievals to constrain vertical temperature structure for uninhabitable planets with strong CO₂ absorption. Additionally, future work could propagate broad posteriors through a climate model to fully depict the range of planetary environments that are possible given the fits to the observed spectra.

Transmission is also sensitive to stratospheric water vapor, and will allow us to use order of magnitude constraints to identify or rule out extremely desiccated planets, but will likely not allow us to discriminate between Venus-like and Earth-like environments using the retrieved water vapor abundance. For the Earth-like TRAPPIST-1 e planet, we retrieve peak stratospheric H₂O abundances of 32 ppm, with a 1- σ distribution spread of 2 – 362 ppm, corresponding to altitudes of 17 – 88 km in the true atmosphere, to which the spectrum is indeed sensitive (c.f., Figure 4.4). In comparison, measurements at the terminator of Venus’s night-side estimated the water vapor abundance to be 0.56 – 2.45 ppm (Chamberlain et al., 2020), while the stratospheric water vapor abundance of the true Earth is $\sim$ 6 ppm. Thus, our broad distribution for the Earth-like planet overlaps with both a Venus-like and an Earth-like scenario. For the desiccated planet corresponding to the most extreme case modeled by Gao et al. (2015), we retrieve a 1- σ upper limit on the H₂O abundance on the order of 167 ppb, which is approximately 12 times smaller than Venus’ maximal stratospheric water abundance. Although the 2- and 3- σ upper limits are higher abundances of 5 and 13 ppm, these stricter upper limits are still orders of magnitude smaller than the stratospheric water vapor abundances (130-302 ppm, c.f., Figure 4.2) that we predict for the self-consistent Earth-like TRAPPIST-1 e. While this upper limit would enable us to distinguish such an extremely dry planet from an Earth-like planet around an M-dwarf, planets that are less desiccated may be more ambiguous.

Characterizing habitability using direct imaging observations — Despite relative insensitivity to temperature at the spectral resolutions considered, our results suggest that direct imaging observations are remarkably sensitive to the near-surface water vapor abundance of a 50% cloudy Earth-twin at 10 pc, and can confidently constrain near-surface water abundances to within 0.1% (at $1\text{-}\sigma$).

In contrast to the transmission cases, we were unable to constrain temperature in any part of the atmosphere at the S/N and spectral resolutions considered. The insensitivity of direct imaging to temperature poses a challenge to future habitability assessments, as it is a key climate modeling parameter, and should be investigated in greater detail. Although direct imaging has been shown to be sensitive to surface pressure (Feng et al., 2018; Damiano & Hu, 2022), we did not fit for surface pressure for the direct imaging cases, and will explore this in future work. A number of techniques have been proposed to determine surface pressure and temperature, including dimer strengths compared to rho-vibrational or electronic states (Misra et al., 2014a), broadening of individual molecular bands (Pallé et al., 2009; Schwinterman et al., 2015), and fitting the Rayleigh scattering slope (Benneke & Seager, 2012). Recent work also considers how we may be able to constrain global temperature from the thermally-dependent behavior of water vapor absorption (Young et al., 2024). However, all of the inference techniques listed above will be impacted by truncation of the atmospheric path length due to clouds (e.g., Earth; Misra et al., 2014a) and haze (e.g., Archean Earth or Venus; Arney et al., 2016; Crisp & Titov, 1997). In addition, absorption from the planetary surface, or clouds and hazes, can also obliterate the pressure signal (Arney et al., 2017).

Planetary albedo could help identify the presence of clouds or help discriminate between ocean and ice surfaces, and although we obtained compelling albedo constraints by assuming strong prior knowledge about the target, multiple factors complicate albedo retrievals. Our cloud-free case yielded a robust albedo constraint (0.200 ± 0.003), but we retrieved a planetary albedo that is 1.8 times the value of the 50% cloud-cover Earth, as another effect of the surface albedo being used by the retrieval to account for clouds. These results assumed tight

priors constraints on planetary radius and phase, which are both degenerate with albedo (Nayak et al., 2017; Guimond & Cowan, 2018), and therefore propagate greater uncertainty into the retrieved surface albedo if unknown (Feng et al., 2018; Carrión-González et al., 2020). This effect has long been appreciated for observations of small bodies in the solar system (Davies et al., 2009), and poses a challenge for future exoplanet direct imaging (Fujii et al., 2017), especially since uncertainties in the radius may propagate to larger errors in surface gravity, surface pressure, and molecular abundances (Feng et al., 2018).

We find that direct imaging observations cannot confidently retrieve CO₂ in the atmosphere of a clear-sky Earth-twin. We retrieve a 1- σ distribution of abundances spanning 17 – 489 ppm with a long tail trailing to longer abundances. We conservatively interpret this distribution as providing a 3- σ upper limit of 91% VMR on the retrieved abundance of CO₂ in the clear-sky case. In the atmosphere of the Earth-twin with 50% cloud coverage, although we detect CO₂ more confidently, we greatly overestimate the abundance of CO₂, and this is again due to the impact of the synthetic data including cloud extinction, which is not in the retrieval forward model. The peak of the retrieved CO₂ distribution is consistent with 5 times more carbon dioxide than is present in the true atmosphere (330 ppm), with a 1- σ distribution spanning 1500 – 2500 ppm. As described above in Section 4.2.1, this bias towards much higher abundances of CO₂ occurs due to overlapping ice cloud and CO₂ absorption features at $\sim 1.6 \mu\text{m}$ and $\sim 1.95 \mu\text{m}$ in the synthetic data. Overestimating the CO₂ abundance in this atmosphere as 3 – 7 times the true value would imply a much stronger greenhouse effect than on the true Earth, and depending on other planetary characteristics, may increase the probability that an environment is incompatible with life, even if the inferred higher abundance does not lead to a runaway greenhouse (Goldblatt et al., 2013). While we retrieved a more accurate CO₂ abundance for the Earth-like TRAPPIST-1 e, we found a broad distribution of abundances that may also be compatible with uninhabitable scenarios, suggesting that both transmission and direct imaging may provide ambiguous assessments of CO₂ for Earth-like planets.

Finally, while our transmission retrievals were not sensitive to near-surface water abundances, our direct imaging retrievals returned 35% of the true surface water abundance for the cloudy Earth-twin with a $1\text{-}\sigma$ distribution spanning abundances of $0.28 - 0.32\%$. Though the retrieved distribution is 65% lower than the true surface abundance (consistent with the true atmosphere’s abundance at ~ 5 km or $\sim 10^4$ Pa, c.f. Figure 4.8), the narrow range of abundances would confidently distinguish this planet from the few ppm expected above planet-wide clouds for a Venus-like environment (Chamberlain et al., 2020). This result complements previous work that shows that reflectance spectra could also be sensitive to surface liquid water via ocean glint at visible and near-IR wavelengths (Williams & Gaidos, 2008; Robinson et al., 2010; Lustig-Yaeger et al., 2018).

However, our retrieved distribution for water vapor will have been impacted by the presence of cloud extinction in the synthetic data that was excluded in the retrieval forward model. As described above in Section 4.2.1, because the model must assume that water vapor is evenly-mixed in the atmosphere, it overestimates the water vapor abundance relative to the distribution of the cloud decks, but underestimates it relative to the surface. It is unclear whether the lower abundance we retrieve is due primarily to the significantly lower water vapor abundance at the cloud-tops, or the bias due to the lack of clouds in the retrieval. However, the inclusion of clouds in future retrievals would likely mitigate the bias towards lower water vapor abundances by allowing the model to correctly infer the near-surface water abundance from the width of the bands, and the cloud-top pressure from the band depths. We note that vertically parameterizing the water vapor abundance in the retrieval model (Damiano & Hu, 2022) may further constrain this problem.

4.2.3 Biosignature Pairs and Discriminants

Below we describe in detail what we were able to conclude about the presence of biosignature gases and false positives in the planets we simulated, and the strengths and weaknesses of using either transmission or direct imaging to assess the presence of life in its environmental

context.

Assessing biosignatures using transmission observations—Our results suggest that a mid-IR transmission telescope with twice the sensitivity of JWST will be adept at detecting and placing constraints on the abundances of the CO_2/CH_4 biosignature pair for a self-consistent M dwarf Earth-like planet, although it may struggle to meaningfully constrain O_3 . False positive discriminants may be present and may be at least, if not more detectable than the putative biosignatures, helping to discriminate between an abiotic or biological origin for these gases.

CO_2/CH_4 : Our results are consistent with previous work that shows carbon dioxide and methane may be a particularly accessible biosignature pair for transmission observations (e.g., [Krissansen-Totton et al., 2018c](#); [Lin et al., 2021](#); [Mikal-Evans, 2022](#); [Rotman et al., 2022](#); [Meadows et al., 2023](#)).

Using the detectability method described by [Lustig-Yaeger et al. \(2019a\)](#), we find that for observations ($t = 40$ hours) of an Earth-like TRAPPIST-1 e with an *Origins*-like telescope, CO_2 absorption bands can be detected with a confidence level of $11\text{-}\sigma$ and CH_4 absorption bands can be detected at $9.4\text{-}\sigma$. Together, these high-significance detections of the individual gases imply a strong detection of the pair. For the clear-sky, Earth-like TRAPPIST-1 e transmission case, we retrieve a broad CO_2 $1\text{-}\sigma$ distribution with abundances spanning $1\% - 37\%$. For the same planet, we retrieve a CH_4 $1\text{-}\sigma$ distribution spanning $43 - 2303$ ppm. When assessing this biosignature pair, the abundance of methane may provide additional evidence to discriminate between geological and biological sources of methane, by determining the inferred surface flux rate needed to maintain the observed abundance against photochemical destruction. A theoretical upper limit for abiotic methane production can be derived from serpentinization, water interacting with ultramafic rock, which is the most efficient known means of producing abiotic methane ([Arney et al., 2016](#)). [Guzmán-Marmolejo et al. \(2013\)](#) calculated that methane abundances exceeding the serpentinization upper limit of 2.5 ppm on an Earth-like planet around a G star require a biological origin. For a planet around an M

dwarf, the abundance of methane from serpentinization may be photochemically enhanced by up to $\sim 3\times$ (Domagal-Goldman et al., 2014), giving a theoretical upper limit of 7.5 ppm.

We can therefore interpret our results assuming that higher abundances of CH_4 combined with low abundances of CO will increase the likelihood that the methane has a biological origin (Krissansen-Totton et al., 2018c; Wogan et al., 2020). Our minimum retrieved CH_4 abundance at $1\text{-}\sigma$ (43 ppm) is nearly $6\times$ larger than the potential M-dwarf-enhanced serpentinization upper limit of 7.5 ppm described above. However, the minimum retrieved abundance at $2\text{-}\sigma$ (8 ppm) roughly equals this serpentinization upper limit, and the minimum retrieved abundance at $3\text{-}\sigma$ (2 ppm) is less than a third of this theoretical abiotic methane upper limit. This suggests that, although biogenic CH_4 is still likely, a conclusive ($> 3\sigma$) interpretation of methane’s origin may be difficult without the simultaneous detection of CO_2 found here. Similarly, the relatively poor constraint on CO, which encompasses hundreds of ppm (0.4 – 662.4 ppm at $1\text{-}\sigma$) could suggest that we cannot rule out abiotic methane produced by volcanism from a reduced mantle (Krissansen-Totton et al., 2018c; Wogan et al., 2020). Tighter constraints, or a stricter upper limit on CO, would be needed to conclusively discriminate between the abiotic and biological source hypotheses, barring the detection of the CO_2/CH_4 pair.

O_3/CH_4 : Both O_3 and CH_4 were detected for the inhabited TRAPPIST-1 e, which, given the short photochemical lifetime of methane in an oxidizing atmosphere, implies a chemical disequilibrium produced by an active flux of methane (Hitchcock & Lovelock, 1967). The indication of an active methane flux on the Earth-like TRAPPIST-1 e planet implies a highly-reducing atmosphere, which decreases the likelihood that the detected O_3 was generated abiotically (Domagal-Goldman et al., 2014). Conversely, the fact that no H-bearing species were detected in the false positive O_3 atmosphere implies that this atmosphere may be in the ultra-oxidizing regime required to generate abundant abiotic O_2/O_3 (Gao et al., 2015). However, the broad O_3 distributions we retrieve for both the living and the dead planets overlap at $1\text{-}\sigma$ (c.f., the O_3 panels in Figure 4.4), implying that the biological O_3 and the

abiotic O_3 cannot be confidently distinguished without additional environmental context.

To strengthen the case for biologically-mediated CH_4 on the inhabited planet, we can use serpentinization as an upper bound on the abiotic production to assess whether the retrieved abundances are consistent with life. However, as described above, we find that our $3\text{-}\sigma$ minimum retrieved CH_4 abundance for the inhabited TRAPPIST-1 e is smaller than the potential 7.5 ppm serpentinization upper limit estimated for M-dwarf stars, suggesting that the methane observed in the atmosphere could be consistent with abiogenesis. Likewise, the $3\text{-}\sigma$ upper limit of 27 ppm we obtain for the CH_4 non-detection in the false positive O_3 TRAPPIST-1 e exceeds the theoretical serpentinization upper limit. Although an abiotic origin is weakly favored ($\sim 2\text{-}\sigma$) by the non-detection, we cannot confidently conclude ($>3\text{-}\sigma$) that traces of methane do not still exist below our detection sensitivity. Future work can explore planets with differing abiotic methane fluxes to further assess the viability of constraining biological CH_4 abundances with transmission observations.

The inherent similarity between the Earth-like and false positive O_3 profiles in the transmission cases presents a challenge for accurate characterization despite relatively accurate retrievals. This implies that additional atmospheric context, such as CO and H_2O abundances, may be required to distinguish the living planet from the lifeless one when CH_4 is not present.

Large abundances of CO have previously been proposed as an abiotic O_3 discriminant (Hu et al., 2012; Domagal-Goldman et al., 2014; Gao et al., 2015; Schwieterman et al., 2018b; Meadows et al., 2018b). We find that the false positive O_3 planet is characterized by a CO abundance that is 20 to potentially $400\times$ greater than the nominal CO abundance of the living planet. High CO abundances have previously been argued to be possible discriminants for abiotic production of O_2 via photolysis of CO_2 (Domagal-Goldman et al., 2014; Gao et al., 2015; Schwieterman et al., 2016), or for abiotic CH_4 via volcanic outgassing from a more reducing mantle (Krissansen-Totton et al., 2018b). The potentially very high CO abundances (1 – 26%) retrieved in the abiotic case make an abiotic source far more likely. Though

previous work showed that living M dwarf planets with photosynthetic biospheres and both reducing or oxidizing atmospheres may also photochemically generate high abundances (100 ppm to several percent) of CO (Schwieterman et al., 2019b), this runaway effect has since been shown to be caused by a model artifact (Ranjan et al., 2023).

Alternatively, strict upper limits on atmospheric water vapor may provide additional context to help identify O₃ produced from abiotic O₂, as extremely low water vapor abundances inhibit the recombination of photolyzed CO₂, leading to a buildup of abiotic O₂, and O₃, in desiccated CO₂-dominated atmospheres (Gao et al., 2015).

While we showed above that the retrieved distribution of water vapor abundances for the Earth-like TRAPPIST-1 e would not allow us to confidently discriminate it from a Venus-like planet, the abiotic planet is extremely desiccated, with a 3- σ upper limit of 13 ppm H₂O, suggesting that it would be easily distinguished from a self-consistent Earth-like planet around an M dwarf with stratospheric water vapor abundances on the order of hundreds of ppm. However, we note that the false positive O₃ atmosphere used in this study is based on the most desiccated case explored by Gao et al. (2015), and abiotic planets with higher water vapor abundances may be more difficult to discriminate from a living planet.

Finally, when considering O₃ as a biosignature in transmitted light, we rely on O₃ to act as a proxy for O₂, but it may be difficult to accurately determine the abundance of the source oxygen, and so whether it is more likely to have a biological or abiotic origin. The conversion is complicated because O₂ and O₃ abundances are not linearly related above O₂ abundances of $\sim 0.2\%$ (Kasting & Donahue, 1980, Fig. 2). Above this limit, a strong degeneracy between ozone and O₂ abundances means that the same O₃ column depth could be generated in atmospheres with O₂ abundances that span orders of magnitude (Kasting & Donahue, 1980), depending on stellar type (Segura et al., 2003; Kozakis et al., 2022). The behavior implies that even highly precise O₃ column abundance constraints may yield imprecise O₂ constraints. In the linear regime, precise conversions to O₂ may still be possible, but the degenerate region spans O₂ abundances that are consistent with photosynthetic

production (Kasting & Donahue, 1980), complicating biosignature interpretation. While gaining insight into the vertical structure of O_3 in the atmosphere may allow photochemical models to better constrain the abundance of the source O_2 , and perhaps help to break the O_2 - O_3 degeneracy, this is unlikely to be effective for transmission observations. By modeling the vertical O_3 structure with photochemistry, it may be possible to back out the abundance of the source O_2 , and thereby deduce whether or not the O_2 is more or less likely to be biological. Unfortunately, we found that transmission is insensitive to the true vertical structure of O_3 for both the Earth-like and the desiccated planets in our experiments. Future transmission observations of atmospheres containing O_3 will require further lines of evidence to discriminate between biological and abiotic oxygen (Domagal-Goldman et al., 2014), such as constraints on the stellar UV flux (Segura et al., 2003; Kozakis et al., 2022), and a retrieval model that includes photochemistry.

Assessing biosignatures using direct imaging observations — Given the caveats that we are using a cloud-free retrieval model to analyze cloudy simulated data, our study shows that a LUVOIR-like direct imaging mission will excel at constraining O_2 and O_3 , but will face challenges constraining modern Earth-like CH_4 fluxes around a G-dwarf. For a full description of the biases that result in overestimated abundances of CO , CO_2 , and CH_4 when wavelength-dependent clouds are included in the data but not in the retrieval, see Section 4.2.1. Although hazes were not considered in our study, scattering slopes and absorption features from hazes would be expected to produce similar biases if present in the data but unaccounted for in the retrieval. We also find that direct imaging will have the advantage of providing precise constraints on O_3 in an oxygenic biosphere. Direct imaging may even be weakly sensitive to the ozone’s vertical structure via the UV Hartley band, potentially providing further context for an O_2 detection. Though we only find an upper limit on CH_4 for our Earth-twin, more informative methane distributions may be retrieved for planets with higher CH_4 abundances. Higher abundances may be due to larger methane surface fluxes, as may be the case for Archean (Arney et al., 2016) and Proterozoic (Segura et al.,

2003; Young et al., 2023) Earth, or longer methane lifetimes in the planetary atmosphere, as may be the case for modern methane fluxes around K, rather than G dwarfs (Arney, 2019). However, we also note that CH_4 may not have risen much above 10 ppm in the Proterozoic due to rapid oxidation of CH_4 in a sulfate rich ocean Olson et al. (2016). Abundant O_2 may also be a strong biosignature on its own, without the presence of CH_4 from a methanogenic biosphere, as long as false-positive mechanisms for the generation of abundant O_2 can be ruled out (Meadows, 2017).

O_2/CH_4 : Although our retrieval places tight constraints on O_2 for the Earth-like G dwarf planet, we do not detect CH_4 (1.86% VMR 3- σ upper limit), and arguments for the probability of the biological origin of the observed O_2 must therefore be based on its abundance alone (Meadows et al., 2018b). For the clear-sky Earth, we retrieve an accurate O_2 abundance of $21 \pm 5\%$ VMR with a significance of 11- σ . For our cloudy Earth-twin, we are able to constrain 60% of the true abundance of O_2 with a 1- σ distribution spanning abundances of 11 – 15% VMR. Regardless of whether clouds are present in the synthetic data, our results allow us to confidently rule out some ocean-loss false-positive atmospheres (Luger & Barnes, 2015; Schaefer et al., 2016) containing >30% VMR O_2 , as well as abiotic mechanisms that predict limited oxygen accumulation (<6% VMR) (Walker, 1977; Selsis et al., 2002; Segura et al., 2007; Harman et al., 2015; Meadows, 2017).

Though our study simulates LUVOR A (B) observations with 10 (40) hours of exposure time, the LUVOR final report simulates longer standard integration times of 100 (230) hours for the average target (LUVOR Team, 2019, their Figure 1-14). However, even these more generous exposure times are not sufficient to detect CH_4 at a 5- σ confidence level on a modern Earth twin at 10 pc. Using the detectability method described in Lustig-Yaeger et al. (2019a), we determine that absorption features from modern Earth methane abundances only become detectable at 5- (3-) σ in 650 (240) hours with LUVOR-A and approximately 2600 (960) hours with LUVOR-B for the direct imaging target studied here. For modern-Earth twins observed with direct imaging, confirmation of the biological origin for any detected O_2

may require ruling out the presence of potential false positive discriminants, like CO, rather than confirming the presence of CH₄. These results are consistent with [Young et al. \(2023\)](#), who show that narrow bandpass observations with the Habitable Worlds Observatory (HWO) of a modern Earth at SNR=10 may only rule out higher Archean methane abundances.

While direct imaging may not be sensitive to modern Earth fluxes of CH₄ around a sun-like star, either larger surface fluxes of CH₄, or longer photochemical lifetimes, could result in higher and potentially more detectable abundances of CH₄. For example, geological data suggesting that Earth once had an organic haze ([Zerke et al., 2012](#)), imply that the Archean required a higher methane surface flux than the modern Earth to sustain the production of this haze ([Arney et al., 2016](#)). Archean-like exoplanets ([Arney et al., 2016](#)) may also have higher fluxes ([Arney et al., 2017](#)), as may Proterozoic exoplanets ([Segura et al., 2003](#); [Young et al., 2023](#)). Even for similar surface fluxes to modern Earth, the incident spectrum of the parent star reduces the destruction rate of methane, potentially allowing it to build up to abundances 100 times greater than modern-Earth levels around a K6V star ([Arney, 2019](#)). Promisingly, the planned HWO will likely target planets around K and M stars ([Mamajek & Stapelfeldt, 2023](#)).

Constraints on O₃ could also provide important additional context for interpreting the planetary environment and an oxygenic biosphere. We retrieve an O₃ abundance that is precise (± 70 ppb), significant ($19\text{-}\sigma$), and consistent within the range of O₃ VMRs in the true vertical structure (see Fig. 4.9), but this abundance is only about half the value of the peak abundance in the stratospheric bulge. When we included vertical O₃ structure in the retrieval, we achieved an improved fit, but with weak statistical evidence in favor of the more complex model (see Table 4.7). Since O₃ is a photochemical byproduct of O₂ ([Ratner & Walker, 1972](#)), accessing its vertical structure could help reveal indirect constraints on the atmospheric chemistry and potentially the biological productivity of the atmosphere through additional modeling steps. In addition, since direct imaging is particularly sensitive to total column abundance ([Robinson, 2017a](#)), we find that retrieving the vertical O₃ structure helps

to mitigate bias that can occur when assuming an evenly mixed O_3 profile, and enable total column abundance inferences that are closer to the true value (c.f., Figure 4.10).

Finally, we note that while an accurate O_2 abundance can rule out false positive processes that generate significantly larger and smaller abundances of abiotic O_2 , for habitable-zone planets around G-dwarf stars, identifying false positive generation of Earth-like O_2 amounts ($\sim 20\%$ VMR) could be quite difficult. For G dwarf planets, the most likely known abiotic mechanism for generating Earth-like abundances of oxygen is the photolysis of stratospheric H_2O in atmospheres with low abundances of non-condensable gases (Wordsworth & Pierrehumbert, 2014). To discriminate between abiotic and biological O_2 in these scenarios, we require constraints on the abundance of non-condensable gases in the planetary atmosphere (Meadows et al., 2018b). However, for a nitrogen-dominated atmosphere, the $\text{N}_2\text{-N}_2$ collision-induced absorption feature occurs at longer wavelengths than those accessible to HWO (Schwieterman et al., 2015; Meadows et al., 2018b), and will likely not be accessible in future direct imaging observations. Instead, it may be possible to constrain the bulk gas by fitting for the mean molecular weight of the atmosphere, and retrieving the abundances of other non-condensable, spectrally active gases (Lustig-Yaeger, 2020).

CO_2/CH_4 : The CO_2/CH_4 disequilibrium pair was first proposed as a biosignature for the Earth’s Archean and for planets orbiting M dwarfs (Krissansen-Totton et al., 2018c, 2016), and we confirm that it is unlikely to be detectable for modern-Earth-like planets orbiting G dwarfs observed using direct imaging, due primarily to the difficulty in detecting modern-Earth-like quantities of CH_4 . For our clear-sky modern Earth, we obtain a modest detection of CO_2 , retrieving a $1\text{-}\sigma$ distribution of CO_2 abundances spanning 17 – 489 ppm in the clear-sky case, which we conservatively interpret as providing a $3\text{-}\sigma$ upper limit of 91% VMR. Similarly, we do not detect CH_4 , obtaining only a $3\text{-}\sigma$ upper limit of 1.86% VMR.

Furthermore, high abundances of CO may indicate that an abiotic volcanic source for CH_4 is more likely (Wogan et al., 2020). However, our retrievals lacked the sensitivity to rule out high CO abundances, as we retrieve a $1\text{-}\sigma$ upper limit of 60 ppm on the abundance

of CO for the clear-sky case, which is 600 times larger than our true value, and 600 times larger than the average CO abundance measured in the modern Earth’s atmosphere (Badr & Probert, 1994). Consequently we confirm that even though the CO₂/CH₄ biosignature pair is a high priority biosignature for transmission observations of M-dwarf targets, and will likely be observable even for a planet with modern-Earth-like fluxes orbiting an M (Segura et al., 2005; Meadows et al., 2023) or K star (Arney, 2019), it will be challenging to detect for a modern Earth-Sun observed in reflected light.

4.3 Model Limitations and Future Work

To simplify our atmospheric retrieval model, we made assumptions related to atmospheric composition and vertical structure, and neglected spectral contamination effects, which may have impacted our results. These model limitations included assuming N₂ as the backfill gas for the atmospheric bulk composition, implementing an isothermal temperature profile, using a single set of absorption coefficients for calculating radiative transfer, excluding clouds from the forward model, and omitting spectral contamination effects from the synthetic data. We discuss impacts and future work for each of these limitations below.

Our model adds N₂ as the bulk component of the retrieved atmospheric composition to achieve the inferred pressure of the terrestrial atmosphere in all cases (Equation 4.4). In atmospheres with small amounts of trace gases, this results in the assumption that N₂ is the predominant bulk gas. The practice of backfilling a terrestrial atmosphere with N₂ is well-established in the retrieval community (e.g., Barstow & Irwin, 2016; Krissansen-Totton et al., 2018b; Changeat et al., 2019; Barstow et al., 2020) as N₂ is a common additional bulk gas in the Solar System’s O₂ and CO₂-dominated terrestrial atmospheres, as it comprises more than 1% of the atmosphere.

However, assuming N₂ as the filling gas in our retrieval imposes a strong prior that N₂ is the bulk gas in the atmosphere due to our use of uniform log priors (Benneke & Seager, 2012, their Figure 1). For our Earth-like cases where N₂ is in fact the dominant gas in

the atmosphere, this assumption of N_2 as the bulk composition does not bias our inferred abundances. However, for the abiotic false positive O_3 TRAPPIST-1 e scenario, we find that this strong prior on N_2 as the bulk gas likely causes us to underestimate the CO_2 abundance by a factor of ~ 2 in this CO_2 -dominated atmosphere. This assumption also causes us to underestimate the mean molecular weight of the atmosphere, since N_2 (28 g/mol) is comparatively lighter than CO_2 (44 g/mol). This bias in the mean molecular weight enhances the inferred scale height of the atmosphere, and the retrieval compensates by decreasing the temperature of the atmosphere, and in some cases, the CO_2 abundance. The pitfalls of assuming a filling gas could be avoided by implementing centered-log ratio (CLR) priors, which allow any gas included in the retrieval to take on high (bulk gas) abundances at equal probability (e.g., [Benneke & Seager, 2012](#); [Damiano & Hu, 2021](#)), or by directly retrieving on the atmospheric mean molecular weight ([Tremblay et al., 2020](#)). These methods will be implemented in subsequent versions of our retrieval code.

Across all of our cases, our retrieval forward model also assumed an isothermal temperature profile to fit data generated with a realistic profile (see Figure 4.2). The use of an isothermal temperature profile did not impact our retrieval of temperature in reflected light, given our inability to constrain it from the spectrum at the spectral resolution investigated. This is likely a reasonable assumption going forward for studies of future direct imaging observations with comparable spectral resolution. In contrast, transmission spectra are more sensitive to temperature, given the dependence of spectral feature size on temperature via the atmospheric scale height ([Seager, 2011](#)). Our constraints on the retrieved temperature for the modern Earth-like planet are consistent with the stratospheric temperature, suggesting that assuming an isothermal temperature profile did not significantly impact our results. This may be an adequate approximation for stratospheres in Earth-like M-dwarf planet atmospheres, which do not show temperature inversions ([Lincowski et al., 2018](#)). Future work could investigate whether or not transmission spectroscopy provides insight to the temperature structure of secondary atmospheres by implementing a more complicated

parameterization of temperature in the forward model (e.g., [Madhusudhan & Seager, 2009](#); [MacDonald et al., 2020](#)).

To increase computational speed, our retrievals did not recalculate absorption coefficients during the retrieval exploration, relying instead on using a single set of absorption coefficients calculated from the realistic simulated atmosphere. While the optical depth in the radiative transfer translates changes in the abundance of each species into observables, the change in the intrinsic line shape and strength are fixed at the assumed line-by-line absorption coefficients throughout the retrieval. This assumption may have contributed to some biases in our CO₂ and H₂O retrieval results, but it does not affect our overall conclusions.

To test the impact of this model assumption, we re-ran absorption coefficients for every state in the transmission retrieval. These tests returned identical posteriors to those using the single set of absorption files generated with the input spectrum, with the exception of CO₂, when it was a bulk component of the atmosphere ($\geq 10\%$). In this case, the CO₂ posterior differed by about 20%, and provided better agreement with the true CO₂ abundance. The only significant effect of the single set of absorption coefficients in this test was a small underestimation of the CO₂ abundance (within $1\text{-}\sigma$ at the simulated precision) in our abiotic false positive O₃ TRAPPIST-1 e case. We additionally tested transmission retrievals run with a single set of absorption coefficients for the input spectrum atmosphere, and retrievals ran with absorption coefficients generated at isothermal/evenly-mixed approximations of the true answer, and these also returned equivalent posteriors, given the finite sampling precision.

Lastly, we quantify the impact of neglecting to recalculate absorption coefficients at each forward model instance by performing a perturbation sensitivity test. We take our best-fit retrieval parameters and perturb the abundances of the critical gases (CO₂ and CH₄ for transmission, and O₂ and H₂O for direct imaging). We then calculate the perturbed spectrum using the unperturbed best-fit absorption coefficients, and calculate it again using recalculated absorption coefficients consistent with the perturbed abundances. Finally, we compare the spectra generated at both sets of absorption coefficients within the simulated noise. We

find that for *Origins*-like transmission observations, when CH_4 and CO_2 abundances are perturbed by up to a factor of 10 without recalculating the absorption coefficients, we observe changes in the spectrum no larger than 1 ppm, which are insignificant ($< 0.5\text{-}\sigma$) for the SNR associated with the $t = 40$ hours exposure time explored here. For LUVIR-A-like direct imaging observations, when oxygen and water abundance is doubled we observe changes to the spectrum on the order of 1/100th the planet-star contrast, which are insignificant given 10 hours of exposure time. However, when water abundance is enhanced by a factor of 10 without recalculating the absorption coefficients, we observe changes to the spectrum on the order of 1/10th the planet-star contrast, which are significant at $4\text{-}\sigma$. This suggests that our retrieved water distributions with high-SNR LUVIR observations may be weakly sensitive to our simplifying line shape assumptions.

Although these model simplifications have not significantly impacted our results, identifying an accurate and computationally efficient means to treat the rotational-vibrational absorption coefficients in terrestrial atmospheres remains important work on the path to performing biosignature assessment with future telescopes. Recent work has shown that differing opacity tables and assumptions may lead to biases in the results of thermal emission retrievals of Earth-twins (Alej et al., 2022) and JWST retrievals of Super-Earth and Warm-Jupiter atmospheres (Niraula et al., 2022). Similarly, the community has identified the growing need for additional laboratory measurements to better inform the opacities that support these critical exoplanet science applications (Fortney et al., 2019).

Another limitation of our retrieval model is that it does not simulate the effect of clouds in the forward model, as discussed in more detail in Section 4.2.1, which can result in overestimates of the surface albedo and retrieved gas abundances. Since the simulated observations had realistic water and ice clouds, which produce wavelength dependent absorption features in the observed spectrum, we find that the cloud-free model attempts to compensate for cloud scattering by enhancing the reflectivity of the planet surface and increasing the absorption of CH_4 , CO , and CO_2 , which have features in the near-IR where the ice clouds also

absorb. The cloud-free simplification consequently results in overestimates of the retrieved CH_4 ($9\times$), CO ($5600\times$), and CO_2 ($5\times$) abundances, compared to their true values, as well as a large overestimate of the grey surface albedo ($1.8\times$) for direct imaging observations of the cloudy Earth-twin. Other retrieval models have successfully included parameterized clouds for transmission (Benneke & Seager, 2012) and direct imaging cases (Feng et al., 2018; Damiano & Hu, 2022), and this capability will also be included in subsequent versions of our retrieval code.

Similar to cloud and condensation treatment in this study, while our simulated data implicitly includes the effect of photochemistry by virtue of our photochemically self-consistent atmospheres, the retrieval itself does not. Though this model simplification is currently the standard in the retrieval community, there is abundant work predicting the detectable effects of photochemistry on observations of exoplanet atmospheres (e.g., Segura et al., 2003, 2005; Hu et al., 2012; Miguel & Kaltenegger, 2013; Rugheimer et al., 2015; Arney et al., 2017; Meadows et al., 2017; Lincowski et al., 2018; Schwieterman et al., 2018b; Arney, 2019). Furthermore, Tsai et al. (2023) recently showed that separate photochemical forward models could be used to explain the retrieval of disequilibrium SO_2 in hot Jupiter atmospheres. In this case, standard retrieval models without photochemistry detected SO_2 , but separate photochemical modeling was required to explain its presence. For rocky planets, Krissansen-Totton et al. (2018b) compared molecular abundance posteriors to surface fluxes as an example of connecting retrieval results to photochemistry. Notably, none of this recent work fully folds photochemistry into the retrieval forward model. We take a step in this direction with this study by attempting to vertically resolve the photochemically-mediated structure of the O_3 profile in an Earth-like atmosphere. Future work could explore methods for explicitly including photochemistry within a retrieval framework, such as implementing photochemistry into the forward model or post-processing retrieval posteriors with a separate photochemistry retrieval.

Finally, we note that we have not included stellar and instrumental spectral contamination

in this study, but previous work has shown that contamination sources may have a significant impact on our ability to characterize exoplanets in transmitted and reflected light. For example, in transmitted light, stellar variability and contamination due to unocculted star spots and faculae may produce a signal that is more than 10 times greater than the spectral transit depths anticipated for terrestrial exoplanets (Rackham et al., 2018). Initial JWST results of TRAPPIST-1 and other M dwarfs suggest that the effects of stellar activity (e.g., transit light source effect, flares, etc.) need to be mitigated in order to access secondary atmospheres for rocky exoplanets transiting M dwarfs (Lim et al., 2023; Moran et al., 2023; May and MacDonald et al., 2023). In this study, we assume this mitigation has been applied prior to our analysis, although work is ongoing to develop such methods. In reflected light, faint background stars may be a significant source of contamination for planets located at low galactic latitudes (Cracraft et al., 2021), and scattered light may be a significant background and noise signal source for large aperture telescopes like HWO (Pfisterer et al., 2018). The impact of stellar activity on reflected light observations has not been extensively studied and must be the subject of future investigations. While outside the scope of this work, future model development is necessary to incorporate more realistic contamination sources into forward-looking retrieval investigations for both transmission and direct imaging observations.

5 Conclusion

In this study, we used a new terrestrial exoplanet retrieval model to compare the capabilities of transmission and direct imaging spectroscopy for habitability assessment and biosignature pair detection in rocky, habitable zone exoplanet atmospheres.

Transmission spectroscopy with a 5.9-m *Origins*-like telescope will not be able to probe the surface and near-surface environment of TRAPPIST-1 e, hampering habitability assessment, but will likely have the sensitivity to detect the CO₂/CH₄ biosignature pair and discriminate

an extremely desiccated O_3 false positive case. By fitting a simplified temperature profile to the transmission data, for both the biotic and abiotic planets we find that the $1\text{-}\sigma$ lower bound on the continuum pressure of $\sim 0.1\text{-bars}$ ($1 \times 10^4 \text{ Pa}$) corresponds to altitudes tens of kilometers above the surface in the true atmosphere, but the high-pressure upper bound of $\sim 4 \text{ bars}$ ($4 \times 10^5 \text{ Pa}$) has no altitude counterpart in the true atmosphere. The broad pressure distributions we retrieved exemplify that the transmission continuum does not disambiguate a solid surface from atmospheric opacity. We find that the inferred isothermal temperature is most representative of the input temperature profile at levels corresponding to the stratosphere (50-70 km). We therefore cannot use surface temperature or near-surface water abundance to inform habitability assessment of the Earth-like TRAPPIST-1 e since transmission does not sufficiently probe these phenomena. Despite these limitations for characterizing the habitability of the self-consistent Earth-like TRAPPIST-1 e, we find that a mid-IR transmission telescope with twice the sensitivity of JWST can detect CO_2 absorption features at $11\text{-}\sigma$ and CH_4 absorption features at $9.4\text{-}\sigma$ in 40 hours of exposure, providing a high-confidence detection of the biosignature pair. Finally, transmission probes broad distributions of O_3 , which overlap for the biotic and abiotic planets and thus cannot be used to discriminate between a photosynthetic biosphere and a CO_2 -dominated abiotic atmosphere. However, in the context of an O_3 detection, we find that the non-detection of water vapor, with a $3\text{-}\sigma$ upper limit of 13 ppm, coupled to the detection of abundant CO ($1\text{-}\sigma$ range of 1-26%), could be used to distinguish between O_3 sourced from CO_2 photolysis and biogenic O_3 sourced from abundant O_2 .

For future direct imaging missions with a 15-m LUVOIR-A like telescope, we cannot constrain surface temperature at the spectral resolution explored here, and although our model was relatively sensitive to planetary albedo, this parameter is degenerate with planetary radius. While insensitivity to temperature and the size/albedo degeneracy complicate our ability to assess habitability, our model was able to constrain the near-surface water abundance with 50% or 0% cloud coverage in the simulated data to within 0.1% of the true

value at high significance ($>10\text{-}\sigma$) even for our assumed exposure time of only 10 hours. We detect O_2 and O_3 with high detection significances of $11\text{-}\sigma$ and $19\text{-}\sigma$, respectively. We also precisely constrain their abundances with ($1\text{-}\sigma$) errors of $\pm 5\%$ VMR for O_2 and ± 70 ppb VMR for O_3 . However, we fail to detect CO_2 and CH_4 , retrieving $3\text{-}\sigma$ upper limits on molecular abundances of 91% and 1.83% , respectively, on the modern Earth-twin. Given our finding that short (~ 10 hour) exposure times with a 15-m direct imaging telescope are insensitive to modern Earth-like CH_4 abundances on planets around G stars, the CO_2/CH_4 and O_2/CH_4 biosignature pairs are difficult to assess for these planetary systems. However, the predicted photochemical enhancement of CH_4 on M- (Segura et al., 2005; Rugheimer et al., 2015) and K-dwarf (Arney, 2019) planets may make the retrieval of the O_2/CH_4 pair more feasible in these systems, and HWO will target 3 M dwarfs and 27 K dwarfs as part of its planned target list for discovering and characterizing Earth-like planets (LUVOIR Team, 2019; Mamajek & Stapelfeldt, 2023). For exposure times of up to 10 (~ 40) hours with a 15-m LUVOIR-A-like (6.7-m LUVOIR-B-like) telescope of a planet at 10 parsecs, we find that direct imaging is weakly sensitive to the vertical structure of the stratospheric ozone bulge via the shape of the Hartley band ($0.2\text{-}0.3\text{ }\mu\text{m}$) when comparing an evenly-mixed and vertically-resolved retrieval model (Bayes factor ~ 1.3). Finally, when clouds are present in the observed spectrum, we see evidence for them in the large residuals at the bottoms of the water-bands (S/N up to 6.5) and from NIR ice absorption features, which may be used to uniquely identify the presence of clouds in the reflected light spectra of Earth-like planets even when not explicitly included in the retrieval model. The profound science objectives of future exoplanet missions—to constrain habitable and inhabited environments—will require further retrieval model developments to fully assess in advance.

