Detecting biosignatures on habitable rocky worlds with ELT/ANDES

Evann Kurzawa Ferrandez [ID],[1] Aaron Bello-Arufe [ID],[1] and Renyu Hu [ID][1,2,3,4]

[1] Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109, USA
[2] Department of Astronomy & Astrophysics, The Pennsylvania State University, University Park, PA 16802, USA
[3] Center for Exoplanets and Habitable Worlds, The Pennsylvania State University, University Park, PA 16802, USA
[4] Institute for Computational and Data Science, The Pennsylvania State University, University Park, PA 16802, USA

ABSTRACT

The search for life beyond the Solar System is at a turning point, transitioning from theoretical predictions to observations enabled by next-generation observatories. The Extremely Large Telescope (ELT) will host the ArmazoNes high Dispersion Echelle Spectrograph (ANDES), optimized for visible to near-infrared high-resolution spectroscopy. We present an end-to-end simulation–detection pipeline to evaluate the ability of ANDES to detect CO_2, H_2O, and potential biosignature gases O_2 and CH_4 in the transmission spectra of potentially habitable rocky exoplanets. Assuming cloud-free, modern Earth-like atmospheres, we model the transmission spectra using realistic noise estimates from the ANDES Exposure Time Calculator. We introduce a novel Bayesian cross-correlation function (CCF) framework that incorporates molecule-specific kernels and a new autoregressive model to account for correlations in the CCF. We apply our framework to 19 potentially habitable transiting exoplanets observable with the ELT and estimate the number of transits required for a very strong detection ($\Delta B \geq 1.5$). We find that CH_4 and H_2O are the most accessible species, potentially detectable in $\lesssim 5$ transits for TRAPPIST-1 d and e. CO_2 requires at least 16 transits in the most favorable target, TRAPPIST-1 d, whereas O_2 remains undetectable in < 38 transits. For eight planets (TRAPPIST-1 d, e, f, g; K2-288 B b; LHS-1140 b, and TOI-700 d, e), at least one gas may be detectable in $\lesssim 20$ transits. Because our framework assumes favorable observing conditions and perfect data detrending, the estimates here should therefore be interpreted as best-case scenarios but provide an initial benchmark for the search for biosignatures with ANDES.

1. INTRODUCTION

With a 39-meter diameter aperture, the ELT will be the largest optical and near-infrared telescope in the world, equipped with a wide range of instruments. These include ANDES (Maiolino et al. 2013; Marconi et al. 2022; Palle et al. 2023; Marconi et al. 2024; Roederer et al. 2024; Martins et al. 2024; D'Odorico et al. 2024), which will consist of three fiber-fed spectrographs ($R \sim 100{,}000$) to simultaneously cover a wavelength range from $0.5\,\mu m$ to $1.8\,\mu m$, with an extended goal of spanning $0.38\,\mu m$ to $2.4\,\mu m$.

One of the main goals of ANDES is to search for detecting biosignature gases in exoplanets through high-resolution transmission spectroscopy (Palle et al. 2023). Among the gases potentially linked to planetary habitability, O_2, CO_2, CH_4 and H_2O are of primary importance. Specifically, CO_2 and H_2O govern greenhouse warming and water cycling for climate stability, while O_2 and CH_4 are short-lived and may signal active biological replenishment (Seager et al. 2012; Schwieterman et al. 2018; Ramirez 2018; Meadows et al. 2018; Thompson et al. 2022). Each of these four gases has distinct absorption features in the visible and/or near-infrared, making them prime candidates for searching for potentially habitable environments with ANDES. The instrument's high resolving power ($R \sim 100{,}000$), combined with the ELT's extremely large aperture, will provide unprecedented-quality data on potentially habitable exoplanets (Kawahara et al. 2012; Snellen et al. 2013), far exceeding the typical spectral resolution of JWST spectrographs ($R \sim 100 - -1000$) (e.g., Morley et al. 2017;

Lustig-Yaeger et al. 2019; Pidhorodetska et al. 2020; Seager et al. 2025).

Over the past decade, significant theoretical and observational advancements have been made, particularly in the use of high-resolution spectroscopy and transit observations to probe the atmospheres of rocky planets orbiting M dwarfs. Early studies, such as Snellen et al. (2013), proposed employing high-resolution spectroscopy ($R \sim 100,000$) and cross-correlation techniques to search for biosignature gases with extremely large telescopes such as the ELT. They estimated that detecting O_2 in the transmission spectrum of an Earth-analog orbiting an M5V star at 12 pc could be feasible after observing 30 transits. However, challenges such as the limited knowledge of the occurrence rate of habitable-zone Earth-sized planets around M-dwarfs and the need to discover bright, nearby targets highlighted the difficulty of such endeavors. Rodler & López-Morales (2014) revisited the estimates from Snellen et al. (2013) and found far less optimistic results, with O_2 requiring up to 175 hr ($\sim$74 transits) to be detected on a target at 10 pc. The authors concluded that close-by M4V stars are the most viable candidates, but even for these stars, achieving a significant detection remains challenging. Serindag & Snellen (2019) reinforced these findings by simulating O_2 detection in real observational data of Proxima Centauri, estimating the need for 20-50 transits for successful detection in optimal scenarios. Later, Currie et al. (2023) examined both inhabited and abiotic scenarios around M dwarfs and found that CO_2 and CH_4 are generally more accessible than O_2, which requires substantially more observational effort, particularly for late-type host stars; they also argued that molecules such as O_3 and C_2H_6 are unlikely to be accessible with ground-based high-resolution transmission spectroscopy.

In parallel, large-scale surveys and theoretical simulations have advanced our understanding of the detectability of biosignature gases. Hardegree-Ullman et al. (2023) expanded on earlier work by creating a catalog of over 286,000 main-sequence stars within 120 pc and simulating a survey of M dwarfs within 20 pc. They concluded that detecting Earth-like O_2 levels within 50 yr is feasible for up to 21% of nearby M dwarf systems with suitable transiting planets, particularly if high-resolution spectrographs ($R \sim 500,000$) are utilized. They also noted that Earth-sized habitable-zone planets within 20 pc are far more numerous in nontransiting configurations than in transiting ones, highlighting the importance of observational techniques beyond the transit method. In response, several studies have explored the potential of direct imaging with next-generation ground-based observatories, focusing on nearby exoplanets. Zhang et al. (2024) demonstrated that instruments such as ELT/METIS and ELT/HARMONI could detect key atmospheric species, such as CH_4, CO_2, H_2O and O_2, in nearby systems. Complementarily, Hardegree-Ullman et al. (2025) investigated the detectability of Earth-like O_2 levels in hypothetical habitable-zone exo-Earth candidates within 20 pc. Under optimistic assumptions and a 10 yr survey baseline, they found that O_2 at Earth-like abundances could be probed for up to $\sim$7 and $\sim$19 candidates using the GMT and ELT, respectively.

A complementary approach is to combine high-resolution spectroscopy with high-contrast imaging, often referred to as high-dispersion coronagraphy (HDC). In this technique, an adaptive-optics coronagraph spatially suppresses stellar light, and a high-resolution spectrograph disentangles the planet's Doppler-shifted molecular lines from the stellar spectrum (Snellen et al. 2015; Wang et al. 2017; Mawet et al. 2017). HDC is being developed on the ground (for example, with Keck/KPIC) and investigated for future space missions such as the Habitable Worlds Observatory (Wang et al. 2021, 2018; Pueyo et al. 2019).

Although ground-based observations are affected by telluric contamination, they provide access to much larger collecting areas and spectral resolutions that remain unmatched by space telescopes. Two strategies for a comprehensive characterization of exoplanetary atmospheres would be (i) combining ground- and space-based observations, as demonstrated by Brogi et al. (2017) and followed in other studies (e.g., Kasper et al. 2023; Smith et al. 2024), and (ii) increasing spectral resolution to facilitate the disentanglement of the planetary signal from the telluric and host star spectra. López-Morales et al. (2019) demonstrated that raising the resolution from $R \sim 100,000$ to $R \sim 300,000 - 400,000$ can double the depth of O_2 lines in Earth-like atmospheres, reducing the required number of transits by over 30%. Similarly, Rukdee (2024) showed that even under the most challenging haze and cloud conditions, increasing the resolution from R $\sim 100,000$ to R $\sim 300,000$ enables robust detections with exposure times reduced by a factor of four, thereby requiring significantly fewer transits.

ANDES represents the most advanced and promising ground-based instrument currently planned for the search of biosignature gases (Palle et al. 2023; Dubey & Majumdar 2024; Dubey et al. 2025). However, an in-depth quantitative study of the number of transits required to detect potential biosignatures on rocky exoplanets with ANDES, while accurately modeling the noise properties of the instrument, has yet to be published. This is the gap that this work aims to address.

Our paper is structured as follows. In Section 2, we present the assumptions and models used to simulate the atmospheres of known transiting habitable-zone terrestrial planets. In Section 3, we generate the transmission spectra using the `petitRADTRANS` library (Mollière et al. 2019). Section 4 describes how noise is added to the spectra and the parameters selected for the ANDES Exposure Time Calculator and the observing setup. In Section 5, we explore the limitations of the CCF analysis, which do not yield conclusive results for our sample. This motivates a more robust approach, introduced in Section 6, which details a custom Bayesian framework with `MultiNest` (Feroz et al. 2009), featuring interpolated kernels that mimic molecule-specific autocorrelation patterns and a CCF-correlation mitigation scheme that inflates uncertainties under a first-order autoregressive noise model. Evidence estimation and Bayes factors are then used to quantify support for atmospheric-signal detection. In Section 7, we apply this method to the planets and gases of interest, and the results are interpreted and compared in Section 8. Finally, in Section 9 we conclude and provide recommendations for ELT/ANDES observing strategies.

2. SELECTED MODEL FOR SIMULATIONS

The Habitable Worlds Catalog[1] (HWC; PHL@UPR-Arecibo, 2024), recently updated in March 2024, identifies 70 potentially habitable exoplanets out of more than 6,000 discovered to date. Our analysis focuses on the 41 HWC planets with confirmed transits, out of a total of 70, and excludes the 11 that never exceed the ELT's 20° altitude cutoff (see Section 2.3), ensuring that all remaining targets reach sufficient altitude for observation. We further exclude 11 planets with radii larger than $2R_\oplus$, as planets above this threshold are unlikely to be rocky (Fulton et al. 2017).

2.1. Global setup

The equilibrium temperature $T_{\rm eq}$ of each planet is estimated from blackbody theory with zero Bond albedo. Bulk planetary properties are listed in Table 1 and are taken from the NASA Exoplanet Archive[2] (Akeson et al. 2013). Planetary masses are estimated using the Bayesian radius-density-mass relation of Parviainen et al. (2024) via the `SPRIGHT` code. We use the maximum a posteriori estimate of the predicted mass distribution, which together with the radius defines the surface gravity g. The planetary reference radius R_p is defined at a reference pressure of 1 bar. The vertical grid spans the

pressures $\log_{10}(p\,({\rm bar})) \in [-7, 0]$ using 100 layers, from 1 bar down to 10^{-7} bar ($\sim$120 km altitude for Earth-like conditions). In practice, we restrict radiative transfer to a single absorber at a time, which means that potential overlaps with other species are neglected.

2.2. Atmospheric model: structure and composition

1. **Temperature profile.** We adopt a two-layer temperature-pressure profile, a standard approximation in radiative-convective equilibrium models (e.g., Kasting et al. 1993; Kopparapu et al. 2013). The tropopause is set at 0.1 bar (Robinson & Catling 2014). Above this, the stratosphere is assumed to be isothermal at the atmospheric skin temperature, described by e.g., Pierrehumbert (2010); Parmentier & Guillot (2014):

$$T_{\rm strat} = T_{\rm skin} \approx 2^{-1/4}\,T_{\rm eq} \qquad (1)$$

Below the tropopause, we assume a dry adiabatic temperature–pressure relation:

$$T(p) = T_{\rm skin}\left(\frac{p}{p_{\rm trop}}\right)^{\kappa} \qquad (2)$$

where $\kappa = R/c_p$ is the adiabatic exponent of the bulk atmosphere, set to 0.286 (Catling & Kasting 2017). This prescription ensures a continuous transition between the radiative and convective layers. The surface temperature is thus the adiabatic extension of $T_{\rm skin}$ to 1 bar.

2. **Atmospheric VMR.** The atmospheric composition is specified as follows:

   • H_2O: the stratosphere is fixed to a constant mixing ratio of of 3 ppmv (e.g., NASA conference 1989; Oman et al. 2008), while in the troposphere the abundance follows the saturation value given by the Clausius-Clapeyron relation, capped at 1% to maintain humidity levels generally consistent with Earth-like conditions (e.g., NOAA 2025). The adopted prescription is:

$$x_{\rm H_2O}(p) = \begin{cases} 3\,{\rm ppmv} & p \le p_{\rm trop} \\ \min\left(\frac{e_s(T(p))}{p}, 1\%\right) & p > p_{\rm trop} \end{cases} \qquad (3)$$

   The saturation vapor pressure $e_s(T)$ is computed as

$$e_s(T) = e_0 \exp\left[\frac{L_v}{R_v}\left(\frac{1}{273.15\,{\rm K}} - \frac{1}{T}\right)\right] \qquad (4)$$

---

[1] http://phl.upr.edu/hwc

[2] https://exoplanetarchive.ipac.caltech.edu/

Table 1. Input parameters for `petitRADTRANS` and the ETC.

Planet	a	i_p	P	T_{14}	T_{eq}	R_p	R_*	M_p	T_*	v_{sys}	v_{bary}	Sp. T	m_J	References
	(au)	(°)	(days)	(hr)	(K)	($R_\oplus$)	($R_\odot$)	($M_\oplus$)	(K)	(km s^{-1})	(km s^{-1})		(Vega)	
K2-288B b	0.164	89.810	31.393	2.100	225[†]	1.900	0.320	4.155[†]	3341	71,60	-30,26[†]	M3V	11.91	1,2
K2-3 d	0.210	89.790	44.556	4.170	313[†]	1.620	0.600	5.217[†]	3835	30,24	29,39[†]	M0V	9.42	3,4,5,6,7
K2-72 e	0.106	89.680	24.159	2.250	297[†]	1.290	0.329	2.211[†]	3498	-42,92	0,69[†]	M2V	11.69	3,4,8,9,10
Kepler-1649 c	0.083	89.339	19.535	1.070	261[†]	1.060	0.232	1.101[†]	3240	0,00[†]	-13,61[†]	M5V	13.38	3,11,12
Kepler-186 f	0.432	89.960	129.943	5.448	211[†]	1.438	0.548	3.306[†]	3876	-60,96	-7,58[†]	M1	12.47	3,4,9,13,14,15,16,17
Kepler-440 b	0.242	89.930	101.111	8.130	311[†]	1.899	0.663	4.140[†]	3891	-20,61	-7,23[†]	K8V[†]	12.96	3,4,9,13,14,15,16
Kepler-442 b	0.409	89.940	112.303	5.630	271[†]	1.395	0.619	2.958[†]	4563	-76,91	-6,72[†]	K4V[†]	13.23	3,4,9,13,14,15,16
Kepler-452 b	1.046	89.806	384.843	10.367	286[†]	1.511	1.119	3.983[†]	5728	-25,49	-7,80[†]	G2	12.26	3,4,9,15,16,18
Kepler-62 e	0.428	89.980	122.386	7.196	305[†]	1.873	0.731	3.971[†]	4842	17,51	-9,65[†]	K2V	12.26	3,4,9,15,19,20
Kepler-62 f	0.720	89.900	267.283	7.460	235[†]	1.540	0.731	4.291[†]	4842	17,51	-9,65[†]	K2V	12.26	3,4,9,19,20,21
LHS 1140 b	0.095	89.860	24.737	2.150	225[†]	1.730	0.216	6.764[†]	3096	-13,74	-28,80[†]	M4,5V	9.61	3,4,22
LP 890-9 c	0.040	89.287	8.457	0.959	271[†]	1.367	0.156	2.740[†]	2850	28,84	19,99[†]	M6V	12.26	3,4,23
TOI-700 d	0.163	89.800	37.424	3.314	268[†]	1.073	0.421	1.146[†]	3459	-4,78	1,78[†]	M2,5V	9.47	3,24,25
TOI-700 e	0.134	89.600	27.810	2.777	296[†]	0.953	0.421	0.757[†]	3459	-4,78	1,78[†]	M2,5V	9.47	3,24,25
TOI-715 b	0.083	89.860	19.288	1.980	252[†]	1.550	0.240	4.395[†]	3075	55,78	-4,49[†]	M4	11.81	3,4,26
TRAPPIST-1 d	0.022	89.896	4.049	0.815	288[†]	0.788	0.119	0.394[†]	2566	-52,00	0,22[†]	M7,5	11.35	3,27,28,29
TRAPPIST-1 e	0.029	89.793	6.101	0.929	251[†]	0.920	0.119	0.671[†]	2566	-52,00	0,22[†]	M7,5	11.35	3,27,28,29
TRAPPIST-1 f	0.038	89.740	9.208	1.048	219[†]	1.045	0.119	1.044[†]	2566	-52,00	0,22[†]	M7,5	11.35	3,27,28,29
TRAPPIST-1 g	0.047	89.742	12.352	1.137	197[†]	1.129	0.119	1.373[†]	2566	-52,00	0,22[†]	M7,5	11.35	3,27,28,29

