

# Retrieval study of cool, directly imaged exoplanet 51 Eri b

Niall Whiteford<sup>1,2,3\*</sup>, Alistair Glasse<sup>2,3,4</sup>, Katy L. Chubb<sup>5</sup>, Daniel Kitzmann<sup>6</sup>,  
 Shrishmoy Ray<sup>7</sup>, Mark W. Phillips<sup>14</sup>, Beth A. Biller<sup>2,3</sup>, Paul I. Palmer<sup>9,3</sup>,  
 Ken Rice<sup>1,2</sup>, Ingo P. Waldmann<sup>8</sup>, Quentin Changeat<sup>12</sup>, Nour Skaf<sup>10,11</sup>,  
 Jason Wang<sup>15</sup>, Billy Edwards<sup>16</sup>, Ahmed Al-Refaie<sup>8</sup>

<sup>1</sup>Department of Astrophysics, American Museum of Natural History, Central Park West at 79th Street, NY 10024, USA

<sup>2</sup>Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh, EH9 3HJ, UK

<sup>3</sup>Centre for Exoplanet Science, University of Edinburgh, UK

<sup>4</sup>UK Astronomy Technology Centre, Royal Observatory, Edinburgh, Blackford Hill, Edinburgh EH9 3HJ, UK

<sup>5</sup>Centre for Exoplanet Science, University of St Andrews, North Haugh, St Andrews, UK

<sup>6</sup>Center for Space and Habitability, University of Bern, Gesellschaftsstrasse 6, Bern, 3012, Switzerland

<sup>7</sup>Astrophysics Group, University of Exeter, EX4 4QL, Exeter, UK

<sup>8</sup>Department of Physics and Astronomy, University College London, Gower Street, WC1E 6BT, UK

<sup>9</sup>School of GeoSciences, University of Edinburgh, Edinburgh, UK

<sup>10</sup>LESIA, Observatoire de Paris, Université PSL, CNRS, Sorbonne Université, Université de Paris, 5 place Jules Janssen, 92195 Meudon, France

<sup>11</sup>Subaru Telescope, National Astronomical Observatory of Japan, 650 North A'Ohoku Place, Hilo, HI 96720, USA

<sup>12</sup>European Space Agency, Space Telescope Science Institute, Baltimore, MD, USA

<sup>13</sup>AIM, CEA, CNRS, Université Paris-Saclay, Université de Paris, F-91191 Gif-sur-Yvette, France

<sup>14</sup>Institute for Astronomy, University of Hawaii, 2680 Woodlawn Drive, Honolulu, HI 96822, USA

<sup>15</sup>Center for Interdisciplinary Exploration and Research in Astrophysics (CIERA) and Department of Physics and Astronomy, Northwestern University, Evanston, IL 60208, USA

17 February 2023

<sup>16</sup>SRON Netherlands Institute for Space Research, Sorbonnelaan 2, 3584 CA, Utrecht, Netherlands

## 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  $2.7\sigma$  confidence, the first indication of ammonia in an 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.

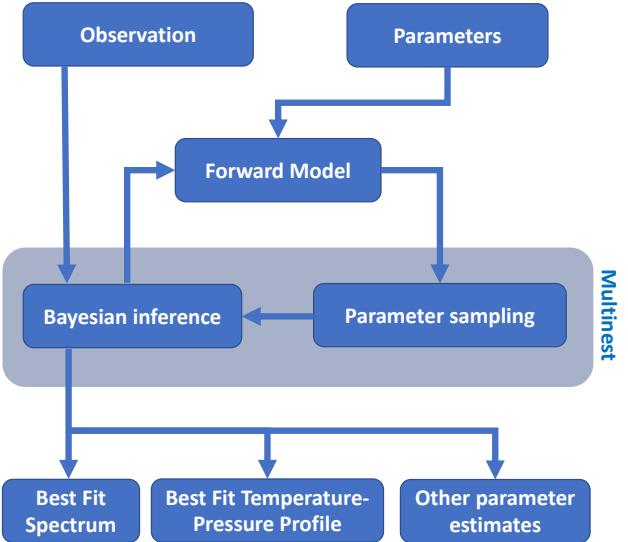
**Key words:** Data analysis – Brown dwarfs – Atmospheres – Gaseous planets

## 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 et al. 2017), have made it possible to start the characterisation and classification effort of directly imaged exoplanet demographics (Nielsen et al. 2019; Vigan et al. 2020). A summary of direct imaging spectroscopy is covered extensively in Biller & Bonnefoy (2018).

