Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **525,** 1375–1400 (2023) https://doi.org/10.1093/mnras/stad670
Advance Access publication 2023 April 17

Retrieval study of cool directly imaged exoplanet 51 Eri b

Niall Whiteford,[1,2,3]★ Alistair Glasse,[2,3,4]★ Katy L. Chubb,[5]★ Daniel Kitzmann,[6] Shrishmoy Ray,[7] Mark W. Phillips,[8] Beth A. Biller,[2,3] Paul I. Palmer,[3,9] Ken Rice,[2,3] Ingo P. Waldmann,[10] Quentin Changeat,[11] Nour Skaf,[12,13] Jason Wang,[14] Billy Edwards[15] and Ahmed Al-Refaie[10]

[1]*Department of Astrophysics, American Museum of Natural History, Central Park West at 79th Street, NY 10024, USA*
[2]*Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh, EH9 3HJ, UK*
[3]*Centre for Exoplanet Science, University of Edinburgh, Edinburgh, EH9 3HJ, UK*
[4]*UK Astronomy Technology Centre, Royal Observatory, Edinburgh, Blackford Hill, Edinburgh, EH9 3HJ, UK*
[5]*Centre for Exoplanet Science, University of St Andrews, North Haugh, St Andrews, KY16 9SS, UK*
[6]*Center for Space and Habitability, University of Bern, Gesellschaftsstrasse 6, Bern, 3012, Switzerland*
[7]*Astrophysics Group, University of Exeter, Exeter, EX4 4QL, UK*
[8]*Institute for Astronomy, University of Hawaii, 2680 Woodlawn Drive, Honolulu, HI 96822, USA*
[9]*School of GeoSciences, University of Edinburgh, Edinburgh, EH8 9XP, UK*
[10]*Department of Physics and Astronomy, University College London, Gower Street, WC1E 6BT, UK*
[11]*European Space Agency, Space Telescope Science Institute, 3700 San Martin Dr, Baltimore, MD 21218 , USA*
[12]*LESIA, Observatoire de Paris, Université PSL, CNRS, Sorbonne Université, Université de Paris, 5 place Jules Janssen, 92195 Meudon, France*
[13]*Subaru Telescope, National Astronomical Observatory of Japan, 650 North A'Ohoku Place, Hilo, HI 96720, USA*
[14]*Center for Interdisciplinary Exploration and Research in Astrophysics (CIERA) and Department of Physics and Astronomy, Northwestern University, Evanston, IL 60208, USA*
[15]*SRON Netherlands Institute for Space Research, Sorbonnelaan 2, 3584 CA, Utrecht, Netherlands*

Accepted 2022 December 21. Received 2022 November 15; in original form 2022 September 12

ABSTRACT

Retrieval methods are a powerful analysis technique for modelling exoplanetary atmospheres by estimating the bulk physical and chemical properties that combine in a forward model to best fit an observed spectrum, and they are increasingly being applied to observations of directly imaged exoplanets. We have adapted TAUREX3, the Bayesian retrieval suite, for the analysis of near-infrared spectrophotometry from directly imaged gas giant exoplanets and brown dwarfs. We demonstrate TAUREX3's applicability to sub-stellar atmospheres by presenting results for brown dwarf benchmark GJ 570D which are consistent with previous retrieval studies, whilst also exhibiting systematic biases associated with the presence of alkali lines. We also present results for the cool exoplanet 51 Eri b, the first application of a free chemistry retrieval analysis to this object, using spectroscopic observations from GPI and SPHERE. While our retrieval analysis is able to explain spectroscopic and photometric observations without employing cloud extinction, we conclude this may be a result of employing a flexible temperature-pressure profile which is able to mimic the presence of clouds. We present Bayesian evidence for an ammonia detection with a 2.7σ confidence, the first indication of ammonia in a directly imaged exoplanetary atmosphere. This is consistent with this molecule being present in brown dwarfs of a similar spectral type. We demonstrate the chemical similarities between 51 Eri b and GJ 570D in relation to their retrieved molecular abundances. Finally, we show that overall retrieval conclusions for 51 Eri b can vary when employing different spectral data and modelling components, such as temperature–pressure and cloud structures.

Key words: data analysis – Brown dwarfs – atmospheres – gaseous planets.

1 INTRODUCTION

While over 5000 exoplanets have been confirmed to date (Akeson et al. 2013; NASA Exoplanet Archive 2022), only a very small fraction have been directly imaged due to the significant technical challenge of detecting a signal from an exoplanet many times fainter than its host star. However, extreme coronagraphic spectrometers, including VLT's Spectro-Polarimetric High-contrast Exoplanet REsearch instrument (SPHERE) (Beuzit et al. 2008), the Gemini Planet Imager (GPI) (Macintosh et al. 2014), and VLT's GRAVITY (Gravity Collaboration 2017), have made it possible to start the characterization and classification effort of directly imaged exoplanet demographics (Nielsen et al. 2019; Vigan et al. 2021). A summary of direct imaging spectroscopy is covered extensively in Biller & Bonnefoy (2018).

The development of extrasolar planetary spectroscopy (see Tinetti, Encrenaz & Coustenis 2013) has mainly been driven by studies of transiting hot-Jupiters and has allowed for unprecedented insight into the diversity of their atmospheres. This led to the expansion and application of inverse atmospheric modelling techniques (outlined in Fig. 1) to exoplanetary spectra (see Line et al. 2013 for a review of early exoplanetary retrieval codes). There are now a variety of

★ E-mail: niallwhiteford@gmail.com (NW); alistair.glasse@stfc.ac.uk (AG); katy.chubb.14@alumni.ucl.ac.uk (KLC)

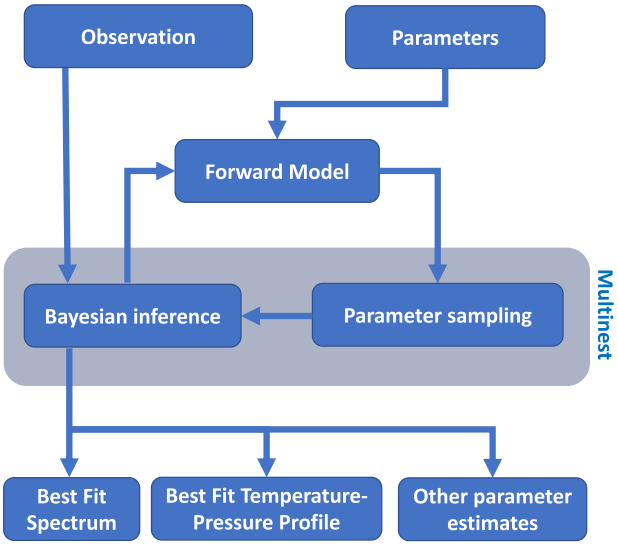


Figure 1. Flowchart of the key components used in inverse retrieval techniques.

retrieval codes developed for exoplanet atmospheric characterization, examples include NEMESIS (Irwin et al. 2008), CHIMERA (Line et al. 2013), BART (Cubillos et al. 20226), SCARLET (Benneke 2015), POSEIDON (MacDonald & Madhusudhan 2017), BREWSTER (Burningham et al. 2017, 2021), HYDRA (Gandhi & Madhusudhan 2018), PETITRADTRANS (Mollière et al. 2019; Nowak et al. 2020; Mollière et al. 2020), PLATON II (Zhang et al. 2020), HELIOS-R2 (Kitzmann et al. 2020), APOLLO (Howe, McElwain & Mandell 2022) and TAUREX3 (Al-Refaie et al. 2021, 2022). In previous studies, TAUREX has been applied to observations of transiting exoplanets (Waldmann et al. 2015a, b; Tsiaras et al. 2016; Rocchetto et al. 2016; Tsiaras et al. 2018, 2019; Changeat et al. 2019; Edwards et al. 2020; Skaf et al. 2020; Pluriel et al. 2020), with a comparative study of TAUREX, CHIMERA, and NEMESIS retrieval codes to be found in Barstow et al. (2020) with a review of the current state-of-the-art in Barstow & Heng (2020) and Madhusudhan (2019).

There is now an abundance of literature outlining the application of the retrieval approach to directly imaged exoplanet and brown dwarf spectroscopy or photometry. This includes HR8799b (Lee, Heng & Irwin 2013), GJ 570D (Line et al. 2014), GJ 570D and HD 3651B (Line et al. 2015), 11 T dwarfs (Line et al. 2017), HR8799b-e (Lavie et al. 2017), 2MASS J05002100+0330501 and 2MASS J2224438−015852 (Burningham et al. 2017), GJ 570D and the Epsilon Indi brown dwarf binary system (Kitzmann et al. 2020), 6 T and 8 Y dwarfs (Zalesky et al. 2019), β Pic b (Gravity Collaboration 2020), HR 8799e (Mollière et al. 2020), HR 8799c (Wang et al. 2020), and the SDSS J1416+1348AB binary (Gonzales et al. 2020). Most recently, it has been employed for analysis of SDSS J125637.13−022452.4 (Gonzales et al. 2021), a 'cloud busting' study of 2MASS 2224−0158, studying the L–T transition (Lueber et al. 2022), analysing high-resolution observations of HD 4747 B (Xuan et al. 2022) and a population analysis of 50 Late-T Dwarfs (Zalesky et al. 2022). Here, we use the TAUREX3 retrieval tool to carry out analysis of directly imaged exoplanet 51 Eridani b (hereafter, 51 Eri b) and brown dwarf benchmark GJ 570D.

Despite the significant development in the field of directly imaged exoplanet spectroscopy in the last decade, upcoming telescopes will prove essential to further our understanding of these objects. The JWST (Gardner et al. 2006) and the soon to be constructed Extremely Large Telescope (ELT) (Udry et al. 2014; Brandl et al. 2014), will lead to increased observational capacity, requiring refined and robust analysis techniques. Retrieval tools will be a corner stone for the analysis of these next generation observations.

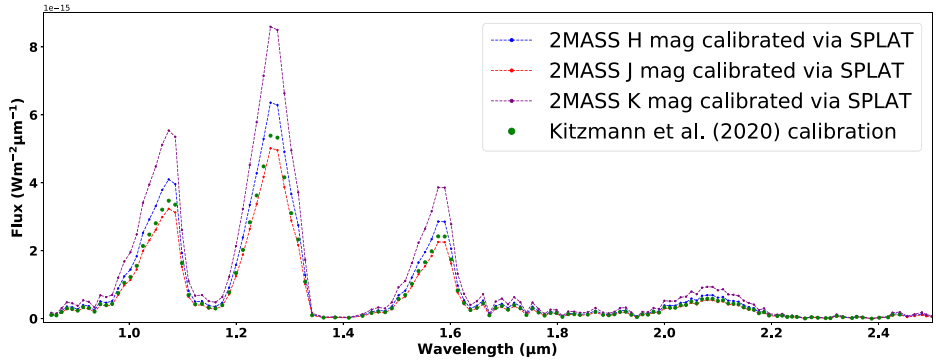
Facilitated by the aforementioned instruments, direct imaging will be a very important technique for the future with the notable benefits that it offers when compared to the currently dominant technique of transmission spectroscopy. These include the ability to view exoplanet and brown dwarf atmospheres as they rotate (Crossfield et al. 2014) (as they are not tidally locked) and being able to probe further into the atmosphere, unlocking more spectral features. The currently observed selection of directly imaged exoplanets are limited to young gas-giants which orbit their host stars at large radial distances. They show similar properties to free-floating planetary mass objects and old field brown dwarfs. As a result, these three subsets of object can have the same spectral types. The youngest low surface gravity objects start out as a hot L spectral type, evolving via cooling firstly to a T type (Kirkpatrick 2005) before finally becoming a very cool Y type (Cushing et al. 2011; Kirkpatrick et al. 2012; Miles et al. 2020), at the limits of current observational capabilities. In this study, we will be focusing on T spectral type objects, with their atmospheric signatures dominated by H_2O and CH_4 absorption.

The importance of cloud modelling for directly imaged exoplanets and brown dwarfs has been well explored and debated (Marley, Saumon & Goldblatt 2010; Morley et al. 2012; Marley et al. 2012; Lee et al. 2013; Morley et al. 2014; Chilcote et al. 2017; Burningham et al. 2017; Charnay et al. 2018; Bowler et al. 2020; Zhou et al. 2020; Lew et al. 2020; Mollière et al. 2020; Burningham et al. 2021). Previous studies of 51 Eri b, for example, used clouds in their grid modelling (Rajan et al. 2017; Samland et al. 2017) to successfully fit the planet's spectral energy distribution (SED). For a recent and extensive review of exoplanet clouds, see Helling (2019). Alternative explanations for the observed SEDs have been explored in Tremblin et al. (2016, 2017), who demonstrated that a reduced atmospheric temperature gradient can reproduce the SEDs of late L and T type brown dwarfs, without the need to invoke clouds. The mechanism reducing the temperature gradient in these atmospheres has been proposed to be diabatic convection triggered by the CO/CH_4 chemical conversion in brown dwarf atmospheres (Tremblin et al. 2019).

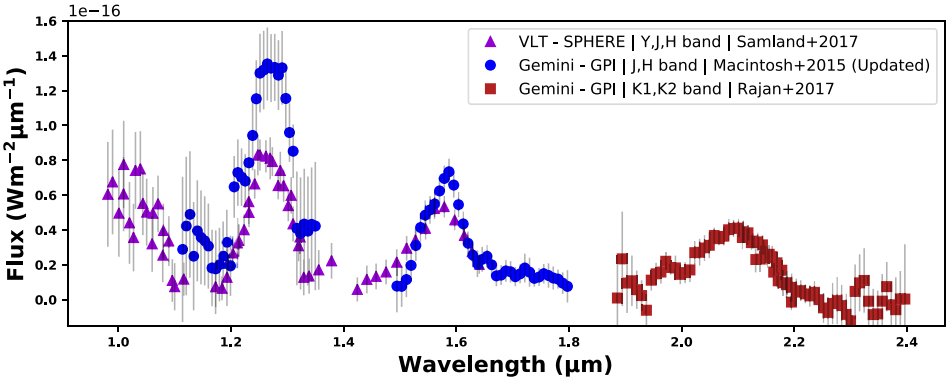
We now describe the specifics of both the retrieval tool and other tools used in our spectral analysis, as applied to spectra of 51 Eri b and GJ 570D.

2 TWO BENCHMARK T DWARFS: GJ 570D AND 51 ERI B OBSERVATIONS

In this section, we give a brief overview of our current knowledge and understanding of GJ 570D and 51 Eri b as well as describing the origin of the data used in their model fitting analysis. We chose to focus on T dwarfs in this first application of TAUREX3 to directly imaged targets as in this temperature regime their SED's are thought to be less influenced by clouds, which are expected to exist below the observable photosphere (Burrows et al. 1997; Burrows & Sharp 1999; Lodders & Fegley 2006). The inclusion of GJ 570D allows us to benchmark TAUREX3 against previous studies using other retrieval codes (Line et al. 2015; Burningham et al. 2017; Kitzmann et al. 2020). 51 Eri b offers a comparable spectral type object but allows us to investigate a completely different mass regime and it has no existing free-chemistry retrieval analysis. We note that clouds seem to be more prominent in the observable atmosphere for low surface gravity objects such as 51 Eri b (Marley et al. 2012; Charnay et al. 2018).



(a) GJ 570D



(b) 51 Eri b

Figure 2. (a): SpeX prism spectrum of GJ 570D, flux calibrated using SPLAT and the object's 2MASS *J*-, *H*- and *K*-band magnitudes. The spectrum used by Kitzmann et al. (2020), produced using a different absolute flux calibration approach, is included for comparison. (b) Published data for 51 Eri b. We include *Y*-, *J*- and *H*-band SPHERE data from Samland et al. (2017), along with the GPI *J* and *H* data from Macintosh et al. (2015) (which is updated using a revised stellar flux and presented in Rajan et al. 2017) along with GPI *K*1 and *K*2 band data from Rajan et al. (2017). There is a clear difference in the *J*-band brightness, and also a difference in the *H*-band brightness, between the GPI and SPHERE observations.

2.1 GJ 570D

GJ 750D (or 2MASS J14571496−2121477) is a cool T7.5 brown dwarf, with an age of 1–5 Gyr (Liu, Leggett & Chiu 2007), and was among the first T dwarf companions to be discovered by Burgasser et al. (2004), Burgasser, Burrows & Kirkpatrick (2006). It is a very wide component in a hierarchical quadruple system, comprising the inner spectroscopic binary companions GJ 570B and C and the primary GJ 570A from which GJ 570D orbits at a projected separation of 1525 ± 25 au (Burgasser et al. 2000). GJ 570D has been included in order to compare TAUREX3 against other retrieval studies as it has become commonly included in novel retrieval approach validations (Line et al. 2014, 2015, 2017; Burningham et al. 2017; Kitzmann et al. 2020; Piette & Madhusudhan 2020). It also offers the opportunity to compare retrieval results against studies using grid model fitting. GJ 570D has a comparable spectral type to the exoplanet 51 Eri b, also included in this study.

2.1.1 *Observations and calibration*

We used observations of GJ 570D taken by the SpeX spectrograph (Rayner et al. 2003), which is mounted on the 3 m NASA InfraRed Telescope Facility. The measured spectrum is part of the SpeX Prism Library (Burgasser 2014) and was first published in Burgasser et al. (2004). The data were reduced using the pipeline described in Cushing, Vacca & Rayner (2004), with the spectrum spanning 0.65 to 2.56 μm at an average spectral resolving power of 120. Using the SpeX Prism Library[1] data analysis toolkit (SPLAT[2]; see Burgasser & Splat Development Team 2017 for details), we flux calibrated the data using photometry from the 2MASS survey[3] (Skrutskie et al. 2006).

The spectra shown in Fig. 2 a have then been calibrated using *J* (15.324 ± 0.05 mag), *H* (15.268 ± 0.09 mag) and *K*-band (15.242 ± 0.16 mag) fluxes. In the following analysis, we used the spectrum calibrated using the *H*-band magnitude. As outlined in Line et al. (2015), neighbouring pixels may not be statistically independent, due to the duplication of flux information. Therefore, when analysing this data set we only include every third data point (pixel) in our model fitting.

2.2 51 Eri b

51 Eri b was the first exoplanet discovered by GPI (Macintosh et al. 2014), and has one of the smallest angular and physical separations

[1]SpeX Prism Library: http://pono.ucsd.edu/~adam/browndwarfs/spexprism/library.html
[2]SPLAT: http://pono.ucsd.edu/~adam/browndwarfs/splat/
[3]2MASS Survey Archive: https://irsa.ipac.caltech.edu/Missions/2mass.html

(∼0.5 arcsec, ∼13 au) of any directly imaged exoplanet. It orbits a young F0-type host, with age estimates of 20 ± 6 Myrs from Macintosh et al. (2014) and 26 ± 3 Myrs from Nielsen et al. (2016). With a spectral type of T6.5 ± 1.5 (Rajan et al. 2017), 51 Eri b is notably the latest spectral type planet yet imaged.

Exhibiting methane absorption (a first for directly imaged exoplanets) with its lower effective temperature (∼700 K) and low mass (<10 M_{Jup}), 51 Eri b defined a new category of directly imaged exoplanets. Further, its SED indicates that the L/T transition occurs at lower temperatures for these lower surface gravity objects compared to the higher surface gravity brown dwarfs (Rajan et al. 2017).

Studies of this exoplanet have included clouds in order to fit the spectroscopic and photometric data. Rajan et al. (2017) used two self-consistent grid models, one with a patchy iron/silicate cloud component, and the other with sulfide/salt cloud to explain the spectral profile, while Samland et al. (2017) used grid models produced using petitCODE (Mollière et al. 2015, 2017) which employed a slightly modified version of the Ackerman & Marley (2001) prescription in their cloud modelling. Samland et al. (2017) also tested the Morley et al. (2012) cloud models against their observations. Samland et al. (2017) could not differentiate between patchy and uniform clouds while Rajan et al. (2017) found a preference for patchy iron/silicate clouds in the model fitting. Both studies concluded that clouds were needed to fit the spectrum well. Neither Samland et al. (2017) or Rajan et al. (2017), however, employed a free chemistry model as we have done in this study.

2.2.1 *Observations*

In this study, we used a combination of observations of 51 Eri b from 2015–2016. These included spectroscopic data taken with GEMINI-GPI's Integral Field Spectrograph (Macintosh et al. 2014) (IFS) in the *J*-, *H*-, *K*1- and *K*2-bands (Rajan et al. 2017) (where *J*- and *H*-band observations are updated from Macintosh et al. 2015) and VLT-SPHERE's IFS (Beuzit et al. 2008, 2019) using its *YJ*, *YH* filters (Samland et al. 2017). The spectra are shown in Fig. 2b, calibrated as outlined in Samland et al. 2017 and Rajan et al. 2017.