References— (1) Feinstein et al. 2019; (2) USNO-B Catalog: Monet et al. 2003; (3) 2MASS Catalog; Cutri et al. 2003; (4) Gaia Collaboration 2023; (5) Fukui et al. 2016; (6) Bonomo et al. 2023; (7) Diamond-Lowe et al. 2022; (8) Dressing et al. 2017; (9) TESS Input Catalog: Stassun et al. 2019; (10) Crossfield et al. 2016; (11) Vanderburg et al. 2020; (12) Hardegree-Ullman et al. 2019; (13) Torres et al. 2015; (14) Gajdoš et al. 2019; (15) Morton et al. 2016; (16) Berger et al. 2018; (17) Quintana et al. 2014; (18) Jenkins et al. 2015; (19) Weiss et al. 2024; (20) Borucki et al. 2013; (21) Borucki et al. 2019; (22) Cadieux et al. 2024a; (23) Delrez et al. 2022; (24) Gilbert et al. 2023; (25) Abdurro'uf et al. 2022; (26) Dransfield et al. 2024; (27) Agol et al. 2021; (28) Jönsson et al. 2020; (29) Gizis et al. 2000.

NOTE—a is the semi-major axis, i_p is the inclination, P is the orbital period, T_{14} is the transit duration, T_{eq} is the equilibrium temperature, R_p is the planet radius, R_* is the host star radius, M_p is the planet mass, T_* is the host star effective temperature, v_{sys} is the planet's host star systemic velocity, v_{bary} is the barycentric correction that maximises $|v_{sys} - v_{bary}|$ over the course of a year, Sp. T is the host star spectral type, and m_J is the Vega magnitude of the host star in the J band. When v_{sys} is unavailable, a default value of $0\,\mathrm{km\,s^{-1}}$ is assumed. Values marked with a dagger (†) were calculated in this work. These data are partially obtained from the SIMBAD database and the NASA Exoplanet Archive.

with e_0 the reference vapor pressure at 273.15 K, L_v the latent heat of vaporization, and R_v the specific gas constant for water vapor.

- O_3: described by a log-normal distribution in pressure as in von Paris et al. (2011, 2013), with peak mixing ratio 10 ppmv centered at $p = 10$ mbar, width $\sigma_{\ln p} = 0.5$, and a floor of 10 ppbv elsewhere (e.g., Bekki & Lefevre 2009; Kuttippurath et al. 2024).

- CH_4: well-mixed in the troposphere at 1.9 ppmv (Lan et al. 2025). Above the tropopause, the abundance decreases exponentially, as $(p/p_{trop})^{0.5}$, as in Bainbridge & Heidt (1966).

- CO: we simplify the CO treatment of Pan et al. (2004); El Amraoui et al. (2014) to a

two-level profile: 100 ppbv in the troposphere and 50 ppbv in the stratosphere.

- O_2: taken as a constant value of 0.2095 at all pressures, corresponding to the standard dry Earth atmosphere

- CO_2: taken as a constant value of 420 ppmv at all pressures (Lan et al. 2025).

- N_2: fills the remaining fraction at each layer so that the total abundance sums to unity.

At each layer, all mixing ratios are renormalized to ensure $\sum_i x_i = 1$, and the mean molecular mass $\mu(p)$ is computed consistently.

3. **Opacity.** Rayleigh scattering by N_2 is included in all simulations; clouds are not modeled. We neglect atmospheric refraction under the conditions

relevant to our targets (temperate-warm, close-in rocky planets around K–M dwarfs). In this regime the large stellar angular size pushes the refractive boundary to pressures/altitudes where gas and Rayleigh opacity dominate the slant optical depth, so refraction makes only a minor contribution and does not set the continuum for our targets. This behavior is captured by analytic refractive-boundary estimates and validated against ray-tracing comparisons (Robinson et al. 2017); for the Earth–Sun case, refraction excludes only the lowest $\sim$13 km and changes the effective radius by at most a few tens of km across the optical/NIR (Bétrémieux & Kaltenegger 2013). The resulting vertical profiles of temperature and volume mixing ratios (VMR) are displayed in Figure 1.

2.3. Visibility constraints and planetary velocity

The 2011 ELT construction proposal[3] specifies a minimum altitude cutoff of 20° (above the local horizon at the ELT site) to avoid rapid increases in atmospheric extinction, dome vignetting, and seeing degradation that arise at lower altitudes. Given the celestial positions of certain transiting stars within the HWC, the first step of our model is to identify and exclude stars whose apparent altitude never exceeds the 20° threshold throughout the year, rendering them permanently unobservable from the ELT site.

To evaluate the observational visibility of each transiting planet from the HWC, we generate a uniform grid of 366 daily epochs spanning January 1, 2025, to January 1, 2026. We then compute the altitude and azimuth of each star, by transforming their coordinates from the International Celestial Reference System (ICRS) to the ELT's Altitude-Azimuth (AltAz) frame. Subsequently, we calculate Earth's barycentric velocity corrections across this temporal grid for each star using a DE430 ephemeris combined with the `barycorrpy` (Kanodia & Wright 2018) Python package, following the algorithm of Wright & Eastman (2014).

We then apply the 20° altitude mask to the resulting time series, retaining only the days when the stellar altitude exceeded this threshold. Consequently, only planets whose host stars met this minimum altitude at least once per year are retained for further analysis. This approach ensures that our simulated spectroscopic observations are limited to stars observable from the ELT site. Coordinate transformations and obser-

vational epoch grids are generated using the Python library `Astropy` (Astropy Collaboration 2022). A full-year barycentric correction time series is generated for each exoplanet whose host star reached the minimum altitude at least once. An example of such a time series is presented in Figure 2.

Finally, we adopt an intentionally optimistic configuration that maximizes the relative Doppler shift between the exoplanet host star and the Earth, to facilitate disentangling exoplanetary signals from telluric lines. Specifically, for each host star we identified, over the course of a year, the time when the magnitude of the barycentric correction relative to the systemic velocity is largest, and we denote this maximum separation by

$$v_{\mathrm{max}} = \max \left| v_{\mathrm{sys}} - v_{\mathrm{bary}}(t) \right| \qquad (5)$$

The systemic velocities and the corresponding maximum-separation barycentric velocity corrections are summarized in Table 1.

3. BUILDING EXOPLANET TRANSMISSION SPECTRA

We generate model transmission spectra with `petitRADTRANS` (Mollière et al. 2019), which performs radiative transfer for planetary atmospheres. For each target, we use published HWC system parameters together with our atmospheric model as inputs.

Because our analysis relies on high-resolution spectroscopy, we adopt molecular opacity data at commensurate resolution. For H_2O, CO_2 and CO, we use the HITEMP line lists (Rothman et al. 2010); for CH_4 we use the HITEMP update by Hargreaves et al. (2020); for O_3, the HITRAN database (Rothman et al. 2013); for O_2, the HITRAN compilation (Gordon et al. 2022); and for N_2, the WCCRMT database (Western et al. 2018). The H_2O, CO_2, CO, O_3, and CH_4 opacity files are obtained in HDF5 format from the Keeper knowledge base of the Max Planck Digital Library,[4] while those for O_2 and N_2 are binary files from the DACE Opacity Database,[5] which employs the open-source `HELIOS-K` code (Grimm & Heng 2015; Grimm et al. 2021) to convert line lists into opacity coefficients suitable for atmospheric simulations. Each binary file corresponds to a specific temperature, pressure, and wavenumber range; which we merge into a single line-by-line file readable by `petitRADTRANS`.

The spectra are calculated over a duration twice that of the transit, centered on the transit midpoint. Prior
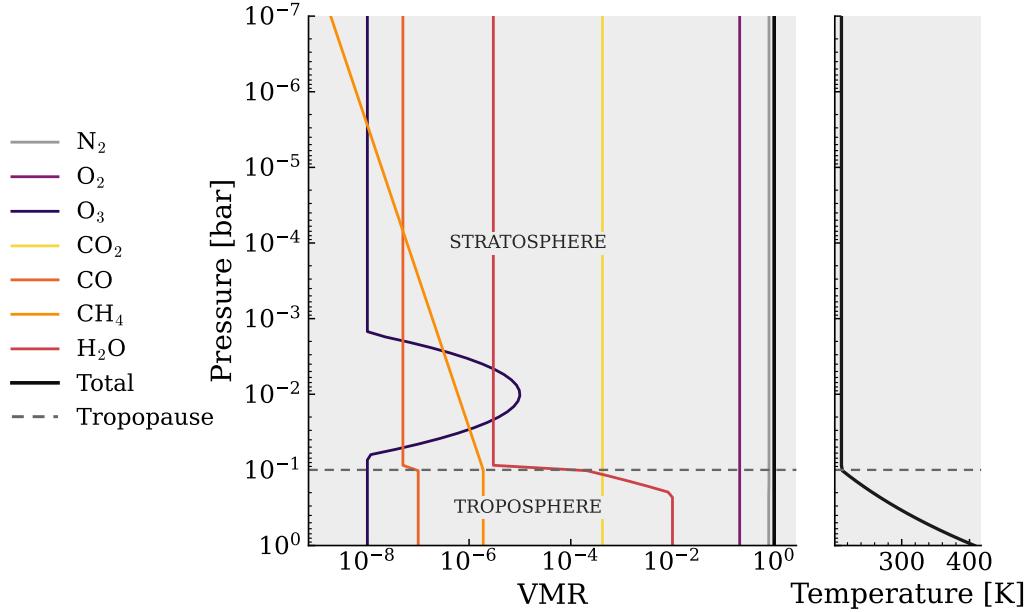

Figure 1. Vertical profiles of temperature and volume mixing ratios for the species included in our atmospheric model, simulated for TRAPPIST-1 e. The profile behavior follows from the equations and assumptions developed in Section 2.2.
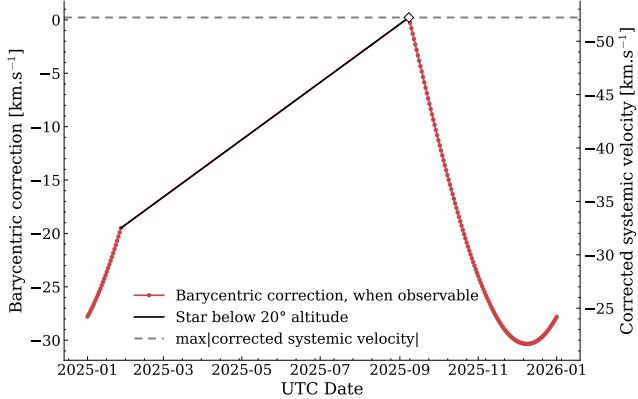

Figure 2. Barycentric velocity correction for TRAPPIST-1 from January 2025 to January 2026. The star is unobservable from the ELT between February and September, when its altitude remains below the $20°$ threshold. Outside this window, the barycentric correction varies between approximately $-30\,\mathrm{km\,s^{-1}}$ and $0\,\mathrm{km\,s^{-1}}$. The sinusoidal modulation during observable periods arises from the projection of Earth's nearly circular orbital velocity onto the line of sight to the star, producing a radial component with an amplitude of $\pm 30\,\mathrm{km\,s^{-1}}$ and a one-year period, phase-shifted according to the star's ecliptic position. The right-hand y-axis shows the relative Doppler separation $|v_{\mathrm{sys}} - v_{\mathrm{bary}}(t)|$, which reaches its maximum for TRAPPIST-1 at a barycentric velocity of $0.22\,\mathrm{km\,s^{-1}}$.

to resampling, we apply instrumental broadening by convolving, with each spectrum with a Gaussian kernel whose width is set by the ANDES resolving power.

The convolution is performed using the `PyAstronomy` library (Czesla et al. 2019). We choose to extend the Gaussian kernel to five standard deviations to ensure computational efficiency while adequately representing the instrument response. After broadening, we convert vacuum wavelengths to air wavelengths using the Edlén (1966) relation. The broadened spectra are then resampled to the ANDES resolving power ($R \sim 100{,}000$) on a grid sampled at three pixels per resolution element (Palle et al. 2023). We ignore other contributions to the broadening or Doppler-shifting of the planet's spectrum (e.g. planetary rotation and atmospheric dynamics).

We Doppler-shift the stellar spectrum by v_{max} and the planet's transmission spectrum by

$$v_{\mathrm{tot}}(t) = v_p(t) + v_{\mathrm{max}} \tag{6}$$

where $v_{\mathrm{tot}}(t)$ is the planet-ELT line-of-sight velocity during transit, $v_p(t)$ is the planet's radial velocity in the star-exoplanet frame, and v_{max} is the constant barycentric-systemic offset at the epoch of maximum Doppler separation (Equation 5). The planet's radial velocity is given by:

$$v_p(t) = K_p \sin(2\pi\phi(t)) \tag{7}$$

where $\phi(t)$ is the orbital phase at time t and K_p is the radial velocity semiamplitude of the planet,

$$K_p = \frac{2\pi a \sin(i_p)}{P\sqrt{1-e^2}} \tag{8}$$

where a is the orbital semi-major axis of the planet, i_p is the orbital inclination, P is the orbital period, and e is the eccentricity. We assume circular orbits in our simulations.

We model the host star spectra using the PHOENIX library (Husser et al. 2013). We do not apply rotational broadening, as the host stars are mostly late-type and generally rotate slowly. We then inject the Doppler-shifted planetary transmission spectrum during the in-transit window. The observed spectrum model is

$$S_{\mathrm{obs}}(\lambda) = S_{\mathrm{star}}(\lambda) \times (1 - S_{\mathrm{planet}}) \qquad (9)$$

where S_{star} is the PHOENIX stellar spectrum Doppler-shifted by v_{max} (Equation 5), and S_{planet} is the transmission spectrum shifted by v_{tot} (Equation 6) and set to zero outside of transit. S_{obs} is the composite {planet + host star} spectrum.

4. THEORETICAL OBSERVATION WITH THE ELT ANDES

After generating the spectral models, we add noise consistent with the expected performance of ANDES. Signal-to-noise ratios (S/N) under our observing conditions are obtained from the ANDES Exposure Time Calculator[6] (ETC; Sanna et al. 2024). We assume operating conditions of 283 K ambient temperature and an airmass of 1.0. For this study, we adopt an integration time of 10 min per observation. For each target we provide the host star's J-band Vega magnitude and spectral type, as listed in Table 1. Because the ETC relies on a finite set of templates from the Pickles (1998) stellar library, we approximate each true spectral type by the closest available option.