The development of extrasolar planetary spectroscopy (see Tinetti et al. 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 modeling techniques (outlined in Figure 1) to exoplanetary spectra (see Line et al. 2013 for a review of early exoplanetary retrieval codes). There are now a variety of retrieval codes developed for exoplanet atmospheric characterisation, examples include Nemesis (Irwin et al. 2008), Chimera (Line et al. 2013), BART (Harrington et al. 2016), SCARLET (Benneke 2015), POSEIDON (MacDonald & Madhusudhan 2017), Brewster (Burningham et al. 2017, 2021), HyDRA (Gandhi & Madhusudhan 2018), petitRADTRANS (Mollière et al. 2019; Nowak, M. et al. 2020; Mollière et al. 2020), Platon II (Zhang et al. 2020), Helios-R2 (Kitzmann et al. 2020), APOLLO (Howe et al. 2022) and TauREx3 (Al-Refaei 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 et al. 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),  $\beta$  Pic b (Gravity Collaboration et al. 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



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

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 James Webb Space Telescope (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  $\text{H}_2\text{O}$  and  $\text{CH}_4$  absorption.

The importance of cloud modelling for directly imaged exoplanets and brown dwarfs has been well explored and debated (Marley et al. 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) and Tremblin et al. (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<sub>4</sub> 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).

### 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 et al. 2007), and was among the first T dwarf companions to be discovered by Burgasser et al. (2004, 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  $1525 \pm 25$  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 3m 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 et al. (2004), with the spectrum spanning 0.65 to 2.56  $\mu\text{m}$  at an average spectral resolving power of 120. Using the SpeX Prism Library<sup>1</sup> data analysis toolkit (SPLAT<sup>2</sup>) (see Burgasser & Splat Development Team (2017) for details), we flux calibrated the data using photometry from the 2MASS survey<sup>3</sup> (Skrutskie et al. 2006).

The spectra shown in Figure 2a have then been calibrated using  $J$  ( $15.324 \pm 0.05$  mag),  $H$  ( $15.268 \pm 0.09$  mag) and  $K$ -band ( $15.242 \pm 0.16$  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 ( $\sim 0.5''$ ,  $\sim 13$  AU) of any directly imaged exoplanet. It orbits a young F0-type host, with age estimates of  $20 \pm 6$  Myrs from Macintosh et al. (2014) and  $26 \pm 3$  Myrs from Nielsen et al. (2016). With a spectral type of T6.5 $\pm$ 1.5 (Rajan et al. 2017), 51 Eri b is notably the latest spectral type planet yet imaged.

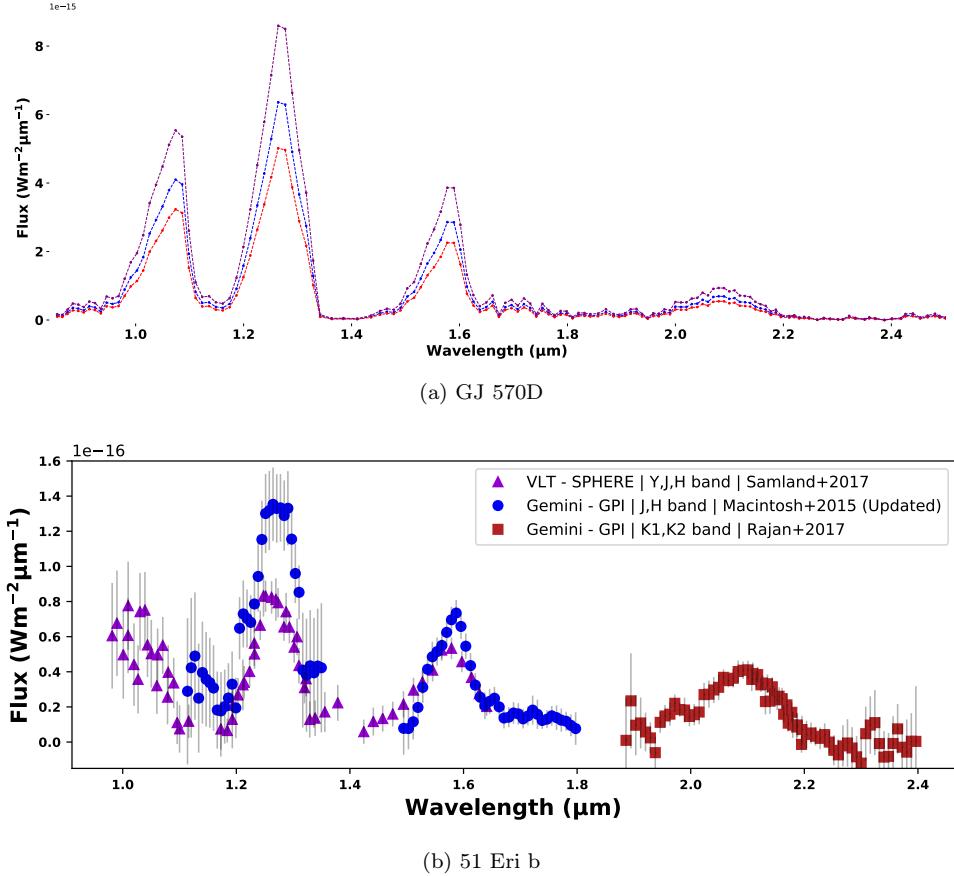
Exhibiting methane absorption (a first for directly-imaged exoplanets) with its lower effective temperature ( $\sim 700$ K) and low mass ( $< 10$  M<sub>Jup</sub>), 51 Eri b defined a new category of directly-imaged exoplanets. Further, its SED indicates that the L/T transition occurs at lower temperatures 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) couldn't 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. We outline previous model fitting results from previous studies in Table 5. Neither Samland et al. (2017) or Rajan