Appendix

Corner Plots of Posterior Distributions

We display the complete set of posteriors for the retrievals presented in this study. Figures 12-19 include the posteriors for the two transmission cases, and Figures 20-27 include the posteriors for the two direct imaging cases. Specifically, figures 12-15 include the complete set of posterior distributions for the transmission retrievals of the Earth-like TRAPPIST-1 e with the evenly-mixed ($\mathcal{M}_{iso}$), three fixed points ($\mathcal{M}_{3P}$), five fixed points ($\mathcal{M}_{5P}$), and the three free points ($\mathcal{M}_{3F}$) forward models. This pattern repeats for the remaining planetary cases investigated here, with the transmission retrievals of the abiotic, false-positive O₃ TRAPPIST-1 e (Figures 16-19), and the direct imaging retrievals of the clear-sky (Figures 20-23) and cloudy Earth (Figures 24-27).

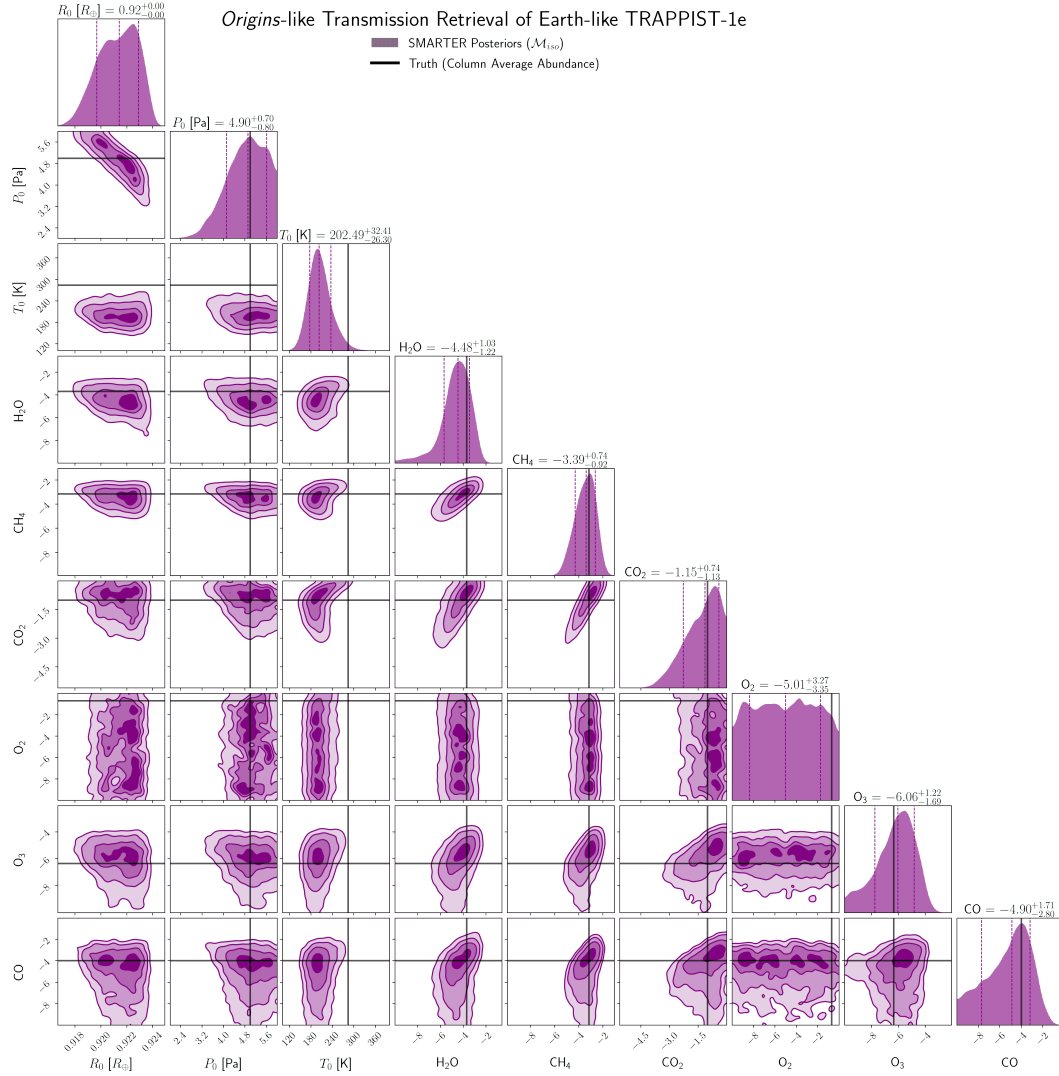


Figure 4.12: The results of the retrieval on the Earth-like TRAPPIST-1 e transmission spectrum with the evenly-mixed forward model, $\mathcal{M}_{iso}$. The plots on the diagonal are the posterior distributions for each free parameter in the retrieval, while the off-diagonal plots indicate the covariances between the retrieval parameters. Truth values are indicated by solid black lines. For the gas abundances, these truth values represent the column average abundance for each gas in the true spectrum.

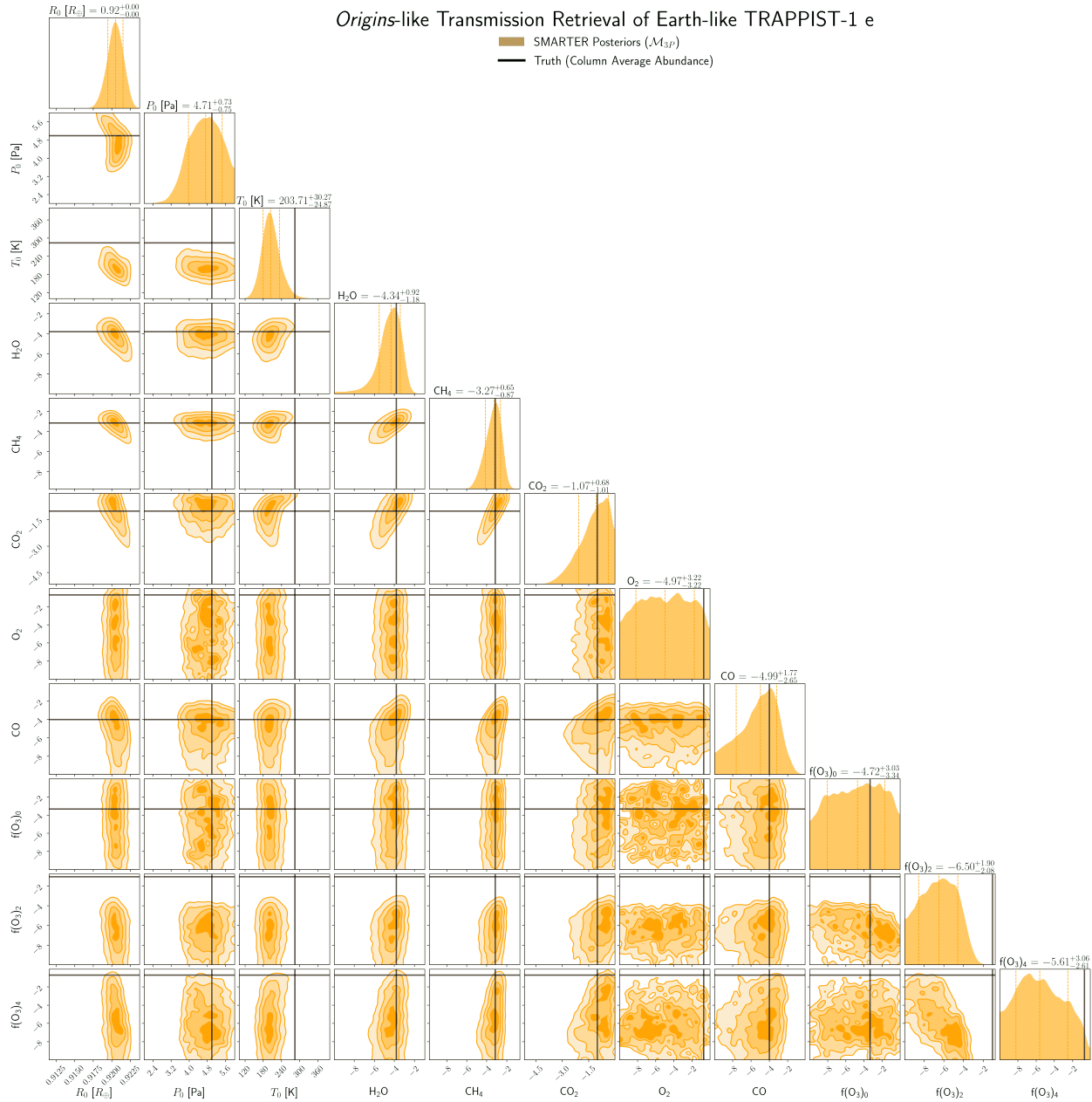


Figure 4.13: The results of the retrieval on the Earth-like TRAPPIST-1 e transmission spectrum with the three-fixed points forward model, $\mathcal{M}_{3P}$. Truth values are indicated by solid black lines. For the non- O_3 gas abundances, these truth values represent the column average abundance for each gas in the true spectrum. For the O_3 gas abundances, the truth values represent the true O_3 value at the corresponding pressure level based on the true profile.

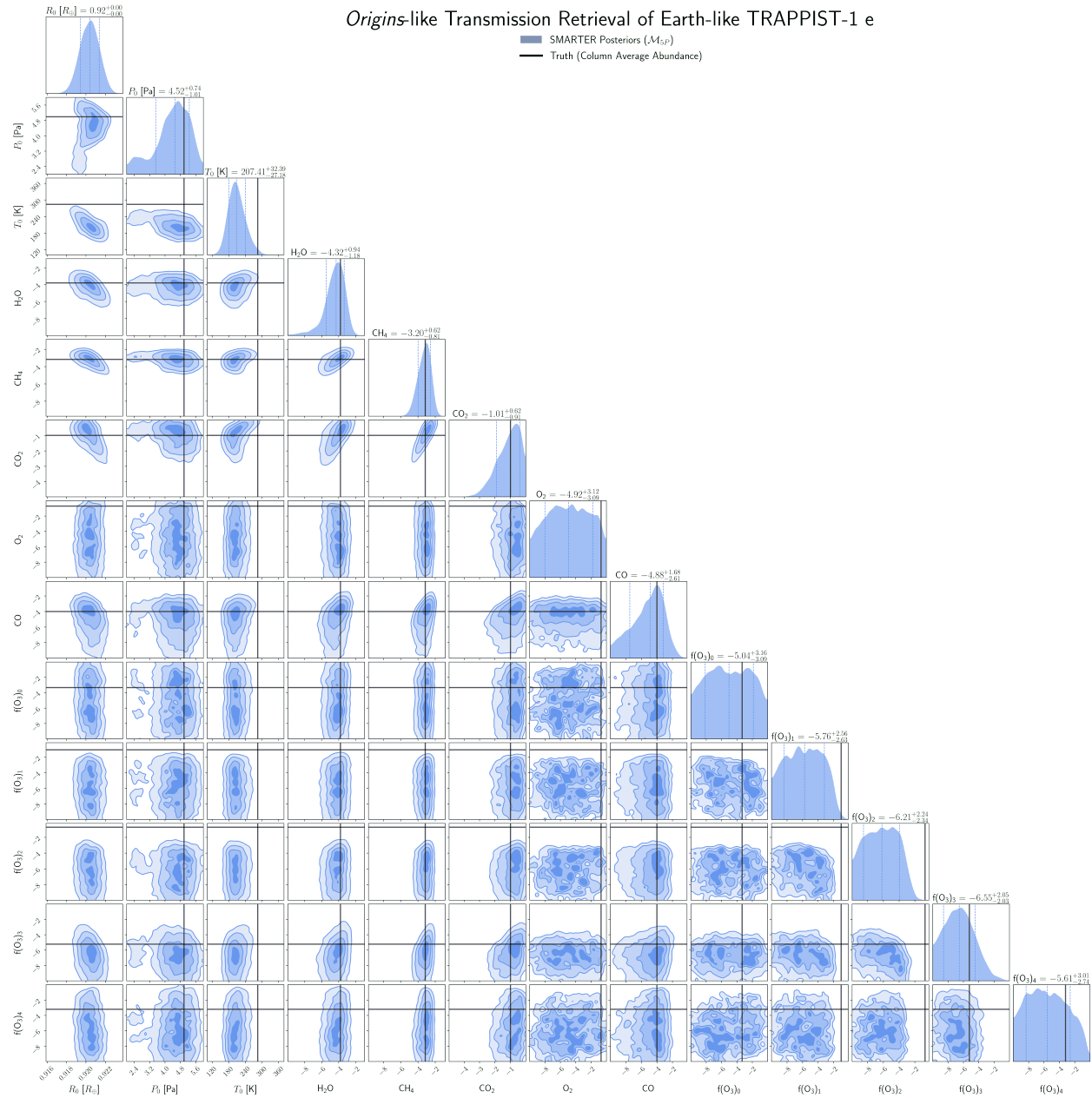


Figure 4.14: The results of the retrieval on the Earth-like TRAPPIST-1 e transmission spectrum with the five-fixed points forward model, $\mathcal{M}_{5P}$.

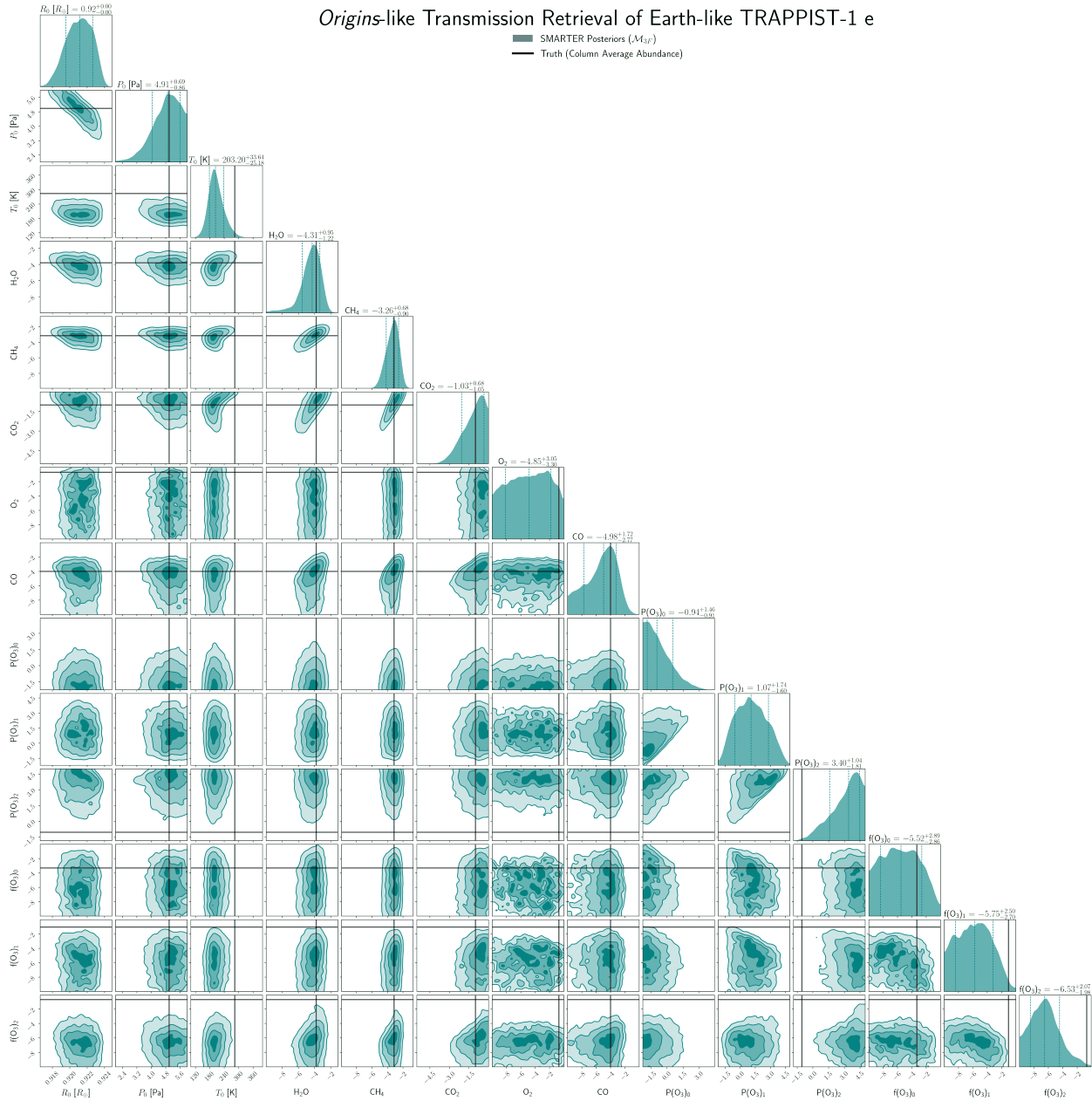


Figure 4.15: The results of the retrieval on the Earth-like TRAPPIST-1 e transmission spectrum with the three-free points forward model, $\mathcal{M}_{3F}$.

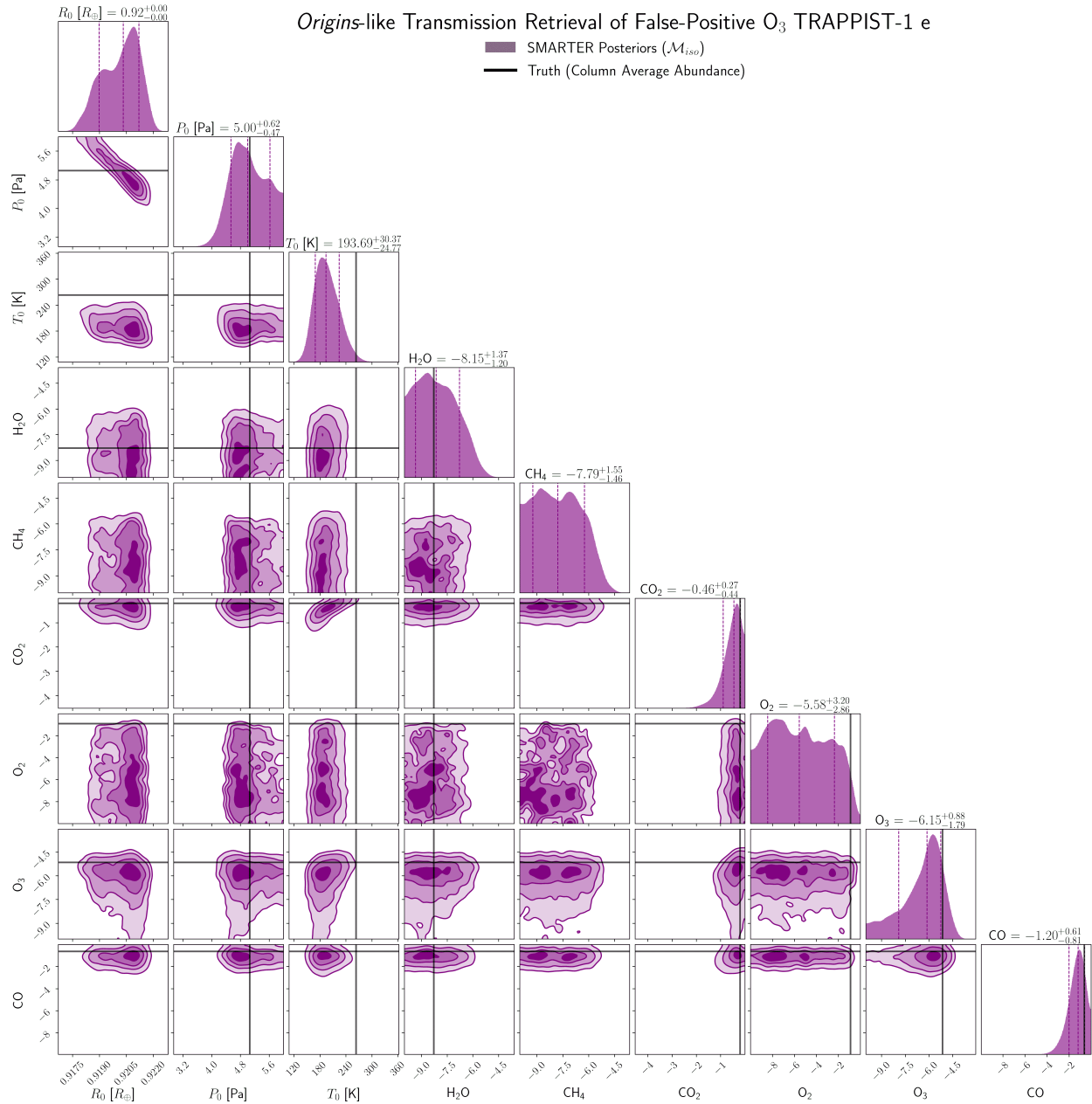


Figure 4.16: The results of the retrieval on the abiotic, false-positive O₃ TRAPPIST-1 e transmission spectrum with the evenly-mixed forward model, $\mathcal{M}_{iso}$.

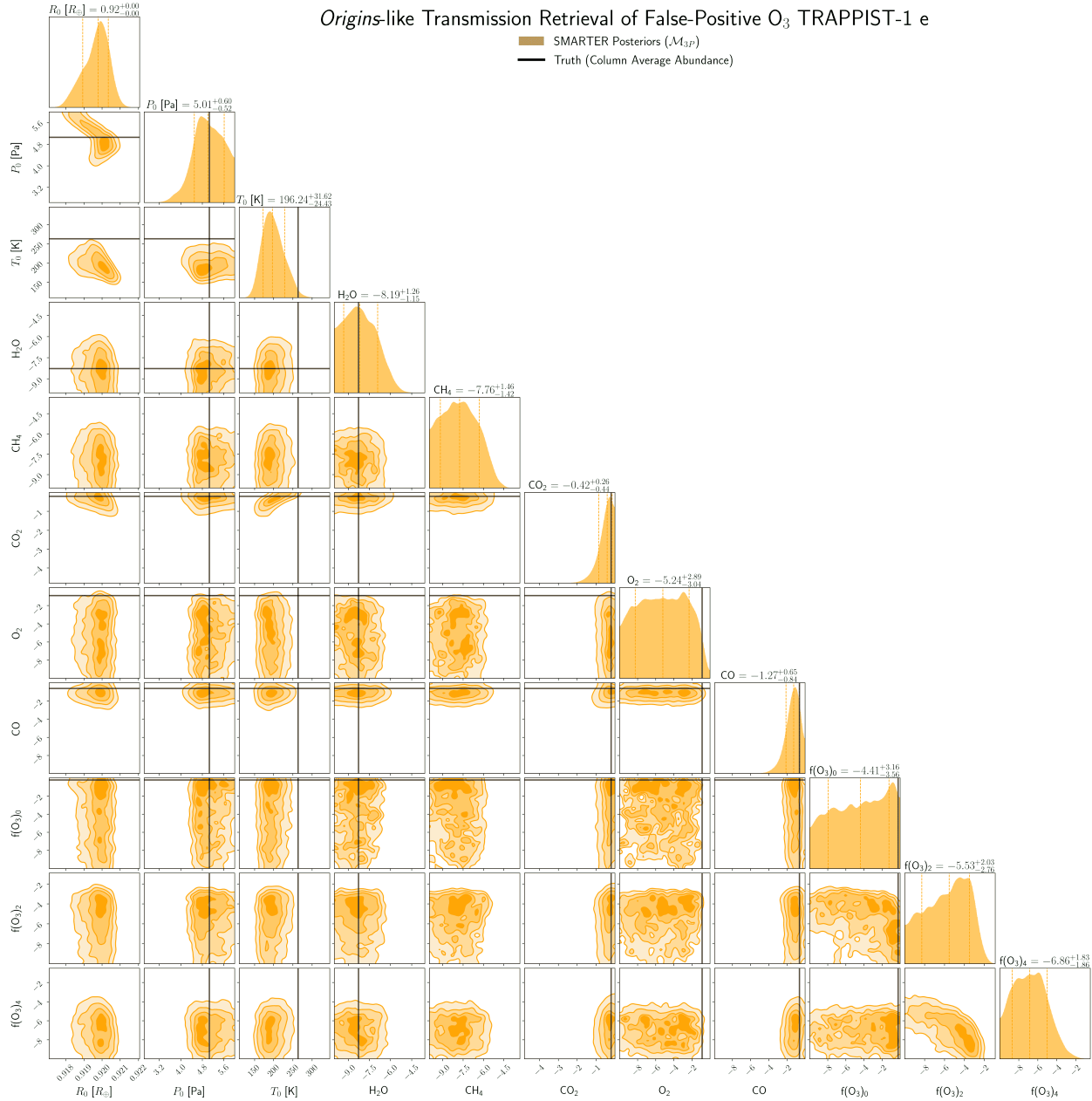


Figure 4.17: The results of the retrieval on the abiotic, false-positive O₃ TRAPPIST-1 e transmission spectrum with the three-fixed points forward model, $\mathcal{M}_{3P}$.

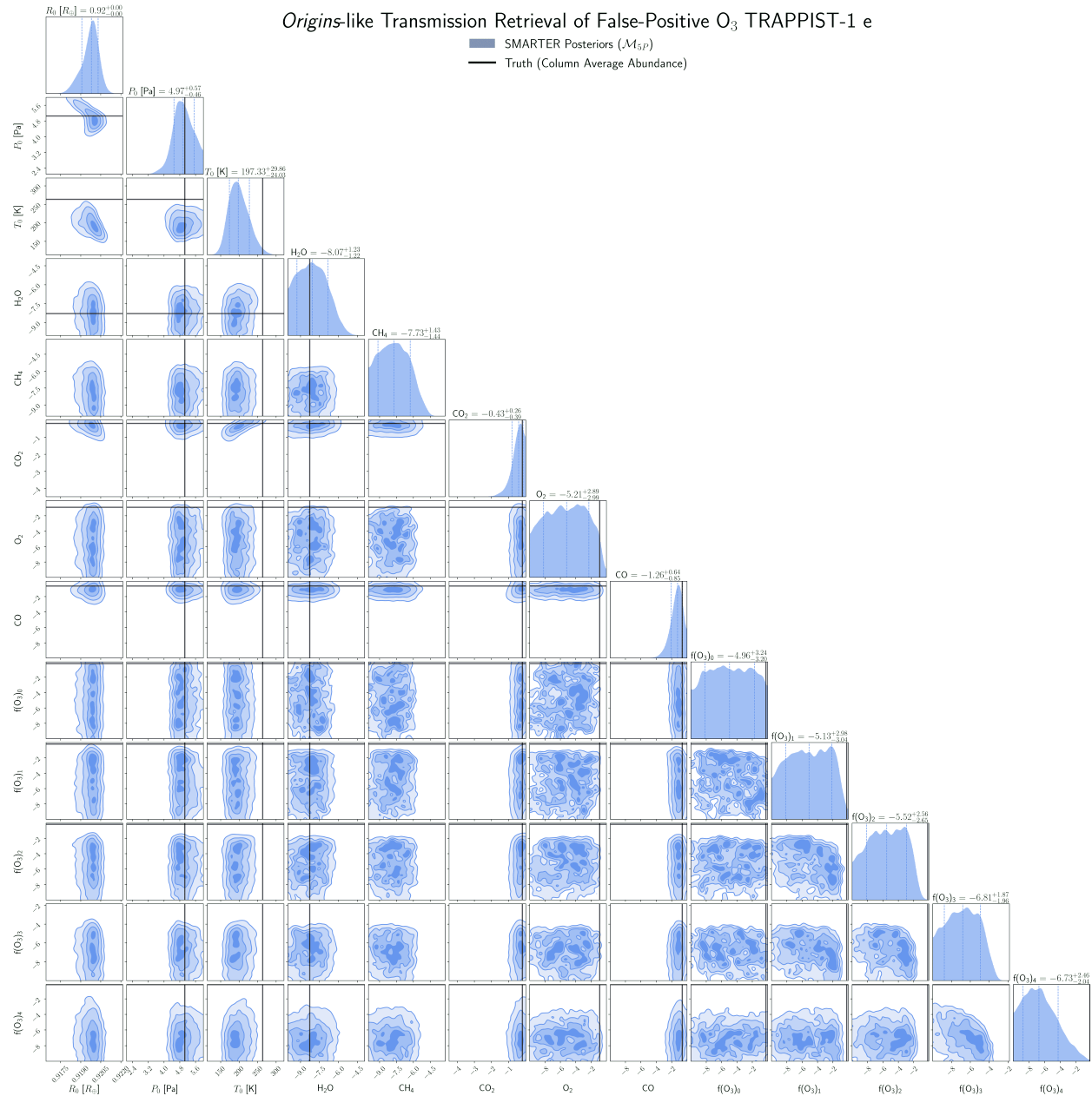


Figure 4.18: The results of the retrieval on the abiotic, false-positive O₃ TRAPPIST-1 e transmission spectrum with the five-fixed points forward model, $\mathcal{M}_{5P}$.

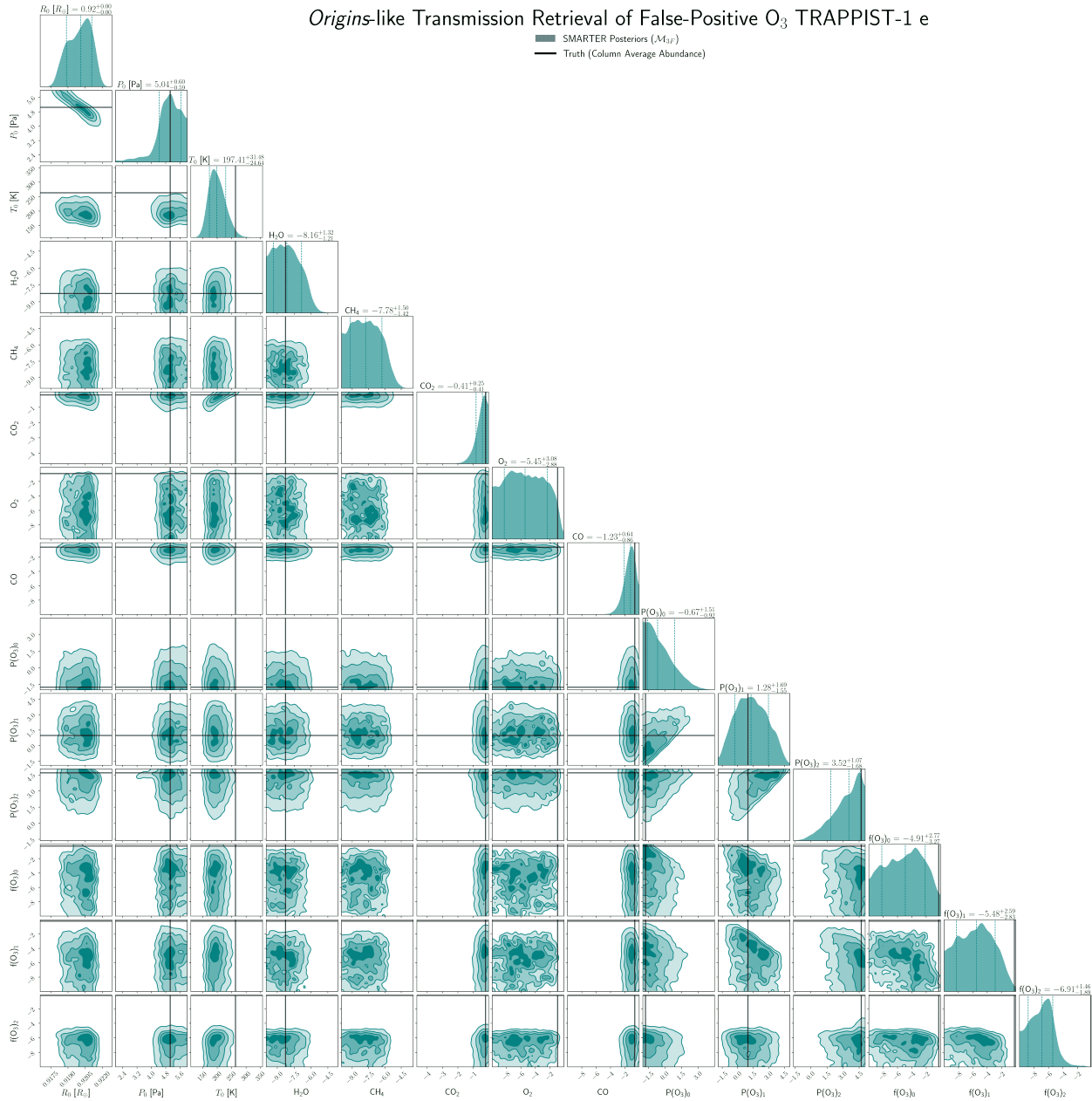


Figure 4.19: The results of the retrieval on the on the abiotic, false-positive O₃ TRAPPIST-1 e transmission spectrum with the three-free points forward model, $\mathcal{M}_{3F}$.

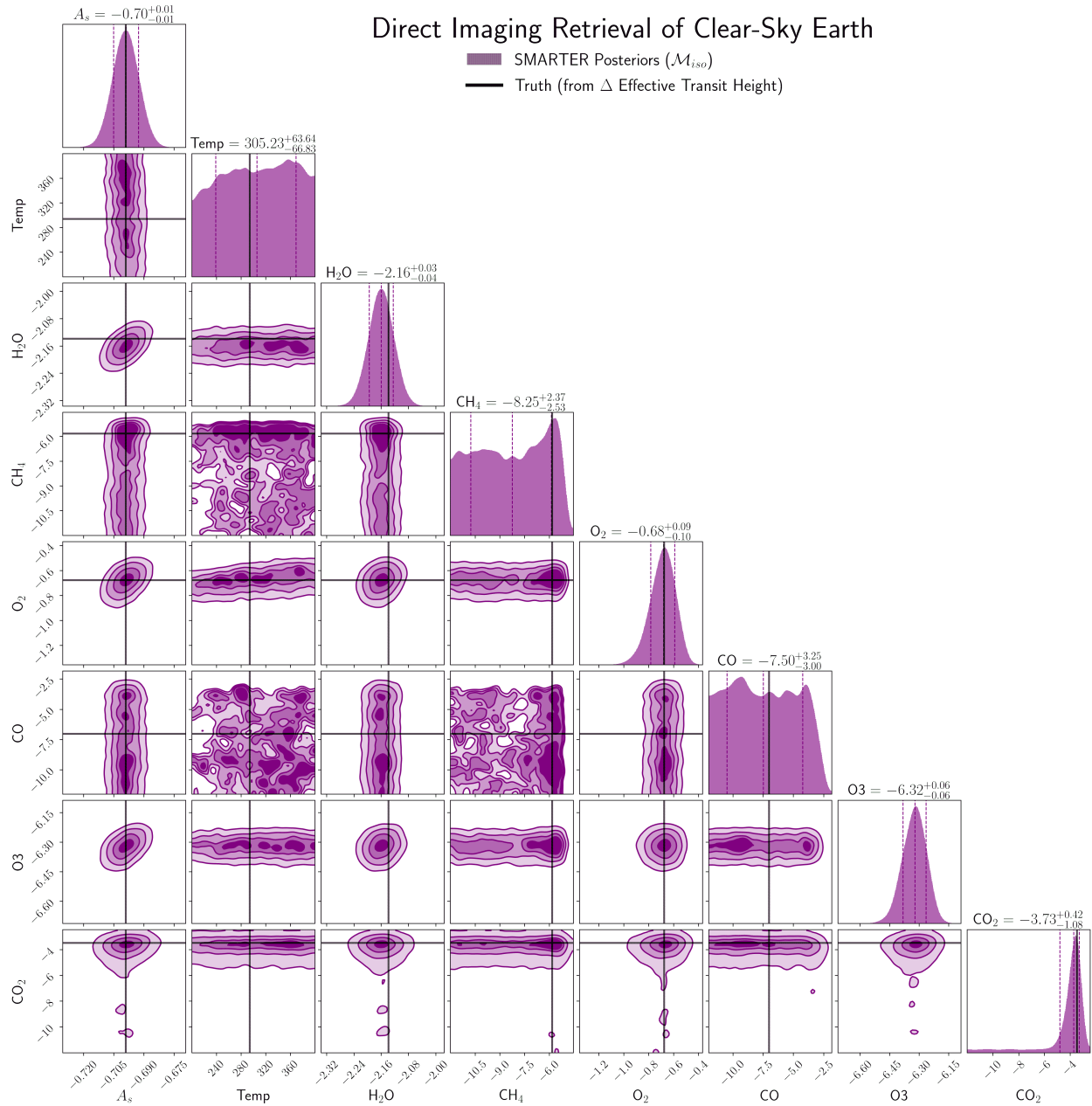


Figure 4.20: The results of the retrieval on the clear-sky Earth direct imaging spectrum with the evenly-mixed forward model, $\mathcal{M}_{iso}$.

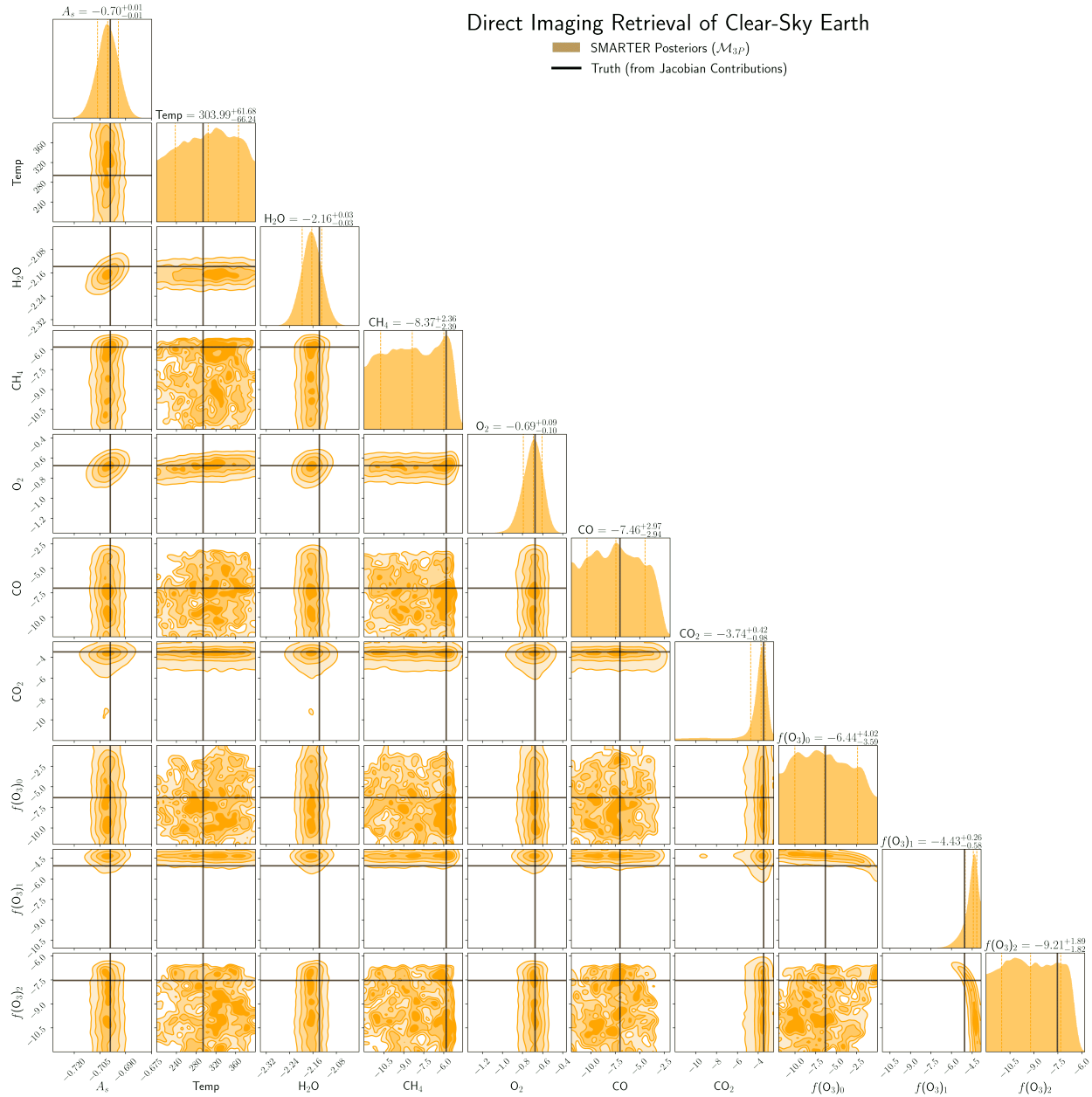


Figure 4.21: The results of the retrieval on the clear-sky Earth direct imaging spectrum with the three-fixed points forward model, $\mathcal{M}_{3P}$.

