



Hot Rocks Survey. V. Secondary Eclipse Photometry of GJ 3473 b with JWST/MIRI

Måns Holmberg¹, Hannah Diamond-Lowe¹, João M. Mendonça^{2,3,4}, Daniel Kitzmann⁵, Néstor Espinoza^{1,6},
Natalie H. Allen⁶, Prune C. August², Mark Fortune⁷, Amélie Gressier¹, Jegug Ih¹, Erik Meier Valdés⁸,
Merlin Zraggen⁹, Lars A. Buchhave², Brice-Olivier Demory^{5,9,10}, Chloe Fisher⁸, Neale P. Gibson⁷,
Kevin Heng^{11,12,13}, Bibiana Prinoth¹⁴, and Adam J. Burgasser¹⁵

¹ Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA; mholmberg@stsci.edu

² Department of Space Research and Space Technology, Technical University of Denmark, Elektrovej 328, 2800 Kgs. Lyngby, Denmark

³ Department of Physics and Astronomy, University of Southampton, Highfield, Southampton SO17 1BJ, UK

⁴ School of Ocean and Earth Science, University of Southampton, Southampton, SO14 3ZH, UK

⁵ Space Research and Planetary Sciences, Physics Institute, University of Bern, Gesellschaftsstrasse 6, 3012 Bern, Switzerland

⁶ Department of Physics and Astronomy, Johns Hopkins University, 3400 N. Charles Street, Baltimore, MD 21218, USA

⁷ School of Physics, Trinity College Dublin, University of Dublin, Dublin 2, Ireland

⁸ Department of Physics, University of Oxford, Keble Road, Oxford, OX1 3RH, UK

⁹ Center for Space and Habitability, University of Bern, Gesellschaftsstrasse 6, 3012 Bern, Switzerland

¹⁰ ARTORG Center for Biomedical Engineering Research, University of Bern, Murtenstrasse 50, CH-3008, Bern, Switzerland

¹¹ Ludwig Maximilian University, Faculty of Physics, Scheinerstr. 1, Munich D-81679, Germany

¹² Munich Center for Geoastrometry, Ludwig Maximilian University, Theresienstrasse 41, D-80333, Munich, Bavaria, Germany

¹³ University College London, Department of Physics & Astronomy, Gower Street, London, WC1E 6BT, UK

¹⁴ Lund Observatory, Division of Astrophysics, Department of Physics, Lund University, Box 118, 221 00 Lund, Sweden

¹⁵ Department of Astronomy & Astrophysics, UC San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA

Received 2025 September 16; revised 2026 February 23; accepted 2026 February 26; published 2026 March 27

Abstract

JWST is transforming our ability to characterize small exoplanets, from sub-Neptunes to rocky worlds. A key open question is whether highly irradiated rocky planets can retain atmospheres or are stripped bare by stellar irradiation—a boundary that remains to be mapped observationally. Here we present the first JWST secondary eclipse observations of the rocky exoplanet GJ 3473 b, obtained with MIRI F1500W photometry. Using four visits, we confidently detect the eclipse at an average depth of 186 ± 45 ppm, somewhat lower than expected for a blackbody. We test a wide range of data reduction and analysis assumptions and provide new insights into MIRI detector settling behavior that will benefit future observations. We model a suite of airless surfaces with varied compositions, textures, and degrees of space weathering, as well as idealized atmospheric scenarios including the possibility of atmospheric collapse. Both atmospheric and bare-rock interpretations remain consistent with the data, but we exclude thick CO₂ atmospheres, placing a 95% credible upper limit of 1.2–6.5 bar on the surface pressure. We also find tentative evidence for visit-to-visit variability in eclipse depth (33–371 ppm), though additional data are required to confirm this. Our results highlight the challenges and intrinsic degeneracies in interpreting MIRI F1500W eclipse measurements of rocky exoplanets, indicating that such observations alone may not uniquely distinguish between bare-rock and atmospheric scenarios. Future spectroscopic or phase-curve observations will be required to determine whether or not GJ 3473 b hosts a substantial atmosphere.

Unified Astronomy Thesaurus concepts: [Exoplanets \(498\)](#); [Extrasolar rocky planets \(511\)](#); [James Webb Space Telescope \(2291\)](#)

1. Introduction

The question of whether rocky exoplanets can retain atmospheres in the face of intense stellar irradiation remains a central challenge in exoplanetary science. Although planets like Earth, Venus, and Mars offer examples of secondary atmospheres shaped by outgassing and surface–atmosphere interactions, extrapolating these processes to exoplanets—especially those orbiting low-mass stars—comes with many uncertainties. The harsh radiation environments of close-in M-dwarf planets may erode primordial and secondary atmospheres through mechanisms of atmospheric escape, such as photoevaporation, stellar wind stripping, and core-powered mass loss (E. D. Lopez & J. J. Fortney 2013; R. Luger et al. 2015; J. E. Owen & Y. Wu 2017; S. Ginzburg et al. 2018).

Based on the solar system, K. J. Zahnle & D. C. Catling (2017) proposed the idea of a “cosmic shoreline,” a boundary that separates bodies with atmospheres from those without depending on the irradiation and escape velocity. While this concept offers a useful organizing principle, recent work has emphasized that the boundary between atmosphere-bearing and airless worlds is not sharp (X. Ji et al. 2025; E. K. Pass et al. 2025; G. Van Looveren et al. 2025, J. Ih et al. 2026, in preparation). Variations in initial volatile inventories, stellar XUV evolution, escape, and outgassing mechanisms all contribute to large uncertainties in where the boundary lies for any given planet. These uncertainties make it difficult to predict atmospheric retention from theory alone. Ultimately, observations are needed to anchor and refine our understanding of the different mechanisms at play.

Recent advances with the James Webb Space Telescope (JWST; J. P. Gardner et al. 2006), particularly using the Mid-Infrared Instrument (MIRI; G. S. Wright et al. 2023), have opened a new window into the atmospheres of small



Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](#). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

exoplanets, from sub-Neptunes to rocky worlds (e.g., T. P. Greene et al. 2023; P. Gao et al. 2023; S. Zieba et al. 2023; R. Hu et al. 2024; Q. Xue et al. 2024; M. Zhang et al. 2024; P. C. August et al. 2025; N. Madhusudhan et al. 2025; N. Tusay et al. 2025). In the mid-infrared, secondary eclipse photometry and spectroscopy enable direct measurements of a planet’s thermal emission, which can be compared with theoretical expectations for bare-rock and atmosphere-bearing scenarios (e.g., R. Hu et al. 2012; D. D. B. Koll et al. 2019; M. Mansfield et al. 2019). Because an atmosphere can redistribute heat or have a nonisothermal temperature profile, deviations from the expected thermal flux of an airless, typically dark surface can serve as indirect evidence for atmospheric retention. In addition, absorption by molecules such as CO_2 can further decrease the thermal flux at specific wavelengths, e.g., at $15\ \mu\text{m}$, if present in the atmosphere. This technique has already been applied to a growing sample of rocky exoplanets orbiting M dwarfs with JWST (see below).

Secondary eclipse photometric observations can serve as a method for identifying candidate atmospheres of hot, tidally locked, rocky exoplanets (D. D. B. Koll et al. 2019), especially when compared to spectroscopic or phase-curve observations, which often require a more significant time investment. This relies on the fact that hot bare-rock surfaces are expected to have a low albedo (D. D. B. Koll et al. 2019; M. Mansfield et al. 2019), coming from materials such as basaltic rock, in addition to the darkening effects of space weathering (e.g., C. M. Pieters et al. 2000; B. Hapke 2001; C. M. Pieters & S. K. Noble 2016). Consequently, a deep eclipse depth, as measured by photometry, may suggest that an exoplanet is airless, while a shallow eclipse could indicate the presence of an atmosphere. This makes secondary eclipse observations a valuable tool for screening potential targets for further study. This is the goal of the Hot Rocks Survey (JWST GO program 3730, PI: H. Diamond-Lowe, Co-PI: J. M. Mendonça), designed to search for atmospheres on nine rocky exoplanets orbiting M dwarfs using secondary eclipse photometry with MIRI.

Recent MIRI photometry of TRAPPIST-1 b and c has shown eclipse depths consistent with bare-rock surfaces, disfavoring the presence of thick CO_2 atmospheres (T. P. Greene et al. 2023; J. Ih et al. 2023; A. P. Lincowski et al. 2023; S. Zieba et al. 2023). However, E. Ducrot et al. (2025) proposed that the thermal emission of TRAPPIST-1 b could also be explained by a CO_2 -dominated atmosphere with a thermal inversion, potentially caused by photochemical hazes. Observations of LHS 1140 c (M. Fortune et al. 2025), TOI-1468 b (E. A. Meier Valdés et al. 2025), and LTT 3780 b (N. H. Allen et al. 2025) are likewise consistent with atmosphere-free, rocky surfaces. In contrast, LHS 1478 b exhibits a notably shallow eclipse depth at $15\ \mu\text{m}$ (P. C. August et al. 2025), potentially suggesting the presence of an atmosphere. However, the two observations of LHS 1478 b were found to be inconsistent due to systematics. This planet is scheduled for follow-up observations with both MIRI LRS and additional F1500W photometry under JWST GO program 7675 (PI: P. August), which aims to determine whether it possesses an atmosphere. In addition, several other hot rocky exoplanets orbiting M-dwarfs—LTT 1445A b (P. Wachiraphan et al. 2025), GJ 1132 b (Q. Xue et al. 2024), GJ 486 b (M. Weiner Mansfield et al. 2024), and GJ 367 b (M. Zhang et al. 2024)—have been observed using

MIRI LRS during secondary eclipse. The thermal-emission spectra for these exoplanets are reported to be consistent with bare-rock scenarios, showing no strong evidence of substantial atmospheres. On the other hand, transmission spectroscopy observations of the small M-dwarf planets L 98-59 b and d have provided tentative evidence for the presence of atmospheres (A. Banerjee et al. 2024; A. Gressier et al. 2024; A. Bello-Arufe et al. 2025). However, several other transmission spectroscopy observations of small planets orbiting M-dwarfs have reported nondetections or inconclusive results regarding atmospheric spectral features (e.g., O. Lim et al. 2023; J. Lustig-Yaeger et al. 2023; E. M. May et al. 2023; S. E. Moran et al. 2023; N. Scarsdale et al. 2024; M. K. Alam et al. 2025; L. Alderson et al. 2024, 2025).

In this work, we present four secondary eclipse observations of GJ 3473 b, a rocky exoplanet orbiting a slowly rotating M dwarf ($T_{\text{eff}} = 3347 \pm 54\ \text{K}$, $J = 9.62$) with an orbital period of 1.2 days, as part of the Hot Rocks Survey. We use JWST MIRI F1500W photometry to assess whether the planet exhibits thermal emission consistent with a bare, atmosphere-free surface, or whether the observed eclipse depth suggests the presence of an atmosphere. GJ 3473 b has a radius of $1.264 \pm 0.050 R_{\oplus}$ and a mass of $1.86 \pm 0.30 M_{\oplus}$ (J. Kemmer et al. 2020), with a zero-albedo equilibrium dayside temperature of approximately $T_{\text{day}} = 1003\ \text{K}$ (assuming no redistribution). A recent analysis by X. Ji et al. (2025) suggests that GJ 3473 b may retain a CO_2 atmosphere, provided it formed with a sufficiently large initial volatile inventory; however, it is less likely to have an atmosphere compared to most planets in their work. Our observations represent the first JWST characterization of GJ 3473 b and provide direct constraints on its dayside emission, contributing to the broader effort to determine if these worlds retain atmospheres under extreme irradiation.

In Section 2, we describe the JWST observations, data reduction, and light-curve fitting. Section 3 presents our forward modeling of the planet’s thermal emission, exploring a wide range of surface and atmospheric scenarios, including space weathering for airless surfaces. In Section 4, we systematically compare these models with the measured eclipse depth of GJ 3473 b to assess the plausibility of each scenario. In Section 5 we discuss the possibility of atmospheric collapse and the potential evidence for eclipse depth variability. Finally, in Section 6, we summarize our work and discuss the broader implications of our findings for atmospheric retention on GJ 3473 b.

2. Observations and Data Analysis

We observed four secondary eclipses of the exoplanet GJ 3473 b using MIRI F1500W photometry as part of the JWST GO program 3730. The observations took place on 2024 March 12, 2024 March 13, 2024 March 30, and 2024 October 20, with each observation lasting 3.4 hr (corresponding to a total charge time of 19.1 hr). We utilized the SUB256 subarray in FASTR1 readout mode with 39 groups per integration, resulting in a total of 1017 integrations per visit. Across the observations, the peak flux was kept below 50% detector saturation. Notably, there was no high-gain antenna movement during any of the visits. In addition to the MIRI observations, we also refined the orbital parameters of GJ 3473 b using TESS sectors 34, 61, and 88, as described in Appendix A. Below, we describe the data reduction and light-curve fitting of

the MIRI data, for which we used a version of the JEXORES pipeline (M. Holmberg & N. Madhusudhan 2023). Leveraging our four visits, we also characterize the detector settling of MIRI to inform aspects of the analysis.

2.1. Data Reduction

We begin by performing Stage 1 and Stage 2 processing using the JWST pipeline (H. Bushouse 2020). This includes the data quality initialization, emicorr, saturation, first-frame, last-frame, linearity, RSCD, dark current, and ramp fitting steps. We omit the jump detection step in Stage 1 and instead identify outliers during Stage 3. Following this, we run the source-type, assign wcs, and flat-field corrections in Stage 2. Stage 3 processing includes outlier correction, background subtraction, and aperture photometry. For the outlier and bad pixel correction step, we first mask all pixels flagged by the pipeline. We then identify temporal outliers by comparing each pixel’s light curve to a median-filtered version, flagging any points that deviate by more than 5σ . To account for cosmic-ray hits that can affect neighboring pixels, we additionally mask adjacent pixels around each flagged outlier. Affected values are interpolated linearly in time. Bad pixels are similarly corrected by interpolating along detector columns.

