The TRAPPIST-1 Habitable Atmosphere Intercomparison (THAI). III. Simulated Observables-the Return of the Spectrum

Thomas J. Fauchez, Geronimo L. Villanueva, Denis E. Sergeev, Martin Turbet, Ian A. Boutle, Kostas Tsigaridis, Michael J. Way, Eric T. Wolf, Shawn D. Domagal-Goldman, François Forget, et al.

The Planetary Science Journal, 3:213 (17pp), 2022 September https://doi.org/10.3847/PSJ/ac6cf1

The TRAPPIST-1 Habitable Atmosphere Intercomparison (THAI). III. Simulated Observables—the Return of the Spectrum

Thomas J. Fauchez[1,2,3,4], Geronimo L. Villanueva[1], Denis E. Sergeev[5], Martin Turbet[6,7], Ian A. Boutle[8,9], Kostas Tsigaridis[10,11], Michael J. Way[4,11,12], Eric T. Wolf[4,13,14], Shawn D. Domagal-Goldman[1,4,14], François Forget[7], Jacob Haqq-Misra[14,15], Ravi K. Kopparapu[1,4,14], James Manners[8], and Nathan J. Mayne[9]
[1] NASA Goddard Space Flight Center, 8800 Greenbelt Road, Greenbelt, MD 20771, USA; thomas.j.fauchez@nasa.gov
[2] Goddard Earth Sciences Technology and Research (GESTAR), Universities Space Research Association (USRA), Columbia, MD 21046, USA
[3] American University, College of Arts and Sciences, Washington, DC 20016, USA
[4] NASA GSFC Sellers Exoplanet Environments Collaboration, USA
[5] Department of Mathematics, College of Engineering, Mathematics, and Physical Sciences, University of Exeter Exeter, EX4 4QF, UK
[6] Observatoire Astronomique de l'Université de Genève, Université de Genève, Chemin des Maillettes 51, 1290 Versoix, Switzerland
[7] Laboratoire de Météorologie Dynamique/IPSL, CNRS, Sorbonne Université, École Normale Supérieure, PSL Research University, École Polytechnique, F-75005 Paris, France
[8] Met Office, FitzRoy Road, Exeter, EX1 3PB, UK
[9] Department of Astrophysics, College of Engineering, Mathematics, and Physical Sciences, University of Exeter, Exeter, EX4 4QL, UK
[10] Center for Climate Systems Research, Columbia University, New York, NY 10025, USA
[11] NASA Goddard Institute for Space Studies, 2880 Broadway, New York, NY 10025, USA
[12] Theoretical Astrophysics, Department of Physics and Astronomy, Uppsala University, Uppsala, Sweden
[13] Laboratory for Atmospheric and Space Physics, University of Colorado Boulder, Boulder, CO 80303, USA
[14] NASA NExSS Virtual Planetary Laboratory, Seattle, WA, 98195, USA
[15] Blue Marble Space Institute of Science, Seattle, WA 98104, USA
Received 2021 September 24; revised 2022 April 7; accepted 2022 May 3; published 2022 September 15

Abstract

The TRAPPIST-1 Habitable Atmosphere Intercomparison (THAI) is a community project that aims to quantify how differences in general circulation models (GCMs) could impact the climate prediction for TRAPPIST-1e and, subsequently, its atmospheric characterization in transit. Four GCMs have participated in THAI: ExoCAM, LMD-Generic, ROCKE-3D, and the UM. This paper, focused on the simulated observations, is the third part of a trilogy, following the analysis of two land planet scenarios (Part I) and two aquaplanet scenarios (Part II). Here we show a robust agreement between the simulated spectra and the number of transits estimated to detect the land planet atmospheres. For the cloudy aquaplanet ones, a 5σ detection of CO_2 could be achieved in about 10 transits if the atmosphere contains at least 1 bar of CO_2. That number can vary by 41%–56% depending on the GCM used to predict the terminator profiles, principally due to differences in the cloud deck altitude, with ExoCAM and LMD-G producing higher clouds than ROCKE-3D and UM. Therefore, for the first time, this work provides "GCM uncertainty error bars" of ∼50% that need to be considered in future analyses of transmission spectra. We also analyzed the intertransit spectral variability. Its magnitude differs significantly between the GCMs, but its impact on the transmission spectra is within the measurement uncertainties. THAI has demonstrated the importance of model intercomparison for exoplanets and also paved the way for a larger project to develop an intercomparison meta-framework, namely, the Climates Using Interactive Suites of Intercomparisons Nested for Exoplanet Studies.

Unified Astronomy Thesaurus concepts: Exoplanet atmospheres (487); Exoplanet atmospheric composition (2021); Exoplanet atmospheric variability (2020); Transmission spectroscopy (2133); Planetary climates (2184)

1. Introduction

> "Here at last...comes the end of our fellowship. I will not say do not weep, for not all tears are an evil." —J.R.R. Tolkien, *The Return of the King* (1955).

At the dawn of terrestrial exoplanet atmospheric characterization with the James Webb Space Telescope (JWST), predicting the detectability of the atmospheres of such planets is crucial in order to prepare for observations and maximize the scientific return. JWST Guaranteed Time Observations (GTOs) and Cycle 1 proposals have already been selected. While CO_2 can be potentially detectable from Cycle 1, it is unlikely that enough transits would be accumulated for any single target to characterize in depth the atmosphere of a terrestrial exoplanet (Fauchez et al. 2019; Lustig-Yaeger et al. 2019; Pidhorodetska et al. 2020) in the habitable zone (HZ; see, e.g., Kopparapu et al. 2013) of M-dwarf stars. The presence of CO_2 has been shown to be the best proxy (Fauchez et al. 2019; Lustig-Yaeger et al. 2019; Turbet et al. 2020) for the detection of a potentially habitable atmosphere owing to its strong absorption band in the mid-infrared (MIR) at 4.3 μm and in the far-infrared at ∼15 μm. However, a 5σ detection under cloudy conditions would likely require more than a dozen transits, even for the most favorable HZ terrestrial planet, TRAPPIST-1e (Fauchez et al. 2019).

TRAPPIST-1e belongs to the system of seven small transiting planets TRAPPIST-1 (Gillon et al. 2016, 2017; Luger et al. 2017) at 12.0 pc away. The star, TRAPPIST-1, is an M8V just slightly larger than Jupiter, which makes it very suitable for

transmission spectroscopy of the atmosphere of small rocky planets. Indeed, the ratio of the surface area of the star's disk blocked out by the planet's disk (including its atmosphere), i.e., the transit depth, is inversely proportional to the square of the star radius. Also, around such a cold and dim star, HZ planets have a very short orbital period, leading to very frequent transits and therefore more accessible data on their atmospheres. The Hubble Space Telescope (HST) has been used to infer the presence of an atmosphere on TRAPPIST-1e and its sibling planets (de Wit et al. 2016, 2018). However, the precision of HST data was only able either to be consistent with the absence of an atmosphere or to rule out clear-sky H_2-dominated atmospheres, while Moran et al. (2018) showed that HST observations could actually be fit by cloudy/hazy H_2 atmospheres. Yet the comparison of TRAPPIST-1e bulk density measurements (Grimm et al. 2018; Agol et al. 2021) to H_2-rich planets mass–radius relationships (Turbet et al. 2020), along with atmospheric escape modeling and gas accretion modeling (Hori & Ogihara 2020), provides accumulating evidence against the presence of H_2-dominated cloudy atmospheres around TRAPPIST-1 planets, including TRAPPIST-1e (see Turbet et al. 2020 and references therein). Furthermore, Krishnamurthy et al. (2021) reported strong upper limit constraints on the absence of helium in the atmosphere of TRAPPIST-1e. More in-depth knowledge about the absence or presence of a high mean molecular weight atmosphere on TRAPPIST-1e would most likely require JWST transit observations. Indeed, even some of the largest planned optical telescopes (Extremely Large Telescopes (ELTs)) are unable—even at the diffraction limit—to separate the light from TRAPPIST-1 and its planets because of their very small angular separation. The same is true for future space observatories such as Roman (Douglas et al. 2020), the Large UV/Optical/Infrared Surveyor (LUVOIR; Team 2019), and the Habitable Exoplanet Observatory (HabEx; Gaudi et al. 2018), for which the inner working angle of their coronagraph would block the light not only from the star but also from the entire system. Also, in the HZ, the planet is relatively too cold to significantly emit thermal infrared radiation, which makes it very challenging to characterize its emission spectrum (Fauchez et al. 2019; Lustig-Yaeger et al. 2019; Kane et al. 2021); more close-in planets, however, will be more sensitive to this technique (Morley et al. 2017; Koll et al. 2019; Lustig-Yaeger et al. 2019; Turbet et al. 2020). Orbital broadband phase-dependent variations in the combined planetary thermal emission and reflected stellar energy can also provide clues about the atmospheric structure and surface properties of the planet (e.g., Selsis et al. 2011; Koll & Abbot 2016; Turbet et al. 2016; von Paris et al. 2016; Haqq-Misra et al. 2018; Wolf et al. 2019), but the level of the phase variation of parts per billion (ppb) is far beyond our current instrumental capabilities (Wolf et al. 2019). Atmospheric characterization of TRAPPIST-1e may be also possible from the ground (Wunderlich et al. 2020) with the planned European Extremely Large Telescope (E-ELT), but this demonstration has been done assuming a clear-sky TRAPPIST-1e atmosphere, which may have led to an overestimation of the E-ELT's capabilities.

Therefore, JWST transit observations are the most viable atmospheric characterization technique for the TRAPPIST-1 planets in the coming decade. Several studies have used either general circulation models (GCMs; Fauchez et al. 2019; Pidhorodetska et al. 2020; May et al. 2021) or 1D radiative convective climate models coupled to photochemistry (Lincowski et al. 2018; Lustig-Yaeger et al. 2019; Wunderlich et al. 2019, 2020; Lin et al. 2021) or analytical models (Morley et al. 2017) to predict the detectability of standard atmospheres such as the modern Earth, the Archean Earth, or a CO_2-dominated atmosphere. While these predictions inherently vary from one model category to another, due to, for instance, the day/night contrast or the presence of clouds and hazes in the simulated atmosphere, models in the same category may also disagree owing to, for example, differences in the atmospheric profiles at the terminator. Evaluating these differences and their impact on synthetic observations in order to optimize JWST observation strategies are the core objectives of the TRAPPIST-1 Habitable Atmosphere Intercomparison (THAI; Fauchez et al. 2020a, 2021). The comparison of these simulated climate systems is described in the companion papers (Turbet et al. 2022, referred to as Part I) for the dry planet benchmark scenarios (Ben 1 and Ben 2) and in (Sergeev et al. 2022, referred to as Part II) for the aquaplanet (Hab 1 and Hab 2) scenarios. The paper is structured as follows: In Section 2, the methods and tools used in this study are described. In Section 3, we present the simulated transmission spectra using each of the GCM outputs, using both time-average and time-dependent terminator profiles. Finally, conclusions are given in Section 4.

2. Method

2.1. The THAI GCM Simulations

In this paper, Part III of a trilogy of THAI papers, we use the same GCM simulations that have been extensively analyzed in Part I (Turbet et al. 2022) and Part II (Sergeev et al. 2022), namely, the Ben 1 and Ben 2 and Hab 1 and Hab 2 cases, respectively. Briefly, the Ben 1 and Ben 2 cases are dry land planet simulations, while the Hab 1 and Hab 2 cases assume that the surface is fully covered by a global (static) ocean and that there are water vapor and clouds in the atmosphere. Ben 1 and Hab 1 atmospheric composition is broadly similar to that of modern Earth (1 bar of N_2, 400 ppmv of CO_2), while the Ben 2 and Hab 2 experiments assume a CO_2-dominated atmosphere (1 bar). A total of 10 orbits (61 Earth days) at a frequency of 6 hr are output for Ben 1 and Ben 2, while 100 orbits (610 Earth days) are output for Hab 1 and Hab 2, in order to smooth out the internal variability with a period of about a dozen orbits induced by clouds.

Each of these simulations has been performed by the four THAI GCMs: ExoCAM (Wolf et al. 2022), the exoplanet branch of the Community Earth System Model (CESM; http://www.cesm.ucar.edu/models/cesm1.2/) version 1.2.1; the LMD-Generic model (LMD-G; see, e.g., Wordsworth et al. 2011; Turbet et al. 2018); the Resolving Orbital and Climate Keys of Earth and Extraterrestrial Environments with Dynamics (ROCKE-3D; Way et al. 2017); and the Met Office Unified Model (UM; see, e.g., Mayne et al. 2014; Boutle et al. 2017). More details on these four GCMs are also provided in the THAI protocol (Sergeev et al. 2022; Turbet et al. 2022) and the THAI workshop report (Fauchez et al. 2021).