For a single wavelength and the observation of a compact source, the ETC predicts the expected signal-to-noise ratio as:

$$S/N = \sqrt{\frac{\left(\frac{e_S \cdot e_I \cdot e_T \cdot \beta \cdot S \cdot t}{R} \times 10^{\frac{16.85 - m_{AB}}{2.5}}\right)^2}{N_b^2 + N_d^2 + \frac{e_S \cdot e_I \cdot e_T \cdot \beta \cdot S \cdot t}{R} \times 10^{\frac{16.85 - m_{AB}}{2.5}}}} \qquad (10)$$

where e_S is the slit efficiency, i.e., the fraction of light from the astronomical target falling inside the spectrometer slit; e_T is the efficiency of the telescope outside the atmosphere; e_I is the efficiency of the instrument; β is the atmospheric transmission; S is the telescope effective collecting area; t is the exposure time; R is the resolving power; m_{AB} is the magnitude in AB units; N_b is the background noise and N_d is the detector noise.

[6] https://andes.inaf.it/instrument/exposure-time-calculator/

The detector noise and the background noise can be calculated using the following relations:

$$N_d = \sqrt{p_{\theta_S} \left(\frac{n \cdot N_{ro}^2}{p_b} + L_{dc} \cdot t\right)} \qquad (11)$$

$$N_b = \sqrt{\frac{\pi}{4} e_I e_T \beta S t \theta_S^2 (\sigma_{sky} + \sigma_{th})} \qquad (12)$$

where σ_{sky} and σ_{th} are defined as follows:

$$\sigma_{sky} = \frac{10^{\left(\frac{16.85 - L_{sky}(M)}{2.5}\right)}}{R} \qquad (13)$$

$$\sigma_{th} = \frac{1{,}4 \times 10^{12} \times L_{tel} \times e^{\left(\frac{-14388}{\lambda T}\right)}}{R\lambda^3} \qquad (14)$$

In these equations, p_{θ_S} is the number of pixels corresponding to the spectrometer aperture, p_b is the on-chip binning factor, n is the number of separate read-outs to achieve the requested exposure time, N_{ro} is the read-out noise of the detector, L_{dc} is the dark current of the detector, L_{sky} is the sky background brightness, M is the airmass ratio, L_{tel} is the total emissivity of the telescope and instrument, θ_S is the sky-projected angular diameter of the spectrometer aperture, and T is the ambient temperature. Although read-out noise is included in our calculations, we ignore the read-out overhead; for the multi-minute exposures considered here, the per-frame overhead is expected to be small (a few percent at most), so this approximation does not materially affect our results.

From the ANDES ETC we obtain per-pixel S/N and stellar spectra in multiple bandpasses spanning 0.35–2.45 μm, corresponding to the target's extended coverage rather than the baseline instrument range of 0.5–1.8 μm. The band-specific arrays are stacked to form a continuous spectrum and the effective exposure time is derived based on the assumed 10 min integration time. We mask the noisiest pixels, corresponding to 3% of the total pixel count, a threshold further discussed in Section 5. We then add synthetic Gaussian noise to the spectrum across the full wavelength range using the S/N values returned by the ANDES ETC.

We normalize each simulated spectrum by sequentially dividing it by its out-of-transit time-averaged stellar baseline, and its median spectrum along wavelength to correct for throughput variations and stellar continuum fluctuations. The corresponding uncertainty from the finite out-of-transit baseline is propagated analytically, including covariance terms for out-of-transit exposures, ensuring that the final variances reflect this additional source of noise. Finally, the noisy spectrum is continuum-normalized as described by Bello-Arufe et al.

(2023). To do so, we first fit the continuum using a Gaussian filter with a standard deviation of 500 pixels. We then divide the spectrum by the continuum such that we isolate high-frequency spectral features while suppressing the low-frequency continuum. We propagate the pixel uncertainties accordingly. This spectrum is subsequently used as input for a cross-correlation analysis to evaluate the detection significance as a function of the number of observed transits. For clarity, the complete procedure for constructing this final spectrum is summarized in Figure 3.

5. CCF ANALYSIS AND ITS LIMITATIONS

After generating the simulated data, we study the minimum number of transits required to detect at least one gas.

Our analysis cross-correlates each exposure's spectrum $S_{\rm obs}(t)$ with a normalized model template $S_T(v)$ Doppler-shifted over $v \in [-200, 200]\,{\rm km\,s^{-1}}$ using 301 velocity bins. The cross-correlation is inverse-variance weighted, with per-pixel variances propagated through the same sequence of normalizations as the spectrum. In particular, during out-of-transit normalization we compute the mean spectrum from out-of-transit pixels, propagate its uncertainty from their individual variances, and include covariance terms capturing the dependence between the spectrum and the out-of-transit mean. This approach accounts for the wavelength dependence of the noise and reduces sensitivity to outliers. Denoting the variance as σ^2, the cross-correlation is computed as follows:

$$c(v, t) = \frac{\sum\limits_{\lambda} \left[S_T(v) - \overline{S_T(v)} \right] \left[S_{\rm obs}(t) - \overline{S_{\rm obs}(t)} \right]}{\sigma^2(t)} \quad (15)$$

We extract the planet signal from a single 1-D CCF at mid-transit, expected at the total planetary velocity (Equation 6). An traditional approach often found in the literature involves extracting the planet's S/N by dividing the CCF peak by the standard deviation of values far from the planet's location (e.g., Brogi et al. 2013; de Kok et al. 2013; Bello-Arufe et al. 2023). However, the noisy pixels associated with regions of high telluric absorption can result in outliers at $v_{\rm sys} = 0\,{\rm km\,s^{-1}}$, which can bias the extracted planet's S/N value. Additionally, because the CCF baseline is non-flat and the noise level changes with velocity, the "far-from-signal" standard deviation is therefore no longer a faithful estimate, biasing the proxy and leading to incorrect rankings even when the underlying planetary contribution is unchanged. This zero-velocity artifact is discussed in more detail in Appendix A.

As a first mitigation of the telluric-induced outliers in the CCF, we mask the $\approx 3\%$ lowest-S/N pixels before cross-correlation. After testing multiple values, we find a threshold of 3% to be optimal: it significantly improves the cross-correlation signal by removing the noisiest pixels while preserving most of the planetary lines.

Even after masking, the CCF exhibits velocity-dependent covariance from residual telluric structure near $v_{\rm sys} = 0\,{\rm km\,s^{-1}}$. These effects violate the common assumption of independent Gaussian noise on a flat baseline, so the far-from-peak variance is not representative of the noise at the CCF peak. We therefore adopt a fully Bayesian treatment that models the entire CCF as the data vector. Building on the HRCCS retrieval framework that recasts the CCF as a likelihood (Brogi & Line 2019; Gibson et al. 2020), we extend it with molecule-specific kernels to capture the expected CCF morphology and an autoregressive AR(1) term to model inter-bin covariance directly in velocity space. We then use `MultiNest` to compute model evidences; Bayes factors provide the quantitative basis for assessing molecular detections.

6. BAYESIAN EVIDENCE ESTIMATION OF ATMOSPHERIC SIGNALS

To determine when a species is detected, we apply Bayesian model comparison to the 1-D CCF, contrasting a flat model with one incorporating the expected molecular signal from the planetary atmosphere, hereafter referred to as the "signal" model. Using `MultiNest`, we compute the evidences $Z_{\rm sig}$ and $Z_{\rm flat}$, where

$$Z_m = \int \mathcal{L}_m(\mathbf{y} \mid \theta_m) \pi_m(\theta_m) {\rm d}\theta_m \quad (16)$$

with likelihood $\mathcal{L}_m$ for the CCF data $\mathbf{y}$ and prior π_m over parameters θ_m of model $m \in \{{\rm sig, flat}\}$. We compare the two models via the Bayes factor B, calculated as

$$\Delta B = \log_{10} Z_{\rm sig} - \log_{10} Z_{\rm flat} \quad (17)$$

with its propagated uncertainty, and we map ΔB to Jeffreys categories for qualitative statements.

6.1. Data and noise handling

For each planet-molecule configuration, we start by generating $N = 100$ independent trials: in trial k we draw a new realization of the noise, process it through the simulation pipeline, and obtain a CCF array $\{y_i^{(k)}\}$ sampled at velocity bins $\{v_i\}$. This ensemble constitutes our Monte Carlo sample for that configuration. Each CCF spans $[-200, 200]\,{\rm km\,s^{-1}}$, sampled on 301 velocity bins ($\Delta v_{\rm native} \approx 1.33\,{\rm km\,s^{-1}}$), typically with a candidate peak near the expected total planetary velocity
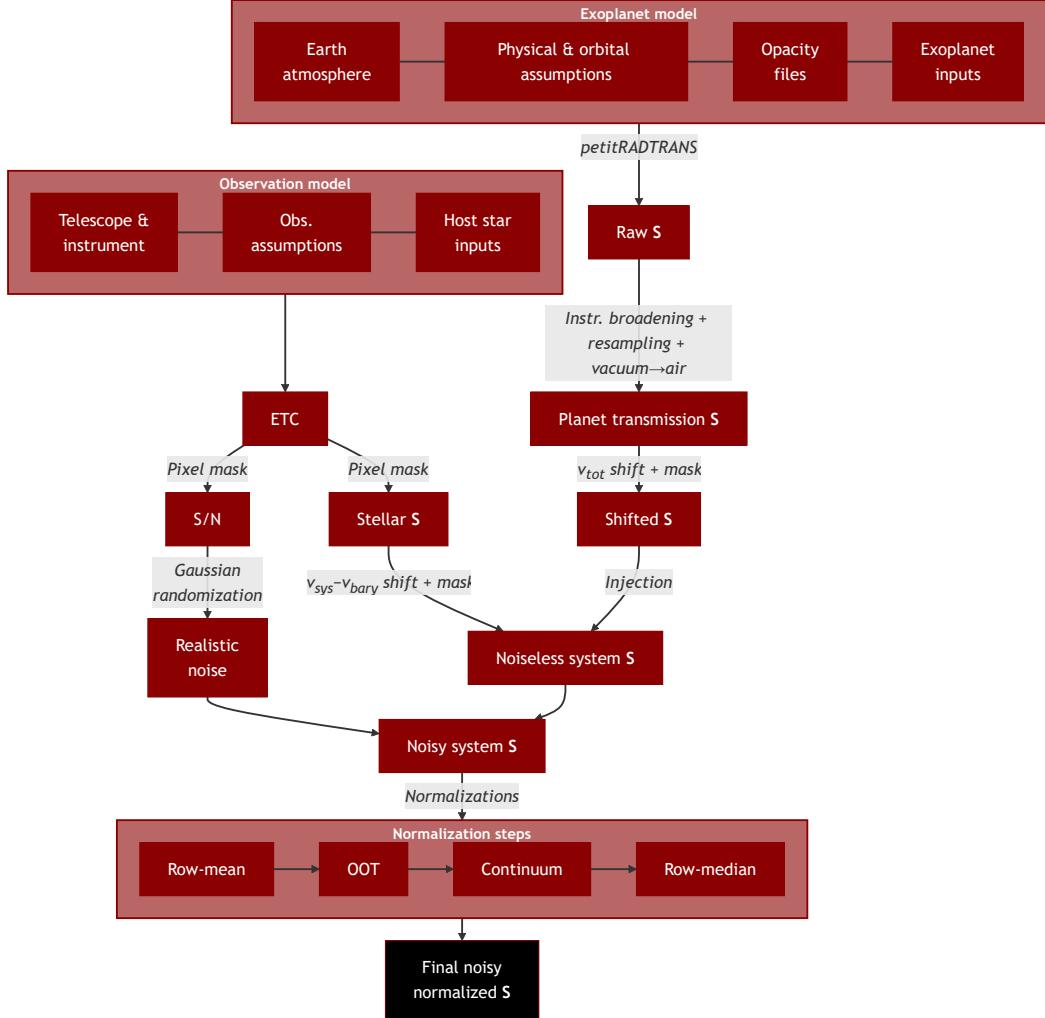
Figure 3. Workflow for constructing the realistic high-resolution transmission spectrum used as input to the CCF analysis for a given molecule-planet pair. Here, S denotes a spectrum

$v_{\rm tot}$ (Equation 6) and, in some cases, a residual telluric feature near $0\,{\rm km\,s^{-1}}$. We then estimate the empirical dispersion across trials,

$$\sigma_i = \sqrt{\frac{1}{N-1}\sum_{k=1}^{N}\left(y_i^{(k)} - \bar{y}_i\right)^2}, \qquad \bar{y}_i = \frac{1}{N}\sum_{k=1}^{N} y_i^{(k)} \tag{18}$$

where $y_i^{(k)}$ is the CCF value at velocity bin v_i in trial k. In practice, adjacent bins within any single CCF are correlated, so σ_i characterizes per-bin scatter across trials but does not encode inter-bin covariance, as discussed in Appendix B.1. Rather than estimating the full covariance (see Appendix B.2), we adopt a mitigation strategy that renders the likelihood conservative and the error model more representative of the data. Specifically, we compute the median empirical lag-1 autocorrelation $\hat{\rho}$ across the N realizations, as motivated

by an AR(1) process (autoregressive model of order 1), and inflate the per-bin uncertainties according to

$$\sigma_{i,{\rm eff}} \leftarrow \sigma_i \sqrt{\frac{1+\hat{\rho}}{1-\hat{\rho}}} \tag{19}$$

with $\hat{\rho}$ clipped to $[-0.95, 0.95]$. This preserves per-bin weighting while widening the effective noise when residual correlation is present. This procedure is further justified and discussed in Appendix B.3.

6.2. *Models*

A straightforward approach would be to fit a Gaussian profile to the CCF maximum. However, such a fit would not address the telluric contamination discussed in Section 5. In addition, some of the species we study exhibit quasi-regular line spacings within a vibrational band, which generate secondary peaks in the CCF at

characteristic lags (Said et al. 2016). In our spectral ranges, O_2 and CO_2 show near-regular line spacing, resulting in side lobes in the CCF (Kiehl & Yamanouchi 1985; Martin & Barker 1932)), as illustrated in Figure 4. By contrast, CH_4 and H_2O possess dense and irregular spectra (e.g., Yurchenko & Tennyson 2014; Yurchenko et al. 2024; Barber et al. 2006; Tennyson et al. 2013), leading to a single dominant CCF maximum with no discernible secondary structure. A simple Gaussian fit, with its flat baseline, cannot capture the side lobes. The baseline then absorbs mid-scale features, biasing the recovered amplitude A toward lower values and reducing the inferred evidence. This behavior is illustrated in the top panel of Figure 4.
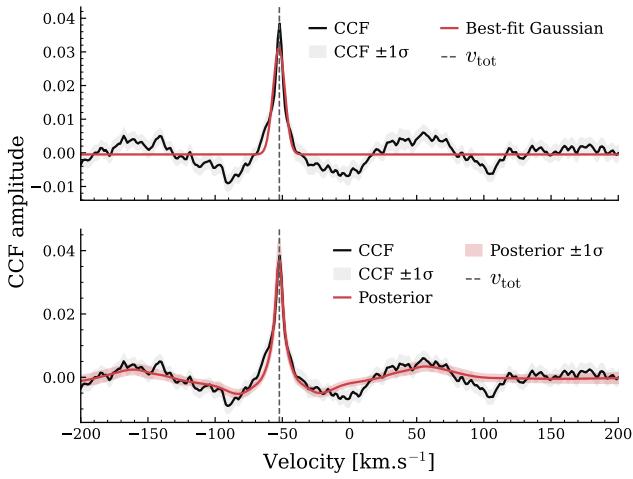
Figure 4. CCF for CO_2 on TRAPPIST-1 d, assuming we observed all available transits in a 30-year window (339 transits), fitted with two different models. **Top:** Gaussian + constant baseline fitted by bounded non-linear least-squares (Trust-Region Reflective algorithm), minimizing the sum of squared residuals. It fails to capture the secondary lobes. **Bottom:** posterior predictive mean of the signal model, from our Bayesian model. The molecule-specific kernel captures the full shape of the CCF, outperforming the simple Gaussian model.