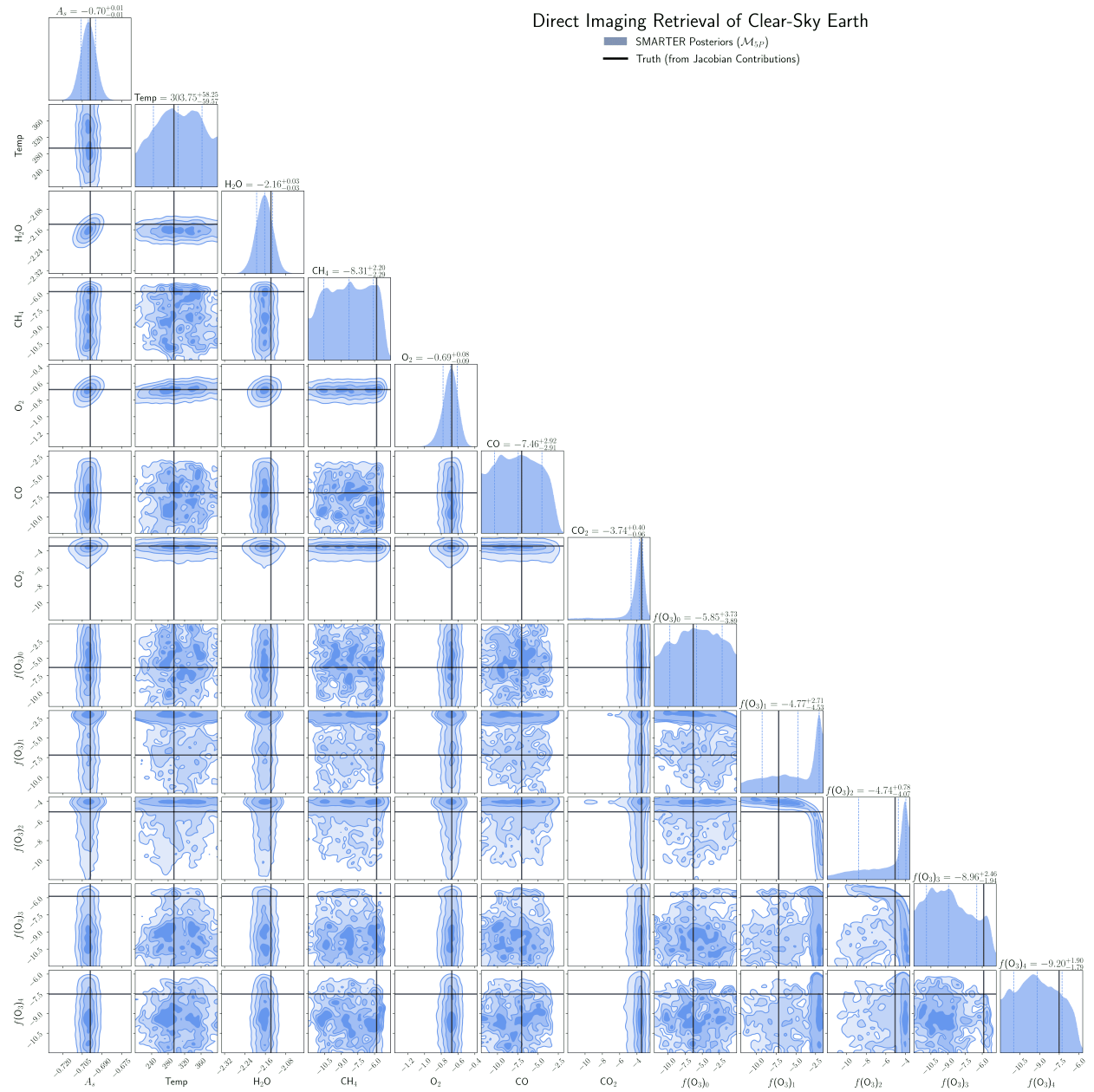


Figure 4.22: The results of the retrieval on the clear-sky Earth direct imaging spectrum with the five-fixed points forward model, $\mathcal{M}_{5P}$.

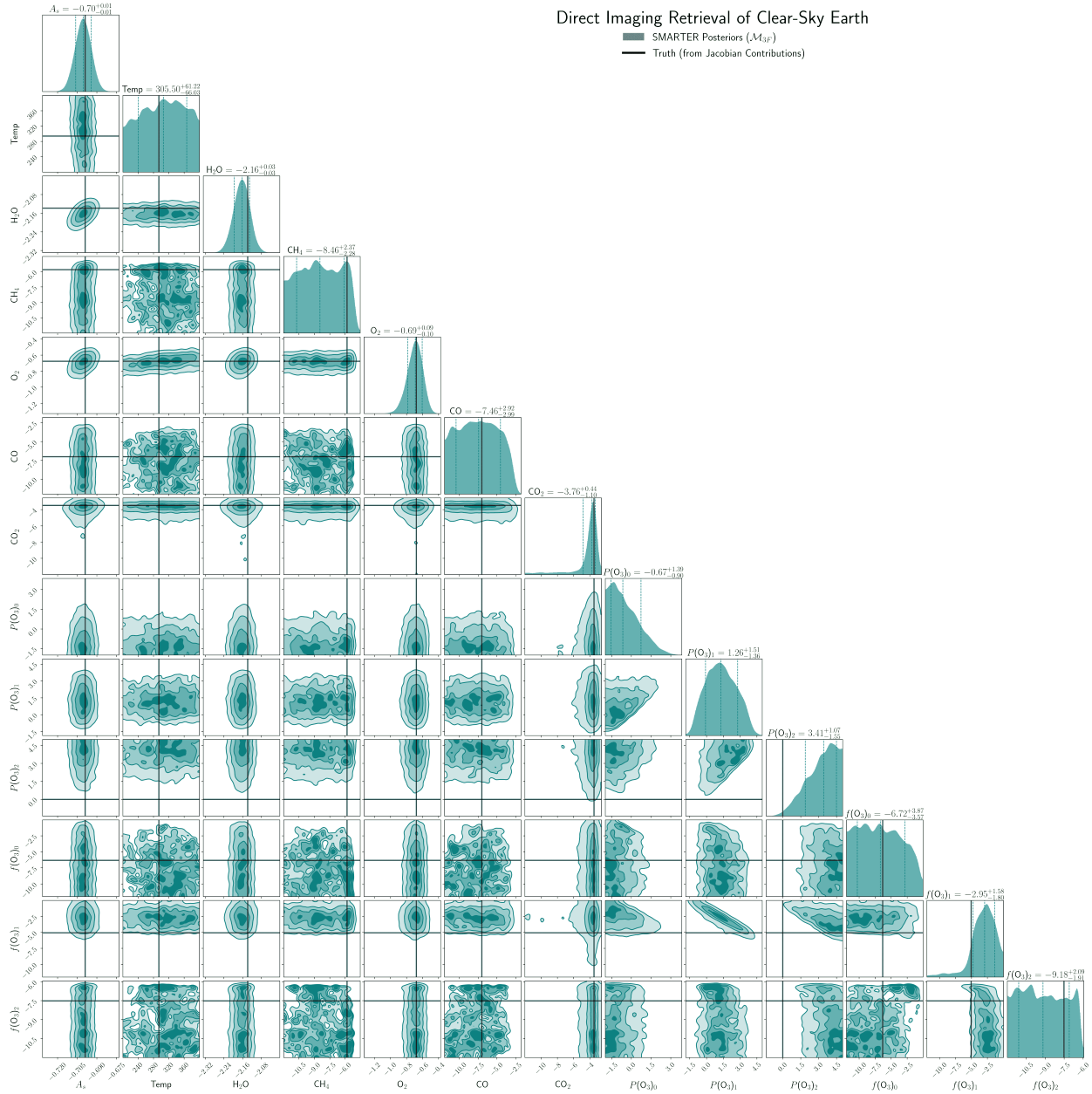


Figure 4.23: The results of the retrieval on the clear-sky Earth direct imaging spectrum with the three-free points forward model, $\mathcal{M}_{3F}$.

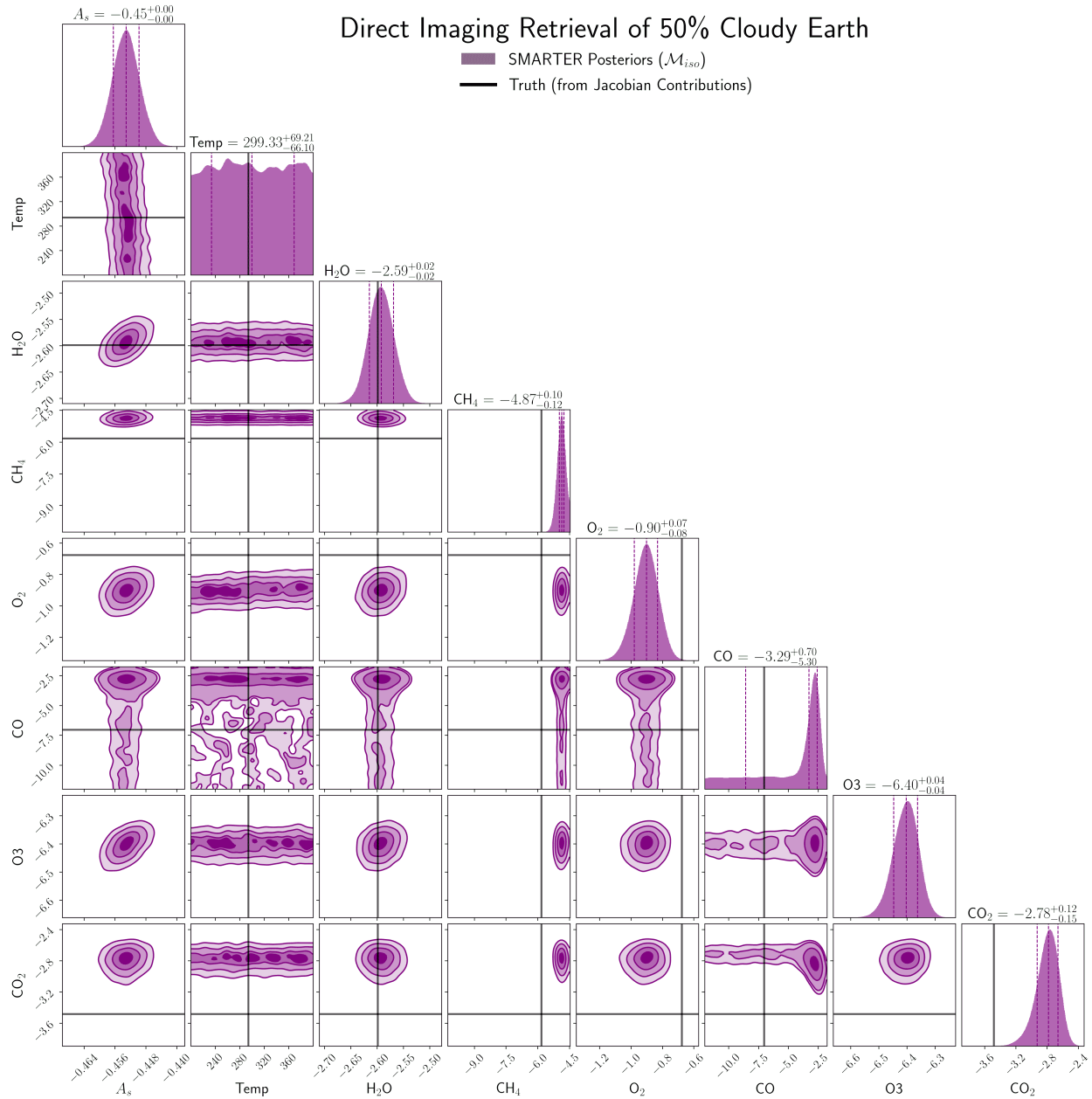


Figure 4.24: The results of the retrieval on the 50% cloudy (25% stratocumulus, 25% cirrus, 50% clear) Earth direct imaging spectrum with the evenly-mixed forward model, $\mathcal{M}_{iso}$.

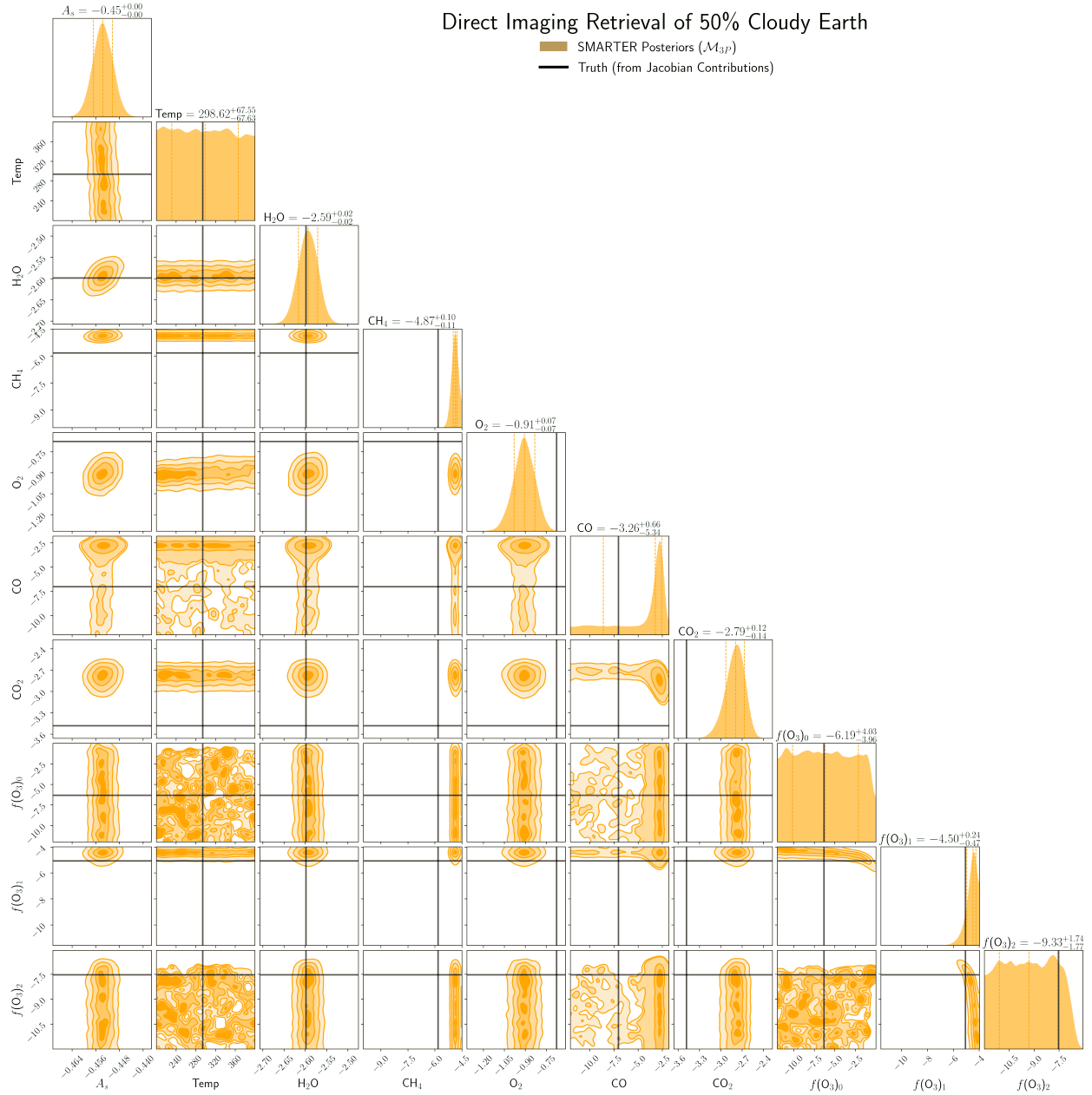


Figure 4.25: The results of the retrieval on the 50% cloudy Earth direct imaging spectrum with the three-fixed points forward model, $\mathcal{M}_{3P}$.

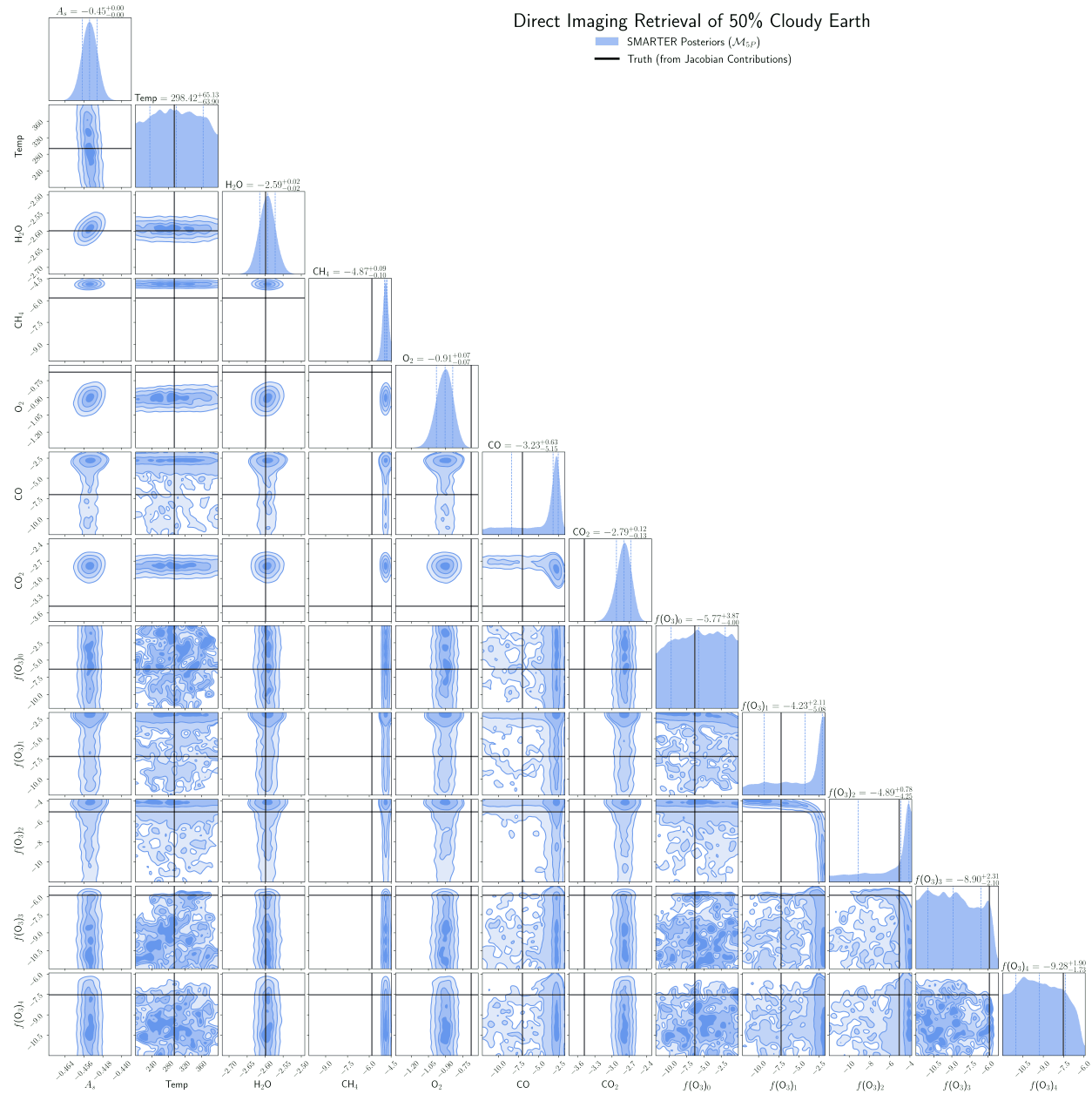


Figure 4.26: The results of the retrieval on the 50% cloudy Earth direct imaging spectrum with the five-fixed points forward model, $\mathcal{M}_{5P}$.

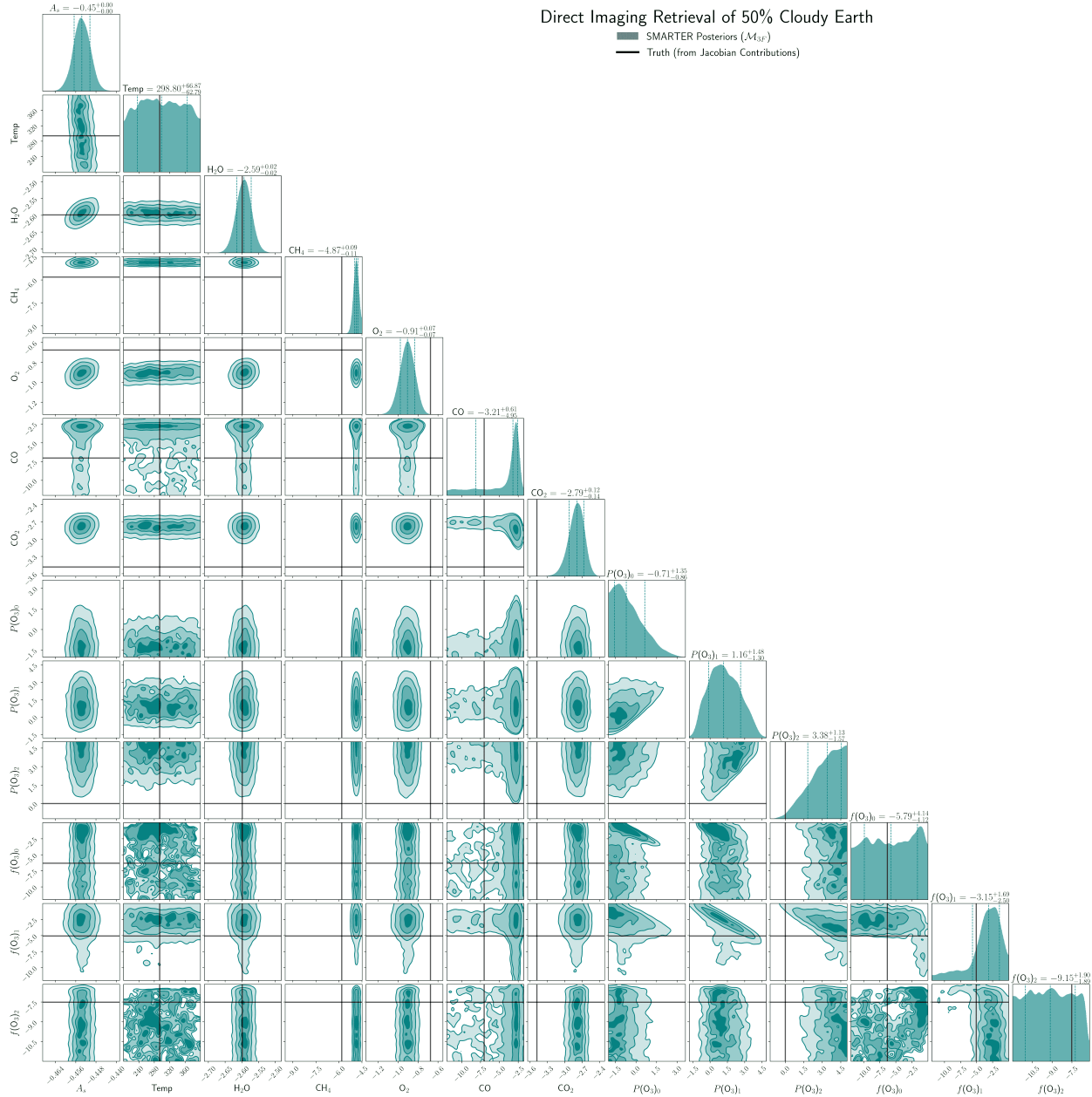


Figure 4.27: The results of the retrieval on the 50% cloudy Earth direct imaging spectrum with the three-free points forward model, $\mathcal{M}_{3F}$.

CHAPTER 5

THE EFFECT OF SPECTRAL RESOLUTION ON BIOSIGNATURE DETECTION VIA REFLECTED OBSERVATIONS OF THE EARTH THROUGH TIME

A version of this chapter has been accepted for publication in The Astrophysical Journal as: Gilbert-Janizek, S., et al., “Resolution Requirements for Life Detection via Reflected Light Spectroscopy of Rocky Exoplanets” (in press). It is reproduced here under the terms of the Creative Commons Attribution 4.0 license. J. Krissansen-Totton and J. Lustig-Yaeger jointly supervised this research; J. Lustig-Yaeger performed the noise calculations; both co-authors provided comments and edits on the manuscript; all other work is my own.

1	Methods	218
1.1	Atmospheric Cases	218
1.2	<code>rfast</code> Forward Model	218

1.3	pyEDITH Exposure Time Calculator	219
1.4	Analytic Relation for Calculating Detectability as a Function of Resolution	222
1.5	rfast Retrieval Model, Setup, & Cases	223
2	Results	227
2.1	Detectability as a Function of Resolution in the Visible and the Near-IR	228
2.2	Retrieval Cases	234
3	Discussion	239
3.1	The Benefits of High Resolution in the Visible Require Lower Detector Noise	240
3.2	The Resolution in the Near-IR Should be Greater than 20 to Avoid Retrieval Ambiguities and Enable CO ₂ Detections	242
3.3	Comparison to Previous Work	244
3.4	Future Work	246
4	Conclusion	250

In this study, we combine forward modeling with analytical detectability calculations and Bayesian retrieval simulations. We generate synthetic HWO observations across a range of spectral resolutions ($R = 20$ –5000 in the visible; $R = 20$ –1000 in the near-infrared) using realistic wavelength-dependent noise models. Our analytical approach efficiently explores the full resolution trade space by calculating minimum exposure times required for confident biosignature detections via Bayesian Information Criterion thresholds. We then validate and extend these findings through full atmospheric retrievals on select cases of particular interest, which reveal how resolution affects not just detectability but our ability to meaningfully constrain gas abundances, including critical upper limits on potential antibiosignatures. This dual approach allows us to answer the following time-critical questions: (1) Can higher visible spectral resolution compensate for the challenges of detecting low O_2 in Proterozoic atmospheres, either by reducing required exposure times or by providing an alternative to UV coronagraphy? (2) Can we reduce near-IR spectral resolution without compromising atmospheric characterization? (3) How do these trade-offs change across different stages of Earth’s atmospheric evolution?

In [section 1](#), we describe our atmospheric test cases spanning Earth’s history, the `rfast` radiative transfer model used to generate synthetic spectra, the `pyEDITH` exposure time calculator for realistic noise modeling, our analytical framework for calculating detectability thresholds, and our retrieval methodology. In [section 2](#), we present analytical calculations of exposure time versus resolution for all Earth-through-time cases, followed by detailed retrieval simulations for the Proterozoic and Phanerozoic (visible wavelengths) and Archean and Phanerozoic (near-infrared wavelengths). We discuss our findings in the context of previous studies and HWO mission design in [section 3](#), and conclude with our spectral resolution recommendations in [section 4](#).

1 Methods

1.1 Atmospheric Cases

We consider the same Earth-through-time cases as described by [Krissansen-Totton et al. \(2025\)](#), which are summarized in [Table 5.1](#). Accordingly, our study complements their investigation of the short-wavelength and long-wave cutoffs needed to accurately characterize the Earth through time. Mixing ratios are chosen based on literature estimates appropriate for each Earth epoch (See [Krissansen-Totton et al., 2025](#), Sec. 2.2 for a thorough description). Generally, our atmospheres are characterized by secularly decreasing CO_2 with $\text{CO} < \text{CH}_4$ due to biogenic CO consumption. Our Archean atmospheres have higher CH_4 abundances (0.5–0.05%) due to the predominance of methanogenic life. Our Proterozoic atmospheres have 0.2% O_2 abundances (1% PAL) due to the emergence of oxygenic life, which rapidly oxidized the atmosphere. The oxidized atmosphere readily destroys CH_4 , leading to lower abundances in the Proterozoic and Phanerozoic eons. For the Proterozoic and Archean epochs, we explore both “high” and “low” CH_4 scenarios. All atmospheres are assumed to be 1 bar.

1.2 `rfast` Forward Model

To calculate all reflected light spectra in this study, we use the `rfast` radiative transfer suite, a 1-D exoplanet atmosphere forward model from [Robinson & Salvador \(2023\)](#). To calculate molecular opacities based on the HITRAN2020 database ([Gordon et al., 2022](#)), `rfast` uses the Line-By-Line ABSorption Coefficients tool (LBLABC, [Meadows & Crisp, 1996](#)). The model also includes a blended ice/liquid water cloud model ([Robinson & Salvador, 2023](#)), which we apply by assuming 50% cloud coverage for all atmospheric cases. All spectra are initially calculated at a resolution of 5000 in the visible and 1000 in the near-IR wavelength ranges, and subsequently degraded to the resolution under investigation via a Gaussian

Table 5.1: Atmospheric composition and properties of Earth through different geological eons. For gases, abundances are reported as volume mixing ratios. For all cases, we simulate planets with $1.0R_{\oplus}$, $1.0M_{\oplus}$, and gray surface albedos of 0.1. Additionally, all atmospheres include 50% cloud coverage, where cloud decks are assumed to be 1×10^4 Pa thick with a cloud-top pressure of 6×10^4 Pa and extinction opacity of $\tau_c = 1$.

	Phanerozoic	Proterozoic (low CH ₄)	Proterozoic (high CH ₄)	Archean (low CH ₄)	Archean (high CH ₄)
N ₂	79.6%	79%	98.8%	80%	94.5%
O ₂	20%	0.2%	0.2%	None	None
O ₃	7×10^{-7}	7×10^{-8}	7×10^{-8}	None	None
CO ₂	0.1%	1%	1%	20%	5%
CO	0.1 ppm	0.1 ppm	0.1 ppm	0.005%	0.05%
CH ₄	1 ppm	5 ppm	100 ppm	0.05%	0.5%
H ₂ O	0.3%	0.3%	0.3%	0.3%	0.3%
Surface Pressure (bar)	1	1	1	1	1

convolution kernel. Based on the conclusions of [Krissansen-Totton et al. \(2025\)](#), we adopt a long-wavelength cutoff in the near-IR of 1.7μ . In contrast to [Krissansen-Totton et al. \(2025\)](#), we adopt a short-wavelength cutoff in the visible of 0.4μ , due to the current UV modeling limitations of our noise model (see below).

1.3 pyEDITH Exposure Time Calculator

We use the Python-based Exposure Direct Imaging Timer for HWO (pyEDITH; [Aleí et al., 2025](#)) to perform noise calculations that are applicable to the current working design of the HWO coronagraph instrument. The pyEDITH exposure time calculator (ETC) has heritage from the ETC of [Stark et al. \(2014\)](#), but represents a new standalone tool for developing and refining exoplanet detection and characterization science capabilities for HWO. pyEDITH features the ability to load EAC YAML files released by the HWO project office that specify the observatory and instrument parameters for each Exploratory Analytic Case (EAC)¹. For more details on pyEDITH, please refer to [Aleí et al. \(2025\)](#) and the online code documentation².

¹<https://github.com/HWO-GOMAP-Working-Groups/Sci-Eng-Interface>

²<https://pyedith.readthedocs.io/en/latest/index.html>

In this work, we use the EAC1 observatory preset, which is provided with the initial release of `pyEDITH`, and we use the IFS observing mode to simulate planetary reflected light spectroscopy using the coronagraph. We provide a list of relevant parameters in [Table 5.2](#) that are used throughout this study unless otherwise noted. As in recent work, we assume that the nominal dark current is $3 \times 10^{-5} \text{ e}^-/\text{pixel}/\text{second}$ ([Stark et al., 2025](#)), however, we explore a broader range of values in [section 2](#) to investigate how detector noise trades with spectral resolution. We omit UV noise modeling from our investigations due to the limited capability of `pyEDITH` to accurately simulate UV observations at the time of writing. Finally, we run `pyEDITH` noise calculations for each of the follow spectral resolving powers: $R_{\text{Vis}}=30, 50, 70, 100, 140, 200, 500, 1000, 2000, 3000, 4000$, and 5000 , and $R_{\text{NIR}}=20, 40, 50, 60, 70, 80, 100, 140, 200, 500$, and 1000 . These fiducial calculations are for an Earth at 10 pc observed with HWO for 10 hours. We rescale the exposure time analytically throughout this work to investigate the relationship between spectral resolution and the minimum required exposure time to detect biosignature gases and habitability indicators in specific bands of interest.

To test the validity of re-scaling a single noise model output, we generated noise instances at the nominal resolution for exposure times $t_{\text{exp}} = 5, 100$, and 1000 hours, and compared the resulting SNR of our re-scaled exposure times to the truth. We found that our analytical relationship accrues errors in the absolute exposure time no larger than 20% relative to the regenerated noise instances. The largest exposure time errors are for the shortest exposure times where fixed overhead due to slew, settling, and digging the dark hole take a significant fraction of the total observation. Cases with longer than our nominal 10 hour exposure time minimize the effect of overheads and approach our analytic calculations. Therefore, while these effects are minimal, we caution that our reported absolute exposure times are less accurate than the relative differences between our various cases. Given that the HWO architecture is not yet set, we argue that relative exposure time differences are the important figure of merit, and as such we focus our interpretation of results on them in [section 3](#).

Table 5.2: Parameters used in the pyEDITH exposure time calculations for the EAC1 observatory preset. All noise instances are generated for a constant exposure time of $t_{\text{exp}} = 10$ hours, and subsequently re-scaled as needed. Where units are not included, parameters are dimensionless.

Parameter	Value	Meaning
Observation		
full wavelength range [μm]	[0.4–1.7]	Full wavelength range of the observations
visible wavelength range [μm]	[0.4–1.0]	Range of wavelengths in the VIS
near-IR wavelength range [μm]	[1.0–1.7]	Range of wavelengths in the NIR
exposure_time [hours]	10	Exposure time for the fiducial observation
CRb_multiplier	2	Factor α for PSF subtraction method (2 for Angular Differential Imaging)
distance [pc]	10	Distance d of the system
stellar_radius [$R_{\odot}$]	1	Stellar radius
nzodis [zodi]	3	Exozodi multiplier
ra	236.01°	Right ascension
dec	2.52°	Declination
angular_separation [arcsec]	0.1	Separation (at d)
Telescope		
diameter [m]	7.226	Circumscribed diameter D_{circ} of telescope aperture
unobscured_area [%]	100.0	Percentage of unobscured collecting area
toverhead_multi	1.100	Multiplicative overhead time τ_{multi} (refining the dark hole)
toverhead_fixed [s]	8250	Static overhead time τ_{static} (slew, settling, digging the dark hole)
telescope_optical_throughput	0.822	Throughput of the telescope optics
temperature [K]	290	Temperature of the optics
T_contamination	1.000	Effective throughput due to contamination
Coronagraph		
Istar	3.076×10^{-16}	Stellar intensity scalar (I_{star} , on-axis PSF), product of the noise floor contrast of the coronagraph (uniform over dark hole) and the peak value of the off-axis PSF
noisefloor	1.025×10^{-17}	Noise floor (NF) of the coronagraph
skytrans	0.798	Coronagraph’s performance when observing an infinitely extended source
coronagraph_optical_throughput	0.438	Throughput of all coronagraphic optics
Detector		
DC [e^- pixel $^{-1}$ s $^{-1}$]	3.000×10^{-5}	Dark current
RN [e^- pixel $^{-1}$ read $^{-1}$]	0	Read noise
tread [s]	1000	Time between reads
CIC [e^- pixel $^{-1}$ ph $^{-1}$]	0	Clock-induced charge
QE	0.750	Quantum efficiency
dQE	0.750	Effective QE due to degradation

1.4 Analytic Relation for Calculating Detectability as a Function of Resolution

We derive an analytic relation for the minimum exposure time required to confidently detect a specific atmospheric gas in an exoplanet spectrum. Our approach uses the Bayesian Information Criterion (BIC) to compare two nested models: one including all gases (model 2) and one excluding the target gas (model 1).

Starting from the standard BIC formulation, where $k = 14$ is the number of model parameters for all cases and N is the number of spectral data points, we require $\Delta\text{BIC} \geq 10$ as our threshold for a gas detection, corresponding to strong Bayesian evidence favoring the complete model. N changes depending on the resolution under investigation. The standard χ^2 metric quantifies data-model goodness-of-fit given the uncertainty, σ . The key insight is how χ^2 scales with observation quality and in turn exposure time. If we assume that HWO observations are in the photon-noise limit, spectral precision scales as $\sigma_1 \propto 1/\sqrt{t}$, so χ^2 is directly proportional to exposure time t . That is, a model that omits a key atmospheric species that is present in the data (model 1) will see a linear increase in χ^2 (a decrease in goodness-of-fit) as exposure time increases.

For model 2 (all gases included), we assume perfect model-data agreement in the absence of random noise, giving $\chi_2^2 = 0$. This means the “data” is effectively the spectrum with all gases included. Model 1 (missing the target gas) will show residuals proportional to that gas’s spectral signature. By expressing uncertainties at a new exposure time t_2 relative to a nominal exposure time t_1 , we have $\sigma_2^2 = \sigma_1^2 \times (t_1/t_2)$. Substituting into ΔBIC and solving for t_2 , we obtain:

$$t_2 = \frac{\Delta\text{BIC}_0 - k \ln(N) + (k-1) \ln(N)}{\left(\sum_{i=1}^N \frac{(\text{data}_i - \text{model}_{1,i})^2}{\sigma_{1,i}^2 \times t_1} - \sum_{i=1}^N \frac{(\text{data}_i - \text{model}_{2,i})^2}{\sigma_{1,i}^2 \times t_1} \right)}. \quad (5.1)$$

Finally, because $\chi_2^2 = 0$ the second sum vanishes, simplifying to:

$$t_2 = \frac{\Delta\text{BIC}_0 - k \ln(N) + (k - 1) \ln(N)}{\left(\sum_{i=1}^N \frac{(\text{data} - \text{model}_1)^2}{\sigma_{1,i}^2 \times t_1} \right)}. \quad (5.2)$$

This expression directly relates the minimum required exposure time to the spectral signature strength of the target gas, providing a computationally efficient way to assess gas detectability in the visible and near-IR wavelength regimes. Accordingly, we generate `pyEDITH` noise instances for each resolution case at a nominal exposure time $t_1 = 10$ hours, and use our analytical relationship to solve for the re-scaled exposure time t_2 that yields a $\Delta\text{BIC} \geq 10$ strong detection of the missing gas. The resulting exposure time t_2 represents the minimum length of observation required to make a significant detection of a key gas in the visible or near-IR wavelength range for a given Earth through time case. We recalculate the minimum exposure time for each gas in each wavelength range to reflect the relative significance of spectral absorption features spanning both wavelength regimes. This is particularly applicable to H_2O , which has broad absorption features in both the visible and the near-IR and O_2 , which has comparatively narrow absorption features at 0.76μ in the visible and at 1.27μ in the near-IR.