2.2. Simulated Spectra

We use the Planetary Spectrum Generator (PSG; Villanueva et al. 2018, 2022) to simulate transmission spectra of

TRAPPIST-1e for each of the THAI scenarios, using data from each of the four GCMs. PSG is an online radiative transfer tool that can be used to simulate planetary spectra observations from any ground- or space-based observatory for various objects of the solar system and beyond, and it also includes a noise calculator.

To simulate and compare time-averaged transmission spectra across the models, we first average the atmospheric properties over the 10 orbits of Ben 1 and Ben 2 and over the 100 orbits of Hab 1 and Hab 2. For our transit calculations, atmospheric profiles were created for each GCM latitude × longitude cell at the terminator of the planet using abundance, pressures, and temperatures as reported by the GCM for that specific terminator cell (vertical parameters of that cell). Then, transit spectra were computed using those profiles across all terminator cells, and the total transit spectrum was computed by the average of all transits across the terminators (weighted by the latitudinal extension of the cell).

Specifically, the radiative transfer is computed employing a layer-by-layer pseudo-spherical refractive ray-tracing algorithm. Rayleigh cross sections are computed as a summation of the individual molecular cross sections (Sneep & Ubachs 2005; Villanueva et al. 2022), which are computed at each wavelength based on the polarizability of the encompassing molecules. PSG employs polarizability values as compiled on the Computational Chemistry Comparison and Benchmark DataBase at NIST (https://cccbdb.nist.gov/pollistx.asp). Collision-induced absorptions (CIAs) generated by inelastic collisions of molecules in a gas are included from considering the latest HITRAN CIA compilations (Gordon et al. 2022) and the HITRAN CIA database (Karman et al. 2019), as well as several other sources as reported in Villanueva et al. (2022), including the MT_CKD water continuum (Mlawer et al. 2012), here in version v3.5 (Payne et al. 2020; Kofman & Villanueva 2021). In the presented spectral range and assumed background abundance, only these CIAs contain notable signatures: CO_2–CO_2, H_2O–H_2O, H_2O–N_2, N_2–N_2. Molecular absorptions were included via correlated-k tables based on the latest HITRAN 2020 line list (Gordon et al. 2022), which were complemented at short wavelengths (<1 μm) with UV cross sections, primarily from the MPI Mainz UV/VIS Spectral Atlas (Keller-Rudek et al. 2013). Aerosol properties are modeled following Mie theory, with water cloud scattering properties as described in Massie & Hervig (2013), while the ice cloud optical property parameterization uses Warren (1984), as also described in Massie & Hervig (2013). The partial abundance of the aerosols at the grid box is explicitly defined by the GCM kg/kg profile at that location, meaning that a profile with zero abundance would correspond to a fully clear scenario. An average of all these spectra is then computed to obtain a limb-averaged spectrum as it would be observed by an instrument. Detailed information regarding the computation of correlated-k tables, Rayleigh scattering, the treatment of CIAs, MT_CKD, the ray-tracing algorithm, and the radiative transfer method can be found in Villanueva et al. (2022).

2.3. Instruments, Noise, and Number of Transits

2.3.1. Instruments

We have simulated JWST observations of TRAPPIST-1e transiting its host star. JWST is a 6.5 m tip-to-tip segmented telescope, equivalent to a full circular aperture of 5.6 m in diameter. Previous studies have shown that the NIRSpec Prism (covering the 0.6–5.3 μm region at resolving power $R = 100$) is the JWST instrument most adapted to characterize the atmosphere of temperate terrestrial planets (Fauchez et al. 2019; Lustig-Yaeger et al. 2019; Pidhorodetska et al. 2020; Wunderlich et al. 2020). Indeed, this spectral region contains various molecular signatures of interest, such as O_2, O_3, H_2O, NO_2, N_2O, CH_4, CO, and CO_2. The latter may likely be the only one with strong enough absorption features to be detectable with JWST in a reasonable number of transits (i.e., achievable in five JWST cycles assuming a constant four-transit observation per cycle as in Cycle 1), when clouds and hazes are present in the atmosphere. It has therefore been suggested as the best proxy to detect the atmosphere of habitable planets (Fauchez et al. 2019; Turbet et al. 2020). Note that JWST GTO proposals have already been awarded for the NIRSpec instrument that will attempt molecular detection with four transit observations (program 1331). In this work we compute spectra from 0.6 to 20 μm, across the range of both JWST NIRSpec Prism and MIRI Medium-Resolution Spectrometer (MRS), and present the figures at $R = 100$ offering the best visibility for multiple spectrum plots. Signal-to-noise ratios (S/Ns) and number of transit estimations are only presented for NIRSpec Prism.

2.3.2. Noise

2.3.3. Estimation of the Detectability of the Atmosphere—Identical Transits

First, we computed noise estimates with PSG and validated these by employing the official JWST Exposure Time Calculator, obtaining very good agreement. For NIRSpec Prism, the effective spectral resolution is 0.022 μm. We have selected the clear filter with the subarray SUB512S and the rapid readout pattern with two groups per integration and 0.225 s per frame. This leads to a partial saturation near the peak of the stellar energy distribution (SED) following Batalha et al. (2018) and Lustig-Yaeger et al. (2019). For MIRI LRS, the effective resolution is 0.0654 μm, and we selected the P750L disperser with the subarray SLITLESSPRISM and a FASTR1 readout pattern with 20 groups per integration with a frame time of 0.15 s.

We assumed a transit time of 3345 s (0.93 hr; Agol et al. 2021). To take into account the noise of the out-of-transit baseline, we used JWST NIRSpec GTO proposal 1331 (Lewis et al. 2018), for which each transit event will be observed for ∼4 hr, therefore leading to an out-of-transit observation of ∼3× transit duration. As the noise adds in quadrature, the single transit noise N, including a 3× out-of-transit baseline, is computed as follows:

$$\begin{aligned} N &= \sqrt{N_{\rm out}^2 + N_{\rm in}^2} \\ &= \sqrt{(1/{\rm time}_{\rm out})^2 + (1/{\rm time}_{\rm in})^2} \\ &= \sqrt{(1/3)^2 + (1/1)^2} \\ &= \sqrt{4/3}. \end{aligned} \quad (1)$$

To estimate the one-transit S/N of CO_2 across the NIRSpec Prism range (0.6–5.3 μm) and the number of transits required to achieve a 5σ detection of CO_2, we proceed following the list below:

1. We compute the spectrum without CO_2 (but keeping the other gases in).
2. We compute the spectrum with CO_2.
3. We compute the difference between steps 1 and 2 across the whole instrument range.
4. We compute the S/N by dividing step 3 by the noise for one transit (3345 s) in each spectral interval.
5. We apply the $\sqrt{4/3} = 1.155$ factor to the noise to take into account the out-of-transit noise.
6. The S/N of CO_2 across the whole instrument range is then computed following Lustig-Yaeger et al. (2019) as

$$\mathrm{S/N} = \sqrt{\sum_{i=0}^{n} \mathrm{S/N_i}^2}, \quad (2)$$

where $\mathrm{S/N}_i$ are the individual S/Ns in each spectral interval.

7. From the S/N, the number of transits N_t required to achieve a 5σ detection is given as in Fauchez et al. (2019, 2020b):

$$N_t = \left(\frac{5}{\mathrm{S/N}}\right)^2. \quad (3)$$

3. Transmission Spectra

3.1. Ben 1 and Ben 2 Cases (Dry Planets)

In Figures 1 and 2 we can see the transmission spectra simulated with PSG using the atmospheric profiles of THAI cases Ben 1 and Ben 2, respectively, provided by each GCM. In panel (a) of both figures, the lowest pressures (highest altitudes) used to compute the spectra correspond to the top of the modeled domains, which are 10^{-5} bars, 4×10^{-5} bars, 14×10^{-5} bars, and 4×10^{-5} bars for Ben 1 and 10^{-5} bars, 4×10^{-5} bars, 14×10^{-5} bars, and 13×10^{-5} bars for Ben 2 for ExoCAM, LMD-G, ROCKE-3D, and the UM, respectively. Note that most GCMs have the domain lid at relatively high pressures for numerical stability reasons and since moving it to lower pressures requires taking into account complex upper atmospheric processes such as non-LTE, molecular diffusion, etc. (see Fauchez et al. 2021, Section 4.1). The lowest pressures usually correspond to a top-of-atmosphere (TOA) altitude of 50−70 km for Earth-like simulations. However, the pressure at this altitude is usually too high to fully capture the transmitted light through the planet's atmosphere, and the strongest atmospheric features can be truncated. This is clearly seen in the Ben 2 case, where the CO_2 strongest absorption lines are truncated (Figure 2(a)). To bypass this limitation, we used PSG to extend the atmosphere to much lower pressures, assuming an isothermal profile and constant volume mixing ratios for the dry gases. This is similar to the so-called "ghost layer" used in Amundsen et al. (2016). We have estimated the TOA pressure that would fully resolve the spectral lines for Ben 1 and Ben 2 as 10^{-7} and 10^{-10} bars, respectively. Lower pressures are required for Ben 2 because the opacity of a pure CO_2 atmosphere remains strong at lower pressures than that of a N_2-dominated atmosphere.

Using the data with extrapolated model top, we have estimated the number of transits that would be required to detect such atmospheres with a 5σ confidence level as presented in Table 1. Two estimations are shown. The first one uses the method presented in Section 2.3.2 and is referred to as 5σ transits. The second method only uses the CO_2 line at 4.3 μm as done in Fauchez et al. (2019), Wunderlich et al. (2019), and Pidhorodetska et al. (2020) and is referred to as 5σ Transit-4.3 μm. Also shown are the S/N for one transit (S/N-1) and four transits (S/N-4), corresponding to JWST Cycle 1 TRAPPIST-1e transit observation (GTO Proposal 1331 by PI Nikole Lewis; Lewis et al. 2017).

First, we can see that for a Ben 1 atmosphere an average of 2.6σ could be achieved from Cycle 1, while an average of 4.3σ could be achieved for a Ben 2 atmosphere with more CO_2. To reach the necessary 5σ threshold, an average of 17 and 6 transits would be needed for Ben 1 and Ben 2, respectively. When using only the CO_2 line at 4.3 μm, these numbers go up to 24 and 25 transits, respectively. This demonstrates that this method strongly overestimates the number of necessary transits, especially if the amount of CO_2, and therefore the number of strong lines, is high. The intermodel differences for 5σ Transit are small in both cases, the maximum difference between the GCMs being 24% and 33% for Ben 1 and Ben 2, respectively, demonstrating that the four GCMs provide similar atmospheric profiles at the terminator to provide consistent simulated spectra and expected number of transits to detect a dry 1 bar atmosphere with a relatively high mean molecular weight. Note that we did not consider the spectral impact of dust that can be lifted from the surface of a land planet and persist in the atmosphere. Dust would likely raise the continuum level, thereby decreasing the amplitude of each spectral line as shown in Boutle et al. (2020). The expected effect would be of the order of 10 ppm.

3.2. Hab 1 and Hab 2 Cases (Aquaplanets)

Rocky exoplanets in the HZ and with surface liquid water will likely have water in a vapor and condensed form in the atmosphere. Clouds have been shown to severely impede atmospheric characterization via transmission spectroscopy (Fauchez et al. 2019; Suissa et al. 2020a; Komacek et al. 2020). Furthermore, clouds are notoriously difficult to represent correctly in GCMs because the characteristic timescale and size of individual clouds are too small to be simulated explicitly and they involve a tremendous number of physical processes. GCMs thus rely on sub-grid-scale parameterizations to represent the formation of clouds that can significantly differ between models (Sergeev et al. 2022). These discrepancies can then lead to different predicted surface temperatures, as was noted in exoplanet GCM simulations by Yang et al. (2019). Details on the differences in GCM predictions, especially for clouds, are given in the companion paper (Sergeev et al. 2022).

Here we use PSG to compute transmission spectra for both the Hab 1 and Hab 2 cases. The atmospheric properties at the terminator have been time-averaged over the 100 orbits in order to smooth out variability that can be introduced by weather patterns and change in clouds at the terminator, as is commonly done for such planets (Fauchez et al. 2019; Suissa et al. 2020a, 2020b; Komacek et al. 2020; Pidhorodetska et al. 2020). Details on the impact of atmospheric variability on transmission spectra are given in Section 3.3.