To address this, we introduce a molecule-specific model that more accurately represents the shape of the CCF. For the Bayesian comparison, both hypotheses (flat and signal) share a quadratic baseline that captures the smooth, low-order curvature of the CCF across the $\pm 200\,\mathrm{km\,s^{-1}}$ range:

$$B(v) = y_0 + b(v - v_{\mathrm{tot}}) + c(v - v_{\mathrm{tot}})^2 \qquad (20)$$

where (y_0, b, c) are nuisance parameters and v_{tot} is the reference velocity defined in Equation 6.

The null model contains only this baseline:

$$m_{\mathrm{null}}(v) = B(v) \qquad (21)$$

while the signal model adds a kernel core derived from the molecular template:

$$m_{\mathrm{sig}}(v) = B(v) + A\,K\left(\frac{v - \mu}{\kappa}\right) \qquad (22)$$

where A is the core amplitude, μ is the signal's peak velocity, and $\kappa > 0$ is a scale factor. The kernel K depends on the molecule and instrument setup and is normalized such that $K(0) = 1$, so that the model's maximum at $v = \mu$ equals A. Operationally, K is defined as the autocorrelation of the model template:

$$K(\Delta v) \propto \int T(u)\,T(u + \Delta v)\,\mathrm{d}u \qquad \max_{\Delta v}|K(\Delta v)| = 1 \qquad (23)$$

tabulated on the CCF grid and interpolated at $\Delta v = (v - \mu)/\kappa$. Using this kernel instead of a Gaussian preserves the deterministic peak-lobe pattern imposed by the molecular line distribution. Because the kernel is the autocorrelation of the template, it faithfully reproduces that structure. The bottom panel of Figure 4 illustrates the result.

6.3. Priors

We choose weakly informative priors on the model parameters:

- **Centroid.**

$$\mu \sim \mathcal{N}\left(v_{\mathrm{tot}}, 1.4\,\mathrm{km\,s^{-1}}\right) \qquad (24)$$

  The standard deviation of this Gaussian prior matches the native bin width ($\approx 1.33\,\mathrm{km\,s^{-1}}$).

- **Kernel stretch.** We place a log-uniform prior on the dimensionless stretch κ,

$$\ln \kappa \sim \mathcal{U}\left(\ln 0.6, \ln 1.6\right) \qquad (25)$$

  so the kernel can broaden or narrow multiplicatively without favoring an absolute scale.

- **Amplitude.** We define a robust amplitude scale $A_\sigma = 4\,\tilde{\sigma}$, where $\tilde{\sigma}$ is the median of the (possibly inflated) per-velocity uncertainties. The amplitude parameter A uses a half-normal prior,

$$A \sim \mathrm{HalfNormal}\left(A_\sigma\right) \qquad (26)$$

  so that typical peaks are a few $\tilde{\sigma}$ while allowing larger values without being overly diffuse.

- **Baseline.** We place normal priors on the offset, slope and curvature of the baseline:

$$y_0 \sim \mathcal{N}(0, \tilde{\sigma}), \quad b \sim \mathcal{N}\left(0, \frac{0.2 A_\sigma}{\Delta v_{\mathrm{span}}}\right), \quad c \sim \mathcal{N}\left(0, \frac{0.1 A_\sigma}{\Delta v_{\mathrm{span}}^2}\right) \qquad (27)$$

6.4. *Likelihood and nested sampling*

All priors are built via inverse cumulative distribution function (CDF) transforms from unit-hypercube draws $u \sim \mathcal{U}(0,1)$ in `MultiNest` and numerically clipped to $\left[10^{-9}, 1 - 10^{-9}\right]$ to avoid infinities. Transforms are implemented with `SciPy` (Virtanen et al. 2020). For a parameter θ:

(i) Normal prior. For $\theta \sim \mathcal{N}(\mu, \sigma^2)$,

$$\theta = \mu + \sigma\sqrt{2}\, \mathrm{erf}^{-1}(2u - 1) \qquad (28)$$

where erf^{-1} is the inverse error function.

(ii) Half-normal prior. For a non-negative parameter with scale σ,

$$\theta = \sigma\sqrt{2}\, \mathrm{erf}^{-1}(u), \qquad \theta \geq 0 \qquad (29)$$

(iii) Log-uniform prior. For bounds $[x_{\min}, x_{\max}]$,

$$\theta = x_{\min} \left(\frac{x_{\max}}{x_{\min}} \right)^u \qquad (30)$$

These mappings ensure that `MultiNest` samples the intended prior measures directly from u while maintaining numerical stability in the likelihood evaluation. Given the velocity grid $\{v_i\}$, observed CCF values y_i, and per-bin uncertainties σ_i (inflated when applicable), we assume independent normal errors:

$$\ln \mathcal{L}(\theta) = -\frac{1}{2} \sum_i \left(\frac{y_i - m(v_i; \theta)}{\sigma_i} \right)^2 + \mathrm{const} \qquad (31)$$

where $m(v_i; \theta)$ is the model prediction at v_i given the parameter vector θ, $\mathcal{L}(\theta)$ is the Gaussian likelihood of $m(v_i; \theta)$, and $\mathrm{const} \equiv -\frac{1}{2} \sum_i \ln(2\pi\sigma_i^2)$ is independent of θ and m, and can be dropped for model comparison.

We fit the flat (no-signal) and atmospheric-signal models with `MultiNest` via `PyMultiNest` (Buchner 2016), obtaining posterior samples and the log-evidence $\ln Z$ for each model. We compute the Bayes factor defined in Equation 17 and propagate its uncertainty from the evidence errors reported by `MultiNest`. Assuming independent evidence estimates,

$$\sigma_{\Delta B} = \frac{1}{\ln 10} \sqrt{\sigma_{\ln Z_{\mathrm{sig}}}^2 + \sigma_{\ln Z_{\mathrm{null}}}^2} \qquad (32)$$

7. RESULTS

7.1. *Initial target screening*

We apply our analysis pipeline[7] to each planet-molecule pair, generating $N \geq 100$ CCF realizations for the following species: O_2, CO_2, CH_4, and H_2O. For every realization, we record ΔB and its associated uncertainty, and assign the corresponding Jeffreys category. We assess the statistical stability sof the median detection strength for $N = 100$ using the stopping criterion described in Appendix C. The procedure is repeated at four cumulative transit-count levels (25%, 50%, 75%, and 100% of the total transits observable within a 30 yr scheduling window) for each planet-molecule pair. Then, we estimate the minimum cumulative observing time required for a "very strong" detection by linear interpolation of the Bayes factors obtained at these four levels. We obtain the planet-molecule pairs that reach a "very strong" detection (Jeffreys scale; $\Delta B \geq 1.5$) within this window. This initial screening excludes eight planets that show no detectable signal within 30 yr; Kepler-1649 c, Kepler-442 b, K2-3 d, Kepler-62 f, Kepler-186 f, Kepler-62 e, Kepler-452 b and Kepler-440 b, leaving eleven candidates.

The maximum cumulative observation period of approximately 30 yr per target is consistent with the current operational plan, which anticipates that the ELT will remain in service for at least three decades[8]. For simplicity, we assume that observations are feasible only half of the time due to daytime sky conditions, and of the remaining time, we further remove half to account for seasonal observability constraints. This effective duration is converted into a number of transits, specific to each planet, by dividing by its orbital period (including an out-of-transit baseline), each transit being characterized by a duration T_{14}.

7.2. *Detectability results*

After the initial screening, eleven planets exhibit at least one "very strong" signal for at least one gas within a 30 yr operational window of ELT/ANDES. We next evaluate whether these gases are detectable within a reasonable number of transits. To this end, we evaluate four exposure times corresponding to 1, 5, 20, and 40 transits, adopting 40 transits as a practical upper limit.

All planets that yield at least one "very strong" signal for at least one gas in fewer than 40 transits are listed in Figure 5. This criterion further reduces the sample

---

[7] https://github.com/EvannKurzawa/ANDES-Forecast
[8] https://elt.eso.org/about/faq/#question_11
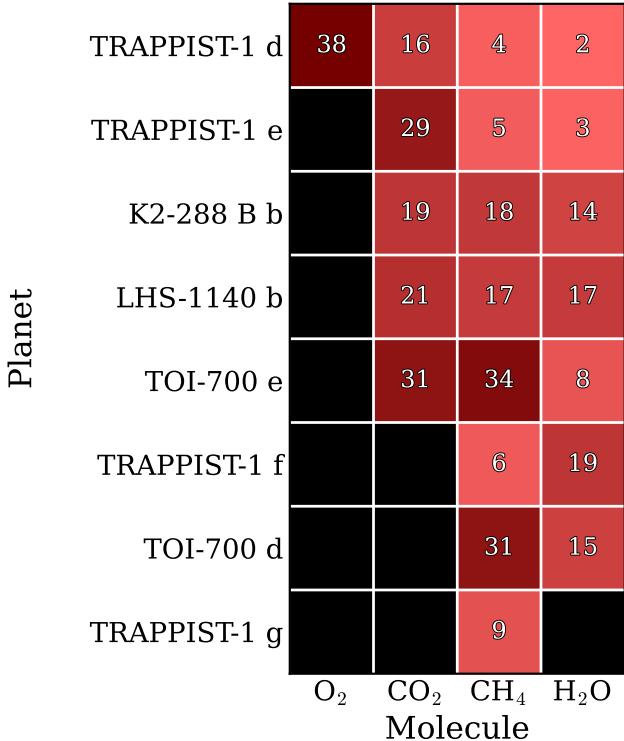
Figure 5. Minimum number of transits required for a "very strong" detection of each gas with ANDES. Planets are ordered from top to bottom by (i) number of detectable molecules, and (ii) smaller total required transits (sum over detected molecules). Empty cells indicate that a "very strong" detection cannot be achieved in 40 transits.

Figure 6. Minimum number of transits required for a "very strong" detection, evaluated per gas using the designated molecular bands in μm (see Table 2). Planets are ordered from top to bottom by (i) number of detectable bands and (ii) smaller total required transits (sum over detected molecules). Empty cells indicate that a "very strong" detection cannot be achieved in 40 transits.

from eleven to eight planets. For these eight planets, we estimate the minimum number of transits required for a "very strong" detection by linear interpolation of the Bayes factors obtained at 1, 5, 20, and 40 transits. Figure 5 compiles these estimates and provides the most planning-relevant result: a concise view of the transit requirement for each gas and retained target. We discuss its implications in Section 8.1.

7.3. Comparing to previous works

While there are no published in-depth studies on the ability of ANDES to detect gases on potentially habitable exoplanets, several works have investigated this topic using general parameters representative of a high-resolution spectrograph on an extremely large telescope (e.g., Fujii et al. 2018; Currie et al. 2023). In our analysis, we obtain one CCF per realization and exposure setting over the full wavelength range of ANDES. This approach is appropriate because ANDES will observe its entire spectral range simultaneously[9]. However, this

broadband metric is not directly comparable to previous studies that assess detectability on individual molecular bands rather than over the full range. To enable a meaningful comparison, we implement an additional per-band pipeline that evaluates molecule-specific spectral windows individually. The adopted bands are listed in Table 2.

For each spectral band, we divide the spectrum into $0.02\,\mu$m bins and, within each bin, discard the lowest 3% of pixels by S/N (as in Section 5) and remove any pixels with zero S/N. A spectral window is deemed valid if at least 200 pixels remain after this cut. We then compute the CCF and evaluate it within our Bayesian framework, yielding a per-band evidence. We obtain "very strong" detectability for eight planets in a subset of our adopted bands. Similarly to 5, we construct Figure 6, which shows, for each retained exoplanet, the number of transits required to achieve a "very strong" detection of each molecule in the bands that yield a measurable signal. The same grid and interpolation procedure as in Figure 5 is used. As expected, the resulting Bayes factors are generally weaker than in the broadband analysis.

Table 2. Spectral bands adopted in the per-band pipeline.

Molecule	Bands (μm)
H_2O	0.90–0.98; 1.10–1.16; 1.30–1.50; 1.78–1.98; 2.38–2.50
CH_4	1.10–1.20; 1.30–1.44; 1.62–1.70; 2.12–2.50
O_2	0.68–0.70; 0.75–0.78; 1.24–1.30
CO_2	1.42–1.46; 1.56–1.62; 1.94–2.10

Finally, given the extensive literature on the TRAPPIST-1 system, we compile Table 3, which lists the number of transits required to achieve a per-band detection of $\Delta B = 5.43$ for TRAPPIST-1 e. Values are reported at 20, 40, 150, and 300 transits using the interpolation in Section 7.2, and the table contains comparisons to previous studies. The specific threshold of $\Delta B = 5.43$ corresponds, via the heuristic relation $n_\sigma \approx \sqrt{2 \ln B}$ as written by Kipping & Benneke (2025) and motivated by the asymptotic connection between Bayes factors and likelihood-ratio statistics discussed by Kass & Raftery (1995), to roughly 5σ. We adopt this only as a heuristic reference to facilitate comparison with earlier work that reported results in σ-units. We stress that such conversions between Bayes factors and frequentist significance are not formally defined and should not be interpreted literally, as discussed by Kipping & Benneke (2025) and Thorngren et al. (2025). Our analysis itself is expressed entirely in terms of Bayes factors. For interpretation, readers should rely on the Bayes factors themselves; the σ proxy is provided solely for approximate comparability. Further discussion is given in Section 8.3.

8. DISCUSSION

8.1. *Interpreting our results*

From the initial pool of 70 targets in our sample, we exclude 29 nontransiting ones, 11 that never rise above the ELT's 20° altitude limit, and 11 with radii larger than $2R_\oplus$. Among the remaining 19 planets, 11 meet the Jeffreys "very strong" detection threshold for at least one molecule with ANDES within an operational time of 30 years, and 8 of these also satisfy the criterion within 40 transits. CH_4 emerges as the most accessible species, meeting the detectability criterion in all eight planets in less than 40 transits. H_2O, detectable for seven planets within 40 transits, is also highly favorable, particularly through its 2.38–2.50 μm band. CO_2 is detectable for five planets, while O_2, detectable for only one target, proves the most challenging.

TRAPPIST-1 d stands out as the overall best candidate, with all four gases detectable, each requiring the lowest number of transits among all targets. It is also the only planet where O_2 is detectable. For TRAPPIST-1 e, relaxing the 40-transit cap to estimate the requirement for O_2 yields 49 transits; above our threshold but indicative of realistic feasibility. More broadly, all four potentially habitable TRAPPIST-1 planets yield at least one "very strong" signal within 40 transits. Other promising targets include the TOI-700 system, as well as the planets LHS 1140 b and K2-288 B b, for which CH_4 and CO_2 are consistently detectable.

Figure 6 indicates that the spectral regions contributing most to the broadband detections are 1.62–1.70 μm and 2.12–2.50 μm for CH_4, 1.56–1.62 μm for CO_2, and 2.38–2.50 μm for H_2O; Table 3 shows 1.24–1.30 μm for O_2. Accordingly, the notably small transit counts for CH_4 and H_2O partly reflect our use of the full targeted ANDES wavelength coverage (0.38–2.4 μm), which includes these strong bands; limiting the configuration to $\lambda \leq 1.8\,\mu$m would reduce their detectability.

8.2. *Interpreting variations in detectability across targets*

The differences in detectability of all gases across planets within the system are explained by variations in scale height. The amplitude of individual planetary lines scales approximately with the atmospheric scale height, as $\Delta D \propto HR_p/R_\star^2$, where $H = kT/(\mu g)$ increases with temperature and decreases with gravity and mean molecular weight (Seager & Sasselov 2000). This explains why planets with higher temperature and lower gravity, such as TRAPPIST-1 d, yield stronger features than cooler, higher-gravity planets like TRAPPIST-1 e.