We perform the background subtraction in two steps. First, we perform a common background subtraction derived from the time-average image. To do this, we iteratively fit a cubic polynomial to each detector column while removing 3σ outliers. To start the process, we initially mask pixels within a 15-pixel radius of the point-spread-function (PSF) center. In the end, we obtain a background model and a mask of background pixels. Second, after subtracting this model from each image, we perform a time-dependent background subtraction using the average of all background pixels. After background subtraction, the standard deviation of the background pixels in the average image is less than 1 DN/s, with the largest deviation being 3.1 DN/s. Thus, in a worst-case scenario, where all pixels in the aperture have an excess flux of 3.1 DN/s, the eclipse depth would be diluted by less than 0.7%.¹⁶ As an additional method, we employed traditional aperture photometry by subtracting the average flux within an annulus that spans 20–40 pixels (e.g., P. C. August et al. 2025) from the centroid position (on a per-integration basis). We find that the eclipse depth is consistent between the two methods, differing only by a few ppm. Thus, we conclude that inaccuracies in background subtraction do not significantly affect the inferred eclipse depth.

Next, we fit the PSF of each image using a Gaussian to obtain an estimate of the centroid and width as a function of time. Using the average centroid position, we then create the extraction mask with a nominal radius of 6 pixels, as discussed below and in Appendix B. After extracting the flux, we remove integrations with more than 1% of the flux masked (within the aperture) to limit the influence of interpolated data. For robustness, we additionally search for outliers in the extracted light curves and mask integrations with 4σ outliers, using the same method as above. This results in an average of 33 discarded integrations per visit. Furthermore, we note that

¹⁶ The dilution factor can be calculated via $1/(1 + F_{\text{bgr}}/F_*)$, where F_{bgr} correspond to the total excess background flux within the aperture, and F_* is the stellar flux.

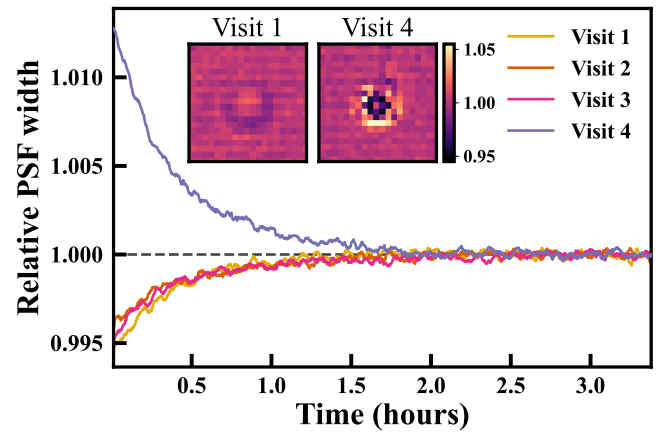


Figure 1. Settling behavior of the point-spread function (PSF) across the four visits. Visits 1–3 exhibit an increase in PSF width over time, while visit 4 shows the opposite trend. The two images display the flux ratio between the average of the first 20 integrations and the average of the final quarter of integrations for visits 1 and 4, respectively, prior to background subtraction. In visits 1–3, the central pixels are initially brighter and the surrounding pixels fainter, resulting in a narrow PSF at the start of the observations. In visit 4, the opposite pattern is observed, with a broader initial PSF that narrows over time.

there is a large cosmic-ray hit at the edge of the aperture in the second observation, occurring after the end of the eclipse. This caused a persistence signal in which the flux of a single pixel increased by 1%–2% for the remainder of the observation (similar events have been noted by Q. Xue et al. 2024; M. Fortune et al. 2025). However, given that the affected pixel accounts for less than 0.1% of the flux within the 6-pixel aperture, it does not significantly affect the inferred eclipse depth.

2.2. Detector Settling

MIRI time-series observations reveal a range of detector settling behaviors (e.g., J. Bouwman et al. 2023; T. P. Greene et al. 2023; T. J. Bell et al. 2024; M. Zhang et al. 2024), which appear as exponential-like ramps in the light curves with either positive or negative amplitudes. In this analysis, we characterize the settling behavior observed in our four observations to enhance our understanding of this phenomenon and to inform the data reduction and light-curve fitting processes.

By inspecting individual pixel light curves, we find that the detector settling behavior varies across the PSF, as mentioned by M. Fortune et al. (2025). Specifically, we find that the settling appears to depend on a pixel’s location within the PSF. We identify three distinct behaviors: (1) pixels at the edge of the PSF, which receive low illumination, show little to no settling; (2) pixels near the PSF core, which are highly illuminated, exhibit strong exponential-like ramps; and (3) pixels in an intermediate annulus display exponential trends with amplitudes opposite in sign to those of the central pixels. This spatial variation results in time-dependent changes in the overall PSF width, either broadening or narrowing, depending on the visit. In observations 1–3, the PSF width increases over time, whereas in observation 4 it decreases, stabilizing after approximately 2 hours, as illustrated in Figure 1. Although the amplitude of this effect varies, the timescale appears similar across visits. We measure an average e-folding time of 22.1 ± 2.1 minutes, in good agreement with the inferred light-curve settling timescale of 24.0 ± 2.6 minutes (from the canonical case; see Section 2.3). Moreover, we find that the

Table 1
The MIRI F1500W Eclipse Depth of GJ 3473 b Given Different Model Assumptions

Case	Trend	GP kernel	Orbit	$\ln B$	Eclipse Depth
Canonical case	Exp	Exp	Circular	0	186_{-45}^{+45} ppm
Effects of trend assumptions	Exp & linear function	Exp	Circular	-3.5	185_{-48}^{+49} ppm
	Exp, separate timescales	Exp	Circular	-4.3	192_{-51}^{+50} ppm
	PSF width	Exp	Circular	-6.7	189_{-54}^{+53} ppm
	Exp & PSF position	Exp	Circular	-17.9	174_{-38}^{+39} ppm
Effects of GP assumptions	Exp	Matern-3/2	Circular	-1.0	185_{-45}^{+45} ppm
	Exp	Exp, separate	Circular	-0.9	160_{-45}^{+46} ppm
	Exp	None	Circular	-1.8	185_{-34}^{+34} ppm
Eccentric orbit	Exp	Exp	Eccentric	-2.4	184_{-52}^{+52} ppm
Variable eclipse depth	Exp	Exp	Circular	2.1	see Table 2
Variable eclipse depth, no GP	Exp	None	Circular	1.5	see Table 2
No-eclipse depth	Exp	Exp	Circular	-5.1	Fixed at 0

Note. The model evidence for the canonical case is $\ln Z_0 = -19443.0$. The natural logarithm of the Bayes factor for each model is defined as $\ln B = \ln Z - \ln Z_0$. We consider $|\ln B| > 2.5$ and $|\ln B| > 5.0$ as moderate and strong evidence (R. Trotta 2008), respectively.

direction of the PSF width evolution correlates with the sign of the light-curve ramp:¹⁷ increasing width corresponds to a negative flux ramp, and vice versa. This suggests that temporal variations in the PSF may be related to the observed systematics in the light curves. Consequently, the PSF width time series may be valuable for light-curve modeling, either as a direct input or to inform priors. Finally, the variation in settling behavior between visits may be linked to persistence effects and differing illumination histories (D. Dicken et al. 2024; M. Fortune et al. 2025). Notably, prior to the fourth observation, the target was observed with the F560W filter for 45 minutes, delivering significantly more flux than with the F1500W filter. In contrast, the other visits were preceded by either the P750L prism or the F1500W filter. This discrepancy may explain why the fourth visit exhibits distinct settling characteristics compared to the others.

Due to the spatially dependent settling effects described above, the choice of aperture size can directly influence the shape of the extracted light curves (N. H. Allen et al. 2025) and consequently, potentially affect the inferred eclipse depth. We find that smaller aperture radii produce stronger exponential ramps, while larger apertures tend to suppress these trends. This occurs because the opposing exponential behaviors seen in different parts of the PSF (e.g., positive near the center and negative in surrounding annuli) partially cancel out when a broader region is integrated. However, this cancellation is incomplete; residual systematics remain even at large aperture sizes. To assess the impact of aperture size on the eclipse depth, we explore a range of aperture radii in Appendix B. We find that the uncertainty in the inferred eclipse depth is larger for smaller apertures, but stabilizes for apertures beyond a certain size. This highlights the importance of aperture selection in the presence of spatially variable systematics, particularly for high-precision eclipse measurements.

¹⁷ We measured the flux at the beginning of each observation to be 1950 ± 220 ppm, 1450 ± 240 ppm, 1900 ± 230 ppm, and -2940 ± 240 ppm relative to the baseline flux for visits 1-4, respectively. These values correspond to p_1 in the first case of Equation (2).

2.3. Light-curve Fitting

To perform the light-curve fitting, we utilize `batman` to model the eclipses (L. Kreidberg 2015) and `MultiNest` (F. Feroz et al. 2009) to sample from the posterior distributions using nested sampling (J. Skilling 2004). We jointly fit the light curves from the four visits using a variety of different assumptions, as outlined in Table 1. For all cases, we sampled the posterior distributions using 1000 live points. We show the systematics-corrected and phase-folded data along with our canonical model fit in Figure 2.

In general, we model the data for each visit as

$$F_{\text{obs}}(t) = F_* \epsilon_{\text{sys}}(t)(f_{\text{eclipse}}(t) + f_{\text{flare}}(t)), \quad (1)$$

where F_* is the stellar flux; ϵ_{sys} corresponds to our systematic model; f_{eclipse} is the eclipse model; and f_{flare} is a flare model (G. Tovar Mendoza et al. 2022), described in Appendix C. To model the detector settling, we employ a range of possible functions:

$$\epsilon_{\text{sys}}(t) = \begin{cases} 1 + p_1 e^{-(t-t_0)/p_2} \\ (1 + p_1 e^{-(t-t_0)/p_2})(1 + p_3(t - t_0)) \\ 1 + p_1 w(t) \\ (1 + p_1 e^{-(t-t_0)/p_2})(1 + p_3 x(t) + p_4 y(t)) \end{cases}, \quad (2)$$

where p_i are free parameters, t_0 is the time at the start of the visit, and w , x , and y are the time series of the position coordinates and width of the PSF,¹⁸ respectively. For our canonical case, we use an exponential function (first case in Equation (2)) with a common timescale across all visits, while allowing the amplitudes to differ, following the discussion in Section 2.2. Such a model is preferred over a model where the timescales are independent with a log Bayes factor of $|\ln B| = 4.3$. Using any of the other functions in Equation (2) yields lower evidence, as shown in Table 1. However, the

¹⁸ These time series are filtered with a median filter to reduce noise, using a window size of 15 integrations.

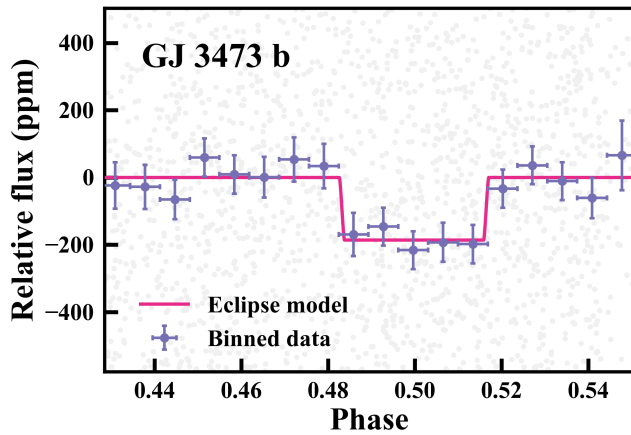


Figure 2. Phase-folded and systematics-corrected MIRI data and secondary eclipse model of GJ 3473 b. The unbinned and binned data are shown in light gray and purple, respectively. To detrend the data, we subtracted the predicted GP models and divided by the systematic models. The pink curve corresponds to an eclipse model with parameters from Table 5, corresponding to our canonical case.

inferred eclipse depth remains consistent among these different choices.

We mask the first few integrations (2 minutes) of each observation to remove any short-lived settling effects and, for the second visit, we also nominally mask integrations numbered 325–524. These additional integrations are masked to mitigate the effects of a temporary ~ 700 ppm increase in the flux just before the start of the eclipse during the second visit, as shown in Figure 3—potentially caused by a stellar flare. As discussed in Appendix C, we also modeled this excess flux with a flare model instead of masking it. However, we find consistent results between the two approaches. Additionally, we investigated the impact of masking different numbers of integrations at the beginning of the observations, as pursued in other works (e.g., P. C. August et al. 2025). We found that removing 2, 15, 30, or 60 minutes (corresponding to 10, 75, 150, and 300 integrations) does not significantly affect the inferred eclipse depth, as these result in average eclipse depths of 186^{+45}_{-45} ppm ($\chi^2_\nu = 1.003$), 192^{+42}_{-41} ppm ($\chi^2_\nu = 1.005$), 200^{+43}_{-41} ppm ($\chi^2_\nu = 1.007$), and 195^{+45}_{-44} ppm ($\chi^2_\nu = 1.009$), respectively. These values are statistically consistent within $< 0.34\sigma$, comparable to the variation introduced by alternative trend parameterizations (Table 1). As shown in Figure 3, the systematics model provides a good fit to the data from the start of each observation. Because removing additional data does not improve the fit quality but reduces the available baseline for constraining the exponential ramps and noise properties, we adopt the minimal 2 minute cut.

We account for potential time-correlated noise by using a Gaussian process (GP) model, implemented via *celerite* (D. Foreman-Mackey et al. 2017). For our canonical case, we use an exponential kernel with a common amplitude and timescale across all visits. We also included a jitter term (added in quadrature to the diagonal elements of the covariance matrix) to account for additional white noise. Given that the light-curve scatter varies between the different visits, ranging from 847 to 879 ppm, we kept the jitter term separate for each visit (i.e., adding four parameters). Doing this, we find that the white noise is about 20% higher compared to the expected uncertainties obtained by propagating the photon and read noise. In context, the measured

white noise corresponds to signal-to-noise ratios (SNRs) of 1140–1180, which is close, but somewhat higher, than the expected SNR of 1050 from the ETC.¹⁹ In addition to our canonical case, we also vary the GP assumptions as described in Table 1. We find a slight preference for an exponential kernel with common parameters compared to using separate kernel parameters for each visit, as well as compared to using a Matern-3/2 kernel. We note that the lower eclipse depth in the case with separate kernel parameters is a result of nonuniform weighting of the visits in combination with the somewhat variable eclipse depth, as discussed in Section 5.2. Ultimately, we determine that the overall preference for time-correlated noise is $|\ln B| = 1.8$. While this evidence is weak, we choose to use a GP due to its more conservative estimate of the eclipse depth uncertainty.