We also employed photometric measurements from KECK-NIRC2's (McLean & Sprayberry 2003) Lp and Ms filters (Rajan et al. 2017), where we used two combinations of data for our analyses: one which combined SPHERE's *Y*-, *J*- and *H*-bands along with GPI's *K*1- and *K*2-band data and the other which combined only the GPI bands. We used this approach as the aforementioned GPI and SPHERE observations differed significantly in brightness in both the *J*- and *H*-bands.

Unlike with the GJ 570D data, we did not exclude any data from the analysis. This was motivated by the data's already low spectral resolution, combined with the relatively large errors, where exclusion of data would severely impact the retrievals ability to constrain parameters. We note that the potential for correlated noise to impact the retrieval is more prominent when using these full data sets.

3 MODELLING

3.1 Retrieval method overview

In this study, we employ the inverse retrieval method. This approach is reviewed in Fortney (2018) and Madhusudhan (2018) and is outlined in Fig. 1. In its simplest form, the technique obtains a best fit to observed spectra using a varying forward model defined by a handful of constraining parameters. Variations in the forward model explore the permitted parameter space while statistically deriving the best fit to an observation. Inverse techniques calculate the posterior distributions of these forward model parameters that best fit the observed data.

The forward model contains a set of input parameters and converts them into the observable format; typically a spectrum at a specified spectral resolution. For the purposes of this paper, we replicate observations of thermal emission in the near and mid-infrared. The model combines fundamental quantities such as the molecular chemistry and the temperature–pressure profile present in the observed atmosphere, the radius and mass of the object, and the existence of clouds or aerosol opacities.

Retrieval models generally use a Bayesian sample approach to select the best-fitting model. As outlined below, the use of Bayesian retrievals allows the formal inclusion of prior knowledge and full exploration of the likelihood probability distribution of the data. Bayesian retrievals have become the norm in atmospheric analyses of transmission and secondary eclipse spectra of transiting exoplanets. However, Bayesian parameter inference has now regularly proven itself as an effective tool in the pursuit of statistically rigorous exoplanet and brown dwarf atmospheric characterizations in the field of direct imaging.

3.2 TAUREX3

TAUREX3 (Tau Retrieval of Exoplanets) is a publicly available[4] Bayesian retrieval code designed to be applied to spectroscopic observations of extrasolar atmospheres (Waldmann et al. 2015a, b; Al-Refaie et al. 2021). It can be employed to analyse emission, transmission, and phase-curve spectroscopic data. Fig. 1 gives an overview of the TAUREX code for emission retrieval. The following subsections outline the key components of TAUREX3.

3.2.1 *Atomic and molecular cross-sections*

TAUREX has its own purpose built molecular and atomic opacities, which can be accessed from the publicly available ExoMolOP data base (Chubb et al. 2020).[5] The line lists used for this data base originate mainly from the ExoMol project (Tennyson & Yurchenko 2012) but also HITEMP (Rothman et al. 2010), HITRAN (Rothman et al. 1987), and MoLLIST (Bernath 2020). This includes the latest line lists for TiO (McKemmish et al. 2019), H_2O (Polyansky et al. 2018), CO (Li et al. 2015), CO_2 (Yurchenko et al. 2020), CH_4 (Yurchenko et al. 2017), VO (McKemmish, Yurchenko & Tennyson 2016), H_2S (Azzam et al. 2016), and NH_3 (Coles et al. 2019). ExoMol provides line lists for extended temperature ranges for a variety of molecules. The pressure and temperature broadened profiles for the resonance doublets of Na and K are computed using methods described in Allard, Spiegelman & Kielkopf (2016) and Allard et al. (2019). All other line data for these atomic species are taken from either the NIST (Kramida, Ralchenko & Reader 2013) or Kurucz (Kurucz & Bell 1995) data base (see Section 4.1). We note here that all the results (Tables and Figures) presented in this study were retrieved using the broadening parameters of Allard et al. and non-resonance lines from the Kurucz data base, unless stated otherwise.The TAUREX3 cross-sections used in this work were sampled at $R = \frac{\lambda}{\Delta\lambda} = 15\,000$ across the 0.3–50 μm wavelength region. For a more detailed discussion of TAUREX's line list library,

[4]TAUREX3: https://github.com/ucl-exoplanets/TauREx3_public
[5]ExoMolOP: http://exomol.com/data/data-types/opacity/

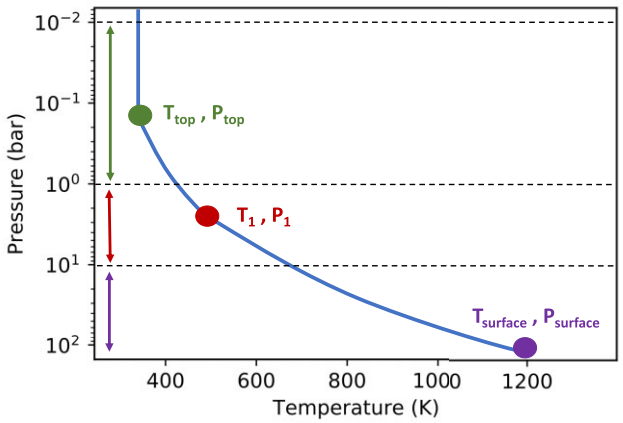


Figure 3. *npoint* temperature–pressure profile used in our analysis. The figure outlines the structure of the n-point profile, with n = 3 in this example.

see Chubb et al. 2020. During the molecular and atomic radiative transfer calculations performed by TAUREX, the model is produced at a much higher resolution than that of the observed spectrum. These high resolution spectra are then binned down to the data resolution in order to calculate the loglikelihood.

3.2.2 *Temperature-pressure profiles*

To accurately model directly imaged emission spectroscopy, an appropriate temperature–pressure parametrization must be adopted. TAUREX offers a variety of temperature–pressure profile options, ranging from radiative two-stream modelling such as the Guillot (2010) prescription (for highly irradiated planets) to more ad hoc geometric approaches in which temperature–pressure nodes are allowed to vary freely. In Fig. 3, we illustrate the $n-point$ temperature–pressure profile adopted in our analysis.

The $n-point$ profile is determined by several parameters including the top of atmosphere temperature, $T_{\rm top}$, and top of atmosphere pressure, $P_{\rm top}$ (set at 10^{-3} bar in this study). The other parameters include the tropopause temperature and pressure, T_1 and P_1, as well as the surface pressure $P_{\rm surf}$ (set at 500 bar for this study) and temperature $T_{\rm surf}$.Temperatures are then linearly interpolated between these temperature–pressure nodes in log space.

In order to be able to compare retrievals with temperature–pressure profiles of differing degrees of freedom, we also added with a less flexible profile following a simple parametrization employed in Lavie et al. (2017). This originated from a reduced version of equation 126 in Heng, Mendonça & Lee (2014):

$$T^4 = \frac{T_{\rm int}^4}{4}\left(\frac{8}{3} + \widetilde{m}\kappa_0\right). \tag{1}$$

where $T_{\rm int}$ is the internal temperature and κ_0 the constant component of the infrared opacity. $\widetilde{m}$ is column density determined via $P_0 = \widetilde{m} \cdot g$ with g being the surface gravity at the bottom of our model atmosphere (500 bar). This simpler profile parametrization only has two free parameters within our retrievals: κ_0 and $T_{\rm int}$.

3.2.3 *Effective temperature calculation*

We have included the effective temperature $T_{\rm eff}$ as a derived parameter which is useful for comparing retrieval results to grid models and evolutionary tracks. For this, we followed the same approach as adopted in Line et al. (2015), integrating the spectrum from 0.1 to 50 μm (at the native resolution of the input cross-sections) to calculate the total emission flux. The effective temperature, as associated uncertainties, is then derived using the Stefan–Boltzmann law and a random sampling of 10 per cent of models ran.

3.2.4 *Bayesian Analysis*

TAUREX employs Bayesian statistics as the cornerstone for the retrieval analysis. Bayes' theorem states that:

$$P(\theta \mid x, \mathcal{M}) = \frac{P(x \mid \theta, \mathcal{M})\, P(\theta, \mathcal{M})}{P(x \mid \mathcal{M})}, \tag{2}$$

where $P(\theta, \mathcal{M})$ is the Bayesian prior, and $\mathcal{M}$ is the forward model. $P(\theta \mid x, \mathcal{M})$ is the posterior probability of the model parameters θ given the data, x, assuming the forward model $\mathcal{M}$. The likelihood, $P(x \mid \theta, \mathcal{M})$ is given by:

$$P(x \mid \theta, \mathcal{M}) = \frac{1}{\mathcal{E}\sqrt{2\pi}} \exp\left[-\frac{1}{2}\sum_{\lambda}^{N}\left(\frac{x_\lambda - \mathcal{M}_\lambda}{\mathcal{E}_\lambda}\right)^2\right], \tag{3}$$

where $\mathcal{E}$ is the error on the input spectral data. This is defined via:

$$\mathcal{E}_\lambda^2 = \sigma_\lambda^2 + 10^b, \tag{4}$$

where σ_λ is the measured error for the λth flux and b is a tolerance factor which is included as a free parameter in the retrieval analysis (Tremaine et al. 2002; Hogg, Bovy & Lang 2010; Foreman-Mackey et al. 2013). This 10^b factor has been used extensively throughout the literature within various retrieval frameworks across many data sets (Line et al. 2015, 2017; Burningham et al. 2017, 2021).

The 10^b error inflation term can account for imperfections in the forward model's capability to fit the observed emission spectrum and/or account for underestimated uncertainties. It also, more demonstrably, allows for the down-weighting of sections of a spectrum (in our case the *K*-band) where the spectral resolution is highest as well as possessing the smallest error bars. Such sections of data can lead to the neglect of other important parameters driving regions of a spectrum, such as in the case of our GJ 570D data. Including the error inflation can therefore allow for a more equally weighted consideration of the whole spectrum when performing the Bayesian evidence calculations. This is discussed more in Section 3.2.9.

3.2.5 *Nested sampling via MULTINEST*

TAUREX includes the implementation of Bayesian statistics via nested sampling (NS) using MULTINEST (Feroz & Hobson 2008; Feroz, Hobson & Bridges 2009; Feroz et al. 2013) via PYMULTINEST (Buchner et al. 2014). NS derives the Bayesian Evidence given by:

$$E = \int P(\theta \mid \mathcal{M}) P(x \mid \theta, \mathcal{M}) \mathrm{d}\theta, \tag{5}$$

where $E = P(x \mid \mathcal{M})$ is the Bayesian Evidence which allows for formal model selection. The statistical results from MULTINEST are then used to derive the parameter estimates which combine to produce the highest Log-Evidence. Using MULTINEST, we sampled the parameter space using 3000–5000 live points at a sampling efficiency of 0.8 which is the default for parameter estimation.

Via the nested sampling Log-Evidence, we can compare model results using the Bayes Factor B:

$$log(b) = \Delta log(Ev) = Log(Ev2) - log(Ev1), \tag{6}$$

This is a ratio of evidence of two competing models (Ev1 and Ev2), allowing for comparison. Table 1, from Kass & Raftery (1995), outlines how $log\,(B)$ can be interpreted.

Table 1. Interpreation of the Bayes ratio outlined in Kass & Raftery 1995.

log (*b*)	Interpretation
0–0.5	*No Evidence*
0.5–1	*Some Evidence*
1–2	*Strong Evidence*
>2	*Decisive*

3.2.6 *Priors*

TAUREX has preset default priors set for all the possible free parameters. This includes all that are necessary for the forward model such as mass, radius, temperature–pressure prescription, and atmospheric trace gases considered. By default TAUREX employs uninformative priors with large prior ranges (e.g. trace-gas abundance priors are log-uniform, log (abundances) $= 1.0 - 1.0 \times 10^{-12}$). The default values can be manually overridden, allowing the user to limit or open up the parameter space. Narrowly defined bounds have the benefit of reducing computational expense but run the risk of being overly restrictive. The priors and prior bounds set for the retrieval analysis performed in this paper were either uniform, log-uniform, or Gaussian priors based on values from previous published studies (when such values were available). See Table 2 for a full overview of the priors set.

Given the lower quality of the 51 Eri b data, we adopted an informative Gaussian prior for our retrievals. This was based on the system age estimate from Rajan et al. (2017) and the evolutionary tracks from Fortney et al. (2008) as shown in Fig. 4. We did not adopt a Gaussian prior on the mass as the reported values in the literature (Macintosh et al. 2015; Rajan et al. 2017; Samland et al. 2017; Nielsen et al. 2019) have a large spread in the planetary mass regime.

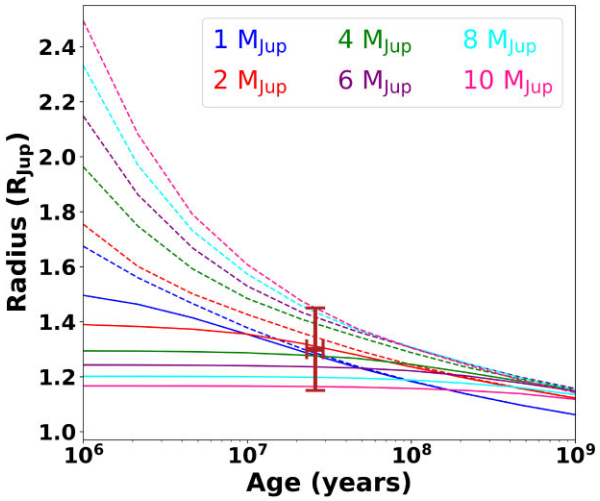


Figure 4. Evolutionary tracks from Fortney et al. (2008) with age uncertainty of 51 Eri system from Rajan et al. (2017) indicated. The Gaussian radius prior we adopt for the 51 Eri b analysis is also indicated.

3.2.7 *Mode for direct imaging*

We modified TAUREX3 to allow us to model directly imaged targets. First, we removed stellar emission from the forward model and added an inverse square law scaling for the exoplanet or brown dwarf

Table 2. Table of retrieval priors. The middle sections list the parameters used for the $n - point$ and Lavie et al. (2017) temperature–pressure profiles, while the bottom section outlines the parameters used in the deck and slab cloud schemes.

Retrieved parameter	Distribution type	GJ 570D Bounds	51 Eri b Bounds
Mixing Ratio	Log-Uniform	1e-12–1e-1	1e-12–1e-1
Radius	Uniform | Gaussian	0.5–2.0 R_{Jup}	$1.3 \pm 0.15\ R_{Jup}^{1,2}$
Mass	Uniform	13–80 M_{Jup}	1–13 M_{Jup}^{3}
Distance	Gaussian	5.8819 ± 0.0029 pc[4]	29.4 ± 0.3 pc[5]
S_{cal}	Gaussian	1.0, 0.1 STD	1.0, 0.1 STD
10^b	Uniform	0.01 x min(σ_λ^2) $\leq 10^b \leq$ 100 x min(σ_λ^2)	–
T_{surf}	Uniform	1250–2500 K	1250–2500 K
P_{surf}	Log-Uniform	5e2–1e1 bar	5e2–1e1 bar
T_1	Uniform	100–2000 K	100–2000 K
P_1	Log-Uniform	1e1–1e-1 bar	1e1–1e-1 bar
T_{top}	Uniform	0–1000 K	0–1000 K
P_{top}	Log-Uniform	1e-1–1e-3 bar	1e-1–1e-3 bar
T_{int}	Uniform	–	10–1500 K
κ_0	Log-Uniform	–	1e-15–1e1
τ_0	Uniform	–	0.01–100
α	Log-Uniform	–	−10–10
P_{top}	Log-Uniform	–	1e2–5e7 bar
P_{bottom}	Log-Uniform	–	1e2–5e7 bar
C_{frac}	Uniform	–	0–1

Notes. [1]Gaussian prior.
[2]We use a Gaussian prior informed using evolutionary models from Fortney et al. (2008) combined with the age presented in Rajan et al. (2017).
[3]We make the assumption of a planetary mass object.
[4]GJ 570D distance comes from Gaia Archive: https://gea.esac.esa.int/archive/.
[5]The 51 Eri b distance comes from Macintosh et al. (2015).

emission:

$$\text{Absolute Flux} = F_{\text{emission}} \cdot S_{\text{cal}} \cdot \frac{R^2}{D^2}, \tag{7}$$

where F_{emission} is the emission flux from the forward model and R is the object radius and D its distance from the Earth. S_{cal} is a scaling calibration factor, to account for imperfect absolute flux calibration assumptions (see Fig. 2a). A calibration factor such as this was used in Oreshenko et al. (2020). Within the retrieval, S_{cal} can also be inversely considered as scaling the observed data to the model derived flux via Obs_{cal}

$$\text{Obs}_{\text{cal}} = \frac{1}{S_{\text{cal}}}, \tag{8}$$

We have added surface gravity $log\,(g)$ as an inferred parameter, determined via Newton's Law of Universal Gravity

$$\log(g) = \log\left[\frac{GM}{R^2}\right], \tag{9}$$

where G is the gravitational constant, M is the object's mass and R is the object's radius. We have also included the calculation of the carbon to oxygen (hereafter C/O) ratio, which for the brown dwarf and exoplanet, we study in this paper, is driven predominantly by the relative abundances of H_2O and CH_4. Therefore, this ratio should really be considered as a CH_4 to H_2O ratio. In the case of our T dwarf analysis the inferred C/O ratio is calculated via:

$$C/O = \frac{\chi_{CH_4} + \chi_{CO}}{\chi_{H_2O} + \chi_{CO}}, \tag{10}$$

where χ is the mixing ratio of the relative molecules.We also add an inferred metallicity via the retrieved abundances. This is approximated by summing metal-containing molecules weighted by the number of metal atoms which is then divided by the abundance of neutral hydrogen. This value is then compared to the summation of the retrieval-traced solar metals relative to hydrogen, using values from Asplund et al. (2009). Metallicity is therefore calculated via

$$M_{\text{object}} = \sum_{\text{molecules}} \frac{f_m \cdot n}{\chi_{H2} \cdot 2}, \tag{11}$$

where f_{m} is the gas fraction of a particular molecule m, n is the number of 'metal' atoms in a given molecule (eg. for CO2 n=3), χ_{H2} is the gas fraction of neutral hydrogen. Therefore, [M/H] is determined via

$$[M/H] = \frac{M_{\text{object}}}{M_{\text{solar}}}, \tag{12}$$

where M_{solar} is determined via all relevant solar elemental abundances relative to solar H.