1.5 `rfast` Retrieval Model, Setup, & Cases

Retrievals enhance our interpretations of the analytical calculations by demonstrating how spectral resolution affects our ability to *constrain* the abundances of key gases. This is particularly important for gases that we do not anticipate to be detectable, as the analytical detectability calculations inherently cannot show the nuances of potential changes to abundance upper limits. When it comes to ruling out the presence of antibiosignature gases, an upper limit can be significant to contextualize the detection of key gases by confirming a biogenic origin. This holds especially true for gases like CO where unrealistically long exposure times are required for detections. Retrievals are the appropriate tool of choice for tracking changes to upper limits on gases with changes to instrument design ([Krissansen-Totton et al.](#),

2025).

We perform retrievals with `rfast`, a 1-D atmospheric retrieval code from [Robinson & Salvador \(2023\)](#) that explores parameter space using the `emcee` Markov Chain Monte Carlo (MCMC) sampler ([Foreman-Mackey et al., 2013](#)). `rfast` enables the retrieval of atmospheric, planetary, and orbital parameters via uniform or Gaussian priors in either log or linear space. In this study, all 14 parameters are retrieved in log 10 space assuming uniform priors. We retrieve seven unknown gas partial pressures, including N_2 , O_2 , H_2O , O_3 , CO_2 , CH_4 , and CO . For N_2 , H_2O , CO_2 , and CH_4 , we impose a log 10 uniform prior ranging from 10^{-2} – 10^7 Pa. For the remaining gases, O_2 , O_3 , and CO , which may be trace species, we adopt a broader log 10 uniform prior ranging from 10^{-12} – 10^7 Pa. We also retrieve surface albedo (prior range 0.01–1.0), planet radius (prior range $10^{-0.5}$ – $10^{0.5} R_\oplus$), planet mass (prior range 0.1 – $10 M_\oplus$), cloud deck thickness (prior range 1 – 10^7 Pa), cloud-top pressure (prior range 1 – 10^7 Pa), cloud extinction optical thickness (prior range 10^{-3} – 10^3), and global cloud fraction (prior range 10^{-3} – 1). To ensure clouds form only above the surface, cloud top pressures smaller than the total atmospheric pressure are prohibited. For all cases a gray surface albedo is assumed.

We set up our retrievals with 100 walkers within `emcee`, where each walker takes 100,000 steps with a burn-in of 10,000 and a thinning parameter of 100 to compress retrieval outputs. To accelerate MCMC convergence, we initialize all walkers in a Gaussian “ball” around the true values, after which all walkers are allowed to explore the entire prior range. We verify that `emcee` has converged by confirming the chain lengths for all parameters exceed $50 \times \tau$, where τ is the autocorrelation time ([Foreman-Mackey et al., 2013](#)).

As in previous studies ([Hall et al., 2023](#); [Krissansen-Totton et al., 2025](#)), we retrieve gas abundances as partial pressures to avoid assuming which gas is the bulk atmospheric constituent. This method is comparable to imposing center log ratio priors for atmospheric abundances ([Benneke & Seager, 2012](#); [Damiano & Hu, 2021, 2022](#)). Since the total atmospheric pressure is the sum of the retrieved gas partial pressures, surface pressure itself is

not explicitly retrieved.

We summarize the retrieval cases included in this study in [Table 5.3](#). We generated unique noise instances at every resolution case for a fixed exposure time of 10 hours. Then, we re-scaled the SNR of the entire spectrum by reducing the wavelength-dependent error bars by a constant factor. The constant factor was chosen to give near-band SNRs consistent with values in the literature. However, as shown in [Figure 5.1](#), the SNR varies significantly with wavelength due to the wavelength-dependent photon flux from the planet and instrument characteristics. This steep gradient in signal-to-noise makes direct comparisons with prior works that assume flat SNR challenging. For instance, an SNR of 20 in the continuum near $\sim 1.1 \mu$ corresponds to $\text{SNR} \lesssim 5$ at $1.5\text{--}1.75 \mu$ ([Figure 5.1b](#)). This wavelength dependence is crucial for understanding detectability across the full spectral range.

Table 5.3: A summary of the retrieval cases explored in this study. For the visible-focused retrievals ($0.4\text{--}1.0 \mu\text{m}$), we examine the Proterozoic and Phanerozoic cases with resolutions of 140, 200, 500, and 1000, while maintaining the near-IR resolution at the nominal value of 70 (in bold). For the near-IR-focused retrievals ($1.0\text{--}1.7 \mu\text{m}$), we examine the late Archean (high CH_4) and Phanerozoic cases with resolutions of 20, 40, 50, 60, and 70, while maintaining the visible resolution at the nominal value of 140 (in bold). We include the SNR_0 at λ_0 used to re-scale the spectral uncertainties, and the corresponding constant exposure time.

Earth through Time Case	Resolutions ($\lambda/\Delta\lambda$)	λ_0 (μm)	SNR_0	Exposure Time (hours)
Archean	$R_{\text{NIR}} = 20, 40, 50, 60, \mathbf{70}$	1.6	40	8163
Proterozoic	$R_{\text{Vis}} = \mathbf{140}, 200, 500, 1000$	0.76	20	82
Phanerozoic	$R_{\text{Vis}} = \mathbf{140}, 200, 500, 1000$	0.76	20	82
	$R_{\text{NIR}} = 20, 40, 50, 60, \mathbf{70}$	1.6	40	8245

We perform retrievals on a set of spectra with differing resolving powers, but the same exposure time to isolate the information gain/lost strictly based on resolution changes. We use the nominal resolution case to determine exposure time scale factors that result in achieving some SNR_0 at λ_0 (see [Table 5.3](#)) based on past work targeting specific molecular bands that typically drive Earth-like characterization. For a given Earth-through-time case, once we determine the noise rescaling factor from the nominal resolution case ($R_{\text{Vis}} = 140$, $R_{\text{NIR}} = 70$), we apply this same scaling factor to all resolution cases. This approach isolates the physical

effect of resolution changes on SNR (while maintaining identical exposure times between resolution cases): when resolution increases, photons are distributed across more spectral bins, reducing the SNR per bin; conversely, when resolution decreases, photons are concentrated into fewer bins, increasing the SNR per bin.

For the cases where we investigate the effect of increasing visible resolution, we scale the SNR of the nominal resolution spectrum for each retrieval case so that there is an $\text{SNR} \geq 8.5$ at the continuum of the O_2 A-band feature at 0.76μ (Feng et al., 2018). The re-scaled $\text{SNR} \sim 20$ for the visible cases corresponds to a constant exposure time of $t_{\text{exp}} \sim 82$ hours. For the cases where we investigate the effect of decreasing the resolution in the near-IR, we scale the SNR of the nominal resolution spectrum so that there is an $\text{SNR} \geq 20$ in the continuum at the CO_2 feature at $\sim 1.5 \mu$ (Krissansen-Totton et al., 2025). The re-scaled $\text{SNR} \sim 40$ for the near-IR cases corresponds to a constant exposure time of $t_{\text{exp}} \sim 8200$ hours. The substantially longer near-IR exposure times arise because the SNR per unit time is significantly lower in this wavelength range – the planet reflects fewer near-IR photons, and detector thermal noise contributions are higher compared to visible wavelengths.

For the visible wavelength investigations, we selected one Proterozoic case to determine whether increased spectral resolution reduces the exposure time required to detect O_2 in low- O_2 Earth-through-time atmospheres. We also examined the Phanerozoic case to establish a reference baseline, as modern Earth represents arguably the most compelling science case for HWO. For the near-IR investigations, we selected the high CH_4 late Archean case, as the elevated abundances of CH_4 and CO_2 during this epoch are expected to produce stronger spectral features compared to the Proterozoic and Phanerozoic cases. We additionally conducted retrievals for the Phanerozoic case in the near-IR to provide a consistent baseline for comparison across wavelength regimes. The near-IR resolution investigation aims to assess whether reduced spectral resolution adversely impacts atmospheric characterization capabilities. Our analytic detectability calculations indicate that exposure time requirements for detecting key gases in the near-IR remain relatively insensitive to resolution degradation,

suggesting potential engineering advantages for instrument development at lower near-IR resolving powers.

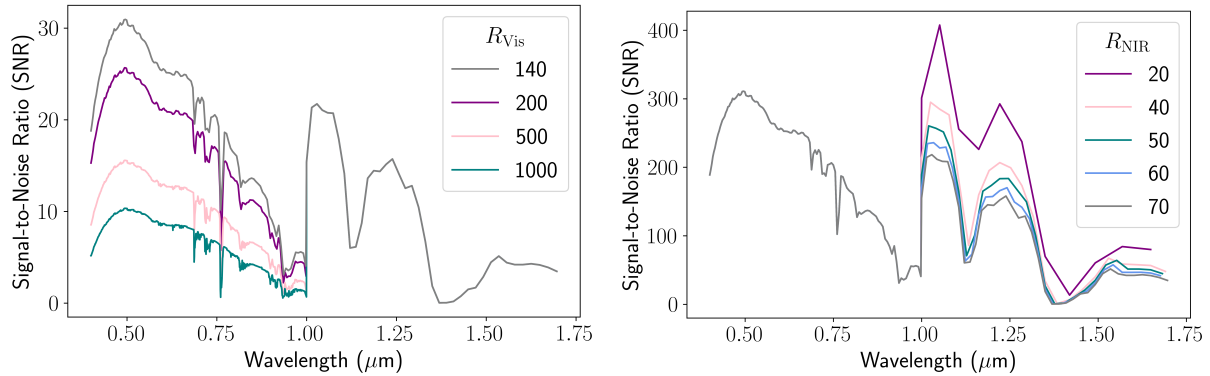


Figure 5.1: SNR as a function of wavelength for the visible and near-infrared resolution retrieval investigations using the exposure time calculator. When visible resolution is under investigation, the resolution in the near-infrared wavelengths is held at the nominal value (and vice versa when the near-infrared resolution is under investigation). Increased resolution leads to decreased SNR, while decreased resolution leads to increased SNR. Steep gradients in the signal-to-noise make it challenging to do an apples-to-apples comparison with prior works that use flat signal-to-noise. For example, for an SNR of 20 in the continuum near $\sim 1.1 \mu\text{m}$, the SNR at $1.5\text{--}1.75 \mu\text{m}$ is <5 .

2 Results

We report the results of our investigation in the form of analytic detectability calculations and posterior distributions from atmospheric retrievals. For the former, we calculate the minimum exposure time required to make detections of key gases for our Earth through time cases as a function of resolution in the visible and near-IR wavelength regimes. We also assess how the magnitude of the dark current affects this minimum exposure time as a function of resolution. Finally, we report an acceptable range of instrument resolutions for detecting each gas in each Earth through time case using an exposure time threshold of 20% from the minimum, which represent the size of the errors in our detectability calculations. For our retrieval cases investigating the effect of increased resolution in the visible wavelength regime, we analyze gas abundance posterior distributions retrieved for the high- CH_4 Proterozoic case

Table 5.4: Retrieval parameters and prior ranges. Pressure is not an independent free parameter as it is the sum of the constituent partial pressures (partial pressures are retrieved parameters). Parameters without explicit units reported are non-dimensional.

Parameter	Prior range
$p\text{N}_2$	$[10^{-2}, 10^7]$ Pa
$p\text{O}_2$	$[10^{-12}, 10^7]$ Pa
$p\text{O}_3$	$[10^{-12}, 10^7]$ Pa
$p\text{CO}_2$	$[10^{-2}, 10^7]$ Pa
$p\text{CO}$	$[10^{-1}, 10^7]$ Pa
$p\text{CH}_4$	$[10^{-2}, 10^7]$ Pa
$p\text{H}_2\text{O}$	$[10^{-2}, 10^7]$ Pa
Surface albedo (A_s)	$[0.01 - 1.0]$
Radius (R_p)	$[10^{-0.5} - 10^{0.5}] R_\oplus$
Mass (M_p)	$[10^{-1} - 10^1] M_\oplus$
Cloud thickness (Δp_c)	$[10^0 - 10^7]$ Pa
Cloud top pressure (p_t)	$[10^0 - 10^7]$ Pa
Cloud extinction (τ_c)	$[10^{-3} - 10^3]$
Cloud fraction (f_c)	$[10^{-3} - 10^0]$

and the Phanerozoic Earth case. To assess the effect of decreased resolution in the near-IR, we show gas abundance posterior distributions retrieved for the high- CH_4 Archean case and the Phanerozoic Earth case.

2.1 Detectability as a Function of Resolution in the Visible and the Near-IR

Figure 5.2 summarizes the relationship between exposure time and resolution for detecting key gases in relevant band-passes for all the Earth through time cases considered here. We indicate the nominal resolution with a vertical dotted line ($R_{\text{Vis}} = 140$ and $R_{\text{NIR}} = 70$ for the visible and near-IR, respectively), while we show our inferred “optimal” resolution, i.e. that which gives the minimal exposure time, with a solid vertical line. The top row shows the exposure time-resolution relationship for gases in the visible, including O_2 , H_2O , and O_3 . The bottom row shows the same relationship for gases in the near-infrared, including CO_2 , H_2O and CH_4 . The vertical offset between the different Earth through time cases occurs

primarily due to differences in abundances. For example, detecting O_2 and O_3 in the visible for both Proterozoic atmospheres requires exposure times several orders of magnitude longer than for the Phanerozoic due to the extremely low oxygen abundances in these atmospheres. Similarly, the exposure time required to detect H_2O in the near-IR is an order of magnitude larger for the late Archean case, as the higher CH_4 abundances lead to strong, overlapping absorption features with broad water vapor bands.

Across all Earth through time cases, the gases in the visible and the near-IR demonstrate a similar pattern for exposure time as a function of spectral resolution. We find three resolution regimes that are associated with under-resolving molecular bands, optimally matching the resolution to the width of the bands, and over-resolving the bands. At insufficiently low resolutions, absorption bands are under-resolved and weakened by blending with the continuum, which requires additional exposure time to detect a given gas. At excessively high resolutions, absorption bands are over-resolved and many resolution elements (and pixels) span their widths, which also requires additional exposure time to build up the necessary signal to noise to detect a given gas due to per-pixel detector noise. We define the “optimal resolution” as that which balances these two effects and minimizes the exposure time. We note the exposure time-resolution relationship for gases in the near-IR is relatively flat in the vicinity of the nominal resolution, and exposure time changes relatively little when resolution is slightly degraded or enhanced.

We identify the optimal resolutions for each gas and band-pass via the solid vertical lines in [Figure 5.2](#). Thus, our results suggest that shorter exposure times could be achieved for detecting O_2 in all Earth through time cases by increasing resolution in the visible wavelengths to $R_{\text{vis}} = 500$. By the same token, H_2O and O_3 appear to be over-resolved in the visible bands, as the nominal resolution exceeds the optimal resolution. However, this is not surprising, given that the resolution in the visible is set not by the detection of these gases, but by the detection of the narrow-band O_2 feature. In the near-IR, our results suggest that shorter exposure times could be achieved for detecting CO_2 and CH_4 by increasing

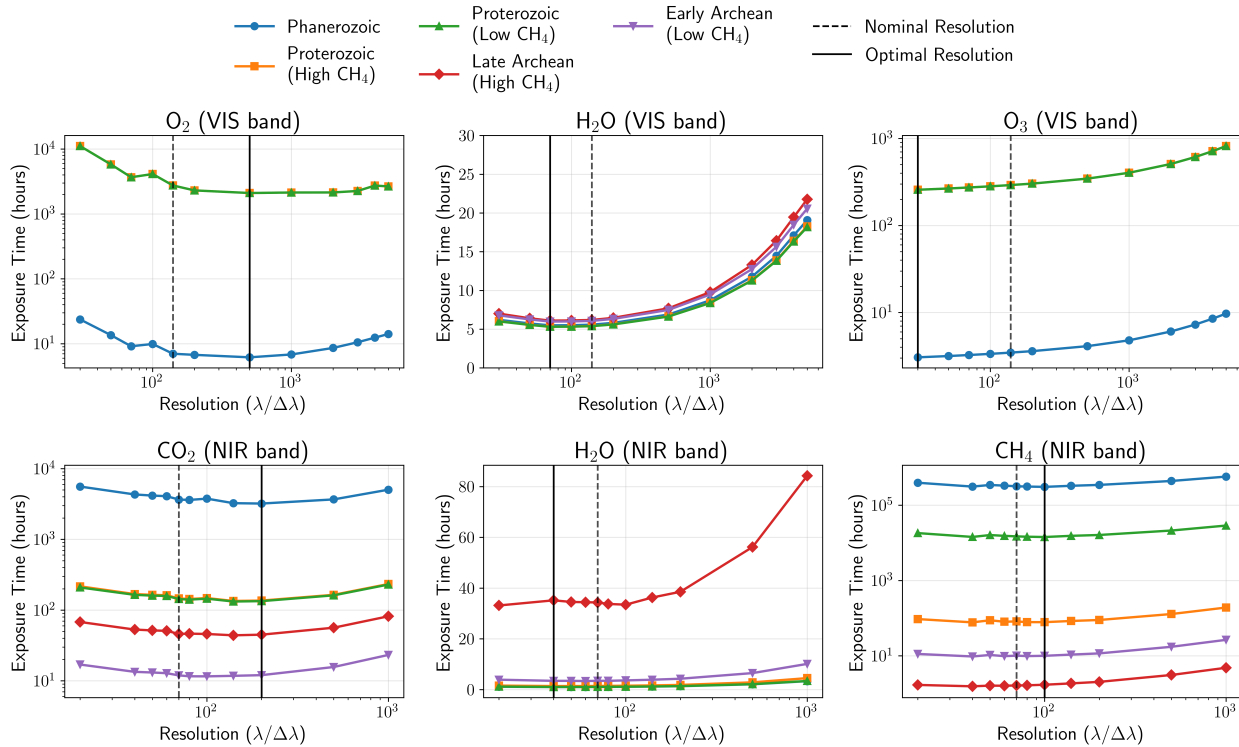


Figure 5.2: Exposure time required for gas detection as a function of resolution for key gases in corresponding band-passes for all Earth through time cases explored in this study, assuming the nominal dark current (3×10^{-5} e-/pixel/s) throughout. Here, “VIS” (top row) is shorthand for the visible band-pass, and “NIR” (bottom row) is shorthand for the near-infrared band-pass. The dashed vertical lines indicate the nominal HWO resolutions in the visible and the near-IR, while the solid vertical lines indicate the “optimal” resolution that gives the minimum exposure time according to our analytical calculations. Despite intrinsic band differences, our results show a generic trend of under-resolving bands where increased resolution decreases the exposure time, and over-resolving bands where increased resolution increases the exposure time by incurring excess pixel noise. In the visible, water, and ozone are broad and have a tendency to be over-resolved for most resolutions considered, whereas the more narrow-band O_2 feature is more likely to be under-resolved.

the spectral resolution to $R_{\text{NIR}} = 200$ and $R_{\text{NIR}} = 100$, respectively. Similar to the visible wavelengths, H_2O remains slightly over-resolved.

While the summary plots in [Figure 5.2](#) appear to show that the “optimal” resolution does not overlap with the current nominal instrument resolutions, point estimates of a single optimal resolution are inherently limited. Realistically, future missions must remain capable across a broad range of possible atmospheres, molecular species, and science cases that we are not fully considering here, but may need to be balanced during the design stage. Resolution ranges that deliver performance within a reasonable threshold of the minimum exposure time are therefore more meaningful and easier to incorporate in downstream design decisions than any single optimum value. Accounting also for the 20% uncertainty in the signal-to-noise scaling calculation, we identify the practical range of resolutions required to detect key gases within a 20% threshold of the minimum possible exposure time in hours. We show these calculated ranges for all Earth through time cases in [Table 5.5](#), with key gases broken down by band-pass. When we account for uncertainty in the table, nearly all acceptable resolution ranges encompass the nominal visible and near-IR resolutions. Our results suggest that only for detecting O_2 in the Proterozoic would it be advantageous to go to slightly higher resolutions in the visible of 171, which is about 20% higher than the nominal visible resolution of 140. However, we find that these gains are likely minimal since they still require very long exposure times of ~ 2000 hours. In the near-IR, our results suggest that resolutions lower than nominal may also be sufficient for characterizing the Earth through time, particularly for Archean-like CH_4 and CO_2 abundances. Overall, our detectability calculation results suggest that the nominal resolutions in both the visible and the near-IR are likely suitable for detecting key gases. Moreover, our calculations suggest that these resolutions need not be strictly prescriptive – there is some flexibility in the tradeoff between resolution and exposure time, and the nominal resolutions may be considered approximate targets within our provided ranges for instrument design considerations seeking to deliver the most scientifically capable telescope.

Our detectability calculation results shown in [Figure 5.2](#) and [Table 5.5](#) assume a nominal dark current value of $3 \times 10^{-5} \text{ e}^-/\text{pixel}/\text{s}$. [Figures 5.3](#) and [5.4](#) demonstrate how varying dark current affects the relationship between resolution and exposure time in the visible band-pass for the Proterozoic (high CH_4) case and the Phanerozoic case, respectively. For each case, we show the effect on the detectability of H_2O , O_2 , and O_3 . On each subplot, the star denotes the case corresponding to the nominal dark current and nominal visible resolution ($R_{\text{Vis}} = 140$), and the contours show combinations of dark current and resolution that give equivalent exposure times. We also show the noise ($1 \times 10^{-3} \text{ e}^-/\text{pixel}/\text{s}$) and spectral resolution ($R = 50$) of the Nancy Grace Roman Telescope’s Coronagraph Instrument (CGI), demarcated as a triangle on the plots. CGI will also provide spectroscopic observations in the visible wavelengths, providing a useful baseline for comparison of current and future coronagraph technologies.

The non-smooth features visible in the contour plots (e.g., at $R_{\text{Vis}} \approx 100$ and $R_{\text{Vis}} \approx 3000$ – 4000) are numerical artifacts arising from pixel-phase aliasing on the discrete ETC wavelength grid, wherein narrow absorption band cores are comparable in width to individual pixels and $\Delta\chi^2$ varies non-monotonically depending on whether pixel centers happen to align with the deepest part of the absorption features as resolution changes. We confirm this interpretation by rigidly shifting the ETC wavelength grid by sub-pixel increments and recomputing $\Delta\chi^2$, finding that the resulting variability in detection metrics is largest where contour irregularities appear and smallest where contours are smooth, consistent with pixel-phase aliasing rather than any physical discontinuity in detectability. More broadly, this sensitivity suggests that for high-priority narrow-band features such as O_2 that are at risk of being under-resolved, instrument design teams should exercise caution to ensure that resolution elements are centered on the band core rather than straddling it, as misalignment between pixel boundaries and narrow absorption features could meaningfully degrade detection significance.

[Figures 5.3](#) and [5.4](#) demonstrate that HWO’s projected dark current represents over an

order of magnitude improvement over CGI. For extremely low-noise detectors with dark current below $1 \times 10^{-6} e^-/\text{pixel}/\text{s}$, higher resolutions grant gas detections in shorter exposure times. More generally, there is a resolution threshold above which dark current must be suppressed to avoid accruing increased exposure time. While higher resolution in the visible reduces the required exposure time to detect O_2 in low- O_2 Proterozoic atmospheres, the minimum exposure time is still on the order of thousands of hours. Thus, any gains from increased spectral resolution in the optical are likely irrelevant based on this case alone. For modern-Earth O_2 in the Phanerozoic, decreasing the visible resolution from the nominal would rapidly increase exposure times regardless of dark current, while increasing resolution would yield comparable exposure times up to $R_{\text{Vis}} \approx 1000$, after which lower dark current detectors would be required to prevent exposure time increases. However, because H_2O is already over-resolved at the nominal resolution, increasing the resolution up to $R_{\text{Vis}} \approx 1000$ would increase exposure times to detect H_2O by $\sim 50\%$ unless detector noise is decreased by nearly an order of magnitude. Thus, an optimal instrument design for detecting one molecule may be suboptimal for another molecule.

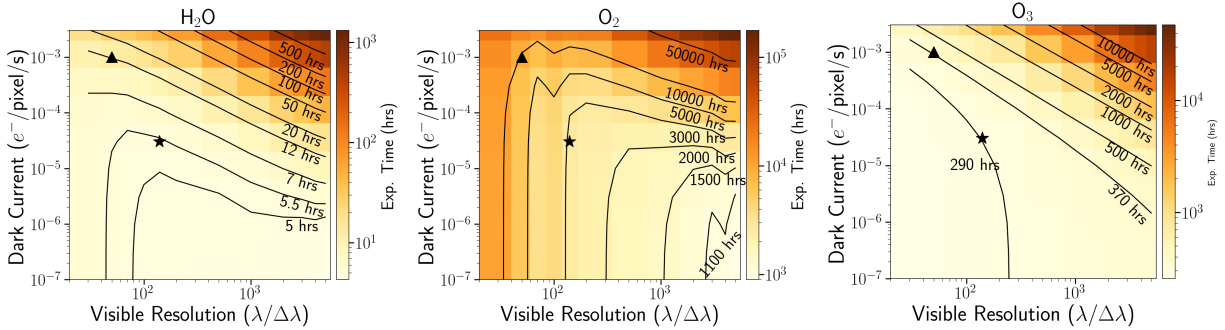


Figure 5.3: Contour plots showing the relationship between dark current, resolution in the visible band-pass, and exposure time required to detect (a) H_2O , (b) O_2 , and (c) O_3 for the high CH_4 Proterozoic Earth. The star shows the nominal values assumed for HWO EAC1 in `pyEDITH`, while the triangle shows the nominal resolution ($R \sim 50$) and expected dark current ($1 \times 10^{-3} e^-/\text{pixel}/\text{s}$) of the Nancy Grace Roman Telescope’s Coronagraph Instrument Spectrograph, which covers comparable wavelengths in the optical. While higher spectral resolution in the visible correlates with lower relative exposure times to detect O_2 at low Proterozoic abundances, it remains infeasible to detect in a reasonable amount of time for all cases, making such gains largely irrelevant.

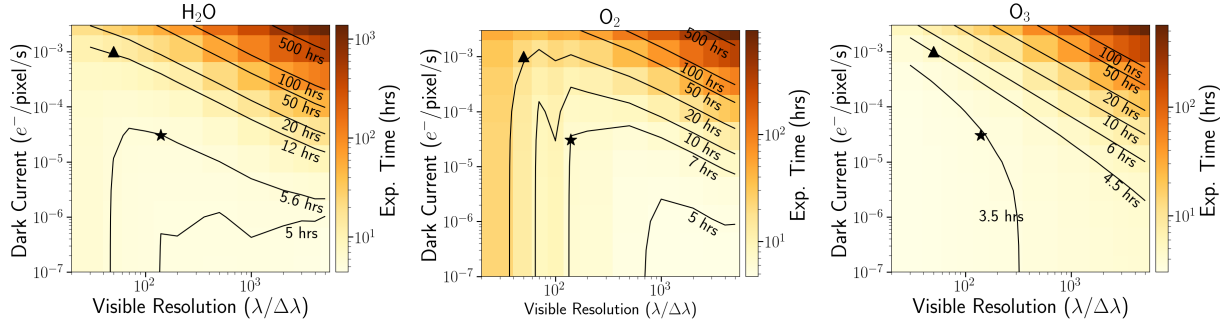


Figure 5.4: Contour plots showing the relationship between dark current, resolution in the visible band-pass, and exposure time required to detect (a) H_2O , (b) O_2 , and (c) O_3 for the Phanerozoic Earth. The star shows the nominal values assumed for HWO EAC1 in `pyEDITH`, while the triangle shows the nominal resolution ($R \sim 50$) and expected dark current ($1 \times 10^{-3} \text{ e}^-/\text{pixel/s}$) of the Nancy Grace Roman Telescope’s Coronagraph Instrument Spectrograph, which covers comparable wavelengths in the optical. Increasing spectral resolution in the visible could lead to lower exposure times to detect modern Earth-like O_2 , but only if detector noise is also decreased significantly from the nominal value.

2.2 Retrieval Cases

We followed up on the results of our analytical exploration of resolution and exposure time by performing retrievals for the most interesting and potentially requirement-driving cases. The implication that higher resolution in the visible may reduce the exposure time required to detect O_2 is particularly compelling in the case of the low- O_2 Proterozoic Earth. Conversely, we further assess the implication that going to lower resolution in the near-infrared would not significantly increase the required exposure time in the case of the Archean Earth. For both the visible and the near-infrared, we also performed retrievals for the Phanerozoic Earth, as this is an important nominal case that will guide HWO instrument design.

We show retrieved posterior distributions for the cases of interest in Figures 5.5, 5.6, 5.7, and 5.8. All retrievals for a given case have a fixed exposure time to isolate how atmospheric constraints vary with spectral resolution in the wavelength region of interest. In each plot, we show the posterior distributions for all atmospheric constituents in terms of log 10 abundances, with different spectral resolutions shown as different colors. The retrieval corresponding to the nominal resolution is shown filled-in gray, while the remaining cases

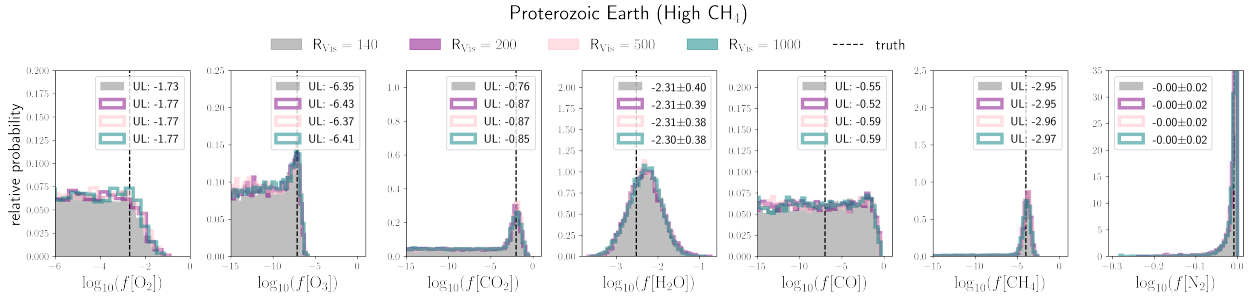


Figure 5.5: Retrieval results for the Proterozoic Earth with varying resolutions in the visible band-pass (NIR resolution is held constant at 70). Constant exposure time ($t = 81.63$ hours) is assumed for each retrieval. We show that, when SNR realistically changes with resolution, no advantage is conferred by going to higher resolution in the visible, where O_2 , O_3 , and H_2O have absorption features.

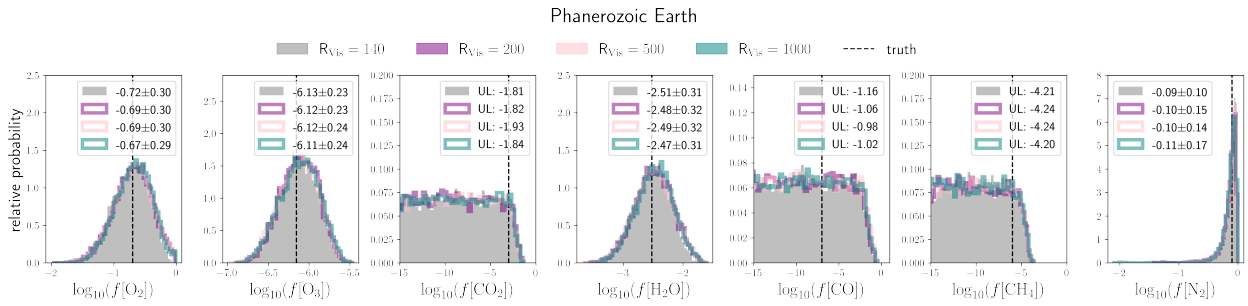


Figure 5.6: Retrieval results for the Phanerozoic Earth with varying resolutions in the visible band-pass (NIR resolution is held constant at 70). Constant exposure time ($t = 81.63$ hours) is assumed for each retrieval. We show that, when SNR realistically changes with resolution, no advantage is conferred by going to higher resolution in the visible, where O_2 , O_3 , and H_2O have absorption features.

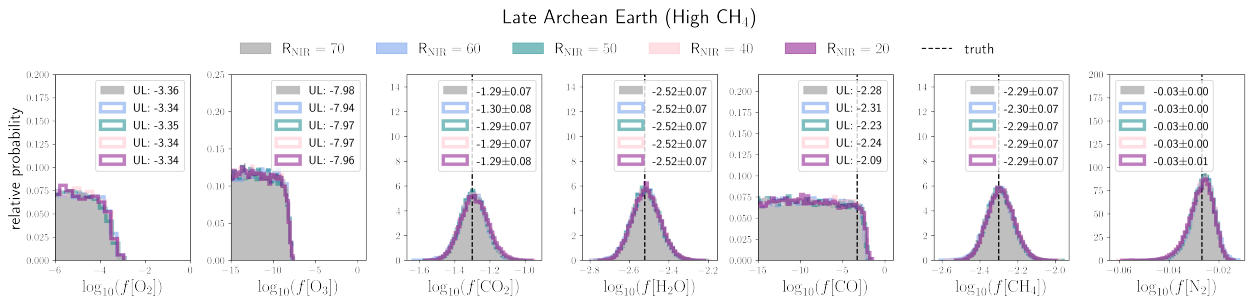


Figure 5.7: Retrieval results for the Archean Earth with varying resolutions in the near-IR band-pass (VIS resolution is held constant at 140). Constant exposure time ($t = 8163.27$ hours) is assumed for each retrieval. We show that, when SNR realistically changes with resolution, no significant advantage is conferred by going to higher resolution in the near-IR, where H_2O , CO_2 , CH_4 , and CO have absorption features.

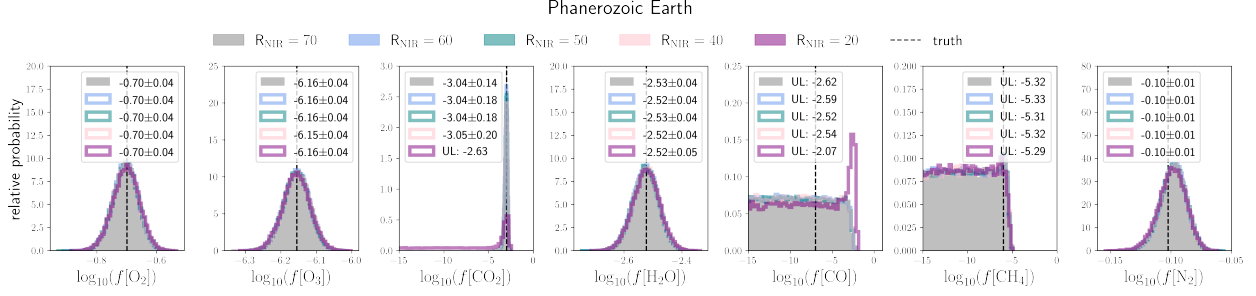


Figure 5.8: Retrieval results for the Phanerozoic Earth with varying resolutions in the near-IR (VIS resolution is held constant at 140). Constant exposure time ($t = 8245.51$ hours) is assumed for each retrieval. We show that there is a clear disadvantage for characterization when the resolution in the near-IR is degraded to $R_{\text{NIR}} = 20$. At this lowest spectral resolution, we lose the detection of CO_2 , an important habitability indicator. The upper limit on CO also increases because CO_2 and CO have neighboring absorption features in the near-IR, and sufficiently low resolving power effectively blurs these features.

are shown as step-style histograms. In each subplot, a dashed vertical line represents the “true” input abundance of a given gas. Finally, a legend in each subplot indicates whether we classify the posteriors (by inspection) as (a) a non-detection upper limit (“UL”) with the 3σ limit given in log-scale, or (b) a detection constraint, with the median retrieved abundance provided with the 1σ standard deviation as an uncertainty, both in log-scale.

For our investigation of spectral resolution in the visible wavelengths, our retrieval results for the Proterozoic (Figure 5.5) and Phanerozoic Earth (Figure 5.6) show slight variations in upper limits that are all within 0.18 dex. For O_2 in the Proterozoic case, variations in upper limits are within ± 0.04 dex. For gas detections in both cases, slight variations in the median retrieved abundances are well within the 1σ uncertainties, with no significant differences in posterior widths. Ozone is not detected in the Proterozoic case due the exclusion of the UV wavelengths. For the Phanerozoic atmosphere, we obtain only upper limits on CO_2 , CH_4 , and CO. This is because our constant exposure time of $t_{\text{exp}} \sim 80$ hours did not yield sufficiently high SNR in the near-IR wavelengths where these gases have significant absorption features.

Similar to the Proterozoic and Phanerozoic cases at visible wavelengths, our retrieval results in the near-infrared for the Archean (Figure 5.7) show a relative insensitivity to lower spectral resolutions. Slight variations in upper limits are all within 0.09 dex, and median

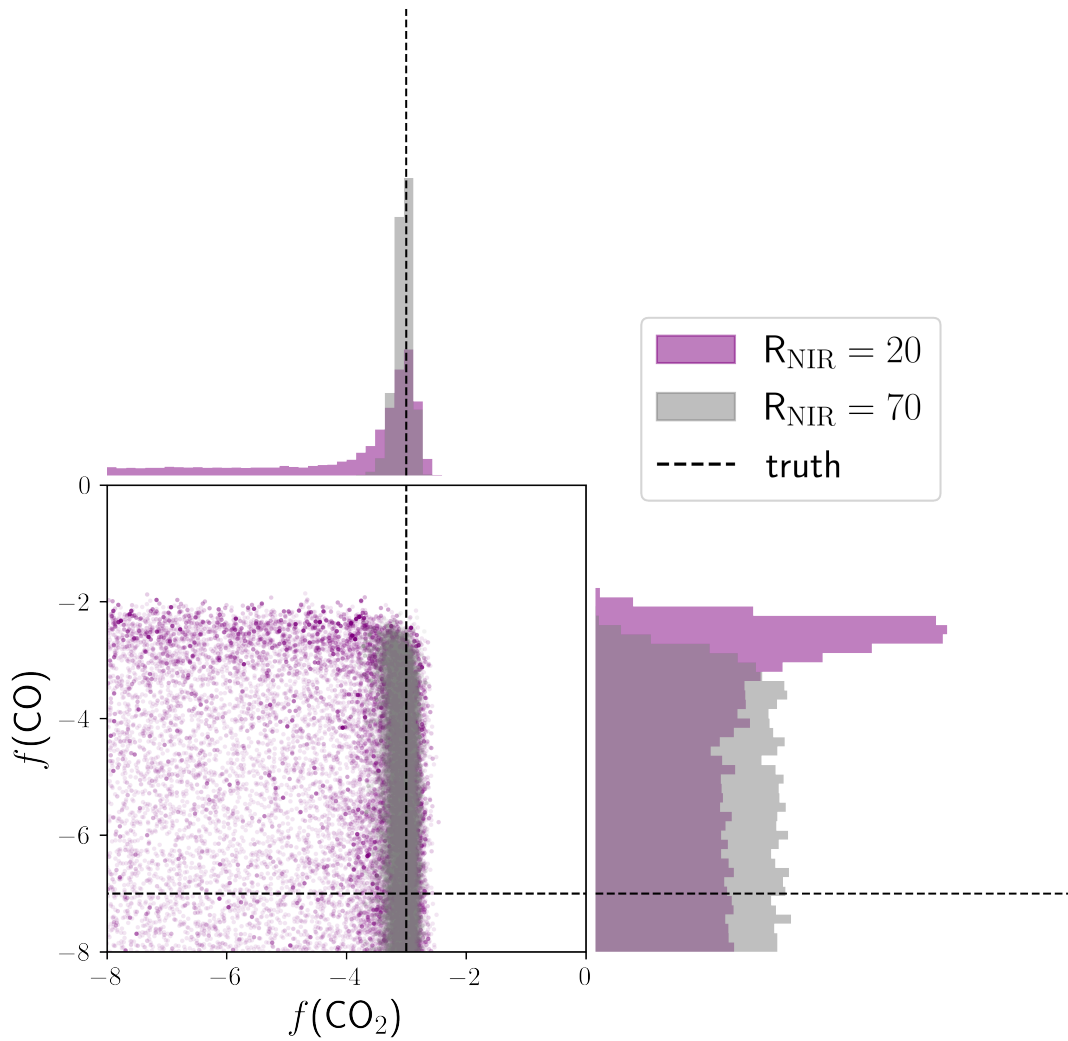


Figure 5.9: A comparison of the marginal and covariance plots of the retrieved CO_2 and CO abundances for the Phanerozoic Earth, with $R_{\text{NIR}} = 20$ and the nominal $R_{\text{NIR}} = 70$. When the spectral resolution in the near-IR is lowered to 20, the retrieval results suggest that the data favors an inaccurately high abundance of CO combined with lower CO_2 abundances.

retrieved abundances overlap within 1σ uncertainties. However, our results for the Phanerozoic Earth (Figure 5.8) show an interesting difference. For the Phanerozoic case, though near-infrared resolutions between 40 and 70 follow a similar pattern, the lowest resolution $R_{\text{NIR}} = 20$ shows a clear difference in the retrieved CO_2 and CO posterior distributions.