We confirm that we detect the secondary eclipse with a log Bayes factor of $|\ln B| = 5.1$ when compared to a model where the eclipse depth is fixed at zero. For this comparison, we assume a circular orbit which, as noted below, is preferred by the data. This level of evidence is considered strong, according to R. Trotta (2008). Additionally, we examine whether the eclipse depth remains consistent across the four observations by allowing it to vary. Among all the tests we considered, this was the only scenario that provided stronger evidence than our canonical model. However, the difference in evidence is relatively weak, with a log Bayes factor of only $|\ln B| = 2.1$. We discuss this further in Section 5.2.

We also examine whether the data is consistent with a circular orbit for GJ 3473 b or if an eccentric orbit is preferred. For this, we fit the eccentricity e and the argument of periastron ω using uniform priors ranging between 0 and 0.5 and $0-2\pi$ (with periodic boundary conditions), respectively. As shown in Table 1, we find that a model with a circular orbit is weakly preferred; however, the eclipse depth remains insensitive to this choice. To accurately model the timing of the eclipse, we consider the effect of light travel time. However, with a stellar radius of $R_* = 0.364 R_\odot$ (J. Kemmer et al. 2020), this translates to a delay of around 16 s, which is only marginally longer than the 12 s between integrations. We find that the eclipse timing is consistent with a circular orbit, since $e \cos(\omega) = -0.0010^{+0.0028}_{-0.0026}$ is consistent with zero (J. N. Winn 2010). Likewise, the eclipse duration is also consistent with that of a circular orbit, given $e \sin(\omega) = 0.005^{+0.085}_{-0.018}$. Overall, we determine an upper limit for the eccentricity, finding $e < 0.36$ at 95% confidence. This upper limit is broad because when $|e \cos(\omega)| \gtrsim 0.1$, the eclipse starts occurring outside of the observing windows, at which point we effectively recover the no-eclipse case. By excluding these solutions, the 95% upper limit becomes $e \lesssim 0.17$, which is primarily set by the upper limit of $e \sin(\omega)$. This result is consistent with previous radial velocity data (J. Kemmer et al. 2020), which also did not show any preference for an eccentric orbit.

Overall, the canonical case has a total of 20 free parameters used to fit all four visits simultaneously. Among these parameters, 12 vary between the different visits, including the stellar fluxes, settling amplitudes, and the jitter terms. The remaining parameters, which are common across the visits, include the settling timescale, GP amplitude, GP timescale, midtransit time, orbital period, normalized semimajor axis, inclination, and eclipse depth. The prior distributions and

¹⁹ We normalized the ETC model flux density to match the measured flux density of 7.974 mJy in the F1500W bandpass, as described in Section 2.4.

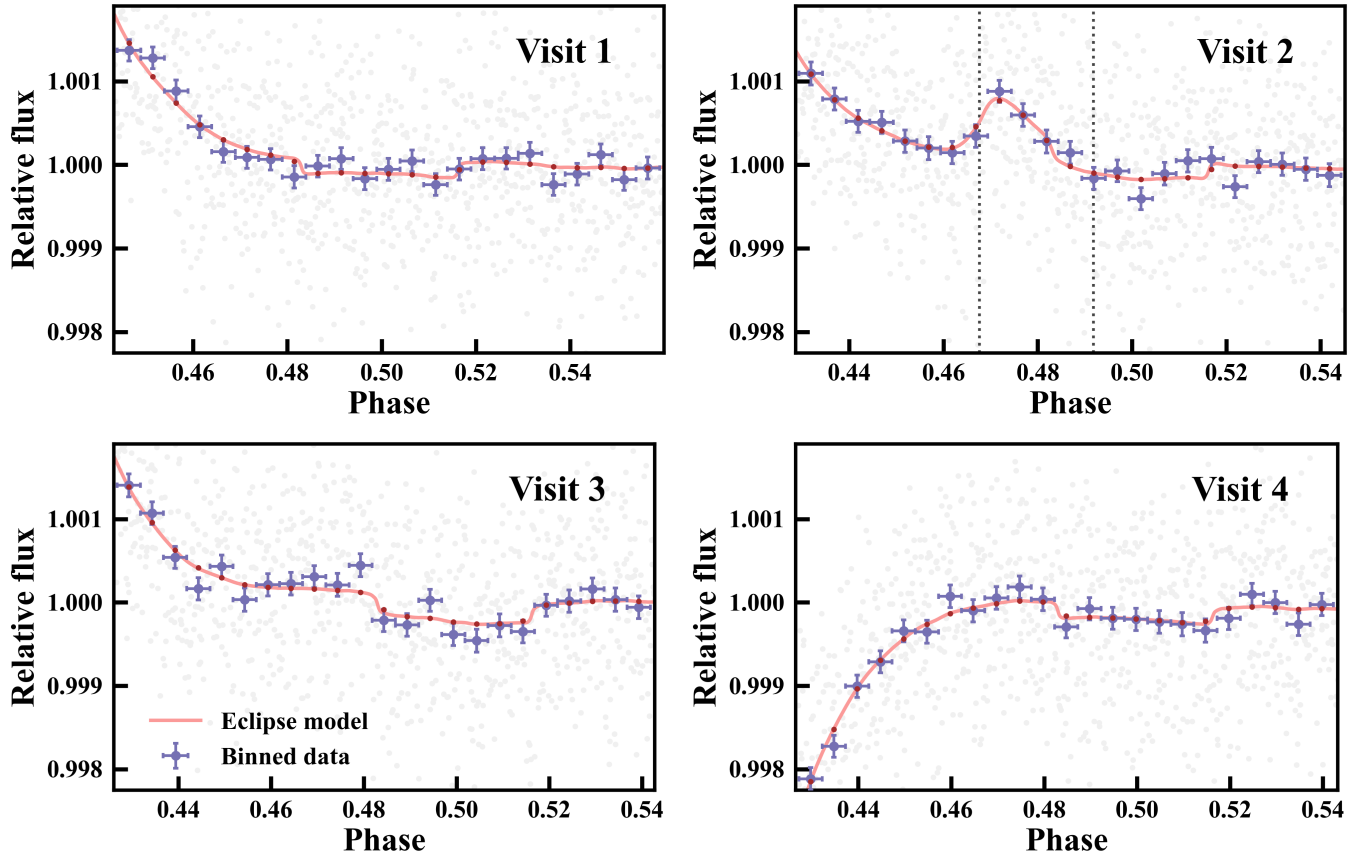


Figure 3. MIRI F1500W light curves of GJ 3473 b at secondary eclipse. The raw and binned data are shown in light gray and purple, respectively. The pink curves correspond to the median model predictions from the joint fit, while the brown data points depict the binned model. The model shown here corresponds to the G. Tovar Mendoza et al. (2022) flare model case, as described in Appendix C, used to model the flare-like feature during the second visit. The dotted vertical lines show the nominally masked region in the second visit. Compared to Figure 2, this figure displays the data without correcting for systematics.

posterior estimates for these parameters are presented in Appendix D.

2.4. Absolute Flux Calibration

In addition to measuring the eclipse depth of GJ 3473 b, we also measure the absolute flux density of the star in the MIRI F1500W bandpass. This allows us to compare the stellar flux with that of a model in the present bandpass to ensure consistency. To perform the absolute flux calibration, we follow the method outlined by K. D. Gordon et al. (2024). This involves using a specific aperture radius of 5.69 pixels and a background annulus range of 8.63–11.45 pixels, along with an aperture correction factor. The resulting measurement is then converted to physical units (DN/s to Jy) using a time-dependent calibration factor. For each visit, we carry out this flux calibration for all integrations and take the median value. The resulting flux density measurements are 7.974 ± 0.038 mJy, 7.976 ± 0.038 mJy, 7.975 ± 0.038 mJy, and 7.965 ± 0.038 mJy, for each visit, respectively. Note that the uncertainty in these measurements arises from a systematic calibration uncertainty of 0.48%, rather than photon noise, which is negligible in comparison. However, the repeatability among these four measurements appears to be much better, at around 0.07%, as computed from the standard deviation of the measurements. Furthermore, compared to K. D. Gordon et al. (2024), our Stage 1 processing included the emicorr step, while we did not use the jump step, as described

in Section 2.1. Nonetheless, we find that these steps do not significantly impact the absolute flux measurements.

Next, we compare our measurements with stellar models. We adopt a BT-Settl stellar model (F. Allard 2014), which we linearly interpolated given the stellar parameters from J. Kemmer et al. (2020). Along with the distance estimate from GAIA DR3 of 27.315 ± 0.018 pc (Gaia Collaboration et al. 2016, 2023) and a stellar radius of $0.364 \pm 0.012 R_{\odot}$ (J. Kemmer et al. 2020), we obtain a model flux density of 8.3 ± 0.5 mJy. For this estimate, we account for uncertainties in the effective temperature, distance, and stellar radius, with the majority of the uncertainty arising from the stellar radius. In the end, we find that the measured flux is in good agreement with the model prediction.

We also compute the predicted flux using a SPHINX stellar model (A. R. Iyer et al. 2023) to evaluate the systematic uncertainty caused by variations in the stellar model. For the SPHINX model, we obtain a model flux density of 8.1 ± 0.5 mJy, which is approximately 3% different from the BT-Settl stellar model. This small discrepancy in the stellar spectrum has only a slight impact on the derived brightness temperature and eclipse depth model predictions, affecting them by a few percent. This change is much smaller than the 24% uncertainty associated with the eclipse depth itself. Therefore, we conclude that the choice of stellar model has minimal influence on the results of this study.

3. Modeling

To assess the likelihood of an atmosphere on GJ 3473 b, we examine various atmospheric and surface model scenarios, assuming that the planet is tidally locked in a 1:1 spin-orbit resonance. Our goal is to explore the potential ranges of eclipse depth within the observed bandpass for these scenarios, allowing us to compare the results with the data. For each scenario, we calculate the eclipse depth as measured in the MIRI F1500W bandpass as follows

$$\frac{F_p}{F_\star} \Big|_{\text{F1500W}} = \left(\frac{R_p}{R_\star} \right)^2 \frac{\int \lambda F_p(\lambda) w_{\text{F1500W}}(\lambda) d\lambda}{\int \lambda F_\star(\lambda) w_{\text{F1500W}}(\lambda) d\lambda}, \quad (3)$$

where w_{F1500W} is the total throughput of JWST/MIRI using the F1500W filter obtained from Pandeia (version 4.0, February 2025; K. M. Pontoppidan et al. 2016), and F_p and F_\star are the planetary and stellar fluxes, respectively.

3.1. Surface Scenarios

We model the dayside thermal emission of an airless planet using the approach of K. Paragas et al. (2025). The surface temperature is set by the balance between the absorbed stellar flux and the emitted thermal radiation:

$$f \int \varepsilon_d F_{\text{inc}} d\lambda = \int \varepsilon_h B_\lambda(T_{\text{day}}) d\lambda, \quad (4)$$

where F_{inc} is the incident stellar irradiance, B_λ is the Planck function, and f is a geometric (redistribution) factor that accounts for the longitudinal temperature gradient across the dayside. Here, $\varepsilon_d = 1 - r_h$ is the directional emissivity given by Kirchhoff's law, with r_h being the directional-hemispherical reflectance (B. Hapke 2001; K. Paragas et al. 2025). The hemispheric emissivity ε_h is the hemispheric average of the directional emissivity. Physically, ε_d characterizes how efficiently the surface absorbs radiation arriving from the single direction of the star, while ε_h describes how efficiently the surface emits thermal radiation averaged over all outgoing directions. Both ε_h and f can be derived from the directional-hemispherical reflectance r_h (B. Hapke 2001; K. Paragas et al. 2025). Once T_{day} is determined, the emergent dayside flux of the planet is

$$F_p = \pi \varepsilon_d B_\lambda(T_{\text{day}}) + r_h F_\star \left(\frac{R_\star}{a} \right)^2, \quad (5)$$

where the first term represents thermal emission and the second term represents reflected starlight. We note that in this formalism r_h is a surface property, whereas reflected light is often expressed in terms of the geometric albedo A_g , an observational quantity defined as the disk-integrated reflectivity at full phase. For a Lambertian surface, these are related by $A_g = 2/3 r_h$. Furthermore, as described in Section 2.4, we use a BT-Settl stellar model (F. Allard 2014), which provides a good match to the measured flux in the MIRI F1500W bandpass. We use the normalized semimajor axis from the TESS data, which gives $a/R_\star = 9.32$, as described in Appendix A.

We consider a wide range of surface compositions and textures mainly drawn from the recent library of K. Paragas et al. (2025). These include Fe-oxidized materials (hematite), ultramafic rocks (dunite xenolith, olivine clinopyroxenite),

mafic rocks (olivine pyroxenite, basaltic andesite, Kilauea basalt, olivine gabbro), felsic rocks (dalmatian granite, orlando gold granite), and a feldspathic lunar anorthosite sample (Feldspathic; L. C. Cheek et al. 1928; R. Hu et al. 2012). Textures range from solid (low reflectance) to crushed or powdered (higher reflectance). While these materials provide a range of possible surface compositions motivated by solar system bodies, this list is not exhaustive. Other materials not considered here may exist on the surface of exoplanets like GJ 3473 b. Furthermore, we note that although we model the surface as being composed of a single material, a realistic surface could consist of a variety of different materials.

In addition to grain-scale texture, macroscopic roughness (e.g., craters) can produce thermal beaming, where thermal radiation is emitted anisotropically, often preferentially toward the star (J. R. Spencer 1990; J. S. V. Lagerros 1998). We do not account for this effect here, in line with previous works (e.g., Q. Xue et al. 2024).