3.2.8 *Clouds*

In order to explore the impact of cloud extinction, we employed the power-law deck and slab cloud paramterizations from Burningham et al. 2017, which are illustrated in Fig. 5. The optical depth of the power-law deck cloud is set by P_{top} pressure where the cloud becomes optically thick (τ=1). The cloud opacity drops off above this pressure via $d\tau/dP \propto \exp(\Delta P/\Phi)$, where ΔP is the height above and below the P_{top} pressure and ϕ is

$$\Phi = \frac{P_{\text{top}} \cdot (10^{\Delta log\,P} - 1)}{10^{\Delta log\,P}} \tag{13}$$

The decay is parametrized by $\Delta log\,P$ and the cloud is made non-grey with the optical depth following $\tau \propto \lambda^{\alpha}$. Therefore, in total, the

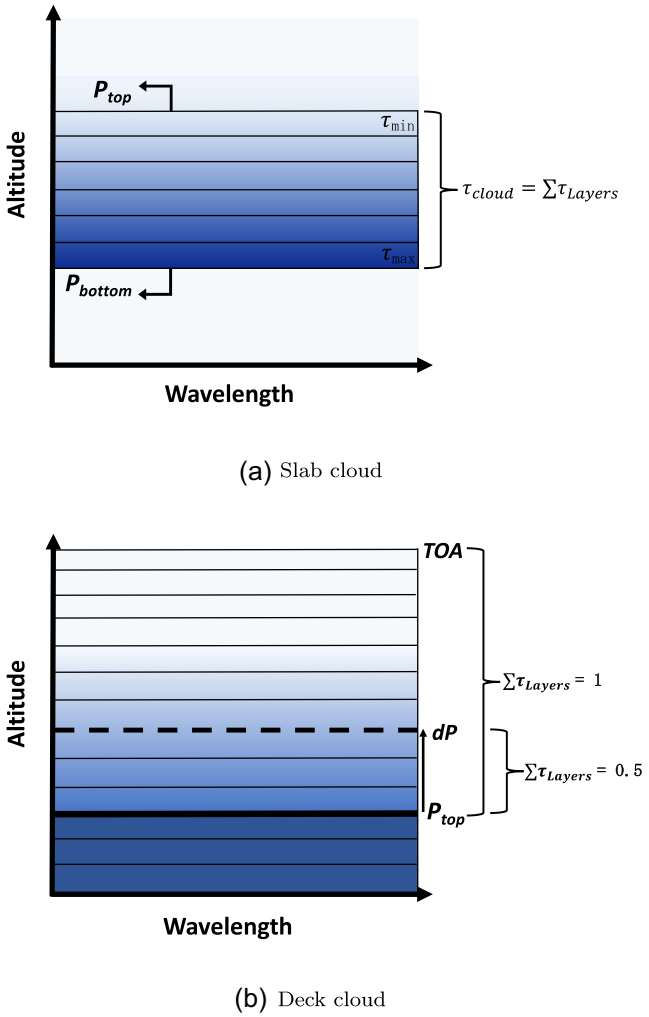


Figure 5. Diagram outlining cloud structures employ in this study. (a) outlines the slab cloud structure. (b) outlined the deck cloud structure.

power-law slab is retrieved via 3 parameters: P_{top}, $\Delta log\,P$ and α. The total optical depth of the slab cloud is determined via

$$\tau_{\text{cloud}} = \tau_0 \left(\frac{\lambda}{\lambda_0}\right)^{\alpha} = \sum \tau_{\text{Layers}}, \tag{14}$$

where λ_0=1 μm. τ_0 and α are the two retrievable components of this cloud prescription as well as the top pressure boundary of P_{top} and the bottom pressure boundary of the cloud P_{bottom}. τ_{cloud} is distributed throughout the layers in the cloud slab pressure boundaries, weighted by $d\tau/\,dP \propto P$ where dP is relative to P_{bottom}. Therefore, the total optical depth is distributed such that the bottom layer has the maximum optical depth while the top layer has the minimum optical depth present. In total, the power-law slab is retrieved via 4 parameters: τ_0, α, P_{top}, and P_{bottom}. These are flexible but simplistic cloud approaches that lack the rigour of the more physically motivated approaches included in Mollière et al. (2020) and Burningham et al. (2021). For example, these parametrizations do not allow us to probe specific cloud species or particle sizes but is still suitable for this study. In an effort to investigate the potential presence and impact of patchy clouds on exoplanet and brown dwarf spectra, the basic patchy cloud consideration from Marley et al. 2010 has been added

$$F_{\text{Tot}} = C_{\text{frac}} \cdot F_{\text{Cloudy}} + (1 - C_{\text{frac}}) \cdot F_{\text{Clear}}, \tag{15}$$

where $F_{\rm Tot}$ is the total flux, $F_{\rm Clear}$ is the flux from regions without clouds, $F_{\rm Cloudy}$ is the flux from regions with clouds and $C_{\rm frac}$ is the fraction of surface area with clouds. The non-cloud properties are identical for the two forward models which are linearly combined. Employing this fractional cloud consideration therefore acts to add an additional retrieved parameter to both the slab and deck cloud parametrizations.

3.2.9 *Forward model set-up*

The atmospheric forward model is described in Al-Refaie et al. (2021). We assume a hydrogen dominated atmosphere with an H_2 and He mixing ratio He/H_2 = 0.17567. We set molecular trace-gas volume mixing ratios to be constant in pressure, which we refer to as isoprofiles. TAUREX3 does allow for pressure dependent abundance profiles (Changeat et al. 2019) but this comes at a significant increase in model complexity and so will be explored in later work.

We use error inflation (as outlined in Section 3.2.5) for our analysis of GJ 570D but not for our 51 Eri b analysis. This is because 51 Eri b's observations have much larger error bars and the observation's resolution is more uniform throughout the spectrum, negating the spectral band weighting issue experienced with the GJ 570D data set (again, see Section 3.2.5). In our GJ 570D data, it is apparent that that *K*-band errors are much smaller than the *J*-band data (see Fig. 2a). The impact of adding this error inflation parameter acted to allow the fit to the *J*-band data to improve without affecting the goodness of fit in the *K* band and also allowed for the overall Log Evidence to increase slightly. This was interpreted as an increase in error size in the *K*-band, while negligible in the *J* band, allowing for a better overall fit by deweighting the small error bars found predominately in the *K* band when performing the Bayesian likelihood calculation.

In this study, a plane-parallel approximation is used to model the atmosphere, with the pressure ranging from 10^{-3} to 500 bar, uniformly sampled in log-space with 100 atmospheric layers. Collision induced absorption (CIA) of H_2–H_2 and H_2–He (Abel et al. 2011; Fletcher, Gustafsson & Orton 2018; Abel et al. 2012) is included.

In the case of the 51 Eri b data analysis, we use multiple scaling factors $S_{\rm cal}$ to account for the inclusion of observations from different instruments. This is employed in the case of the SPHERE *Y*, *J*, and *H* data ($S_{{\rm cal}\,SPH}$) being combined with the GPI *K*1 and *K*2 band data ($S_{{\rm cal}\,GPI}$). When employing data from a single instrument, we simply use one scaling $S_{\rm cal}$ factor.

3.3 ATMO 2020

We compare our cloudless retrievals to self-consistent radiative-convective grid models. For this, we use the recently published ATMO 2020 set of atmosphere and evolutionary models for cool brown dwarfs and self-luminous giant exoplanets (Phillips et al. 2020).

The ATMO code is a 1D radiative-convective equilibrium model, and has been most recently described in Phillips et al. (2020) and Goyal et al. (2020). Briefly, ATMO defines the TP-profile of an atmosphere on a logarithmic optical depth grid with 100 model levels. The outer boundary condition in the first model level is fixed at a pressure of 10^{-5} bar and is given an optical depth of $\tau \sim 10^{-4}$–10^{-7} depending on surface gravity. The inner boundary condition in the last model level is not fixed in pressure and is given an optical depth of $\tau = 1000$. The model then iterates the pressure and temperature in each model level towards radiative-convective and hydrostatic equilibrium using a Newton–Raphson solver. On each iteration chemical equilibrium abundances are calculated for the current TP-profile using a Gibbs energy minimization scheme based on that of Gordon & McBride (1994). ATMO also has the ability to calculate non-equilibrium chemical abundances self-consistently with the TP-profile, using kinetic networks or relaxation schemes (Drummond et al. 2016; Phillips et al. 2020). Once the chemical abundances have been computed, the opacities used by ATMO can be obtained from pre-computed correlated-*k* tables for individual gases (Amundsen et al. 2014), and are combined within the code using the random overlap to obtain the total mixture opacity consistently with the pressure, temperature, and abundances in each iteration (Amundsen et al. 2017). The radiative flux is computed by solving the integral form of the radiative transfer equation in 1D plane-parallel geometry including isotropic scattering following Bueno & Bendicho (1995). The convective flux is computed using mixing length theory using the same method as Gustafsson et al. (2008), with the adiabatic gradient computed using equation of state tables from Saumon, Chabrier & van Horn (1995).

This grid includes solar metallicity atmosphere models spanning $T_{\rm eff} = 200 - 3000$ K and $\log(g) = 2.5$–5.5 (g in units of cm s^{-2}), with steps of 100 K for $T_{\rm eff} > 600$ K, 50 K for $T_{\rm eff} < 600$ K, and 0.5 in $\log(g)$. The ATMO 2020 model set consists of three atmosphere model grids spanning this parameter range. The first is calculated assuming chemical equilibrium, and the second and third are calculated assuming non-equilibrium chemistry with different strengths of vertical mixing. Each model in the grid is generated with the ATMO code and consists of a TP-profile, chemical abundance profiles, and a spectrum of the emergent flux from the top of the atmosphere, which are publicly available for download.[6]

3.3.1 *Sampling using Markov Chain Monte Carlo*

To calculate the best fits from the ATMO 2020 grid to the spectrophotometry of 51 Eri b (see section 2.2), we used a Markov chain Monte Carlo (MCMC) method utilising the EMCEE PYTHON package (Foreman-Mackey et al. 2013). We generated each independent model using an interpolation to the ATMO 2020 grid with temperatures ranging from 200 K to 3000 K and $\log(g)$ from 2.5 cm s^{-2} to 5.5 cm s^{-2} for models assuming chemical equilibrium, and temperature ranging from 350 K to 1800 K and $\log(g)$ from 3.0 cm s^{-2} to 5.5 cm s^{-2} for models assuming non-equilibrium chemistry due to vertical mixing. The radius was constrained between 0.07 $R_\odot$ ($\sim 0.7\,R_{\rm Jup}$) and 0.2 $R_\odot$ ($\sim 2\,R_{\rm Jup}$) for both cases, using a rough estimation from the ATMO evolutionary tracks, given the system's age. With this grid, the MCMC was set up with 100 walkers and was executed for 500 steps. The posteriors were constructed after discarding the first 200 steps, to account for the 'burn-in'. This eliminates any bias caused by the initial values supplied to the MCMC as a starting point in the parameter space. All results are reported with an uncertainty of 1σ.

4 RESULTS: GJ 570D

In order to evaluate TAUREX3's emission model against brown dwarf observations, we perform retrieval analysis on the Spex observations of GJ 570D. We compare our results with previous studies which employed other retrieval codes, with the aim of determining if the results were consistent with these previous studies. The results of the comparison are shown in Table 3, with the retrieval priors used in the analysis listed in Table 2.

[6] ATMO 2020: http://opendata.erc-atmo.eu

Table 3. Summary of retrieval bulk parameters for GJ 570D along with values from previous studies.

	Mass (M_{Jup})	Radius (R_{Jup})	log(g) (cm s^{-2})	T_{eff} (K)	C/O	[M/H]
This work (TAUREX3)	$48.00^{+13.03}_{-11.87}$	$1.17^{+0.08}_{-0.08}$	$4.93^{+0.11}_{-0.12}$	722^{+23}_{-26}	$0.87^{+0.08}_{-0.07}$	$-0.19^{+0.05}_{-0.03}$
This work (ATMO 2020 – EC FM)	–	$0.71^{+0.04}_{-0.02}$	$4.64^{+0.34}_{-0.30}$	$826.34^{+12.88}_{-17.21}$	–	–
This work (ATMO 2020 – NEC FM)	–	$0.72^{+0.06}_{-0.03}$	$4.63^{+0.16}_{-0.10}$	$813.33^{+14.01}_{-27.19}$	–	–
Kitzmann et al. 2020 (FCR)	53^{+24}_{-20}	$1.13^{+0.05}_{-0.06}$	$5.01^{+0.13}_{-0.19}$	703^{+17}_{-30}	$1.11^{+0.09}_{-0.09}$	$-0.13^{+0.06}_{-0.08}$
Kitzmann et al. 2020 (ECR)	$17^{+3.8}_{-3.0}$	$1.00^{+0.10}_{-0.09}$	$4.61^{+0.08}_{-0.08}$	730^{+18}_{-17}	$0.83^{+0.09}_{-0.08}$	$-0.15^{+0.05}_{-0.04}$
Burningham et al. 2017 (FCR)	$19.80^{+28.60}_{-15.96}$	$0.96^{+0.80}_{-0.11}$	$4.73^{+0.31}_{-1.17}$	$752.25^{+35.51}_{-82.10}$	–	–
Line et al. 2015 (FCR)	$30.90^{+26.64}_{-15.76}$	$1.14^{+0.10}_{-0.09}$	$4.76^{+0.27}_{-0.28}$	$714.11^{+20.19}_{-23.15}$	$1.09^{+0.16}_{-0.14}$	$-0.25^{+0.13}_{-0.12}$
Oreshenko et al. 2020: Sonora (SML)	–	–	$4.93^{+0.38}_{-0.55}$	808^{+43}_{-27}	–	–
Oreshenko et al. 2020: AMES-cond (SML)	–	–	$5.27^{+0.43}_{-0.67}$	878^{+23}_{-78}	–	–
Oreshenko et al. 2020: HELIOS (SML)	–	–	$5.08^{+0.62}_{-0.68}$	800^{+14}_{-100}	–	–
Samland et al. 2017 (FM)	–	$0.94^{+0.04}_{-0.04}$	$4.67^{+0.04}_{-0.04}$	769^{+14}_{-13}	–	–
Filippazzo et al. 2015 (EM)	$37.28^{+24.05}_{-24.05}$	$0.94^{+0.16}_{-0.16}$	$4.90^{+0.50}_{-0.50}$	759^{+63}_{-63}	–	–
Testi 2009 (FM)	–	–	5.0	900	–	–
Del Burgo et al. 2009 (FM)	–	–	$4.5^{+0.5}_{-0.5}$	948^{+58}_{-58}	–	–
Saumon et al. 2006 (EM, FM)	$42.5^{+4.5}_{-4.5}$	$0.855^{+0.023}_{-0.023}$	5.09–5.23	800–820	–	–
Burgasser et al. 2006 (EM)	–	–	5.1	780–820	–	–

Notes. [1]EC FM is Equilibrium Chemistry Forward Model
[2]NEC FM is Non-Equilibrium Chemistry Forward Model
[3]FCR is Free Chemistry Retrieval
[4]ECR is Equilibrium Chemistry Model
[5]SML is Supervised Machine Learning
[6]EM = Evolutionary Model
[7]FM = Forward Model

4.1 Na + K systematic model bias

We encountered a systematic bias in the retrieved estimates for mass and radius when attempting to fit the 0.85–2.5 μm spectrum of GJ 570D using a flat prior. This bias resulted in a non-physical radius (see Fortney et al. 2008 and Chabrier et al. 2009 for typical radii) of 1.4 to 1.55 R_{Jup}, along with a mass value that converged to the prior's upper boundary. The mass was also found to increase with the radius, likely in an effort to maintain the best-fitting surface gravity. Both retrievals and grid-modelling approaches have encountered the problem of mass values converging to the upper boundary of the model prior space (Line et al. 2015; Zalesky et al. 2019; Schneider et al. 2015).

To investigate this effect, we ran retrievals with varying absolute flux calibrations and also employed the Madhusudhan & Seager (2009) temperature–pressure profile. Neither of these approaches negated the systematic bias. The application of a tight Gaussian prior on the radius was also tested, but in this case, the mass was still seen to converge to the upper boundary of its flat prior. We find this systematic issue to be sensitive to the sodium and potassium (Na + K) cross sections, a dominating source of contribution in near-infrared model fitting as shown in Line et al. (2015), Burningham et al. (2017), and Oreshenko et al. (2020).

It is noteworthy that this issue seems most prevalent when fitting the whole 0.85–2.5 μm spectrum. The resonance doublets of K and Na are at ∼0.77 μm and ∼0.59 μm, respectively. We encountered examples when the bias issue would not be present when fitting only 0.85–1.2 μm (∼0.77 μm K / ∼0.59 μm Na resonance doublet impacted region), or 1.2–2.5 μm (non-resonance lines/resonance doublet line wings region). This indicates a potential issue with either the combination of the resonance doublets and non-resonance lines within the Na + K cross sections, or with the extent of the broadening of the resonance doublets. Some different combinations of computing the cross sections of the resonance doublet and non-resonance lines of K are illustrated in Fig. 6. The resonance doublets tested in the present study were either treated using the broadening parameters of Burrows et al. (Burrows & Volobuyev 2003) or Allard et al. (Allard et al. 2016; Allard et al. 2019). The non-resonance lines from both the NIST (Kramida et al. 2013) or Kurucz (Kurucz & Bell 1995) databases were also tested. Testing various combinations didn't negate the aforementioned bias. We again note that the results presented in this study (Tables and Figures) were retrieved using the broadening parameters of Allard et al. and non-resonance lines from the Kurucz data base. The issues related to the Na and K cross sections are discussed further in Section 6.

We therefore present two separate retrieval analyses for GJ 570D. First, to avoid the impact of this systematic bias but to still attain a set of values for the scaling factors (radius, distance and S_{cal}) along with the mass (and by extension the inferred surface gravity) we ran a retrieval fitting only the 1.2–2.5 μm part of the spectrum. This cutoff of the potassium resonance doublet impacted region of the spectrum allowed for physically credible results for the mass and radius using flat priors. We then used these values as fixed (non-fitted) priors in a subsequent retrieval to infer the chemical properties of the atmosphere. This was necessary as extending the fit of the 1.2–2.5 μm retrieval to the 0.85–1.2 μm data showed a significant mismatch between the model fit and the observed SED in this region, as shown in Fig. 7. This two-step approach leads to the most literature consistent values for the retrieved parameters but does lead to a very slightly lower Bayesian Evidence value (see Section 3.2.6) due to a slightly worse fit of the *J*-band peak.

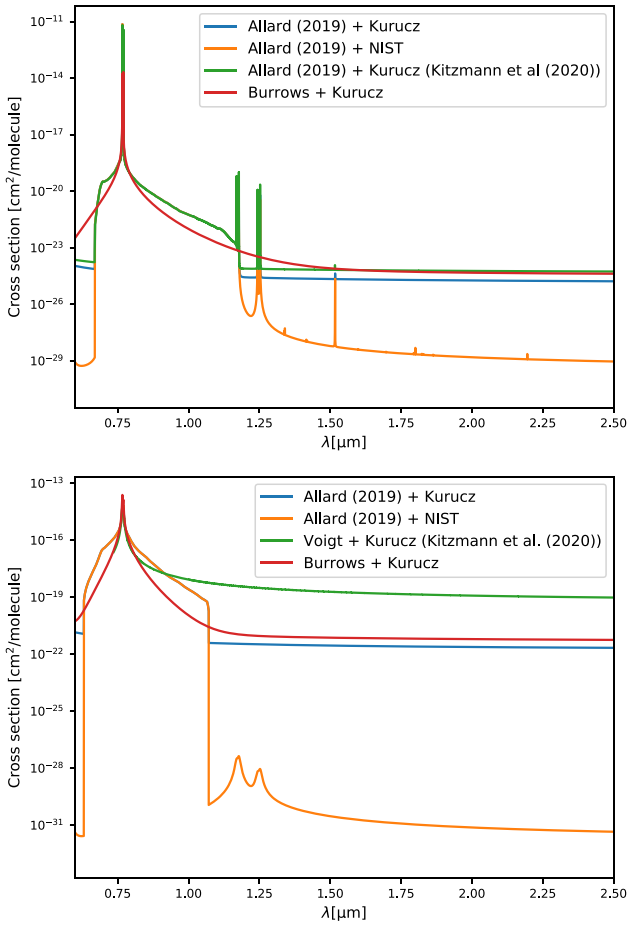


Figure 6. A comparison of different methods used to compute the resonance doublet and non-resonance lines of Na and K. The first panel gives cross sections for K computed at $T = 1000$ K, $P = 0.1$ bar, and the second panel the cross sections for K computed at $T = 600$ K, $P = 10$ bar. The cross sections in green are those which were used in Kitzmann et al. (2020). All other combinations shown (using either Burrows et al. (Burrows & Volobuyev 2003) or Allard et al. (Allard et al. 2016; Allard et al. 2019) for the computation of the resonance doublets, and either NIST (Kramida et al. 2013) or Kurucz (Kurucz & Bell 1995) for the non-resonance lines) were tested in the present study.

While this strategy did derive results consistent with previous studies, it does have its limitations and imperfections. First, the assumptions of flat priors while also truncating the data is not an ideal approach. The temperature–pressure profile, which is fit in the second retrieval, is significantly constrained as the scaling factors, with which it is intricately linked, are fixed. The same can be said for the alkali abundance, which is strongly correlated to surface gravity. While our approach derives an alkali abundance consistent with previous studies, likely as a result of being able to make use of the alkali dominated wavelength region, we acknowledge this has been driven to an extent by our constraint on this parameter.