Figure 3 shows the Hab 1 transmission spectra calculated using the output from ExoCAM, LMD-G, ROCKE-3D, and the UM. The extension of the top of the model has a negligible impact on the spectra, except within the strongest 4.3 and 15 μm CO_2 absorption bands. The differences between the spectra from the different GCMs are mostly noticeable in the continuum and for the weakest absorption bands. Indeed, as

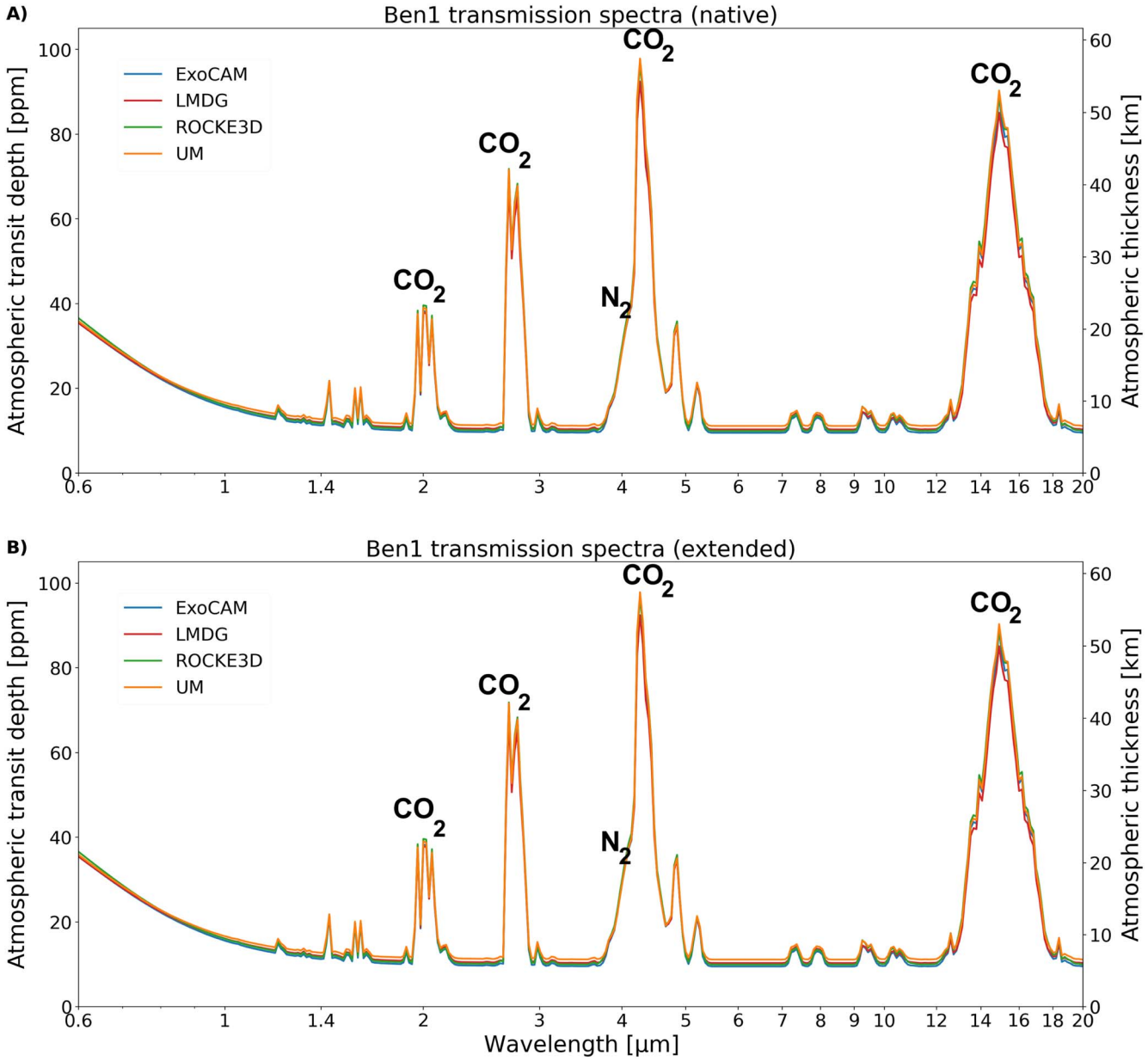

Figure 1. Ben 1 transmission spectra simulated with PSG using the terminator atmospheric profiles from the ExoCAM, LMD-G, ROCKE-3D, and UM simulations. In panel (a) the atmospheric profiles are limited up to the lowest pressure at the top of the GCMs, namely, 10^{-5} bars, 4×10^{-5} bars, 14×10^{-5} bars, and 4×10^{-5} bars for ExoCAM, LMD-G, ROCKE-3D, and UM, respectively, while in panel (b) the atmospheric profiles have been extended up to 10^{-7} bars assuming an isothermal atmosphere and fixed mixing ratios for the dry gases. The differences between the transmission spectra are extremely small.

already shown by Fauchez et al. (2019) and Suissa et al. (2020a), the continuum level in a cloudy atmosphere is raised to the altitude of the cloud deck. Strong bands like CO_2 at 4.3 μm are less affected by clouds because even if the denser, most absorbing part of the atmosphere is under them, the efficiency of absorption is so strong that the small CO_2 partial pressure remaining above the cloud deck is large enough to saturate the line. The ExoCAM continuum level is the highest, followed by LMD-G, ROCKE-3D, and the UM. This is explained by the fact that the liquid water cloud mixing ratio and the altitude of the cloud deck at the west and east terminators, respectively, are much higher in ExoCAM than in LMD-G, ROCKE-3D, and the UM, in that order (Figures 4(e), (f)). Regarding ice clouds, while the mixing ratios predicted by each model are comparable, the average altitude of the clouds is different, with ExoCAM producing the highest ice clouds, followed by LMD-G, ROCKE-3D, and the UM (Figures 4(g), (h)).

Figure 3 also shows that the LMD-G water band around 6 μm is significantly weaker than that predicted by other models. This is due to the fact that LMD-G simulations have a much drier upper atmosphere (Figures 4(c), (d)). The amount of water above the tropopause in LMD-G is primarily controlled by the tropopause temperature at the substellar point (where water is injected by deep moist convection), which is the coldest in LMD-G, especially at the western terminator (see the companion paper, Sergeev et al. 2022, Figure 4). As a result, the detectability of water in Hab 1 simulations for LMD-G is even more challenging than for the other GCMs.

Figure 5 is the same as Figure 3, but for Hab 2 simulations. First, we can see that because the CO_2 mixing ratio is much

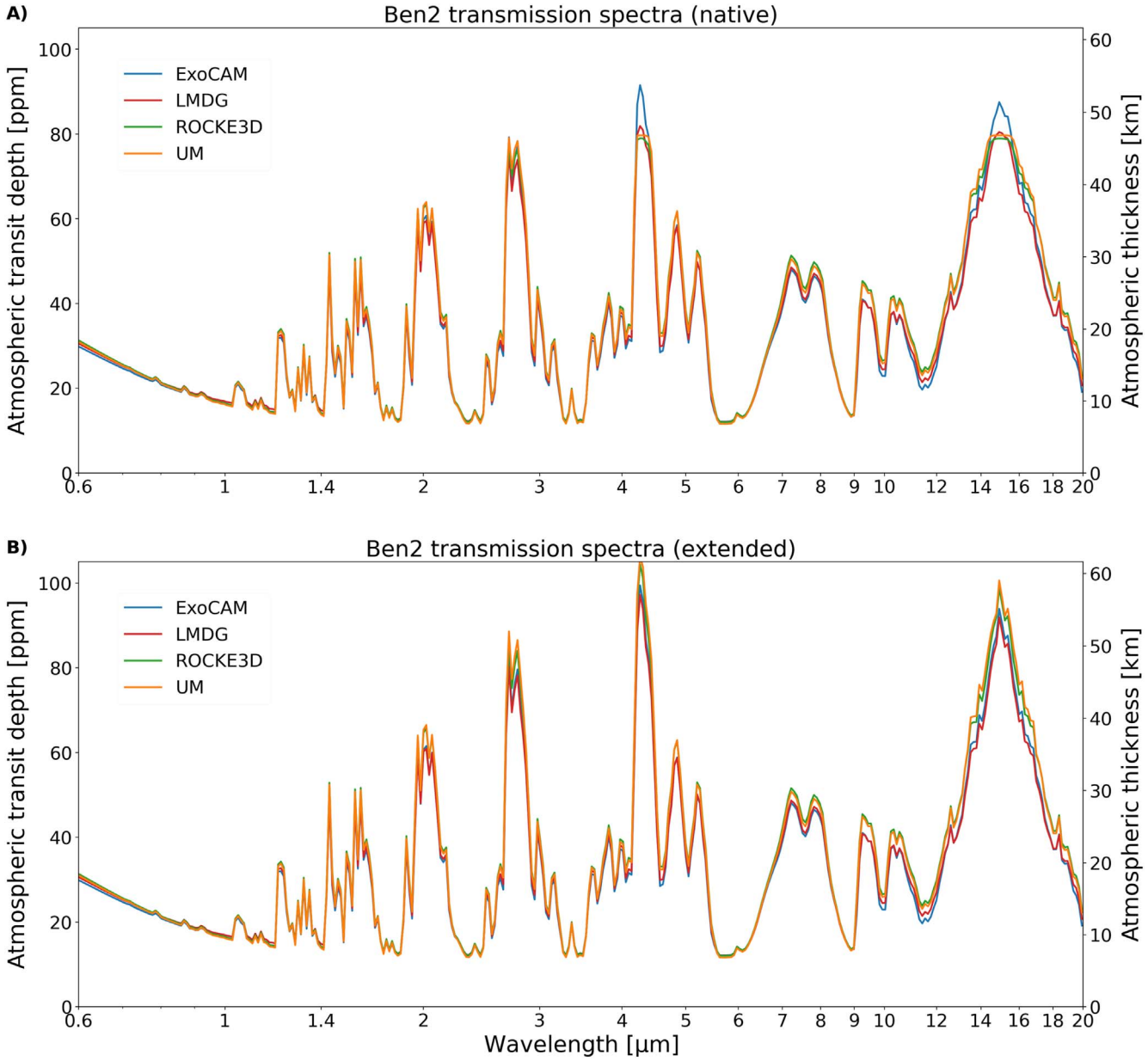

Figure 2. Ben 2 transmission spectra simulated with PSG using the terminator atmospheric profiles from the ExoCAM, LMD-G, ROCKE-3D, and UM simulations. In panel (a) the atmospheric profiles are limited up to the lowest pressure at the top of the GCMs, namely, 10^{-5} bars, 4×10^{-5} bars, 14×10^{-5} bars, and 13×10^{-5} bars for ExoCAM, LMD-G, ROCKE-3D, and the UM, respectively, while in panel (B) the atmospheric profiles have been extended up to 10^{-10} bars assuming an isothermal atmosphere and fixed mixing ratios for the dry gases. CO_2 features are truncated if the atmospheres are not vertically extended. Visible differences between the spectra appear in the CO_2 bands in both the native (panel (a)) and extended (panel (b)) scenarios. CO_2 labels are not shown, as the spectra are exclusively CO_2.

higher in the Hab 2 simulations (from about 400 ppm for Hab 1 to nearly 100% for Hab 2), strong absorption lines are more easily truncated by a low model top. Similarly to the Hab 1 case, the ExoCAM continuum is higher than that of the other three GCMs. This time, however, the continuum level in ROCKE-3D is slightly above that in LMD-G. Note that LMD-G's absorption peaks are the smallest among the four GCMs.

In the warm and humid atmosphere of the Hab 2 case (Figures 6(a), (b)) there is no clear temperature inversion at the tropopause except for a decrease in the lapse rate from 100 hPa and lower (for more details see Sergeev et al. 2022). The specific humidity closely follows the temperature profiles: colder temperature profiles correspond to drier profiles (Figures 6(c), (d)). Furthermore, the warm atmosphere of Hab 2 results in the liquid water cloud mass mixing ratio being comparable to and even larger than the ice cloud mixing ratio (Figure 6). When the altitudes of both cloud types are combined, ExoCAM has on average higher and thicker clouds, followed by ROCKE-3D, then LMD-G, and finally the UM. It is interesting to note that in the warmer, moister, and cloudier Hab 2 case, the relative difference in cloudiness between LMD-G, ROCKE-3D, and the UM is smaller than that for Hab 1. Only ExoCAM persistently produces higher clouds. More detailed discussion about the differences of cloud coverage produced by the THAI models is given in Sergeev et al. (2022).