<sup>1</sup> SpeX Prism Library: [http://pono.ucsd.edu/~adam/brown\\_dwarfs/speX\\_prism/library.html](http://pono.ucsd.edu/~adam/brown_dwarfs/speX_prism/library.html)

<sup>2</sup> SPLAT: [http://pono.ucsd.edu/~adam/brown\\_dwarfs/splat/](http://pono.ucsd.edu/~adam/brown_dwarfs/splat/)

<sup>3</sup> 2MASS Survey Archive: <https://irsa.ipac.caltech.edu/Missions/2mass.html>



**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  $K1$  and  $K2$  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.

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  $J$ ,  $H$ ,  $K1$  and  $K2$  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 Figure 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  $K1$  and  $K2$  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 Figure 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 fit 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 characterisations in the field of direct imaging.

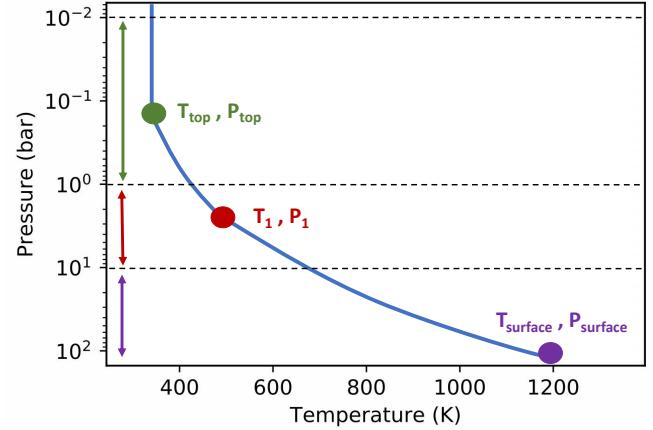
### 3.2 TauREx3

TauREx3 (Tau Retrieval of Exoplanets) is a publicly available<sup>4</sup> 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. Figure 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 database (Chubb et al. 2020)<sup>5</sup>. The line lists used for this database 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<sub>2</sub>O (Polyansky et al. 2018), CO (Li et al. 2015), CO<sub>2</sub> (Yurchenko et al. 2020), CH<sub>4</sub> (Yurchenko et al. 2017), VO (McKemmish et al. 2016), H<sub>2</sub>S (Azzam et al. 2016) and NH<sub>3</sub> (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 et al. (2016) and Allard et al. (2019). All other line data for these atomic species are taken from either the NIST (Kramida et al. 2013) or Kurucz (Kurucz & Bell 1995) database (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 database, 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  $\mu\text{m}$  wavelength region. For a more detailed discussion of TauREx's line list library see Chubb et al. (2020). During

<sup>4</sup> TauREx3: [https://github.com/ucl-exoplanets/TauREx3\\_public](https://github.com/ucl-exoplanets/TauREx3_public)



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

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 log-likelihood.

#### 3.2.2 Temperature-pressure profiles

To accurately model directly-imaged emission spectroscopy, an appropriate temperature-pressure parameterisation 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 Figure 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,  $T_{\text{top}}$ , and top of atmosphere pressure,  $P_{\text{top}}$  (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  $P_{\text{surf}}$  (set at 500 bar for this study) and temperature  $T_{\text{surf}}$ . 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 parameterisation employed in Lavie et al. (2017). This originated from a reduced version of equation 126 in Heng et al. (2014):

$$T^4 = \frac{T_{\text{int}}^4}{4} \left( \frac{8}{3} + \tilde{m} \kappa_0 \right) \quad (1)$$

where  $T_{\text{int}}$  is the internal temperature and  $\kappa_0$  the constant component of the infrared opacity.  $\tilde{m}$  is column density determined via  $P_0 = \tilde{m} \cdot g$  with  $g$  being the surface gravity at the bottom of our model atmosphere (500 bar). This simpler profile parameterisation only has two free parameters within our retrievals:  $\kappa_0$  and  $T_{\text{int}}$ .

### 3.2.3 Effective Temperature Calculation

We have included the effective temperature  $T_{\text{eff}}$  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  $\mu\text{m}$  (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% of models ran.

### 3.2.4 Bayesian Analysis

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

$$P(\theta | x, \mathcal{M}) = \frac{P(x | \theta, \mathcal{M}) P(\theta, \mathcal{M})}{P(x | \mathcal{M})}, \quad (2)$$

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

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

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

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

where  $\sigma_{\lambda}$  is the measured error for the  $\lambda$ th flux and  $b$  is a tolerance factor which is included as a free parameter in the retrieval analysis ([Tremaine et al. 2002](#); [Hogg et al. 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 et al. 2009, 2013](#)) via PyMultinest ([Buchner et al. 2014](#)). NS derives the Bayesian Evidence given by:

$$E = \int P(\theta | \mathcal{M}) P(x | \theta, \mathcal{M}) d\theta, \quad (5)$$

**Table 1.** Interpretation of the Bayes ratio outlined in [Kass & Raftery 1995](#)

$\log(b)$	Interpretation
0 - 0.5	<i>No Evidence</i>
0.5 - 1	<i>Some Evidence</i>
1 - 2	<i>Strong Evidence</i>
$>2$	<i>Decisive</i>

where  $E = P(x | \mathcal{M})$  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:

$$\log(B) = \Delta \log(Ev) = \log(Ev2) - \log(Ev1), \quad (6)$$

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.