4.2 Scaling factors and bulk parameters

The model posteriors for the mass, radius, $S_{\rm cal}$ and distance, along with the inferred surface gravity, can be seen in Fig. 8, along with the spectral fit to the data used. These results are also summarized in Table 3. In general, the retrieved parameter values are consistent with previous studies. Values from previous studies can also be seen

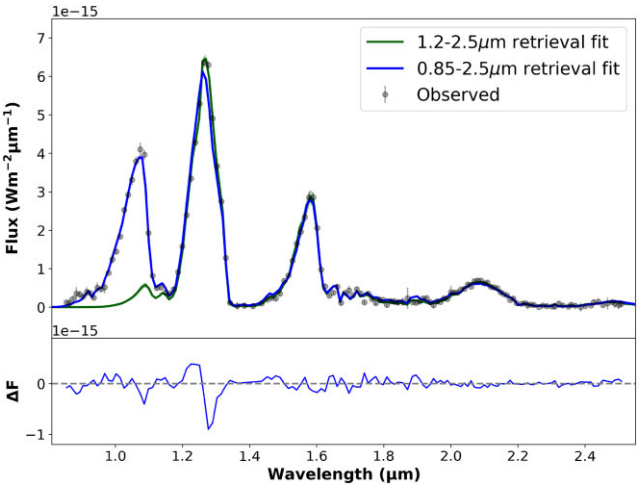


Figure 7. GJ 570D retrieval spectral fit.

in Table 3. Mass is consistent with all previous studies outlined in the table, apart from the equilibrium chemistry retrieval presented in Kitzmann et al. (2020). Radius is consistent with all previous free chemistry retrievals quoted in the table, and is 2σ consistent with the slightly lower radii presented in the equilibrium chemistry retrieval from Kitzmann et al. (2020) and non-retrieval analysis conducted in previous studies. As mass and radius are largely consistent with previous studies, so too is the inferred surface gravity. As the distance prior is so tightly constrained because of the precise Gaia measurements (Gaia Collaboration 2016, 2018), the distance parameter does not play a significant role in the scaling of the spectrum. Our retrieved effective temperature matches well with all previously conducted retrieval studies, whilst some other studies such as Saumon et al. (2006) and Burgasser et al. (2006) have obtained slightly higher values for this parameter.

4.3 Abundances

The posterior distributions for the retrieved abundances are shown in Fig. 9, and listed in Table 4. The resulting SED fit, derived combining these retrieved abundances along with the locked scaling parameters outlined previously, is shown in Fig. 7. These show that the three most abundant molecules are H_2O, CH_4, and NH_3, whilst Na + K is also well constrained.

The abundance for Na + K that we retrieve is similar to that from Kitzmann et al. (2020) but noticeably different from the values presented in Line et al. (2015) and Burningham et al. (2017), which we ascribe to the use of the broadening coefficients from Allard et al. (2016, 2019) in our analysis and that from Kitzmann et al. (2020).

Overall, these abundances (and by extension the C/O and [M/H] ratio) are similar to those from previous retrieval studies of this object presented in Line et al. (2015), Burningham et al. (2017), and Kitzmann et al. (2020). Our super-solar $0.87^{+0.08}_{-0.07}$ C/O ratio for GJ 570D is in good agreement with the reported 0.65–0.97 C/O for its host star presented in Line et al. (2015). Our value is slightly lower than that derived in Line et al. (2015)'s and Kitzmann et al. (2020)'s free chemistry retrievals, but is consistent with Kitzmann et al. (2020)'s equilibrium chemistry model. We do note however that this comparison is imperfect, as our inferred C/O value only considers the pure gas phase and this neglects the elemental losses dues to condensation.

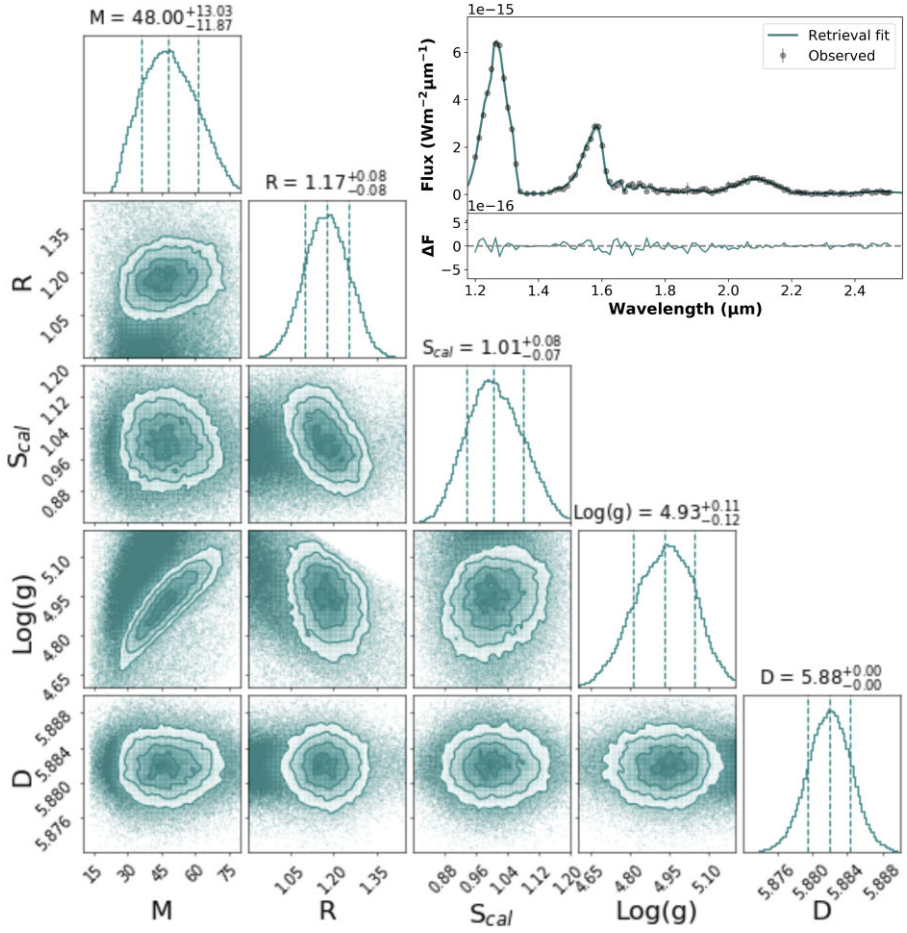


Figure 8. GJ 570D bulk parameter posterior probability distributions for the spectral fit of the 1.2–2.5 μm data used.

4.4 Temperature–Pressure profile

Our retrieved temperature–pressure profile is very similar to that obtained in the Kitzmann et al. (2020) study (see Fig. 9 for comparison, where the blue band marks the one sigma error on our derived profile). The agreement in the 1–10 bar pressure region is particularly close, as expected in this region which contributes most to the spectral emission profile. We are further encouraged that this good agreement continues up into the stratospheric region where the constraining influence of the spectral emission is smaller.

5 RESULTS: 51 ERI B

In this section, we outline our retrieval results for 51 Eri b, compared to previous studies, all of which required clouds to produce the observed SED. Here, we find that inverse retrieval methods can recreate the observed SED with cloud-free atmospheres but using a more flexible (*npoint*) temperature–pressure profile. These results include evidence of a tentative ammonia detection. We also compare our results for 51 Eri b to those for GJ 570 D (which is a close match in spectral type). Finally, we outline retrievals which included a power-law deck or slab cloud combined with a less flexible temperature–pressure profile (Lavie et al. 2017).

We ran retrievals on the SPHERE *Y*-, *J*-, and *H*-band data and separately on the GPI *J*- and *H*-band data. For both retrievals, we adopted the GPI *K*1 and *K*2 data. We present the spectral fits of these combinations in Fig. 10 and the comparison posteriors in Fig. 11, with the results from the individual data sets shown in Figs A3 and A4. As Samland et al. (2017) did not fit the GPI *K*1 and *K*2 data in their study, we also present retrieval results using only the SPHERE *Y*-, *J*-, and *H*-band data. The posteriors for these results are presented in Fig. A5. We do note here, however, that Samland et al. (2017) included SPHERE and GPI photometry in their fitting which was a driving component of the high metallicity they derive.

The retrieval priors used in this analysis are presented in Table 2, with an overall summary of the retrieval results in Tables 5 and 6. The following subsections focus on the retrievals which used the *npoint* temperature–pressure profile and omitted clouds as these derived the highest (or comparably indistinguishable) Log (Ev). We then discuss the cloudy retrievals in a subsequent subsection.

5.1 Scaling factors and bulk parameters

Our highest Log (Ev) posterior probability distributions for the Mass, Radius, S_{cal}, and Distance, along with the inferred surface gravity are presented in Fig. 11. We find that our retrieval analysis is able to produce excellent fits to 51 Eri b's observed SED (see Fig. 9) while deriving physically credible mass and radius values. This is the case when analysing each data set as outlined previously.

The S_{cal} factors derived indicate a preference for a brighter *K* band absolute flux calibration in both retrievals where this data is employed. In the cases of the SPHERE data being employed within the retrieval, an $S_{cal} \sim 1$ is derived, indicating a model preference for

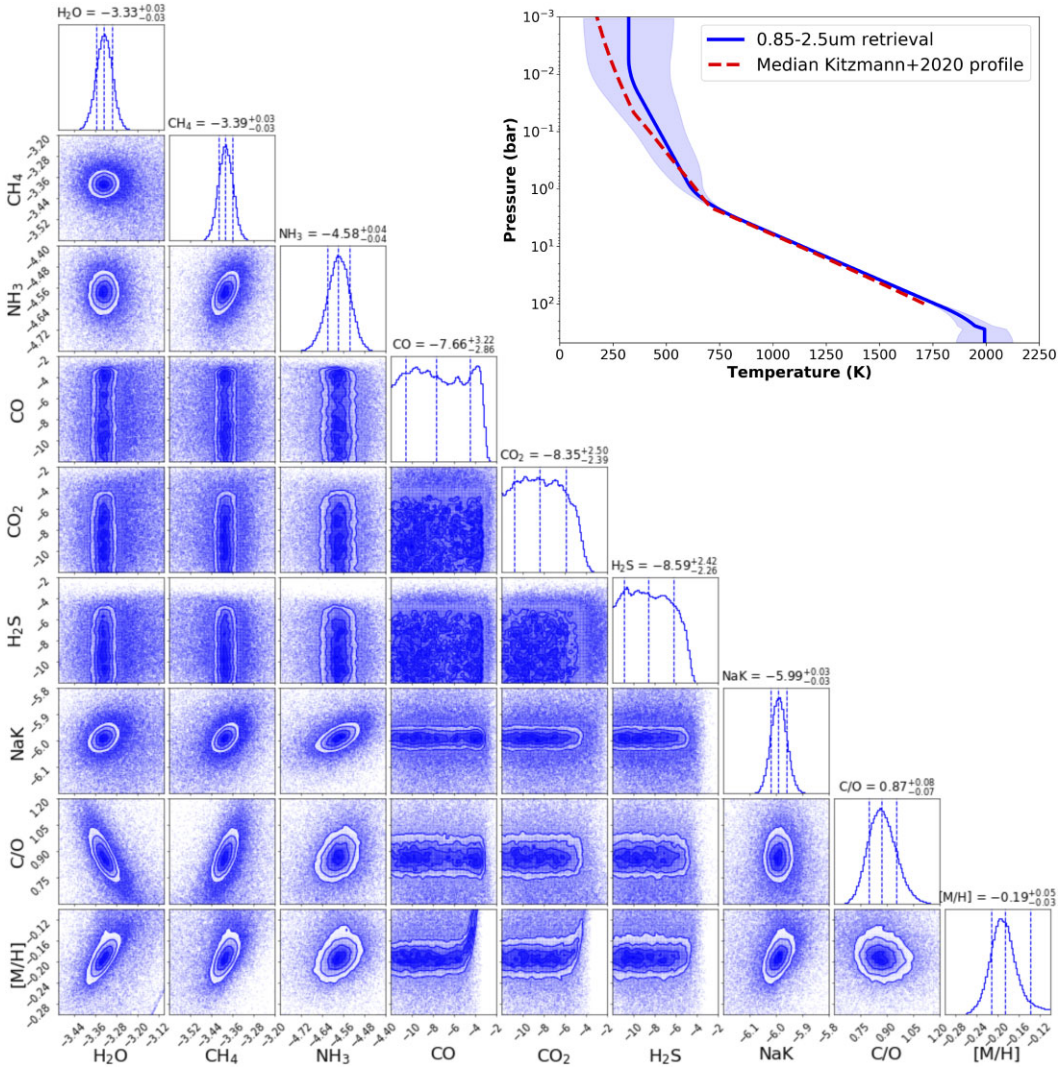


Figure 9. GJ 570D mixing ratio posteriors. C/O and [M/H] posteriors are inferred parameters, while all the other parameters are sampled as part of the retrieval. The retrieved temperature–pressure profile is also shown along with a comparison to the median profile retrieved in the Kitzmann et al. (2020) study.

Table 4. Summary of GJ 570D retrieved molecular abundances along with a comparison to previous studies. TW = This work.

	TW,0.85–2.5 μm	TW, 1.2–2.5 μm	Kitzmann et al. (2020)	Burningham et al. (2017)	Line et al. (2015)
log (H_20)	$-3.33^{+0.03}_{-0.03}$	$-3.11^{+0.05}_{-0.05}$	$-3.33^{+0.05}_{-0.06}$	$-3.42^{+0.16}_{-0.22}$	$-3.40^{+0.13}_{-0.13}$
log (CH_4)	$-3.39^{+0.03}_{-0.03}$	$-3.34^{+0.05}_{-0.06}$	$-3.28^{+0.06}_{-0.09}$	$-3.44^{+0.20}_{-0.31}$	$-3.45^{+0.10}_{-0.10}$
log (NH_3)	$-4.58^{+0.04}_{-0.04}$	$-4.69^{+0.07}_{-0.09}$	$-4.38^{+0.07}_{-0.10}$	$-4.82^{+0.26}_{-2.47}$	$-4.64^{+0.15}_{-0.15}$
log (CO)	$-7.66^{+3.22}_{-2.86}$	$-7.06^{+4.12}_{-3.12}$	$-7.70^{+2.7}_{-2.4}$	$-7.47^{+3.05}_{-3.04}$	$-7.53^{+2.65}_{-3.07}$
log (CO_2)	$-8.35^{+2.50}_{-2.39}$	$-8.58^{+2.12}_{-2.09}$	$-7.70^{+2.7}_{-2.4}$	$-7.86^{+2.67}_{-2.66}$	$-7.76^{+2.23}_{-2.89}$
log (H_2S)	$-8.59^{+2.42}_{-2.26}$	$-3.86^{+0.12}_{-2.26}$	-8.47^{+2}_{-2}	$-8.74^{+2.68}_{-2.20}$	$-8.94^{+2.22}_{-2.11}$
log (Na + K)	$-5.99^{+0.03}_{-0.03}$	$-4.37^{+0.06}_{-0.06}$	$-5.86^{+0.04}_{-0.03}$	$-5.47^{+0.09}_{-0.30}$	$-5.45^{+0.06}_{-0.06}$

this absolute flux calibration given the priors set. All the derived S_{cal} values can be found in Fig. A3, A4, and A5.

We note that the cloudless models used in the previous studies did not fit the SED particularly well. The cloudless models (Saumon & Marley 2008) in Macintosh et al. (2015) derived a barely sub-stellar mass of 67 M_{Jup} with a low radius of 0.76 R_{Jup} while Samland et al. (2017)'s cloudless model (Mollière et al. 2015, 2017) derived a mass that was 1σ consistent with that of a planetary mass object, but had an improbably small radius of 0.40 R_{Jup} for a Jovian exoplanet, violating electron degeneracy pressure laws for an object such as this (Chabrier et al. 2009). We attempted to fit the SED of 51 Eri b using a cloudless ATMO grid model as shown in Figs 10 and A1, illustrating that these grid models are unable to explain the SED of this object or to constrain its

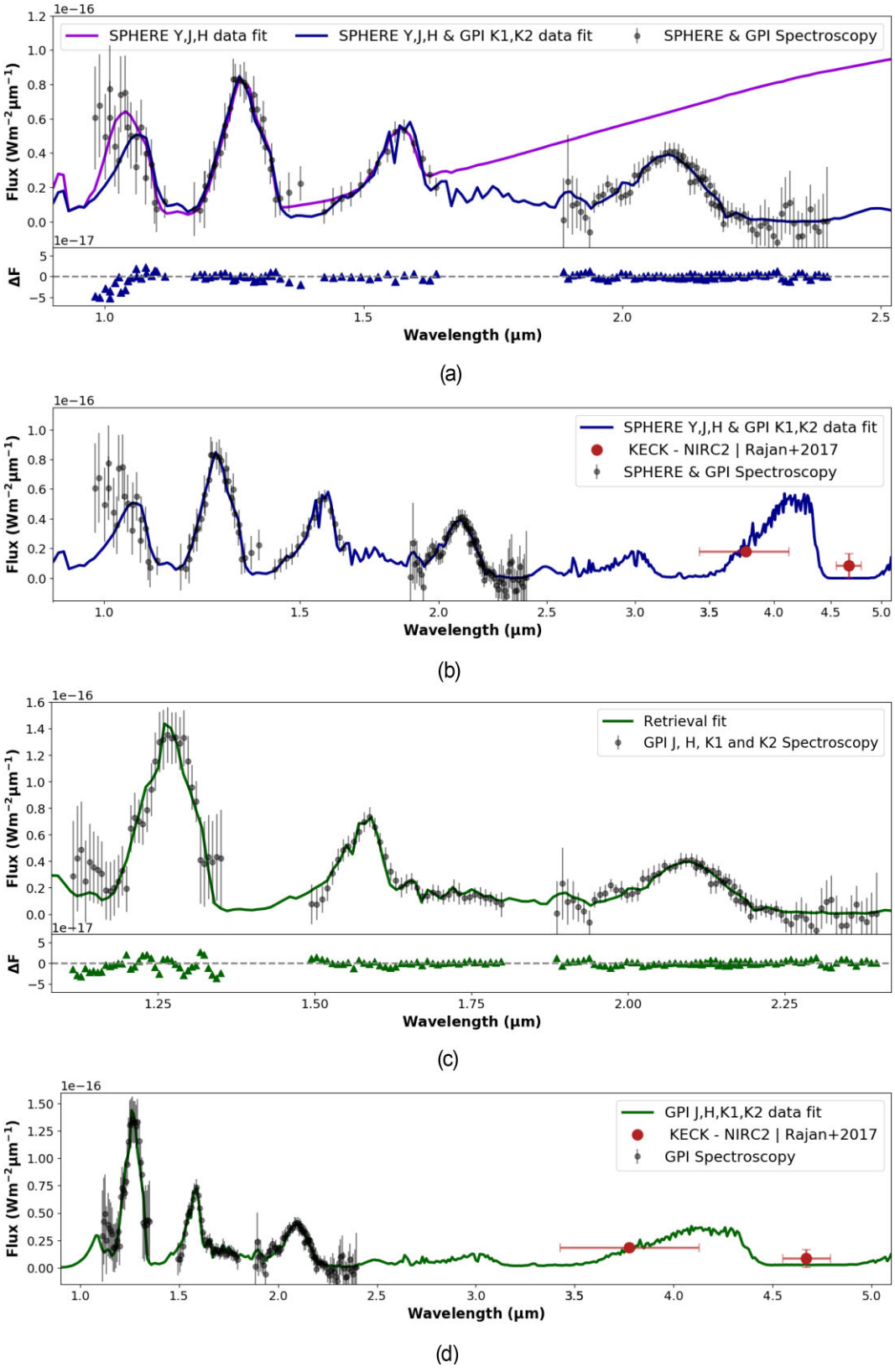


Figure 10. 51 Eri b SED fits via cloudless retrievals. (a) illustrates the model fit for retrievals including the SPHERE data where the dark violet fit shows the retrieval fit to only the SPHERE *Y*-, *J*-, and *K*-band data and the dark blue fit shows the retrieval fit when the SPHERE data is combined with the GPI *K*1 and *K*2 data. (b) shows the SPHERE *Y*, *J*, *H* and GPI *K*1, and *K*2 data fit extrapolated to longer wavelengths, with the inclusion of KECK-NIRC2 photometry. (c) illustrates the model fit for the retrieval using the GPI *J*, *H*, *K*1, and *K*2 data. (b) shows the GPI *J*, *H*, *K*1, *K*2 data fit extrapolated to longer wavelengths, with the inclusion of KECK-NIRC2 photometry.

Table 5. Summary of retrived bulk parameters for 51 Eri b. Values from previous studies are also included for comparison. L17: Lavie et al. (2017). GPI all: GPI *J*, *H*, *K*1, *K*2 data. SPHERE & GPI: SPHERE *Y*, *J*, *H*& GPI *K*1, *K*2 data. SPHERE only: SPHERE *Y*, *J*, *H* data. U: Uniform clouds coverage. P: Patchy cloud coverage.