The contrast in molecular detectability patterns across TRAPPIST-1, LHS 1140, and K2-288 B is primarily driven by the spectral energy distributions of their host stars. CO_2 absorption is probed mainly near 1.6 μm, whereas CH_4 and H_2O are best detected at longer wavelengths, (2.1-2.5 μm, see Figure 6). Cooler stars like TRAPPIST-1 emit a larger fraction of their energy in the K band than earlier M dwarfs such as LHS 1140 or K2-288 B, leading to a higher photon budget in the relevant CH_4/H_2O bands. Synthetic atmosphere models confirm that late-M dwarfs concentrate a larger share of their flux at longer wavelengths, with flux escaping primarily through the J, H, and K bands in that order of decreasing strength (Allard & Hauschildt 1995). As a result, CH_4 and H_2O are significantly easier to detect in TRAPPIST-1 than in the earlier-type systems LHS 1140 and K2-288 B, even though CO_2 detectability remains comparable across all three.

Table 3. Per-band transit requirements for detecting different gases on TRAPPIST-1 e.

Band	Transits required (this work)	Currie et al. (2023)	Wunderlich et al. (2020)
(μm)	($\Delta B = 5.43$)	(5σ)	(5σ)
CH_4; 2.12–2.50	31	-	26
H_2O; 2.38–2.50	47	-	-
CO_2; 1.56–1.62	108	34	33
CH_4; 1.62–1.70	115	33	-
O_2; 1.24–1.30	148	200	910
H_2O; 0.90–0.98	> 300	> 300	1224

NOTE—Number of transits required to reach a per-band detection threshold of $\Delta B = 5.43$ (used here as a rough proxy for "5σ"; see Section 7.3). Values are reported for each relevant spectral band from our per-band pipeline, with additional columns comparing to estimates from Currie et al. (2023) and Wunderlich et al. (2020). As discussed in Section 7.3, such σ-based comparisons should be treated with caution: we interpret results via Bayes factors, whereas the cited works adopt frequentist significance.

8.3. *Comparing to previous studies*

Establishing direct comparisons with previous studies is challenging. While a limited number of works have explored the potential of the ELT (or other extremely large telescopes) for characterizing the atmospheres of potentially habitable rocky planets around M-dwarfs (e.g., Currie et al. 2023; Wunderlich et al. 2020), these typically consider hypothetical planets at distances that differ from those of our sample (often 5-15 pc). Most of these studies targeted individual absorption bands, and none specifically investigated the performance of AN-DES. Nonetheless, with our per-band pipeline, we can make some approximate comparisons with earlier studies, particularly those focusing on the TRAPPIST-1 system, which has been extensively investigated.

Currie et al. (2023) simulated transmission spectroscopy of an Earth-like planet transiting TRAPPIST-1, with an orbital configuration analogous to TRAPPIST-1 e. They reported that the ELT could detect CO_2 at $1.56\,\mu$m in 33 transits, CH_4 at $1.6\,\mu$m in 34 transits, H_2O at $0.9\,\mu$m in > 300 transits, and O_2 at $1.27\,\mu$m in 200 transits (all at a 5σ significance level). Using our per-band pipeline applied to TRAPPIST-1 e up to 300 transits (see Table 3), we find a "5σ" detection for CO_2 at $1.56\,\mu$m in 108 transits, CH_4 at $1.6\,\mu$m in 115 transits, H_2O at $0.9\,\mu$m in > 300 transits, and O_2 at $1.27\,\mu$m in 148 transits. To achieve a "decisive" (Jeffreys scale; $\Delta B \geq 2.0$) detection of CO_2 and CH_4 on TRAPPIST-1 e with our pipeline, ANDES requires 40 and 46 transits at $1.56\,\mu$m (CO_2) and $1.6\,\mu$m (CH_4), respectively. The results are in good agreement, despite the differences in methodology and instrument parameters (see Section 7.3).

Wunderlich et al. (2020) modeled a TRAPPIST-1 e atmosphere containing 0.1 bar of CO_2 and estimated the number of transits required for a 5σ detection for similar spectral bands. Their estimates are more conservative, with 1,224 transits needed for the H_2O band at 0.90–0.98 μm. However, there is reasonable agreement for CH_4, with 26 transits required for the 2.12–2.50 μm band in Wunderlich et al. (2020) and 31 in our analysis. Table 3 again illustrates that the two bands beyond 2 μm carry much of the diagnostic power: the number of transits needed to detect H_2O on TRAPPIST-1 e with our pipeline, is about ten times smaller for the 2.38–2.50 μm band than for 0.90–0.98 μm.

We adopt a "very strong" Bayes factor as our default detection threshold throughout this study but use the more stringent $\Delta B = 5.43$ level when comparing to 5σ-based results, which correspond to a higher detection confidence. We emphasize, however, that these comparisons mix different detection metrics: our analysis quantifies detectability via the Bayes factor (evidence ratio), whereas other works use an S/N-based σ threshold; any mapping between the two is heuristic (see Section 7.3 and Kipping & Benneke 2025).

Finally, comparison with previous target-prioritization studies reveals strong convergence on the most amenable targets for atmospheric characterization. Both Hill et al. (2023) and Bohl et al. (2025) identify TRAPPIST-1 d as one of the most favorable habitable-zone planets, consistent with our findings, where TRAPPIST-1 d also ranks highest (see Figure 5). In our framework, the remaining potentially habitable TRAPPIST-1 planets (e, f, g) follow closely, similar to the sequence reported by Bohl et al. (2025). Other planets that appear prominently in at least one of these

works, as well as in ours, include LHS 1140 b, K2-288 B b, TOI-700 d, and TOI-700 e, reinforcing their status as strong secondary candidates. However, the radius of K2-288 B b falls within the radius valley (Fulton et al. 2017), and the lack of a measured mass means that this planet might not be rocky. Similarly, the low density of LHS 1140 b may indicate a water world composition instead of a purely rocky one (Damiano et al. 2024; Cadieux et al. 2024a), although the planet might still retain habitable conditions on its surface.

8.4. *Preliminary JWST results for these targets*

Nearly all of the planets identified in our final sample have been or will be observed by the James Webb Space Telescope (JWST), and early results are beginning to inform the atmospheric properties of these temperate terrestrial worlds. Initial JWST observations appear inconsistent with significant atmospheres on most warm-to-hot rocky exoplanets around M dwarfs (e.g. Greene et al. 2023; Zieba et al. 2024; Meier Valdés et al. 2025; Allen et al. 2025; Fortune et al. 2025; Luque et al. 2025; but see also Gressier et al. 2024; August et al. 2025; Bello-Arufe et al. 2025), but it is still unclear whether this trend extends to the temperate population (Damiano et al. 2024; Cadieux et al. 2024b; Espinoza et al. 2025; Piaulet-Ghorayeb et al. 2025; Glidden et al. 2025).

JWST has devoted significant attention to the TRAPPIST-1 system. For the inner planets, TRAPPIST-1 b and c, dayside emission spectra show high brightness temperatures inconsistent with dense, heat-redistributing atmospheres, effectively ruling out Earth-like or Venus-like conditions (Greene et al. 2023; Zieba et al. 2024). For TRAPPIST-1 d, our top-ranked target, JWST/NIRSpec transmission spectra across four transits reveal a flat spectrum, ruling out a modern Earth-like, cloud-free atmosphere like the one we model in this work (Piaulet-Ghorayeb et al. 2025). Instead, the data are consistent with either a tenuous Mars-like atmosphere, a high-altitude Venus-like cloud deck, or no atmosphere at all.

In contrast, the outer TRAPPIST-1 planets remain promising. For TRAPPIST-1 e, early JWST transmission spectra exclude H_2-dominated atmospheres but remain consistent with high-mean-molecular-weight atmospheres such as N_2 with trace CO_2 or CH_4 (Espinoza et al. 2025; Glidden et al. 2025). These data do not yet confirm the presence of an atmosphere, but they leave open the possibility of the types of atmospheres we simulate. The outer planets may be better candidates for retaining secondary atmospheres (Krissansen-Totton 2023; Piaulet-Ghorayeb et al. 2025), but confirming the presence of such atmospheres remains an open challenge.

Among non-TRAPPIST-1 targets, LHS 1140 b stands out as one of the most observable targets with ANDES. Our study finds strong CH_4 and CO_2 detectability assuming an Earth-like atmosphere, and recent JWST observations provide the strongest evidence to date that a temperate terrestrial exoplanet may indeed retain a substantial secondary atmosphere. Spectroscopic observations with JWST indicate a featureless or weakly sloped transmission spectrum, ruling out a clear H_2/He-dominated envelope and pointing instead to a high-mean-molecular-weight atmosphere (Damiano et al. 2024). Modeling of the available data favors a scenario in which LHS 1140 b is a water-rich world, potentially with an N_2-dominated atmosphere and some greenhouse agents (e.g., H_2O or CO_2). Tentative evidence for molecular nitrogen has been reported, and while not yet conclusive, such a detection would represent a major milestone in exoplanet science (Cadieux et al. 2024b). These findings reinforce LHS 1140 b as a top-priority target for future characterization. Follow-up observations with JWST are underway through the GO 7073 (Lustig-Yaeger et al. 2025) and the "Rocky Worlds"[10] DDT programs, which will obtain transmission and secondary eclipse observations, respectively, to further constrain the planet's atmospheric composition.

The TOI-700 system hosts multiple temperate, Earth-sized planets that are considered strong candidates for atmospheric retention and future characterization (Gilbert et al. 2020; Suissa et al. 2020; Gilbert et al. 2023). Its quiet M-dwarf host star, combined with the planets' favorable sizes and irradiation levels, make TOI-700 d and e particularly amenable targets. JWST Cycle 3 GO program 6193 (Pass et al. 2024), which has been approved to observe both planets with NIRISS/SOSS, may directly test our model assumptions. Other favorable planets from our sample, such as K2-288 B b, have not yet been observed by JWST but might become viable candidates in future cycles.

Taken together, these JWST results suggest that modern, cloud-free Earth-like atmospheres may be rare among the most irradiated terrestrial exoplanets. At the same time, modern Earth-like atmospheres remain plausible for several key targets, including LHS 1140 b and TRAPPIST-1 e. While our simulations present optimistic lower limits, the ongoing synergy between JWST and upcoming facilities like ANDES will be crucial in refining our understanding of atmospheric diversity and biosignature detectability on rocky exoplanets.

---

[10] https://rockyworlds.stsci.edu/rw-website-targets.html

8.5. *Important caveats*

Below, we summarize the most relevant caveats.

1. **Telluric correction.** We assume that telluric absorption lines can be removed down to the photon-noise level. This remains a significant challenge, particularly in the near-infrared where telluric contamination is severe and time-variable. Current methods such as `molecfit` (Smette et al. 2015; Kausch et al. 2015) and data-driven principal component analysis (PCA; e.g., Meech et al. 2022) can achieve excellent results under favorable conditions, but full correction at the noise limit is rarely attained. The achievable precision depends on the instrument stability, the accuracy of the telluric models, and the temporal variability of the atmosphere during the observing sequence. Residual telluric structures may therefore introduce additional systematic noise that increases the effective number of required transits.

2. **Planet-star spectral separation.** We assume that the planetary spectrum can be isolated from the stellar and telluric signals without degrading the exoplanetary features. This is a strong assumption, especially for slowly moving planets such as those in the TRAPPIST-1 system, where the radial-velocity change during transit is only a few tens of meters per second. Traditional detrending techniques, including SYSREM (Tamuz et al. 2005; Mazeh et al. 2007) or PCA-based filtering, rely on the assumption that the planetary lines shift sufficiently in velocity space to avoid being removed together with the quasi-stationary stellar and telluric lines. In practice, this assumption can fail for terrestrial systems, causing significant attenuation of the planetary signal. To address this limitation, Piskunov et al. (2025) recently introduced *TSD* (Transmission Spectroscopy Decomposition), an inverse-problem framework that simultaneously solves for the stellar, telluric, and planetary spectra directly from the time series, thereby avoiding template cross-correlation and minimizing signal loss. Although promising, the method requires resolving the planet's velocity change during transit and may therefore remain challenging for slowly orbiting rocky planets. Complementary approaches based on machine learning are also emerging; for instance, Kjærsgaard et al. (2023) proposed *TAU*, a neural-network-based telluric and stellar correction framework that can disentangle overlapping spectral components without relying on linear decomposition. These techniques represent an important step beyond classical detrending but have yet to be extensively validated in the low-velocity, low-signal-to-noise regime relevant to temperate terrestrial planets.

3. **Clouds and hazes.** We assume cloud-free atmospheres, although clouds and hazes may be present on these planets (e.g., Yang et al. 2013), potentially muting spectral features and complicating atmospheric characterization (Lustig-Yaeger et al. 2019). High-resolution transmission spectroscopy, however, probes the cores of individual lines, which form at low pressures, and is therefore less sensitive to clouds and hazes than its low-resolution counterpart (e.g., JWST) (Gandhi et al. 2020).

4. **Spectral range of ANDES.** We adopt the extended ANDES wavelength coverage (0.38-2.4 μm) rather than the baseline design. Extending the range to 2.4 μm is particularly important for the detection of CH_4 and H_2O, as demonstrated in Section 8.1. If the final implementation does not reach this range, the achievable sensitivity to these species will be significantly reduced.

5. **Stacking and noise scaling.** We assume that stacking multiple transits improves the signal-to-noise ratio as $\sqrt{N_{transits}}$, but it is unclear whether this will will be true for ELT/ANDES. The effective scaling will need to be established empirically during commissioning and early surveys; until then, our estimates should be considered optimistic lower limits.

6. **Stellar contamination.** We do not explicitly account for stellar variability, active regions, or flares, which are common among M dwarfs. Such activity can alter the effective photospheric spectrum during transit and bias molecular detections (the so-called transit light source effect, Rackham et al. 2018; Rathcke et al. 2025). While high-resolution spectroscopy provides some mitigation by leveraging the planet's orbital motion to disentangle the planetary and stellar components (Cauley et al. 2018), the impact of stellar contamination on high-resolution transmission spectroscopy of M dwarf systems remains poorly constrained.

7. **Additional simplifying assumptions,** including equilibrium temperatures with zero Bond albedo, circular orbits, equilibrium chemistry with

Earth-like composition and surface pressure, and a fixed readout time of 30 min per observation. We also neglect potential instrument systematics. In addition, because we apply the optimal barycentric correction to maximize Doppler separation between exoplanetary and telluric spectral lines, our results represent an optimistic scenario. In practice, observers are encouraged to schedule observations such that the exoplanet spectral lines are Doppler shifted away from the noisy telluric lines in order to increase the S/N of molecular features.

Taken together, these assumptions imply that our results likely underestimate the true number of transits required for detection. The goal of this work is to provide an overview of what could be achievable purely on a S/N basis and quantify the minimum number of transits required for detection for the most promising targets. Future efforts could aim to relax these assumptions and refine the physical modeling and pipeline implementation, and we encourage the community to focus on these challenges before ANDES starts operations, to maximize scientific return.

9. SUMMARY AND CONCLUSIONS

We developed a comprehensive framework to estimate the number of transits required to detect four key molecules, two of them potential biosignature gases, on rocky exoplanets with the ELT/ANDES. For each planet, we modeled a cloud-free, Earth-like, two-layer atmosphere and simulated the transmission spectra with `petitRADTRANS`, adding realistic noise with the ANDES ETC. We cross correlated the resulting spectra with model templates, and we quantified the detection significance using a Bayesian framework featuring a custom kernel scheme to emulate the CCFs and an autoregressive AR(1) model to mitigate correlation across velocity bins. We then applied the full pipeline to the most observable planets in our sample, which yielded evidence scores and the required transit numbers to reach detectability.

Our simulations indicate that methane and water vapor are the most accessible species for ground-based detection with ANDES, requiring fewer than ~5 transits for the most favorable targets (e.g., TRAPPIST-1 d and e). Carbon dioxide is moderately accessible, while oxygen remains the most challenging species to detect, even in the best cases. Eight planets show at least one molecule detectable within $\lesssim20$ transits, assuming ideal observing and no systematics introduced during detrending: TRAPPIST-1 d, e, f, g; K2-288 B b; LHS 1140 b; and TOI-700 d and e. TRAPPIST-1 d emerges as the most favorable target overall, with all

four gases detectable in fewer than 40 transits. These results should be regarded as optimistic lower limits, built on near-photon-noise performance, effective telluric removal, and cloud-free atmospheres.