### 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,  $\log(\text{abundance}) = 1.0 - 1.0 \times 10^{-12}$ ). The default values can be manually over-ridden, 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 Figure 4. We didn't 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.

### 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

**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 $R_{Jup}$	$1.3 \pm 0.15 R_{Jup}$ <sup>1,2</sup>
Mass	Uniform	13 - 80 $M_{Jup}$	$1 - 13 M_{Jup}$ <sup>3</sup>
Distance	Gaussian	$5.8819 \pm 0.0029$ pc <sup>4</sup>	$29.4 \pm 0.3$ pc <sup>5</sup>
$S_{cal}$	Gaussian	1.0, 0.1 STD	1.0, 0.1 STD
$10^b$	Uniform	$0.01 \times \min(\sigma_\lambda^2) \leq 10^b \leq 100 \times \min(\sigma_\lambda^2)$	-
$T_{surf}$	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
$\kappa_0$	Log-Uniform	-	$1e-15 - 1e1$
$\tau_0$	Uniform	-	0.01 - 100
$\alpha$	Log-Uniform	-	-10 - 10
$P_{top}$	Log-Uniform	-	$1e2 - 5e7$ bar
$P_{bottom}$	Log-Uniform	-	$1e2 - 5e7$ bar
$C_{frac}$	Uniform	-	0 - 1

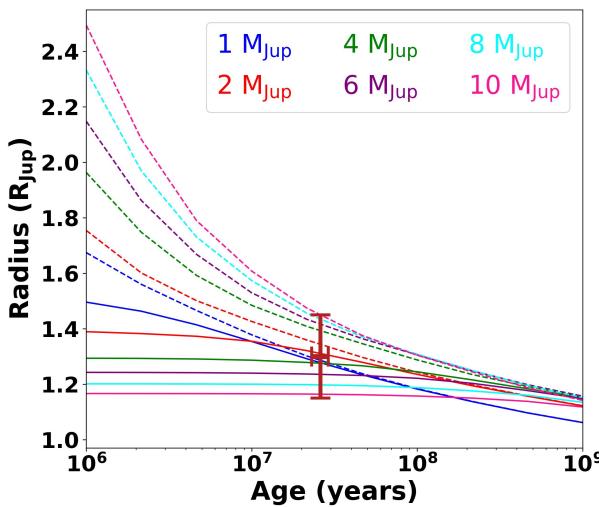
<sup>1</sup> Note: Gaussian prior.

<sup>2</sup> Note: We use a Gaussian prior informed using evolutionary models from [Fortney et al. \(2008\)](#) combined with the age presented in [Rajan et al. \(2017\)](#).

<sup>3</sup> Note: We make the assumption of a planetary mass object.

<sup>4</sup> Note: GJ 570D distance comes from Gaia Archive: <https://gea.esac.esa.int/archive/>.

<sup>5</sup> Note: The 51 Eri b distance comes from [Macintosh et al. \(2015\)](#).



**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.

model and added an inverse square law scaling for the exo-planet or brown dwarf emission:

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

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

$$Obs_{\text{cal}} = \frac{1}{S_{\text{cal}}}, \quad (8)$$

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

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

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  $\text{H}_2\text{O}$  and  $\text{CH}_4$ . Therefore, this ratio should really be considered as a  $\text{CH}_4$  to  $\text{H}_2\text{O}$  ratio. In the case of our T dwarf analysis the inferred C/O ratio is calculated via:

$$\text{C/O} = \frac{\chi_{\text{CH}_4} + \chi_{\text{CO}}}{\chi_{\text{H}_2\text{O}} + \chi_{\text{CO}}}, \quad (10)$$

where  $\chi$  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:

$$M_{\text{object}} = \sum_{\text{molecules}} \frac{f_m \cdot n}{\chi_{\text{H}_2} \cdot 2}, \quad (11)$$

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  $\text{CO}_2$   $n=3$ ),  $\chi_{\text{H}_2}$  is the gas fraction of neutral hydrogen. Therefore  $[\text{M}/\text{H}]$  is determined via:

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

where  $M_{\text{solar}}$  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 parameterisations from [\(Burningham et al. 2017\)](#) which are illustrated in Figure 5. The optical depth of the power law deck cloud is set by  $P_{\text{top}}$  pressure where the cloud becomes optically thick ( $\tau=1$ ). The cloud opacity drops off above this pressure via  $d\tau/dP \propto \exp(\Delta P/\Phi)$  where  $\Delta P$  is the height above and below the  $P_{\text{top}}$  pressure and  $\phi$  is:

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

The decay is parameterised by  $\Delta \log P$  and the cloud is made non grey with the optical depth following  $\tau \propto \lambda^\alpha$ . Therefore, in total, the power law slab is retrieved via 3 parameters:  $P_{\text{top}}$ ,  $\Delta \log P$  and  $\alpha$ . The total optical depth of the slab cloud is determined via:

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

where  $\lambda_0=1\mu\text{m}$ .  $\tau_0$  and  $\alpha$  are the two retrievable components of this cloud prescription as well as the top pressure boundary of  $P_{\text{top}}$  and the bottom pressure boundary of the cloud  $P_{\text{bottom}}$ .  $\tau_{\text{cloud}}$  is distributed throughout the layers in the cloud slab pressure boundaries, weighted by  $d\tau/dP \propto$

$dP$  where  $dP$  is relative to  $P_{\text{bottom}}$ . 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:  $\tau_0$ ,  $\alpha$ ,  $P_{\text{top}}$  and  $P_{\text{bottom}}$ .

These are flexible but simplistic cloud approach that lack the rigour of the more physically motivated approaches included in [Mollière et al. \(2020\)](#) and [Burningham et al. \(2021\)](#). For example, these parameterisations 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:

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

where  $F_{\text{Tot}}$  is the total flux,  $F_{\text{Clear}}$  is the flux from regions without clouds,  $F_{\text{Cloudy}}$  is the flux from regions with clouds and  $C_{\text{frac}}$  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 parameterisations.

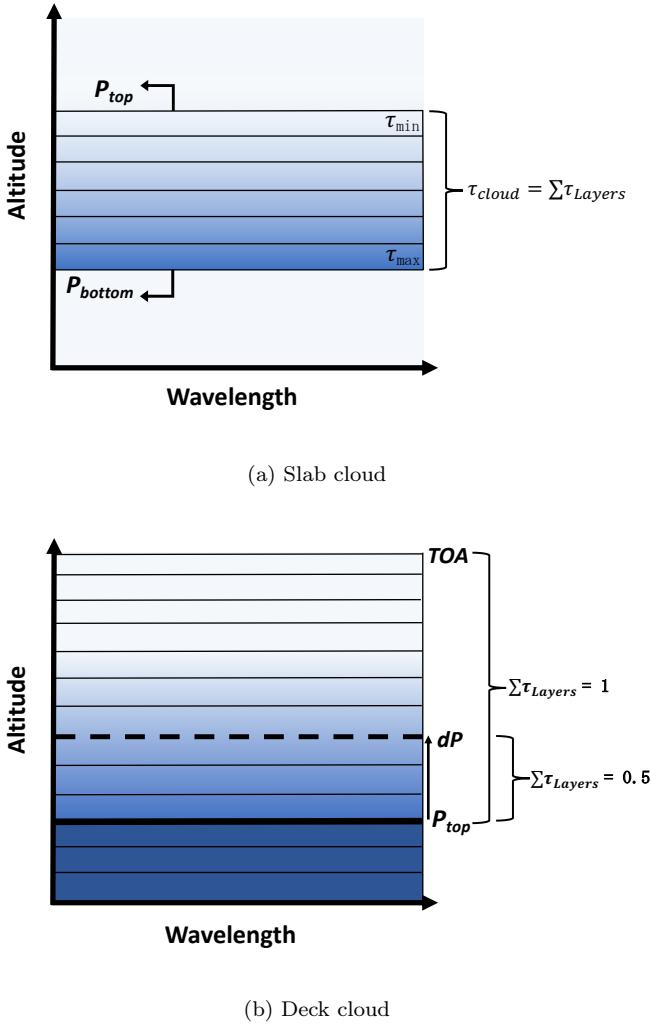
### 3.2.9 Forward model setup

The atmospheric forward model is described in [Al-Refaei et al. \(2021\)](#). We assume a hydrogen dominated atmosphere with a  $\text{H}_2$  and He mixing ratio  $\text{He}/\text{H}_2 = 0.17567$ . 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 3.2.4) 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 3.2.4). In our GJ 570D data it is apparent that that  $K$  band errors are much smaller than the  $J$  band data (see Figure 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 de-weighting 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  $\text{H}_2-\text{H}_2$  and  $\text{H}_2-\text{He}$  ([Abel et al. 2011; Fletcher et al. 2018; Abel et al. 2012](#)) is included.

In the case of the 51 Eri b data analysis we use multiple scaling factors  $S_{\text{cal}}$  to account for the inclusion of observations from different instruments. This is employed in the case of the SPHERE  $Y$ ,  $J$  and  $H$  data ( $S_{\text{cal, SPH}}$ ) being combined with the GPI  $K1$  and  $K2$  band data ( $S_{\text{cal, GPI}}$ ). When



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

employing data from a single instrument we simply use one scaling  $S_{cal}$  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 minimisation 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 et al. (1995).