	TP Profile Type	Log (Ev)	Mass (M_{Jup})	Radius (R_{Jup})	log (g)	T_{eff} (K)	C/O	[M/H]
TW, Cloudless (GPI *J*, *H*, *K*1, *K*2 data)	*npoint*, flexible	5596.02	$8.50^{+2.86}_{-3.44}$	$1.09^{+0.11}_{-0.11}$	$4.26^{+0.16}_{-0.23}$	769^{+37}_{-41}	$0.92^{+0.19}_{-0.27}$	$-0.26^{+0.66}_{-0.18}$
TW, Cloudless (SPHERE *Y*, *J*, *H*& GPI *K*1, *K*2 data)	*npoint*, flexible	5141.33	$7.93^{+3.22}_{-3.54}$	$1.18^{+0.12}_{-0.12}$	$4.16^{+0.18}_{-0.26}$	700^{+42}_{-45}	$0.97^{+0.09}_{-0.20}$	$-0.04^{+0.95}_{-0.49}$
TW, Cloudless (SPHERE *Y*, *J*, *H* data)	*npoint*, flexible	2330.45	$8.25^{+3.01}_{-3.33}$	$1.31^{+0.12}_{-0.11}$	$4.09^{+0.15}_{-0.23}$	909^{+37}_{-50}	$0.40^{+0.26}_{-0.15}$	$-0.66^{+0.14}_{-0.11}$
Nielsen et al. 2019			$2.6^{+0.3}_{-0.3}$	–	–	–	–	–
Samland et al. 2017 (PTC-uniform clouds)[1]			$9.1^{+4.9}_{-3.3}$	$1.11^{+0.16}_{-0.14}$	$4.26^{+0.25}_{-0.25}$	760^{+20}_{-20}	–	–
Samland et al. 2017 (PTC-patchy clouds)[1]			$14.5^{+8.7}_{-5.6}$	$1.11^{+0.16}_{-0.14}$	$4.47^{+0.24}_{-0.26}$	757^{+24}_{-24}	–	–
Samland et al. 2017 (PTC-clear)[1]			$14.5^{+4.7}_{-3.1}$	$0.40^{+0.02}_{-0.02}$	$5.35^{+0.15}_{-0.12}$	982^{+18}_{-15}	–	–
Samland et al. 2017 (Morley et al. 2012 clouds)			$64.9^{+19.1}_{-15.6}$	$1.01^{+0.07}_{-0.06}$	$5.19^{+0.10}_{-0.11}$	684^{+16}_{-20}	–	$1.03^{+0.102}_{-0.11}$
Rajan et al. 2017 (iron-silicate, patchy clouds)			–	$0.68^{+0.13}_{-0.14}$	3.25	737^{+39}_{-46}	–	–
Rajan et al. 2017 (Sulfide, salt, uniform clouds)			–	$0.90^{+0.23}_{-0.26}$	$4.05^{+0.36}_{-0.35}$	605^{+61}_{-66}	–	–
Macintosh et al. 2015 (cloud-free)			67	0.76	5.5	750	–	–
Macintosh et al. 2015 (partial-cloud)			2	1	3.5	700	–	–

Table 6. Summary of 51 Eri b retrieved molecular abundances, C/O ratio and metallicity [M/H] with a comparison to our retrieved values for GJ 570D. These are the values from the highest Log(Ev) (cloudlesss, *npoint* TP) retrievals.

	51 Eri b (1)[1]	51 Eri b (2)[2]	GJ 570D
log (H_20)	$-3.52^{+0.16}_{-0.16}$	$-3.50^{+0.16}_{-0.19}$	$-3.33^{+0.03}_{-0.03}$
log (CH_4)	$-3.63^{+0.12}_{-0.13}$	$-3.60^{+0.09}_{-0.11}$	$-3.39^{+0.03}_{-0.03}$
log (NH_3)	$-4.85^{+0.15}_{-0.18}$	$-4.61^{+0.11}_{-0.14}$	$-4.58^{+0.04}_{-0.04}$
log (CO)	$-3.32^{+1.13}_{-5.68}$	$-5.10^{+2.27}_{-4.54}$	$-7.66^{+3.22}_{-2.86}$
log(Na + K)	$-9.52^{+1.69}_{-1.59}$	$-7.65^{+2.58}_{-2.83}$	$-5.99^{+0.03}_{-0.03}$
C/O	$0.97^{+0.09}_{-0.20}$	$0.92^{+0.19}_{-0.27}$	$0.87^{+0.08}_{-0.07}$
[M/H]	$-0.04^{+0.95}_{-0.49}$	$-0.26^{+0.66}_{-0.18}$	$-0.19^{+0.05}_{-0.03}$

Note. [1]51 Eri b (1) refers to results retrieved using SPHERE *Y*-, *J*-, *H*- and GPI *K*1-, *K*2-band data. [2]51 Eri b (2) refers to results retrieved using GPI *J*-, *H*-, *K*1-, and *K*2-band data

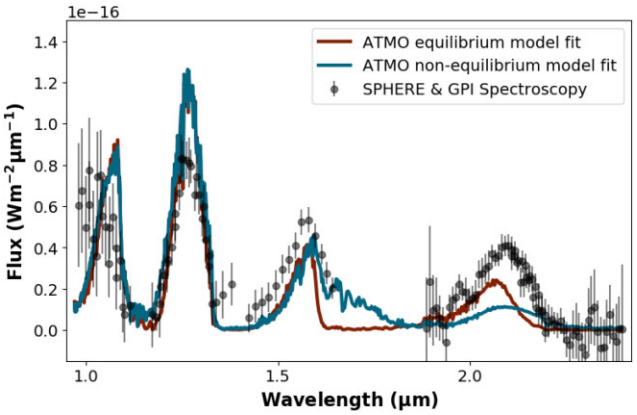


Figure 11. 51 Eri b spectral fit using ATMO.

surface gravity, radius, or effective temperature, using both chemical equilibrium and chemical disequilibrium assumptions as shown in Figs A1 and A2.

Our retrieved effective temperature values are consistent with expectations for a T dwarf except in the case of the retrieval using only the SPHERE data as longer wavelength data are neglected in this instance. This resulting SED fit is, however, inaccurate when extrapolated to the *K* band as shown in Fig. 10a.

5.2 Abundances, tentative ammonia detection

The higest Log(Ev) posterior distributions for the retrieved abundances are shown in Fig. 11 while retrieved abundances are shown in Table 6. Comparison with GJ 570D shows that the abundances of 51 Eri b and GJ 570D match to within 1σ, not unexpected given their similar spectral types (51 Eri b: T6.5 $\pm$ 1.5, GJ 570D: T7.5). We see that the derived [M/H] values for 51 Eri b, while consistent with GJ 570 D, have large uncertainties. This can also be seen in the large posterior tails for [M/H] shown in Fig. 11. This appears to be a result of the large uncertainties seen in the (unconstrained) retrieval CO abundances.

As presented in Table 6, the retrieved Na + K abundance for 51 Eri b is the only abundance which is not 1σ consistent with that retrieved for GJ 570D. This could either be a physical effect due to 51 Eri b's much lower surface gravity, or it could be related to the Na and K cross sections used in our retrievals. It could also be an impact

of absent data below ∼1 μm where this species plays a key role in contribution. We combine the Na and K cross-sections together at solar abundance ratios, which could be an incorrect assumption for one or both of these objects. However, we found a minimal change in the retrieval results when separate Na and K cross-sections (not combined at solar ratios) were used.

We report a tentative detection of ammonia in the atmosphere of 51 Eri b. This is another example of similar characteristics between 51 Eri b and GJ 570 D. This detection is at a confidence of $\sim 2.7\sigma$ ($log\,(b) = 2.36$) for the data set combining SPHERE and GPI obervations, and at 2.5σ confidence ($log\,(b) = 1.95$) for the data set employing only GPI observations. This was done using a Bayes factor to sigma conversion (Trotta 2008). We note that these specific confidence test retrievals were conducted with a radius prior of $R_{jup} = 1.2 \pm 0.2$ with only one scaling factor S_{cal} per data combination. If verified, this would be the first detection and constraint on the presence of ammonia in a directly imaged exoplanet. This molecular species is present in planet forming, or protoplanetary discs (Salinas et al. 2016) and has long been included in models of substellar atmosphere (Ackerman & Marley 2001; Saumon et al. 2012). It is also shown to be present in Jupiter's atmosphere (Becker et al. 2020).

We note, however, that this detection in both retrievals presented in Fig. 11 and Table 6 are driven by the GPI *K*1- and *K*2-band data. We do not detect ammonia when only analysing the *Y*-, *J*-, and *H*-band data as shown in Fig. A5. Also, as noted previously, this analysis does not account for potential cross-correlated noise, which could reduce the confidence of this detection.

5.3 Temperature–pressure profile

The derived *npoint* temperature–pressure profiles, retrieved with each data set are shown in Fig. 11. Despite the differing spectral data inputs, the results are similar and are consistent at the 2-sigma level. We attribute the hotter profile between ∼20 to 100 bar when using only GPI data due to this having a brighter *J*-band peak compared to the SPHERE *J*-band peak as shown in Fig. 2b.

We do not include the temperature–pressure profile retrieved using only the SPHERE data as this is an imperfect solution, as mentioned previously (given its inability to explain the GPI *K*1 and *K*2 data). This is a symptom of neglecting data, photometric, and spectroscopic, at the longer wavelengths in the case of this particular retrieval. Samland et al. (2017) avoided such an issue by employing photometric data points at longer wavelengths. The similarities in atmospheric properties between 51 Eri b and GJ 570D, as highlighted in the previous subsection, also encompass the temperature–pressure profile. This is shown in Fig. 12, where the temperature gradients of both objects are similar but 51 Eri b has a slightly steeper, and thus more isothermal, temperature gradient in the photosphere.

In Fig. 13, we show how our retrieved temperature–pressure profile differs mainly in the *Y*- and *J*-band photospheric contributions regions when compared to the radiative-convective equilibrium profile from the ATMO 2020 grid models. In other words, our retrieval analysis derives cooler *Y*- and *J*-band photospheric temperatures.

The differences between profiles derived for 51 Eri b compared to the GJ 570 D retrieval (see Fig. 12) and ATMO 2020 fitting (see Fig. 13) may indicate the presence of an unmodelled cloud as the profile departs from an adiabatic gradient and becomes more isothermal. This kind of behaviour has been noted in previous retrieval studies such as Burningham et al. (2017) and Mollière et al. (2020) when a cloudless retrieval attempted to account for clouds included in mock data by making the profile more isothermal. As we noted above, our retrieved profile acts to produce a cooler *Y* and *J* photosphere and as such may be inadvertently mimicking the presence of a cloud layer. Alternatively, reduced, non-adiabatic temperature gradients triggered by chemical transitions have been suggested as an explanation for the SEDs of brown dwarfs (Tremblin et al. 2016, 2019). Thus, the retrieved non-adiabatic temperature profile could also be indicative of thermocompositional convection taking place in the atmosphere of 51 Eri b.

5.4 The question of formation

Parameters derived from retrieval analysis can allow us to peer into the formation history of exoplanets (Gravity Collaboration 2020; Mollière et al. 2022). Here, one can attempt to differentiate between possible formation mechanisms for 51 Eri b, primarily gravitational instability (GI) (Bodenheimer 1974; Boss 1997; Durisen et al. 2007) or core accretion (Pollack et al. 1996; Lissauer & Stevenson 2007). GI is a rapid mechanism that has similarities with the general star formation process. When the system is very young, the disc may become massive enough to become gravitationally unstable, producing spiral density waves that may collapse to form bound objects, which could then slowly contract to produce planetary-mass bodies. Core accretion occurs when an initial solid core forms, then slowly accretes gas from the surrounding disc. If the mass gets high enough, it may enter a phase of 'runaway gas accretion' where the protoplanet very rapidly gains a significant amount of gas. Overall, the time-scale of core accretion is much longer than that of GI.

All our derived radii values are consistent with both the classical cold start and and hot start planetary thermal evolution models from Fortney et al. (2008) (updated models from Marley et al. 2007) as outlined in Fig. 15, using age estimates from Macintosh et al. (2015) or Rajan et al. (2017). Fig. 15 also shows that our derived surface gravity values are consistent with both classical cold start and hot start model predictions at the 2-sigma level. As such, with the current uncertainties derived from retrievals such as that presented in this study, we are unable to differentiate between formation pathways using these models.

However, using carbon and oxygen abundances from Luck (2017) (Identifier: c Eri, Carbon log $\varepsilon = 8.41$, Oxygen log $\varepsilon = 8.80$) we derive a C/O ratio of ∼0.41 for 51 Eri. Therefore, the large mass retrieved for 51 Eri b, its ∼13 au separation, both coupled with a super-stellar C/O ratio could point towards formation via GI (e.g. Vigan et al. 2017). A core accretion pathway would happen on a much longer time-scale resulting in planetesimal enrichment (Mordasini et al. 2016), thus lowering the initial C/O ratio (Espinoza et al. 2017). However, Ilee et al. (2017) illustrate that even GI could produce a wide range of possible atmospheric abundances and so one should interpret the C/O ratio with caution. We also note that a super-stellar C/O ratio for a T dwarf could also be due to oxygen depletion via condensate processes and the formation of clouds below the photosphere (Burrows & Sharp 1999; Lodders & Fegley 2006). Therefore, the use of inferred C/O ratio informing on possible formation pathways should be approached with caution for T dwarf exoplanets.

5.5 Cloudy versus cloudless retrievals

In order to perform a cloudy vs cloudless comparison we used the simpler but less inflexible temperature-pressure profile parametrization from Lavie et al. (2017). This was combined with both the uniform and patchy power law deck and slab clouds outlined in Section 3.2.8. This analysis was motivated by the cloudy results of Samland et al. (2017) and Rajan et al. (2017) as well as the the well documented trend of flexible temperature–pressure structures being

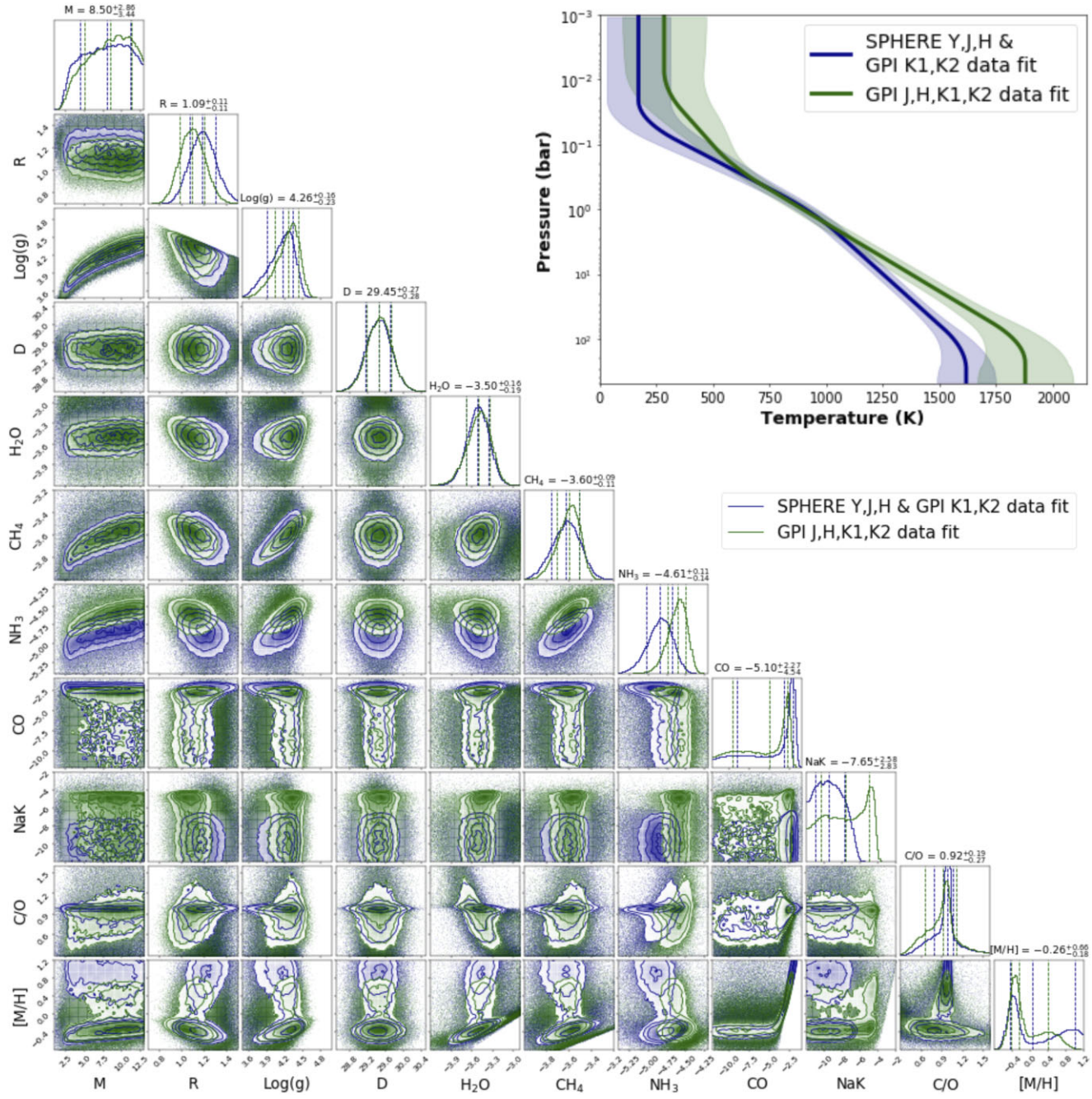


Figure 12. 51 Eri b posteriors. Blue indicates the retrieved values from the SPHERE *Y*, *J*, *H* and GPI *K*1, *K*2 data set. Green indicates the retrieved values from the GPI *J*, *H*, *K*1, *K*2 data set. Log (*g*), C/O, and [M/H] posteriors are inferred parameters, while all the other parameters are sample as part of the retrieval. A comparison of the retrieved temperature–pressure profiles from each respective data set are also shown.

able to mimic spectral imprint of cloud opacity. An overview of out cloudy retrieval evidences are outlined in Table 7.

When considering the use of the Lavie et al. (2017) profile only, the Log (Ev) derived a strong preference for clouds to be included (Table 7), matching the results of Samland et al. (2017) and Rajan et al. (2017). We also match Samland et al. (2017) and Rajan et al. (2017) in relation to no preference for patchy clouds versus a preference for patchy clouds, relative to the data they used when performing model fitting analysis. However, in the cases of both broad-band spectroscopy combinations used in our analysis, the retrievals determined a higher or comparable Log (Ev) when clouds were omitted and the more flexible *npoint* temperature–pressure profile was employed (see Table 7). These results, however, do not mean that 51 Eri b is cloud free. In fact, many studies (Burningham et al. 2017; Mollière et al. 2020) have documented that retrievals can often use flexible profile to mimic the presence and spectral contributions of clouds, even when retrieving on synthetic data where clouds were included. As can be seen in Fig. 14, the retrieved temperature–pressure structure is significantly different in the contributing photosphere where the cloud contribution is needed to counteract the enforcement of a more adiabatic, less isothermal and hotter profile. Burningham et al. (2021) have shown that longer wavelength data, particularly in the mid-

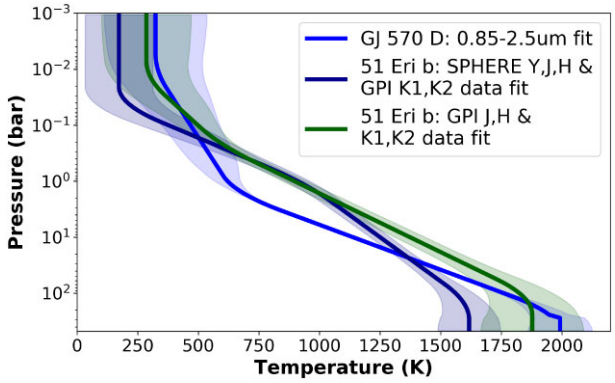


Figure 13. Retrived 51 Eri b TP profile compared to the GJ 570D TP profile.

Table 7. Comparison of 51 Eri b retrieval Log Evidences when employing different model set-ups. L17: Lavie et al. (2017). GPI all: GPI *J*, *H*, *K*1, *K*2 data. SPHERE & GPI: SPHERE *Y* , *J* , *H* & GPI *K*1, *K*2 data. SPHERE only: SPHERE *Y*, *J*, *H* data. Uniform: Uniform clouds coverage. Patchy: Patchy cloud coverage.