APPENDIX

A. ZERO-VELOCITY ARTIFACT

We adopt a one-dimensional CCF analysis, which in several cases reveals a narrow peak at $v \approx 0 \, \mathrm{km \, s^{-1}}$ in the observer frame. This manifests as either a positive or negative peak, as shown in Figure A1 and accounts for most of the variance near zero lag. Consequently, "far-from-signal" standard-deviation estimates cease to be representative of the baseline, and classical S/N metrics become unreliable. The residual telluric structure can be explained by:

1. **Stationarity in the observer frame.** Unlike the stellar and the planetary spectra, telluric lines are stationary in the observer's rest frame. Any residual structure therefore accumulates at a fixed velocity, $v \approx 0 \, \mathrm{km \, s^{-1}}$, and can imprint spurious peaks near zero lag.

2. **Molecular overlap with searched gases.** The dense line lists of O_2, H_2O, CO_2 and CH_4 overlap strongly with Earth's atmosphere, so even small residual mismatches can mimic the spectral patterns used to search for exoplanet signals.
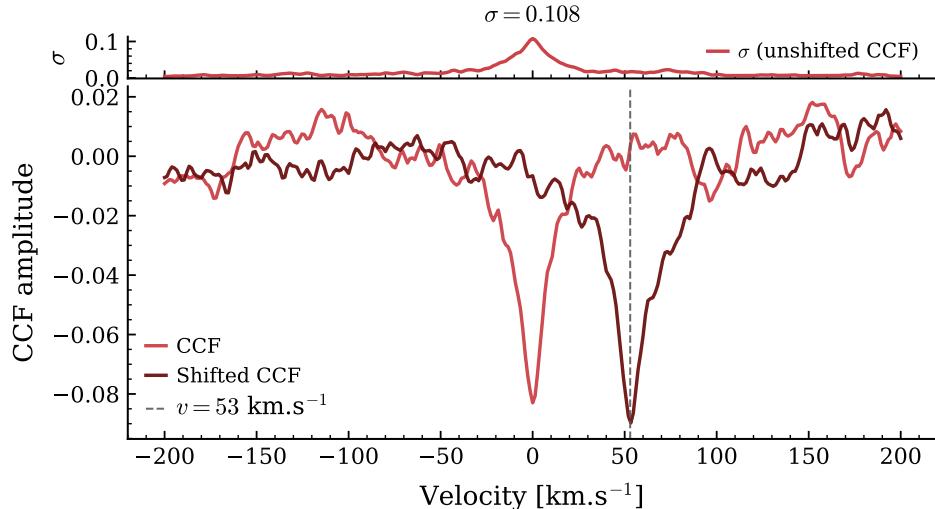

Figure A1. CCFs for Kepler-1649 c with low pixel masking. Two CCFs are shown: the standard one, and one where the stellar, planetary and telluric are shifted independently. In the shifted case, the feature follows the telluric component at $53 \, \mathrm{km \, s^{-1}}$, as discussed in Appendix A. For illustration, we display the standard deviation curve from the unshifted case, which shows a pronounced uncertainty peak at $0 \, \mathrm{km \, s^{-1}}$, further supporting the telluric origin of the artifact.

We perform the following diagnostics, which directly link the zero-velocity feature to telluric residuals:

1. **Artificial velocity-shift experiment.** The observed spectrum can be decomposed into contributions from the star, the planet, and Earth's atmosphere. To test their individual contributions, we artificially shift each component by arbitrary velocities: $53 \, \mathrm{km \, s^{-1}}$ for the telluric spectrum, $29 \, \mathrm{km \, s^{-1}}$ for the stellar spectrum, and $97 \, \mathrm{km \, s^{-1}}$ for the planetary spectrum. In the resulting CCFs, the spurious feature shifted exactly with the telluric component ($53 \, \mathrm{km \, s^{-1}}$), whereas structures associated with the stellar and planetary spectra did not follow. This isolates the terrestrial component as the origin of the artifact.

2. **Monte Carlo uncertainty analysis.** We construct the CCF standard deviation array by repeating the calculation hundreds of times with Monte Carlo sampling. This reveals a localized peak in uncertainty centered at $0 \, \mathrm{km \, s^{-1}}$, consistent with excess variance introduced by telluric features, supporting the interpretation of a telluric origin.

A representative example of these two diagnostics is shown in Figure A1.

B. CCF VELOCITY BINS CORRELATION: ORIGIN, IDEAL TREATMENT, AND OUR PRACTICAL APPROACH

Here we derive the CCF correlation structure, using classical time-series results on AR(1) processes and variance inflation (e.g., Hamilton 1994; Wilks 2011).

B.1. *Correlation origin and behavior with Δv*

Let the continuum-normalized spectrum at the pixel index p be

$$d_p = s_p + n_p \tag{B1}$$

with signal s_p and noise n_p (zero mean). For a velocity v, we correlate the data with a Doppler-shifted template $t_p(v)$, applying inverse-variance weights w_p to each pixel:

$$y(v) = \sum_{p=1}^{P} \underbrace{\left[w_p\, t_p(v) \right]}_{\phi_p(v)} d_p = \boldsymbol{\phi}(v)^\top \mathbf{d} \tag{B2}$$

Here $\boldsymbol{\phi}(v) \in \mathbb{R}^P$ is the velocity-dependent weight vector, $y(v)$ is the CCF between the data and the template, while $\mathbf{d}$ denotes the vector of artificially observed spectrum values. Sampling $y(v)$ on a velocity grid $\{v_i\}_{i=1}^N$ yields the CCF vector $\mathbf{y} \in \mathbb{R}^N$ with components $y_i = y(v_i)$. Because nearby weight vectors $\boldsymbol{\phi}(v_i)$ and $\boldsymbol{\phi}(v_j)$ are shifted versions of the same template, they overlap strongly. Using the covariance of linear combinations of a random vector, this overlap couples y_i and y_j:

$$\mathrm{Cov}(y_i, y_j) = \mathrm{Cov}\!\left(\boldsymbol{\phi}_i^\top \mathbf{d}, \boldsymbol{\phi}_j^\top \mathbf{d} \right) = \boldsymbol{\phi}_i^\top \mathbf{C}_d\, \boldsymbol{\phi}_j \tag{B3}$$

where $\mathbf{C}_d = \mathrm{Cov}(\mathbf{d})$ is the data covariance matrix, and $\boldsymbol{\phi}_i = \boldsymbol{\phi}(v_i)$.

If we consider white noise (same variance for each pixel, no correlations) for the data, then the data covariance matrix is diagonal: $\mathbf{C}_d = \sigma_n^2 \mathbf{I}$, hence

$$\mathrm{Cov}(y_i, y_j) = \sigma_n^2\, \boldsymbol{\phi}_i^\top \boldsymbol{\phi}_j, \qquad \mathrm{Var}(y_i) = \sigma_n^2 \|\boldsymbol{\phi}_i\|^2 \tag{B4}$$

Using the Pearson correlation coefficient, the correlation between CCF bins is

$$\rho_{ij} = \frac{\mathrm{Cov}(y_i, y_j)}{\sqrt{\mathrm{Var}(y_i)\,\mathrm{Var}(y_j)}} = \frac{\boldsymbol{\phi}_i^\top \boldsymbol{\phi}_j}{\|\boldsymbol{\phi}_i\|\,\|\boldsymbol{\phi}_j\|} \tag{B5}$$

Thus, ρ_{ij} is the normalized dot product between two weight vectors, such that $\rho_{ii} = 1$. If $\boldsymbol{\phi}_j$ is a velocity-shifted version of $\boldsymbol{\phi}_i$, the correlation is the normalized overlap of the template with itself at that shift. Writing a continuous weight function $\phi(x)$ and a velocity lag $\Delta v = v_j - v_i$, the continuum analogue of Equation B5 is the normalized inner product:

$$\rho(\Delta v) = \frac{\int \phi(x)\phi(x - \Delta v)dx}{\int \phi(x)^2 dx} \tag{B6}$$

For a Gaussian-shaped weight, let $\phi(x) = \exp[-x^2/(2\sigma^2)]$. Then

$$\int \phi(x)\phi(x - \Delta v)\, dx = \int \exp\left[-\frac{x^2 + (x - \Delta v)^2}{2\sigma^2} \right] dx \tag{B7}$$

$$= \exp\left[-\frac{(\Delta v)^2}{4\sigma^2} \right] \int \exp\left[-\frac{(x - \Delta v/2)^2}{\sigma^2} \right] dx \tag{B8}$$

and since shifting the center of a Gaussian doesn't change its total area,

$$\boxed{ \rho(\Delta v) \propto \exp\left[-\frac{(\Delta v)^2}{4\sigma^2} \right] } \tag{B9}$$

In practice, our kernel is the autocorrelation of the molecular template, not an ideal Gaussian. For CH_4 and H_2O, the shape is close to Gaussian, while for CO_2 and O_2 it resembles a Gaussian with secondary lobes (see Section 6.2). Nevertheless, $\rho(\Delta v)$ still decays with increasing Δv, as expected from the diminishing overlap between shifted kernels, with the Gaussian case of Equation B9 providing a useful approximation to the general behavior.

B.2. *Ideal treatment: full covariance across velocity bins*

With N CCF samples in $\mathbf{y} \in \mathbb{R}^N$, let $\mathbf{m}(\boldsymbol{\theta})$ be the CCF model with parameters $\boldsymbol{\theta}$. Since the CCF bins are linear combinations of pixel Gaussian noise, they are themselves Gaussian. The correct distribution for a vector of correlated Gaussian variables is the multivariate normal distribution, whose covariance matrix is $\boldsymbol{\Sigma}$. The likelihood is then just the standard formula for a multivariate normal distribution (e.g., Bishop & Nasrabadi 2007):

$$\ln \mathcal{L}(\mathbf{y} \mid \boldsymbol{\theta}) = -\tfrac{1}{2} \Big[\ln |\boldsymbol{\Sigma}| + (\mathbf{y} - \mathbf{m}(\boldsymbol{\theta}))^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y} - \mathbf{m}(\boldsymbol{\theta})) + N \ln(2\pi) \Big] \tag{B10}$$

From Equation B3, let's consider multiple exposures indexed by e. Each exposure has a matrix $\mathbf{F}^{(e)}$, whose i-th row is $\boldsymbol{\phi}^{(e)}(v_i)^\top$, and a noise covariance matrix $\mathbf{C}^{(e)}$ for its spectrum. The combined CCF covariance is then

$$\boldsymbol{\Sigma} = \sum_e \mathbf{F}^{(e)} \mathbf{C}^{(e)} \mathbf{F}^{(e)\top} \tag{B11}$$

This expression follows directly from the linearity of the CCF, and therefore holds analytically. It naturally incorporates both variations in noise across pixels and correlations between neighboring pixels.

Evaluating Equation B10 requires applying $\boldsymbol{\Sigma}^{-1}$ and computing $\ln |\boldsymbol{\Sigma}|$. The standard way is to use a Cholesky factorization $\boldsymbol{\Sigma} = \mathbf{L}\mathbf{L}^\top$, which avoids explicitly inverting $\boldsymbol{\Sigma}$ and allows stable evaluation of the likelihood. A Cholesky factorization costs $\mathcal{O}(N^3)$ time per likelihood call (e.g., Krishnamoorthy & Menon 2011). More efficient strategies can reduce this cost, and we adopt such methods in practice.

B.3. *Practical approach: AR(1) noise inflation*

To obtain reliable uncertainty estimates without explicitly constructing $\boldsymbol{\Sigma}$, we inflate the errors to account for residual correlations. Let $e_i = y_i - m_i$ denote the residuals along the velocity grid, which we approximate as a stationary AR(1) process:

$$e_i = \rho \, e_{i-1} + \varepsilon_i, \qquad |\rho| < 1 \tag{B12}$$

where ε_i is random noise with variance σ_ε^2.

We choose an AR(1) process because each residual bin depends on the previous one, and the condition $|\rho| < 1$ ensures the sequence remains stable. From the variance recursion, we obtain

$$\mathrm{Var}(e_i) = \sigma^2 = \frac{\sigma_\varepsilon^2}{1 - \rho^2}$$

and by induction on h, the lag-h correlation is

$$\mathrm{Corr}(e_i, e_{i+h}) = \rho^h$$

By iterating the AR(1) relation, we further obtain

$$\mathrm{Cov}(e_i, e_j) = \sigma^2 \rho^{|i-j|}$$

Then, we want the variance of the sample mean

$$\bar{e} = \frac{1}{N} \sum_{i=1}^N e_i$$

By definition,

$$\mathrm{Var}(\bar{e}) = \mathrm{Var}\Big(\tfrac{1}{N} \sum_{i=1}^N e_i \Big) = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \mathrm{Cov}(e_i, e_j)$$

We now split the double sum into two cases:
- If $i = j$ (diagonal terms), then $\mathrm{Cov}(e_i, e_j) = \mathrm{Var}(e_i) = \sigma^2$. There are N such terms, contributing $N\sigma^2$.
- If $i \neq j$ (off-diagonal terms), then $\mathrm{Cov}(e_i, e_j) = \sigma^2 \rho^{|i-j|}$. Grouping by lag $h = |i - j|$, for each $h = 1, 2, \ldots, N-1$ there are $(N - h)$ such pairs, and symmetry doubles the contribution.

Hence

$$\text{Var}(\bar{e}) = \frac{1}{N^2}\Big[N\sigma^2 + 2\sum_{h=1}^{N-1}(N-h)\sigma^2\rho^h\Big] = \frac{\sigma^2}{N^2}\Big[N + 2\sum_{h=1}^{N-1}(N-h)\rho^h\Big] = \frac{\sigma^2}{N^2}\Big[N + 2N\sum_{h=1}^{N-1}\rho^h - 2\sum_{h=1}^{N-1}h\rho^h\Big]$$

As $N \to \infty$,

$$\sum_{h=1}^{N-1}\rho^h \to \frac{\rho}{1-\rho}, \qquad \sum_{h=1}^{N-1}h\rho^h \to \frac{\rho}{(1-\rho)^2}$$

The third term is a constant (independent of N), thus, for large N we can neglect it and obtain

$$\text{Var}(\bar{e}) \approx \frac{\sigma^2}{N^2}\Big[N + 2N \cdot \frac{\rho}{1-\rho}\Big]$$

Simplifying gives

$$\text{Var}(\bar{e}) \approx \frac{\sigma^2}{N}\frac{1+\rho}{1-\rho}$$

Hence the standard-deviation inflation needed to compensate for AR(1) correlation is

$$\text{infl}(\rho) = \sqrt{\frac{1+\rho}{1-\rho}} \geq 1 \tag{B13}$$

To estimate the correlation parameter ρ directly from the data, we use the lag-1 sample autocorrelation, following the general definition (Equation B5) applied to the residual series e_i under stationarity, which gives:

$$\rho(1) = \frac{\text{Cov}(e_i, e_{i+1})}{\text{Var}(e_i)}$$

We use the sample autocovariances at lag 0 and 1;

$$\hat{\gamma}(0) = \frac{1}{N}\sum_{i=1}^{N}(e_i - \bar{e})^2, \qquad \hat{\gamma}(1) = \frac{1}{N-1}\sum_{i=1}^{N-1}(e_i - \bar{e})(e_{i+1} - \bar{e})$$

and the lag-1 sample autocorrelation is:

$$\hat{\rho} = \frac{\hat{\gamma}(1)}{\hat{\gamma}(0)}$$

The normalization factors cancel in the ratio for large N, giving the practical formula

$$\hat{\rho} = \frac{\sum_{i=1}^{N-1}(e_i - \bar{e})(e_{i+1} - \bar{e})}{\sum_{i=1}^{N}(e_i - \bar{e})^2}$$

Intuitively, this compares each residual e_i to its immediate neighbor e_{i+1}, normalized by the overall variance of the series. If the residuals are strongly correlated, $\hat{\rho}$ will be close to $+1$, if they are nearly independent, $\hat{\rho} \approx 0$ and if they tend to alternate in sign, $\hat{\rho}$ will be negative.