At the lowest resolution that we explored in the near-infrared ($R_{\text{NIR}} = 20$), the retrieval results show that we lose confidence in the unambiguous detection of CO_2 (as evidenced by the extended tail to low CO_2 abundances). At the same time, the 3σ upper limit on CO increases by 0.55 dex relative to the nominal resolution case. Both posteriors imply a worsening ambiguity in the abundances of CO and CO_2 , where the data slightly favors an unrealistically high CO abundance with correspondingly low CO_2 abundances. This ambiguity arises because the extremely coarse resolution effectively blends the weak CO_2 and CO near-infrared features, resulting in a covariance favoring low CO_2 abundances with high CO abundances. To show the underlying degeneracy between CO and CO_2 , in Figure 5.9 we compare the 1D and 2D marginalized posteriors of CO_2 and CO for the Phanerozoic case for both $R_{\text{NIR}} = 20$ and the nominal $R_{\text{NIR}} = 70$. The CO- CO_2 covariance shows that the long tail extending to lower CO_2 abundances results in a corresponding probability build-up at higher CO abundances, creating a visible peak in the CO marginal distribution that could mislead interpretations.

While decreased resolution appears to have little to no effect on the retrieved posteriors for the Archean case, the Phanerozoic case demonstrates a clear disadvantage of excessively degrading the resolution in the near-infrared. This is likely because the Phanerozoic has lower abundances of CO_2 , and thus weaker absorption features. At higher, Archean-like CO_2 abundances, additional absorption features that do not overlap with CO help break this degeneracy, allowing for a confident detection of CO_2 .

3 Discussion

Our study produces results in two forms: analytical detectability calculations, and retrieved posterior distributions. Together, these methods are complementary. The detectability calculations provide the minimum spectral resolution, dark current, and exposure time required to confidently detect gases via isolated absorption features, whereas retrievals demonstrate how resolution trade space affects our ability to accurately and meaningfully characterize the planetary atmosphere. Given that the detectability calculations are less nuanced, instrument parameters inferred from this analysis represent general guidance specific to the science cases investigated here. Conversely, retrievals provide the most rigorous assessment of how spectral resolution may impact the search for life on Earth-like exoplanets.

In the visible wavelengths, we are primarily concerned with the spectral resolution required to detect and constrain O_2 , O_3 , and H_2O . While our detectability calculation results (e.g., [Table 5.5](#)) show that O_3 and H_2O may be detected at lower resolving powers ($R_{\text{Vis}} \geq 30$), the spectral resolution in the visible should be set by the narrow O_2 -A band, which requires higher spectral resolution for detection ($R_{\text{Vis}} \geq 135$). These constraints are consistent with findings from past works ([Des Marais et al., 2002](#); [Brandt & Spiegel, 2014](#); [Feng et al., 2018](#); [Damiano & Hu, 2022](#)). While R_{Vis} should not drop below this recommendation, [Table 5.5](#) and [Figures 5.2, 5.3, and 5.4](#) suggest that increased spectral resolution may slightly reduce the exposure time required to detect O_2 , and we further discuss this implication in [subsection 3.1](#).

In the near-IR, we are primarily concerned with detecting and/or constraining H_2O , CH_4 , CO_2 , and CO . While our detectability calculation results (e.g., [Table 5.5](#)) show that $R_{\text{NIR}} \geq 20$ may be sufficient for detecting H_2O and CH_4 , CO_2 requires higher spectral resolving power ($R_{\text{NIR}} \geq 38$ in the early Archean, and $R_{\text{NIR}} \geq 66$ in the Phanerozoic). Though CO seems to similarly require higher spectral resolutions for detection, we note that the minimum exposure time is uniformly very high across all Earth through time cases ($t_{\text{exp}} \geq 10^6$ hours). For these cases, we lean more on the interpretation of our retrieval simulations,

where information on upper limits is accessible. Our findings are consistent with those from previous works (Des Marais et al., 2002; Damiano & Hu, 2022). Figure 5.8 and Figure 5.9 show that spectral resolutions $R_{\text{NIR}} < 40$ may introduce a concerning degeneracy between CO_2 and CO . This degeneracy results in the loss of a confident CO_2 detection, important for characterizing planetary habitability, and a worsening upper limit constraints on CO , an important antibioticsignature for methanogenic life. We further discuss these results in subsection 3.2. A detailed comparison of our results to previous works is included in subsection 3.3. Finally, we discuss future research directions in subsection 3.4.

3.1 The Benefits of High Resolution in the Visible Require Lower Detector Noise

At visible wavelengths, our detectability calculations and retrieval simulations suggest that the nominal resolution of $R_{\text{Vis}} = 140$ is sufficient for detecting and characterizing O_2 for the Phanerozoic Earth. Though the detectability calculations suggest that H_2O and O_3 are generally over-resolved, this does not appreciably degrade our ability to constrain the abundance of these gases. Likewise, in our Phanerozoic retrieval simulations where O_2 , O_3 , and H_2O are confidently detected across all resolution cases, higher spectral resolution does not significantly enhance the precision on abundance constraints – all 1σ uncertainties fall within ± 0.01 dex.

For the low- O_2 Proterozoic atmospheres, our detectability calculations suggest that a minimum resolution of $R_{\text{Vis}} = 171$ is required to detect O_2 for exposure times within 20% of the minimum, suggesting that the nominal $R_{\text{Vis}} = 140$ is suboptimal for detecting Proterozoic-like O_2 abundances. Accordingly, the nominal dark current and nominal resolution yield a minimum required exposure time of ~ 3000 hours for detecting O_2 at 0.76μ for such low abundances. To reduce the minimum required exposure time to detect O_2 to 1000 hours, our contour plots (Figure 5.3) suggest that the dark current would need to be reduced by at least an order of magnitude from the nominal HWO value (3×10^{-5} to 1×10^{-6}), and

the spectral resolution increased to nearly $R_{\text{Vis}} = 5000$. For more realistic exposure times that correspond to searches for modern-Earth O_2 levels ($t_{\text{exp}} \sim 80$ hours; $\text{SNR} = 20$ at 0.76μ for $R_{\text{Vis}} = 140$), our Proterozoic retrieval simulations suggest that higher spectral resolution does not meaningfully improve the precision of upper limits obtained on O_2 or O_3 .

The detectors under consideration for HWO are already projected to improve upon Roman CGI by nearly two orders of magnitude (Stark et al., 2025). Thus, the dark current we determined is required to detect O_2 in low- O_2 Proterozoic-Earth-like atmospheres at higher resolutions (~ 3 order of magnitude improvement over CGI) may ultimately be challenging with anticipated detector technology. Furthermore, even if such detectors are developed, the minimum required exposure time would only be reduced by at most a factor of 3. Assuming comparable exposure times required for characterizing the Phanerozoic Earth, increasing the spectral resolution at the nominal dark current does not meaningfully improve our characterization of the atmosphere. For example, at the nominal resolution, the 3σ upper limit corresponds to an O_2 abundance that is a factor of 9.3 larger than the true abundance. At the higher resolutions considered in our study, the 3σ upper limit corresponds to an O_2 abundance that is a factor of 8.5 larger than the true abundance. In either case, the retrieved upper limits would not indicate whether O_2 is present in this atmosphere.

It is also important to recognize that spectral resolution decisions do not occur in a vacuum: resolution affects the detectability of all gases simultaneously, not just the target species of interest. For example, were R_{Vis} increased to ≥ 1000 to improve sensitivity to low O_2 abundances, the exposure time required to detect H_2O would also increase by up to a factor of ~ 2 at the nominal dark current (Figure 5.3). This is a meaningful penalty, given that H_2O is expected to be a near-ubiquitous feature of habitable-zone planets and would be a primary observational target across a full planetary survey. Young et al. (2024) outline a decision-tree framework in which H_2O detection serves as the essential first step for establishing planetary habitability, with all subsequent observations, including searches for O_2 , CH_4 , and CO_2 , contingent on that initial detection. Any potential gains for low-

O₂ worlds would therefore come at the cost of substantially more expensive observations for the more general case of planets with atmospheric water vapor. On the other hand, were R_{Vis} decreased to ~ 30 to optimize for detecting H₂O, O₂ would be under-resolved and our ability to accurately identify worlds with oxygenic life would be seriously compromised. Thus, optimizing spectral resolution for characterizing one gas may degrade the efficiency of characterizing others, and resolution requirements must be balanced across multiple species to inform instrument design.

We conclude that the gains from higher spectral resolution in the visible are marginal at best, and that the nominal resolution in the visible ($R_{\text{Vis}} = 140$) is likely adequate for characterizing O₂ on the the Earth through time, despite challenges associated with probing O₂ in Proterozoic-Earth-like atmospheres. Furthermore, our analysis implies that such low-O₂ atmospheres ($\leq 1\%$ PAL) would require more stringent detector noise properties to achieve the theoretical gains from high resolution, which could complicate technology risk assessments. Instead, potentially the most efficient path is to indirectly infer the presence of O₂ via O₃, which may be readily detectable in the ultraviolet ($0.25 < \lambda < 0.40 \mu$) at low spectral resolution ($R_{\text{UV}} \sim 7$) across a range of abundances (Schwieterman et al., 2018a; Gilbert-Janizek et al., 2024; Krissansen-Totton et al., 2025).

3.2 The Resolution in the Near-IR Should be Greater than 20 to Avoid Retrieval Ambiguities and Enable CO₂ Detections

In the near-IR wavelengths, our detectability calculation results suggest that spectral resolutions of $R_{\text{NIR}} \geq 42$ are required to detect CO₂ for the early-through-late Archean Earth within our 20% exposure time threshold. Given this exposure time threshold, we find that slightly higher spectral resolutions of $R_{\text{NIR}} \geq 61$ and $R_{\text{NIR}} \geq 66$ are better for detecting CO₂ for the Proterozoic and Phanerozoic Earth, respectively. This suggests that the nominal $R_{\text{NIR}} = 70$ is sufficient for detecting Proterozoic and Phanerozoic-like CO₂ abundances.

Overall, our results suggest that degrading the spectral resolution in the near-IR has

modest effects on the characterization of Archean-like atmospheres. Our fixed-exposure-time retrieval simulations (Figure 5.7), where CO₂, CH₄, and H₂O are confidently detected across all resolution cases, show that degrading the spectral resolution to as low as $R_{\text{NIR}} = 20$ does not significantly affect any of the abundance constraints on these gases – our 1σ uncertainties never exceed ± 0.01 dex. Thus, though the detectability calculations suggest that H₂O and CH₄ are generally over-resolved, this does not negatively impact our ability to constrain the abundance of these gases. Interestingly, the 3σ CO upper limit generally increases with decreased spectral resolution. The 3σ upper limit of the coarsest resolution considered ($R_{\text{NIR}} = 20$) is 0.19 dex larger than that of the nominal resolution case ($R_{\text{NIR}} = 70$). These correspond to upper limits of 8.1×10^3 ppm for the lowest resolution, and 1.6×10^3 ppm for the nominal resolution. This trend for this particular Earth-through-time case suggests that lower spectral resolution may induce larger uncertainties on the abundance of undetected gases.

Our Phanerozoic retrieval simulations most clearly demonstrate how degrading the spectral resolution in the near-IR may lead to larger uncertainties on gas abundances. While decreasing the spectral resolution does not significantly impact our constraint on H₂O (within ± 0.01 dex) nor our upper limit on CH₄ (within ± 0.03 dex), it does ultimately affect our ability to confidently detect CO₂ and place upper limits on CO abundances. As spectral resolution decreases, the 1σ uncertainties in the CO₂ abundance constraints generally increase from ± 0.14 ($R_{\text{NIR}} = 70$) to ± 0.20 dex ($R_{\text{NIR}} = 40$). At the lowest resolution we investigated, $R_{\text{NIR}} = 20$, this increasing uncertainty ultimately gives rise to a concerning degeneracy between CO₂ and CO abundances. At this coarse spectral resolution, the CO₂ detection degrades to an upper limit, and the 3σ upper limit on CO increases by 0.55 dex relative to the nominal case. This effect occurs due to spectral degeneracies between CO₂ and CO absorption and a resulting covariance between their abundances (Figure 5.9). The lack of a CO₂ detection in this case is concerning, as it is an important habitability indicator (Catling et al., 2018). Furthermore, the poorer upper limit on CO results in worsening

constraints on the ratio of CO/CH₄. We note that the 3σ upper limits on both gases for the lowest spectral resolution case do not allow us to rule out scenarios where CO > CH₄, an important biosignature false positive diagnostic (Ragsdale, 2004; Zahnle et al., 2008, 2011; Gao et al., 2015; Wang et al., 2016; Wogan & Catling, 2020; Krissansen-Totton et al., 2022; Ranjan et al., 2023).

We conclude that the nominal resolution in the near-IR ($R_{\text{NIR}} = 70$) is more than adequate for characterizing CH₄, CO₂, and H₂O on the Earth through time, and that spectral resolutions as low as $R_{\text{NIR}} = 40$ may be sufficient in this wavelength range. Our results also suggest that degrading the resolution in the near-IR below $R_{\text{NIR}} = 40$ may incur additional uncertainties in the abundance of CO₂ and CO. At lower spectral resolutions ($R_{\text{NIR}} = 20$), these additional uncertainties result in an ambiguity between low abundances of CO₂ and high abundances of CO which misrepresent the planetary atmosphere. We therefore conclude that resolutions $R_{\text{NIR}} < 40$ are demonstrably disadvantageous for accurately characterizing Earth-through-time atmospheres.

3.3 Comparison to Previous Work

Our study assesses the spectral resolution needed to characterize Earth-through-time atmospheres using the same cases considered by Krissansen-Totton et al. (2025), with a long-wavelength cutoff of 1.7μ . Our retrieval results are broadly consistent despite key differences in study design. While Krissansen-Totton et al. (2025) assumed constant SNR across the full wavelength range, we assumed constant exposure times. Additionally, we excluded UV wavelengths that they included.

For the Proterozoic Earth, Krissansen-Totton et al. (2025) detected CO₂ and CH₄ with SNR > 10 whereas our visible-focused retrievals obtained only an upper limit on both gases, likely because our $t_{\text{exp}} \sim 80$ hour simulations achieve SNR < 5 in the near-IR, as shown in Figure 5.1b. Both studies obtain upper limits on CO, but neither detects O₂. We do not detect O₃ due to our exclusion of UV wavelengths where O₃ has strong absorption

features. For the Archean Earth in the near-IR, both studies find that CO_2 and CH_4 are well-constrained. For the Phanerozoic Earth, both studies obtain upper limits on CH_4 and constrain O_2 and O_3 . Finally, our CO_2 results are consistent with [Krissansen-Totton et al. \(2025\)](#) when accounting for SNR: in our visible-focused retrieval with low SNR in the near-IR, we obtain only upper limits as they do. In our near-IR-focused retrieval with $\text{SNR} > 40$, we confidently detect CO_2 , likely because our near-IR SNR exceeded the highest SNR case considered in their study.

In terms of spectral resolution, previous work argued that resolution in the visible should be set by the detection of the narrow-band O_2 feature at 0.76μ ([Des Marais et al., 2002](#); [Brandt & Spiegel, 2014](#); [Feng et al., 2018](#); [LUVOIR Team, 2019](#); [Gaudi et al., 2020](#); [Damiano & Hu, 2022](#)), while the resolution in the near-IR should be set by the detection of CH_4 and CO_2 features ([Des Marais et al., 2002](#); [LUVOIR Team, 2019](#); [Gaudi et al., 2020](#); [Damiano & Hu, 2022](#)). Here we compare our results to that of previous studies. Whereas our retrieval simulations are most comparable to those of [Feng et al. \(2018\)](#) and [Damiano & Hu \(2022\)](#), our analytical calculations are more similar to those in [Brandt & Spiegel \(2014\)](#). Our retrieval simulations suggest that the nominal spectral resolution in the visible ($R_{\text{vis}} = 140$) is sufficient to meaningfully constrain O_2 for atmospheres akin to the Phanerozoic Earth. While our analytical calculations suggest that spectral resolutions up to 1,300 would also be acceptable for the nominal dark current, our retrieval results suggest that the information gained from higher resolutions is negligible compared to lower resolutions, given identical exposure times. Moreover, [Figures 5.4 and 5.3](#) suggest that the required dark current needs to be at least an order of magnitude smaller for higher resolution to ensure equivalent oxygen detectability, further increasing the burden of detector complexity. These results are in agreement with [Feng et al. \(2018\)](#) and [Damiano & Hu \(2022\)](#), who found that $R_{\text{vis}} = 140$ (with $\text{SNR} = 20$) is required in the optical wavelengths to ensure meaningful constraints on O_2 abundance. Our results are also in agreement with the detectability calculations of [Brandt & Spiegel \(2014\)](#), which found that a resolution of 150 is required to detect O_2 . By

contrast, [Des Marais et al. \(2002\)](#) found that the required resolution to detect O_2 in the optical and the near-IR is 69–72 by measuring the location and width of various atmospheric absorption features. The figure found by [Feng et al. \(2018\)](#); [Damiano & Hu \(2022\)](#), and this work is likely twice as large due to the consideration of Nyquist sampling.

In the near-IR, we found that a spectral resolution of $R_{\text{NIR}} = 40$ is sufficient for characterizing the Archean and Phanerozoic Earth atmospheres, given adequately long exposure times. However, this likely represents a lower limit on the spectral resolution in this wavelength range, as we found that $R_{\text{NIR}} = 20$ leads to ambiguities in the retrieved CO and CO_2 abundances. The LUVIR final report ([LUVIR Team, 2019](#)) recommended $R_{\text{NIR}} = 70$ for CO_2 detection with an integral field spectrograph and $R_{\text{NIR}} = 200$ for single-point spectrographs. Our detectability calculations for the Phanerozoic Earth are consistent with these recommendations, finding that a range of resolutions ($R_{\text{NIR}} = 66\text{--}564$) would yield CO_2 detection in comparable exposure times ($t_{\text{exp}} \approx 3248$ hours). Rather than conducting retrieval simulations across this entire range, we focused on exploring the lower resolution limit suggested by our Archean Earth calculations ($R_{\text{NIR}} = 38$), finding that resolutions below $R_{\text{NIR}} = 40$ introduce degeneracies in atmospheric characterization. These results are consistent with those of [Damiano & Hu \(2022\)](#), who showed that a resolution of $R_{\text{NIR}} = 40$ in the near-IR is sufficient to characterize Archean/Modern Earths and discriminate CO_2 -dominated atmospheres. Though our results are broadly consistent with the range of acceptable resolutions found by [Des Marais et al. \(2002\)](#), $R_{\text{NIR}} = 11\text{--}40$, ($R_{\text{NIR}} = 22\text{--}80$ with Nyquist sampling) we rule out $R_{\text{NIR}} < 40$ for accurately characterizing the Phanerozoic Earth. [Brandt & Spiegel \(2014\)](#) and [Feng et al. \(2018\)](#) did not assess spectral resolution in the near-IR, so we are not able to compare our results to these prior works.

3.4 Future Work

In this study, we explored the effects of assumed spectral resolution in the visible and near-IR wavelengths for a limited set of atmospheric cases representing the Earth through time.

We did not explore the effect of spectral resolution in the UV, and we did not consider a broader range of atmospheric cases, which include abiotic false positive atmospheres. We also assumed photon-limited observations throughout, such that our absolute exposure time estimates do not account for time-independent noise sources like speckle noise that could introduce a noise floor at long integration times; our predictions of relative changes in exposure time across resolution cases are therefore more robust than the absolute values. Finally, we did not investigate the sensitivity of spectral resolution for surface characterization, including surface biosignatures. Future work should address these gaps to ensure HWO spectroscopy is fully capable of identifying living planets and discriminating their imitators.

Future work should explore the spectral resolution in the UV needed to detect and constrain O_3 abundances. As a photochemical byproduct of O_2 , O_3 may indirectly indicate the presence of O_2 in low- O_2 atmospheres like the Proterozoic Earth (Krissansen-Totton et al., 2025). Though it has weak absorption features in the visible wavelengths, the Hartley-Huggins bands in the UV are most sensitive to changes in abundance and the presence of a stratospheric bulge (Gilbert-Janizek et al., 2024). Given our results showing that higher spectral resolution in the visible cannot improve upper limits on O_2 abundance, UV capabilities may be particularly important for characterizing Proterozoic-like, low- O_2 atmospheres (Damiano et al., 2023; Young et al., 2026; Washington III et al., 2026).

Our study did not assess spectral resolution for a broad range of atmospheric cases and species despite the potential diversity of terrestrial exoplanet atmospheres to be discovered. Future work should extend this assessment to additional atmospheric cases, including false positive atmospheres. For example, recent work has shown that false positive atmospheres can inform HWO instrument requirements like the short-wavelength cutoff in the UV and the long-wavelength cutoff in the near-IR (Krissansen-Totton et al., 2025). Similarly, future work should assess how spectral resolution affects our ability to constrain SO_2 , an intermediate product of sulfate metabolisms with absorption features that overlap with those of O_3 in the near-UV (Young et al., 2026; Washington III et al., 2026). Another species that should

be considered in future works is N_2O , a waste product of microbial nitrogen metabolisms with absorption features that overlap with other near-IR absorbers (Tokadjian et al., 2024). Broadening the atmospheres and species under investigation is particularly important for a complete assessment of how spectral resolution affects our sensitivity to spectral degeneracies.

Our exposure time calculations assumed photon-limited observations, but real observations may deviate from this regime in ways that affect the absolute exposure times reported here, though the relative changes in exposure time across resolution cases are likely more robust. Time-independent noise sources, such as speckle noise from residual starlight, introduce a noise floor that photon-limited scaling cannot capture. At sufficiently long exposure times, such a noise floor could prevent detection of species that already require exceedingly long integration times (> 100 hours) at the nominal resolutions. This caveat is particularly relevant to our Proterozoic results, where the exposure times required to detect O_2 ($t_{\text{exp}} \sim 2000$ hours) already exceed practical limits. In this context, our predictions of *relative* changes in exposure time across resolution cases are more robust than the absolute exposure times themselves, which will also vary with target distance and telescope aperture. Future work should incorporate time-independent noise components to assess whether spectral resolution choices remain optimal in the presence of a noise floor, and to refine exposure time estimates for specific HWO target stars.

Speckles can also generate correlated noise across the detector as a function of wavelength, which can mimic or obscure absorption features. While we do not account for correlated noise in this study, see Ruffio et al. (2026) for a detailed investigation. Despite neglecting speckles, our findings agree with Ruffio et al. (2026) in that sufficiently low detector noise can enable more sensitive, high resolution ($R \sim 1000$) observations, particularly for narrow-band absorption features. Ruffio et al. (2026) highlight that ultra-stable wavefront control minimizes correlated speckle noise, but this level of performance poses a significant technological challenge for HWO. Future work could explore the trade space between relaxed (stricter) wavefront stability requirements and stricter (relaxed) detector noise requirements needed

for high (low) resolution observations (e.g., using [Steiger et al., 2026](#)).

Our investigation also assumed gray surface albedos throughout, but recent work has demonstrated that reflected light observations can be sensitive to surface composition ([Ulses et al., 2025](#)), including biological pigments ([Hegde et al., 2015](#); [Schwieterman et al., 2018b](#); [Gomez Barrientos et al., 2023](#); [Borges et al., 2024](#)). [Gomez Barrientos et al. \(2023\)](#) showed that wavelength-dependent retrieval models can accurately identify spectral edge features, including the “vegetative red edge” – a sharp albedo transition near 750 nm caused by vegetation. Constraining the surface composition of a planet may also help discriminate abiotic scenarios. For example, [Ulses et al. \(2025\)](#) demonstrated that reflected light observations spanning 0.3–1.0 μ can reveal the presence of land surfaces, which would help rule out water-world false positives. Future work should investigate how spectral resolution affects the detectability of wavelength-dependent albedo features and determine the UV, visible, and near-IR resolutions needed to break potential spectral degeneracies, such as the O₃-land degeneracy identified by [Ulses et al. \(2025\)](#).

Finally, our retrieval simulations also adopted simplifying assumptions about clouds: we assumed a constant cloud fraction across all atmospheric cases and treated clouds as gray absorbers. While these choices provide a useful controlled baseline, real cloud populations exhibit wavelength-dependent absorption features that can be degenerate with gas-phase absorbers. For example, [Gilbert-Janizek et al. \(2024\)](#) demonstrated that ice cloud absorption features in the near-IR can produce false positive detections of gases such as CO. In principle, higher spectral resolution may help break such degeneracies, since the sharp pressure-broadened line structure of gas-phase species is readily distinguishable from the broad, featureless absorption of condensed-phase particles at sufficient resolution. Future work should investigate how varying cloud fraction, cloud particle composition, and particle size affect retrieval outcomes, and assess whether the resolution requirements identified here remain robust across a broader range of cloud scenarios.

4 Conclusion

We assessed how spectral resolution affects HWO’s ability to detect and characterize biosignature gases in Earth-like exoplanet atmospheres using complementary detectability calculations and atmospheric retrievals. Our analysis focused on O_2 , O_3 , H_2O , CH_4 , CO_2 , and CO across visible and near-IR wavelengths for atmospheres representing Earth’s evolution from the Archean through the Phanerozoic.

In the visible, we find that the nominal resolution of $R_{\text{Vis}} = 140$ is sufficient for detecting O_2 in Phanerozoic-like atmospheres. While higher resolutions could theoretically reduce the minimum exposure time needed to detect O_2 in low- O_2 Proterozoic atmospheres, the required improvements in detector noise ($> 10\times$ reduction in dark current from our nominal value assumed for HWO) may increase technological risk. Furthermore, spectral resolution decisions do not occur in a vacuum: increasing the spectral resolution in the visible to improve sensitivity for low- O_2 worlds could increase the exposure time required to detect H_2O by up to a factor of ~ 2 , penalizing the foundational habitability constraint that anchors downstream biosignature search strategies. The marginal improvements in exposure time seen for O_2 by increasing R_{Vis} likely do not justify significant increases in spectral resolution beyond the nominal value, although small increases beyond 140 have no major downsides. In the near-IR, we find that $R_{\text{NIR}} \geq 40$ is necessary to avoid a concerning degeneracy between CO_2 and CO that could produce false positive detections of abundant CO . The nominal resolution of $R_{\text{NIR}} = 70$ provides robust constraints on CO_2 and H_2O across all Earth-through-time cases.

These results support HWO’s current baseline resolution choices. Future work should extend this analysis to UV wavelengths to assess O_3 detection in low- O_2 atmospheres and to a broader range of atmospheric scenarios including abiotic false positive cases. Our findings provide actionable guidance, including some flexibility, for finalizing HWO’s spectrometer requirements while maintaining the technical capability necessary for the search for life on exoplanets.

Appendix

In Figures 5.10, 5.11, 5.12, and 5.13 we show the full corner plots for each atmospheric case explored in the retrieval portion of this investigation. Each plot overlays the different spectral resolution cases for a given Earth-through-time atmosphere, including the corresponding “truth” values for each retrieval parameter. In addition to the atmospheric abundances shown in the main text, the full corner plots include surface albedo (A_0), planet radius (R_p) and mass (M_p) in Earth units, cloud layer thickness (Δp_c) in Pa, cloud-top pressure (p_t) in Pa, cloud opacity (τ_c), cloud fraction (f_c), and total atmospheric surface pressure (p_0) in Pa. Since all gases are retrieved originally in partial pressure space, the total surface pressure posterior represents the sum-total of the gas posteriors. A brief discussion of key cloud parameters is included in the main text, with a more thorough discussion below.

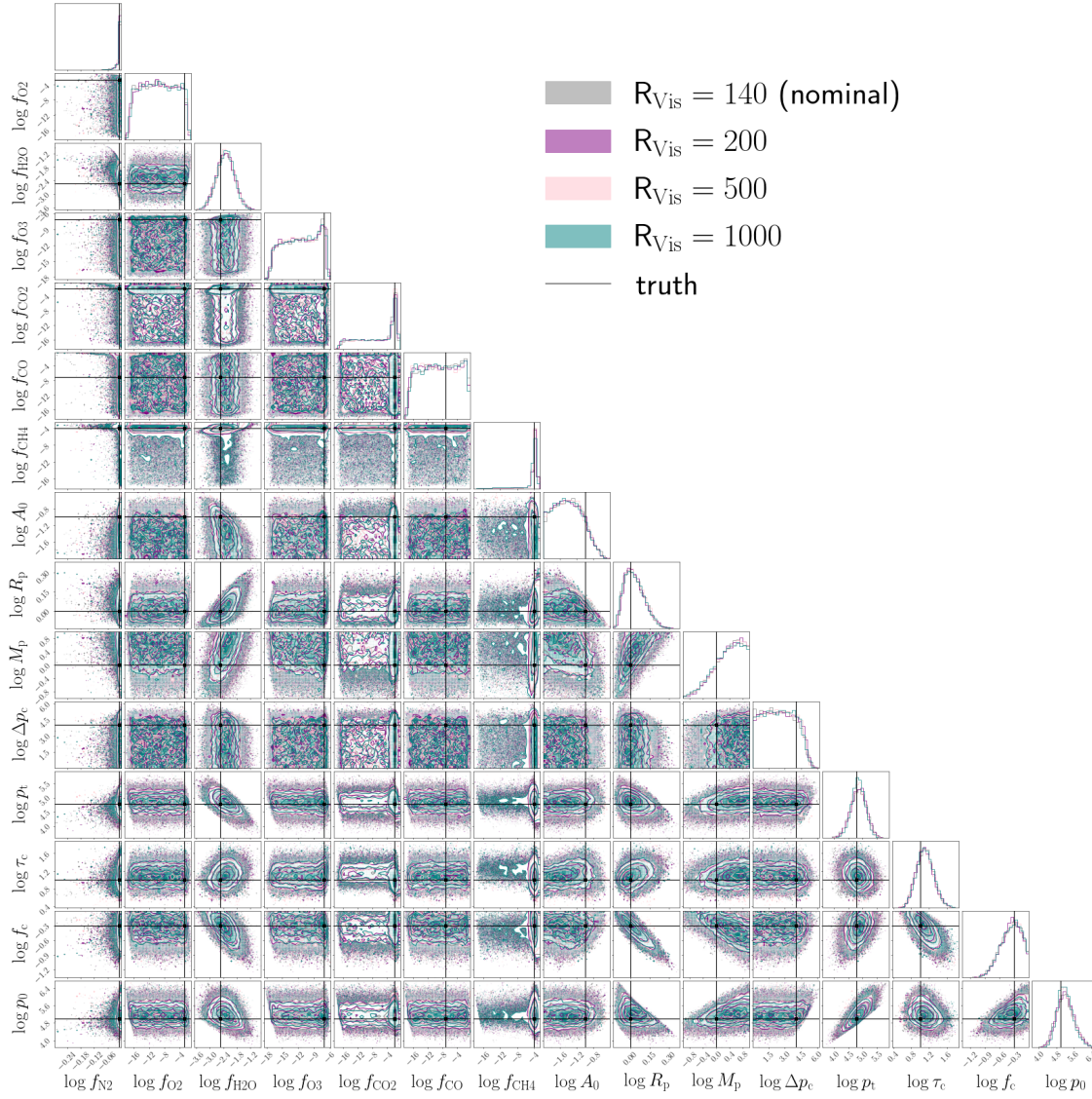


Figure 5.10: Full corner plot for the high-CH₄ Proterozoic case, including the nominal spectral resolution ($R_{vis} = 140$) and higher resolutions ($R_{vis} = 200, 500$, and 1000) in the visible (VIS) wavelengths.

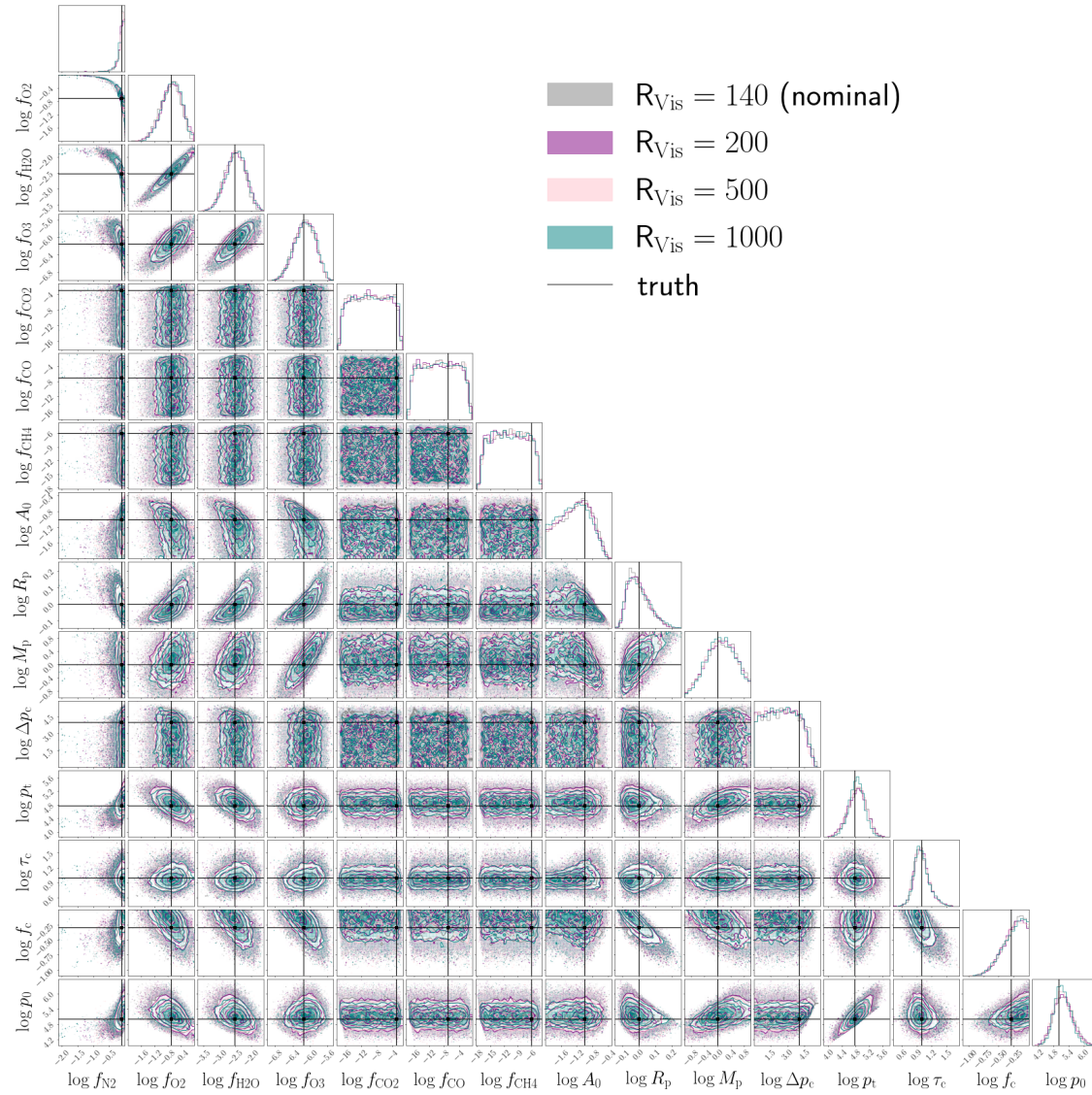


Figure 5.11: Full corner plot for the Phanerozoic case, including the nominal spectral resolution ($R_{Vis} = 140$) and higher resolutions ($R_{Vis} = 200, 500$, and 1000) in the visible (VIS) wavelengths.

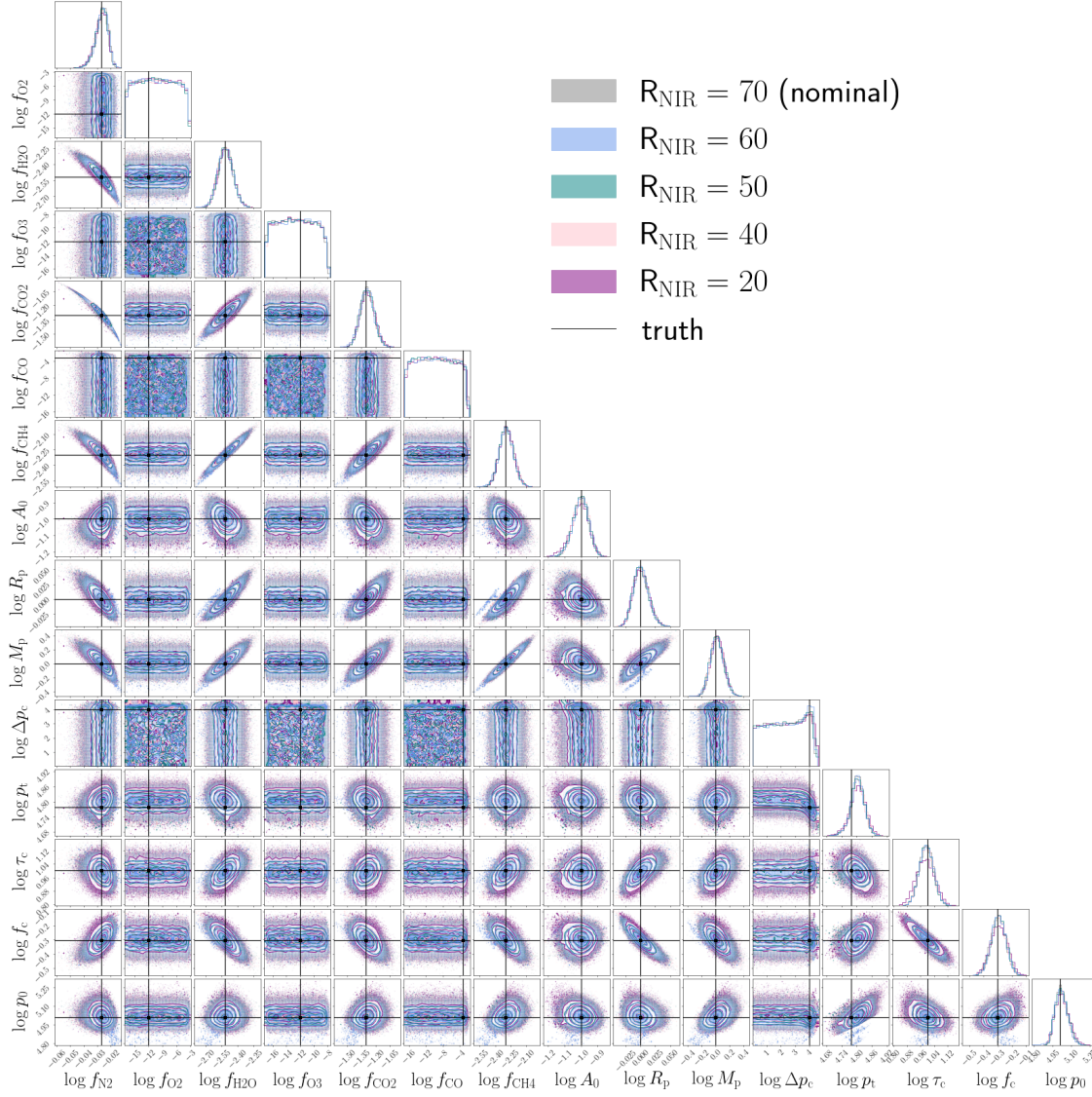


Figure 5.12: Full corner plot for the Late Archean case, including the nominal spectral resolution ($R_{NIR} = 70$) and lower resolutions ($R_{Vis} = 60, 50, 40$, and 20) in the near-infrared (NIR) wavelengths.