Cloud Type	TP Profile Type	Log(Ev)
Deck cloud, Patchy (GPI all)	L17, inflexible	5596.29
Cloudless (GPI all)	*npoint*, flexible	5596.02
Slab cloud, Patchy (GPI all)	L17, inflexible	5593.38
Deck cloud, Uniform (GPI all)	L17, inflexible	5591.43
cloudless (GPI all)	L17, inflexible	5588.28
Slab cloud, Uniform (GPI all)	L17, inflexible	5587.58
Cloudless (SPHERE & GPI)	*npoint*, flexible	5141.33
Deck cloud, Uniform (SPHERE & GPI)	L17, inflexible	5132.55
Slab cloud, Uniform (SPHERE & GPI)	L17, inflexible	5130.97
Deck cloud, Patchy (SPHERE & GPI)	L17, inflexible	5130.05
Slab cloud, Patchy (SPHERE & GPI)	L17, inflexible	5128.25
cloudless (SPHERE & GPI)	L17, inflexible	5105.07

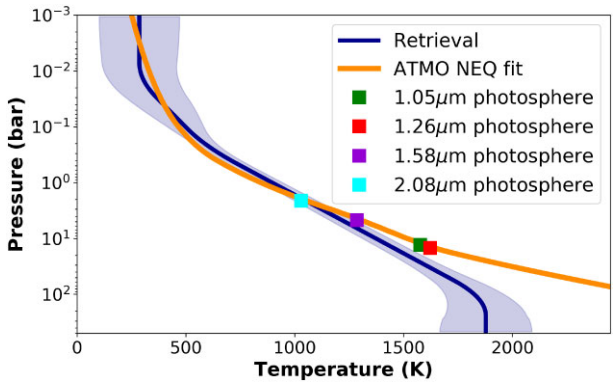


Figure 14. Retrieved temperature pressure profiles using the npoint and Lavie et al. (2017) prescriptions. Also indicated is the retrieved cloud position when using the Lavie et al. (2017) profile. Top: Using SPHERE *Y*-, *J*-, *H*-, and GPI *K*1-, *K*2-band data. Bottom: Using GPI *J*-, *H*-, *K*1-, and *K*2-band data.

IR, is crucial for accurately constraining the temperature–pressure profile when using a reasonably flexible temperature–pressure profile parametrization. It is also very likely that the simple and flexible power-law cloud parametrizations included in our analysis were not the most suitable compared to more physically motivated and condensate specific cloud modelling.

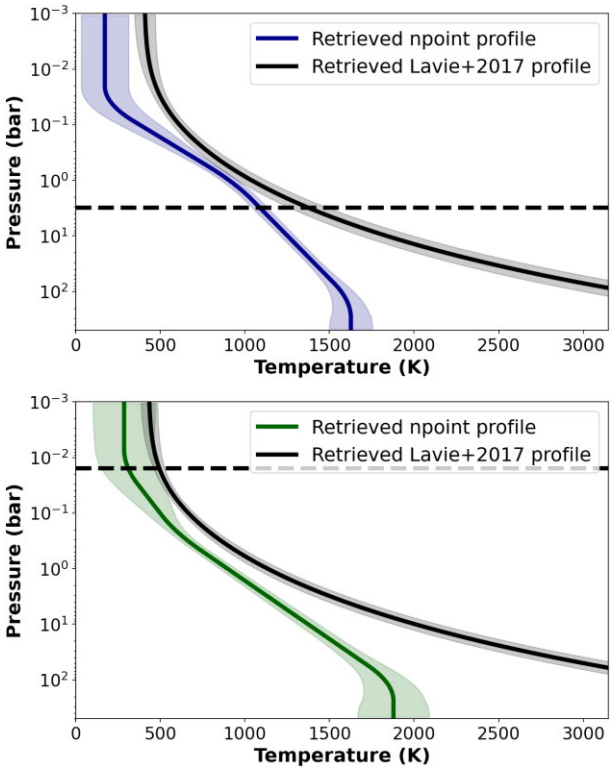


Figure 15. Temperature–pressure profile comparison between profiles derived by TAUREX3 and ATMO 2020 for 51 Eri b.

6 DISCUSSION

We have presented retrieval results which are consistent with previous studies and often provide improvements relative to forward models used in non-retrieval studies. This can mainly be attributed to the increased flexibility of model parameters, especially in the free chemistry retrievals. However, consistency between retrieval studies is encouraging when taking into account the use of different samplers, temperature–pressure prescriptions and differing cross-section inputs. The success of these studies demonstrates the scope for application of these tools to both the extensive archival data and future planned observations of brown dwarfs and directly imaged giant exoplanets.

There are, of course, limitations and imperfections in our retrieval analysis as we make assumptions such as isoprofile (constant) mixing ratios, something not expected to be the case in real atmospheres. However, adding additional capabilities to existing retrieval frameworks, such as non-isoprofile mixing ratios, will certainly be probed in future work using both archival and future observation of directly imaged exoplanets and brown dwarfs. In fact, the current quality and quantity of brown dwarf observations offer a perfect testbed for new modelling parametrizations.

Retrieval analysis is also quite computationally intensive, often requiring computing clusters to run within a reasonable time frame when compared to simply iterating over a grid of forward models. This could become increasingly problematic when significantly higher resolution and increased spectral coverage observations from JWST allow for further parameters to be probed, increasing the overall parameter space and, hence, the computational expense. Recently, however, efforts have been made to use machine learning for the model selection, showing the possibilities for significant gains in computational efficiency (Zingales & Waldmann 2018).

Based on our analysis of the GJ 570D spectrum, we found that different approaches when considering the Na + K line broadening can have a significant effect on the retrieved abundances. This seems to have knock-on effects with other retrieved parameters, such as radius and mass, seemingly in an attempt to preserve surface gravity for a larger object whilst driving up the metallicity. These parameters have been found to be degenerate in other studies such as Kitzmann et al. (2020). Some potential reasons for the issues caused by the Na and K cross-sections are outlined below.

(i) The profiles of Allard et al. (2016, 2019) are only considered valid up to a H_2 density of $10^{21}\,cm^{-3}$. They therefore break down at pressures above 10–100 bar. Kitzmann et al. (2020) took the approach here of switching back to Voigt profiles at these high pressures. The difference in cross-sections computed using these varying approaches is illustrated in Fig. 6. It can be seen that the divergence between the computed cross-sections used in this work and those of Kitzmann et al. (2020) are much higher at larger pressures for this reason. We would not expect the deviations at such larger pressures to have such an impact, but it is worth looking into this more in the future.

(ii) The line cores of Na and K, based on the data of Allard et al. (2016, 2019), are computed considering Lorentzian broadening only, with the effects of Doppler broadening not taken into consideration. It is possible this has some effect and is worthy of further investigation.

(iii) We considered the effects due to using both different sources and different line-wing cutoffs for the non-resonance lines of Na and K. We compared using lines from the NIST (Kramida et al. 2013) and the Kurucz (Kurucz & Bell 1995) data bases, and found Kurucz contains more lines for both Na and K. The effects of these various approaches for different pressures and temperatures can be seen in Fig. 6. The larger number of non-resonance lines in the Kurucz data base leads to a larger overall opacity when pressure broadening is taken into account, with more pronounced effects at higher pressures. However, the use of the different sources for the non-resonance data was found to have negligible effect on the retrieval results.

(iv) We tried using a completely different scheme for treating the line profiles of the Na and K resonance doublets; that of Burrows & Volobuyev (2003). The use of these cross-sections did show some effects in terms of the retrieved parameters of Na + K abundance, radius, and mass. However, they still did not give physically plausible radius and mass values, which led us to proceed with the method of splitting the spectra into two regions, as outlined in Section 4.1

(v) We only use H_2-broadening and not He-broadening for the Na and K resonance doublets. As the contribution from H_2-broadening is much higher than from He, this is thought to be a good approximation. However, it would be worth looking into He-broadening in the future, as outlined in Peach et al. (2020) and Peach (2017). We did not implement an instrument profile within our analysis as used in Kitzmann et al. (2020). Such a profile can account for flux being spread across instrument pixels. TAUREX3 does not account for such a spread when binning the higher resolution forward model to the resolution of the observation. Such an instrument profile could explain why Helios-R (Kitzmann et al. 2020) was able to fit the heavily alkali influenced *J*-band peak more successfully than TAUREX3 and thus may have helped negate the bias issue we experienced in this study. This was also one of the few differences in our retrieval approach and that outlined in the (Kitzmann et al. 2020) study.

We could not identify the exact source of the issue causing unrealistic radius and mass values to be retrieved when the full wavelength coverage spectrum is used. However, it is apparent that the Na + K cross-sections used in the retrievals can have a significant impact on the retrieved parameters. This suggests that the *Y*-band sodium and potassium lines could be driving both retrieval and grid-modelling approaches to derive masses which converge towards the upper bounds of the prior space (Line et al. 2015; Schneider et al. 2015; Zalesky et al. 2019).

We note there are now several studies that have used the updated broadening coefficients from Allard et al. (2016, 2019) in analysis of T dwarf spectra. One of these studies, (Kitzmann et al. 2020), did not encounter this issue for GJ 570D. Oreshenko et al. (2020) negated these known issues when modelling the 0.85–1.2 μm region by neglecting this wavelength region in their analysis. Piette & Madhusudhan (2020) modulated their K cross-sections with a multiplicative factor within their retrieval while only analysing data >1.1 μm. This topic warrants further investigation in the future, but we note that it may be less prevalent when studying data from JWST, which will benefit from having wider wavelength spectral coverage, down-weighting the problematic Na + K dominated region when carrying out retrieval analysis. Further studies of the broadening behaviour of Na and K lines in laboratory settings would likely prove invaluable.

More dynamical (model independent) constraints for directly imaged exoplanets and brown dwarfs will help reduce the volume of parameter space explored by the retrieval method. Such measurements have been carried out for Gl 229B (Brandt et al. 2019), ultracool binaries (Dupuy & Liu 2017) and Beta Pic b (Snellen & Brown 2018; Dupuy et al. 2019), with *HST* monitoring campaigns also underway for cool brown dwarfs (Dupuy 2018; Dupuy et al. 2020). This would significantly improve constraints on retrieval mass priors, and may also help constrain the radius values retrieved in various studies as surface gravity plays a key role in shaping the SED. Retrieval analyses have, quite often, returned physically improbable radius values, both in the results presented here (which we attribute to the issues of the Na + K opacities) and in other studies (Kitzmann et al. 2020). Additionally, the temperature–pressure structure would also likely be better constrained, as radius and effective temperature are inversely correlated. During this study, we have seen examples of pressure–temperature profiles changing as a result of varying radii whilst maintaining a similar surface gravity, demonstrating a significant and problematic degeneracy. Dynamical and model independent mass measurements for objects in the directly imaged regime will help constrain the parameter space significantly. Better parallax measurements, such as from *Gaia*, help constrain the scaling factor (radius) further. For example, both our and the Kitzmann et al. (2020) study benefited from better distance constraints versus that of Line et al. (2015) and Burningham et al. (2017). The narrowing of parameter space for these model *drivers* may result in the ability to better probe other, more elusive, properties, and will also reduce the computational expense of retrieval analyses.

We retrieved very similar effective temperatures and abundances for both 51 Eri b and GJ 570D. This further supports the use of brown dwarfs as proxies for the harder-to-observe cohort of planetary mass companions. Another example of a close exoplanet analogue is PSO J318.5−22, a free-floating planetary mass brown dwarf with a spectrum which closely matches those of the atmospheres of the HR 8799 planets (Liu et al. 2013; Bonnefoy et al. 2016; Miles et al. 2018). These free floating objects are much easier to observe and can offer a window into their characteristic counterpart exoplanets, as we can make use of the superior quality of spectral data availability for these objects. Therefore, in the same way, PSO 318 has long been documented to have overlapping properties with the same spectral type HR 8799 planets, 51 Eri b also has striking chemical similarities to the benchmark T dwarf GJ 570D and other late T dwarfs from the Line et al. (2015, 2017) studies.

The atmospheric similarities between the bona fide exoplanet 51 Eri b and late-T field brown dwarfs extend to mixing ratios, most notably that of ammonia. We acknowledge, though, that such a tentative detection, motivated by the GPI *K*-band data, needs further observations to provide a higher confidence detection. This could, perhaps, be achieved using VLT-GRAVITY (Gravity Collaboration 2017), Subaru-REACH (Lozi et al. 2018; Kotani et al. 2018), or KECK-KPIC (Pezzato et al. 2019). These instruments deliver higher-resolution observations than that provided by SPHERE and GPI. This would allow us to detect more subtle features. The high-resolution data from REACH and KPIC would allow us to probe individual lines using both retrievals and cross-correlation methods (Hoeijmakers et al. 2018; Brogi & Line 2019).

We only employed a single scaling S_{cal} factor for 51 Eri b when considering data take from a single instrument and two in the case of the SPHERE plus GPI combination. However, this may be an imperfect approach in the case of using only the GPI data as this spectrum is stitched together from different bands which can employ different data reduction pipelines and photometric calibrations. Such an approach of allowing each band to scale independently was a successfully strategy adopted in Nowak et al. (2020) when combining observations of Beta Pic b. Crucially, such flexibility appears employable when using a high quality data set, as in the case of the GRAVITY Beta Pic b data used in the Nowak et al. (2020), with this data appearing to anchor the model and deriving a very small uncertainty for the GRAVITY data scaling factor. Our 51 Eri b data quality from GPI data is such that we did not find this necessary, given the large uncertainties present in the data we analyse in this study. Future studies should be able to allow for scaling factors in each band when improved data become available for this exoplanet. The S_{cal} factor is directly correlated to the retrieved radii and can act to help the retrievals to maintain a physically sensible and higher radii instead of purely accounting for possible calibration imperfections, creating a degeneracy. This behaviour is likely exacerbated by the trend of retrievals deriving small radii (Burningham et al. 2021). Our experience of this factor with the data sets used in this study is that it commonly acted to scale down the model ($S_{cal} < 1$) which can then be counterbalanced by a higher radius, especially given the priors we applied in the case of 51 Eri b. This is why we placed a Gaussian prior on this parameter when also employing one on the radius parameter, in an effort to restrict this degeneracy and the ability for the retrieval to simply use S_{cal} to retrieve our set radius prior. This will likely be a continued issue for retrievals going forward, where scaling factors designed for flux calibration and possible variability considerations could mask the documented inability of models to derive expected radii values, especially when using flat priors.

Unlike previous studies, we were able to fit the spectral profile of 51 Eri b without clouds. This is an interesting and important result as previous studies all employed cloud models within grid modelling, often based on more rigidly parametrized temperature–pressure profile assumptions (e.g. radiative–convective equilibrium) and chemistry. We acknowledge and stress, however, that our ability to fit the data with a preference for a cloudless modelling may be due to our flexible temperature–pressure profile being able to mimic and account for the presence of an unmodelled cloud. Our result matches with that from Burningham et al. (2017) and Mollière et al. (2020), where synthetic data of cloudy L dwarfs was successfully fit due to the use of a flexible temperature–pressure profile. Mollière et al. (2020) also showed that when an incorrect cloud model was employed to fit synthetic cloudy data the retrieval determines a preference for a cloudless fit. This may be indicative that the power-law deck and slab clouds were insufficient for analysis of this target and future work on 51 Eri b should include a more diverse set of cloud modelling. We attempted to explore a more physically motivated cloud prescription, compared to the the power-law parametrization, using the cloud parametrization outlined in Lee et al. (2013) and employed in Lavie et al. (2017). However, the data quality for 51 Eri b is such that these cloud parameters could not be constrained to a useful extent and thus this analysis was inconclusive and omitted from this article. The degenerate ability for a flexible temperature–pressure profile to account for clouds in the absence of any cloud modelling within a retrieval may be negated in the future by employing data across a wider wavelength range, when such data becomes available. Retrieval analysis including clouds will be further explored in further work when improved data becomes available for 51 Eri b, such as the further photometric points from JWST Program #1412 which may assist in breaking model degeneracies.

Our ability to fit the 51 Eri b data without clouds may have also been assisted by the free chemistry nature of the retrieval, where grid-models are often much more constrained based on employment of coarse parameter sampling, solar abundance ratios and chemical equilibrium. In the case of abundances, for example, exoplanets have been shown to possess a variety of chemical compositions, often deviating from norms seen in our own Solar system. For example, exoplanets can possess C/O ratios much higher than that present in our Solar system (Madhusudhan, Lee & Mousis 2012; Moses et al. 2013). This is further shown by the super-stellar C/O ratio we retrieved for 51 Eri b. This parameter allows us to hypothesize possible formation pathways. Due to 51 Eri b's large retrieved mass, measured orbital separation and retrieved C/O ratio, we suggest this may hint at formation via gravitational instability (Vigan et al. 2017). Further spectral observations of 51 Eri b, using instruments such as GRAVITY, may help further constrain the C/O ratio and permit a more in depth analysis of possible formation scenarios for this exoplanet.

Overall, we suggest the best approach is testing and exploring a suite of free-retrieval set-ups along with self-consistent modelling, as performed in this study, when characterizing self-luminous objects. Ideally, when further data become available from instruments aboard JWST, GPI2, and SPHERE+, the results derived from these different approaches should converge to agreement.

7 SUMMARY

We introduce TAUREX3 which we have modified to be suitable for directly imaged objects, and apply it to the benchmark brown dwarf GJ 570D and the cool exoplanet 51 Eri b.

We discuss issues with the Na + K cross-sections when applied to T dwarf spectra. The retrievals converged to a high mass and radius, likely due to biases introduced by the methods used to compute these cross-sections. This issue was overcome by splitting the retrieval into two parts. Part 1 retrieved the mass, radius, distance, and S_{cal} using the 1.2–2.5 μm data, while part 2 retrieved the chemical profile of the atmosphere using the 0.85–2.5 μm data. This allowed for more plausible results.

We compared our GJ 570 D results with other studies that performed retrieval analyses of this object (Line et al. 2015; Burningham et al. 2017; Kitzmann et al. 2020). The different analysis of GJ 570D, across various retrieval codes, shows an encouraging stability of most parameters, especially relating to the atmospheric chemistry as well as the temperature–pressure profile. We therefore successfully demonstrate TAUREX3's suitability for brown dwarf emission analysis.

We also carried out free chemistry and cloudless retrieval analyses on all published spectroscopy observations of 51 Eri b, while

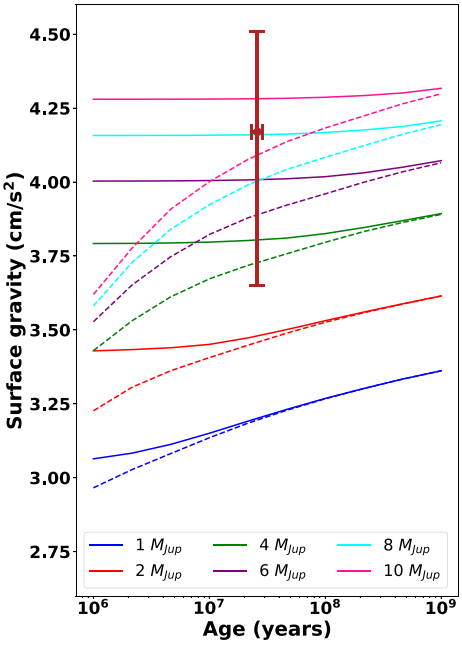


Figure 16. Planetary thermal evolution tracks for different planet masses from Fortney et al. (2008), updated from Marley et al. (2007). Dotted lines indicate hot start planets. Solid lines indicate cold start planets. The error bar indicates the retrieved 2 sigma confidence boundary for surface gravity from the SPHERE *Y*, *J*, *H*, and GPI *K*1, *K*2 data set with the age estimate from Rajan et al. (2017).

comparing our results to previous studies that used grid modelling. The main results of our 51 Eri b retrieval analysis are

(i) Our retrievals result in excellent fits to the observations without requiring cloud extinction, deriving a higher Log (Ev) when compared to retrievals including power-law clouds. This is in contrast to the cloudy atmosphere conclusions made in all previous studies (Macintosh et al. 2015; Samland et al. 2017; Rajan et al. 2017) who employed grid model fitting. However, this could be due to our flexible temperature–pressure profile being able to account for unmodelled clouds with this behaviour also being seen in Burningham et al. (2017) and Mollière et al. (2020).

(ii) We confirm and constrain the presence of H_2O and CH_4.

(iii) We find tentative evidence of NH3 in the atmosphere of 51 Eri b, to a ∼2.7 sigma confidence. Further observations are required to confirm this.

(iv) We retrieve a supersolar C/O ratio, and a solar consistent [M/H] for 51 Eri b.