We finally inflate all per-bin uncertainties used in the Gaussian likelihood by

$$\boxed{\sigma_{i,\text{eff}} = \text{infl}(\hat{\rho})\sigma_i = \sqrt{\frac{1+\hat{\rho}}{1-\hat{\rho}}}\sigma_i} \tag{B14}$$

Putting it together, our pipeline constructs the CCF on the velocity grid, measures the residual lag-1 correlation $\hat{\rho}$, and inflates all per-bin uncertainties by $\sqrt{(1+\hat{\rho})/(1-\hat{\rho})}$ before performing Bayesian model comparison under the assumption of independent Gaussian errors. This strategy does not require building or inverting Σ. The outcome is deliberately conservative Bayes factors, while sidestepping the high computational cost of full covariance treatments and keeping the analysis simple and reproducible.

C. MONTE-CARLO STABILITY AND STOPPING CRITERION

To get an estimate of the scatter in the final S/N, we run the pipeline 100 times per configuration. Here, we assess whether $N = 100$ is statistically adequate.

To verify that $N = 100$ realizations yield a statistically stable estimate of the underlying median detection strength, we compute the standard error (SE) of the sample median and its associated margin of error (MoE) at 95% confidence. Let s denote the Monte Carlo standard deviation (i.e., the between-realization scatter) of the detection statistic. For approximately normal sampling distributions, the standard error of the median is

$$\text{SE} \approx 1.253 \, \frac{s}{\sqrt{N}} \tag{C15}$$

which follows from the asymptotic variance of sample quantiles evaluated for a normal parent distribution (e.g., Walck 1996). For a two-sided 95% confidence interval, we use the standard normal quantile $z_{0.975} = 1.96$:

$$\text{MoE}_{95} = 1.96 \, \text{SE} \quad \text{for } N \geq 100 \tag{C16}$$

For each configuration, we compute MoE_{95} and define a relative, scale-aware metric,

$$\text{rMoE}_{95} = \frac{\text{MoE}_{95}}{\max(|\tilde{x}|, 0.5)} \tag{C17}$$

where $\tilde{x}$ is the sample median of ΔB. We adopt the following stopping rule: if $\text{rMoE}_{95} < 0.15$, the estimate of the median ΔB is deemed sufficiently precise for classification; otherwise we increase N and recompute until this criterion is met. This ensures that the 95% confidence interval for $\tilde{x}$ spans at most 30% of a relevant scale: either the signal level itself (for large $|\tilde{x}|$) or one Jeffreys step (0.5) (for small $|\tilde{x}|$). As a result, strong signals are not over-penalized, while weak signals do not render the ratio in Equation C17 uninformative near zero, ensuring that category assignments are not driven by Monte Carlo noise.

Since $\text{MoE}_{95} \propto N^{-1/2}$ for fixed variance, rMoE_{95} decreases on average with the number of realizations N. The procedure is guaranteed to converge once the target precision is reached, as shown in Figure C2.
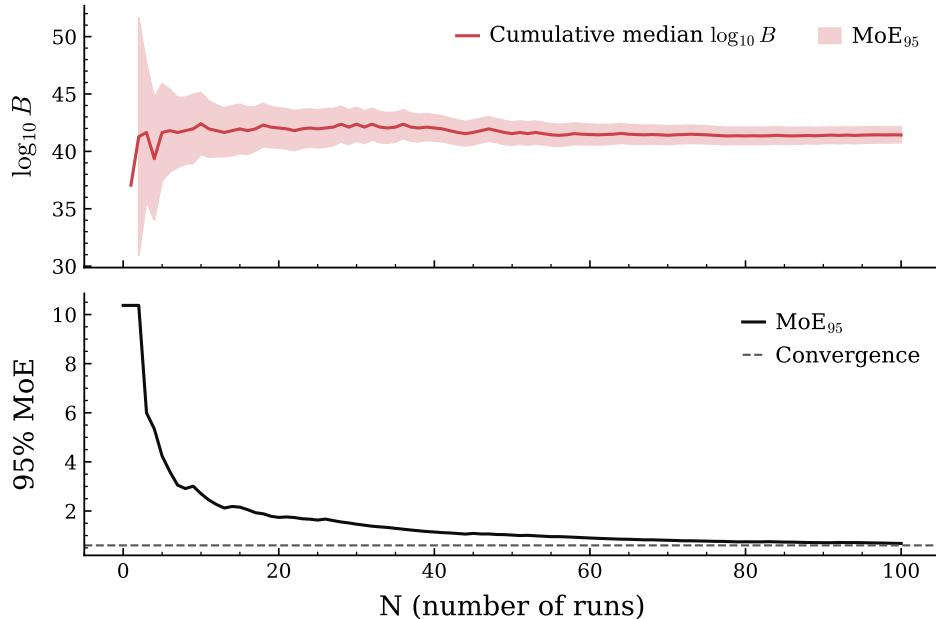

Figure C2. N realizations of the simulation pipeline followed by the Bayesian framework, for CH_4 on TRAPPIST-1 g, assuming we observed all available transits in a 30-year window (112 transits). **Top:** The cumulative median of ΔB stabilizes by $N \approx 80$ realizations while the 95% confidence interval narrows. **Bottom:** The 95% MoE on $\log_{10} B$ decreases as $N^{-1/2}$ and plateaus by $N \approx 100$. At that point the convergence criterion is met, yielding a statistically stable estimate of ΔB.

REFERENCES

Abdurro'uf, Accetta, K., Aerts, C., et al. 2022, ApJS, 259, 35, doi: 10.3847/1538-4365/ac4414

Agol, E., Dorn, C., Grimm, S. L., et al. 2021, PSJ, 2, 1, doi: 10.3847/PSJ/abd022

Akeson, R. L., Chen, X., Ciardi, D., et al. 2013, PASP, 125, 989, doi: 10.1086/672273

Allard, F., & Hauschildt, P. H. 1995, ApJ, 445, 433, doi: 10.1086/175708

Allen, N. H., Espinoza, N., Diamond-Lowe, H., et al. 2025, AJ, 170, 240, doi: 10.3847/1538-3881/adfc51

Astropy Collaboration. 2022, The Astrophysical Journal, 935, 167, doi: 10.3847/1538-4357/ac7c74

August, P. C., Buchhave, L. A., Diamond-Lowe, H., et al. 2025, A&A, 695, A171, doi: 10.1051/0004-6361/202452611

Bainbridge, A. E., & Heidt, L. E. 1966, Tellus, 18, 221, doi: 10.1111/j.2153-3490.1966.tb00230.x10.3402/tellusa.v18i2-3.9677

Barber, R. J., Tennyson, J., Harris, G. J., & Tolchenov, R. N. 2006, MNRAS, 368, 1087, doi: 10.1111/j.1365-2966.2006.10184.x

Bekki, S., & Lefevre, F. 2009, in European Physical Journal Web of Conferences, Vol. 1, European Physical Journal Web of Conferences (EDP), 113, doi: 10.1140/epjconf/e2009-00914-y

Bello-Arufe, A., Knutson, H. A., Mendonça, J. M., et al. 2023, AJ, 166, 69, doi: 10.3847/1538-3881/acd935

Bello-Arufe, A., Damiano, M., Bennett, K. A., et al. 2025, ApJL, 980, L26, doi: 10.3847/2041-8213/adaf22

Berger, T. A., Huber, D., Gaidos, E., & van Saders, J. L. 2018, ApJ, 866, 99, doi: 10.3847/1538-4357/aada83

Bétrémieux, Y., & Kaltenegger, L. 2013, ApJL, 772, L31, doi: 10.1088/2041-8205/772/2/L31

Bishop, C. M., & Nasrabadi, N. M. 2007, Journal of Electronic Imaging, 16, 049901, doi: 10.1117/1.2819119

Bohl, A., Lawrence, L., Lowry, G., & Kaltenegger, L. 2025, arXiv e-prints, arXiv:2501.14054, doi: 10.48550/arXiv.2501.14054

Bonomo, A. S., Dumusque, X., Massa, A., et al. 2023, A&A, 677, A33, doi: 10.1051/0004-6361/202346211

Borucki, W., Thompson, S. E., Agol, E., & Hedges, C. 2019, arXiv e-prints, arXiv:1905.05719, doi: 10.48550/arXiv.1905.05719

Borucki, W. J., Agol, E., Fressin, F., et al. 2013, Science, 340, 587, doi: 10.1126/science.1234702

Brogi, M., Line, M., Bean, J., Désert, J. M., & Schwarz, H. 2017, ApJL, 839, L2, doi: 10.3847/2041-8213/aa6933

Brogi, M., & Line, M. R. 2019, AJ, 157, 114, doi: 10.3847/1538-3881/aaffd3

Brogi, M., Snellen, I. A. G., de Kok, R. J., et al. 2013, The Astrophysical Journal, 767, 27, doi: 10.1088/0004-637X/767/1/27

Buchner, J. 2016, PyMultiNest: Python interface for MultiNest, Astrophysics Source Code Library, record ascl:1606.005

Cadieux, C., Plotnykov, M., Doyon, R., et al. 2024a, ApJL, 960, L3, doi: 10.3847/2041-8213/ad1691

Cadieux, C., Doyon, R., MacDonald, R. J., et al. 2024b, ApJL, 970, L2, doi: 10.3847/2041-8213/ad5afa

Catling, D. C., & Kasting, J. F. 2017, The Structure of Planetary Atmospheres (Cambridge University Press), 3–26

Cauley, P. W., Kuckein, C., Redfield, S., et al. 2018, AJ, 156, 189, doi: 10.3847/1538-3881/aaddf9

Crossfield, I. J. M., Ciardi, D. R., Petigura, E. A., et al. 2016, ApJS, 226, 7, doi: 10.3847/0067-0049/226/1/7

Currie, M. H., Meadows, V. S., & Rasmussen, K. C. 2023, PSJ, 4, 83, doi: 10.3847/PSJ/accf86

Cutri, R., Skrutskie, M., Dyk, S., et al. 2003, VizieR Online Data Catalog

Czesla, S., Schröter, S., Schneider, C. P., et al. 2019, PyA: Python astronomy-related packages, Astrophysics Source Code Library, record ascl:1906.010

Damiano, M., Bello-Arufe, A., Yang, J., & Hu, R. 2024, ApJL, 968, L22, doi: 10.3847/2041-8213/ad5204

de Kok, R. J., Brogi, M., Snellen, I. A. G., et al. 2013, A&A, 554, A82, doi: 10.1051/0004-6361/201321381

Delrez, L., Murray, C. A., Pozuelos, F. J., et al. 2022, A&A, 667, A59, doi: 10.1051/0004-6361/202244041

Diamond-Lowe, H., Kreidberg, L., Harman, C. E., et al. 2022, AJ, 164, 172, doi: 10.3847/1538-3881/ac7807

D'Odorico, V., Bolton, J. S., Christensen, L., et al. 2024, Experimental Astronomy, 58, 21, doi: 10.1007/s10686-024-09967-3

Dransfield, G., Timmermans, M., Triaud, A. H. M. J., et al. 2024, MNRAS, 527, 35, doi: 10.1093/mnras/stad1439

Dressing, C. D., Vanderburg, A., Schlieder, J. E., et al. 2017, AJ, 154, 207, doi: 10.3847/1538-3881/aa89f2

Dubey, D., & Majumdar, L. 2024, ApJ, 972, 165, doi: 10.3847/1538-4357/ad5a97

Dubey, D., Majumdar, L., Beichman, C., et al. 2025, The Astrophysical Journal Supplement Series, 278, 19, doi: 10.3847/1538-4365/adbf05

Edlén, B. 1966, Metrologia, 2, 71, doi: 10.1088/0026-1394/2/2/002

El Amraoui, L., Attié, J. L., Ricaud, P., et al. 2014, Atmospheric Measurement Techniques, 7, 3035, doi: 10.5194/amt-7-3035-201410.5194/amtd-6-6517-2013

24

Espinoza, N., Allen, N. H., Glidden, A., et al. 2025, ApJL, 990, L52, doi: 10.3847/2041-8213/adf42e

Feinstein, A. D., Schlieder, J. E., Livingston, J. H., et al. 2019, AJ, 157, 40, doi: 10.3847/1538-3881/aafa70

Feroz, F., Hobson, M. P., & Bridges, M. 2009, Monthly Notices of the Royal Astronomical Society, 398, 1601, doi: 10.1111/j.1365-2966.2009.14548.x

Fortune, M., Gibson, N. P., Diamond-Lowe, H., et al. 2025, A&A, 701, A25, doi: 10.1051/0004-6361/202554198

Fujii, Y., Angerhausen, D., Deitrick, R., et al. 2018, Astrobiology, 18, 739, doi: 10.1089/ast.2017.1733

Fukui, A., Livingston, J., Narita, N., et al. 2016, AJ, 152, 171, doi: 10.3847/0004-6256/152/6/171

Fulton, B. J., Petigura, E. A., Howard, A. W., et al. 2017, AJ, 154, 109, doi: 10.3847/1538-3881/aa80eb

Gaia Collaboration. 2023, A&A, 674, A1, doi: 10.1051/0004-6361/202243940

Gajdoš, P., Vaňko, M., & Parimucha, Š. 2019, Research in Astronomy and Astrophysics, 19, 041, doi: 10.1088/1674-4527/19/3/41

Gandhi, S., Brogi, M., & Webb, R. K. 2020, MNRAS, 498, 194, doi: 10.1093/mnras/staa2424

Gibson, N. P., Merritt, S., Nugroho, S. K., et al. 2020, MNRAS, 493, 2215, doi: 10.1093/mnras/staa228

Gilbert, E. A., Barclay, T., Schlieder, J. E., et al. 2020, AJ, 160, 116, doi: 10.3847/1538-3881/aba4b2

Gilbert, E. A., Vanderburg, A., Rodriguez, J. E., et al. 2023, ApJL, 944, L35, doi: 10.3847/2041-8213/acb599

Gizis, J. E., Monet, D. G., Reid, I. N., et al. 2000, AJ, 120, 1085, doi: 10.1086/301456

Glidden, A., Ranjan, S., Seager, S., et al. 2025, ApJL, 990, L53, doi: 10.3847/2041-8213/adf62e

Gordon, I. E., Rothman, L. S., Hargreaves, R. J., et al. 2022, JQSRT, 277, 107949, doi: 10.1016/j.jqsrt.2021.107949

Greene, T. P., Bell, T. J., Ducrot, E., et al. 2023, Nature, 618, 39, doi: 10.1038/s41586-023-05951-7

Gressier, A., Espinoza, N., Allen, N. H., et al. 2024, ApJL, 975, L10, doi: 10.3847/2041-8213/ad73d1

Grimm, S. L., & Heng, K. 2015, ApJ, 808, 182, doi: 10.1088/0004-637X/808/2/182

Grimm, S. L., Malik, M., Kitzmann, D., et al. 2021, ApJS, 253, 30, doi: 10.3847/1538-4365/abd773

Hamilton, J. D. 1994, Time Series Analysis (Princeton University Press). http://www.jstor.org/stable/j.ctv14jx6sm

Hardegree-Ullman, K. K., Apai, D., Bergsten, G. J., Pascucci, I., & López-Morales, M. 2023, AJ, 165, 267, doi: 10.3847/1538-3881/acd1ec

Hardegree-Ullman, K. K., Apai, D., Haffert, S. Y., et al. 2025, AJ, 169, 171, doi: 10.3847/1538-3881/adb02f

Hardegree-Ullman, K. K., Cushing, M. C., Muirhead, P. S., & Christiansen, J. L. 2019, AJ, 158, 75, doi: 10.3847/1538-3881/ab21d2

Hargreaves, R. J., Gordon, I. E., Rey, M., et al. 2020, ApJS, 247, 55, doi: 10.3847/1538-4365/ab7a1a

Hill, M. L., Bott, K., Dalba, P. A., et al. 2023, AJ, 165, 34, doi: 10.3847/1538-3881/aca1c0

Husser, T. O., Wende-von Berg, S., Dreizler, S., et al. 2013, A&A, 553, A6, doi: 10.1051/0004-6361/201219058