Similar to the Ben 1 and Ben 2 experiments, we extrapolated the model top for the PSG calculation, which gave the number of transits required for atmospheric detection with a 5σ confidence level by JWST using the full NIRSpec Prism spectral range (5σ Transit) and with the CO_2 line at 4.3 μm

Table 1
Table Summarizing the Signal-to-noise Ratio Achieved from One Transit (S/N-1), That from the Four JWST Cycle 1 Transits (S/N-4), Number of Transits to Reach a 5σ Detection Using All Available CO_2 Lines (5σ Transit), and That Using only the 4.3 μm CO_2 Line

			Ben 1	
Model	S/N-1	S/N-4	5σ Transit	5σ Transit-4.3 μm
ExoCAM	1.3	2.6	16	23
LMD-G	1.2	2.4	19	25
ROCKE-3D	1.3	2.6	15	23
UM	1.3	2.6	16	23
Average	1.3	2.6	17	24
Maximum difference (%)	8	8	24	8
			Ben 2	
Model	S/N-1	S/N-4	5σ Transit	5σ Transit-4.3 μm
ExoCAM	2.2	4.4	5	23
LMD-G	2.0	4.0	7	28
ROCKE-3D	2.2	4.4	5	23
UM	2.2	4.4	5	24
Average	2.2	4.3	6	25
Maximum difference (%)	9	9	33	20
			Hab 1	
Model	S/N-1	S/N-4	5σ Transit	5σ Transit-4.3 μm
ExoCAM	0.9	1.8	35	39
LMD-G	0.9	1.8	29	37
ROCKE-3D	1.0	2.0	38	31
UM	1.0	2.0	23	24
Average	1.0	2.0	29	33
Maximum difference (%)	10	10	41	45
			Hab 2	
Model	S/N-1	S/N-4	5σ Transit	5σ Transit-4.3 μm
ExoCAM	1.5	3.0	12	36
LMD-G	1.7	3.4	9	31
ROCKE-3D	1.7	3.4	8	29
UM	2.0	4.0	7	23
Average	1.7	3.4	9	30
Maximum difference (%)	29	29	56	43

only (5σ Transit-4.3 μm). For Hab 1, we found that 35, 29, 38, and 23 transits are required, while for Hab 2 we found that 12, 9, 8, and 7 transits are required (for ExoCAM, LMD-G, ROCKE-3D, and the UM, respectively). The average number of transits for Hab 1 is 29, with a maximum difference of 41%, and for Hab 2 it is 9, with a maximum difference of 56%. Similarly to the Ben cases, the use of only the 4.3 μm CO_2 line overestimates the predicted numbers, especially for the 1 bar CO_2 case. Note that the reason why the overall required number of JWST observed transits is higher for the Hab scenarios is the presence of clouds that raise the continuum level up to the cloud deck altitude, shrinking each absorption line from the bottom and therefore reducing their detectability. This can clearly be seen in Figure 7, where the left panels show the Ben 1 simulations and the right panels the Hab 1 simulations displaying a change in the continuum level. That change is significantly larger for ExoCAM, which has higher clouds, than the UM. We can also see in Figure 7 that one and four transits (green and magenta error bars, respectively) are far from being enough to detect the CO_2 features at 5σ but that 20 transits (blue error bars) may be enough for Ben 1, while 40 transits (red error bars) may be necessary for Hab 1. Finally, we can also see that for the Hab 1 cases water lines around 1.4 μm are too small to be detectable, even for the UM showing stronger H_2O lines due to its lower cloud deck.

The differences between the predicted transits from each GCM terminator atmosphere are controlled to the first order by the altitude of the cloud deck and to the second order by the temperature profile. Those differences between ExoCAM and LMD-G on one hand and ROCKE-3D and the UM on the other are significant and could have consequences for the number of hours requested for a JWST proposal and for the interpretation of future data using retrieval algorithms. However, it seems clear that regardless of the GCM used to produce the atmospheric data, at least seven transit observations would be required to detect at a 5σ confidence level a high molecular weight atmosphere on TRAPPIST-1e with a cloudy sky. Note that during JWST Cycle 1, four transits with NIRSpec Prism will be observed as part of the GTO program 1331 led by PI Nikole Lewis. On average, this would lead to a 2σ detection for Hab 1 and a 3.4σ detection for Hab 2 (see the third column of Table 1).

3.3. Intertransit Variability

Many previous modeling studies estimating the detectability of an exoplanet through transmission spectroscopy have assumed that each planetary transit will be constant through time (Fauchez et al. 2019; Lustig-Yaeger et al. 2019; Suissa et al. 2020a, 2020b; Komacek et al. 2020; Pidhorodetska et al. 2020). However, this is not a realistic assumption, as weather patterns and clouds, if present, are likely to change from one transit to another. Previous work on hot Jupiters by Komacek & Showman (2019) has shown that temporal variability could be already detectable using either secondary eclipse observations with JWST or phase curve observations and/or Doppler wind speed measurements with high-resolution spectrographs. More recently, May et al. (2021) simulated TRAPPIST-1e with ExoCAM for various concentrations of CO_2 and looked at the atmospheric variability between 10 transits induced by ice water clouds. The amplitudes of the transit variability for their 10^{-4} bars of CO_2 (comparable to Hab 1) and 1 bar of CO_2 (comparable to Hab 2) are very similar, of the order of 10 and 20 ppm (May et al. 2021, their Figure 4), respectively. However, they computed one transmission spectrum per day, while in our study we compute it at the exact time of the transit, i.e., every 6.1 days, potentially leading to larger atmospheric differences. The main conclusion is that the time variability of the spectra does not affect retrieved abundances at detectable levels. However, the findings of May et al. (2021) are likely to be dependent on the GCM (ExoCAM). Here we analyze the intertransit variability produced by three more GCMs: LMD-G, ROCKE-3D, and the UM.

Figure 8 shows the standard deviation of the atmospheric transit depth and of the transit atmospheric thickness over 100 transits. This variability is wavelength dependent: it is the largest in the continuum, as the transmitted light is closer to the surface, where clouds are present, and the smallest for the strongest absorption lines like the CO_2 at 2.7, 4.3, and 15 μm. There are significant differences between the GCMs. In general, the cloudier the simulation is, the more variable the

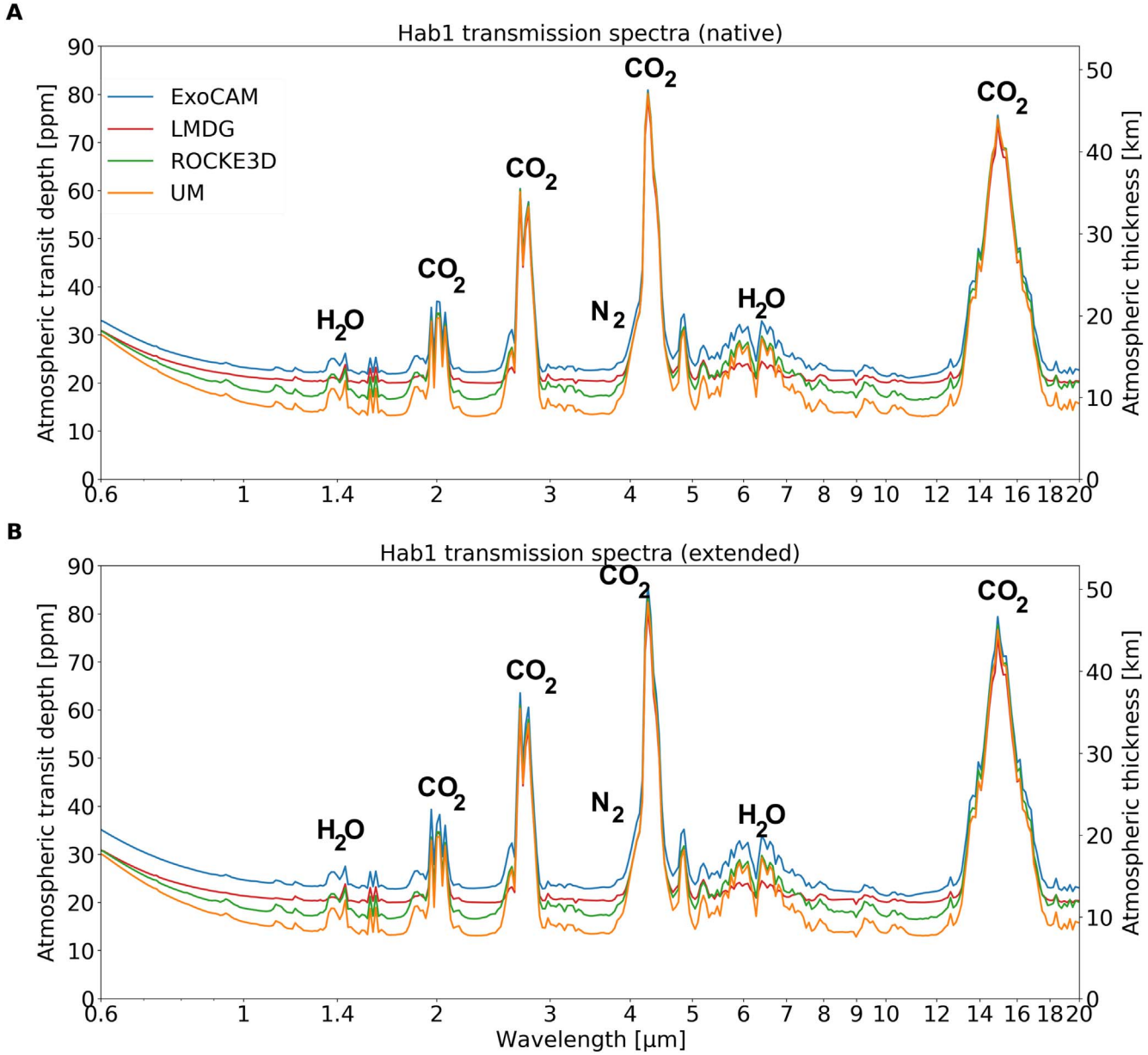

Figure 3. Hab 1 transmission spectra simulated with PSG using the terminator atmospheric profiles from the ExoCAM, LMD-G, ROCKE-3D, and UM simulations. In panel (a) the atmospheric profiles are limited up to the lowest pressure at the top of the GCMs, namely, 10^{-5} bars, 4×10^{-5} bars, 14×10^{-5} bars, and 4×10^{-5} bars for ExoCAM, LMD-G, ROCKE-3D, and the UM, respectively, while in panel (b) the atmospheric profiles have been extended up to 10^{-7} bars assuming an isothermal atmosphere and fixed mixing ratios for the dry gases.

transmission spectrum is. For LMD-G and ROCKE-3D, the time variability in both Hab 1 and Hab 2 is remarkably similar, while for ExoCAM and the UM it differs. This difference is due to the change in the average altitude of clouds between Hab 1 (Figure 4) and Hab 2 (Figure 6). In LMD-G and ROCKE-3D simulations, the average altitude of liquid water and ice water clouds does not change substantially between Hab 1 and Hab 2, while for ExoCAM and the UM it increases sharply in Hab 2. For instance, for ExoCAM the east terminator water ice cloud maximum density peaks at 500 hPa for Hab 1 and at 250 hPa for Hab 2. We hypothesize that stronger winds at this lower pressure relative to those deeper in the atmosphere lead to higher cloud variability. Overall, the standard deviation of the continuum level in the Hab 1 case for ExoCAM, LMD-G, ROCKE-3D, and the UM is about 3, 3, 2, and 1 ppm, respectively, leading to a median value of ∼2 ppm. For Hab 2 it is about 5, 3, 2.5, and 2 ppm, respectively, leading to a median value of ∼3 ppm. These values are low relative to the JWST expected 1σ noise of 10–25 ppm as assumed in Fauchez et al. (2019) and near the upper limit value (14 ppm) estimated by NIRSpec lab time series in Rustamkulov et al. (2022). It is also interesting to note that those values are comparable to the relative transit depth of H_2O or O_2 (Pierrehumbert 2010; Fauchez et al. 2019; Lustig-Yaeger et al. 2019; Wunderlich et al. 2020). This means that even if one assumes no noise floor, atmospheric variability would produce a continuum fluctuation that would swamp those highly important but weak absorption lines.