3.2. Space Weathering

We additionally include effects of space weathering (e.g., C. M. Pieters & S. K. Noble 2016), whereby a surface not protected by an atmosphere is altered by a variety of processes, such as micrometeorite impacts and solar wind irradiation. In the solar system, these processes often lower the albedo of airless bodies—darkening the surface and imparting a red slope to the visible and near-infrared reflectance, depending on the mechanism. To capture this effect, we follow B. Hapke (2001) and X. Lyu et al. (2024), modeling space weathering as the addition of small fractions of nanophase metallic iron (npFe⁰) or graphite into the host surface material. In the solar system, iron explains the low albedo of the Moon (W. Cassidy & B. Hapke 1975; C. M. Pieters et al. 2000), while graphite has been invoked to explain the darkening of Mercury's surface (M. B. Syal et al. 2015). Weathering alters the reflectance of a surface, depending on the volume fraction of iron or graphite, the refractive index of the host material, and the mean photon path length D , which we assume is 10 μm for all materials—somewhat smaller than the average grain size of the powdered samples ($\sim 40 \mu\text{m}$; K. Paragas et al. 2025). The strength of weathering is determined by the volume fraction of these inclusions (up to 5%; X. Lyu et al. 2024). We note that the range of eclipse depths at 15 μm resulting from different levels of space weathering is not strongly dependent on the choice of D . For the refractive indices, we adopt values for iron and graphite from the Refractive Index database (M. N. Polyanskiy 2024). For the different surface types, we obtain the refractive indices from the Aerosol Refractive Index Archive.²⁰ For ultramafic materials, we used the refractive index of olivine, while for mafic materials, we used the refractive index of basalt. For lunar anorthosite, we instead used a constant (real) refractive index of 1.7 (B. Hapke 2001), since we could not find its wavelength-dependent refractive index. In Figure 4, we show the effects of iron space weathering on the single scattering albedo of basalt powder. In this case, as well as for other materials with a high optical albedo, the reddening caused by iron results in a reduction of the albedo at wavelengths where the incident stellar flux is significant. Meanwhile, it maintains high reflectance (i.e., low

²⁰ <https://eodg.atm.ox.ac.uk/ARIA/>

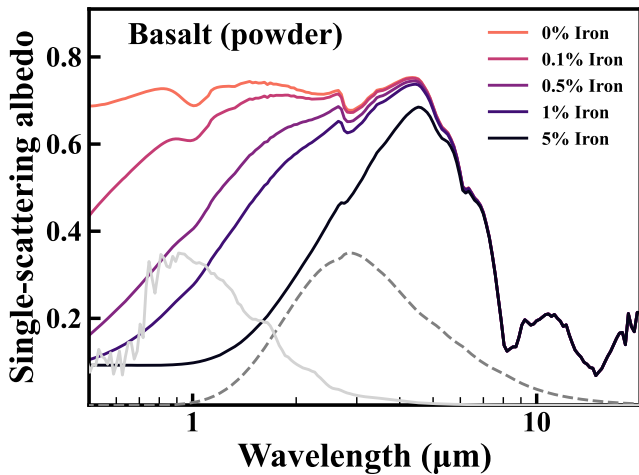


Figure 4. Single-scattering albedo of Basalt powder at varying degrees of space weathering via the formation of nanophase metallic iron particles. No added space weathering is represented by 0% iron. The spectra of the stellar irradiation and planetary flux, with 0.5% iron, are represented by solid light gray and dashed gray lines, respectively; however, these are not to scale for clarity.

emissivity) at longer wavelengths where the planet emits, causing the surface temperature to potentially exceed that of a simple blackbody.

3.3. Atmospheric Scenarios

For the atmospheric scenarios, we utilize HELIOS (M. Malik et al. 2017, 2019a, 2019b) to calculate the temperature structure and emission spectrum of the planet. In this work, we consider well-mixed atmospheres composed entirely of either pure CO₂ or H₂O, representing end-member cases for potential atmospheric compositions. The model employs k-distribution tables for the opacities, calculated using the HELIOS-K opacity calculator (S. L. Grimm & K. Heng 2015; S. L. Grimm et al. 2021), while integrating radiative fluxes over 386 spectral bands and 20 Gaussian points. We used cross-sections for CO₂ (L. S. Rothman et al. 2010) and H₂O (R. J. Barber et al. 2006). Additionally, we incorporate Rayleigh scattering (A. N. Cox 2000; M. Snee & W. Ubachs 2005; R. Thalman et al. 2014) and collision-induced absorption cross-sections for CO₂ and H₂O (M. Gruszka & A. Borysov 1997; Y. I. Baranov et al. 2004; E. J. Mlawer et al. 2023). In terms of the surface pressure, we consider a wide range of between 0.1 mbar and 100 bar. At the surface, we adopt a constant albedo of 0.1 (e.g., P. C. August et al. 2025). We nominally select an albedo of 0.1, approximately corresponding to basalt, rather than zero, since an atmosphere would protect the surface against space weathering. For the heat redistribution factor, we use the approximate scaling relation by D. D. B. Koll (2022). The stellar spectrum is derived from the BT-Settl model grid (F. Allard 2014), which is linearly interpolated based on the stellar parameters. We also take into account the effect of a nonzero Bond albedo. Rather than using a cloud or haze prescription, we simulate a Bond albedo of $A_B = 0.3$ (similar to that of Earth) by increasing the semimajor axis. Finally, we note that other atmospheres not considered here are certainly possible, in terms of different gas species, possible contributions from clouds or hazes, or dynamics, which could impact the temperature structure and the resulting emission spectrum. However, given the data

quality, we treat these idealized cases as distinct end-member scenarios, similar to our approach with the bare-rock scenarios.

4. Results

In this work, we provide the first measurement of the secondary eclipse depth of GJ 3473 b to assess the likelihood of the planet possessing an atmosphere. Using our four visits with JWST, we find an average eclipse depth of $F_p/F_\star = 186 \pm 45$ ppm in the MIRI F1500W bandpass, corresponding to a brightness temperature of 820 ± 120 K,²¹ which we obtained by solving for T_{day} in Equation (3). The eclipse is detected with a log Bayes factor of $\ln B = 5.1$, corresponding to strong evidence (R. Trotta 2008). We find that the ratio of the brightness temperature at 15 μm to that of a blackbody is $R = 0.82 \pm 0.12$. In this section, we evaluate if the measured eclipse depth of GJ 3473 b is consistent with the planet having a bare-rock surface, an atmosphere, or if the data is inconclusive.

4.1. Implications for an Atmosphere

In order to evaluate the likelihood of an atmosphere of GJ 3473 b, we use forward modeling to explore a range of possible atmosphere and surface compositions. The goal is to assess the range of the eclipse depths allowed by different scenarios, both for an atmosphere and a bare-rock surface, in order to compare these with our measurement of $F_p/F_\star = 186 \pm 45$ ppm in the MIRI F1500W bandpass. For this, we assume that GJ 3473 b is tidally locked.

Starting with the bare-rock scenarios, we compute the predicted eclipse depths for each of the different surface categories: Fe-oxidized, ultramafic, feldspathic, mafic, and felsic, as described in Section 3.1. These categories contain a range of materials with different textures (K. Paragas et al. 2025). For each of these, we consider different levels of space weathering by iron and graphite (between 0% and 5% volume fraction, as described in Section 3.2). We then take the minimum and maximum eclipse depths of these cases to represent the possible ranges for each of the surface categories. These limits, along with our measurement, are shown in Figure 5. Note that the eclipse depths occasionally exceed that of blackbody for some highly reflective materials when adding space weathering, as explained in Section 3.2. We find that the measured eclipse depth of GJ 3473 b is consistent with a range of different surface compositions. In fact, all surface compositions that we assess are consistent with the data within 2σ . However, surfaces that allow for a lower brightness temperature, such as ultramafic, feldspathic, and felsic materials, are somewhat preferred. Among the considered surface compositions, we find that granite, being highly reflective, allows for the lowest brightness temperature in the MIRI F1500W bandpass—with an eclipse depth of 130 ppm. We note that other highly reflective materials may also be possible (M. Hammond et al. 2025). Furthermore, while we consider surfaces of a single composition, this may not be realistic. For instance, granite, which is found only on Earth among solar system bodies, is not distributed uniformly on the planet. Overall, we cannot rule out the possibility of GJ 3473 b being a bare-rock given our photometric data alone.

²¹ In the absence of an atmosphere, the brightness temperature serves as a lower limit of the dayside surface temperature, unless the emissivity in the observed bandpass is equal to one, in which case they are the same.

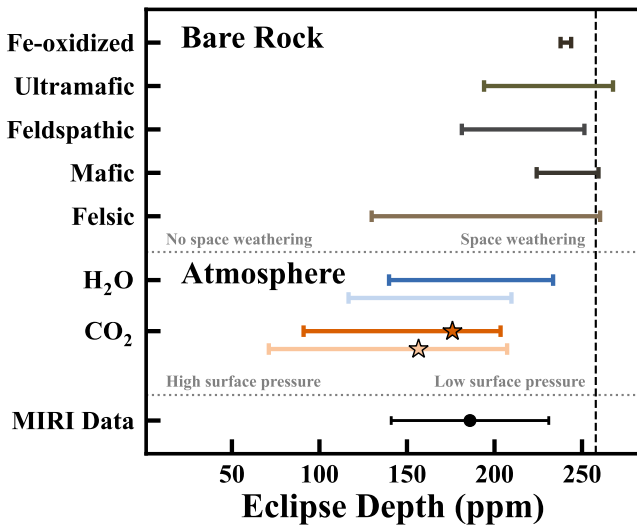


Figure 5. The span of eclipse depths of GJ 3473 b in the MIRI F1500W bandpass for different bare-rock and atmosphere scenarios, assuming that the planet is tidally locked in a 1:1 spin-orbit resonance. In the bare-rock cases, the limits arise from different materials, textures, and varying degrees of space weathering, as described in Section 3.1. In contrast, in the atmospheric scenarios, the limits are determined by varying surface pressures, which range from 0.1 mbar to 100 bar, as outlined in Section 3.3. The two shades of color represent Bond albedos of 0 (darker) and 0.3 (lighter), respectively. The stars indicate the critical surface pressure (0.6–1 mbar) below which a CO₂ atmosphere is susceptible to collapse, as discussed in Section 5.1. The dashed vertical line represents the eclipse depth of a blackbody. The measured average eclipse depth of GJ 3473 b is shown in black at the bottom of the figure.

We also explore a range of different atmospheric scenarios. We consider two types of atmospheres: 100% CO₂ or H₂O with surface pressures between 0.1 mbar and 100 bar, as described in Section 3.3. For the redistribution factor, we consider the scaling relation by D. D. B. Koll (2022), resulting in low redistribution at low surface pressures and higher redistribution at higher surface pressures. Out of these scenarios, a CO₂ atmosphere has been suggested to potentially be possible on GJ 3473 b, given a sufficiently high initial volatile content (X. Ji et al. 2025); however, this is less likely compared to most planets in their study. The same study found that N₂, H₂O, and CH₄ atmospheres are less likely to persist compared to CO₂. We additionally consider cases that approximate a lower Bond albedo of 0.3, as detailed in Section 3.3. The ranges on the possible eclipse depths, in the MIRI F1500W bandpass, for these different atmospheric compositions are shown in Figure 5. The upper limits are set by atmospheres with low surface pressure, while the lower limits are reached for the highest pressures. Even though we cannot differentiate between a bare-rock surface or an atmosphere, the data allows us to rule out a thick 100% CO₂ atmosphere, placing a 95% credible upper limit of 1.2–6.5 bar on the surface pressure, depending on the Bond albedo (0–0.3). All other atmospheric scenarios that we consider are consistent with the data. Moreover, while pure atmospheres consisting of a single gas are not entirely realistic, they serve as end-member cases given our limited data. In the end, additional observations are required to determine if GJ 3473 b has an atmosphere or not.

Finally, GJ 3473 b might be in a higher-order spin-orbit resonance, similar to Mercury’s 3:2 resonance, or it may not be tidally locked. In such a case, stellar irradiation would be

distributed more evenly across the planet’s surface, depending on its rotation rate and the thermal inertia of its surface or atmosphere (X. Lyu et al. 2024). This could reduce the dayside temperature and lower the secondary eclipse depth. As a result, a low measured eclipse depth could be mistakenly interpreted as evidence for an atmosphere, when it may instead reflect the effects of nonsynchronous rotation. Although short-period planets are generally assumed to be tidally locked due to the typically rapid timescales for tidal synchronization, this assumption depends on several factors, including orbital eccentricity, internal structure, atmospheric presence, and potential gravitational perturbations from other planets (R. Barnes 2017). Notably, a higher-order spin-orbit resonance requires a nonzero eccentricity—typically above $e \gtrsim 0.01$ (M. Turbet et al. 2018)—which remains fully within our 95% upper limit of $e \lesssim 0.17$ (see Section 2.3). Future phase-curve observations of GJ 3473 b could help break this degeneracy by constraining the longitudinal temperature distribution, as demonstrated for the rocky exoplanet LHS 3844 b, which was shown to be tidally locked in a 1:1 resonance (X. Lyu et al. 2024).

4.2. A Bayesian Approach

In Figure 5, we present the eclipse depth ranges of different bare-rock and atmosphere scenarios in the MIRI F1500W bandpass, suggesting that many of these cases are degenerate with only one photometric data point. However, the ranges indicated solely reflect the maximum and minimum eclipse depths for each scenario. To better quantify the preference for an atmosphere compared to a bare-rock surface, given our data, we employ a Bayesian model comparison framework. From this analysis, we can derive the posterior probability of an atmosphere by summing the probabilities of our atmospheric models and comparing this total to that of the bare-rock models, under the assumption that the planet is tidally locked in a 1:1 spin-orbit resonance.

For each case, we compute the Bayesian evidence by integrating the likelihood over the prior, normalized by the prior volume. For the bare-rock surface cases, we vary the level of space weathering, with log-uniform priors between 0.1 ppm and 5%. For the atmospheric cases, we vary the surface pressure with a log-uniform distribution between 0.1 mbar and 100 bar (linearly interpolating between the model grid). We show some of the best-fit spectra in Figure 6. We then calculate the probability p'_i of each model i , using the (log) Bayesian evidence $\ln Z_i$, via

$$p'_i = \pi_i \exp(\ln Z_i - \ln Z_{\max}), \quad (6)$$

where $\ln Z_{\max}$ corresponds to the maximum evidence²² and where π_i is the prior probability of the model itself. Since it is difficult to quantify the prior probability of each model, we construct a set of agnostic priors that ensures (1) an equal probability of having either an atmosphere or a bare-rock surface, (2) an equal probability for each type of atmosphere or surface (e.g., a mafic and an ultramafic surface are given the same prior probability), and (3) an equal probability for each model within each type. For example, the last condition means that the atmospheric cases with bond albedos of 0 and 0.3 are

²² We subtract $\ln Z_{\max}$ for numerical stability.