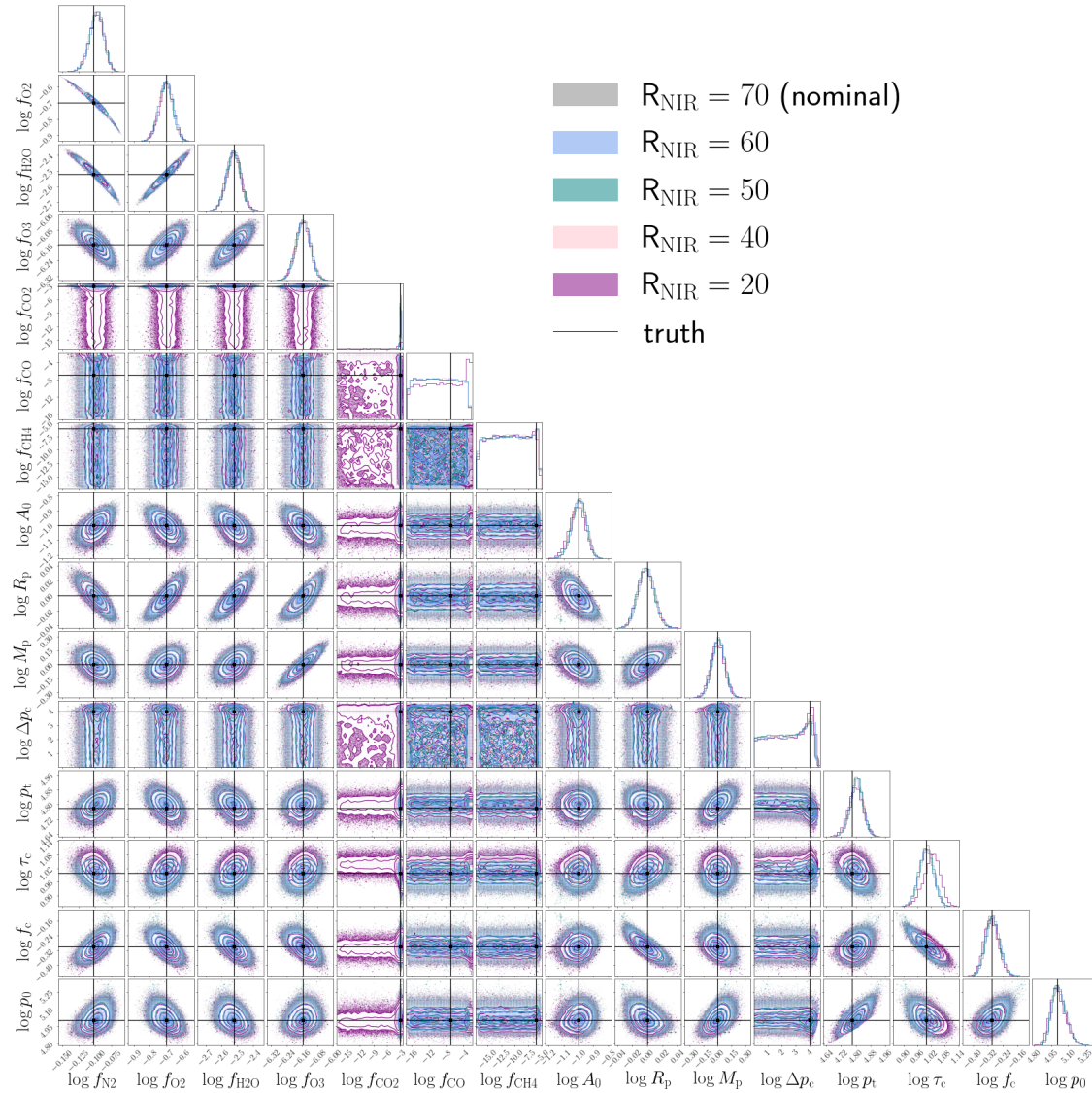


Figure 5.13: Full corner plot for the Phanerozoic case, including the nominal spectral resolution ($R_{NIR} = 70$) and lower resolutions ($R_{Vis} = 60, 50, 40$, and 20) in the near-infrared (NIR) wavelengths.

Figures 5.14 and 5.15 highlight minor trends in the retrieved cloud parameters, which are similar to those detailed for the gas abundance parameters in the main text. We show the retrieval accuracy (median $\bar{\ell}$ - truth ℓ_{true}), posterior standard deviation (σ), and the relative precision ($\sigma/\ell_{\text{true}}$) for select cloud parameters. Across all the Earth-through-time cases in the visible and near-IR, we find ~ 0.03 – 0.08 dex improvement in the standard deviation of the retrieved cloud top pressure posterior, corresponding to a 0.1–1% improvement in relative precision, as a function of increased spectral resolution. For the Late Archean and Phanerozoic atmospheres explored in the near-IR, the standard deviation (Figure 5.14, left panel) of the retrieved cloud opacity posterior remains mostly flat across the spectral resolutions under investigation. However, for the lowest resolution case ($R_{\text{NIR}}=20$) for the Phanerozoic and Archean atmospheres, we find that the posteriors are less precise than those of the nominal resolution ($R_{\text{NIR}}=70$) case by 0.005 and 0.01 dex in standard deviation, corresponding to 0.5% and 1.2% worse relative precision, respectively.

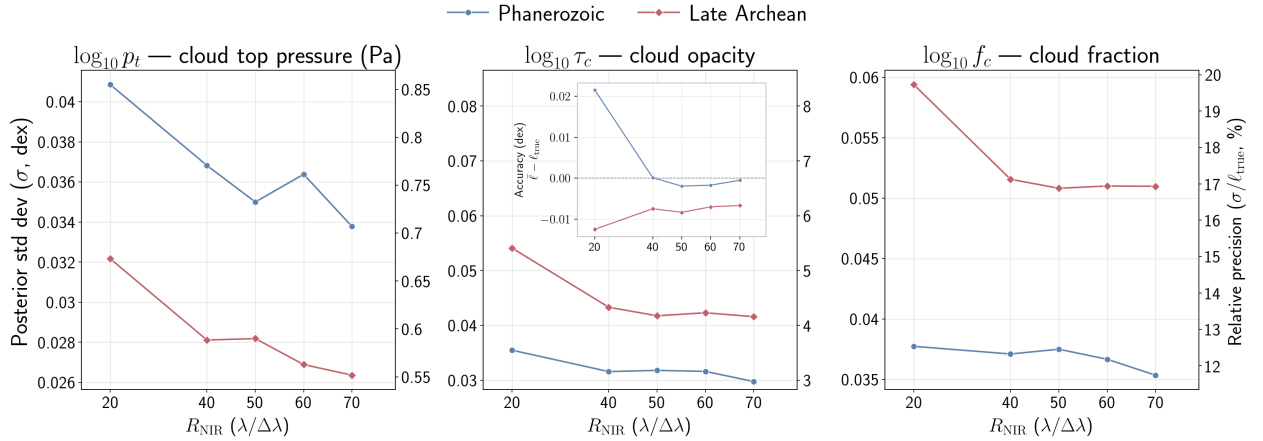


Figure 5.14: Cloud parameter retrieval diagnostics for the near-IR channel as a function of spectral resolution for the Phanerozoic (blue circles) and Late Archean (red diamonds) atmospheric cases. Each column shows a different cloud parameter: cloud-top pressure (p_t), cloud opacity (τ_c), and cloud fraction (f_c). The left-hand y-axis shows the posterior standard deviation and the right-hand y-axis shows the relative precision ($\sigma/\ell_{\text{true}}$) of the retrieved parameter, both computed in $\log_{10}$ space. The inset in the τ_c standard deviation panel shows the retrieval accuracy ($\bar{\ell} - \ell_{\text{true}}$) for cloud opacity, highlighting the minor systematic bias that emerges at the lowest resolution ($R_{\text{NIR}} = 20$) for the Phanerozoic case. The nominal resolution ($R_{\text{NIR}} = 70$) is marked on the x-axis of each panel for reference.

The Phanerozoic cloud opacity also exhibits a larger systematic bias at the lowest resolution, with the posterior median deviating from the true value by 2.2% at $R_{\text{NIR}}=20$ compared to only 0.1% at the nominal resolution, suggesting that coarse spectral resolution may introduce modest biases in cloud retrievals in addition to the precision losses described above (Figure 5.14, middle panel, inset). Finally, while Figure 5.14 (right panel) shows that the standard deviation of the retrieved cloud fraction posterior remains flat with spectral resolution for the Phanerozoic atmosphere in the near-IR, the late Archean case shows a loss of precision of 0.08 dex in standard deviation corresponding to a loss of 3% relative precision for the lowest resolution case relative to the nominal resolution. Although Figures 5.14 and 5.15 do reveal improvements in the retrieved precision of the cloud parameters at higher resolutions, these gains are sufficiently small so as to be nearly imperceptible in the corner plots (Figures 5.10, 5.11, 5.12, and 5.13) where these posteriors are over-plotted. Therefore, it is not evident that such gains would lead to statistically significant improvements in practice.

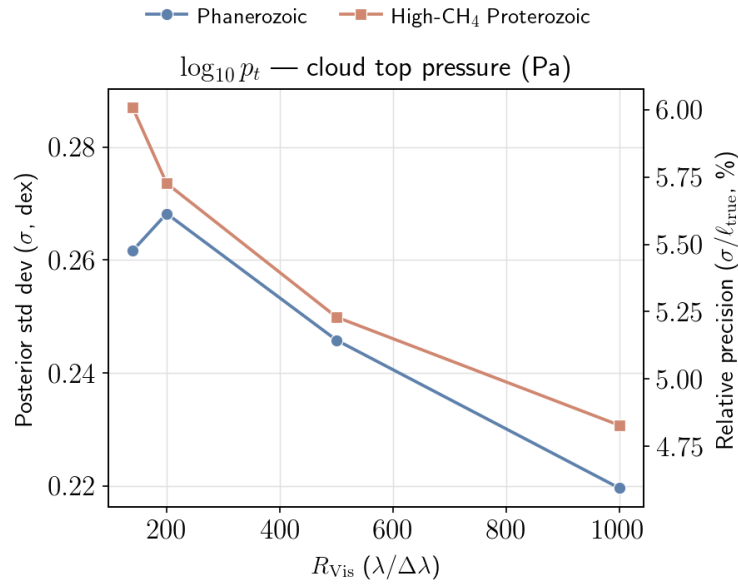


Figure 5.15: Cloud-top pressure (p_t) retrieval diagnostics for the visible channel as a function of spectral resolution for the Phanerozoic (blue circles) and High-CH₄ Proterozoic (orange squares) atmospheric cases. The left-hand y-axis shows posterior standard deviation and the right-hand y-axis shows relative precision ($\sigma/\ell_{\text{true}}$), both computed in $\log_{10}$ space. Both atmospheric cases show modest but consistent improvements in retrieval precision with increasing spectral resolution.

Table 5.5: For each Earth through time case, we show the permitted resolution ranges for detecting key gases within a 20% threshold of the minimum possible exposure time in hours, assuming the nominal dark current (3.5×10^{-5} e⁻/pixel/s) throughout. The nominal resolution is 140 in the visible (0.4–1.0 μm) and 70 in the near-IR (1.0–1.7 μm). In the “interpretation” column, we indicate for each gas whether the nominal resolution falls in the acceptable range, and whether the corresponding minimum exposure time is reasonable ($t < 1000$ hours). “SR” (green) means that the nominal resolution is sufficient and yields a reasonable exposure time, “SU” (yellow) means that the nominal resolution is sufficient but yields an unrealistically long exposure time, and “ISU” (red) means that the nominal resolution is insufficient and the required resolution range yields an unrealistically long exposure time. **Note that we do not find that altering the nominal resolution in the visible or near-IR wavelength ranges would significantly shorten the exposure times required to detect any of the gases across the Earth through time cases studied here.**

Earth through Time Case	Gas	Bandpass	Min. Exposure Time (hours)	Resolution Range ($\lambda/\Delta\lambda$)	Interpretation
Early Archean (Low CH ₄)	H ₂ O	Visible	5.98	30–420	SR
		Near-IR	3.38	20–164	SR
	CO ₂	Near-IR	11.56	38–352	SR
	CO	Near-IR	5.5×10^8	197–782	SU
	CH ₄	Near-IR	9.61	20–195	SR
Late Archean (High CH ₄)	H ₂ O	Visible	6.10	30–410	SR
		Near-IR	33.16	20–221	SR
	CO ₂	Near-IR	43.99	42–406	SR
	CO	Near-IR	1.8×10^6	57–706	SU
	CH ₄	Near-IR	1.55	20–142	SR
Proterozoic (Low CH ₄)	H ₂ O	Visible	5.30	30–422	SR
		Near-IR	1.04	20–144	SR
	O ₂	Visible	2099.93	171–3548	ISU
		Near-IR	1.3×10^6	60–339	SU
	O ₃	Visible	257.27	30–242	SR
	CO ₂	Near-IR	132.13	60–466	SR
	CO	Near-IR	1.1×10^{13}	58–407	SU
	CH ₄	Near-IR	1.4×10^4	26–256	SU
Proterozoic (High CH ₄)	H ₂ O	Visible	5.32	30–421	SR
		Near-IR	1.41	20–147	SR
	O ₂	Visible	2100.57	171–3549	ISU
		Near-IR	1.3×10^6	60–339	SU
	O ₃	Visible	257.30	30–242	SR
	CO ₂	Near-IR	134.60	61–468	SR
	CO	Near-IR	1.1×10^{13}	58–407	SU
	CH ₄	Near-IR	77.34	22–225	SR
Phanerozoic	H ₂ O	Visible	5.49	30–416	SR
		Near-IR	1.00	20–140	SR
	O ₂	Visible	6.17	135–1322	SR
		Near-IR	262.85	50–299	SR
	O ₃	Visible	3.06	30–241	SR
	CO ₂	Near-IR	3248.13	66–564	SU
	CO	Near-IR	5.9×10^{12}	60–366	SU
	CH ₄	Near-IR	3.1×10^5	27–269	SU

CHAPTER 6

CONCLUSIONS AND FUTURE WORK

1	Summary of Key Findings	260
1.1	Whole-Planet Modeling and the Abiotic Earth Baseline	260
1.2	Stellar Type and Habitable Zone Evolution	261
1.3	Transmission vs. Direct Imaging for Habitability and Biosignature Assessment	262
1.4	Spectral Resolution Requirements for HWO	263
2	Future Directions	264
2.1	Expanding the Abiotic Planet Parameter Space	264
2.2	Estimating the HZs of HWO FGK Stars	265
2.3	Retrieval Model Development	266
2.4	Mission Trade Studies and Target Prioritization	266
3	Concluding Remarks	267

1 Summary of Key Findings

This thesis has developed and validated tools for interpreting spectroscopic observations of Earth-like exoplanets with next-generation space telescopes, with a particular focus on NASA’s Habitable Worlds Observatory (HWO). By treating Earth as an exoplanet and applying the same remote sensing techniques to our own world that we will use to search for life elsewhere, we have established quantitative frameworks for predicting and assessing both habitability and biosignatures in exoplanet atmospheres.

1.1 Whole-Planet Modeling and the Abiotic Earth Baseline

In Chapter 2, we presented a coupled core-mantle-crust-climate model that reproduces crucial geophysical, climate, and ocean properties of the pre-industrial Earth. Our model predicts that life is not required to maintain a habitable planet resembling Earth-like conditions over billion-year timescales. This work contributes to the ongoing debate regarding the Gaia hypothesis by suggesting that life and habitability may be somewhat decoupled, with profound implications for exoplanet characterization: the detection of habitable conditions alone cannot be considered evidence for life.

Our model demonstrates that the evolution of the planetary interior plays an important role in determining the long-term habitability of Earth-like exoplanets. Though exoplanet interiors themselves will likely never be directly observable, the effects of interior evolution are apparent in the composition of the atmosphere and the albedo of the surface environment, both of which HWO will observe on Earth-like planets around Sun-like stars via reflected light spectroscopy. We have demonstrated that atmospheric states from our model can generate realistic reflected light spectra, establishing an abiotic baseline for biosignature interpretation.

1.2 Stellar Type and Habitable Zone Evolution

In Chapter 3, we presented an improved version of our whole-planet abiotic Earth model, including a treatment for fractional ice coverage, and assessed the sensitivity of our model to assumptions about seafloor and continental weathering. We then applied two versions of the model to exploring the HZ boundaries of selected F-, G-, and K-type stars.

We find that the occurrence of limit cycles strongly depends on our assumptions about the evolving carbon cycle. In particular, we find that limit cycles occur only if we assume that seafloor weathering cannot continue during global glaciations. Our evolving HZ boundaries for a $1.2M_{\odot}$ F-type star, solar-type G-dwarf, and $0.8M_{\odot}$ K-star are broadly consistent with those found by [Kadoya & Tajika \(2019\)](#). However, we find slightly more generous inner edges across all stellar types; less generous outer edges; and slightly broader HZs for G- and K-dwarfs at the conclusion of their main sequence lifetimes. For the F- and K-type stars, some of these differences may be attributed to differences in stellar mass, but we also identified key differences in our wavelength-independent treatment of surface albedo and our cloud parameterizations. Future work can further explore the sensitivity of model results to these parameterization choices.

Our findings have direct implications for HWO target prioritization. The strong dependence of outer HZ climate stability on poorly constrained geophysical processes – particularly seafloor weathering during glaciated states – introduces fundamental uncertainty in predicting which outer HZ planets are habitable at any given point in their evolution. Planets near the inner HZ edge, where the predominant control on habitability may be atmospheric water loss or the runaway greenhouse limit, may therefore represent higher priority targets. Among stellar types, K-dwarfs present a nuanced case: their slow luminosity evolution and long main sequence lifetimes are observationally advantageous, but our results suggest outer HZ K-dwarf planets are particularly susceptible to permanent glaciation. Assuming Earth-like internal heat budgets and degassing histories, inner HZ K-dwarf planets may therefore

offer the best combination of long-term habitability and observational accessibility, though this pessimistic HZ is notably extremely narrow.

1.3 Transmission vs. Direct Imaging for Habitability and Biosignature Assessment

In Chapter 4, we used a novel terrestrial exoplanet retrieval model to compare the capabilities of transmission and direct imaging spectroscopy for habitability assessment and biosignature detection in rocky, habitable zone exoplanet atmospheres.

We found that transmission spectroscopy with a 5.9-m *Origins*-like telescope will not be able to probe the surface and near-surface environment of M-dwarf planets like TRAPPIST-1 e, hampering direct habitability assessment, but will likely have the sensitivity to detect the CO₂/CH₄ biosignature pair in 40-hour exposures. We find that transmission probes the stratosphere (50–70 km altitudes) rather than surface conditions, with continuum pressure constraints providing only broad lower bounds on surface pressure. Despite these limitations for characterizing habitability, transmission can discriminate between biological and false-positive O₃ scenarios: in the context of an O₃ detection, the non-detection of water vapor (3- σ upper limit of 13 ppm) coupled to the detection of abundant CO (1- σ range of 1-26%) could distinguish O₃ sourced from CO₂ photolysis from biogenic O₃ sourced from abundant O₂.

For direct imaging missions with a 15-m LUVOIR-A-like telescope, we cannot constrain surface temperature at the spectral resolutions explored here, and planetary albedo is degenerate with planetary radius. However, our model constrains near-surface water abundance to within 0.1% of the true value at high significance (>10- σ) in only 10-hour exposures, providing a critical habitability indicator. We detect O₂ and O₃ with high significances (11- σ and 19- σ , respectively) and precisely constrain their abundances, but fail to detect modern-Earth-like CH₄ abundances on planets around G stars. The predicted photochemical enhancement of CH₄ on M- and K-dwarf planets may make the O₂/CH₄ biosignature

pair more feasible in these systems, which constitute 30 of HWO’s planned targets. We find weak sensitivity to the vertical structure of the stratospheric ozone bulge via the near-UV Hartley band (Bayes factor ~ 1.3), and evidence for water clouds in large residuals at the bottoms of water bands and NIR ice absorption features.

1.4 Spectral Resolution Requirements for HWO

In Chapter 5, we assessed how spectral resolution affects HWO’s ability to detect and characterize biosignature gases using complementary detectability calculations and atmospheric retrievals across Earth-through-time atmospheres. In the visible, the nominal resolution $R_{\text{vis}} = 140$ is sufficient for detecting O_2 in Phanerozoic-like atmospheres. While higher resolutions could theoretically reduce exposure times for low- O_2 Proterozoic atmospheres, the required detector noise improvements ($>10\times$ reduction in dark current) may increase technological risk. Furthermore, increasing visible spectral resolution could increase H_2O detection exposure times by up to $\sim 2\times$, penalizing the foundational habitability constraint that anchors downstream biosignature searches.

In the near-IR, we found that $R_{\text{NIR}} \geq 40$ is necessary to avoid a degeneracy between CO_2 and CO that could produce false positive detections of abundant CO – a critical anti-biosignature discriminant. The nominal resolution $R_{\text{NIR}} = 70$ provides robust constraints on CO_2 and H_2O across all Earth-through-time cases. These results support HWO’s current baseline resolution choices while providing actionable guidance and flexibility for finalizing spectrometer requirements.

2 Future Directions

2.1 Expanding the Abiotic Planet Parameter Space

Our core-mantle-crust-climate model represents significant progress toward whole-planet modeling, but additional development is required to fully quantify the abiotic null hypothesis.

Future work should explore:

- **Non-Earth-like compositions:** Rocky exoplanets likely span a range of bulk compositions, volatile inventories, and ocean masses. Systematic exploration of how these parameters affect long-term climate stability and atmospheric composition will illuminate the diversity of potential equilibrium states for rocky planets.
- **Additional species and evolving mantle redox state:** Our current work focuses on the cycling of CO_2 and H_2O for an oxidized mantle. Future work should incorporate reduced species like CH_4 and mantle redox balance.
- **Enhanced coupling:** Our model is initialized after the solidification of the magma ocean and neglects atmospheric escape and energy from early impact events. Future work should allow for continuous evolution through the magma ocean into the solid mantle, as this evolution establishes the early atmospheric composition. Similarly, future work should couple our model to giant impacts during formation and atmospheric escape, as both volatile-loss processes play a significant role in shaping long-term atmospheric evolution.
- **Direct comparison to 3-D models:** Our model is a simplified framework for predicting globally-averaged quantities over time. Future work should assess the robustness of these predictions by comparing to GCMs. Though GCMs cannot currently span the billion-year timescales of interest, evolutionary “snapshots” in time from our results would facilitate direct comparison with GCM climate predictions.

- **Predicting abiotic planetary evolution:** Our whole-planet model is well-suited to exploring the long-term atmospheric evolution of uninhabited worlds and how their corresponding reflected light spectra change over time. Thus, our model can be used to help quantify an abiotic baseline for predicting and later interpreting HWO spectroscopic observations.

2.2 Estimating the HZs of HWO FGK Stars

Applying our whole-planet model to evolving Sun-like stars will shed light on the evolving HZ boundaries of key HWO target stars. However, further work is required to ensure our conclusions are robust to assumptions about continental and seafloor weathering on Earth-like planets and external forces we neglected. Future work should explore:

- **Re-tuning the weathering model:** We recommend identifying the weathering parameterizations that reproduce early Earth history within measurement uncertainties. This will help us reduce some of the uncertainty associated with our modeling choices. However, there will likely still be some uncertainty associated with the values of activation energies and other key parameters that control the weathering rates. We should marginalize over these uncertainties by using MCMC to explore a range of values for such parameters (see next point).
- **Broader parameter exploration:** We should sweep over key properties of the interior (e.g., budget of radionuclides, energy from early impact events). This expansion will build on previous work that directly varied the CO₂ degassing rate by allowing it to scale self-consistently with realistic interior dynamics. A broader parameter exploration will also help quantify and assess the unique role of the interior in shaping the evolution of habitable planets.

2.3 Retrieval Model Development

Future areas of research for atmospheric retrieval models include:

- **Cloud parameterization:** Our finding that water cloud features may be detectable in NIR ice absorption and water band residuals motivates explicit wavelength-dependent cloud modeling in retrieval frameworks.
- **UV wavelength analysis:** Extending our spectral resolution analysis to UV wavelengths (0.2–0.4 μm) is critical for assessing O_3 detection strategies in low- O_2 atmospheres, where indirect biosignature inference may be more efficient than direct O_2 detection.
- **Abiotic false positive scenarios:** Our retrieval work focused primarily on Earth-through-time atmospheres. Systematic application to abiotic false positive cases (e.g., CO_2 -dominated atmospheres with photolytic O_2/O_3 , post-ocean-loss desiccated worlds) will establish the observational discriminants needed to rule out these scenarios.
- **Multi-epoch observations of 3-D planets:** Though this work has consistently used 1-D modeling techniques, planets are three-dimensional objects. While HWO observations will represent a spectrally-integrated disk, it is unclear how 3-D, planetary-scale effects may affect our ability to properly interpret spectral observations and accurately characterize exoplanet atmospheres. Temporal variability in cloud coverage, surface features, and atmospheric composition could provide additional constraints. Future work should assess how repeated observations over orbital phases or multiple years improve retrieval constraints.

2.4 Mission Trade Studies and Target Prioritization

Our findings inform several aspects of HWO mission development. Future work should explore:

- **The K-dwarf advantage?:** K-dwarf planets near the inner HZ edge may represent more robust long-term habitability targets than outer HZ planets, where the climate trajectory is highly sensitive to interior evolution. Future work should continue to investigate the evolving inner HZ of cooler K stars for a broader range of stellar masses and rotation rates, and confirm whether the HZ of K stars is as pessimistically narrow as our predictions.
- **Complementarity with transmission missions:** Our comparative analysis of transmission and direct imaging reveals complementary strengths. Future work should explore observing strategies that combine both techniques, potentially using transmission observations from ground-based extremely large telescopes or space missions for initial screening followed by HWO direct imaging for detailed characterization.
- **Exposure time optimization:** Our spectral resolution work reveals trade-offs between different science goals (O_2 detection vs. H_2O characterization). Mission-level optimization should consider these trade-offs in the context of the full target list to maximize overall science return.
- **Other science cases:** HWO will serve a broad range of science outside of the search for life on exoplanets. While this science priority will likely be used to establish the baseline telescope capabilities, future work should investigate the possible benefits of higher spectral resolving power for other exciting science cases, including characterizing the atmospheres of Super-Earths and Gas Giants.

3 Concluding Remarks

This thesis has treated Earth as an exoplanet, using our own world to establish the theoretical and observational frameworks needed to search for life elsewhere. We have shown that:

1. Habitable conditions can persist without life, requiring that we distinguish inhabited

from merely habitable worlds through spectroscopic biosignature detection rather than habitability assessment alone.

2. Planetary interiors influence atmospheric evolution and long-term climate stability, with different stellar types producing characteristically different habitability trajectories.
3. Transmission and direct imaging probe different atmospheric regions and excel at detecting different molecular species, with implications for biosignature search strategies.
4. HWO's baseline spectral resolution choices are well-motivated for detecting biosignatures across a range of Earth-like atmospheric compositions.

By developing whole-planet models validated against Earth's pre-industrial state, extending these models to HWO target stars, and establishing the retrieval capabilities needed to interpret future observations, this work contributes directly to HWO mission development and the broader goal of detecting life beyond Earth. The search for extraterrestrial life is no longer a purely theoretical exercise, but a scientific and engineering challenge with concrete observational requirements and testable predictions. The framework established here provides actionable guidance for converting that challenge into meaningful discovery as we approach the next frontier of space exploration.

BIBLIOGRAPHY

- Abbot, D. S.** 2016, *Analytical investigation of the decrease in the size of the habitable zone due to a limited CO₂ outgassing rate*, The Astrophysical Journal, 827, 117
- Abbot, D. S., Voigt, A., Li, D., et al.** 2013, *Robust elements of Snowball Earth atmospheric circulation and oases for life*, Journal of Geophysical Research: Atmospheres, 118, 6017
- Ackerman, A. S., & Marley, M. S.** 2001, *Precipitating condensation clouds in substellar atmospheres*, The Astrophysical Journal, 556, 872
- Aitchison, J.** 1982, *The statistical analysis of compositional data*, Journal of the Royal Statistical Society: Series B (Methodological), 44, 139
- Al Asad, M., & Lau, H. C.** 2024, *Coupled fates of Earth’s mantle and core: Early sluggish-lid tectonics and a long-lived geodynamo*, Science Advances, 10, eadp1991
- Alcabes, O. D., Olson, S., & Abbot, D. S.** 2020, *Robustness of Gaian feedbacks to climate perturbations*, Monthly Notices of the Royal Astronomical Society, 492, 2572
- Alei, E., Konrad, B., Mollière, P., et al.** 2022, in *Space Telescopes and Instrumentation 2022: Optical, Infrared, and Millimeter Wave*, ed. L. E. Coyle, M. D. Perrin, & S. Matsuura (SPIE), doi: [10.1117/12.2631692](https://doi.org/10.1117/12.2631692)
- Alei, E., Mandell, A. M., Currie, M. H., et al.** 2025, *Multi-bandpass Photometry for Exoplanet Atmosphere Reconnaissance (MPEAR) with the Habitable Worlds Observatory (HWO) – I. Differentiating Earth from Neptunes During Discovery*, arXiv e-prints, arXiv:2512.05279, doi: [10.48550/arXiv.2512.05279](https://doi.org/10.48550/arXiv.2512.05279)
- Alfè, D., Gillan, M., & Price, G. D.** 2002, *Composition and temperature of the Earth’s core constrained by combining ab initio calculations and seismic data*, Earth and Planetary Science Letters, 195, 91
- Allan, R. P., Willett, K. M., John, V. O., & Trent, T.** 2022, *Global changes in water vapor 1979–2020*, Journal of Geophysical Research: Atmospheres, 127, e2022JD036728
- Anglada-Escudé, G., Amado, P. J., Barnes, J., et al.** 2016, *A terrestrial planet candidate in a temperate orbit around Proxima Centauri*, Nature, 536, 437
- Armann, M., & Tackley, P. J.** 2012, *Simulating the thermochemical magmatic and tectonic evolution of Venus’s mantle and lithosphere: Two-dimensional models*, Journal of Geophysical Research: Planets, 117
- Arney, G., Domagal-Goldman, S. D., & Meadows, V. S.** 2018, *Organic haze as a biosignature in anoxic Earth-like atmospheres*, Astrobiology, 18, 311
- Arney, G., Meadows, V., Crisp, D., et al.** 2014, *Spatially resolved measurements of H₂O, HCl, CO, OCS, SO₂, cloud opacity, and acid concentration in the Venus near-infrared spectral windows*, Journal of Geophysical Research (Planets), 119, 1860, doi: [10.1002/2014JE004662](https://doi.org/10.1002/2014JE004662)
- Arney, G., Domagal-Goldman, S. D., Meadows, V. S., et al.** 2016, *The pale orange dot: the spectrum and habitability of hazy Archean Earth*, Astrobiology, 16, 873

- Arney, G. N. 2019, *The K dwarf advantage for biosignatures on directly imaged exoplanets*, The Astrophysical Journal Letters, 873, L7
- Arney, G. N., Meadows, V. S., Domagal-Goldman, S. D., et al. 2017, *Pale orange dots: the impact of organic haze on the habitability and detectability of Earthlike exoplanets*, The Astrophysical Journal, 836, 49
- Atri, D., & Mogan, S. R. C. 2021, *Stellar flares versus luminosity: XUV-induced atmospheric escape and planetary habitability*, Monthly Notices of the Royal Astronomical Society: Letters, 500, L1
- Badr, O., & Probert, S. 1994, *Carbon monoxide concentration in the Earth's atmosphere*, Applied Energy, 49, 99
- Baraffe, I., Homeier, D., Allard, F., & Chabrier, G. 2015, *New evolutionary models for pre-main sequence and main sequence low-mass stars down to the hydrogen-burning limit*, Astronomy & Astrophysics, 577, A42
- Barley, M. E., Bekker, A., & Krapež, B. 2005, *Late Archean to Early Paleoproterozoic global tectonics, environmental change and the rise of atmospheric oxygen*, Earth and Planetary Science Letters, 238, 156
- Barnes, R., Luger, R., Deitrick, R., et al. 2020, *VPLANET: the virtual planet simulator*, Publications of the Astronomical Society of the Pacific, 132, 024502
- Barstow, J. K., Changeat, Q., Garland, R., et al. 2020, *A comparison of exoplanet spectroscopic retrieval tools*, Monthly Notices of the Royal Astronomical Society, 493, 4884
- Barstow, J. K., & Irwin, P. G. 2016, *Habitable worlds with JWST: transit spectroscopy of the TRAPPIST-1 system?*, Monthly Notices of the Royal Astronomical Society: Letters, 461, L92
- Bartman, F. L. 1980, *A time variable model of earth's albedo*, Tech. rep., NASA
- Batalha, N. E., Lewis, N. K., Line, M. R., Valenti, J., & Stevenson, K. 2018, *Strategies for constraining the atmospheres of temperate terrestrial planets with JWST*, The Astrophysical Journal Letters, 856, L34
- Batalha, N. E., Marley, M. S., Lewis, N. K., & Fortney, J. J. 2019, *Exoplanet reflected-light spectroscopy with PICASO*, The Astrophysical Journal, 878, 70
- Beatty, T. G., & Gaudi, B. S. 2008, *Predicting the yields of photometric surveys for transiting extrasolar planets*, The Astrophysical Journal, 686, 1302
- Bellefroid, E. J., Hood, A. v. S., Hoffman, P. F., et al. 2018, *Constraints on Paleoproterozoic atmospheric oxygen levels*, Proceedings of the National Academy of Sciences, 115, 8104
- Benneke, B., & Seager, S. 2012, *Atmospheric retrieval for super-Earths: uniquely constraining the atmospheric composition with transmission spectroscopy*, The Astrophysical Journal, 753, 100
- . 2013, *How to distinguish between cloudy mini-Neptunes and water/volatile-dominated super-Earths*, The Astrophysical Journal, 778, 153
- Benneke, B., Wong, I., Piaulet, C., et al. 2019, *Water vapor and clouds on the habitable-zone sub-Neptune exoplanet K2-18b*, The Astrophysical Journal Letters, 887, L14
- Berner, R. A., Lasaga, A. C., & Garrels, R. M. 1983, *Carbonate-silicate geochemical cycle and its effect on atmospheric carbon dioxide over the past 100 million years*, Am. J. Sci.:(United States), 283
- Bétrémieux, Y., & Kaltenegger, L. 2014, *Impact of atmospheric refraction: how deeply can we probe exo-Earth's atmospheres during primary eclipse observations?*, The Astrophysical Journal, 791, 7
- Bialik, O. M., Sisma-Ventura, G., Vogt-Vincent, N., Silverman, J., & Katz, T. 2022, *Role of oceanic abiotic carbonate precipitation in future atmospheric CO₂ regulation*, Scientific Reports, 12, 15970, doi: [10.1038/s41598-022-20446-7](https://doi.org/10.1038/s41598-022-20446-7)
- Blake, R. E., Chang, S. J., & Lepland, A. 2010, *Phosphate oxygen isotopic evidence for a temperate and biologically active Archean ocean*, Nature, 464, 1029

BIBLIOGRAPHY

- Blanchard, I., Siebert, J., Borensztajn, S., & Badro, J. 2017, *The solubility of heat-producing elements in Earth's core*, *Geochemical Perspectives Letters*, 5, 1
- Boehler, R. 1996, *Melting temperature of the Earth's mantle and core: Earth's thermal structure*, *Annual Review of Earth and Planetary Sciences*, 24, 15
- Bolcar, M. R., Feinberg, L., France, K., et al. 2016, in *Space telescopes and instrumentation 2016: Optical, infrared, and millimeter wave*, Vol. 9904, SPIE, 192–203
- Borges, S. R., Jones, G. G., & Robinson, T. D. 2024, *Detectability of Surface Biosignatures for Directly Imaged Rocky Exoplanets*, *Astrobiology*, 24, 283
- Boschi, R., Lucarini, V., & Pascale, S. 2013, *Bistability of the climate around the habitable zone: A thermodynamic investigation*, *Icarus*, 226, 1724
- Boukaré, C.-É., Badro, J., & Samuel, H. 2025, *Solidification of Earth's mantle led inevitably to a basal magma ocean*, *Nature*, 1
- Boyd, P. W., Claustre, H., Levy, M., Siegel, D. A., & Weber, T. 2019, *Multi-faceted particle pumps drive carbon sequestration in the ocean*, *Nature*, 568, 327
- Brandt, T. D., & Spiegel, D. S. 2014, *Prospects for detecting oxygen, water, and chlorophyll on an exo-Earth*, *Proceedings of the National Academy of Sciences*, 111, 13278
- Brent, R. P. 2013, *Algorithms for minimization without derivatives* (Courier Corporation)
- Brewer, A. 1949, *Evidence for a world circulation provided by the measurements of helium and water vapour distribution in the stratosphere*, *Quarterly Journal of the Royal Meteorological Society*, 75, 351
- Buchner, J., Georgakakis, A., Nandra, K., et al. 2014, *X-ray spectral modelling of the AGN obscuring region in the CDFS: Bayesian model selection and catalogue*, *Astronomy & Astrophysics*, 564, A125
- Caldas, A., Leconte, J., Selsis, F., et al. 2019, *Effects of a fully 3D atmospheric structure on exoplanet transmission spectra: retrieval biases due to day–night temperature gradients*, *Astronomy & Astrophysics*, 623, A161
- Cameron, E. M., Blum, T. B., Cavosie, A. J., et al. 2024, *Evidence for oceans pre-4300 Ma confirmed by preserved igneous compositions in Hadean zircon*, *American Mineralogist*, 109, 1670
- Campbell, I. H., & O'Neill, H. S. C. 2012, *Evidence against a chondritic Earth*, *Nature*, 483, 553
- Carlson, R. W., Garnero, E., Harrison, T. M., et al. 2014, *How did early Earth become our modern world?*, *Annual Review of Earth and Planetary Sciences*, 42, 151
- Carone, L., Barnes, R., Noack, L., et al. 2025, *From CO₂-to H₂O-dominated atmospheres and back-How mixed outgassing changes the volatile distribution in magma oceans around M dwarf stars*, *Astronomy & Astrophysics*, 693, A303
- Carpenter, S. J., & Lohmann, K. C. 1992, *SrMg ratios of modern marine calcite: Empirical indicators of ocean chemistry and precipitation rate*, *Geochimica et Cosmochimica Acta*, 56, 1837
- Carrión-González, Ó., García Muñoz, A., Cabrera, J., et al. 2020, *Directly imaged exoplanets in reflected starlight: the importance of knowing the planet radius*, *A&A*, 640, A136, doi: [10.1051/0004-6361/202038101](https://doi.org/10.1051/0004-6361/202038101)
- Carroll, D., Menemenlis, D., Dutkiewicz, S., et al. 2022, *Attribution of space-time variability in global-ocean dissolved inorganic carbon*, *Global Biogeochemical Cycles*, 36, e2021GB007162
- Catling, D. C., Kiang Nancy, Y., Robinson Tyler, D., et al. 2018, *Exoplanet biosignatures: a framework for their assessment*, *Astrobiology*
- Catling, D. C., & Zahnle, K. J. 2020, *The Archean atmosphere*, *Science advances*, 6, eaax1420
- Catling, D. C., Zahnle, K. J., & McKay, C. P. 2001, *Biogenic methane, hydrogen escape, and the irreversible oxidation of early Earth*, *Science*, 293, 839