Figure 9 shows the spectra as the ratio in percentage of the standard deviation of the variability with respect to the measurement noise for 100 (blue), 50 (orange), 25 (green), and 10 (red) transits for the atmospheric transit depth (left *Y*-axis) and

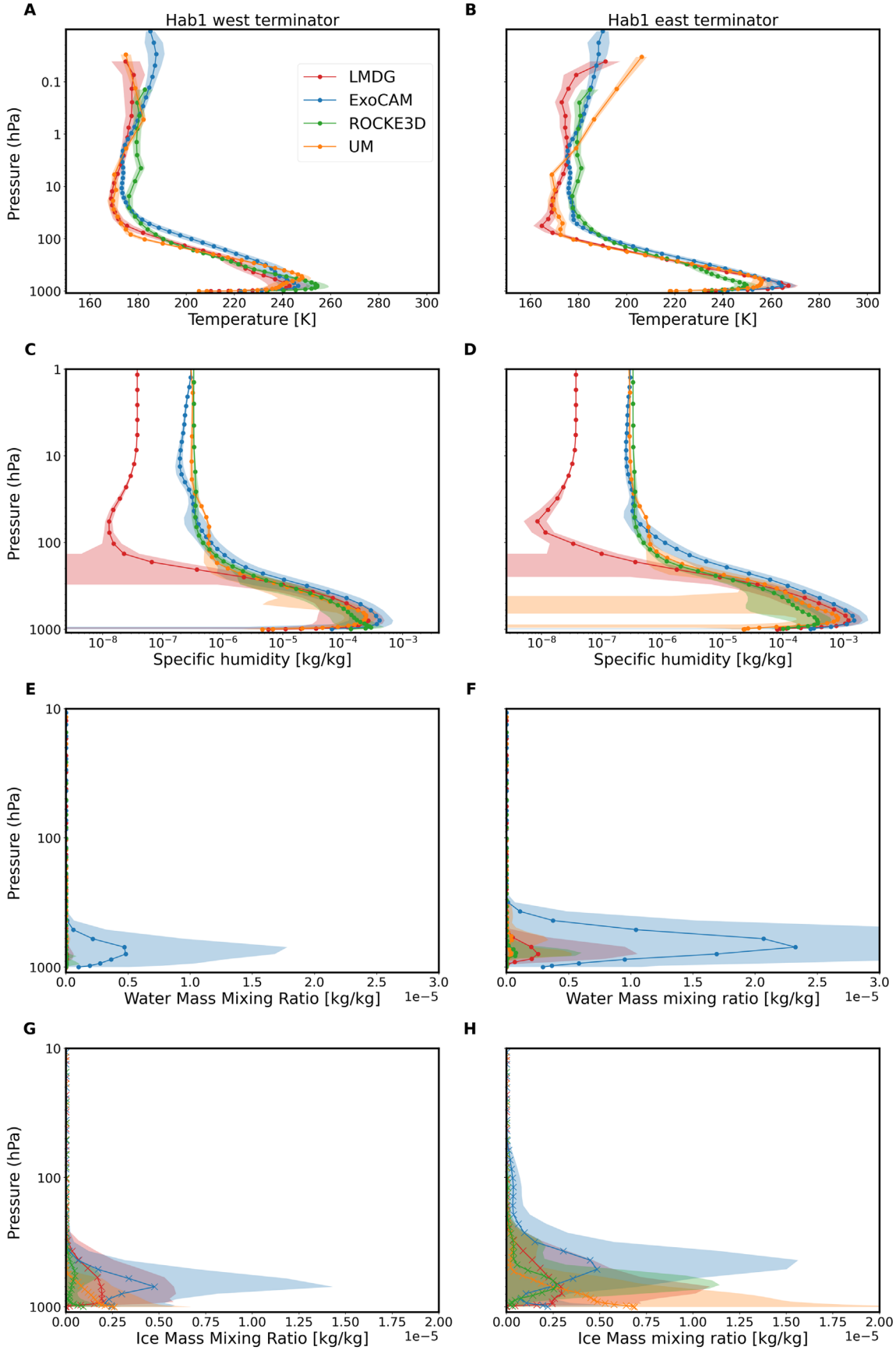

Figure 4. Hab 1 terminator atmospheric profiles predicted by ExoCAM (blue), LMD-G (red), ROCKE-3D (green), and the UM (orange). From top to bottom: temperature, specific humidity, the mass mixing ratio of liquid water, and the mass mixing ratio of ice water for the west terminator (left column) and east terminator (right column). Time-averaged values are represented by thick lines, while the 1σ deviations are represented by shaded regions.

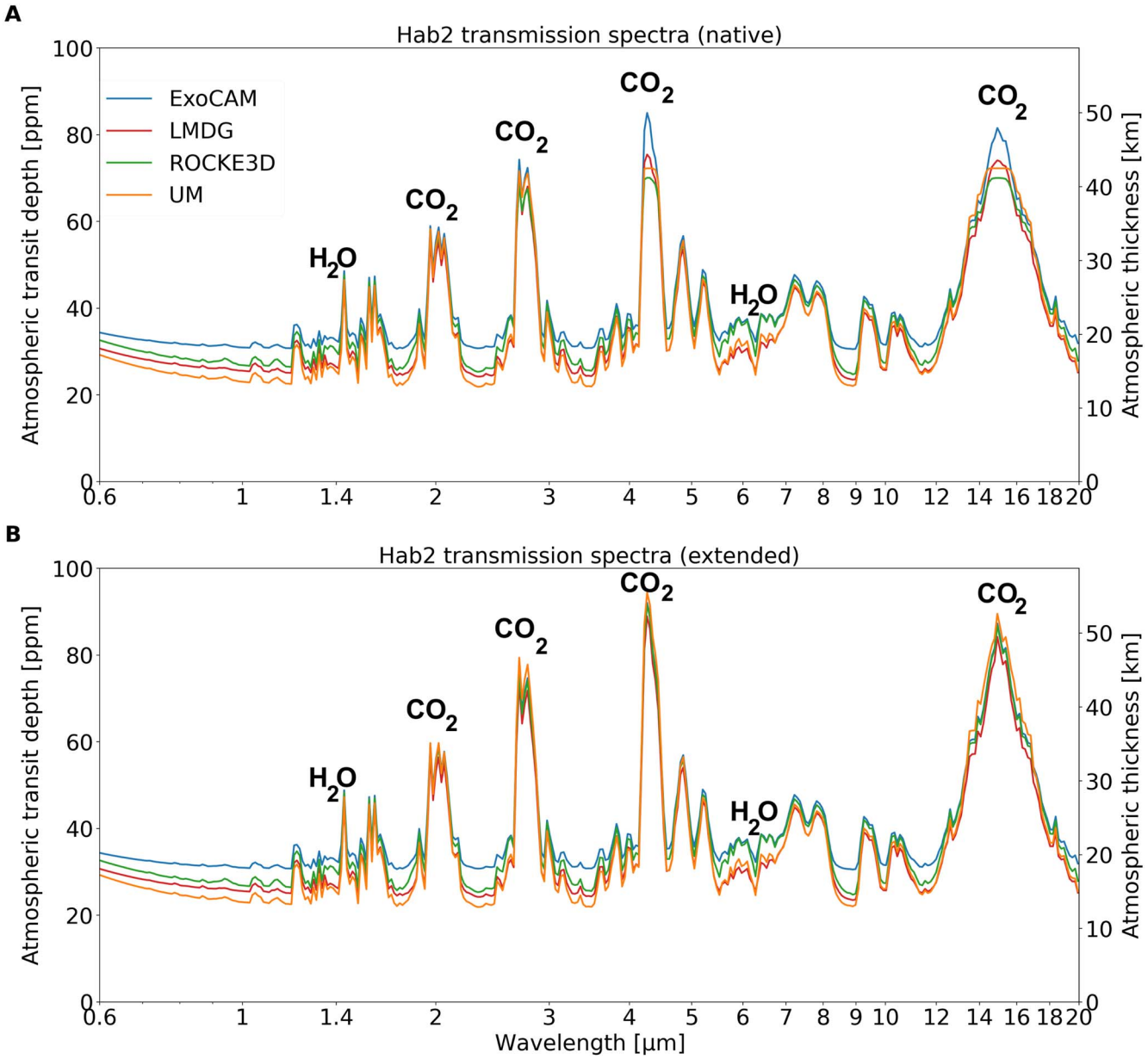

Figure 5. Hab 2 transmission spectra simulated with PSG using the terminator atmospheric profiles from the ExoCAM, LMD-G, ROCKE-3D, and UM simulations. In panel (a) the atmospheric profiles are limited up to the lowest pressure at the top of the GCMs, namely, 10^{-5} bars, 4×10^{-5} bars, 14×10^{-5} bars, and 13×10^{-5} bars for ExoCAM, LMD-G, ROCKE-3D, and the UM, respectively, while in panel (b) the atmospheric profiles have been extended up to 10^{-10} bars assuming an isothermal atmosphere and fixed mixing ratios for the dry gases. CO_2 features are truncated if the atmospheres are not extended. Differences between the spectra due to clouds are mostly seen in the continuum and for the weakest absorption bands.

atmospheric thickness (right *Y*-axis). The larger the number of transits, the lower the noise and therefore the higher the variability-to-noise ratio (%). Interestingly, these spectra can be reminiscent of emission spectra (Morley et al. 2017; Fauchez et al. 2019; Lustig-Yaeger et al. 2019). The minimum values correspond to the absorption-line peak, while the maxima correspond to the continuum. Only the Hab 2 atmosphere simulated by ExoCAM would lead to a transit depth and atmospheric thickness variability higher than the measurement noise if 100 transits or more are acquired with JWST. Considering observational constraints and science priority, this number is likely too high. Fauchez et al. (2019) have used the https://exoctk.stsci.edu/contam_visibility tool to estimate the number of times TRAPPIST-1e will be visible transiting over JWST 5 yr of nominal lifetime and found 85 (17 transits per year). The recent successful launch of JWST and optimized Ariane V trajectory saving large amounts of fuel has allowed us to extend JWST's potential lifetime up to 20 yr. If every single transit is effectively observed, this is an upper limit of 340 available transits. However, for Cycle 1, only four transits are going to be observed; if we assume that number constant over 20 yr, it will bring up to 80 transits. In our work, we have therefore considered 50 nonconsecutive transits as a realistic but conservative estimate. In a more realistic scenario of 50 nonconsecutive transits accumulated over the lifetime of the JWST, the impact of atmospheric variability relative to the noise for Hab 1 and Hab 2 would be of about 50% and 80% for ExoCAM, 50% and 50% for LMD-G, 40% and 40% for ROCKE-3D, and 20% and 40% for the UM, respectively, and will therefore not be of concern.

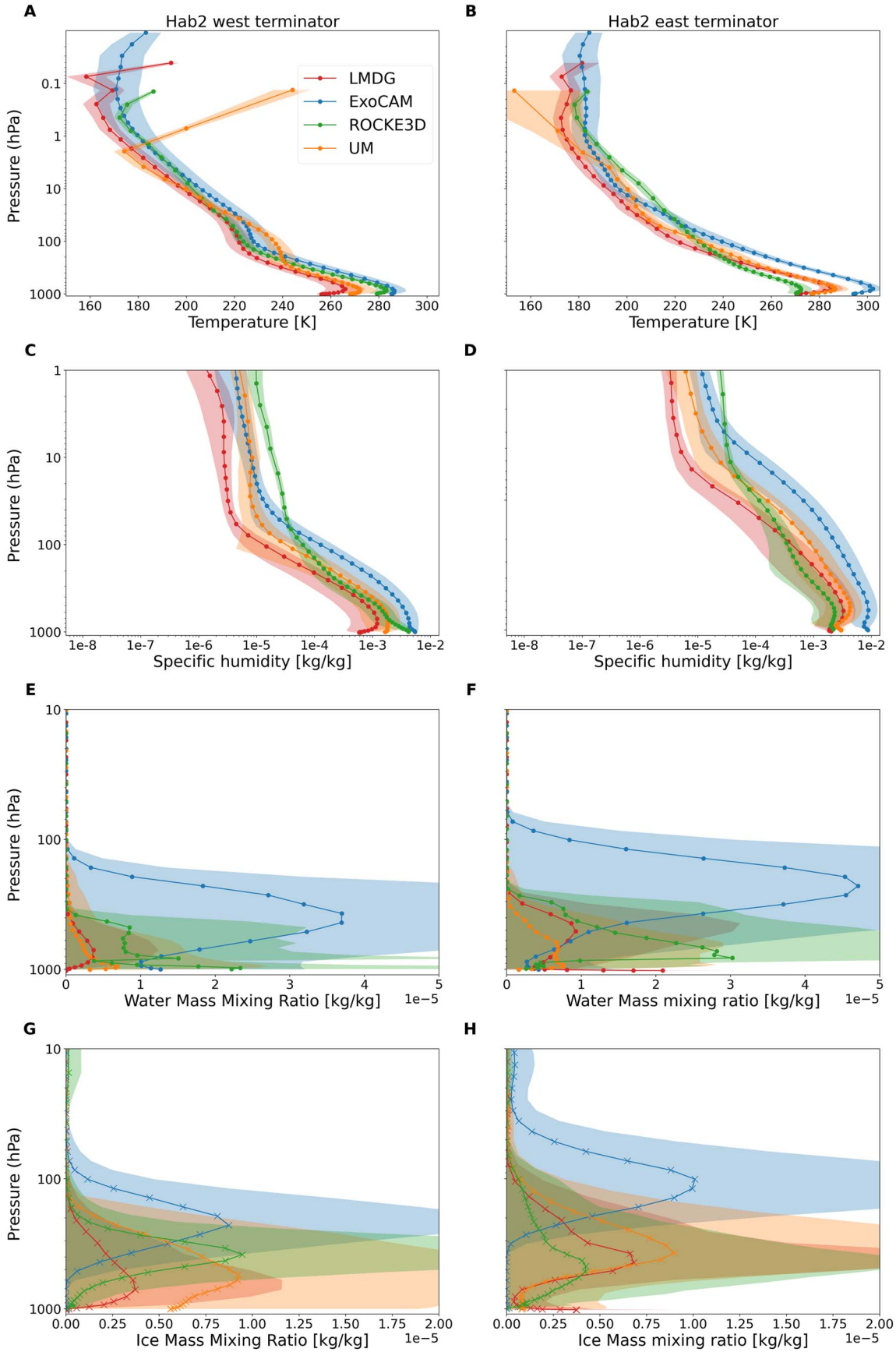

Figure 6. Hab 2 terminator atmospheric profiles simulated by ExoCAM (blue), LMD-G (red), ROCKE-3D (green), and the UM (orange). From top to bottom: temperature, specific humidity, mass mixing ratio of liquid water, and mass mixing ratio of ice water for the west terminator (left column) and east terminator (right column). Time-averaged values are represented by thick lines, while the 1σ deviations are represented by shaded regions.