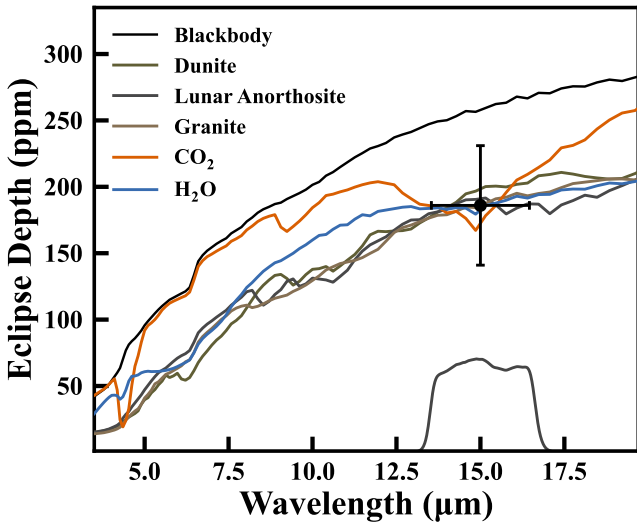


Figure 6. The eclipse depth of GJ 3473 b as a function of wavelength for a range of best-fit bare-rock and atmosphere scenarios that are consistent with the data, as described in Section 4.2. The data for the dunite (ultramafic) and granite (felsic) surfaces are obtained from K. Paragas et al. (2025), while the lunar anorthosite (feldspathic) data are from R. Hu et al. (2012). The free parameter for the bare-rock surfaces was the level of space weathering (in this case, iron), whereas the surface pressure was the free parameter for the atmospheric scenarios (with best-fit values of 0.4 mbar and 34 mbar for CO₂ and H₂O, respectively). For comparison, the eclipse depth of a blackbody is shown in black.

equally likely. We express this as follows

$$\pi_i = \frac{1}{2} \frac{1}{N_{i,\text{types}}} \frac{1}{N_{i,\text{models}}}, \quad (7)$$

where $N_{i,\text{types}}$ are the number of atmosphere/surface types (2 and 5, respectively), and where $N_{i,\text{models}}$ are the number of models in each type. For example, for mafic surfaces, we have 20 models that correspond to 10 different samples from K. Paragas et al. (2025), with both iron and graphite space weathering. Next, we compute the overall probability of an atmosphere and a bare-rock surface as

$$P'_{\text{atm}} = \sum_{i \in \text{atm}} p'_i, \quad P'_{\text{rock}} = \sum_{i \in \text{rock}} p'_i, \quad (8)$$

which we then normalize,

$$P_{\text{atm}} = \frac{P'_{\text{atm}}}{P'_{\text{atm}} + P'_{\text{rock}}}, \quad P_{\text{rock}} = 1 - P_{\text{atm}}. \quad (9)$$

For convenience, we convert the probability of an atmosphere to the number of standard deviations σ away from the mean of a normal distribution, via $\sigma = \Phi^{-1}(P_{\text{atm}})$, where Φ^{-1} is the inverse cumulative distribution function. This defines a one-sided significance, such that $\sigma = 0$ corresponds to equal posterior probabilities, $P_{\text{atm}} = P_{\text{rock}} = 1/2$, while positive values indicate a preference for an atmosphere, and negative values indicate a preference for a bare-rock surface. We refer to $\pm\sigma$ as the detection significance for either an atmosphere or a bare-rock surface. Note that a one-sided significance of 2.5σ , as defined above, corresponds to an odds ratio of $\sim 150:1$, which we adopt as the threshold for strong evidence (R. Trotta 2008).

Next, we explore various model and prior considerations. Our first case consists of including all our model scenarios,

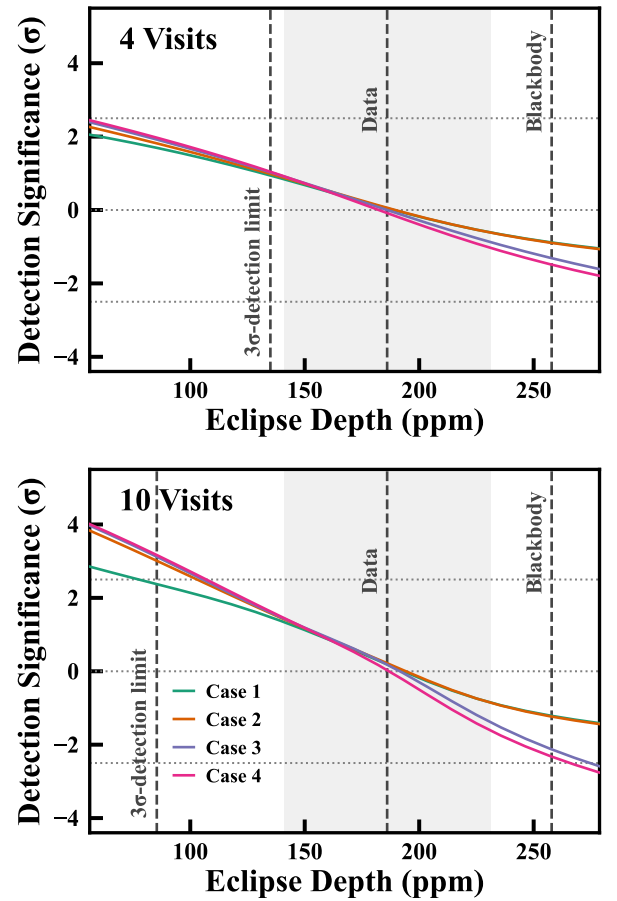


Figure 7. Detection significance of an atmosphere on GJ 3473 b as a function of the eclipse depth in the MIRI F1500W bandpass, with 4 (top) and 10 (bottom) visits. Positive values correspond to the detection significance of an atmosphere, while negative values correspond to the detection significance of a bare rock. At a significance of zero, these two scenarios are equally likely. The four cases are described in Section 4.2, representing different prior considerations. The vertical dashed lines denote the required eclipse depth for a 3σ detection of the eclipse itself, the actual measured eclipse depth, and the eclipse depth of a blackbody, respectively. The light-gray region represents the uncertainty of the present measurement. The horizontal dotted lines show the detection significances at 0 and $\pm 2.5\sigma$.

using the five surface types and two atmospheric scenarios discussed above and shown in Figure 5. For the second case, we remove felsic surfaces, i.e., highly reflective granitoids, as these have been suggested to be unlikely to form at the high temperatures of GJ 3473 b (M. Mansfield et al. 2019). As a third case, we only consider a CO₂ atmosphere while also excluding the felsic surfaces (consistent with case 2). We include this case because CO₂ atmospheres are expected to be more resilient to the high irradiation compared to other atmospheric compositions, such as H₂O or N₂ (e.g., X. Ji et al. 2025). In our fourth case, we begin with the conditions of case 3 but restrict the lower prior range of the surface pressure to 0.6 and 1 mbar, for bond albedos of 0 and 0.3, respectively, rather than the nominal 0.1 mbar. These pressures correspond to the critical surface pressures for the collapse of a CO₂ atmosphere, as discussed later in Section 5.1. Thus, these cases progress from considering all scenarios equally to imposing several constraints on the prior space based on theoretical expectations. The purpose of examining these cases is to assess how these different choices influence the potential inference of an atmosphere on GJ 3473 b. Figure 7 shows the detection

significance of an atmosphere for a range of possible eclipse depths, with either 4 or 10 visits.

Based on the measured eclipse depth of 186 ± 45 ppm using four visits, we find that it is not possible to distinguish whether GJ 3473 b has an atmosphere or a bare-rock surface. The posterior probability of the planet having an atmosphere is estimated to be between 47% and 52% (across the four cases), which aligns with the assumed prior probability of 50%. This finding is consistent with the scenarios presented in Figure 5, showing that various atmospheric and bare-rock scenarios are compatible with the data. Furthermore, by considering a range of potential eclipse depths, we find that four observations alone are insufficient to provide a posterior probability that exceeds 2.5σ for either an atmosphere or a bare-rock surface on GJ 3473 b, regardless of the measurement outcome. While a low eclipse depth could yield a 2σ preference for the presence of an atmosphere, in such a scenario, the eclipse itself would remain undetected. This could lead to potential ambiguity regarding the existence of an atmosphere and the timing of the eclipse itself, e.g., due to nonzero orbital eccentricity.

If we instead had a total of 10 MIRI F1500W visits, it would be possible to achieve a stronger preference for an atmosphere. For this experiment, we scaled the measured eclipse depth uncertainty by $\sqrt{4}/\sqrt{10}$, resulting in an uncertainty of 28 ppm with 10 visits. Among the four cases considered, we find that case 1 produces the lowest detection significance for an atmosphere, due to the degeneracy between an atmosphere and a high-albedo surface. As a result, the preference for an atmosphere increases at low eclipse depths when we assume that highly reflective felsic surfaces are not possible (cases 2–4). Furthermore, with 10 observations, we find that it is impossible to infer a bare-rock surface with high confidence if we consider both H₂O and CO₂ atmospheres down to a surface pressure of 0.1 mbar (cases 1–2). However, if we can rule out H₂O atmospheres in advance and limit the surface pressure to be above 0.6–1 mbar (case 4), we could establish a 2σ – 2.5σ preference for a bare-rock surface, provided that the measured eclipse depth is sufficiently high.

These results emphasize the opportunity and challenges of detecting an atmosphere using MIRI F1500W photometry alone and the importance of carefully considering prior knowledge. By allowing for highly reflective surfaces, atmospheric compositions beyond just CO₂, and for surface pressures down to 0.1 mbar, we demonstrate that distinguishing between an atmosphere and a bare-rock surface on GJ 3473 b is difficult, even with 10 MIRI F1500W observations. Future research can incorporate more complex models, including atmospheres with a mixture of gases, a broader range of temperature structures, potential clouds or hazes, and additional surface types or a combination of different surfaces. These factors may change the overall constraining power of the data. Additionally, we note that it might be easier to infer the presence of an atmosphere on rocky exoplanets other than GJ 3473 b. Although we show that confidently detecting an atmosphere with only MIRI F1500W photometry can be challenging, our framework provides a useful metric, namely P_{atm} , to identify promising targets for follow-up observations. Additional observations, e.g., using MIRI LRS, could provide stronger constraints on the presence or absence of an atmosphere (e.g., A. A. Piette et al. 2022).

5. Discussions

In this section, we evaluate whether the eclipse depths from our four visits are consistent and explore whether the potential for atmospheric collapse can be used to constrain the prior on the surface pressure of a CO₂ atmosphere for GJ 3473 b or other rocky exoplanets.

5.1. The Possibility of Atmospheric Collapse

Given that GJ 3473 b is likely tidally locked (J. Leconte et al. 2015), we consider the possibility of atmospheric collapse in the case of a pure CO₂ atmosphere. Collapse can occur if the nightside temperature falls below the condensation temperature of CO₂, turning the atmosphere into surface ice on the planet’s permanent nightside (J. F. Kasting et al. 1993). While GJ 3473 b is strongly irradiated, collapse may still occur if heat redistribution becomes sufficiently inefficient at low surface pressures, allowing the nightside to become a cold trap, assuming no other sources of heating and negligible abundances of other gases. This could provide a useful lower limit of the surface pressure for CO₂ atmospheres.

To estimate the critical surface pressure below which the atmosphere is susceptible to collapse, we approximate the nightside temperature following R. Wordsworth (2015):

$$T_{\text{night}} = \left(\frac{(1 - A_B) \kappa p_s}{4g} \right)^{1/4} \sqrt{\frac{R_*}{a}} T_{\text{eff}}, \quad (10)$$

where $\kappa = 1.6 \times 10^{-4} \text{ m}^2 \text{ kg}^{-1}$ is the approximate infrared opacity for CO₂, and $g = 11.15 \text{ m/s}^2$ is the surface gravity of GJ 3473 b. Solving for the surface pressure using the CO₂ condensation curve from R. Wordsworth (2015), we find a critical threshold of 0.6–1 mbar for Bond albedos between 0 and 0.3. These pressures correspond to eclipse depths between 157 and 176 ppm in the MIRI F1500W bandpass, as illustrated in Figure 5. Since the eclipse depth corresponding to the critical surface pressure is consistent with the data, we cannot rule out the presence of a thin CO₂ atmosphere. However, we note that more recent work suggests that the critical surface pressure can exceed the purely radiative estimate used here (P. Auclair-Desrotour & K. Heng 2020). For other close-in rocky planets with lower irradiation, the critical surface pressure would be higher than for GJ 3473 b, potentially providing a valuable constraint when interpreting observations.

Among the planets in our program, we find that LHS 1140 c exhibits the highest critical surface pressure, ranging from 13 to 19 mbar for Bond albedos between 0 and 0.3. Notably, this aligns with the 3σ upper limit of 10 mbar for a CO₂ atmosphere reported by M. Fortune et al. (2025). Consequently, since lower pressures are not expected to be possible, we may be able to rule out a CO₂ atmosphere on LHS 1140 c, if the planet is tidally locked. However, this conclusion depends on the specific temperature structure of the atmosphere, which is affected by dynamics, internal heating, clouds/hazes and the amount of noncondensable gas.