This grid includes solar metallicity atmosphere models spanning  $T_{\text{eff}} = 200 - 3000$  K and  $\log(g) = 2.5 - 5.5$  ( $g$  in units of  $\text{cm/s}^2$ ), with steps of 100 K for  $T_{\text{eff}} > 600$  K, 50 K for  $T_{\text{eff}} < 600$  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<sup>6</sup>.

#### 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  $2.5 \text{ cm/s}^2$  to  $5.5 \text{ cm/s}^2$  for models assuming chemical equilibrium, and temperature ranging from 350 K to 1800 K and  $\log(g)$  from  $3.0 \text{ cm/s}^2$  to  $5.5 \text{ cm/s}^2$  for models assuming non-equilibrium chemistry due to vertical mixing. The radius was constrained between  $0.07 R_{\odot}$  ( $\sim 0.7 R_{\text{Jup}}$ ) and  $0.2 R_{\odot}$  ( $\sim 2 R_{\text{Jup}}$ ) 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  $1\sigma$ .

<sup>6</sup> ATMO 2020: <http://opendata.erc-atmo.eu>

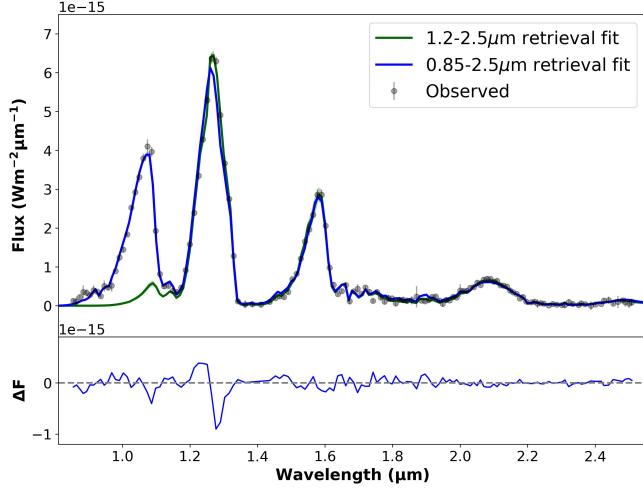


Figure 6. GJ 570D retrieval spectral fit

## 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.

### 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 $\mu\text{m}$  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<sub>Jup</sub>, 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 fit 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

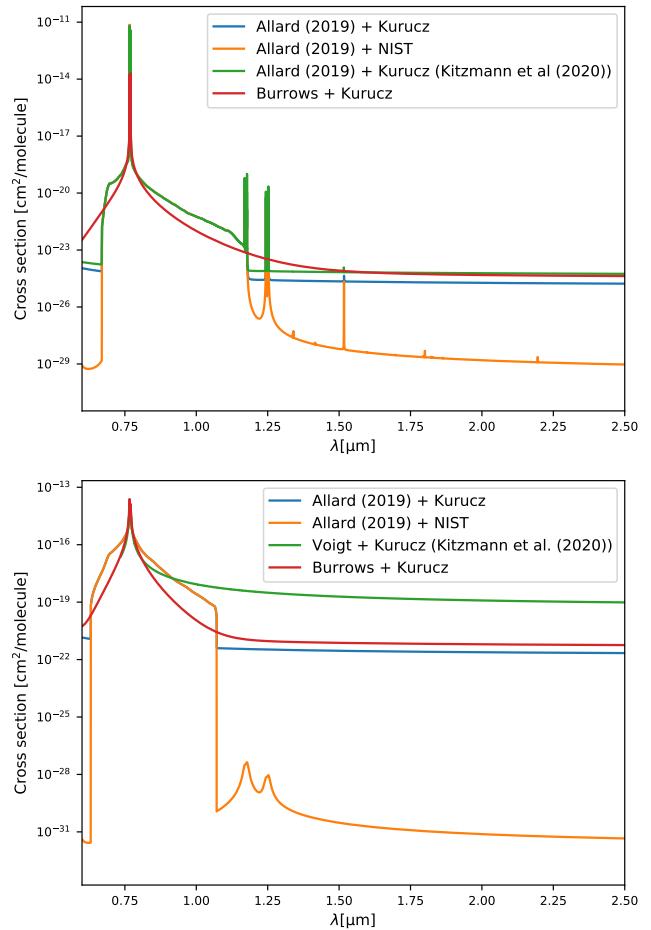


Figure 7. A comparison of different methods used to compute the resonance doublet and non-resonance lines of Na and K. The first panel gives cross sections for K computed at  $T = 1000$  K,  $P = 0.1$  bar, and the second panel the cross sections for K computed at  $T = 600$  K,  $P = 10$  bar. The cross 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, 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.

fitting the whole 0.85 - 2.5  $\mu\text{m}$  spectrum. The resonance doublets of K and Na are at  $\sim 0.77 \mu\text{m}$  and  $\sim 0.59 \mu\text{m}$  respectively. We encountered examples when the bias issue would not be present when fitting only 0.85 - 1.2  $\mu\text{m}$  ( $\sim 0.77 \mu\text{m}$  K /  $\sim 0.59 \mu\text{m}$  Na resonance doublet impacted region), or 1.2 - 2.5  $\mu\text{m}$  (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 Figure 7. 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, 2019). The non-resonance lines from both the NIST (Kramida et al. 2013) or Kurucz (Kurucz & Bell