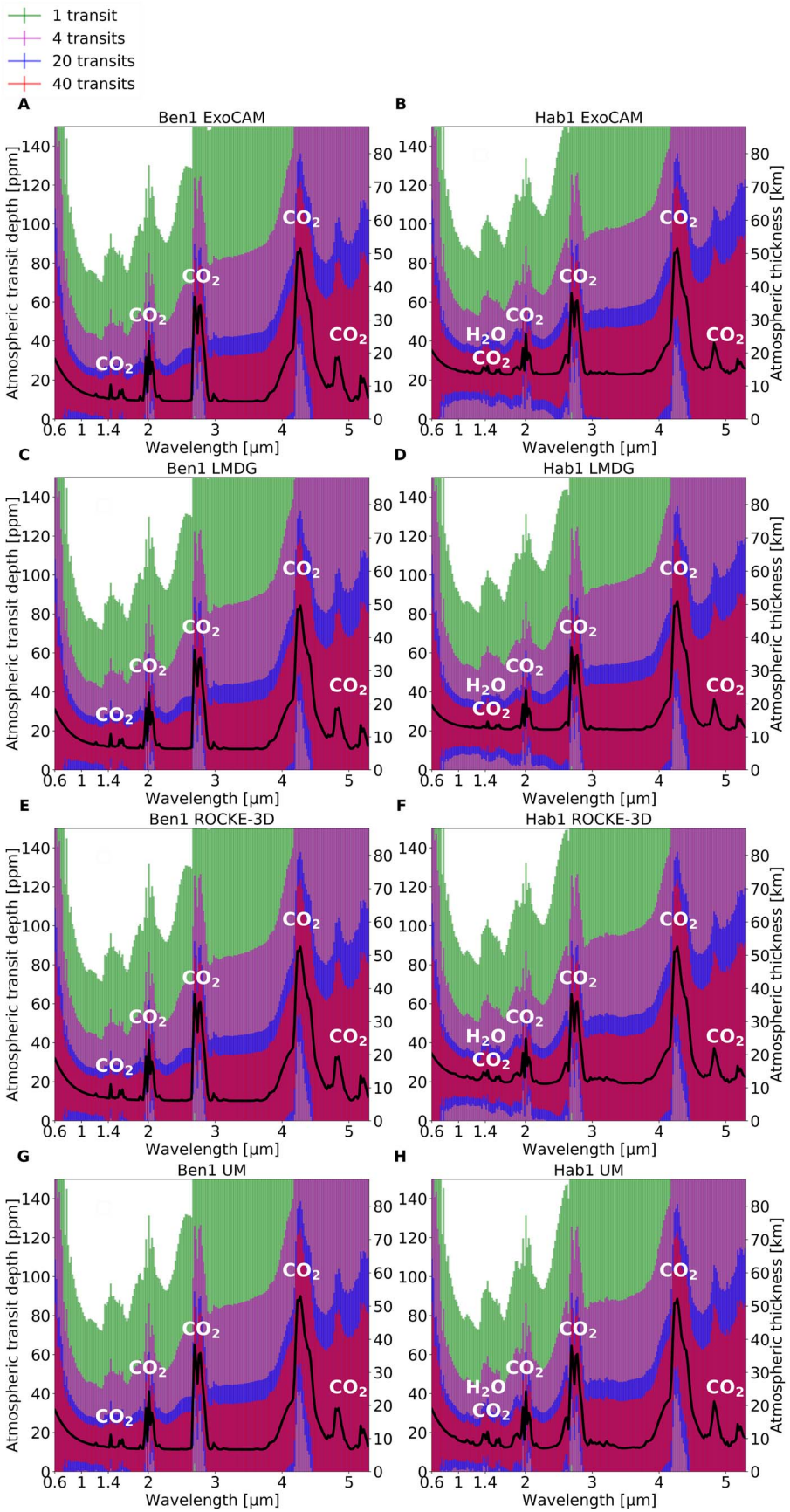

Figure 7. Ben 1 (left column) and Hab 1 (right column) transmission spectra for ExoCAM (panels (a) and (b)), LMD-G (panels (c) and (d)), ROCKE-3D (panels (e) and (f)), and the UM (panels (g) and (h)). The error bars are shown for the native NIRSpec spectral resolution and for 1 transit (green), 4 transits (magenta, NIRSpec Cycle 1 proposal 1331; Lewis et al. 2017), 20 transits (blue, ∼5 JWST years at a constant four transits per year), and 40 transits (red, ∼10 JWST years at a constant four transits per year). Relevant absorption features are written in white.

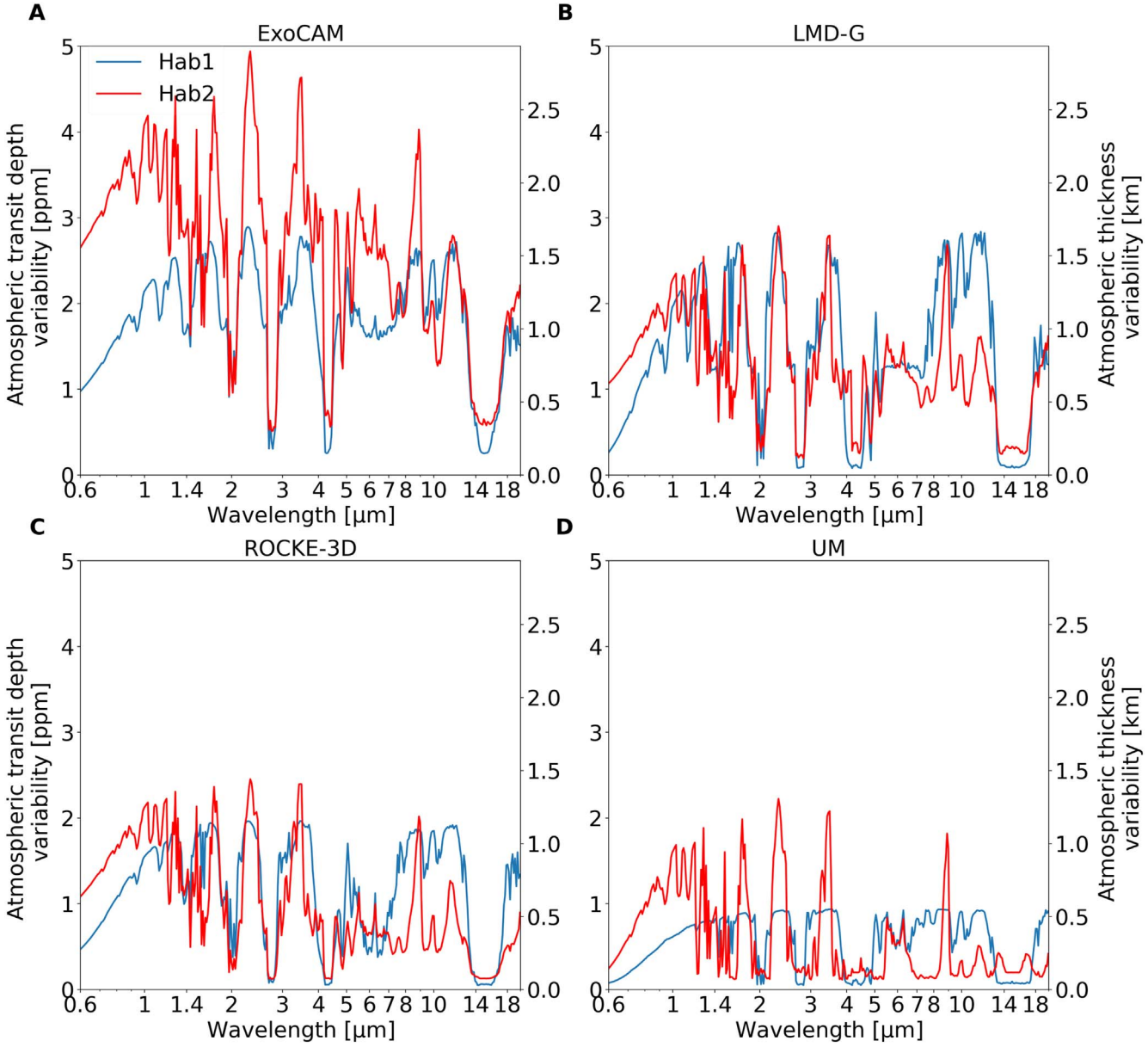

Figure 8. Standard deviation of the atmospheric transit depth [ppm] and of the atmospheric thickness [km] as a function of wavelength for Hab 1 (blue) and Hab 2 (red) and for ExoCAM (panel (a)), LMD-G (panel (b)), ROCKE-3D (panel (c)), and the UM (panel (d)).

Comparisons between all GCMs in a single panel are shown in Figure 10. In panels (a) (Hab 1) and (b) (Hab 2) are shown each ExoCAM transmission spectrum (black lines) and the median spectrum (blue line) with the associated 1σ error (blue error bars). ExoCAM was selected for this example, as it is the GCM showing the largest variability. In panels (c) (Hab 1) and (d) (Hab 2) the spectra of the variability relative to noise for 50 transits and for the four GCMs are shown.

As a summary, in the case of TRAPPIST-1e it seems that predictions of the atmospheric variability introduced by clouds for a N_2- or CO_2-dominated atmosphere are within the measurement noise for a reasonable number of transits (<50) regardless of the GCM used to simulate this temporal and spatial variability. However, if in the fortunate event (albeit unlikely; see the discussion in Gillon et al. 2020) that a similar exoplanet were to be found closer to Earth, the noise will be reduced. In that case the atmospheric variability could be a possible proxy for the presence of clouds via the temporal changes of the continuum level relative to the relatively stationary strong absorption peaks. Without such variability, a continuum level corresponding to a cloud deck, an atmospheric refraction limit, or a planet's surface may not be discernible. Additionally, ExoCAM simulations produce systematically higher variability in the synthetic spectra compared to that in the three other models, while the UM tends to produce the lowest variability. LMD-G and ROCKE-3D are in the middle. We recommend that these differences in variability are taken into account when using a single GCM for future studies. Finally, Sergeev et al. (2022) have noticed that the timescale of this periodic variability differs between the models (see their Figure 12). The tendency is the same for Hab 1 and Hab 2, with ExoCAM showing the longest period of 12.5 orbits for both cases, followed by LMD-G with 11.1 and 7.7 orbits, respectively, ROCKE-3D with 2.3 and 3 orbits, respectively, and the UM with 1.1 and 2.5 orbits, respectively. With real atmospheric data it is unknown what this period would be within this 1–12.5 orbit range. As a result, it is not clear whether observing consecutive