- Chadney, J., Galand, M., Unruh, Y., Koskinen, T., & Sanz-Forcada, J. 2015, *XUV-driven mass loss from extrasolar giant planets orbiting active stars*, *Icarus*, 250, 357
- Chamberlain, S., Mahieux, A., Robert, S., et al. 2020, *SOIR/VEx observations of water vapor at the terminator in the Venus mesosphere*, *Icarus*, 346, 113819
- Changeat, Q., Edwards, B., Waldmann, I., & Tinetti, G. 2019, *Toward a more complex description of chemical profiles in exoplanet retrievals: A two-layer parameterization*, *The Astrophysical Journal*, 886, 39
- Chao, K.-H., deGraffenried, R., Lach, M., et al. 2021, *Lava worlds: From early earth to exoplanets*, *Geochemistry*, 81, 125735
- Charette, M. A., & Smith, W. H. 2010, *The volume of Earth's ocean*, *Oceanography*, 23, 112
- Charnay, B., Forget, F., Wordsworth, R., et al. 2013, *Exploring the faint young Sun problem and the possible climates of the Archean Earth with a 3-D GCM*, *Journal of Geophysical Research: Atmospheres*, 118, 10
- Charnay, B., Meadows, V., Misra, A., Leconte, J., & Arney, G. 2015, *3D modeling of GJ1214b's atmosphere: formation of inhomogeneous high clouds and observational implications*, *The Astrophysical Journal Letters*, 813, L1
- Chave, K. E., & Suess, E. 1970, *Calcium carbonate saturation in seawater: effects of dissolved organic matter*, *Limnology and Oceanography*, 15, 633
- Christiansen, J. L., McElroy, D. L., Harbut, M., et al. 2025, *The NASA exoplanet archive and exoplanet follow-up observing program: Data, tools, and usage*, *The Planetary Science Journal*, 6, 186
- Chu, D. 2019, in *Remote sensing of land use and land cover in mountain region: a comprehensive study at the central Tibetan Plateau* (Springer), 195–207
- Clark, R. N., Swayze, G. A., Wise, R., et al. 2003, *USGS Digital Spectral Library splib05a*, Open File Report 03-395, U.S. Geological Survey. <http://pubs.usgs.gov/of/2003/ofr-03-395/datatable.html>
- Cochran, M. F., & Berner, R. A. 1996, *Promotion of chemical weathering by higher plants: field observations on Hawaiian basalts*, *Chemical Geology*, 132, 71
- Cockell, C. S. 2014, *Habitable worlds with no signs of life*, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 372, 20130082
- Cogné, J.-P., & Humler, E. 2004, *Temporal variation of oceanic spreading and crustal production rates during the last 180 My*, *Earth and Planetary Science Letters*, 227, 427
- Condie, K. C., Aster, R. C., & Van Hunen, J. 2016, *A great thermal divergence in the mantle beginning 2.5 Ga: Geochemical constraints from greenstone basalts and komatiites*, *Geoscience Frontiers*, 7, 543
- Constantinou, T., Shorttle, O., Cranmer, M., & Rimmer, P. B. 2025, *Comparative Biosignatures*, arXiv preprint arXiv:2505.01512
- Costa, A., Caricchi, L., & Bagdassarov, N. 2009, *A model for the rheology of particle-bearing suspensions and partially molten rocks*, *Geochemistry, Geophysics, Geosystems*, 10
- Cracraft, M., De Rosa, R., Sparks, W., Bailey, V., & Turnbull, M. 2021, *Characterizing the Galactic and Extragalactic Background Near Exoplanet Direct Imaging Targets*, arXiv preprint arXiv:2110.08097
- Crisp, D., & Titov, D. 1997, in *Venus II: Geology, Geophysics, Atmosphere, and Solar Wind Environment*, ed. S. W. Bougher, D. M. Hunten, & R. J. Phillips, 353
- Damiano, M., & Hu, R. 2021, *Reflected spectroscopy of small exoplanets I: determining the atmospheric composition of sub-Neptunes planets*, *The Astronomical Journal*, 162, 200
- . 2022, *Reflected spectroscopy of small exoplanets II: characterization of terrestrial exoplanets*, *The Astronomical Journal*, 163, 299

BIBLIOGRAPHY

- Damiano, M., Hu, R., & Mennesson, B. 2023, *Reflected spectroscopy of small exoplanets. III. Probing the UV band to measure biosignature gases*, The Astronomical Journal, 166, 157
- Davies, J. K., Wright, G. S., & Glasse, A. C. 2009, *Solar System Observations with MIRI, The Mid InfraRed Instrument on the James Webb Space Telescope*, Earth, Moon, and Planets, 105, 73
- de Pater, I., & Lissauer, J. 2010, Planetary Sciences (Cambridge University Press,)
- De Wit, M. J., & Furnes, H. 2016, *3.5-Ga hydrothermal fields and diamictites in the Barberton Greenstone Belt—Paleoarchean crust in cold environments*, Science advances, 2, e1500368
- Deitrick, R., Barnes, R., Bitz, C., et al. 2018, *Exo-Milankovitch cycles. II. Climates of G-dwarf planets in dynamically hot systems*, The astronomical journal, 155, 266
- Des Marais, D. J., Harwit, M. O., Jucks, K. W., et al. 2002, *Remote sensing of planetary properties and biosignatures on extrasolar terrestrial planets*, Astrobiology, 2, 153
- Dessert, C., Dupré, B., Gaillardet, J., François, L. M., & Allègre, C. J. 2003, *Basalt weathering laws and the impact of basalt weathering on the global carbon cycle*, Chemical Geology, 202, 257
- Dittmann, J. A., Irwin, J. M., Charbonneau, D., et al. 2017, *A temperate rocky super-Earth transiting a nearby cool star*, Nature, 544, 333
- Dlugokencky, E. J., Nisbet, E. G., Fisher, R., & Lowry, D. 2011, *Global atmospheric methane: budget, changes and dangers*, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 369, 2058
- Dole, S. H., & Asimov, I. 1964, *Planets for man*, New York
- Domagal-Goldman, S. D., Segura, A., Claire, M. W., Robinson, T. D., & Meadows, V. S. 2014, *Abiotic ozone and oxygen in atmospheres similar to prebiotic Earth*, The Astrophysical Journal, 792, 90
- Driese, S. G., Jirsa, M. A., Ren, M., et al. 2011, *Neoproterozoic paleoweathering of tonalite and metabasalt: Implications for reconstructions of 2.69 Ga early terrestrial ecosystems and paleoatmospheric chemistry*, Precambrian Research, 189, 1
- Driscoll, P., & Bercovici, D. 2013, *Divergent evolution of Earth and Venus: influence of degassing, tectonics, and magnetic fields*, Icarus, 226, 1447
- . 2014, *On the thermal and magnetic histories of Earth and Venus: Influences of melting, radioactivity, and conductivity*, Physics of the Earth and Planetary Interiors, 236, 36
- Driscoll, P. E., & Barnes, R. 2015, *Tidal heating of Earth-like exoplanets around M stars: thermal, magnetic, and orbital evolutions*, Astrobiology, 15, 739
- Dziewonski, A. M., & Anderson, D. L. 1981, *Preliminary reference Earth model*, Physics of the earth and planetary interiors, 25, 297
- Ebadirad, S., Lyons, T. W., & Fournier, G. P. 2025, *Archean Methane Cycling and Life's Co-Evolution: Intertwining Early Biogeochemical Processes and Ancient Microbial Metabolism*, arXiv preprint arXiv:2503.21919
- Eide, M., Olsen, A., Ninnemann, U. S., & Johannessen, T. 2017, *A global ocean climatology of preindustrial and modern ocean $\delta^{13}C$* , Global Biogeochemical Cycles, 31, 515
- Elkins-Tanton, L. T. 2008, *Linked magma ocean solidification and atmospheric growth for Earth and Mars*, Earth and Planetary Science Letters, 271, 181
- . 2012, *Magma oceans in the inner solar system*, Annual Review of Earth and Planetary Sciences, 40, 113
- Etheridge, D. M., Steele, L., Langenfelds, R. L., et al. 1996, *Natural and anthropogenic changes in atmospheric CO₂ over the last 1000 years from air in Antarctic ice and firn*, Journal of Geophysical Research: Atmospheres, 101, 4115
- Etiopo, G., & Sherwood Lollar, B. 2013, *Abiotic methane on Earth*, Reviews of Geophysics, 51, 276

- Fan, S., Li, C., Li, J.-Z., et al. 2019, *Earth as an exoplanet: A two-dimensional alien map*, The Astrophysical Journal Letters, 882, L1
- Farquhar, J., Zerkle, A. L., & Bekker, A. 2011, *Geological constraints on the origin of oxygenic photosynthesis*, Photosynthesis research, 107, 11
- Fauchez, T., Arney, G., Kopparapu, R. K., & Goldman, S. D. 2018, *Explicit cloud representation in the Atmos 1D climate model for Earth and rocky planet applications*, arXiv preprint arXiv:1811.10759
- Fauchez, T. J., Turbet, M., Villanueva, G. L., et al. 2019, *Impact of clouds and hazes on the simulated JWST transmission spectra of habitable zone planets in the TRAPPIST-1 system*, The Astrophysical Journal, 887, 194
- Faure, P., Bouhifd, M. A., Boyet, M., et al. 2020, *Uranium and thorium partitioning in the bulk silicate Earth and the oxygen content of Earth's core*, Geochimica et Cosmochimica Acta, 275, 83
- Feely, R. A., Sabine, C. L., Takahashi, T., Wanninkhof, R., et al. 2001, *Uptake and storage of carbon dioxide in the ocean: The global CO_2 survey*, Oceanography, 14, 18
- Feinberg, L. D., Sitariski, B. N., McElwain, M. W., et al. 2026, *Habitable Worlds Observatory's Concept and Technology Maturation: Initial Feasibility and Trade Space Exploration*, arXiv preprint arXiv:2601.11803
- Feng, Y. K., Line, M. R., & Fortney, J. J. 2020, *2D retrieval frameworks for hot Jupiter phase curves*, The Astronomical Journal, 160, 137
- Feng, Y. K., Line, M. R., Fortney, J. J., et al. 2016, *The Impact of Non-uniform Thermal Structure on the Interpretation of Exoplanet Emission Spectra*, ApJ, 829, 52, doi: [10.3847/0004-637X/829/1/52](https://doi.org/10.3847/0004-637X/829/1/52)
- Feng, Y. K., Robinson, T. D., Fortney, J. J., et al. 2018, *Characterizing Earth analogs in reflected light: atmospheric retrieval studies for future space telescopes*, The Astronomical Journal, 155, 200
- Feroz, F., Hobson, M., & Bridges, M. 2009, *MultiNest: an efficient and robust Bayesian inference tool for cosmology and particle physics*, Monthly Notices of the Royal Astronomical Society, 398, 1601
- Feroz, F., & Hobson, M. P. 2008, *Multimodal nested sampling: an efficient and robust alternative to Markov Chain Monte Carlo methods for astronomical data analyses*, Monthly Notices of the Royal Astronomical Society, 384, 449
- Feroz, F., Hobson, M. P., Cameron, E., & Pettitt, A. N. 2013, *Importance nested sampling and the MultiNest algorithm*, arXiv preprint arXiv:1306.2144
- Fischer, T. P. 2008, *Fluxes of volatiles (H_2O , CO_2 , N_2 , Cl , F) from arc volcanoes*, Geochemical Journal, 42, 21
- Fleming, J. R. 1998, *Historical perspectives on climate change* (Oxford University Press)
- Foley, B. J. 2015, *The role of plate tectonic-climate coupling and exposed land area in the development of habitable climates on rocky planets*, The Astrophysical Journal, 812, 36
- Foley, B. J., & Driscoll, P. E. 2016, *Whole planet coupling between climate, mantle, and core: Implications for rocky planet evolution*, Geochemistry, Geophysics, Geosystems, 17, 1885
- Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, *emcee: the MCMC hammer*, Publications of the Astronomical Society of the Pacific, 125, 306
- Fortney, J., Robinson, T., Domagal-Goldman, S., et al. 2019, *The Need for Laboratory Measurements and Ab Initio Studies to Aid Understanding of Exoplanetary Atmospheres*, Astro2020: Decadal Survey on Astronomy and Astrophysics, 2020, 146
- Fortney, J. J. 2005, *The effect of condensates on the characterization of transiting planet atmospheres with transmission spectroscopy*, Monthly Notices of the Royal Astronomical Society, 364, 649
- Frank, E. A., Meyer, B. S., & Mojzsis, S. J. 2014, *A radiogenic heating evolution model for cosmochemically Earth-like exoplanets*, Icarus, 243, 274

BIBLIOGRAPHY

- Fujii, Y., Lustig-Yaeger, J., & Cowan, N. B. 2017, *Rotational spectral unmixing of exoplanets: degeneracies between surface colors and geography*, The Astronomical Journal, 154, 189
- Gaillard, F., Bernadou, F., Roskosz, M., et al. 2022, *Redox controls during magma ocean degassing*, Earth and Planetary Science Letters, 577, 117255
- Gale, A., Dalton, C. A., Langmuir, C. H., Su, Y., & Schilling, J.-G. 2013, *The mean composition of ocean ridge basalts*, Geochemistry, Geophysics, Geosystems, 14, 489
- Gandhi, S., & Madhusudhan, N. 2018, *Retrieval of exoplanet emission spectra with HyDRA*, Monthly Notices of the Royal Astronomical Society, 474, 271
- Gao, P., Hu, R., Robinson, T. D., Li, C., & Yung, Y. L. 2015, *Stability of CO₂ atmospheres on desiccated M dwarf exoplanets*, The Astrophysical Journal, 806, 249
- Garcia, R., Barnes, R. K., Driscoll, P. E., Gialluca, M. T., & Meadows, V. S. 2026, *Investigation of Venus' thermal history, crustal evolution, and core dynamics with a coupled interior-lithosphere-atmosphere model*, The Planetary Science Journal, submitted
- Gaudi, B. S., Seager, S., Mennesson, B., et al. 2020, *The Habitable Exoplanet Observatory (HabEx) mission concept study final report*, arXiv preprint arXiv:2001.06683
- Gebauer, S., Grenfell, J., Stock, J. W., et al. 2017, *Evolution of Earth-like extrasolar planetary atmospheres: assessing the atmospheres and biospheres of early Earth analog planets with a coupled atmosphere biogeochemical model*, Astrobiology, 17, 27
- Geboy, N. J., Kaufman, A. J., Walker, R. J., et al. 2013, *Re-Os age constraints and new observations of Proterozoic glacial deposits in the Vazante Group, Brazil*, Precambrian Research, 238, 199
- Genda, H. 2016, *Origin of Earth's oceans: An assessment of the total amount, history and supply of water*, Geochemical Journal, 50, 27
- Gialluca, M. T., Barnes, R., Meadows, V. S., et al. 2024, *The Implications of Thermal Hydrodynamic Atmospheric Escape on the TRAPPIST-1 Planets*, The Planetary Science Journal, 5, 137
- Gilbert-Janizek, S., Meadows, V. S., & Lustig-Yaeger, J. 2024, *Retrieved Atmospheres and Inferred Surface Properties for Terrestrial Exoplanets Using Transmission and Reflected-light Spectroscopy*, The Planetary Science Journal, 5, 148
- Gillmann, C., Way, M. J., Avicé, G., et al. 2022, *The Long-Term Evolution of the Atmosphere of Venus: Processes and Feedback Mechanisms: Interior-Exterior Exchanges*, Space Science Reviews, 218, 56
- Gillon, M., Jehin, E., Lederer, S. M., et al. 2016, *Temperate Earth-sized planets transiting a nearby ultracool dwarf star*, Nature, 533, 221
- Gillon, M., Triaud, A. H., Demory, B.-O., et al. 2017, *Seven temperate terrestrial planets around the nearby ultracool dwarf star TRAPPIST-1*, Nature, 542, 456
- Glidden, A., Ranjan, S., Seager, S., et al. 2025, *JWST-TST DREAMS: secondary atmosphere constraints for the habitable zone planet TRAPPIST-1 e*, The Astrophysical Journal Letters, 990, L53
- Godolt, M., Grenfell, J., Kitzmann, D., et al. 2016, *Assessing the habitability of planets with Earth-like atmospheres with 1D and 3D climate modeling*, Astronomy & Astrophysics, 592, A36
- Godolt, M., Grenfell, J., Hamann-Reinus, A., et al. 2015, *3D climate modeling of Earth-like extrasolar planets orbiting different types of host stars*, Planetary and Space Science, 111, 62
- Goldblatt, C., Claire, M. W., Lenton, T. M., et al. 2009, *Nitrogen-enhanced greenhouse warming on early Earth*, Nature Geoscience, 2, 891

- Goldblatt, C., Robinson, T. D., Zahnle, K. J., & Crisp, D. 2013, *Low simulated radiation limit for runaway greenhouse climates*, *Nature Geoscience*, 6, 661
- Gomez Barrientos, J., MacDonald, R. J., Lewis, N. K., & Kaltenegger, L. 2023, *In search of the edge: A Bayesian exploration of the detectability of red edges in exoplanet reflection spectra*, *The Astrophysical Journal*, 946, 96
- Gordon, I. E., Rothman, L. S., Hill, C., et al. 2017, *The HITRAN2016 molecular spectroscopic database*, *Journal of quantitative spectroscopy and radiative transfer*, 203, 3
- Gordon, I. E., Rothman, L. S., Hargreaves, R., et al. 2022, *The HITRAN2020 molecular spectroscopic database*, *Journal of quantitative spectroscopy and radiative transfer*, 277, 107949
- Gough, D. 1981, *Solar interior structure and luminosity variations*, *Solar Physics*, 74, 21
- Gri  meier, J.-M., Stadelmann, A., Penz, T., et al. 2004, *The effect of tidal locking on the magnetospheric and atmospheric evolution of “Hot Jupiters”*, *Astronomy & Astrophysics*, 425, 753
- Gronoff, G., Arras, P., Baraka, S., et al. 2020, *Atmospheric escape processes and planetary atmospheric evolution*, *Journal of Geophysical Research: Space Physics*, 125, e2019JA027639
- Grotzinger, J. P. 1989, *Facies and evolution of Precambrian carbonate depositional systems: Emergence of the modern platform archetype*, *SEPM Special Publication*, 44, 79
- Grotzinger, J. P., & James, N. P. 2000, in *Carbonate Sedimentation and Diagenesis in the Evolving Precambrian World*, ed. J. P. Grotzinger & N. P. James (SEPM Society for Sedimentary Geology), 3–20
- Grotzinger, J. P., & Kasting, J. F. 1993, *New constraints on Precambrian ocean composition*, *Journal of Geology*, 101, 235, doi: 10.1086/648218
- Guimond, C. M., & Cowan, N. B. 2018, *The direct imaging search for Earth 2.0: Quantifying biases and planetary false positives*, *The Astronomical Journal*, 155, 230
- Gunell, H., Maggilo, R., Nilsson, H., et al. 2018, *Why an intrinsic magnetic field does not protect a planet against atmospheric escape*, *Astronomy & Astrophysics*, 614, L3
- Guyon, O., Martinache, F., Cady, E. J., et al. 2012, in *Adaptive Optics Systems III*, Vol. 8447, International Society for Optics and Photonics, 84471X
- Guzm  n-Marmolejo, A., Segura, A., & Escobar-Briones, E. 2013, *Abiotic production of methane in terrestrial planets*, *Astrobiology*, 13, 550
- Hagee, C., Latouf, N., Mandell, A. M., et al. 2026
- Halevy, I., & Bachan, A. 2017, *The geologic history of seawater pH*, *Science*, 355, 1069
- Hall, S., Krissansen-Totton, J., Robinson, T., Salvador, A., & Fortney, J. J. 2023, *Constraining background N₂ inventories on directly imaged terrestrial exoplanets to rule out O₂ false positives*, *The Astronomical Journal*, 166, 254
- Hamano, K., Abe, Y., & Genda, H. 2013, *Emergence of two types of terrestrial planet on solidification of magma ocean*, *Nature*, 497, 607
- Haqq-Misra, J., Kopparapu, R. K., Batalha, N. E., Harman, C. E., & Kasting, J. F. 2016, *Limit cycles can reduce the width of the habitable zone*, *The Astrophysical Journal*, 827, 120
- Haqq-Misra, J. D., Domagal-Goldman, S. D., Kasting, P. J., & Kasting, J. F. 2008, *A revised, hazy methane greenhouse for the Archean Earth*, *Astrobiology*, 8, 1127
- Harman, C., Schwieterman, E., Schottelkotte, J. C., & Kasting, J. 2015, *Abiotic O₂ levels on planets around F, G, K, and M stars: possible false positives for life?*, *The Astrophysical Journal*, 812, 137
- Harman, C. E., & Domagal-Goldman, S. 2018, *Biosignature false positives*, Tech. rep., Springer International Publishing

BIBLIOGRAPHY

- Harrison, E. F., Minnis, P., Barkstrom, B., et al. 1990, *Seasonal variation of cloud radiative forcing derived from the Earth Radiation Budget Experiment*, Journal of Geophysical Research: Atmospheres, 95, 18687
- Hashimoto, G. L., & Abe, Y. 2001, *Predictions of a simple cloud model for water vapor cloud albedo feedback on Venus*, Journal of Geophysical Research: Planets, 106, 14675
- Hawkins, E., & Sutton, R. 2016, *Connecting climate model projections of global temperature change with the real world*, Bulletin of the American Meteorological Society, 97, 963
- Hawkins, E., Ortega, P., Suckling, E., et al. 2017, *Estimating changes in global temperature since the preindustrial period*, Bulletin of the American Meteorological Society, 98, 1841
- Hegde, S., Paulino-Lima, I. G., Kent, R., Kaltenegger, L., & Rothschild, L. 2015, *Surface biosignatures of exo-Earths: remote detection of extraterrestrial life*, Proceedings of the National Academy of Sciences, 112, 3886
- Held, I. M., & Soden, B. J. 2000, *Water vapor feedback and global warming*, Annual review of energy and the environment, 25, 441
- Helling, C. 2019, *Exoplanet clouds*, Annual Review of Earth and Planetary Sciences, 47, 583
- Herzberg, C., Condie, K., & Korenaga, J. 2010, *Thermal history of the Earth and its petrological expression*, Earth and Planetary Science Letters, 292, 79
- Hirose, K., Labrosse, S., & Hernlund, J. 2013, *Composition and state of the core*, Annual Review of Earth and Planetary Sciences, 41, 657
- Hitchcock, D. R., & Lovelock, J. E. 1967, *Life detection by atmospheric analysis*, Icarus, 7, 149
- Hoffman, P. F., Kaufman, A. J., Halverson, G. P., & Schrag, D. P. 1998, *A Neoproterozoic snowball earth*, science, 281, 1342
- Hoffman, P. F., Abbot, D. S., Ashkenazy, Y., et al. 2017, *Snowball Earth climate dynamics and Cryogenian geology-geobiology*, Science Advances, 3, e1600983
- Hren, M., Tice, M., & Chamberlain, C. 2009, *Oxygen and hydrogen isotope evidence for a temperate climate 3.42 billion years ago*, Nature, 462, 205
- Hu, R., Peterson, L., & Wolf, E. T. 2020, *O₂-and CO-rich atmospheres for potentially habitable environments on TRAPPIST-1 planets*, The Astrophysical Journal, 888, 122
- Hu, R., Seager, S., & Bains, W. 2012, *Photochemistry in terrestrial exoplanet atmospheres. I. Photochemistry model and benchmark cases*, The Astrophysical Journal, 761, 166
- Huybers, P., & Tziperman, E. 2008, *Integrated summer insolation forcing and 40,000-year glacial cycles: The perspective from an ice-sheet/energy-balance model*, Paleoceanography, 23
- Irwin, P. G. J., Teanby, N. A., de Kok, R., et al. 2008, *The NEMESIS planetary atmosphere radiative transfer and retrieval tool*, J. Quant. Spec. Radiat. Transf., 109, 1136, doi: 10.1016/j.jqsrt.2007.11.006
- Jackson, M. G., Carlson, R. W., Kurz, M. D., et al. 2010, *Evidence for the survival of the oldest terrestrial mantle reservoir*, Nature, 466, 853
- Jaupart, C., Labrosse, S., Lucazeau, F., & Mareschal, J. 2015, *Treatise on geophysics. Temperatures, heat and energy in the mantle of the Earth*, Elsevier, Oxford
- Jeanloz, R., & Morris, S. 1986, *Temperature distribution in the crust and mantle*, IN: Annual review of earth and planetary sciences. Volume 14 (A87-13190 03-46). Palo Alto, CA, Annual Reviews, Inc., 1986, p. 377-415. NSF-NASA-supported research., 14, 377
- Jeffreys, H. 1998, *The theory of probability* (OUP Oxford)
- Jet Propulsion Laboratory. 1999, *ASTER Spectral Library*. <http://speclib.jpl.nasa.gov>

- Jiang, L.-Q., Carter, B. R., Feely, R. A., Lauvset, S. K., & Olsen, A. 2019, *Surface ocean pH and buffer capacity: past, present and future*, Scientific reports, 9, 18624
- Jones, E., Oliphant, T., Peterson, P., et al. 2001, *SciPy: Open source scientific tools for Python*, <https://www.scipy.org>
- Kadoya, S., & Tajika, E. 2015, *Evolutionary climate tracks of Earth-like planets*, The Astrophysical Journal Letters, 815, L7
- . 2019, *Outer limits of the habitable zones in terms of climate mode and climate evolution of earth-like planets*, The Astrophysical Journal, 875, 7
- Kah, L. C., & Riding, R. 2007, *Mesoproterozoic carbon dioxide levels inferred from calcified cyanobacteria*, Geology, 35, 799
- Kaltenegger, L., MacDonald, R. J., Kozakis, T., et al. 2020, *The White Dwarf Opportunity: Robust Detections of Molecules in Earth-like Exoplanet Atmospheres with the James Webb Space Telescope*, ApJ, 901, L1, doi: [10.3847/2041-8213/aba9d3](https://doi.org/10.3847/2041-8213/aba9d3)
- Kaltenegger, L., Traub, W. A., & Jucks, K. W. 2007, *Spectral evolution of an Earth-like planet*, The Astrophysical Journal, 658, 598
- Kanzaki, Y., & Murakami, T. 2015, *Estimates of atmospheric CO₂ in the Neoproterozoic–Paleoproterozoic from paleosols*, Geochimica et Cosmochimica Acta, 159, 190
- Kargel, J. S., Leonard, G. J., Bishop, M. P., Kääb, A., & Raup, B. H. 2014, *Global land ice measurements from space* (Springer)
- Kasting, J. F. 2005, *Methane and climate during the Precambrian era*, Precambrian Research, 137, 119
- Kasting, J. F., & Donahue, T. M. 1980, *The evolution of atmospheric ozone*, J. Geophys. Res., 85, 3255, doi: [10.1029/JC085iC06p03255](https://doi.org/10.1029/JC085iC06p03255)
- Kasting, J. F., Pollack, J. B., & Ackerman, T. P. 1984, *Response of Earth's atmosphere to increases in solar flux and implications for loss of water from Venus*, Icarus, 57, 335
- Kasting, J. F., Whitmire, D. P., & Reynolds, R. T. 1993, *Habitable zones around main sequence stars*, Icarus, 101, 108
- Katsura, T., Yamada, H., Shinmei, T., et al. 2003, *Post-spinel transition in Mg₂SiO₄ determined by high P–T in situ X-ray diffraction*, Physics of the Earth and Planetary Interiors, 136, 11
- Katsura, T., Yamada, H., Nishikawa, O., et al. 2004, *Olivine-wadsleyite transition in the system (Mg, Fe) ₂SiO₄*, Journal of Geophysical Research: Solid Earth, 109
- Kerrick, D. M. 2001, *Present and past nonanthropogenic CO₂ degassing from the solid Earth*, Reviews of Geophysics, 39, 565
- Kharecha, P., Kasting, J., & Siefert, J. 2005, *A coupled atmosphere–ecosystem model of the early Archean Earth*, Geobiology, 3, 53
- Kim, S. J., Jung, A., Sim, C., et al. 2011, *Retrieval and tentative identification of the 3 μ m spectral feature in Titan's haze*, Planetary and Space Science, 59, 699
- Kinzler, R. J., & Grove, T. L. 1992, *Primary magmas of mid-ocean ridge basalts 1. Experiments and methods*, Journal of Geophysical Research: Solid Earth, 97, 6885
- Kipping, D. 2025, *Strong Evidence that Abiogenesis Is a Rapid Process on Earth Analogs*, Astrobiology, 25, 323
- Kipping, D. M., & Sandford, E. 2016, *Observational biases for transiting planets*, Monthly Notices of the Royal Astronomical Society, 463, 1323
- Kirchner, J. W. 1989, *The Gaia hypothesis: Can it be tested?*, Reviews of Geophysics, 27, 223
- . 2002, *The Gaia hypothesis: fact, theory, and wishful thinking*, Climatic change, 52, 391

BIBLIOGRAPHY

- . 2003, *The Gaia hypothesis: conjectures and refutations*, Climatic Change, 58, 21
- Kivelson, M. G., & Bagenal, F. 2014, in *Encyclopedia of the solar system* (Elsevier), 137–157
- Klein, E. M., & Langmuir, C. H. 1987, *Global correlations of ocean ridge basalt chemistry with axial depth and crustal thickness*, Journal of Geophysical Research: Solid Earth, 92, 8089
- Kleinböhl, A., Schofield, J. T., Kass, D. M., et al. 2009, *Mars Climate Sounder limb profile retrieval of atmospheric temperature, pressure, and dust and water ice opacity*, Journal of Geophysical Research (Planets), 114, E10006, doi: 10.1029/2009JE003358
- Komacek, T. D., & Abbot, D. S. 2019, *The atmospheric circulation and climate of terrestrial planets orbiting Sun-like and M dwarf stars over a broad range of planetary parameters*, The Astrophysical Journal, 871, 245
- Komacek, T. D., Fauchez, T. J., Wolf, E. T., & Abbot, D. S. 2020, *Clouds will likely prevent the detection of water vapor in JWST transmission spectra of terrestrial exoplanets*, The Astrophysical Journal Letters, 888, L20
- Kopparapu, R., Arney, G., Haqq-Misra, J., Lustig-Yaeger, J., & Villanueva, G. 2021, *Nitrogen dioxide pollution as a signature of extraterrestrial technology*, The Astrophysical Journal, 908, 164
- Kopparapu, R. K., Ramirez, R., Kasting, J. F., et al. 2013, *Habitable zones around main-sequence stars: new estimates*, The Astrophysical Journal, 765, 131
- Korenaga, J. 2006, *Archean geodynamics and the thermal evolution of Earth*, Geophysical Monograph-American Geophysical Union, 164, 7
- . 2008a, *Plate tectonics, flood basalts and the evolution of Earth's oceans*, Terra Nova, 20, 419
- . 2008b, *Urey ratio and the structure and evolution of Earth's mantle*, Reviews of Geophysics, 46
- Kozakis, T., Mendonça, J. M., & Buchhave, L. A. 2022, *Ozone as a Proxy for Molecular Oxygen*, Bulletin of the American Astronomical Society, 54, 102
- Kreidberg, L., Bean, J. L., Désert, J.-M., et al. 2014, *Clouds in the atmosphere of the super-Earth exoplanet GJ 1214b*, Nature, 505, 69
- Kreidberg, L., Line, M. R., Bean, J. L., et al. 2015, *A detection of water in the transmission spectrum of the hot Jupiter WASP-12b and implications for its atmospheric composition*, The Astrophysical Journal, 814, 66
- Krissansen-Totton, J., Arney, G. N., & Catling, D. C. 2018a, *Constraining the climate and ocean pH of the early Earth with a geological carbon cycle model*, Proceedings of the National Academy of Sciences, 115, 4105
- Krissansen-Totton, J., Bergsman, D. S., & Catling, D. C. 2016, *On detecting biospheres from chemical thermodynamic disequilibrium in planetary atmospheres*, Astrobiology, 16, 39
- Krissansen-Totton, J., Fortney, J. J., Nimmo, F., & Wogan, N. 2021, *Oxygen false positives on habitable zone planets around sun-like stars*, AGU Advances, 2, e2020AV000294
- Krissansen-Totton, J., Garland, R., Irwin, P., & Catling, D. C. 2018b, *Detectability of biosignatures in anoxic atmospheres with the James Webb Space Telescope: A TRAPPIST-1e case study*, The Astronomical Journal, 156, 114
- Krissansen-Totton, J., Olson, S., & Catling, D. C. 2018c, *Disequilibrium biosignatures over Earth history and implications for detecting exoplanet life*, Science advances, 4, eaao5747
- Krissansen-Totton, J., Thompson, M., Galloway, M. L., & Fortney, J. J. 2022, *Understanding planetary context to enable life detection on exoplanets and test the Copernican principle*, Nature Astronomy, 6, 189
- Krissansen-Totton, J., Ulses, A. G., Frissell, M., et al. 2025, *Wavelength Requirements for Life Detection via Reflected Light Spectroscopy of Rocky Exoplanets*, arXiv preprint arXiv:2507.14771
- Kuipers, G., Beunk, F. F., & van der Wateren, F. M. 2013, *Periglacial evidence for a 1.91–1.89 Ga old glacial period at low latitude, Central Sweden*, Geology Today, 29, 218

- Kump, L. R. 2008, *The rise of atmospheric oxygen*, *Nature*, 451, 277
- Labrosse, S. 2003, *Thermal and magnetic evolution of the Earth's core*, *Physics of the Earth and Planetary Interiors*, 140, 127
- Labrosse, S., Hernlund, J., & Coltice, N. 2007, *A crystallizing dense magma ocean at the base of the Earth's mantle*, *Nature*, 450, 866
- Lacis, A. A., Schmidt, G. A., Rind, D., & Ruedy, R. A. 2010, *Atmospheric CO₂: Principal control knob governing Earth's temperature*, *Science*, 330, 356
- Lammer, H., Zerkle, A. L., Gebauer, S., et al. 2018, *Origin and evolution of the atmospheres of early Venus, Earth and Mars*, *The Astronomy and Astrophysics Review*, 26, 2
- Laneuville, M., Kameya, M., & Cleaves, H. J. 2018, *Earth without life: A systems model of a global abiotic nitrogen cycle*, *Astrobiology*, 18, 897
- Latouf, N., Himes, M. D., Mandell, A. M., et al. 2025, *BARBIE. Bayesian Analysis for Remote Biosignature Identification on exoEarths. III. Introducing the KEN*, *The Astronomical Journal*, 169, 50
- Latouf, N., Mandell, A. M., Villanueva, G. L., et al. 2023a, *Bayesian analysis for remote biosignature identification on exoEarths (BARBIE). I. Using grid-based nested sampling in coronagraphy observation simulations for H₂O*, *The Astronomical Journal*, 166, 129
- . 2023b, *Bayesian analysis for remote biosignature identification on exoEarths (BARBIE). II. Using grid-based nested sampling in coronagraphy observation simulations for O₂ and O₃*, *The Astronomical Journal*, 167, 27
- Latour, B. 2017, *Facing Gaia: Eight lectures on the new climatic regime* (John Wiley & Sons)
- Lavie, B., Mendonça, J. M., Mordasini, C., et al. 2017, *HELIOS-RETRIEVAL: an open-source, nested sampling atmospheric retrieval code; application to the HR 8799 exoplanets and inferred constraints for planet formation*, *The Astronomical Journal*, 154, 91
- Le Moigne, F. A. 2019, *Pathways of organic carbon downward transport by the oceanic biological carbon pump*, *Frontiers in Marine Science*, 6, 634
- Li, J., Jian, H., Chen, Y. J., et al. 2015, *Seismic observation of an extremely magmatic accretion at the ultraslow spreading Southwest Indian Ridge*, *Geophysical Research Letters*, 42, 2656
- Lichtenberg, T., Bower, D. J., Hammond, M., et al. 2021, *Vertically resolved magma ocean–protoatmosphere evolution: H₂, H₂O, CO₂, CH₄, CO, O₂, and N₂ as primary absorbers*, *Journal of Geophysical Research: Planets*, 126, e2020JE006711
- Lim, O., Benneke, B., Doyon, R., et al. 2023, *Atmospheric Reconnaissance of TRAPPIST-1 b with JWST/NIRISS: Evidence for Strong Stellar Contamination in the Transmission Spectra*, *The Astrophysical Journal Letters*, 955, L22
- Lin, Z., MacDonald, R. J., Kaltenegger, L., & Wilson, D. J. 2021, *Differentiating modern and prebiotic Earth scenarios for TRAPPIST-1e: high-resolution transmission spectra and predictions for JWST*, *MNRAS*, 505, 3562, doi: 10.1093/mnras/stab1486
- Lin, Z., MacDonald, R. J., Kaltenegger, L., & Wilson, D. J. 2021, *Differentiating modern and prebiotic Earth scenarios for TRAPPIST-1e: high-resolution transmission spectra and predictions for JWST*, *Monthly Notices of the Royal Astronomical Society*, 505, 3562
- Lincowski, A. P., Lustig-Yaeger, J., & Meadows, V. S. 2019, *Observing isotopologue bands in terrestrial exoplanet atmospheres with the James Webb Space Telescope: implications for identifying past atmospheric and ocean loss*, *The Astronomical Journal*, 158, 26
- Lincowski, A. P., Meadows, V. S., Crisp, D., et al. 2018, *Evolved climates and observational discriminants for the TRAPPIST-1 planetary system*, *The Astrophysical Journal*, 867, 76

BIBLIOGRAPHY

- Line, M. R., & Parmentier, V.** 2016, *The Influence of Nonuniform Cloud Cover on Transit Transmission Spectra*, ApJ, 820, 78, doi: [10.3847/0004-637X/820/1/78](https://doi.org/10.3847/0004-637X/820/1/78)
- Line, M. R., Brogi, M., Bean, J. L., et al.** 2021, *A solar C/O and sub-solar metallicity in a hot Jupiter atmosphere*, Nature, 598, 580
- Loeb, N. G., Wielicki, B. A., Doelling, D. R., et al.** 2009, *Toward optimal closure of the Earth's top-of-atmosphere radiation budget*, Journal of Climate, 22, 748
- López-Morales, M., Currie, T., Teske, J., et al.** 2019, *Detecting Earth-like biosignatures on rocky exoplanets around nearby stars with ground-based extremely large telescopes*, arXiv preprint arXiv:1903.09523
- Lovelock, J.** 2000, *The ages of Gaia: A biography of our living earth* (Oxford University Press, USA)
- . 2016, *Gaia: A new look at life on earth* (Oxford University Press)
- Lovelock, J. E.** 1983, in *Biomining and biological metal accumulation: biological and geological perspectives papers presented at the fourth international symposium on biomineralization*, renaissance, The Netherlands, june 2–5, 1982, Springer, 15–25
- Lovelock, J. E., & Margulis, L.** 1974, *Atmospheric homeostasis by and for the biosphere: the Gaia hypothesis*, Tellus, 26, 2
- Luger, R., & Barnes, R.** 2015, *Extreme water loss and abiotic O₂ buildup on planets throughout the habitable zones of M dwarfs*, Astrobiology, 15, 119
- Lustig-Yaeger, J.** 2020, *The Detection, Characterization, and Retrieval of Terrestrial Exoplanet Atmospheres* (University of Washington)
- Lustig-Yaeger, J., Meadows, V. S., Crisp, D., Line, M. R., & Robinson, T. D.** 2023a, *Earth as a transiting exoplanet: A validation of transmission spectroscopy and atmospheric retrieval methodologies for terrestrial exoplanets*, The Planetary Science Journal, 4, 170
- . 2023b, *Earth as a Transiting Exoplanet: A Validation of Transmission Spectroscopy and Atmospheric Retrieval Methodologies for Terrestrial Exoplanets*, The Planetary Science Journal, 4, 170
- Lustig-Yaeger, J., Meadows, V. S., & Lincowski, A. P.** 2019a, *The detectability and characterization of the TRAPPIST-1 exoplanet atmospheres with JWST*, The Astronomical Journal, 158, 27
- . 2019b, *A mirage of the cosmic shoreline: Venus-like clouds as a statistical false positive for exoplanet atmospheric erosion*, The Astrophysical Journal Letters, 887, L11
- Lustig-Yaeger, J., Meadows, V. S., Mendoza, G. T., et al.** 2018, *Detecting ocean glint on exoplanets using multiphase mapping*, The Astronomical Journal, 156, 301
- Lustig-Yaeger, J., Robinson, T. D., & Arney, G.** 2019c, “coronagraph”: Telescope Noise Modeling for Exoplanets in Python, Journal of Open Source Software, 4, 1387
- Lustig-Yaeger, J., Sotzen, K. S., Stevenson, K. B., et al.** 2022, *Hierarchical Bayesian Atmospheric Retrieval Modeling for Population Studies of Exoplanet Atmospheres: A Case Study on the Habitable Zone*, AJ, 163, 140, doi: [10.3847/1538-3881/ac5034](https://doi.org/10.3847/1538-3881/ac5034)
- Lustig-Yaeger, J., Fu, G., May, E., et al.** 2023c, *A JWST transmission spectrum of the nearby Earth-sized exoplanet LHS 475 b*, Nature Astronomy, 7, 1317
- LUVOIR Team.** 2019, *The LUVOIR mission concept study final report*, arXiv preprint arXiv:1912.06219
- Lyons, T. W., Reinhard, C. T., & Planavsky, N. J.** 2014, *The rise of oxygen in Earth's early ocean and atmosphere*, Nature, 506, 307
- MacDonald, R. J., Goyal, J. M., & Lewis, N. K.** 2020, *Why is it so cold in here? explaining the cold temperatures retrieved from transmission spectra of exoplanet atmospheres*, The Astrophysical Journal Letters, 893, L43