(v) Our surface gravity values are consistent with both classical hot-start and cold-start planetary thermal evolution models from Fortney et al. (2008) as shown in Fig. 16.

(vi) We demonstrate the importance of the *K*-band observations for constraining the effective temperature and temperature–pressure profile.

(vii) Our highest Log (Ev) retrieval literature consistent radius values of $1.18^{+0.12}_{-0.12}\,R_{Jup}$ and $1.09^{+0.11}_{-0.11}\,R_{Jup}$ for our two data sets. This is despite not employing cloud modelling, something previous studies struggled to do.

(viii) Our analysis highlights strong similarities between the retrieved molecular mixing ratios and temperature–pressure profiles of 51 Eri b and GJ 570D. The slight gradient differences in temperature–pressure profiles is attributed to possibly accounting for an unmodelled cloud structure in the case of 51 Eri b's retrieval by adopting a more isothermal gradient.

(ix) Our retrieved *npoint* temperature–pressure profiles for 51 Eri b adopts a much more isothermal profile compared to the adiabatic profile employed in the unsuccessful ATMO 2020 grid model fit. This more isothermal profile, again, could account for the impact of an unmodelled photospheric cloud structure, or alternatively could be indicative of diabatic convection triggered by the CO/CH_4 chemical transition (Tremblin et al. 2016, 2019) in the atmosphere of 51 Eri b.

(x) Our retrieved super-stellar C/O ratio, coupled with our retrieved mass and previously measured orbital separation, hints at a possible formation pathway of gravitational instability for 51 Eri b. However, this conclusion is tentative and higher quality data is required for a more thorough analysis of the possible formation history of 51 Eri b.

(xi) When including clouds along with a less flexible temperature–pressure profile, our retrievals derived a strong preference for the inclusion of clouds. However, the Log (Ev) never surpassed that of the cloudless retrieval when employing the more flexible *npoint* temperature–pressure profile. Table 7 shows that the conclusions of our retrievals, in relation to cloudy versus cloud-free and patchy clouds versus uniform clouds, can be biased by model set-up and which data are employed. This points to the need for diverse and rigours testing of retrieval model set-ups and data combinations, particularly when analysing high-contract imaging data with challenging SNRs.

ACKNOWLEDGEMENTS

NW thanks the referee of this paper for providing very constructive and insightful feedback, helping to improve this resulting manuscript. NW thanks Ben Burningham for the discussion and assistance with implementing cloud models. NW thanks Jackie Faherty for comments provided on this article. NW also thanks the BDNYC retrieval group for useful discussions regarding this work. NW acknowledges postgraduate funding from the UK's Science and Technology Facilities Council (STFC), NSF Award 1909776 and NASA Award 80NSSC22K0142. IW acknowledges funding from the UK's Science and Technology Facilities Council (STFC) and from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 758892, ExoAI). Furthermore, IW acknowledges funding by the Science and Technology Funding Council (STFC) grants: ST/K502406/1, ST/P000282/1, ST/P002153/1, and ST/S002634/1. NS acknowledges funding from the OCAV-PSL project. The authors wish to acknowledge the very significant cultural role and reverence that the summit of Mauna Kea has always had within the indigenous Hawaiian community. We are most fortunate to have the opportunity to use observations conducted from this mountain. Some of the data used herein were obtained at the W. M. Keck Observatory, which is operated as a scientific partnership among the California Institute of Technology, the University of California, and the National Aeronautics and Space Administration. The Observatory was made possible by the generous financial support of the W. M. Keck Foundation. This publication makes use of data products from the Two Micron All Sky Survey, which is a joint project of the University of Massachusetts and the Infrared Processing and Analysis Center/California Institute of Technology, funded by the National Aeronautics and Space Administration and the National Science Foundation. This work used the Cirrus UK National Tier-2 HPC Service at EPCC (http://www.cirrus.ac.uk) funded by the University of Edinburgh and EPSRC (EP/P020267/1)

DATA AVAILABILITY

GJ 570D data can be accessed via the SpeX Prism Library (see footnote 1). Data used for 51 Eridani can be accessed via supplementary materials from Rajan et al. (2017) and Samland et al. (2017).

REFERENCES

Abel M., Frommhold L., Li X., Hunt K. L., 2011, J. Phy. Chem. A, 115, 6805
Abel M., Frommhold L., Li X., Hunt K. L., 2012, J. Chem. Phys., 136, 044319
Ackerman A. S., Marley M. S., 2001, ApJ, 556, 872
Akeson R. L. et al., 2013, PASP, 125, 989
Al-Refaie A. F., Changeat Q., Waldmann I. P., Tinetti G., 2021, ApJ, 917, 37
Al-Refaie A. F., Changeat Q., Venot O., Waldmann I. P., Tinetti G., 2022, ApJ, 932, 123
Allard N. F., Spiegelman F., Kielkopf J. F., 2016, A&A, 589, A21
Allard N. F., Spiegelman F., Leininger T., Mollière P., 2019, A&A, 628, A120
Amundsen D. S., Baraffe I., Tremblin P., Manners J., Hayek W., Mayne N. J., Acreman D. M., 2014, A&A, 564, A59
Amundsen D. S., Tremblin P., Manners J., Baraffe I., Mayne N. J., 2017, A&A, 598, A97
Asplund M., Grevesse N., Sauval A. J., Scott P., 2009, ARA&A, 47, 481
Azzam A. A. A., Yurchenko S. N., Tennyson J., Naumenko O. V., 2016, MNRAS, 460, 4063
Barstow J. K., Heng K., 2020, Space Sci. Rev., 216, 82
Barstow J. K., Changeat Q., Garland R., Line M. R., Rocchetto M., Waldmann I. P., 2020, MNRAS, 493, 4884
Becker H. N. et al., 2020, Nature, 584, 55
Benneke B., 2015, preprint (arXiv:1504.07655)
Bernath P. F., 2020, J. Quant. Spectrosc. Radiat. Transf., 240, 106687
Beuzit J.-L. et al., 2008, in McLean I. S., Casali M. M., eds, Proc. SPIE Conf. Ser. Vol. 7014, Ground-based and Airborne Instrumentation for Astronomy II. SPIE, Bellingham, p. 701418
Beuzit J. L. et al., 2019, A&A, 631, A155
Biller B. A., Bonnefoy M., 2018, Exoplanet Atmosphere Measurements from Direct Imaging. Springer, Berlin, 101
Bodenheimer P., 1974, Icarus, 23, 319
Bonnefoy M. et al., 2016, A&A, 587, A58
Boss A. P., 1997, Sci., 276, 1836
Bowler B. P., Zhou Y., Morley C. V., Kataria T., Bryan M. L., Benneke B., Batygin K., 2020, ApJ, 893, L30
Brandl B. R. et al., 2014, in Ramsay S. K., McLean I. S., Takami H., eds, Proc. SPIE Conf. Ser. Vol. 9147, Ground-based and Airborne Instrumentation for Astronomy V. SPIE, Bellingham, p. 914721
Brandt T. D., Dupuy T. J., Bowler B. P., Bardalez Gagliuffi D. C., Faherty J., Mirek Brandt G., Michalik D., 2019, AJ, 160, 196
Brogi M., Line M. R., 2019, AJ, 157, 114
Buchner J. et al., 2014, A&A, 564, A125
Bueno T. J., Bendicho F. P., 1995, ApJ, 455, 646
Burgasser A. J., 2014, in Signh H.P., Prugniel P. , Vauglin I., eds, ASI Conf. Ser. Vol. 11, International Workshop on Stellar Spectral Libraries. p. 7
Burgasser A. J., Splat Development Team, 2017, in Coelho P., Martins L., Griffin E., eds, ASI Conf. Ser. Vol. 14, p. 7
Burgasser A. J. et al., 2000, ApJ, 531, L57
Burgasser A. J., McElwain M. W., Kirkpatrick J. D., Cruz K. L., Tinney C. G., Reid I. N., 2004, AJ, 127, 2856
Burgasser A. J., Burrows A., Kirkpatrick J. D., 2006, ApJ, 639, 1095
Burningham B., et al., 2021, MNRAS, 506, 1944
Burningham B., Marley M. S., Line M. R., Lupu R., Visscher C., Morley C. V., Saumon D., Freedman R., 2017, MNRAS, 470, 1177
Burrows A. et al., 1997, ApJ, 491, 856
Burrows A., Sharp C. M., 1999, ApJ, 512, 843
Burrows A., Volobuyev M., 2003, ApJ, 583, 985
Chabrier G., Baraffe I., Leconte J., Gallardo J., Barman T., 2009, in Stempels E., ed., AIP Conf. Ser. Proc. Vol. 1094, 15th Cambridge Workshop on Cool Stars, Stellar Systems, and the Sun. Am. Inst. Phys., New York, p. 102
Changeat Q., Edwards B., Waldmann I. P., Tinetti G., 2019, ApJ, 886, 39
Charnay B., Bézard B., Baudino J. L., Bonnefoy M., Boccaletti A., Galicher R., 2018, ApJ, 854, 172
Chilcote J. et al., 2017, AJ, 153, 182
Chubb K. L. et al., 2020, A&A, 639, A3
Coles P. A., Yurchenko S. N., Tennyson J., 2019, MNRAS, 490, 4638
Crossfield I. J. M. et al., 2014, Nature, 505, 654
Cushing M. C., Vacca W. D., Rayner J. T., 2004, PASP, 116, 362
Cushing M. C. et al., 2011, ApJ, 743, 50
Del Burgo C., Martín E. L., Zapatero Osorio M. R., Hauschildt P. H., 2009, A&A, 501, 1059
Drummond B., Tremblin P., Baraffe I., Amundsen D. S., Mayne N. J., Venot O., Goyal J., 2016, A&A, 594, A69
Dupuy T. J., 2018, The Coolest Sample of Brown Dwarf Dynamical Masses, HST Proposal, Cycle 27, ID. #15944
Dupuy T. J., Liu M. C., 2017, ApJS, 231, 15
Dupuy T. J., Brandt T. D., Kratter K. M., Bowler B. P., 2019, ApJ, 871, L4
Dupuy T. J., Allers K., Bowler B., Line M., Liu M. C., Marley M. S., Morley C., Zhang Z., 2020, Resolving Mass Benchmarks for Ultracool Atmospheres, HST Proposal, Cycle 28, ID. #16268
Durisen R. H., Boss A. P., Mayer L., Nelson A. F., Quinn T., Rice W. K. M., 2007, in Reipurth B., Jewitt D., Keil K., eds, Protostars and Planets V. University of Arizona Press, Tucson, p. 607
Edwards B. et al., 2020, AJ, 160, 8
Espinoza N., Fortney J. J., Miguel Y., Thorngren D., Murray-Clay R., 2017, ApJ, 838, L9
Feroz F., Hobson M. P., 2008, MNRAS, 384, 449
Feroz F., Hobson M. P., Bridges M., 2009, MNRAS, 398, 1601
Feroz F., Hobson M. P., Cameron E., Pettitt A. N., 2013, Open J. Astrophys., 2, 10
Filippazzo J. C., Rice E. L., Faherty J., Cruz K. L., Van Gordon M. M., Looper D. L., 2015, ApJ, 810, 158
Fletcher L. N., Gustafsson M., Orton G. S., 2018, ApJS, 235, 24
Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306
Fortney J. J., 2018, in Bozza V., Mancini L., Sozzetti A., eds, Astrophysics of Exoplanetary Atmospheres. Springer
Fortney J. J., Marley M. S., Saumon D., Lodders K., 2008, ApJ, 683, 1104
Gaia Collaboration, 2016, A&A, 595, 36
Gaia Collaboration, 2018, A&A, 616, 22
Gandhi S., Madhusudhan N., 2018, MNRAS, 474, 271
Gardner J. P. et al., 2006, Space Sci. Rev., 123, 485
Gonzales E., Burningham B., Faherty J., Cleary C., Visscher C., Marley M., Lupu R., Freedman R., 2020, ApJ, 905, 46
Gonzales E. C., Burningham B., Faherty J. K., Visscher C., Marley M., Lupu R., Freedman R., Lewis N. K., 2021, ApJ, 923, 19
Gordon S., McBride B. J., 1994, Computer Program for Calculation of Complex Chemical Equilibrium Compositions and Applications. NASA Reference Publ.
Goyal J. M. et al., 2020, MNRAS, 498, 4680
Gravity Collaboration, 2017, A&A, 602, 23
Gravity Collaboration, 2020, A&A, 633, 19
Guillot T., 2010, A&A, 520, A27
Gustafsson B., Edvardsson B., Eriksson K., Jørgensen U. G., Nordlund Å., Plez B., 2008, A&A, 486, 951
Cubillos P. E. et al., 2016, Planet. Sci. J, 3, 81
Helling C., 2019, Annu. Rev. Earth Planet. Sci., 47, 583
Heng K., Mendonça J. M., Lee J.-M., 2014, ApJS, 215, 4
Hoeijmakers H. J. et al., 2018, Nature, 560, 453
Hogg D. W., Bovy J., Lang D., 2010, preprint (arXiv:1008.4686)
Howe A. R., McElwain M. W., Mandell A. M., 2022, ApJ, 935, 107
Ilee J. D. et al., 2017, MNRAS, 472, 189
Irwin P. G. J. et al., 2008, J. Quant. Spec. Radiat. Transf., 109, 1136
Kass R. E., Raftery A. E., 1995, J. Am. Stat. Assoc., 90, 773
Kirkpatrick J. D. et al., 2012, ApJ, 753, 156
Kirkpatrick J. D., 2005, ARA&A, 43, 195
Kitzmann D., Heng K., Oreshenko M., Grimm S. L., Apai D., Bowler B. P., Burgasser A. J., Marley M. S., 2020, ApJ, 890, 174

Kotani T. et al., 2018, Proc. SPIE Conf. Ser. Vol. 10702, Ground-based and Airborne Instrumentation for Astronomy VII. SPIE, Bellingham, p. 1070211
Kramida A., Ralchenko Y., Reader J., 2013, NIST Atomic Spectra Database – Ver. 5, available at http://www.nist.gov/pml/data/asd.cfm
Kurucz R., Bell B., 1995, Kurucz molecular database
Lavie B. et al., 2017, AJ, 154, 91
Lee J.-M., Heng K., Irwin P. G. J., 2013, ApJ, 778, 97
Lew B. W. P. et al., 2020, ApJ, 903, 15
Li G. et al., 2015, ApJS, 216, 15
Line M. R. et al., 2013, ApJ, 775, 137
Line M. R. et al., 2017, ApJ, 848, 83
Line M. R., Fortney J. J., Marley M. S., Sorahana S., 2014, ApJ, 793, 33
Line M. R., Teske J., Burningham B., Fortney J. J., Marley M. S., 2015, ApJ, 807, 183
Lissauer J. J., Stevenson D. J., 2007, in Reipurth B., Jewitt D., Keil K., eds, Protostars and Planets V. Univ. Arizona Press, Tucson, p. 591
Liu M. C. et al., 2013, ApJ, 777, L20
Liu M. C., Leggett S. K., Chiu K., 2007, ApJ, 660, 1507
Lodders K., Fegley B. J., 2006, in Mason J. W., ed., Astrophysics Update 2, Springer Praxis Books. Praxis Publishing Ltd, Chichester, UK, p. 1
Lozi J. et al., 2018, Proc. SPIE Conf. Ser. Vol.10703, Adaptive Optics Systems VI. SPIE, Bellingham, p. 1070359
Luck R. E., 2017, AJ, 153, 21
Lueber A., Kitzmann D., Bowler B. P., Burgasser A. J., Heng K., 2022, ApJ, 930, 136
MacDonald R. J., Madhusudhan N., 2017, MNRAS, 469, 1979
Macintosh B. et al., 2014, Proc. Nat. Acad. Sci., 111, 12661
Macintosh B. et al., 2015, Sci., 350, 64
Madhusudhan N., 2018, in Deeg H. J., Belmonte J. A., eds, Handbook of Exoplanets. Springer International Publ. AG, p. 104
Madhusudhan N., 2019, ARA&A, 57, 617
Madhusudhan N., Seager S., 2009, ApJ, 707, 24
Madhusudhan N., Lee K. K. M., Mousis O., 2012, ApJ, 759, L40
Marley M. S., Fortney J. J., Hubickyj O., Bodenheimer P., Lissauer J. J., 2007, ApJ, 655, 541
Marley M. S., Saumon D., Goldblatt C., 2010, ApJ, 723, L117
Marley M. S., Saumon D., Cushing M., Ackerman A. S., Fortney J. J., Freedman R., 2012, ApJ, 754, 135
McKemmish L. K., Yurchenko S. N., Tennyson J., 2016, MNRAS, 463, 771
McKemmish L. K., Masseron T., Hoeijmakers H. J., Pérez-Mesa V., Grimm S. L., Yurchenko S. N., Tennyson J., 2019, MNRAS, 488, 2836
McLean I. S., Sprayberry D., 2003, in Iye M., Moorwood A. F. M., eds, Proc. SPIE Conf. Ser. Vol. 4841, Instrument Design and Performance for Optical/Infrared Ground-based Telescopes. SPIE, Bellingham, p. 1
Miles B. E., Skemer A. J., Barman T. S., Allers K. N., Stone J. M., 2018, ApJ, 869, 18
Miles B. E. et al., 2020, AJ, 160, 63
Mollière P., van Boekel R., Dullemond C., Henning T., Mordasini C., 2015, ApJ, 813, 47
Mollière P., van Boekel R., Bouwman J., Henning T., Lagage P. O., Min M., 2017, A&A, 600, A10
Mollière P., Wardenier J. P., van Boekel R., Henning T., Molaverdikhani K., Snellen I. A. G., 2019, A&A, 627, A67
Mollière P. et al., 2020, A&A, 640, A131
Mollière P. et al., 2022, ApJ, 934, 74
Mordasini C., van Boekel R., Mollière P., Henning T., Benneke B., 2016, ApJ, 832, 41
Morley C. V., Fortney J. J., Marley M. S., Visscher C., Saumon D., Leggett S. K., 2012, ApJ, 756, 172
Morley C. V., Marley M. S., Fortney J. J., Lupu R., Saumon D., Greene T., Lodders K., 2014, ApJ, 787, 78
Moses J. I., Madhusudhan N., Visscher C., Freedman R. S., 2013, ApJ, 763, 25
NASA Exoplanet Archive, 2022, NASA Exoplanet Archive, NASA Exoplanet Science Institute, available at https://exoplanetarchive.ipac.caltech.edu/docs/counts_detail.html
Nielsen E. L. et al., 2016, AJ, 152, 175
Nielsen E. L. et al., 2019, AJ, 158, 13
Nowak M. et al., 2020, A&A, 642, L2
Oreshenko M. et al., 2020, AJ, 159, 6
Peach G., 2017, Open Astron., 20, 516
Peach G., Yurchenko S., Chubb K., Baraffe I., Phillips M. W., Tremblin P., 2020, Contrib. Astron. Obs. Skalnaté Pleso, 50, 193
Pezzato J. et al., 2019, Proc SPIE Conf. Ser. Vol. 11117, Techniques and Instrumentation for Detection of Exoplanets IX, SPIE, Bellingham, p. 12
Phillips M. W. et al., 2020, A&A, 637, A38
Piette A. A. A., Madhusudhan N., 2020, MNRAS, 497, 5136
Pluriel W. et al., 2020, AJ, 160, 112
Pollack J. B., Hubickyj O., Bodenheimer P., Lissauer J. J., Podolak M., Greenzweig Y., 1996, Icarus, 124, 62
Polyansky O. L., Kyuberis A. A., Zobov N. F., Tennyson J., Yurchenko S. N., Lodi L., 2018, MNRAS, 480, 2597
Rajan A. et al., 2017, AJ, 154, 10
Rayner J. T., Toomey D. W., Onaka P. M., Denault A. J., Stahlberger W. E., Vacca W. D., Cushing M. C., Wang S., 2003, PASP, 115, 362
Rocchetto M., Waldmann I. P., Venot O., Lagage P. O., Tinetti G., 2016, ApJ, 833, 120
Rothman L. S. et al., 1987, Appl. Opt., 26, 4058
Rothman L. S. et al., 2010, J. Quant. Spec. Radiat. Transf., 111, 2139
Salinas V. N. et al., 2016, A&A, 591, A122
Samland M. et al., 2017, A&A, 603, A57
Saumon D., Marley M. S., 2008, ApJ, 689, 1327
Saumon D., Chabrier G., van Horn H. M., 1995, ApJS, 99, 713
Saumon D., Marley M. S., Cushing M. C., Leggett S. K., Roellig T. L., Lodders K., Freedman R. S., 2006, ApJ, 647, 552
Saumon D., Marley M. S., Abel M., Frommhold L., Freedman R. S., 2012, ApJ, 750, 74
Schneider A. C. et al., 2015, ApJ, 804, 92
Skaf N. et al., 2020, AJ, 160, 109
Skrutskie M. F. et al., 2006, AJ, 131, 1163
Snellen I. A. G., Brown A. G. A., 2018, Nat. Astron., 2, 883
Tennyson J., Yurchenko S. N., 2012, MNRAS, 425, 21
Testi L., 2009, A&A, 503, 639
Tinetti G., Encrenaz T., Coustenis A., 2013, A&AR, 21, 63
Tremaine S. et al., 2002, ApJ, 574, 740
Tremblin P. et al., 2017, ApJ, 850, 46
Tremblin P. et al., 2019, ApJ, 876, 144
Tremblin P., Amundsen D. S., Chabrier G., Baraffe I., Drummond B., Hinkley S., Mourier P., Venot O., 2016, ApJ, 817, L19
Trotta R., 2008, Contemp. Phys., 49, 71
Tsiaras A. et al., 2016, ApJ, 820, 99
Tsiaras A. et al., 2018, AJ, 155, 156
Tsiaras A., Waldmann I. P., Tinetti G., Tennyson J., Yurchenko S. N., 2019, Nat. Astron., 3, 1086
Udry S. et al., 2014, preprint (arXiv:1412.1048)
Vigan A. et al., 2017, A&A, 603, A3
Vigan A. et al., 2021, A&A, 651, A72
Waldmann I. P., Tinetti G., Rocchetto M., Barton E. J., Yurchenko S. N., Tennyson J., 2015a, ApJ, 802, 107
Waldmann I. P., Rocchetto M., Tinetti G., Barton E. J., Yurchenko S. N., Tennyson J., 2015b, ApJ, 813, 13
Wang J. et al., 2020, AJ, 160, 150
Xuan J. W. et al., 2022, ApJ, 937, 54
Yurchenko S. N., Amundsen D. S., Tennyson J., Waldmann I. P., 2017, A&A, 605, A95
Yurchenko S. N., Mellor T. M., Freedman R. S., Tennyson J., 2020, MNRAS, 496, 5282
Zalesky J. A., Line M. R., Schneider A. C., Patience J., 2019, ApJ, 877, 24
Zalesky J. A., Saboi K., Line M. R., Zhang Z., Schneider A. C., Liu M. C., Best W. M. J., Marley M. S., 2022, ApJ, 936, 44
Zhang M., Chachan Y., Kempton E. M. R., Knutson H., 2020, ApJ, 899, 27
Zhou Y., Bowler B. P., Morley C. V., Apai D., Kataria T., Bryan M. L., Benneke B., 2020, AJ, 160, 77
Zingales T., Waldmann I. P., 2018, ApJ, 156, 156

APPENDIX A: APPENDIX FIGURES

Here, we present our posterior distributions for the ATMO 2020 grid modelling fits of properties of GJ 570D and 51 Eri b. We also present the individual posterior distributions for our retrieval analysis of different data sets.