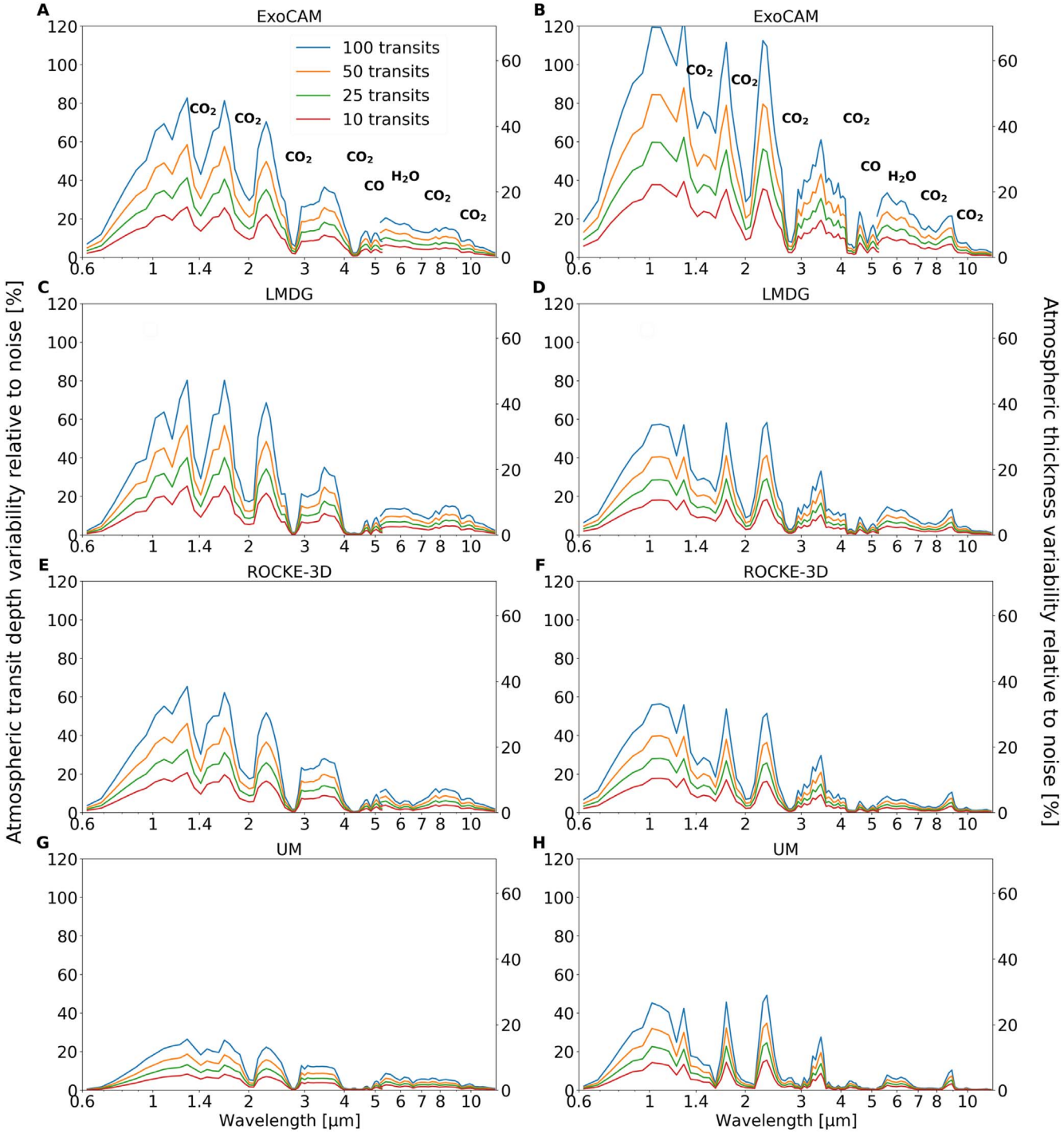

Figure 9. Atmospheric transit depth variability relative to noise (left *Y*-axis) and atmospheric thickness variability relative to noise (right *Y*-axis) for 100 (blue), 50 (orange), 25 (green), and 10 (red) transits for ExoCAM (panels (a) and (b)), LMD-G (panels (c) and (d)), ROCKE-3D (panels (e) and (f)), and the UM (panels (g) and (h)). Hab 1 is in the left column, while Hab 2 is in the right column. The noise is computed at the native instrument resolution and later binned down by a factor 3 for this figure.

transits or scattered ones would have any impact on the atmospheric characterization. Finally, it is worth noting that while TRAPPIST-1e lies between the fast and Rhines rotation regimes (Sergeev et al. 2022; Turbet et al. 2022), planets with longer orbital periods that remain in synchronous rotation will transition from the Rhines rotation regime to a slow rotation regime, which increases the symmetry of temperature and winds about the substellar point (Haqq-Misra et al. 2018). However, how the atmospheric variability across transits would be affected by the change of atmospheric regime remains to be explored.

4. Conclusions

In this third and last part of the THAI paper series we analyzed how the prediction of the detectability of N_2- and CO_2-dominated atmospheres on TRAPPIST-1e is sensitive to the choice of a 3D GCM used to simulate its atmosphere.

First, we have simulated the transmission spectra for the Ben 1 and Ben 2 scenarios (dry land planets for which the comparison of the predicted atmosphere is presented in Part I; Turbet et al. 2022) using the PSG (Villanueva et al. 2018). We

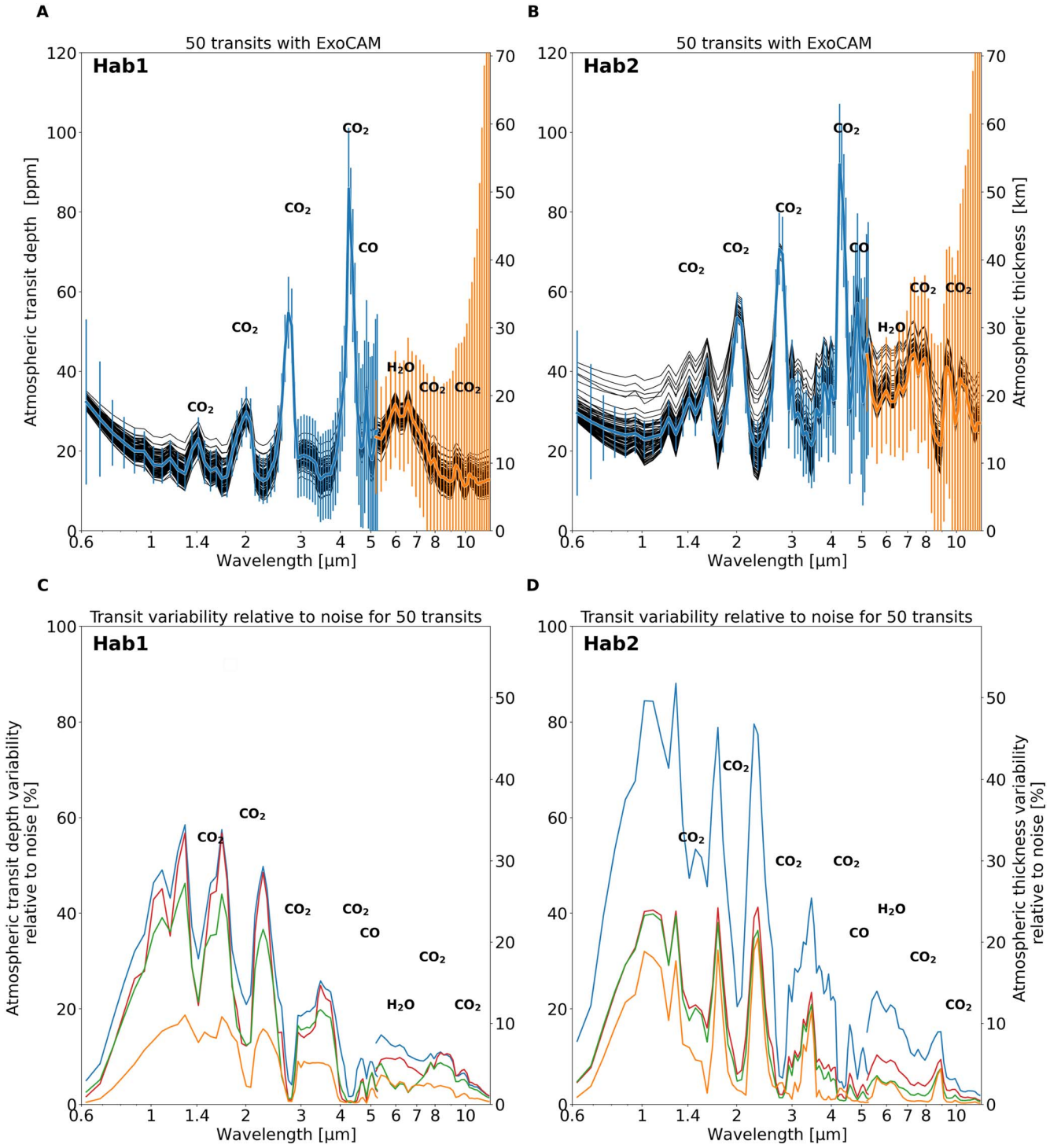

Figure 10. (a, b) Transmission spectra for Hab 1 and Hab 2, respectively, obtained from ExoCAM simulations. Each fine black trace represents 1 of the 100 individual transits, and the thick blue line represents the median transit. The noise for 50 transits and for both NIRSpec Prism (blue) and MIRI (orange) has been binned down by a factor of 3 to maximize the number of photons per spectral bin while preserving the line shapes and is represented by the vertical error bars. (c, d) Atmospheric transit depth variability relative to noise (left *Y*-axis) and atmospheric thickness variability relative to noise (right *Y*-axis) for 50 transits and each of the four GCMs for Hab 1 and Hab 2, respectively.

have shown that the predicted spectra are similar between the GCMs and the number of transits to detect an atmosphere with a 5σ confidence level is close for the Ben 1 (within four transits or 24%) and Ben 2 (within two transits or 33%) cases.

Concerning the aquaplanet scenarios, presented in Part II (Sergeev et al. 2022) of this work, we have shown that in the terminator region, in the Hab 1 case, the ExoCAM water cloud mixing ratio (only true for liquid water) and altitude of the cloud deck are much higher than for LMD-G, ROCKE-3D, and the UM, in that order. In the Hab 2 case, ExoCAM has on average higher thick clouds, followed by ROCKE-3D and closely by LMD-G and finally the UM, with the relative difference between the three latter models being smaller than for the Hab 1 case. The large cloud mixing ratio for LMD-G

was at first counterintuitive, as it employs a convective adjustment that notoriously produces fewer convective clouds than the mass-flux scheme used by the three other models, as also shown in Yang et al. (2013) and Sergeev et al. (2022). Here, at the terminator, the fact that LMD-G is cloudier than ROCKE-3D and the UM is likely due to a larger production of stratiform clouds. Unfortunately, the THAI protocol did not include separate output for convective and stratiform clouds. This differentiation will be investigated in a future study.

The differences in the simulated cloud coverage between the models, along with changes in temperature and water vapor profile, lead to 41%–56% differences between the number of transits predicted to detect (at 5σ) molecular species in the Hab 1 and Hab 2 cases, respectively, using the atmospheric profile of a GCM or another. These differences are non-negligible, as they can change by about 1σ the confidence level of a predicted detectability of an atmosphere or increase the observing time by 41%–56%, potentially making a given observation proposal unfeasible. Without observational data, we do not know which model is closer to the truth, but comparing them against each other indicates whether the detectability estimate is optimistic or pessimistic. This work therefore provides for the first time a "GCM uncertainty error bar" of ∼50% that needs to be considered in future analysis of JWST spectra of TRAPPIST-1e. Namely, simulations of temperate rocky exoplanets with ExoCAM would likely produce higher and thicker clouds relative to other GCMs. As a result, a tool like PSG would give a higher number of transits required to detect such an atmosphere. On the other hand, using ROCKE-3D or the UM would give a lower number of transits because they would likely produce lower and thinner clouds. As for LMD-G, the number would be comparable to that of ExoCAM owing to LMD-G's colder upper atmosphere. It also seems that these relative differences increase with the atmospheric concentration of the gas (here CO_2) that is being retrieved. This is because more minor lines appear, and because they are shallow, they are more sensitive to the cloud properties predicted by a given GCM. However, a larger gas concentration generally leads to a lower average absolute number of transits. Note that with the four transits expected for NIRSpec Prism in the JWST GTO (program 1331, PI Nikole Lewis), we can expect an average of 2.6σ and 4.3σ for the dry atmospheres of Ben 1 and Ben 2 and an average of 2.0σ and 3.4σ for the moist and cloudy Hab 1 and Hab 2 atmospheres, respectively.

THAI has been well received by the community, as demonstrated by the attendance of 125 people at the THAI workshop and the 35 authors of the THAI workshop report (Fauchez et al. 2021). Due to extreme paucity of observational data, it is important for the exoplanet community to develop and maintain intercomparison frameworks to benchmark atmospheric models, improve physical parameterizations, and evaluate their sensitivity. In this context, THAI is the first step toward a larger framework of intercomparison for exoplanets, the Climates Using Interactive Suites of Intercomparisons Nested for Exoplanet Studies (CUISINES). Within the CUISINES framework, we hope to develop intercomparison projects similar to THAI for exoplanets other than TRAPPIST-1e using a hierarchy of numerical models. Ultimately, the goal of CUISINES is to provide the exoplanet community, on both the modeling and observational ends of the spectrum, with model benchmarks and recommendations for comparison with existing observations and for planning future ones.