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

	Mass (M <sub>Jup</sub> )	Radius (R <sub>Jup</sub> )	log(g) (cm/s <sup>2</sup> )	T <sub>eff</sub> (K)	C/O	[M/H]
This work (TauREX3)	48.00 <sup>+13.03</sup> <sub>-11.87</sub>	1.17 <sup>+0.08</sup> <sub>-0.08</sub>	4.93 <sup>+0.11</sup> <sub>-0.12</sub>	722 <sup>+23</sup> <sub>-26</sub>	0.87 <sup>+0.08</sup> <sub>-0.07</sub>	-0.19 <sup>+0.05</sup> <sub>-0.03</sub>
This work (ATMO 2020 - EC FM)	-	0.71 <sup>+0.04</sup> <sub>-0.02</sub>	4.64 <sup>+0.34</sup> <sub>-0.30</sub>	826.34 <sup>+12.88</sup> <sub>-17.21</sub>	-	-
This work (ATMO 2020 - NEC FM)	-	0.72 <sup>+0.06</sup> <sub>-0.03</sub>	4.63 <sup>+0.16</sup> <sub>-0.10</sub>	813.33 <sup>+14.01</sup> <sub>-27.19</sub>	-	-
Kitzmann et al. 2020 (FCR)	53 <sup>+24</sup> <sub>-20</sub>	1.13 <sup>+0.05</sup> <sub>-0.06</sub>	5.01 <sup>+0.13</sup> <sub>-0.19</sub>	703 <sup>+17</sup> <sub>-30</sub>	1.11 <sup>+0.09</sup> <sub>-0.09</sub>	-0.13 <sup>+0.06</sup> <sub>-0.08</sub>
Kitzmann et al. 2020 (ECR)	17 <sup>+3.8</sup> <sub>-3.0</sub>	1.00 <sup>+0.10</sup> <sub>-0.09</sub>	4.61 <sup>+0.08</sup> <sub>-0.08</sub>	730 <sup>+18</sup> <sub>-17</sub>	0.83 <sup>+0.09</sup> <sub>-0.08</sub>	-0.15 <sup>+0.05</sup> <sub>-0.04</sub>
Burningham et al. 2017 (FCR)	19.80 <sup>+28.60</sup> <sub>-15.96</sub>	0.96 <sup>+0.80</sup> <sub>-0.11</sub>	4.73 <sup>+0.31</sup> <sub>-1.17</sub>	752.25 <sup>+35.51</sup> <sub>-82.10</sub>	-	-
Line et al. 2015 (FCR)	30.90 <sup>+26.64</sup> <sub>-15.76</sub>	1.14 <sup>+0.10</sup> <sub>-0.09</sub>	4.76 <sup>+0.27</sup> <sub>-0.28</sub>	714.11 <sup>+20.19</sup> <sub>-23.15</sub>	1.09 <sup>+0.16</sup> <sub>-0.14</sub>	-0.25 <sup>+0.13</sup> <sub>-0.12</sub>
Oreshenko et al. 2020: Sonora (SML)	-	-	4.93 <sup>+0.38</sup> <sub>-0.55</sub>	808 <sup>+43</sup> <sub>-27</sub>	-	-
Oreshenko et al. 2020: AMES-cond (SML)	-	-	5.27 <sup>+0.43</sup> <sub>-0.67</sub>	878 <sup>+23</sup> <sub>-78</sub>	-	-
Oreshenko et al. 2020: HELIOS (SML)	-	-	5.08 <sup>+0.62</sup> <sub>-0.68</sub>	800 <sup>+14</sup> <sub>-100</sub>	-	-
Saumon et al. 2017 (FM)	-	0.94 <sup>+0.04</sup> <sub>-0.04</sub>	4.67 <sup>+0.04</sup> <sub>-0.04</sub>	769 <sup>+14</sup> <sub>-13</sub>	-	-
Filippazzo et al. 2015 (EM)	37.28 <sup>+24.05</sup> <sub>-24.05</sub>	0.94 <sup>+0.16</sup> <sub>-0.16</sub>	4.90 <sup>+0.50</sup> <sub>-0.50</sub>	759 <sup>+63</sup> <sub>-63</sub>	-	-
Testi 2009 (FM)	-	-	5.0	900	-	-
Del Burgo et al. 2009 (FM)	-	-	4.5 <sup>+0.5</sup> <sub>-0.5</sub>	948 <sup>+58</sup> <sub>-58</sub>	-	-
Saumon et al. 2006 (EM, FM)	42.5 <sup>+4.5</sup> <sub>-4.5</sub>	0.855 <sup>+0.023</sup> <sub>-0.023</sub>	5.09 <sup>+5.23</sup> <sub>-5.23</sub>	800-820	-	-
Burgasser et al. 2006 (EM)	-	-	5.1	780-820	-	-