5.2. Eclipse Depth Variability

We find some evidence for a variable eclipse depth of GJ 3473 b at $15 \mu\text{m}$, although with a log Bayes factor of $\ln B = 2.1$ (compared to a constant eclipse depth), this evidence is weak at best (R. Trotta 2008). The evidence of

Table 2

The Eclipse Depth of GJ 3473 b for Each of the Four Visits, With and Without Gaussian Processes (GP)

Date	Eclipse Depth (GP)	Eclipse Depth (no GP)
12 March 2024	33^{+80}_{-83} ppm	47^{+62}_{-62} ppm
13 March 2024	103^{+93}_{-95} ppm	114^{+72}_{-71} ppm
30 March 2024	371^{+83}_{-81} ppm	357^{+67}_{-67} ppm
20 October 2024	224^{+86}_{-84} ppm	218^{+67}_{-70} ppm

Note. The visits are provided in the same order as in Figure 8. For these cases, we jointly fit the light curves while allowing for different eclipse depths with a uniform prior between ± 1000 ppm. Although the eclipse depths are somewhat different, the evidence for variability is weak, as described in Section 5.2.

this model, with varying eclipse depths, was computed using a uniform prior range of 0–500 ppm for the eclipse depths—the same as in the canonical case. To measure the eclipse depths of the different visits, we instead adopted a uniform prior range of ± 1000 ppm to not be influenced by the edges of the priors. The measured eclipse depths for each visit are listed in Table 2 and shown in Figure 8. The weighted average of these depths is 186 ± 43 ppm, consistent with the result from the constant eclipse depth fit (canonical case). Using this average, we obtained $\chi^2 = 9.53$ with three degrees of freedom. This corresponds to a p-value of 2.3%, indicating some evidence against a constant eclipse depth. We also note that although the eclipse depths from visits 1 and 3 differ from the canonical average by approximately 2σ ; excluding these does not significantly affect the overall average. When we consider only visits 2 and 4, we obtain an average of 170 ± 63 ppm, which is consistent with the canonical result.

This apparent variability, if not a statistical outlier, could be planetary in nature or be due to unaccounted systematic errors, such as stellar activity (e.g., minor flares) or unknown instrumental effects. For instance, P. C. August et al. (2025) found that two observations of LHS 1478 b—also using MIRI F1500W photometry—were inconsistent due to unknown systematics during the second visit. In our case, however, the low evidence for time-correlated noise ($\ln B = 1.8$, with a nominal aperture of 6 pixels), the absence of correlation with the PSF positions ($\ln B = -17.9$), and the stability of the stellar flux between visits ($\sim 0.07\%$) suggest that the quality of our observations is high. It is noteworthy that the visits with the most significant difference in eclipse depth—visits 1 and 3—exhibit nearly identical detector settling behavior. Both observations show an excess flux of around 1900 ppm at the start of the visits (see Section 2.2). Therefore, we conclude that the differences in the detector settling ramps are unlikely to affect the measured eclipse depths. Moreover, we note that the evidence for time-correlated noise drops from $\ln B = 1.8$ to $\ln B = 0.6$ when allowing for a variable eclipse depth, indicating some degeneracy between the presence of time-correlated noise and variable eclipse depth. We also find that the evidence for variability increases to $\ln B = 3.3$ when we do not use a GP. However, out of all the cases, the model with the highest evidence includes both time-correlated noise and variability in eclipse depth, as described in Section 2.3. In the end, the overall evidence for a variable eclipse depth across our four visits remains weak and may therefore not be real.

Eclipse depth variability has previously been reported for 55 Cnc e using Spitzer photometry (B.-O. Demory et al. 2015;

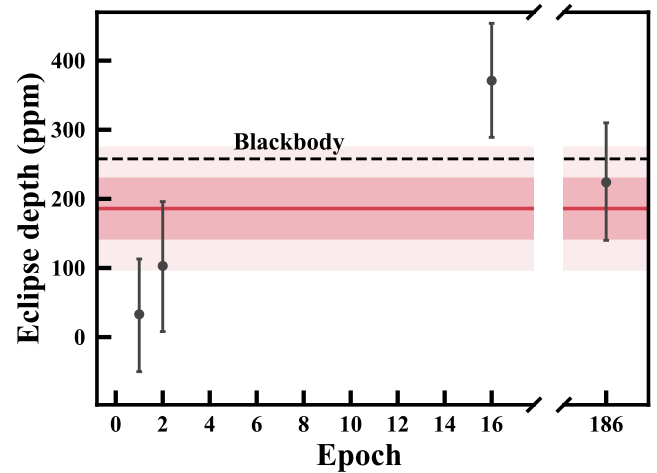


Figure 8. The eclipse depth of GJ 3473 b as a function of epoch, starting on 12 March 2024. The solid red line corresponds to the average eclipse depth obtained from the canonical case. The two lighter shades of red illustrate the 1σ and 2σ regions. The dashed black line shows the eclipse depth of a blackbody.

P. Tamburo et al. 2018) and recently with JWST (J. A. Patel et al. 2024). For 55 Cnc e, the cause of the variability has not been robustly established; however, several possibilities have been proposed, such as volcanic plumes or circumstellar/circumplanetary material (B.-O. Demory et al. 2015; K. Heng 2023). If the apparent variability of GJ 3473 b is planetary in nature, this could provide evidence for an atmosphere on the planet, as it is unlikely that a bare-rock surface could vary significantly over the time spanned by our observations. On the other hand, circumstellar/circumplanetary material, such as Io’s plasma torus, may act to change the apparent eclipse depth over time. Furthermore, the fact that the eclipse depths of the first two observations (only separated by one orbital period) remain consistent within 1σ , while the other observations (separated by $\gtrsim 20$ days) are somewhat different, may be evidence for a gradual change over time. We also note that the measured eclipse depth appears to vary by a factor of a few (e.g., ~ 3.6 between visits 2 and 3), similar to the variability reported for 55 Cnc e (B.-O. Demory et al. 2015). Ultimately, more observations are needed to reliably demonstrate or rule out a variable eclipse depth of GJ 3473 b.

6. Conclusions

We present the first JWST secondary eclipse observations of GJ 3473 b, a highly irradiated rocky exoplanet orbiting a mid-M dwarf, as part of the Hot Rocks Survey (JWST GO program 3730). Using four MIRI F1500W photometry visits, we detect the eclipse at high confidence ($\ln B = 5.1$), with an average eclipse depth of 186 ± 45 ppm. To ensure the robustness of our analysis, we explore an extensive range of data reduction and light-curve fitting assumptions, including variations in background subtraction, extraction methods, and the modeling of systematics, as outlined in Section 2. In addition, in the second visit we identify a ~ 700 ppm flare-like feature (Appendix C), for which we estimate an energy of $3.24^{+1.08}_{-0.86} \times 10^{28}$ erg in the F1500W bandpass. We find that the data are consistent with the planet having a circular orbit, and place an upper limit on the eccentricity of $e \lesssim 0.17$ at 95% confidence. Our observations provide the first constraints on

the thermal emission of GJ 3473 b, allowing us to empirically assess the likelihood of an atmosphere.

We find that the measured eclipse depth is consistent with both bare-rock or atmospheric scenarios. Forward modeling of airless surfaces, accounting for a wide range of compositions, textures, and degrees of space weathering, indicates that a variety of materials—including mafic, ultramafic, feldspathic, and felsic surfaces—can reproduce the observed eclipse depth within the measurement uncertainties (at less than 2σ). Notably, highly reflective materials like granite yield the lowest brightness temperatures among the surfaces that we consider, and is compatible with the data. However, we caution that such compositions may be geologically unrealistic for planets like GJ 3473 b (e.g., M. Mansfield et al. 2019). However, other surface compositions not considered here may be present, potentially widening the range spanned by the bare-rock scenario. Finally, if the planet is not tidally locked in a 1:1 spin-orbit resonance, the dayside temperature could be significantly cooler, making it more difficult to differentiate between the various scenarios.

Additionally, we also explore a suite of idealized atmospheric scenarios, including CO₂ and H₂O atmospheres over a broad range of surface pressures. While the data do not require GJ 3473 b to have an atmosphere, we are able to rule out a thick CO₂ atmosphere, placing a 95% credible upper limit of 1.2–6.5 bar on the surface pressure. Conversely, thin CO₂ atmospheres and atmospheres with other compositions remain consistent with the observed eclipse depth. Without further constraints of the prior parameter space, it may be challenging to confidently detect an atmosphere using MIRI F1500W photometry alone, due to the degeneracies and many unknowns involved. Nevertheless, a low eclipse depth may still identify promising candidates for follow-up observations. Ultimately, detecting atmospheric features through spectroscopy can help definitively determine whether an atmosphere is present.

To further constrain the atmospheric parameter space, we consider the possibility of atmospheric collapse for a CO₂-dominated atmosphere, assuming that GJ 3473 b is tidally locked in a 1:1 spin-orbit resonance. We estimate that a surface pressure of approximately 1 mbar is required to prevent atmospheric collapse on GJ 3473 b. However, this limit does not allow us to rule out a thin CO₂ atmosphere on GJ 3473 b. On the other hand, for exoplanets receiving lower stellar irradiation, the critical pressure would be higher and could therefore serve as a useful prior. In cases where the critical pressure is higher than the limit placed by observations, this could allow us to eliminating CO₂-rich scenarios altogether. This approach may be especially valuable for planets where CO₂ is the only atmospheric constituent expected to survive long-term atmospheric escape (X. Ji et al. 2025), offering a path to rule out the presence of thin atmospheres.

We also find some evidence, albeit weak, of eclipse depth variability across the four visits, with a log Bayes factor of $\ln B = 2.1$. Although not conclusive with the present data, this result raises the possibility of temporal variability in the dayside emission, potentially linked to dynamic atmospheric processes, volcanism, or circumstellar/circumplanetary material (as proposed for 55 Cnc e; B.-O. Demory et al. 2015). However, it is also possible that instrumental systematics or stellar activity (e.g., in the form of unresolved flares) could be contributing to this apparent variability. If the variability is not

a statistical outlier, it could bias the interpretations of similar data unless several epochs are observed, regardless of the variability’s origin. Additional observations will be essential to establish the nature and origin of this variability, if it is real.

The presence or absence of an atmosphere on GJ 3473 b remains an open question, as the measured eclipse depth is consistent with both a bare-rock surface and a secondary atmosphere. The resulting degeneracy between surface and atmospheric emission illustrates the limitations of interpreting MIRI F1500W eclipse photometry in isolation, even with a significant thermal emission detection. Our analysis demonstrates the importance of explicitly accounting for these degeneracies when interpreting current and future MIRI F1500W observations of rocky exoplanets. For GJ 3473 b, additional observations, such as secondary eclipse spectroscopy or a thermal phase curve, will be required to break these degeneracies and place more robust constraints on the planet’s atmospheric or surface composition. While our work highlights the difficulties of detecting atmospheres on highly irradiated rocky exoplanets using single-band photometry alone, it establishes a practical framework for interpreting such data and for guiding future follow-up observations.

Acknowledgments

This work is based on observations with the NASA/ESA/CSA James Webb Space Telescope obtained at the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Incorporated, under NASA contract NAS5-03127. Support for program No. 3730 was provided through a grant from the STScI under NASA contract NAS5-03127. J.M.M. acknowledges support from the Horizon Europe Guarantee Fund, grant EP/Z00330X/1. N.H.A. acknowledges support by the National Science Foundation Graduate Research Fellowship under grant No. DGE1746891. P.C.A. acknowledges support from the Carlsberg Foundation, grant CF22-1254. B.-O. D. acknowledges support from the Swiss State Secretariat for Education, Research and Innovation (SERI) under contract No. MB22.00046. C.E.F. acknowledges support from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation program under grant agreement No. 805445. N.P.G. gratefully acknowledges support from Science Foundation Ireland and the Royal Society through a University Research Fellowship (URF\R\201032). B.P. acknowledges support from the Walter Gyllenberg Foundation. M.H. thanks Guadalupe Tovar Mendoza for useful discussions regarding stellar flares. We acknowledge the input from the following individuals to the GO 3730 proposal: Kathryn Jones, Jens Hoeijmakers, Anna Lueber, Alexander Rathcke, Can Akin, Matthew Hooton, Andrea Guzmán Mesa, Nicholas Borsato, Jens Hoeijmakers, Meng Tian, Mette Baungaard.

The JWST data presented in this article were obtained from the Mikulski Archive for Space Telescopes (MAST) at the Space Telescope Science Institute. The specific observations analyzed can be accessed via DOI: [10.17909/n43h-6185](https://doi.org/10.17909/n43h-6185).

Facility: JWST (MIRI).

Appendix A TESS Light-curve Fitting

We improve the orbital parameters of GJ 3473 b by utilizing data from TESS sectors 34, 61, and 88. First, we disregard the

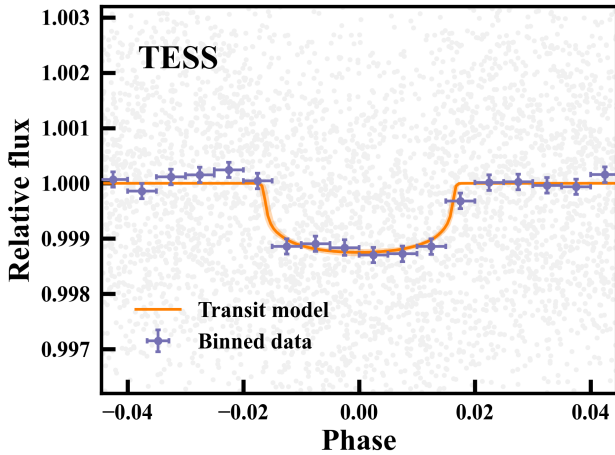


Figure 9. Phase-folded TESS data and transit model of GJ 3473 b. Prior to phase folding, we subtracted the median GP model. The purple data points are binned for visual clarity. The shaded orange region corresponds to the 1σ contour.