The authors thank the two anonymous reviewers, whose comments helped to improve the quality and clarity of this manuscript. T.J.F., G.L.V., and M.J.W. acknowledge support from the GSFC Sellers Exoplanet Environments Collaboration (SEEC), which is funded in part by the NASA Planetary Science Divisions Internal Scientist Funding Model. D.E.S., I.A.B., J.M., and N.J.M. acknowledge use of the Monsoon system, a collaborative facility supplied under the Joint Weather and Climate Research Programme, a strategic partnership between the Met Office and the Natural Environment Research Council. We acknowledge support of the Met Office Academic Partnership secondment program. This work was partly supported by a Science and Technology Facilities Council Consolidated grant (ST/R000395/1), UKRI Future Leaders Fellowship (MR/T040866/1), and the Leverhulme Trust (RPG-2020-82).

This project has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Sklodowska-Curie grant agreement No. 832738/ESCAPE. M.T. thanks the Gruber Foundation for its generous support to this research. M.T. acknowledges support from the PORTAL BRAIN-be 2.0 BELSPO project. M.T. was granted access to the High-Performance Computing (HPC) resources of Centre Informatique National de l'Enseignement Supérieur (CINES) under the allocations Nos. A0020101167 and A0040110391 made by Grand Équipement National de Calcul Intensif (GENCI). This work has been carried out within the framework of the National Centre of Competence in Research PlanetS supported by the Swiss National Science Foundation. M.T. acknowledges the financial support of the SNSF. M.T. and F.F. thank the LMD-Generic Global Climate team for the teamwork development and improvement of the model. J.H.M. acknowledges funding from the NASA Habitable Worlds program under award 80NSSC 20K0230.

The authors acknowledge the help of Andrew Ackerman to set up the cloud diagnostics in ROCKE-3D.

T.F. thanks Avi Mandell for his helpful discussion on the observation noise. The THAI GCM intercomparison team is grateful to the Anong's Thai Cuisine restaurant in Laramie for hosting its first meeting on 2017 November 15.

Numerical experiments performed for this study required the use of supercomputers, which are energy-intensive facilities and thus have nonnegligible greenhouse gas emissions associated with them.

Software: MATPLOTLIB (Hunter 2007). PSG (Villanueva et al. 2018) is available online at https://psg.gsfc.nasa.gov/index.php. ExoCAM (Wolf & Toon 2015) is available on Github: https://github.com/storyofthewolf/ExoCAM. The Met Office Unified Model is available for use under licence; see http://www.metoffice.gov.uk/research/modeling-systems/unified-model. ROCKE-3D is public domain software and available for download for free from https://simplex.giss.nasa.gov/gcm/ROCKE-3D/. Annual tutorials for new users take place annually, whose recordings are freely available online at https://www.youtube.com/user/NASAGISStv/playlists?view=50&sort=dd&shelf_id=15. LMD-G is available upon request from Martin Turbet (martin.turbet@lmd.jussieu.fr) and François Forget (francois.forget@lmd.jussieu.fr).

Appendix
Data Accessibility

All our GCM THAI data are permanently available for download at https://ckan.emac.gsfc.nasa.gov/organization/thai, with variables described for each data set. If you use

those data, please cite this current paper and add the following statement: "THAI data have been obtained from https://ckan.emac.gsfc.nasa.gov/organization/thai, a data repository of the Sellers Exoplanet Environments Collaboration (SEEC), which is funded in part by the NASA Planetary Science Divisions Internal Scientist Funding Model."

Scripts to process the THAI data are available on GitHub: https://github.com/projectcuisines.

Scripts to generate PSG/GlobES spectra are available on GitHub: https://github.com/nasapsg/globes.

It has been approved by editor Brian Jackson. THAI terminator atmospheric outputs (ASCII) averaged over the terminator latitudes, the 100 orbits and the four GCMs can be found, along with their corresponding averaged transmission spectra, at https://ckan.emac.gsfc.nasa.gov/organization/thai.

ORCID iDs

Thomas J. Fauchez https://orcid.org/0000-0002-5967-9631
Geronimo L. Villanueva https://orcid.org/0000-0002-2662-5776
Denis E. Sergeev https://orcid.org/0000-0001-8832-5288
Martin Turbet https://orcid.org/0000-0003-2260-9856
Ian A. Boutle https://orcid.org/0000-0002-1485-4475
Kostas Tsigaridis https://orcid.org/0000-0001-5328-819X
Michael J. Way https://orcid.org/0000-0003-3728-0475
Eric T. Wolf https://orcid.org/0000-0002-7188-1648
Shawn D. Domagal-Goldman https://orcid.org/0000-0003-0354-9325
François Forget https://orcid.org/0000-0002-3262-4366
Jacob Haqq-Misra https://orcid.org/0000-0003-4346-2611
Ravi K. Kopparapu https://orcid.org/0000-0002-5893-2471
James Manners https://orcid.org/0000-0003-4402-6811
Nathan J. Mayne https://orcid.org/0000-0001-6707-4563

References

Agol, E., Dorn, C., Grimm, S. L., et al. 2021, PSJ, 2, 1
Amundsen, D. S., Mayne, N. J., Baraffe, I., et al. 2016, A&A, 595, A36
Batalha, N. E., Lewis, N. K., Line, M. R., Valenti, J., & Stevenson, K. 2018, ApJ, 856, L34
Boutle, I. A., Joshi, M., Lambert, F. H., et al. 2020, NatCo, 11, 2731
Boutle, I. A., Mayne, N. J., Drummond, B., et al. 2017, A&A, 601, A120
de Wit, J., Wakeford, H. R., Gillon, M., et al. 2016, Natur, 537, 69
de Wit, J., Wakeford, H. R., Lewis, N. K., et al. 2018, NatAs, 2, 214
Douglas, E. S., Ashcraft, J. N., Belikov, R., et al. 2020, Proc. SPIE, 11443, 1144338
Fauchez, T. J., Turbet, M., Sergeev, D. E., et al. 2021, PSJ, 2, 106
Fauchez, T. J., Turbet, M., Villanueva, G. L., et al. 2019, ApJ, 887, 194
Fauchez, T. J., Turbet, M., Wolf, E. T., et al. 2020a, GMD, 13, 707
Fauchez, T. J., Villanueva, G. L., Schwieterman, E. W., et al. 2020b, NatAs, 4, 372
Gaudi, B. S., Seager, S., Mennesson, B., et al. 2018, arXiv:1809.09674
Gillon, M., Jehin, E., Lederer, S. M., et al. 2016, Natur, 533, 221
Gillon, M., Meadows, V., Agol, E., et al. 2020, arXiv:2002.04798
Gillon, M., Triaud, A. H. M. J., Demory, B.-O., et al. 2017, Natur, 542, 456
Gordon, I., Rothman, L., Hargreaves, R., et al. 2022, JQSRT, 277, 107949
Grimm, S. L., Demory, B.-O., Gillon, M., et al. 2018, A&A, 613, A68
Haqq-Misra, J., Wolf, E. T., Joshi, M., Zhang, X., & Kopparapu, R. K. 2018, ApJ, 852, 67
Hori, Y., & Ogihara, M. 2020, ApJ, 889, 77
Hunter, J. D. 2007, CSE, 9, 90
Kane, S. R., Jansen, T., Fauchez, T., Selsis, F., & Ceja, A. Y. 2021, AJ, 161, 53
Karman, T., Gordon, I. E., van der Avoird, A., et al. 2019, Icar, 328, 160
Keller-Rudek, H., Moortgat, G. K., Sander, R., & Sörensen, R. 2013, ESSD, 5, 365
Kofman, V., & Villanueva, G. L. 2021, JQSRT, 270, 107708
Koll, D. D. B., & Abbot, D. S. 2016, ApJ, 825, 99
Koll, D. D. B., Malik, M., Mansfield, M., et al. 2019, ApJ, 886, 140
Komacek, T. D., Fauchez, T. J., Wolf, E. T., & Abbot, D. S. 2020, ApJ, 888, L20
Komacek, T. D., & Showman, A. P. 2019, ApJ, 888, 2
Kopparapu, R. K., Ramirez, R., Kasting, J. F., et al. 2013, ApJ, 765, 131
Krishnamurthy, V., Hirano, T., Stefánsson, G., et al. 2021, AJ, 162, 82
Lewis, N., Clampin, M., Mountain, M., et al. 2017, JWST Proposal. Cycle 1, ID. #1331
Lewis, N. T., Lambert, F. H., Boutle, I. A., et al. 2018, ApJ, 854, 171
Lin, Z., MacDonald, R. J., Kaltenegger, L., & Wilson, D. J. 2021, MNRAS, 505, 3562
Lincowski, A. P., Meadows, V. S., Crisp, D., et al. 2018, ApJ, 867, 76
Luger, R., Sestovic, M., Kruse, E., et al. 2017, NatAs, 1, 0129
Lustig-Yaeger, J., Meadows, V. S., & Lincowski, A. P. 2019, AJ, 158, 27
Massie, S. T., & Hervig, M. 2013, JQSRT, 130, 373
May, E. M., Taylor, J., Komacek, T. D., Line, M. R., & Parmentier, V. 2021, ApJL, 911, L30
Mayne, N. J., Baraffe, I., Acreman, D. M., et al. 2014, GMD, 7, 3059
Mlawer, E. J., Payne, V. H., Moncet, J. L., et al. 2012, RSPTA, 370, 2520
Moran, S. E., Hörst, S. M., Batalha, N. E., Lewis, N. K., & Wakeford, H. R. 2018, AJ, 156, 252
Morley, C. V., Kreidberg, L., Rustamkulov, Z., Robinson, T., & Fortney, J. J. 2017, ApJ, 850, 121
Payne, V. H., Drouin, B. J., Oyafuso, F., et al. 2020, JQSRT, 255, 107217
Pidhorodetska, D., Fauchez, T. J., Villanueva, G. L., Domagal-Goldman, S. D., & Kopparapu, R. K. 2020, ApJ, 898, L33
Pierrehumbert, R. T. 2010, Principles of Planetary Climate (Cambridge: Cambridge Univ. Press)
Rustamkulov, Z., Sing, D., Liu, R., & Wang, A. 2022, ApJL, 928, L7
Selsis, F., Wordsworth, R. D., & Forget, F. 2011, A&A, 532, A1
Sergeev, D. E., Fauchez, T. J., Turbet, M., et al. 2022, PSJ, 3, 212
Sneep, M., & Ubachs, W. 2005, JQSRT, 92, 293
Suissa, G., Mandell, A. M., Wolf, E. T., et al. 2020a, ApJ, 891, 58
Suissa, G., Wolf, E. T., Kumar Kopparapu, R., et al. 2020b, AJ, 160, 118
Team, T. L. 2019, arXiv:1912.06219
Tolkien, J. R. R. 1955, The Return of the King (London: George Allen & Unwin)
Turbet, M., Bolmont, E., Bourrier, V., et al. 2020, SSRv, 216, 100
Turbet, M., Bolmont, E., Leconte, J., et al. 2018, A&A, 612, A86
Turbet, M., Fauchez, T. J., Sergeev, D. E., et al. 2022, PSJ, 3, 211
Turbet, M., Leconte, J., Selsis, F., et al. 2016, A&A, 596, A112
Villanueva, G. L., Liuzzi, G., Faggi, S., et al. 2022, Fundamentals of the Planetary Spectrum Generator (Geronimo Villanueva)
Villanueva, G. L., Smith, M. D., Protopapa, S., Faggi, S., & Mandell, A. M. 2018, JQSRT, 217, 86
von Paris, P., Gratier, P., Bordé, P., & Selsis, F. 2016, A&A, 587, A149
Warren, S. G. 1984, ApOpt, 23, 1206
Way, M. J., Aleinov, I., Amundsen, D. S., et al. 2017, ApJS, 231, 12
Wolf, E. T., Kopparapu, R., Haqq-Misra, J., & Fauchez, T. J. 2022, PSJ, 3, 7
Wolf, E. T., Kopparapu, R. K., & Haqq-Misra, J. 2019, ApJ, 877, 35
Wolf, E. T., & Toon, O. B. 2015, JGRD, 120, 5775
Wordsworth, R. D., Forget, F., Selsis, F., et al. 2011, ApJL, 733, L48
Wunderlich, F., Godolt, M., Grenfell, J. L., et al. 2019, A&A, 624, A49
Wunderlich, F., Scheucher, M., Godolt, M., et al. 2020, ApJ, 901, 126
Yang, J., Cowan, N. B., & Abbot, D. S. 2013, ApJL, 771, L45
Yang, J., Leconte, J., Wolf, E. T., et al. 2019, ApJ, 875, 46