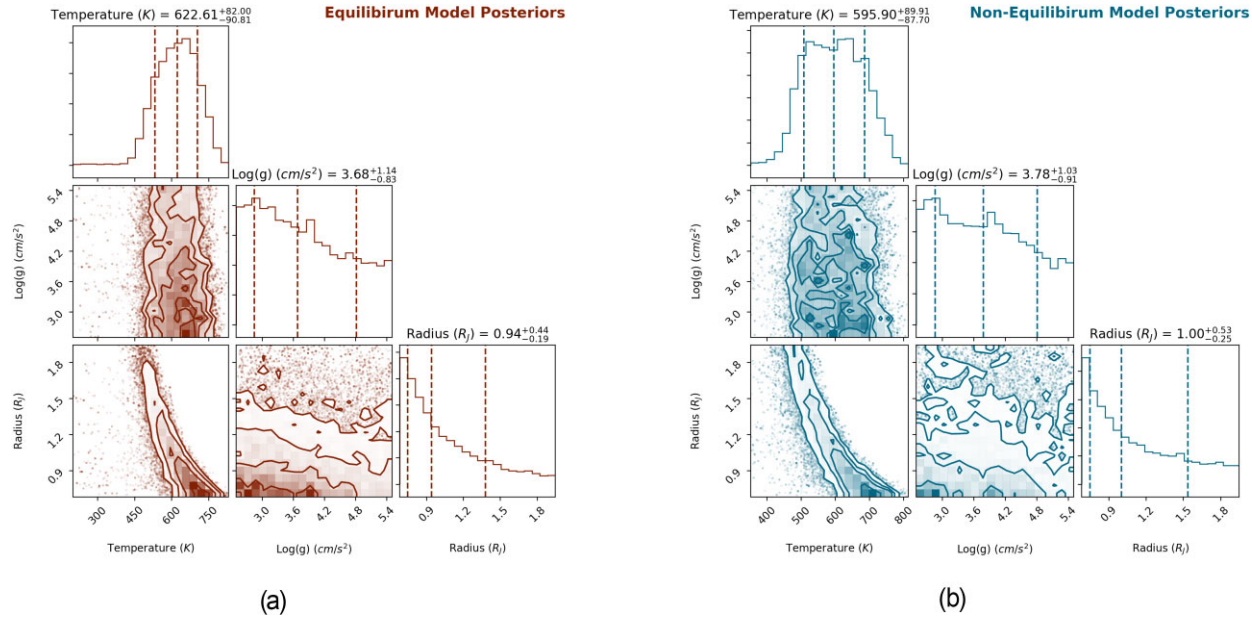


Figure A1. ATMO posteriors plots for 51 Eri b. Left: Equilibrium chemistry model. Right: Non-equilibrium chemistry model.

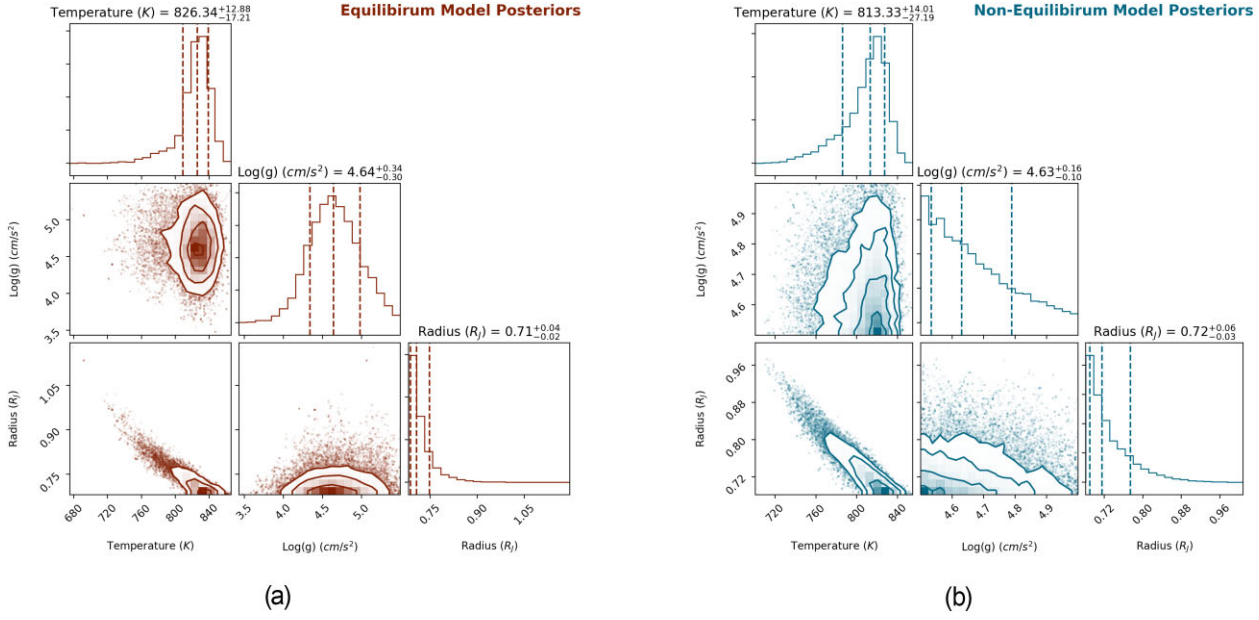


Figure A2. ATMO posteriors plots for GJ 570D. Left: Equilibrium chemistry model. Right: Non-equilibrium chemistry model.

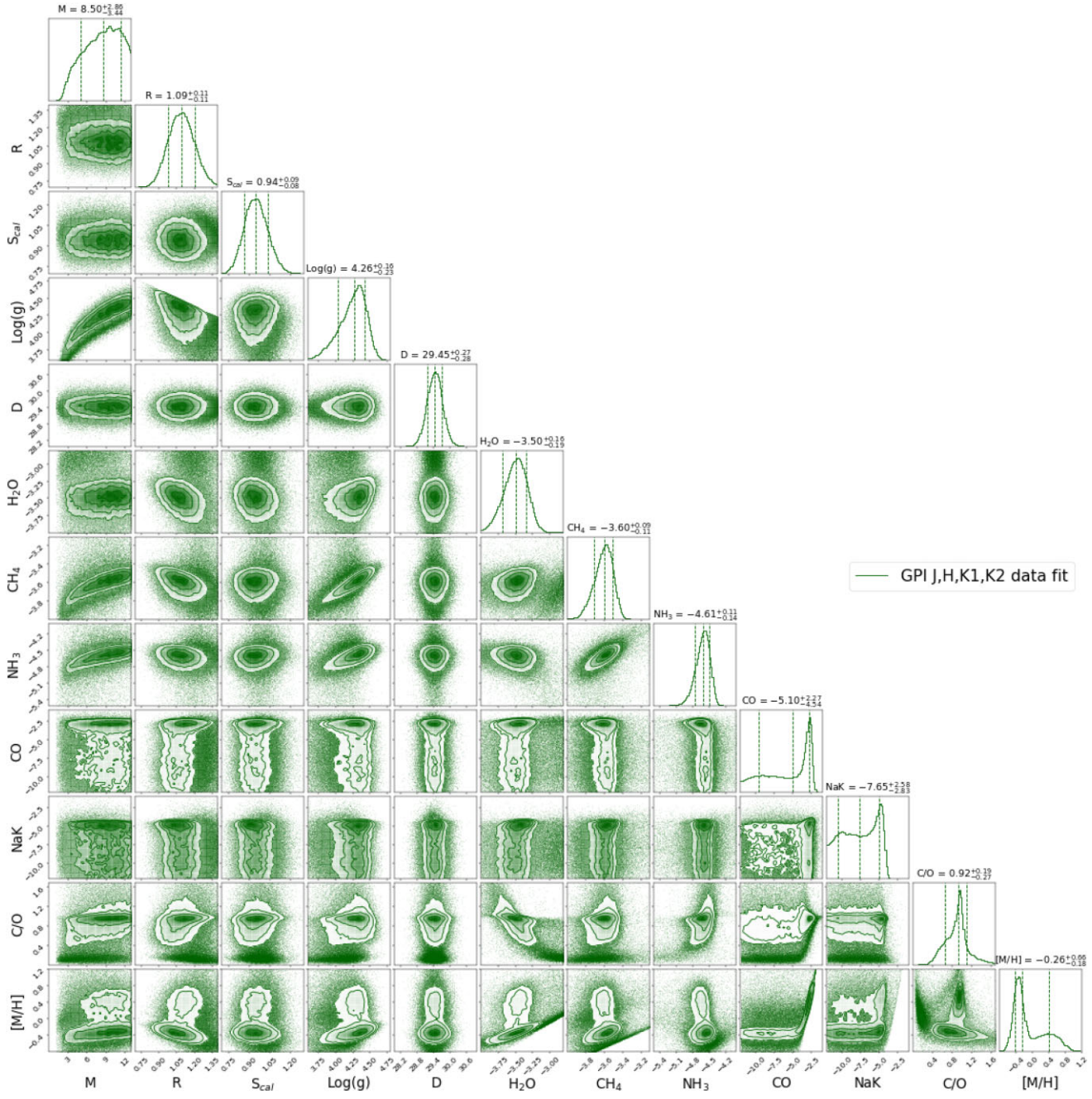


Figure A3. 51 Eri b posteriors for GPI *J*-, *H*-, and *K*-band data. Log (g), C/O and [M/H] posteriors are inferred parameters, while all the other parameters are sample as part of the retrieval.

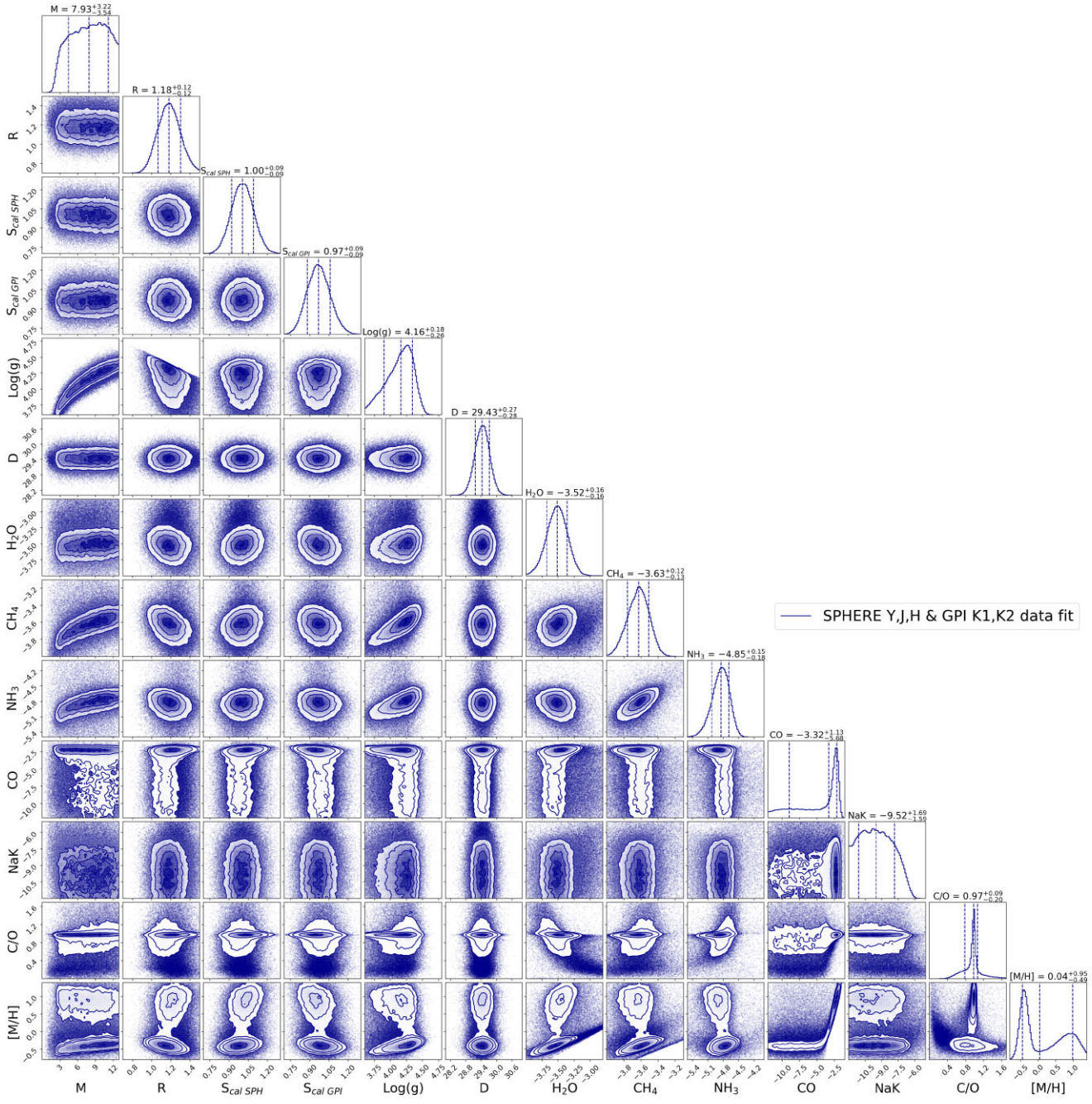


Figure A4. 51 Eri b posteriors for SPEHRE *Y*-, *J*-, *H*-, and GPI *K*-band data. Log(g), C/O and [M/H] posteriors are inferred parameters, while all the other parameters are sample as part of the retrieval.

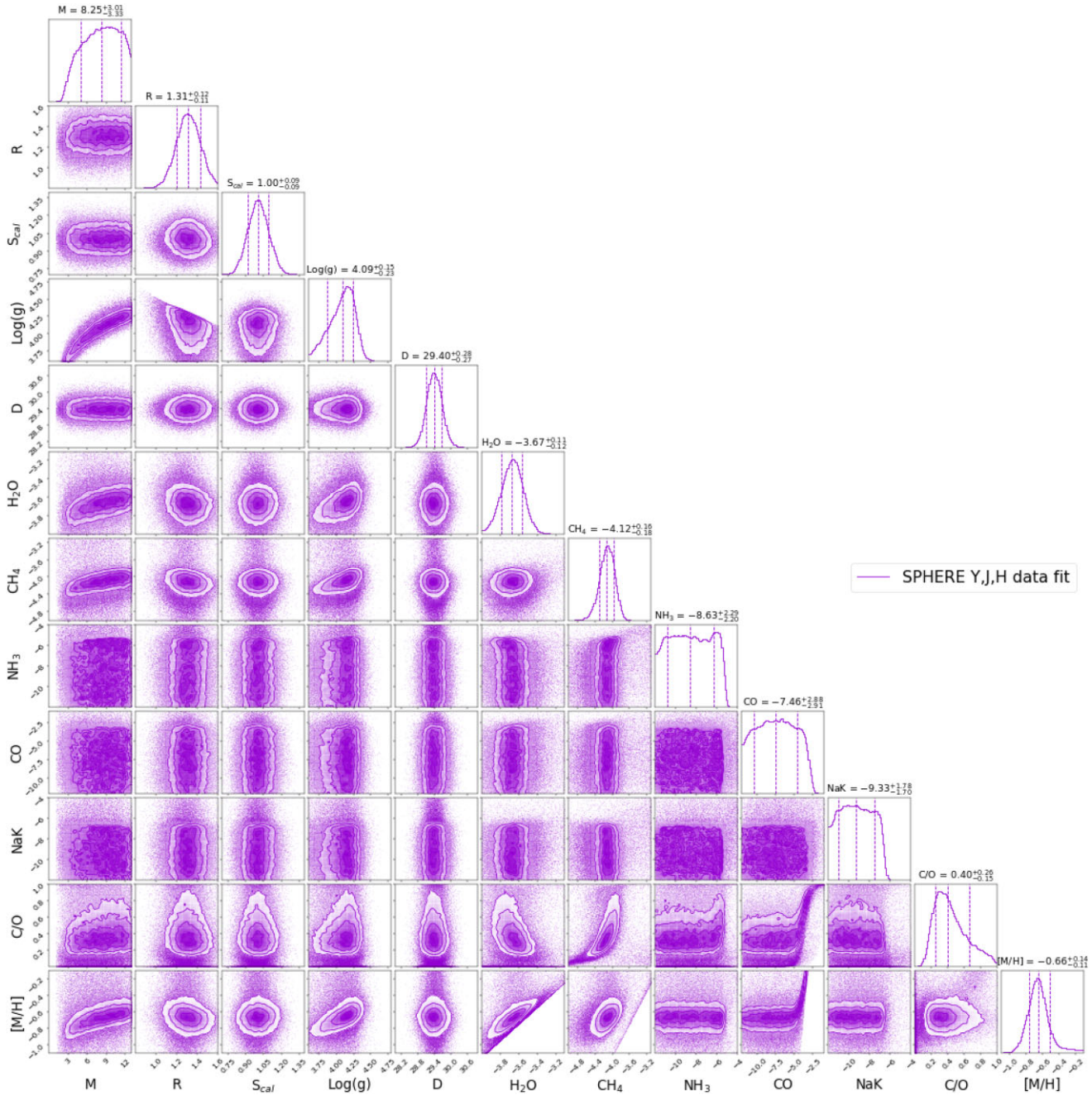


Figure A5. 51 Eri b posteriors for *Y*-, *J*- and *H*-band data. Log(g), C/O and [M/H] posteriors are inferred parameters, while all the other parameters are samples as part of the retrieval.

This paper has been typeset from a TEX/LATEX file prepared by the author.