- MacDonald, R. J., & Lewis, N. K. 2021, *TRIDENT: A Rapid 3D Radiative Transfer Model for Exoplanet Transmission Spectra*, arXiv preprint arXiv:2111.05862
- Macintosh, B., Graham, J., Barman, T., et al. 2015, *Discovery and spectroscopy of the young jovian planet 51 Eri b with the Gemini Planet Imager*, *Science*, 350, 64
- Madhusudhan, N., & Seager, S. 2009, *A temperature and abundance retrieval method for exoplanet atmospheres*, *The Astrophysical Journal*, 707, 24
- Mahieux, A., Vandaele, A. C., Neefs, E., et al. 2010, *Densities and temperatures in the Venus mesosphere and lower thermosphere retrieved from SOIR on board Venus Express: Retrieval technique*, *Journal of Geophysical Research (Planets)*, 115, E12014, doi: 10.1029/2010JE003589
- Mamajek, E., & Stapelfeldt, K. 2023, *NASA ExEP Mission Star List for the Habitable Worlds Observatory*, Tech. rep., NASA Jet Propulsion Laboratory and California Institute of Technology
- Marois, C., Macintosh, B., Barman, T., et al. 2008, *Direct imaging of multiple planets orbiting the star HR 8799*, *science*, 322, 1348
- Marois, C., Zuckerman, B., Konopacky, Q. M., Macintosh, B., & Barman, T. 2010, *Images of a fourth planet orbiting HR 8799*, *Nature*, 468, 1080
- Marty, B., Zimmermann, L., Pujol, M., Burgess, R., & Philippot, P. 2013, *Nitrogen isotopic composition and density of the Archean atmosphere*, *Science*, 342, 101
- Maruyama, S., Santosh, M., & Azuma, S. 2018, *Initiation of plate tectonics in the Hadean: Eclogitization triggered by the ABEL Bombardment*, *Geoscience Frontiers*, 9, 1033
- Mauritsen, T., Stevens, B., Roeckner, E., et al. 2012, *Tuning the climate of a global model*, *Journal of advances in modeling Earth systems*, 4
- May, E. M., Taylor, J., Komacek, T. D., Line, M. R., & Parmentier, V. 2021, *Water Ice Cloud Variability and Multi-epoch Transmission Spectra of TRAPPIST-1e*, *ApJ*, 911, L30, doi: 10.3847/2041-8213/abeff
- May and MacDonald, Bennett, K. A., Moran, S. E., et al. 2023, *Double Trouble: Two Transits of the Super-Earth GJ 1132 b Observed with JWST NIRSpec G395H*, arXiv preprint arXiv:2310.10711
- McCullom, T. M. 2003, *Formation of meteorite hydrocarbons from thermal decomposition of siderite (FeCO₃)*, *Geochimica et Cosmochimica Acta*, 67, 311
- . 2016, *Abiotic methane formation during experimental serpentinization of olivine*, *Proceedings of the National Academy of Sciences*, 113, 13965
- McDonough, W. F. 2025, *Earth's composition: origin, evolution and energy budget*, arXiv preprint arXiv:2505.02641
- Meadows, V., Arney, G., Schwieterman, E., et al. 2017, in *American Astronomical Society Meeting Abstracts# 229*, Vol. 229, 120–03
- Meadows, V. S. 2017, *Reflections on O₂ as a biosignature in exoplanetary atmospheres*, *Astrobiology*, 17, 1022
- Meadows, V. S., & Crisp, D. 1996, *Ground-based near-infrared observations of the Venus nightside: The thermal structure and water abundance near the surface*, *Journal of Geophysical Research: Planets*, 101, 4595
- Meadows, V. S., Lincowski, A. P., & Lustig-Yaeger, J. 2023, *The Feasibility of Detecting Biosignatures in the TRAPPIST-1 Planetary System with JWST*, *The Planetary Science Journal*, 4, 192
- Meadows, V. S., Arney, G. N., Schwieterman, E. W., et al. 2018a, *The habitability of Proxima Centauri b: environmental states and observational discriminants*, *Astrobiology*, 18, 133
- Meadows, V. S., Reinhard, C. T., Arney, G. N., et al. 2018b, *Exoplanet biosignatures: understanding oxygen as a biosignature in the context of its environment*, *Astrobiology*, 18, 630

BIBLIOGRAPHY

- Meixner, M., Cooray, A., Leisawitz, D., et al. 2019, *Origins space telescope mission concept study report*, arXiv preprint arXiv:1912.06213
- Mennesson, B., Gaudi, S., Seager, S., et al. 2016, in *Space telescopes and instrumentation 2016: Optical, infrared, and millimeter wave*, Vol. 9904, SPIE, 212–221
- Métivier, L., Caron, L., Greff-Lefftz, M., et al. 2016, *Evidence for postglacial signatures in gravity gradients: A clue in lower mantle viscosity*, *Earth and Planetary Science Letters*, 452, 146
- Miguel, Y., & Kaltenegger, L. 2013, *Exploring atmospheres of hot mini-Neptunes and extrasolar giant planets orbiting different stars with application to HD 97658b, WASP-12b, CoRoT-2b, XO-1b, and HD 189733b*, *The Astrophysical Journal*, 780, 166
- Mikal-Evans, T. 2022, *Detecting the proposed CH₄–CO₂ biosignature pair with the James Webb Space Telescope: TRAPPIST-1e and the effect of cloud/haze*, *Monthly Notices of the Royal Astronomical Society*, 510, 980
- Millero, F. J. 2005, *Chemical oceanography*, Vol. 30 (CRC press)
- Mills, B., Watson, A. J., Goldblatt, C., Boyle, R., & Lenton, T. M. 2011, *Timing of Neoproterozoic glaciations linked to transport-limited global weathering*, *Nature geoscience*, 4, 861
- Misra, A., Meadows, V., Claire, M., & Crisp, D. 2014a, *Using dimers to measure biosignatures and atmospheric pressure for terrestrial exoplanets*, *Astrobiology*, 14, 67
- Misra, A., Meadows, V., & Crisp, D. 2014b, *The effects of refraction on transit transmission spectroscopy: application to Earth-like exoplanets*, *The Astrophysical Journal*, 792, 61
- Mockler, S. 1995, *Water vapor in the climate system*, Special Report, American, Geophysical Union
- Mollière, P., Wardenier, J. P., van Boekel, R., et al. 2019, *petitRADTRANS. A Python radiative transfer package for exoplanet characterization and retrieval*, *A&A*, 627, A67, doi: [10.1051/0004-6361/201935470](https://doi.org/10.1051/0004-6361/201935470)
- Moran, S. E., Stevenson, K. B., Sing, D. K., et al. 2023, *High Tide or Riptide on the Cosmic Shoreline? A Water-rich Atmosphere or Stellar Contamination for the Warm Super-Earth GJ 486b from JWST Observations*, *The Astrophysical Journal Letters*, 948, L11
- Muñoz, A. G., Osorio, M. Z., Barrena, R., et al. 2012, *Glancing views of the Earth: from a lunar eclipse to an exoplanetary transit*, *The Astrophysical Journal*, 755, 103
- Murante, G., Provenzale, A., Vladilo, G., et al. 2020, *Climate bistability of Earth-like exoplanets*, *Monthly Notices of the Royal Astronomical Society*, 492, 2638
- Nakagawa, T., & Tackley, P. J. 2012, *Influence of magmatism on mantle cooling, surface heat flow and Urey ratio*, *Earth and Planetary Science Letters*, 329, 1
- National Academies of Sciences, Engineering, and Medicine. 2021, *Pathways to Discovery in Astronomy and Astrophysics for the 2020s* (The National Academies Press)
- Nayak, M., Lupu, R., Marley, M. S., et al. 2017, *Atmospheric retrieval for direct imaging spectroscopy of gas giants in reflected light. II. Orbital phase and planetary radius*, *Publications of the Astronomical Society of the Pacific*, 129, 034401
- Nelder, J. A., & Mead, R. 1965, *A Simplex Method for Function Minimization*, *The Computer Journal*, 7, 308, doi: [10.1093/comjnl/7.4.308](https://doi.org/10.1093/comjnl/7.4.308)
- Newman, M. J., & Rood, R. T. 1977, *Implications of solar evolution for the Earth's early atmosphere*, *Science*, 198, 1035
- Nicholson, A. E., Wilkinson, D. M., Williams, H. T., & Lenton, T. M. 2018, *Gaia bottlenecks and planetary habitability maintained by evolving model biospheres: The ExoGaia model*, *Monthly Notices of the Royal Astronomical Society*, 477, 727

- Nikolaou, A., Katyal, N., Tosi, N., et al. 2019, *What factors affect the duration and outgassing of the terrestrial magma ocean?*, The Astrophysical Journal, 875, 11
- Nimmo, F. 2002, *Why does Venus lack a magnetic field?*, Geology, 30, 987
- Niraula, P., de Wit, J., Gordon, I. E., et al. 2022, *The impending opacity challenge in exoplanet atmospheric characterization*, Nature Astronomy, doi: 10.1038/s41550-022-01773-1
- Nixon, M. C., & Madhusudhan, N. 2022, *AURA-3D: A Three-dimensional Atmospheric Retrieval Framework for Exoplanet Transmission Spectra*, arXiv preprint arXiv:2201.03532
- North, G. R., & Coakley Jr, J. A. 1979, *Differences between seasonal and mean annual energy balance model calculations of climate and climate sensitivity*, Journal of Atmospheric Sciences, 36, 1189
- Nowicki, M., DeVries, T., & Siegel, D. A. 2022, *Quantifying the carbon export and sequestration pathways of the ocean's biological carbon pump*, Global Biogeochemical Cycles, 36, e2021GB007083
- Ojakangas, R. W., Srinivasan, R., Hegde, V., Chandrakant, S., & Srikantia, S. 2014, *The Talya conglomerate: An Archean (2.7 Ga) glaciomarine formation, Western Dharwar Craton, Southern India*, Current Science, 387
- O'Leary, M. H. 1981, *Carbon isotope fractionation in plants*, Phytochemistry, 20, 553
- Olson, S. L., Reinhard, C. T., & Lyons, T. W. 2016, *Limited role for methane in the mid-Proterozoic greenhouse*, Proceedings of the National Academy of Sciences, 113, 11447
- Pallé, E., Osorio, M. R. Z., Barrena, R., Montañés-Rodríguez, P., & Martín, E. L. 2009, *Earth's transmission spectrum from lunar eclipse observations*, Nature, 459, 814
- Passow, U., & Carlson, C. A. 2012, *The biological pump in a high CO₂ world*, Marine Ecology Progress Series, 470, 249, doi: 10.3354/meps09985
- Patel, V. K., & Kuttippurath, J. 2023, *Increase in tropospheric water vapor amplifies global warming and climate change*, Ocean-Land-Atmosphere Research, 2, 0015
- Paulson, A., Zhong, S., & Wahr, J. 2005, *Modelling post-glacial rebound with lateral viscosity variations*, Geophysical Journal International, 163, 357
- Pavlov, A. A., Brown, L. L., & Kasting, J. F. 2001, *UV shielding of NH₃ and O₂ by organic hazes in the Archean atmosphere*, Journal of Geophysical Research: Planets, 106, 23267
- Pavlov, A. A., Hurtgen, M. T., Kasting, J. F., & Arthur, M. A. 2003, *Methane-rich Proterozoic atmosphere?*, Geology, 31, 87
- Pavlov, A. A., Kasting, J. F., Brown, L. L., Rages, K. A., & Freedman, R. 2000, *Greenhouse warming by CH₄ in the atmosphere of early Earth*, Journal of Geophysical Research: Planets, 105, 11981
- Pfisterer, R. N., Harvey, J. E., & Breckinridge, J. B. 2018, in *Space Telescopes and Instrumentation 2018: Optical, Infrared, and Millimeter Wave*, Vol. 10698, International Society for Optics and Photonics, 106985F
- Pilson, M. E. 2012, *An Introduction to the Chemistry of the Sea* (cambridge university press)
- Planavsky, N. J., Reinhard, C. T., Wang, X., et al. 2014, *Low Mid-Proterozoic atmospheric oxygen levels and the delayed rise of animals*, science, 346, 635
- Pluriel, W., Whiteford, N., Edwards, B., et al. 2020, *ARES. III. Unveiling the two faces of KELT-7 b with HST WFC3*, The Astronomical Journal, 160, 112
- Quanz, S. P., Ottiger, M., Fontanet, E., et al. 2021, *Large Interferometer For Exoplanets (LIFE): I. Improved exoplanet detection yield estimates for a large mid-infrared space-interferometer mission*, arXiv preprint arXiv:2101.07500
- Quay, P., Sonnerup, R., Westby, T., Stutsman, J., & McNichol, A. 2003, *Changes in the 13C/12C of dissolved inorganic carbon in the ocean as a tracer of anthropogenic CO₂ uptake*, Global Biogeochemical Cycles, 17, 4

BIBLIOGRAPHY

- Rackham, B. V., Apai, D., & Giampapa, M. S. 2018, *The transit light source effect: false spectral features and incorrect densities for M-dwarf transiting planets*, The Astrophysical Journal, 853, 122
- Radica, M., Piaulet-Ghorayeb, C., Taylor, J., et al. 2025, *Promise and peril: stellar contamination and strict limits on the atmosphere composition of TRAPPIST-1 c from JWST NIRISS transmission spectra*, The Astrophysical Journal Letters, 979, L5
- Ragsdale, S. W. 2004, *Life with carbon monoxide*, Critical reviews in biochemistry and molecular biology, 39, 165
- Ramanathan, V., Cess, R., Harrison, E., et al. 1989, *Cloud-radiative forcing and climate: Results from the Earth Radiation Budget Experiment*, Science, 243, 57
- Randel, W., & Park, M. 2019, *Diagnosing observed stratospheric water vapor relationships to the cold point tropical tropopause*, Journal of Geophysical Research: Atmospheres, 124, 7018
- Ranjan, S., Schwieterman, E. W., Leung, M., Harman, C. E., & Hu, R. 2023, *The Importance of the Upper Atmosphere to CO/O₂ Runaway on Habitable Planets Orbiting Low-mass Stars*, The Astrophysical Journal Letters, 958, L15
- Ratner, M. I., & Walker, J. C. 1972, *Atmospheric ozone and the history of life*, Journal of Atmospheric Sciences, 29, 803
- Ray, R. D., Eanes, R., & Chao, B. F. 1996, *Detection of tidal dissipation in the solid Earth by satellite tracking and altimetry*, Nature, 381, 595
- Roberson, A. L., Roadt, J., Halevy, I., & Kasting, J. 2011, *Greenhouse warming by nitrous oxide and methane in the Proterozoic Eon*, Geobiology, 9, 313
- Robinson, T. D. 2017a, *Characterizing exoplanet habitability*, arXiv preprint arXiv:1701.05205
- . 2017b, *A theory of exoplanet transits with light scattering*, The Astrophysical Journal, 836, 236
- . 2018, in Handbook of Exoplanets (Springer International Publishing), 3137–3157
- Robinson, T. D., & Catling, D. C. 2014, *Common 0.1 bar tropopause in thick atmospheres set by pressure-dependent infrared transparency*, Nature Geoscience, 7, 12
- Robinson, T. D., Meadows, V. S., & Crisp, D. 2010, *Detecting oceans on extrasolar planets using the glint effect*, The Astrophysical Journal Letters, 721, L67
- Robinson, T. D., & Salvador, A. 2023, *Exploring and validating exoplanet atmospheric retrievals with solar system analog observations*, The Planetary Science Journal, 4, 10
- Robinson, T. D., Stapelfeldt, K. R., & Marley, M. S. 2016, *Characterizing rocky and gaseous exoplanets with 2 m class space-based coronagraphs*, Publications of the Astronomical Society of the Pacific, 128, 025003
- Robinson, T. D., Meadows, V. S., Crisp, D., et al. 2011, *Earth as an extrasolar planet: Earth model validation using EPOXI Earth observations*, Astrobiology, 11, 393
- Rodler, F., & López-Morales, M. 2014, *Feasibility Studies for the Detection of O₂ in an Earth-like Exoplanet*, The Astrophysical Journal, 781, 54
- Rose, B. E., Cronin, T. W., & Bitz, C. M. 2017, *Ice Caps and Ice Belts: The Effects of Obliquity on Ice- Albedo Feedback*, The Astrophysical Journal, 846, 28
- Rothman, L. S., Gordon, I., Barber, R., et al. 2010, *HITEMP, the high-temperature molecular spectroscopic database*, Journal of Quantitative Spectroscopy and Radiative Transfer, 111, 2139
- Rothman, L. S., Gordon, I. E., Babikov, Y., et al. 2013, *The HITRAN2012 molecular spectroscopic database*, Journal of Quantitative Spectroscopy and Radiative Transfer, 130, 4
- Rotman, Y., Komacek, T. D., Villanueva, G. L., Fauchez, T. J., & May, E. M. 2022, *General Circulation Model*

- Constraints on the Detectability of the CO₂-CH₄ Biosignature Pair on TRAPPIST-1e with JWST*, The Astrophysical Journal Letters, 942, L4
- Ruffio, J.-B., Steiger, S., Spohn, C., et al. 2026, *Characterizing Earth analogs may require a moderate or high-resolution spectrograph*, arXiv preprint arXiv:2604.17554
- Rugheimer, S., & Kaltenegger, L. 2018, *Spectra of Earth-like planets through geological evolution around FGKM stars*, The Astrophysical Journal, 854, 19
- Rugheimer, S., Kaltenegger, L., Segura, A., Linsky, J., & Mohanty, S. 2015, *Effect of UV radiation on the spectral fingerprints of Earth-like planets orbiting M stars*, The Astrophysical Journal, 809, 57
- Rugheimer, S., Kaltenegger, L., Zsom, A., Segura, A., & Sasselov, D. 2013, *Spectral fingerprints of Earth-like planets around FGK stars*, Astrobiology, 13, 251
- Sagan, C., & Mullen, G. 1972, *Earth and Mars: Evolution of atmospheres and surface temperatures*, Science, 177, 52
- Santerne, A., Brugger, B., Armstrong, D., et al. 2018, *An Earth-sized exoplanet with a Mercury-like composition*, Nature Astronomy, 2, 393
- Saunio, M., Stavert, A. R., Poulter, B., et al. 2019, *The global methane budget 2000–2017*, Earth System Science Data Discussions, 2019, 1
- Schaefer, L., & Elkins-Tanton, L. T. 2018, *Magma oceans as a critical stage in the tectonic development of rocky planets*, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 376, 20180109
- Schaefer, L., Wordsworth, R. D., Berta-Thompson, Z., & Sasselov, D. 2016, *Predictions of the atmospheric composition of GJ 1132b*, The Astrophysical Journal, 829, 63
- Schwartzman, E., Reinhard, C., Olson, S., & Lyons, T. 2018a, *The importance of UV capabilities for identifying inhabited exoplanets with next generation space telescopes*, arXiv preprint arXiv:1801.02744
- Schwartzman, E. W., & Leung, M. 2024, *An overview of exoplanet biosignatures*, Reviews in Mineralogy and Geochemistry, 90, 465
- Schwartzman, E. W., Reinhard, C. T., Olson, S. L., Harman, C. E., & Lyons, T. W. 2019a, *A limited habitable zone for complex life*, The Astrophysical Journal, 878, 19
- Schwartzman, E. W., Reinhard, C. T., Olson, S. L., et al. 2019b, *Rethinking CO antibiosignatures in the search for life beyond the solar system*, The Astrophysical Journal, 874, 9
- Schwartzman, E. W., Robinson, T. D., Meadows, V. S., Misra, A., & Domagal-Goldman, S. 2015, *Detecting and constraining N₂ abundances in planetary atmospheres using collisional pairs*, The Astrophysical Journal, 810, 57
- Schwartzman, E. W., Meadows, V. S., Domagal-Goldman, S. D., et al. 2016, *Identifying planetary biosignature impostors: spectral features of CO and O₄ resulting from abiotic O₂/O₃ production*, The Astrophysical Journal Letters, 819, L13
- Schwartzman, E. W., Kiang, N. Y., Parenteau, M. N., et al. 2018b, *Exoplanet biosignatures: a review of remotely detectable signs of life*, Astrobiology, 18, 663
- Seager, S. 2011, *Exoplanets* (University of Arizona Press)
- Seager, S., Turner, E. L., Schafer, J., & Ford, E. B. 2005, *Vegetation's red edge: a possible spectroscopic biosignature of extraterrestrial plants*, Astrobiology, 5, 372
- Seales, J., & Lenardic, A. 2020, *Deep Water Cycling and the Multi-Stage Cooling of the Earth*, Geochemistry, Geophysics, Geosystems, 21, e2020GC009106
- Segura, A., Kasting, J. F., Meadows, V., et al. 2005, *Biosignatures from Earth-like planets around M dwarfs*, Astrobiology, 5, 706

BIBLIOGRAPHY

- Segura, A., Krellove, K., Kasting, J. F., et al. 2003, *Ozone concentrations and ultraviolet fluxes on Earth-like planets around other stars*, *Astrobiology*, 3, 689
- Segura, A., Meadows, V., Kasting, J., Crisp, D., & Cohen, M. 2007, *Abiotic formation of O₂ and O₃ in high-CO₂ terrestrial atmospheres*, *Astronomy & Astrophysics*, 472, 665
- Segura, A., Walkowicz, L. M., Meadows, V., Kasting, J., & Hawley, S. 2010, *The effect of a strong stellar flare on the atmospheric chemistry of an Earth-like planet orbiting an M dwarf*, *Astrobiology*, 10, 751
- Selsis, F., Despois, D., & Parisot, J.-P. 2002, *Signature of life on exoplanets: Can Darwin produce false positive detections?*, *Astronomy & Astrophysics*, 388, 985
- Shaik, I., Nagamani, P., Yadav, S., Manmode, Y., & Rao, G. S. 2025, *Advanced deep learning technique for estimating global surface ocean calcium carbonate saturation (Ω_{cal})*, *Marine Chemistry*, 268, 104483
- Sheldon, N. D. 2006, *Precambrian paleosols and atmospheric CO₂ levels*, *Precambrian Research*, 147, 148
- Shields, A. L., Meadows, V. S., Bitz, C. M., et al. 2013, *The effect of host star spectral energy distribution and ice-albedo feedback on the climate of extrasolar planets*, *Astrobiology*, 13, 715
- Shim, S.-H. 2014, *Un-Earth-Like Interiors of Earth-Like Exoplanets*, *Search for Life Beyond the Solar System. Exoplanets, Biosignatures & Instruments*, 2
- Shue, J.-H., Song, P., Russell, C., et al. 1998, *Magnetopause location under extreme solar wind conditions*, *Journal of Geophysical Research: Space Physics*, 103, 17691
- Siegel, D. A., DeVries, T., Cetinić, I., & Bisson, K. M. 2023, *Quantifying the ocean's biological pump and its carbon cycle impacts on global scales*, *Annual review of marine science*, 15, 329
- Skilling, J. 2004, in *AIP Conference Proceedings*, Vol. 735, American Institute of Physics, 395–405
- Skilling, J. 2006, *Nested sampling for general Bayesian computation*, *Bayesian analysis*, 1, 833
- Sleep, N. H., & Zahnle, K. 2001, *Carbon dioxide cycling and implications for climate on ancient Earth*, *Journal of Geophysical Research: Planets*, 106, 1373
- Snellen, I., De Kok, R., Le Poole, R., Brogi, M., & Birkby, J. 2013, *Finding extraterrestrial life using ground-based high-dispersion spectroscopy*, *The Astrophysical Journal*, 764, 182
- Solomon, S., Qin, D., Manning, M., et al. 2007, *IPCC fourth assessment report (AR4)*, *Climate change*, 374
- Spurr, R. J. D., Kurosu, T. P., & Chance, K. V. 2001, *A linearized discrete ordinate radiative transfer model for atmospheric remote-sensing retrieval*, *J. Quant. Spec. Radiat. Transf.*, 68, 689, doi: [10.1016/S0022-4073\(00\)00055-8](https://doi.org/10.1016/S0022-4073(00)00055-8)
- Stalport, F., Coll, P., Cabane, M., et al. 2005, *Search for past life on Mars: Physical and chemical characterization of minerals of biotic and abiotic origin: Part 1-calcite*, *Geophysical Research Letters*, 32
- Stamnes, K., Tsay, S.-C., Wiscombe, W., & Jayaweera, K. 1988, *Numerically stable algorithm for discrete-ordinate-method radiative transfer in multiple scattering and emitting layered media*, *Applied optics*, 27, 2502
- Stark, C. C., Roberge, A., Mandell, A., & Robinson, T. D. 2014, *Maximizing the ExoEarth Candidate Yield from a Future Direct Imaging Mission*, *ApJ*, 795, 122, doi: [10.1088/0004-637X/795/2/122](https://doi.org/10.1088/0004-637X/795/2/122)
- Stark, C. C., Belikov, R., Bolcar, M. R., et al. 2019, *ExoEarth yield landscape for future direct imaging space telescopes*, *Journal of Astronomical Telescopes, Instruments, and Systems*, 5, 024009
- Stark, C. C., Steiger, S., Tokadjian, A., et al. 2025, *Cross-Model Validation of Coronagraphic Exposure Time Calculators for the Habitable Worlds Observatory: A Report from the Exoplanet Science Yield sub-Working Group*, *arXiv preprint arXiv:2502.18556*
- Steiger, S., Chen, P., & Pueyo, L. 2026, *Incorporating wavefront error, wavefront sensing and control, and sensitivities*

- into exposure time calculations for future space missions with the error budget software, *Journal of Astronomical Telescopes, Instruments, and Systems*, 12, 041005
- Stueeken, E. E., Som, S., Claire, M., et al. 2020, *Mission to planet Earth: the first two billion years*, *Space Science Reviews*, 216, 31
- Suissa, G., Mandell, A. M., Wolf, E. T., et al. 2020, *Dim prospects for transmission spectra of ocean Earths around M stars*, *The Astrophysical Journal*, 891, 58
- Sumner, D. Y., & Grotzinger, J. P. 1996, *Were kinetics of Archean calcium carbonate precipitation related to oxygen concentration?*, *Geology*, 24, 119, doi: [10.1130/0091-7613\(1996\)024<0119:WKOACC>2.3.CO;2](https://doi.org/10.1130/0091-7613(1996)024<0119:WKOACC>2.3.CO;2)
- Tajika, E. 2007, *Long-term stability of climate and global glaciations throughout the evolution of the Earth*, *Earth, planets and space*, 59, 293
- Taylor, J., Parmentier, V., Irwin, P. G. J., et al. 2020, *Understanding and mitigating biases when studying inhomogeneous emission spectra with JWST*, *MNRAS*, 493, 4342, doi: [10.1093/mnras/staa552](https://doi.org/10.1093/mnras/staa552)
- Thompson, M. A., Krissansen-Totton, J., Wogan, N., Telus, M., & Fortney, J. J. 2022, *The case and context for atmospheric methane as an exoplanet biosignature*, *Proceedings of the National Academy of Sciences*, 119, e2117933119
- Tian, F., Chassefière, E., Leblanc, F., & Brain, D. 2013, *Atmosphere escape and climate evolution of terrestrial planets*, *Comparative climatology of terrestrial planets*, chapitre
- Tokadjian, A., Hu, R., & Damiano, M. 2024, *The Detectability of CH₄/CO₂/CO and N₂O Biosignatures Through Reflection Spectroscopy of Terrestrial Exoplanets*, *The Astronomical Journal*, 168, 292
- Toon, O. B., McKay, C., Ackerman, T., & Santhanam, K. 1989, *Rapid calculation of radiative heating rates and photodissociation rates in inhomogeneous multiple scattering atmospheres*, *Journal of Geophysical Research: Atmospheres*, 94, 16287
- Tremblay, L., Line, M. R., Stevenson, K., et al. 2020, *The Detectability and Constraints of Biosignature Gases in the Near- and Mid-infrared from Transit Transmission Spectroscopy*, *AJ*, 159, 117, doi: [10.3847/1538-3881/ab64dd](https://doi.org/10.3847/1538-3881/ab64dd)
- Trenberth, K. E. 1998, *Atmospheric moisture residence times and cycling: Implications for rainfall rates and climate change*, *Climatic change*, 39, 667
- Trenberth, K. E., Christy, J. R., & Olson, J. G. 1987, *Global atmospheric mass, surface pressure, and water vapor variations*, *Journal of Geophysical Research: Atmospheres*, 92, 14815
- Tsai, S.-M., Lee, E. K., Powell, D., et al. 2023, *Photochemically produced SO₂ in the atmosphere of WASP-39b*, *Nature*, 617, 483
- Tsamalis, C., Chédin, A., Pelon, J., & Capelle, V. 2013, *The seasonal vertical distribution of the Saharan Air Layer and its modulation by the wind*, *Atmospheric Chemistry and Physics*, 13, 11235
- Tsiaras, A., Rocchetto, M., Waldmann, I., et al. 2016, *Detection of an atmosphere around the super-Earth 55 Cancri e*, *The Astrophysical Journal*, 820, 99
- Tuchow, N. W., Harada, C. K., Mamajek, E. E., et al. 2025, *HWO Target Stars and Systems: A Prioritized Community List of Potential Stellar Targets for the Habitable Worlds Observatory's ExoEarth Survey*, *Publications of the Astronomical Society of the Pacific*, 137, 104402
- Turbet, M., Bolmont, E., Bourrier, V., et al. 2020, *A review of possible planetary atmospheres in the TRAPPIST-1 system*, *Space science reviews*, 216, 1
- Turbet, M., Forget, F., Leconte, J., Charnay, B., & Tobie, G. 2017, *CO₂ condensation is a serious limit to the deglaciation of Earth-like planets*, *Earth and Planetary Science Letters*, 476, 11
- Turcotte, D. L., & Schubert, G. 2002, *Geodynamics* (Cambridge university press)

BIBLIOGRAPHY

- Tyndall, J.** 1861, *XXIII. On the absorption and radiation of heat by gases and vapours, and on the physical connexion of radiation, absorption, and conduction.—The bakerian lecture*, The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, 22, 169
- Ulses, A. G., Krissansen-Totton, J., Robinson, T. D., et al.** 2025, *Detecting Land with Reflected-light Spectroscopy to Rule Out Waterworld O₂ Biosignature False Positives*, The Astrophysical Journal, 990, 48
- United States Committee on Extension to the Standard Atmosphere.** 1976, US standard atmosphere, 1976 (National Oceanic and Atmospheric Administration)
- Vial, J., Dufresne, J.-L., & Bony, S.** 2013, *On the interpretation of inter-model spread in CMIP5 climate sensitivity estimates*, Climate Dynamics, 41, 3339
- Vinatier, S., Bézard, B., Fouchet, T., et al.** 2007, *Vertical abundance profiles of hydrocarbons in Titan's atmosphere at 15 S and 80 N retrieved from Cassini/CIRS spectra*, Icarus, 188, 120
- Vladilo, G., Murante, G., Silva, L., et al.** 2013, *The habitable zone of Earth-like planets with different levels of atmospheric pressure*, The Astrophysical Journal, 767, 65
- Volk, T.** 2002, *Toward a future for Gaia theory*, Climatic Change, 52, 423
- Wakeford, H. R., & Sing, D. K.** 2015, *Transmission spectral properties of clouds for hot Jupiter exoplanets*, Astronomy & Astrophysics, 573, A122
- Wakeford, H. R., Sing, D. K., Deming, D., et al.** 2017, *The complete transmission spectrum of WASP-39b with a precise water constraint*, The Astronomical Journal, 155, 29
- Waldmann, I. P., Tinetti, G., Rocchetto, M., et al.** 2015, *Tau-REx I: A next generation retrieval code for exoplanetary atmospheres*, The Astrophysical Journal, 802, 107
- Walker, J.** 1977, *Evolution of the Atmosphere* (Macmillan Co., New York)
- Walker, J. C., Hays, P., & Kasting, J. F.** 1981, *A negative feedback mechanism for the long-term stabilization of Earth's surface temperature*, Journal of Geophysical Research: Oceans, 86, 9776
- Walter, M., Buick, R., & Dunlop, J.** 1980, *Stromatolites 3,400–3,500 Myr old from the North pole area, Western Australia*, Nature, 284, 443
- Wang, C., Forget, F., Bertrand, T., et al.** 2018, *Parameterization of rocket dust storms on Mars in the LMD Martian GCM: Modeling details and validation*, Journal of Geophysical Research: Planets, 123, 982
- Wang, Y., Tian, F., Li, T., & Hu, Y.** 2016, *On the detection of carbon monoxide as an anti-biosignature in exoplanetary atmospheres*, Icarus, 266, 15
- Washington III, R. C., Gao, P., Wogan, N., Saxena, P., & Wilkins, J.** 2026
- Watanabe, K., Ohtani, E., Kamada, S., et al.** 2014, *The abundance of potassium in the Earth's core*, Physics of the Earth and Planetary Interiors, 237, 65
- Watson, A. J., Donahue, T. M., & Walker, J. C.** 1981, *The dynamics of a rapidly escaping atmosphere: applications to the evolution of Earth and Venus*, Icarus, 48, 150
- Way, M. J., Del Genio, A. D., Aleinov, I., et al.** 2018, *Climates of warm Earth-like planets. I. 3D model simulations*, The Astrophysical Journal Supplement Series, 239, 24
- West, A. J.** 2012, *Thickness of the chemical weathering zone and implications for erosional and climatic drivers of weathering and for carbon-cycle feedbacks*, Geology, 40, 811
- Wielicki, B. A., Barkstrom, B. R., Harrison, E. F., et al.** 1996, *Clouds and the Earth's Radiant Energy System (CERES): An earth observing system experiment*, Bulletin of the American Meteorological Society, 77, 853

- Wild, B., Gerrits, R., & Bonneville, S. 2022, *The contribution of living organisms to rock weathering in the critical zone*, npj Materials degradation, 6, 98
- Wilhelm, C., Barnes, R., Deitrick, R., & Mellman, R. 2022, *The Ice Coverage of Earth-like Planets Orbiting FGK Stars*, The Planetary Science Journal, 3, 13
- Williams, D. M., & Gaidos, E. 2008, *Detecting the glint of starlight on the oceans of distant planets*, Icarus, 195, 927
- Williams, G. E. 2005, *Subglacial meltwater channels and glaciofluvial deposits in the Kimberley Basin, Western Australia: 1.8 Ga low-latitude glaciation coeval with continental assembly*, Journal of the Geological Society, 162, 111
- Windsor, J. D., Robinson, T. D., kumar Kopparapu, R., et al. 2023, *A radiative-convective model for terrestrial planets with self-consistent patchy clouds*, The Planetary Science Journal, 4, 94
- Wogan, N., Krissansen-Totton, J., & Catling, D. C. 2020, *Abundant atmospheric methane from volcanism on terrestrial planets is unlikely and strengthens the case for methane as a biosignature*, The Planetary Science Journal, 1, 58
- Wogan, N., Lopez, J., Batalha, N., & Fortney, J. 2025a, *Resampled Opacity Databases for PICASO from the Photochem model*, v1.0.0, Zenodo, doi: [10.5281/zenodo.17381172](https://doi.org/10.5281/zenodo.17381172)
- Wogan, N. F., & Catling, D. C. 2020, *When is chemical disequilibrium in Earth-like planetary atmospheres a biosignature versus an anti-biosignature? Disequilibria from dead to living worlds*, The Astrophysical Journal, 892, 127
- Wogan, N. F., Catling, D. C., Zahnle, K. J., & Claire, M. W. 2022, *Rapid timescale for an oxic transition during the Great Oxidation Event and the instability of low atmospheric O₂*, Proceedings of the National Academy of Sciences, 119, e2205618119
- Wogan, N. F., Batalha, N. E., Zahnle, K., et al. 2025b, *The Open-source Photochem Code: A General Chemical and Climate Model for Interpreting (Exo) Planet Observations*, The Planetary Science Journal, 6, 256
- Wolf, E. T., Kopparapu, R., Haqq-Misra, J., & Fauchez, T. J. 2022, *ExoCAM: a 3D climate model for exoplanet atmospheres*, The Planetary Science Journal, 3, 7
- Wolf, E. T., Shields, A. L., Kopparapu, R. K., Haqq-Misra, J., & Toon, O. B. 2017, *Constraints on climate and habitability for Earth-like exoplanets determined from a general circulation model*, The Astrophysical Journal, 837, 107
- Wordsworth, R., & Pierrehumbert, R. 2014, *Abiotic oxygen-dominated atmospheres on terrestrial habitable zone planets*, The Astrophysical Journal Letters, 785, L20
- Wunderlich, F., Godolt, M., Grenfell, J. L., et al. 2019, *Detectability of atmospheric features of Earth-like planets in the habitable zone around M dwarfs*, Astronomy & Astrophysics, 624, A49
- Yang, J., Leconte, J., Wolf, E. T., et al. 2016, *Differences in water vapor radiative transfer among 1D models can significantly affect the inner edge of the habitable zone*, The Astrophysical Journal, 826, 222
- Young, A., Arney, G., Robinson, T. D., et al. 2026
- Young, A. V., Crouse, J., Arney, G., et al. 2024, *Retrievals Applied to a Decision Tree Framework Can Characterize Earthlike Exoplanet Analogs*, The Planetary Science Journal, 5, 7
- Young, A. V., Robinson, T. D., Krissansen-Totton, J., et al. 2025, *Modern Earth-like Chemical Disequilibrium Biosignatures are Challenging to Constrain through Spectroscopic Retrievals*, The Astrophysical Journal, 986, 206
- . 2023, *Inferring Chemical Disequilibrium Biosignatures for Proterozoic Earth-Like Exoplanets*, arXiv preprint arXiv:2311.06083
- Young, G. M., Brunn, V. V., Gold, D. J., & Minter, W. 1998, *Earth's oldest reported glaciation: physical and chemical evidence from the Archean Mozaan Group (~2.9 Ga) of South Africa*, The Journal of Geology, 106, 523
- Zahnle, K., Freedman, R. S., & Catling, D. C. 2011, *Is there methane on Mars?*, Icarus, 212, 493

BIBLIOGRAPHY

- Zahnle, K., Haberle, R. M., Catling, D. C., & Kasting, J. F.** 2008, *Photochemical instability of the ancient Martian atmosphere*, Journal of Geophysical Research: Planets, 113
- Zahnle, K. J., Gacesa, M., & Catling, D. C.** 2019, *Strange messenger: A new history of hydrogen on Earth, as told by Xenon*, Geochimica et Cosmochimica Acta, 244, 56
- Zahnle, K. J., Lupu, R., Catling, D. C., & Wogan, N.** 2020, *Creation and evolution of impact-generated reduced atmospheres of early Earth*, The Planetary Science Journal, 1, 11
- Zerkle, A. L., Claire, M. W., Domagal-Goldman, S. D., Farquhar, J., & Poulton, S. W.** 2012, *A bistable organic-rich atmosphere on the Neoarchaeon Earth*, Nature Geoscience, 5, 359
- Zhuang, D., Yan, H., Tucker, M. E., et al.** 2018, *Calcite precipitation induced by Bacillus cereus MRR2 cultured at different Ca²⁺ concentrations: Further insights into biotic and abiotic calcite*, Chemical Geology, 500, 64
- Zieba, S., Kreidberg, L., Ducrot, E., et al.** 2023, *No thick carbon dioxide atmosphere on the rocky exoplanet TRAPPIST-1 c*, Nature, 620, 746
- Zsom, A., Kaltenegger, L., & Goldblatt, C.** 2012, *A 1D microphysical cloud model for Earth, and Earth-like exoplanets: Liquid water and water ice clouds in the convective troposphere*, Icarus, 221, 603