Jenkins, J. M., Twicken, J. D., Batalha, N. M., et al. 2015, AJ, 150, 56, doi: 10.1088/0004-6256/150/2/56

Jönsson, H., Holtzman, J. A., Allende Prieto, C., et al. 2020, AJ, 160, 120, doi: 10.3847/1538-3881/aba592

Kanodia, S., & Wright, J. 2018, Research Notes of the AAS, 2, 4, doi: 10.3847/2515-5172/aaa4b7

Kasper, D., Bean, J. L., Line, M. R., et al. 2023, AJ, 165, 7, doi: 10.3847/1538-3881/ac9f40

Kass, R. E., & Raftery, A. E. 1995, Journal of the American Statistical Association, 90, 773. http://www.jstor.org/stable/2291091

Kasting, J. F., Whitmire, D. P., & Reynolds, R. T. 1993, Icarus, 101, 108, doi: 10.1006/icar.1993.1010

Kausch, W., Noll, S., Smette, A., et al. 2015, A&A, 576, A78, doi: 10.1051/0004-6361/201423909

Kawahara, H., Matsuo, T., Takami, M., et al. 2012, ApJ, 758, 13, doi: 10.1088/0004-637X/758/1/13

Kiehl, J. T., & Yamanouchi, T. 1985, Tellus Series B Chemical and Physical Meteorology B, 37, 1, doi: 10.3402/tellusb.v37i1.14986

Kipping, D., & Benneke, B. 2025, arXiv e-prints, arXiv:2506.05392, doi: 10.48550/arXiv.2506.05392

Kjærsgaard, R. D., Bello-Arufe, A., Rathcke, A. D., Buchhave, L. A., & Clemmensen, L. K. H. 2023, A&A, 677, A120, doi: 10.1051/0004-6361/202346652

Kopparapu, R. K., Ramirez, R., Kasting, J. F., et al. 2013, ApJ, 765, 131, doi: 10.1088/0004-637X/765/2/131

Krishnamoorthy, A., & Menon, D. 2011, arXiv e-prints, arXiv:1111.4144, doi: 10.48550/arXiv.1111.4144

Krissansen-Totton, J. 2023, ApJL, 951, L39, doi: 10.3847/2041-8213/acdc26

Kuttippurath, J., Pillai Gopikrishnan, G., Müller, R., Godin-Beekmann, S., & Brioude, J. 2024, Atmospheric Chemistry & Physics, 24, 6743, doi: 10.5194/acp-24-6743-2024

Lan, X., Thoning, K., & Dlugokencky, E. 2025, Trends in globally-averaged CH4, N2O, and SF6 determined from NOAA Global Monitoring Laboratory measurements., https://gml.noaa.gov/ccgg/trends/, doi: 10.15138/P8XG-AA10

López-Morales, M., Ben-Ami, S., Gonzalez-Abad, G., et al. 2019, AJ, 158, 24, doi: 10.3847/1538-3881/ab21d7

Luque, R., Coy, B. P., Xue, Q., et al. 2025, AJ, 170, 49, doi: 10.3847/1538-3881/addb40

Lustig-Yaeger, J., Meadows, V. S., & Lincowski, A. P. 2019, AJ, 158, 27, doi: 10.3847/1538-3881/ab21e0

Lustig-Yaeger, J., Stevenson, K., Alam, M., et al. 2025, Charting the Cosmic Shoreline, JWST Proposal. Cycle 4, ID. #7073

Maiolino, R., Haehnelt, M., Murphy, M. T., et al. 2013, arXiv e-prints, arXiv:1310.3163, doi: 10.48550/arXiv.1310.3163

Marconi, A., Abreu, M., Adibekyan, V., et al. 2022, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 12184, Ground-based and Airborne Instrumentation for Astronomy IX, ed. C. J. Evans, J. J. Bryant, & K. Motohara, 1218424, doi: 10.1117/12.2628689

Marconi, A., Abreu, M., Adibekyan, V., et al. 2024, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 13096, Ground-based and Airborne Instrumentation for Astronomy X, ed. J. J. Bryant, K. Motohara, & J. R. D. Vernet, 1309613, doi: 10.1117/12.3017966

Martin, P. E., & Barker, E. F. 1932, Physical Review, 41, 291, doi: 10.1103/PhysRev.41.291

Martins, C. J. A. P., Cooke, R., Liske, J., et al. 2024, Experimental Astronomy, 57, 5, doi: 10.1007/s10686-024-09928-w

Mawet, D., Ruane, G., Xuan, W., et al. 2017, ApJ, 838, 92, doi: 10.3847/1538-4357/aa647f

Mazeh, T., Tamuz, O., & Zucker, S. 2007, in Astronomical Society of the Pacific Conference Series, Vol. 366, Transiting Extrapolar Planets Workshop, ed. C. Afonso, D. Weldrake, & T. Henning, 119, doi: 10.48550/arXiv.astro-ph/0612418

Meadows, V. S., Reinhard, C. T., Arney, G. N., et al. 2018, Astrobiology, 18, 630, doi: 10.1089/ast.2017.1727

Meech, A., Aigrain, S., Brogi, M., & Birkby, J. L. 2022, MNRAS, 512, 2604, doi: 10.1093/mnras/stac662

Meier Valdés, E. A., Demory, B. O., Diamond-Lowe, H., et al. 2025, A&A, 698, A68, doi: 10.1051/0004-6361/202453449

Mollière, P., Wardenier, J. P., van Boekel, R., et al. 2019, A&A, 627, A67, doi: 10.1051/0004-6361/201935470

Monet, D. G., Levine, S. E., Canzian, B., et al. 2003, AJ, 125, 984, doi: 10.1086/345888

Morley, C. V., Kreidberg, L., Rustamkulov, Z., Robinson, T., & Fortney, J. J. 2017, ApJ, 850, 121, doi: 10.3847/1538-4357/aa927b

Morton, T. D., Bryson, S. T., Coughlin, J. L., et al. 2016, ApJ, 822, 86, doi: 10.3847/0004-637X/822/2/86

NASA conference. 1989, in NASA Conference Publication No. 3023, ed. M. J. Prather (Washington, D.C.: National Aeronautics and Space Administration, Office of Management, Scientific and Technical Information Division). https://ntrs.nasa.gov/api/citations/19890011217/downloads/19890011217.pdf

NOAA. 2025, The Atmosphere. https://www.noaa.gov/jetstream/atmosphere

Oman, L., Waugh, D. W., Pawson, S., Stolarski, R. S., & Nielsen, J. E. 2008, Journal of the Atmospheric Sciences, 65, 3278, doi: 10.1175/2008JAS2696.1

Palle, E., Biazzo, K., Bolmont, E., et al. 2023, arXiv e-prints, arXiv:2311.17075, doi: 10.48550/arXiv.2311.17075

Pan, L. L., Randel, W. J., Gary, B. L., Mahoney, M. J., & Hintsa, E. J. 2004, Journal of Geophysical Research (Atmospheres), 109, D23103, doi: 10.1029/2004JD004982

Parmentier, V., & Guillot, T. 2014, A&A, 562, A133, doi: 10.1051/0004-6361/201322342

Parviainen, H., Luque, R., & Palle, E. 2024, MNRAS, 527, 5693, doi: 10.1093/mnras/stad3504

Pass, E., Bean, J. L., Charbonneau, D., Cherubim, C., & Garcia-Mejia, J. 2024, A Search for Exoplanet Satellites that are the Same Size as the Earth's Moon, JWST Proposal. Cycle 3, ID. #6193

PHL@UPR-Arecibo. 2024. http://phl.upr.edu/hwc

Piaulet-Ghorayeb, C., Benneke, B., Turbet, M., et al. 2025, ApJ, 989, 181, doi: 10.3847/1538-4357/adf207

Pickles, A. J. 1998, PASP, 110, 863, doi: 10.1086/316197

Pidhorodetska, D., Fauchez, T. J., Villanueva, G. L., Domagal-Goldman, S. D., & Kopparapu, R. K. 2020, ApJL, 898, L33, doi: 10.3847/2041-8213/aba4a1

Pierrehumbert, R. T. 2010, Principles of Planetary Climate (Cambridge University Press)

Piskunov, N., Rains, A. D., & Boldt-Christmas, L. 2025, arXiv e-prints, arXiv:2509.12737, doi: 10.48550/arXiv.2509.12737

Pueyo, L., Stark, C., Juanola-Parramon, R., et al. 2019, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 11117, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 1111703, doi: 10.1117/12.2530722

26

Quintana, E. V., Barclay, T., Raymond, S. N., et al. 2014, Science, 344, 277, doi: 10.1126/science.1249403

Rackham, B. V., Apai, D., & Giampapa, M. S. 2018, ApJ, 853, 122, doi: 10.3847/1538-4357/aaa08c

Ramirez, R. M. 2018, Geosciences, 8, 280, doi: 10.3390/geosciences8080280

Rathcke, A. D., Buchhave, L. A., de Wit, J., et al. 2025, ApJL, 979, L19, doi: 10.3847/2041-8213/ada5c7

Robinson, T. D., & Catling, D. C. 2014, Nature Geoscience, 7, 12, doi: 10.1038/ngeo2020

Robinson, T. D., Fortney, J. J., & Hubbard, W. B. 2017, ApJ, 850, 128, doi: 10.3847/1538-4357/aa951e

Rodler, F., & López-Morales, M. 2014, ApJ, 781, 54, doi: 10.1088/0004-637X/781/1/54

Roederer, I. U., Alvarado-Gómez, J. D., Allende Prieto, C., et al. 2024, Experimental Astronomy, 57, 17, doi: 10.1007/s10686-024-09938-8

Rothman, L. S., Gordon, I. E., Barber, R. J., et al. 2010, JQSRT, 111, 2139, doi: 10.1016/j.jqsrt.2010.05.001

Rothman, L. S., Gordon, I. E., Babikov, Y., et al. 2013, JQSRT, 130, 4, doi: 10.1016/j.jqsrt.2013.07.002

Rukdee, S. 2024, Scientific Reports, 14, doi: 10.1038/s41598-024-78071-5

Said, A., Davizón, Y. A., Espino-Román, P., Rodríguez-Said, R., & Hernández-Santos, C. 2016, Symmetry, 8, 78, doi: 10.3390/sym8080078

Sanna, N., Canto Martins, B. L., Martins, A. M., et al. 2024, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 13096, Ground-based and Airborne Instrumentation for Astronomy X, ed. J. J. Bryant, K. Motohara, & J. R. D. Vernet, 130964G, doi: 10.1117/12.3018720

Schwieterman, E. W., Kiang, N. Y., Parenteau, M. N., et al. 2018, Astrobiology, 18, 663, doi: 10.1089/ast.2017.1729

Seager, S., & Sasselov, D. D. 2000, ApJ, 537, 916, doi: 10.1086/309088

Seager, S., Schrenk, M., & Bains, W. 2012, Astrobiology, 12, 61, doi: 10.1089/ast.2010.0489

Seager, S., Welbanks, L., Ellerbroek, L., Bains, W., & Petkowski, J. J. 2025, arXiv e-prints, arXiv:2504.12946, doi: 10.48550/arXiv.2504.12946

Serindag, D. B., & Snellen, I. A. G. 2019, ApJL, 871, L7, doi: 10.3847/2041-8213/aafa1f

Smette, A., Sana, H., Noll, S., et al. 2015, A&A, 576, A77, doi: 10.1051/0004-6361/201423932

Smith, P. C. B., Line, M. R., Bean, J. L., et al. 2024, AJ, 167, 110, doi: 10.3847/1538-3881/ad17bf

Snellen, I., de Kok, R., Birkby, J. L., et al. 2015, A&A, 576, A59, doi: 10.1051/0004-6361/201425018

Snellen, I. A. G., de Kok, R. J., le Poole, R., Brogi, M., & Birkby, J. 2013, ApJ, 764, 182, doi: 10.1088/0004-637X/764/2/182

Stassun, K. G., Oelkers, R. J., Paegert, M., et al. 2019, AJ, 158, 138, doi: 10.3847/1538-3881/ab3467

Suissa, G., Wolf, E. T., Kopparapu, R. k., et al. 2020, AJ, 160, 118, doi: 10.3847/1538-3881/aba4b4

Tamuz, O., Mazeh, T., & Zucker, S. 2005, MNRAS, 356, 1466, doi: 10.1111/j.1365-2966.2004.08585.x

Tennyson, J., Bernath, P. F., Brown, L. R., et al. 2013, JQSRT, 117, 29, doi: 10.1016/j.jqsrt.2012.10.002

Thompson, M. A., Krissansen-Totton, J., Wogan, N., Telus, M., & Fortney, J. J. 2022, Proceedings of the National Academy of Science, 119, e2117933119, doi: 10.1073/pnas.2117933119

Thorngren, D. P., Sing, D. K., & Mukherjee, S. 2025, arXiv e-prints, arXiv:2510.00169, doi: 10.48550/arXiv.2510.00169

Torres, G., Kipping, D. M., Fressin, F., et al. 2015, ApJ, 800, 99, doi: 10.1088/0004-637X/800/2/99

Vanderburg, A., Rowden, P., Bryson, S., et al. 2020, ApJL, 893, L27, doi: 10.3847/2041-8213/ab84e5

Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, Nature Methods, 17, 261, doi: 10.1038/s41592-019-0686-2

von Paris, P., Cabrera, J., Godolt, M., et al. 2011, A&A, 534, A26, doi: 10.1051/0004-6361/201117091

von Paris, P., Hedelt, P., Selsis, F., Schreier, F., & Trautmann, T. 2013, A&A, 551, A120, doi: 10.1051/0004-6361/201220009

Walck, C. 1996, Hand-book on statistical distributions for experimentalists (Stockholms universitet). https://api.semanticscholar.org/CorpusID:122813337

Wang, J., Mawet, D., Hu, R., et al. 2018, Journal of Astronomical Telescopes, Instruments, and Systems, 4, 035001, doi: 10.1117/1.JATIS.4.3.035001

Wang, J., Mawet, D., Ruane, G., Hu, R., & Benneke, B. 2017, AJ, 153, 183, doi: 10.3847/1538-3881/aa6474

Wang, J. J., Delorme, J.-R., Ruffio, J.-B., et al. 2021, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 11823, Techniques and Instrumentation for Detection of Exoplanets X, ed. S. B. Shaklan & G. J. Ruane, 1182302, doi: 10.1117/12.2596484

Weiss, L. M., Isaacson, H., Howard, A. W., et al. 2024, ApJS, 270, 8, doi: 10.3847/1538-4365/ad0cab

Western, C. M., Carter-Blatchford, L., Crozet, P., et al. 2018, JQSRT, 219, 127, doi: 10.1016/j.jqsrt.2018.07.017

Wilks, D. S. 2011, Statistical methods in the atmospheric sciences (Amsterdam; Boston: Elsevier Academic Press)

Wright, J. T., & Eastman, J. D. 2014, Publications of the Astronomical Society of the Pacific, 126, 838–852, doi: 10.1086/678541

Wunderlich, F., Scheucher, M., Godolt, M., et al. 2020, ApJ, 901, 126, doi: 10.3847/1538-4357/aba59c

Yang, J., Cowan, N. B., & Abbot, D. S. 2013, ApJL, 771, L45, doi: 10.1088/2041-8205/771/2/L45

Yurchenko, S. N., Owens, A., Kefala, K., & Tennyson, J. 2024, MNRAS, 528, 3719, doi: 10.1093/mnras/stae148

Yurchenko, S. N., & Tennyson, J. 2014, MNRAS, 440, 1649, doi: 10.1093/mnras/stu326

Zhang, H., Wang, J., & Plummer, M. K. 2024, AJ, 167, 37, doi: 10.3847/1538-3881/ad109e

Zieba, S., Kreidberg, L., Ducrot, E., et al. 2024, in European Geosciences Union General Assembly 2024 (EGU24), EGU General Assembly Conference Abstracts, 18489, doi: 10.5194/egusphere-egu24-18489