Table 3
Parameters Estimated from the Light-curve Analysis of TESS

Parameter	Prior	Value
P (days)	$N(1.1980035, 1.9 \times 10^{-6})$	$1.19800484^{+5.3 \times 10^{-7}}_{-5.2 \times 10^{-7}}$
T_0 (BJD)	$N(58491.70408, 0.00043)$	$58491.70415^{+0.00039}_{-0.00042}$
a/R_*	$N(9.39, 0.21)$	$9.32^{+0.18}_{-0.19}$
b	$N(0.336, 0.074)$	$0.300^{+0.055}_{-0.060}$
i (deg)	...	$88.16^{+0.38}_{-0.36}$
R_p/R_*	$N(0.03184, 0.00069)$	$0.03220^{+0.00055}_{-0.00055}$
q_1	$U(0, 1)$	$0.59^{+0.27}_{-0.28}$
q_2	$U(0, 1)$	$0.23^{+0.28}_{-0.16}$
σ_{GP} (ppm)	$\log U(1, 10^4)$	247^{+32}_{-24} ppm
ρ_{GP} (days)	$\log U(0.01, 10.0)$	$0.82^{+0.17}_{-0.13}$

Note. The period P , midtransit time T_0 (BJD -2400000.5 days), and GP timescale ρ_{GP} are given in days. The priors on P , T_0 , a/R_* , b , and R_p/R_* are taken from the constraints by J. Kemmer et al. (2020).

first 2 hours from each sector and remove outliers deviating more than 4σ from a median-filtered version of the light curves. Next, we employ the `juliet` framework (N. Espinoza et al. 2019) to jointly fit the transit light curves. For this, we use priors from J. Kemmer et al. (2020) for the period P , midtransit time, R_p/R_* , normalized semimajor axis a/R_* , and impact parameter b . This is why we did not include TESS sector 7 in the fit, as that data was used to derive the priors. We assumed a circular orbit of the planet, in line with J. Kemmer et al. (2020). This assumption is also consistent with the present MIRI observations, as discussed in Section 2. Additionally, we employed Gaussian processes, using a Matern-3/2 kernel, to account for stellar variability. The refined parameters are presented in Table 3, and Figure 9 illustrates the phase-folded TESS data and transit model.

Appendix B Aperture Radius Selection

An important consideration when performing photometry is to appropriately select an aperture size. If we assume that each pixel is independent and that the noise follows a normal

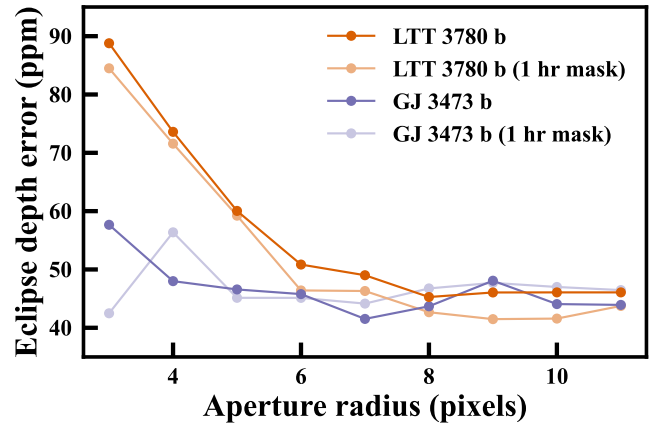


Figure 10. Eclipse depth uncertainty for a range of aperture radii. The orange points correspond to the data of LTT 3780 b, while the purple points correspond to GJ 3473 b. The lighter-colored points represent uncertainties where we masked the first hour of each observation to minimize detector settling effects.

distribution, we can effectively address this issue by instead using optimal extraction (K. Horne 1986). However, if these assumptions are not valid, it becomes challenging to determine how to properly weigh each pixel without a realistic model of the system. Therefore, we consider aperture extraction and explore a range of different aperture radii.

Two competing factors influence the determination of the optimal aperture size. On the one hand, if the aperture is too small we might get inaccurate results due to PSF jitter or other effects that could cause neighboring pixels to become related, such as wave front distortions or charge migration. For this reason, we consider an aperture larger than approximately 5–6 pixels, which encompasses the Airy disk, to minimize these systematic errors. Conversely, an aperture that is too large would introduce unnecessary read noise. To identify the optimal size, we investigate the uncertainty in eclipse depth for a variety of aperture sizes, as shown in Figure 10. In this analysis, we not only include the target of this study but also two additional secondary eclipse observations of LTT 3780 b (N. H. Allen et al. 2025). LTT 3780 b is the brightest target in our program using the SUB256 subarray and has 22 groups per integration, while GJ 3473 b is the dimmest target in the same subarray, with 39 groups per integration. We observe that the uncertainty in eclipse depth is larger for smaller apertures, particularly for LTT 3780 b, and stabilizes after around 6–8 pixels. This trend remains consistent even when we mask the first hour of the observations, suggesting that the underlying mechanism does not strongly depend on the effects of detector settling. Consequently, we select an aperture size of 6 pixels for our analysis.

For completeness, we present the eclipse depths as a function of aperture radius in Figure 11. We find that the inferred eclipse depths are largely consistent across different apertures, given the measurement uncertainty. The standard deviation among these inferences is 24 ppm, which, when added in quadrature, increases the eclipse depth uncertainty of the canonical case from 45 to 51 ppm. However, this argument only works if we consider these choices equally valid which, as discussed above, is not the case. Additionally, a change of just 6 ppm is likely within the uncertainty of the uncertainty itself, given the range of 34–51 ppm from different model

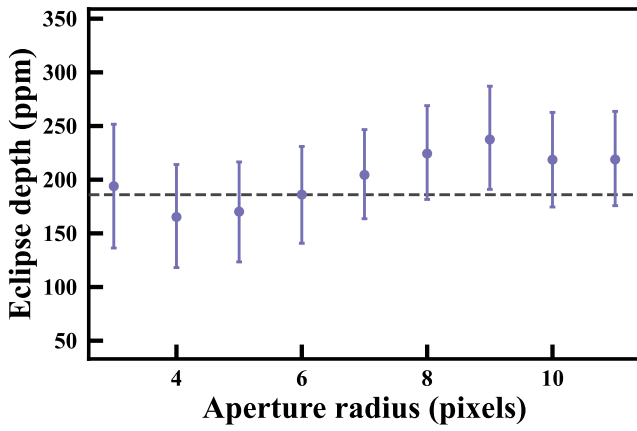


Figure 11. Eclipse depth of GJ 3473 b as a function of aperture size. The dashed horizontal line indicates the eclipse depth for an aperture radius of 6 pixels, which is the standard size used in our analysis.

assumptions that we cannot rule out with high confidence, as shown in Table 1.

Appendix C A Potential Flare

We observe a temporary increase in flux just before the secondary eclipse of the second visit, as seen in Figure 3. The peak flux of this flare-like feature is around 700 ppm, which exceeds the depth of the secondary eclipse. Therefore, we consider it unlikely that this increase is of planetary origin. Another possibility is that it could be caused by instrument systematics. However, we did not detect any changes in the point-spread function (PSF), either in its position or width, that correlate with the temporary increase in flux. Additionally, the background flux remained stable for the entire duration of the observation. We also visually inspected each image for signs of asteroids, which could potentially contaminate the extracted stellar flux, but found no evidence of this. Another potential cause could be a persistence signal resulting from a large cosmic-ray hit, as noted by M. Fortune et al. (2025). In fact, we do detect a persistence signal in a pixel at the edge of the aperture, occurring toward the end of the second observation, as described in Section 2. However, the timing does not align, and we find that the overall contribution of this effect is negligible. Therefore, we conclude that the signal is likely astrophysical, possibly resulting from a stellar flare (similar events are reported by M. Gillon et al. 2025).

In this section, we examine various methods for addressing this flare-like feature, either by incorporating a flare model or by masking the affected integrations. For the first approach, we adopt the analytic single-flare model proposed by G. Tovar Mendoza et al. (2022), which we add to the overall eclipse model following Equation (1). This flare model consists of the convolution of a Gaussian function (representing heating) with one or more exponential functions (representing cooling, with different timescales). Here, we consider a single timescale for the cooling rate. The model is given by

$$f_{\text{flare}}(t) = \int_{-\infty}^t g(\tau)h(t-\tau)d\tau = \frac{AD}{2}e^{D(B-t)+\left(\frac{CD}{2}\right)^2} \operatorname{erfc}\left(\frac{B-t}{C} + \frac{DC}{2}\right), \quad (\text{C1})$$

Table 4

The Eclipse Depth of GJ 3473 b Under Different Treatments of the Flare-like Feature Observed During the Second Visit (See Figure 3)

Case	$\ln B$	Eclipse Depth
Masked region (canonical case)	...	186_{-45}^{+45} ppm
Gaussian model	0	192_{-48}^{+49} ppm
G. Tovar Mendoza et al. (2022) model	-0.3	195_{-45}^{+46} ppm
No flare model	-2.8	235_{-64}^{+62} ppm

Note. The affected integrations are included in all cases except the canonical one, where they are masked. As a result, Bayes factor comparisons between the canonical and other cases are not meaningful.

with

$$g(t) = \frac{A}{\sqrt{\pi}C}e^{-\frac{(t-B)^2}{C^2}}, h(t) = D e^{-Dt}. \quad (\text{C2})$$

Expressed in this form, $A = \int f_{\text{flare}} dt$ represents the integral of the relative flux excess, also called the equivalent duration. Parameters B and C correspond to the midpoint and width of the Gaussian heating impulse, respectively, while D denotes the inverse timescale of the exponential cooling ($C, D > 0$). Note that we identify a minor inconsistency in the previous reporting of this expression, leading to Equation C1 having a somewhat different normalization from what is presented in G. Tovar Mendoza et al. (2022).

We report the eclipse depth obtained by jointly fitting the light curves and the above flare model in Table 4. For this, we do not mask the affected integrations, unlike in the canonical case. For the priors on C and D , we adopt uniform priors between 1.4-84.9 min (corresponding to a standard deviation of 1-60 min) and between 1 min^{-1} and $1/12 \text{ hr}^{-1}$, respectively. Given the apparent symmetric shape of the flare-like feature, we also fit a Gaussian function alone, without including exponential cooling. We find that the resulting difference in evidence between these two approaches is negligible and that the eclipse depths are consistent with the canonical case. We obtain consistent values for the equivalent duration A of $1.17_{-0.31}^{+0.39}$ s and $1.16_{-0.31}^{+0.41}$ s for the G. Tovar Mendoza et al. (2022) model and Gaussian model, respectively. Additionally, we estimate the full width at half maximum to be 23_{-7}^{+9} min and 26_{-6}^{+10} min for the two scenarios, respectively.

In contrast, if we opt not to fit a flare model while still including the affected integrations, we get lower evidence and a somewhat higher eclipse depth, as detailed in Table 4. In this scenario, the uncertainty in the eclipse depth increases due to larger residuals, which necessitates an increase in the amplitude of the Gaussian process (GP) kernel. Specifically, the evidence for the inclusion of a GP kernel is $|\ln B| = 15.5$ when we do not include a flare model and do not mask the affected integrations. This is significantly higher than what we obtain when masking the affected integrations ($|\ln B| = 1.8$), as outlined in Table 1. Overall, we find that the eclipse depth is not sensitive to the treatment of the flare-like feature, whether we model the excess flux or mask the affected integrations. For this reason, we adopt the latter approach in this work, which involves fewer free parameters.

Using the absolute-calibrated flux described in Section 2, and assuming that the excess flux is indeed caused by a flare, we determine that it had an energy of $3.24_{-0.86}^{+1.08} \times 10^{28}$ erg in

the F1500W bandpass. We calculated this using (e.g., R. O. P. Loyd et al. 2018)

$$E_{\text{flare,F1500W}} = 4\pi d^2 F_* \Delta\nu A, \quad (\text{C3})$$

where $d = 27.315 \pm 0.018$ pc is the distance (Gaia Collaboration et al. 2016, 2023), $F_* = 7.976 \pm 0.038$ mJy is the stellar flux density during the second visit, and $\Delta\nu = 3.89 \times 10^{12}$ Hz is the effective bandwidth of the total MIRI F1500W throughput.

Appendix D List of Parameters

Here we provide a list of all priors and constraints for the canonical case described in Section 2.3. These correspond to 20 free parameters, which are detailed in Table 5. The priors for the period P , midtransit time T_0 , normalized semimajor axis a/R_* , and inclination i are derived from the analysis of the TESS data in Appendix A. Similarly, we fix the planet-to-star radius ratio R_p/R_* to the value obtained from TESS. The stellar radius R_* , which is used to compute the light travel time, is obtained from J. Kemmer et al. (2020). The remaining fitted parameters are used to describe the stellar flux, the systematic model (as described in Equation (2)), and the noise characteristics.

The prior for the settling timescale, p_2 , follows a log-uniform distribution ranging from 2 minutes (equivalent to 10 integrations) to 6 hr (roughly twice the observation time per visit). For the noise model (i.e., Gaussian process), we use the

Table 5
Parameters Estimates and Priors for the MIRI Light-curve Analysis

Parameter	Prior	Value
P (days)	$N(1.19800484, 5.3 \times 10^{-7})$	$1.19800453^{+5.3 \times 10^{-7}}_{-4.6 \times 10^{-7}}$
T_0 (BJD)	$N(58491.70415, 0.00042)$	$58491.70404^{+0.00040}_{-0.00040}$
a/R_*	$N(9.32, 0.19)$	$9.23^{+0.18}_{-0.18}$
i (deg)	$N(88.16, 0.38)$	$88.32^{+0.35}_{-0.36}$
F_p/F_* (ppm)	$U(0, 500)$	186^{+45}_{-45}
R_p/R_*	fixed	0.03220
R_* (R_\odot)	fixed	0.364
e	fixed	0
ω (deg)	fixed	90
$F_*^{(1)}$ (DN/s)	$U(0.98\bar{F}, 1.02\bar{F})$	$51766.0^{+7.8}_{-9.0}$
$F_*^{(2)}$ (DN/s)	$U(0.98\bar{F}, 1.02\bar{F})$	$51769.0^{+8.0}_{-8.7}$
$F_*^{(3)}$ (DN/s)	$U(0.98\bar{F}, 1.02\bar{F})$	$51742.1^{+7.6}_{-9.2}$
$F_*^{(4)}$ (DN/s)	$U(0.98\bar{F}, 1.02\bar{F})$	$51674.9^{+8.4}_{-8.9}$
$p_1^{(1)}$	$U(-0.02, 0.02)$	$0.00195^{+0.00022}_{-0.00022}$
$p_1^{(2)}$	$U(-0.02, 0.02)$	$0.00145^{+0.00022}_{-0.00024}$
$p_1^{(3)}$	$U(-0.02, 0.02)$	$0.00190^{+0.00023}_{-0.00022}$
$p_1^{(4)}$	$U(-0.02, 0.02)$	$-0.00294^{+0.00024}_{-0.00024}$
$\log(p_2/\text{days})$	$U(-2.86, -0.60)$	$-1.779^{+0.047}_{-0.042}$
$\log(\alpha)$	$U(-4, 0.3)$	$-1.38^{+0.03}_{-0.59}$
$\log(\beta/\text{days})$	$U(-2.86, 1.0)$	$-0.96^{+1.13}_{-0.85}$
$\log(\gamma^{(1)})$	$U(-4, 0.3)$	$-0.221^{+0.034}_{-0.040}$
$\log(\gamma^{(2)})$	$U(-4, 0.3)$	$-0.199^{+0.037}_{-0.041}$
$\log(\gamma^{(3)})$	$U(-4, 0.3)$	$-0.180^{+0.031}_{-0.035}$
$\log(\gamma^{(4)})$	$U(-4, 0.3)$	$-0.226^{+0.036}_{-0.040}$

Note. These parameters correspond to our canonical case, described in Section 2.3. Here, \bar{F} represents the median flux (in DN/s) of each visit, used to construct the prior of the stellar flux parameters.

following parameterization of the covariance matrix

$$\Sigma_{i,j} = \sigma_i^2 \delta_{i,j} + \bar{\sigma}^2 (\alpha^2 e^{-|t_i - t_j|/\beta} + \gamma^2 \delta_{i,j}). \quad (\text{D1})$$

Here, σ_i represents the predicted uncertainty, derived from the propagation of photon and read noise; $\bar{\sigma}$ is the median of σ_i ; α and β denote the amplitude and timescale of the time-correlated noise (exponential kernel), respectively; and γ indicates the strength of additional white noise (jitter term). We set the prior of β to be log-uniform between 2 minutes and 1 day. Although this upper limit exceeds the observation duration, we still consider such long timescales due to the constraints from TESS, which indicate time-correlated noise with a timescale of about 1 day, as presented in Table 3.

ORCID iDs

Måns Holmberg  <https://orcid.org/0000-0002-0931-735X>
 Hannah Diamond-Lowe  <https://orcid.org/0000-0001-8274-6639>
 João M. Mendonça  <https://orcid.org/0000-0002-6907-4476>
 Daniel Kitzmann  <https://orcid.org/0000-0003-4269-3311>
 Néstor Espinoza  <https://orcid.org/0000-0001-9513-1449>
 Natalie H. Allen  <https://orcid.org/0000-0002-0832-710X>
 Prune C. August  <https://orcid.org/0000-0003-3829-8554>
 Mark Fortune  <https://orcid.org/0000-0002-8938-9715>
 Amélie Gressier  <https://orcid.org/0000-0003-0854-3002>
 Jegug Ih  <https://orcid.org/0000-0003-2775-653X>
 Erik Meier Valdés  <https://orcid.org/0000-0002-2160-8782>
 Merlin Zraggen  <https://orcid.org/0009-0001-6868-6171>
 Lars A. Buchhave  <https://orcid.org/0000-0003-1605-5666>
 Brice-Olivier Demory  <https://orcid.org/0000-0002-9355-5165>
 Chloe Fisher  <https://orcid.org/0000-0003-0652-2902>
 Neale P. Gibson  <https://orcid.org/0000-0002-9308-2353>
 Kevin Heng  <https://orcid.org/0000-0003-1907-5910>
 Bibiana Prinoth  <https://orcid.org/0000-0001-7216-4846>
 Adam J. Burgasser  <https://orcid.org/0000-0002-6523-9536>

References

- Alam, M. K., Gao, P., Adams Redai, J., et al. 2025, *AJ*, 169, 15
 Alderson, L., Batalha, N. E., Wakeford, H. R., et al. 2024, *AJ*, 167, 216
 Alderson, L., Moran, S. E., Wallack, N. L., et al. 2025, *AJ*, 169, 142
 Allard, F. 2014, *IAUS*, 299, 271
 Allen, N. H., Espinoza, N., Diamond-Lowe, H., et al. 2025, *AJ*, 170, 240
 Auclair-Desrotour, P., & Heng, K. 2020, *A&A*, 638, A77
 August, P. C., Buchhave, L. A., Diamond-Lowe, H., et al. 2025, *A&A*, 695, A171
 Banerjee, A., Barstow, J. K., Gressier, A., et al. 2024, *ApJL*, 975, L11
 Baranov, Y. I., Lafferty, W. J., & Fraser, G. T. 2004, *JMoSp*, 228, 432
 Barber, R. J., Tennyson, J., Harris, G. J., & Tolchenov, R. N. 2006, *MNRAS*, 368, 1087
 Barnes, R. 2017, *CeMDA*, 129, 509
 Bell, T. J., Crouzet, N., Cubillos, P. E., et al. 2024, *NatAs*, 8, 879
 Bello-Arufe, A., Damiano, M., Bennett, K. A., et al. 2025, *ApJL*, 980, L26
 Bouwman, J., Kendrew, S., Greene, T. P., et al. 2023, *PASP*, 135, 038002
 Bushouse, H. 2020, *ASP*, 527, 583
 Cassidy, W., & Hapke, B. 1975, *Icar*, 25, 371
 Cheek, L. C., Pieters, C. M., Dyar, M. D., & Milam, K. A. 2009, Revisiting Plagioclase Optical Properties for Lunar Exploration, in 40th Annual Lunar and Planetary Science Conference, Lunar and Planetary Science Conference, 1928
 Cox, A. N. 2000, *Allen's Astrophysical Quantities* (AIP Press)
 Demory, B.-O., Gillon, M., Madhusudhan, N., & Queloz, D. 2015, *MNRAS*, 455, 2018
 Dicken, D., Marín, M. G., Shivaee, I., et al. 2024, *A&A*, 689, A5
 Ducrot, E., Lagage, P.-O., Min, M., et al. 2025, *NatAs*, 9, 358

- Espinoza, N., Kossakowski, D., & Brahm, R. 2019, *MNRAS*, **490**, 2262
- Feroz, F., Hobson, M. P., & Bridges, M. 2009, *MNRAS*, **398**, 1601
- Foreman-Mackey, D., Agol, E., Ambikasaran, S., & Angus, R. 2017, *AJ*, **154**, 220
- Fortune, M., Gibson, N. P., Diamond-Lowe, H., et al. 2025, *A&A*, **701**, 25
- Gaia Collaboration, Prusti, T., de Bruijne, J., H., J., et al. 2016, *A&A*, **595**, A1
- Gaia Collaboration, Vallenari, A., Brown, A. G. A., et al. 2023, *A&A*, **674**, A1
- Gao, P., Piette, A. A. A., Steinrueck, M. E., et al. 2023, *ApJ*, **951**, 96
- Gardner, J. P., Mather, J. C., Clampin, M., et al. 2006, *SSRv*, **123**, 485
- Gillon, M., Ducrot, E., Bell, T. J., et al. 2025, arXiv:2509.02128
- Ginzburg, S., Schlichting, H. E., & Sari, R. 2018, *MNRAS*, **476**, 759
- Gordon, K. D., Sloan, G. C., Garcia Marin, M., et al. 2024, *AJ*, **169**, 6
- Greene, T. P., Bell, T. J., Ducrot, E., et al. 2023, *Natur*, **618**, 39
- Gressier, A., Espinoza, N., Allen, N. H., et al. 2024, *ApJL*, **975**, L10
- Grimm, S. L., & Heng, K. 2015, *ApJ*, **808**, 182
- Grimm, S. L., Malik, M., Kitzmann, D., et al. 2021, *ApJS*, **253**, 30
- Gruszka, M., & Borysow, A. 1997, *Icar*, **129**, 172
- Hammond, M., Guimond, C. M., Lichtenberg, T., et al. 2025, *ApJL*, **978**, L40
- Hapke, B. 2001, *JGR*, **106**, 10039
- Heng, K. 2023, *ApJL*, **956**, L20
- Holmberg, M., & Madhusudhan, N. 2023, *MNRAS*, **524**, 377
- Horne, K. 1986, *PASP*, **98**, 609
- Hu, R., Bello-Arufe, A., Zhang, M., et al. 2024, *Natur*, **630**, 609
- Hu, R., Ehlmann, B. L., & Seager, S. 2012, *ApJ*, **752**, 7
- Ih, J., Kempton, E. M. R., Whittaker, E. A., & Lessard, M. 2023, *ApJL*, **952**, L4
- Iyer, A. R., Line, M. R., Muirhead, P. S., Fortney, J. J., & Gharib-Nezhad, E. 2023, *ApJ*, **944**, 41
- Ji, X., Chatterjee, R. D., Park Coy, B., & Kite, E. S. 2025, *ApJ*, **992**, 198
- Kasting, J. F., Whitmire, D. P., & Reynolds, R. T. 1993, *Icar*, **101**, 108
- Kemmer, J., Stock, S., Kossakowski, D., et al. 2020, *A&A*, **642**, A236
- Koll, D. D. B. 2022, *ApJ*, **924**, 134
- Koll, D. D. B., Malik, M., Mansfield, M., et al. 2019, *ApJ*, **886**, 140
- Kreidberg, L. 2015, *PASP*, **127**, 1161
- Lagerros, J. S. V. 1998, *A&A*, **332**, 1123
- Leconte, J., Wu, H., Menou, K., & Murray, N. 2015, *Sci*, **347**, 632
- Lim, O., Benneke, B., Doyon, R., et al. 2023, *ApJL*, **955**, L22
- Lincowski, A. P., Meadows, V. S., Zieba, S., et al. 2023, *ApJL*, **955**, L7
- Lopez, E. D., & Fortney, J. J. 2013, *ApJ*, **776**, 2
- Lloyd, R. O. P., France, K., Youngblood, A., et al. 2018, *ApJ*, **867**, 71
- Luger, R., Barnes, R., Lopez, E., et al. 2015, *AsBio*, **15**, 57
- Lustig-Yaeger, J., Fu, G., May, E. M., et al. 2023, *NatAs*, **7**, 1317
- Lyu, X., Koll, D. D. B., Cowan, N. B., et al. 2024, *ApJ*, **964**, 152
- Madhusudhan, N., Constantinou, S., Holmberg, M., et al. 2025, *ApJL*, **983**, L40
- Malik, M., Grosheintz, L., Mendonça, J. M., et al. 2017, *AJ*, **153**, 56
- Malik, M., Kempton, E. M. R., Koll, D. D. B., et al. 2019a, *ApJ*, **886**, 142
- Malik, M., Kitzmann, D., Mendonça, J. M., et al. 2019b, *AJ*, **157**, 170
- Mansfield, M., Kite, E. S., Hu, R., et al. 2019, *ApJ*, **886**, 141
- May, E. M., MacDonald, R. J., Bennett, K. A., et al. 2023, *ApJL*, **959**, L9
- Meier Valdés, E. A., Demory, B. O., Diamond-Lowe, H., et al. 2025, *A&A*, **698**, A68
- Mlawer, E. J., Cady-Pereira, K. E., Mascio, J., & Gordon, I. E. 2023, *JQSRT*, **306**, 108645
- Moran, S. E., Stevenson, K. B., Sing, D. K., et al. 2023, *ApJL*, **948**, L11
- Owen, J. E., & Wu, Y. 2017, *ApJ*, **847**, 29
- Paragas, K., Knutson, H. A., Hu, R., et al. 2025, *ApJ*, **981**, 130
- Pass, E. K., Charbonneau, D., & Vanderburg, A. 2025, *ApJL*, **986**, L3
- Patel, J. A., Brandeker, A., Kitzmann, D., et al. 2024, *A&A*, **690**, A159
- Pieters, C. M., & Noble, S. K. 2016, *JGRE*, **121**, 1865
- Pieters, C. M., Taylor, L. A., Noble, S. K., et al. 2000, *M&PS*, **35**, 1101
- Piette, A. A. A., Madhusudhan, N., & Mandell, A. M. 2022, *MNRAS*, **511**, 2565
- Polyanskiy, M. N. 2024, *Scientific Data*, **11**, 94
- Pontoppidan, K. M., Pickering, T. E., Laidler, V. G., et al. 2016, *SPIE*, **9910**, 991016
- Rothman, L. S., Gordon, I. E., Barber, R. J., et al. 2010, *JQSRT*, **111**, 2139
- Skilling, J. 2004, *AIPC*, **735**, 395
- Scarsdale, N., Wogan, N., Wakeford, H. R., et al. 2024, *AJ*, **168**, 276
- Sneep, M., & Ubachs, W. 2005, *JQSRT*, **92**, 293
- Spencer, J. R. 1990, *Icar*, **83**, 27
- Syal, M. B., Schultz, P. H., & Riner, M. A. 2015, *NatGe*, **8**, 352
- Tamburo, P., Mandell, A., Deming, D., & Garhart, E. 2018, *AJ*, **155**, 221
- Thalman, R., Zarzana, K. J., Tolbert, M. A., & Volkamer, R. 2014, *JQSRT*, **147**, 171
- Tovar Mendoza, G., Davenport, J. R. A., Agol, E., Jackman, J. A. G., & Hawley, S. L. 2022, *AJ*, **164**, 17
- Trotta, R. 2008, *ConPh*, **49**, 71
- Turbet, M., Bolmont, E., Leconte, J., et al. 2018, *A&A*, **612**, A86
- Tusay, N., Wright, J. T., Beatty, T. G., et al. 2025, *ApJ*, **987**, 6
- Van Looveren, G., Boro Saikia, S., Herbort, O., et al. 2025, *A&A*, **694**, A310
- Wachiraphan, P., Berta-Thompson, Z. K., Diamond-Lowe, H., et al. 2025, *AJ*, **169**, 311
- Weiner Mansfield, M., Xue, Q., Zhang, M., et al. 2024, *ApJL*, **975**, L22
- Winn, J. N. 2010, in *Exoplanets*, ed. S. Seager (Univ. Arizona Press)
- Wordsworth, R. 2015, *ApJ*, **806**, 180
- Wright, G. S., Rieke, G. H., Glasse, A., et al. 2023, *PASP*, **135**, 048003
- Xue, Q., Bean, J. L., Zhang, M., et al. 2024, *ApJL*, **973**, L8
- Zahnle, K. J., & Catling, D. C. 2017, *ApJ*, **843**, 122
- Zhang, M., Hu, R., Inglis, J., et al. 2024, *ApJL*, **961**, L44
- Zieba, S., Kreidberg, L., Ducrot, E., et al. 2023, *Natur*, **620**, 746