<sup>1</sup> EC FM is Equilibrium Chemistry Forward Model<sup>2</sup> NEC FM is Non-Equilibrium Chemistry Forward Model<sup>3</sup> FCR is Free Chemistry Retrieval<sup>4</sup> ECR is Equilibrium Chemistry Model<sup>5</sup> SML is Supervised Machine Learning<sup>6</sup> EM = Evolutionary Model<sup>7</sup> FM = Forward Model

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 database. 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  $S_{\text{cal}}$ ) along with the mass (and by extension the

inferred surface gravity) we ran a retrieval fitting only the 1.2 - 2.5  $\mu\text{m}$  part of the spectrum. This cut-off 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  $\mu\text{m}$  retrieval to the 0.85 - 1.2  $\mu\text{m}$  data showed a significant mismatch between the model fit and the observed SED in this region, as shown in Figure 6. This two-step approach leads to the most literature consistent values

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

	TW, 0.85–2.5 $\mu$ m	TW, 1.2–2.5 $\mu$ m	Kitzmann et al. (2020)	Burningham et al. (2017)	Line et al. (2015)
log(H <sub>2</sub> O)	−3.33 <sup>+0.03</sup> <sub>−0.03</sub>	−3.11 <sup>+0.05</sup> <sub>−0.05</sub>	−3.33 <sup>+0.05</sup> <sub>−0.06</sub>	−3.42 <sup>+0.16</sup> <sub>−0.22</sub>	−3.40 <sup>+0.13</sup> <sub>−0.13</sub>
log(CH <sub>4</sub> )	−3.39 <sup>+0.03</sup> <sub>−0.03</sub>	−3.34 <sup>+0.05</sup> <sub>−0.06</sub>	−3.28 <sup>+0.06</sup> <sub>−0.09</sub>	−3.44 <sup>+0.20</sup> <sub>−0.31</sub>	−3.45 <sup>+0.10</sup> <sub>−0.10</sub>
log(NH <sub>3</sub> )	−4.58 <sup>+0.04</sup> <sub>−0.04</sub>	−4.69 <sup>+0.07</sup> <sub>−0.09</sub>	−4.38 <sup>+0.07</sup> <sub>−0.10</sub>	−4.82 <sup>+0.26</sup> <sub>−2.47</sub>	−4.64 <sup>+0.15</sup> <sub>−0.15</sub>
log(CO)	−7.66 <sup>+3.22</sup> <sub>−2.86</sub>	−7.06 <sup>+4.12</sup> <sub>−3.12</sub>	−7.70 <sup>+2.7</sup> <sub>−2.4</sub>	−7.47 <sup>+3.05</sup> <sub>−3.04</sub>	−7.53 <sup>+2.65</sup> <sub>−3.07</sub>
log(CO <sub>2</sub> )	−8.35 <sup>+2.50</sup> <sub>−2.39</sub>	−8.58 <sup>+2.12</sup> <sub>−2.09</sub>	−7.70 <sup>+2.7</sup> <sub>−2.4</sub>	−7.86 <sup>+2.67</sup> <sub>−2.66</sub>	−7.76 <sup>+2.23</sup> <sub>−2.89</sub>
log(H <sub>2</sub> S)	−8.59 <sup>+2.42</sup> <sub>−2.26</sub>	−3.86 <sup>+0.12</sup> <sub>−2.26</sub>	−8.47 <sup>+2</sup> <sub>−2</sub>	−8.74 <sup>+2.68</sup> <sub>−2.20</sub>	−8.94 <sup>+2.22</sup> <sub>−2.11</sub>
log(Na+K)	−5.99 <sup>+0.03</sup> <sub>−0.03</sub>	−4.37 <sup>+0.06</sup> <sub>−0.06</sub>	−5.86 <sup>+0.04</sup> <sub>−0.03</sub>	−5.47 <sup>+0.09</sup> <sub>−0.30</sub>	−5.45 <sup>+0.06</sup> <sub>−0.06</sub>

for the retrieved parameters but does lead to a very slightly lower Bayesian Evidence value (see section 3.2.5) due to a slightly worse fit of the *J* band peak.

While this strategy did derive results consistent with previous studies it does have its limitations and imperfections. Firstly, 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,  $S_{cal}$  and distance, along with the inferred surface gravity, can be seen in Figure 8, along with the spectral fit to the data used. These results are also summarised in Table 3. In general, the retrieved parameter values are consistent with previous studies. Values from previous studies can also be seen 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  $2\sigma$  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 et al. 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 Burgress et al. (2006) have obtained slightly higher values for this parameter.

## 4.3 Abundances

The posterior distributions for the retrieved abundances are shown in Figure 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 Figure 6. These show that the three most abundant molecules are H<sub>2</sub>O, CH<sub>4</sub> and NH<sub>3</sub>, 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) and Allard et al. (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 due to condensation.

## 4.4 Temperature-Pressure profile

Our retrieved temperature-pressure profile is very similar to that obtained in the Kitzmann et al. (2020) study (see Figure 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.

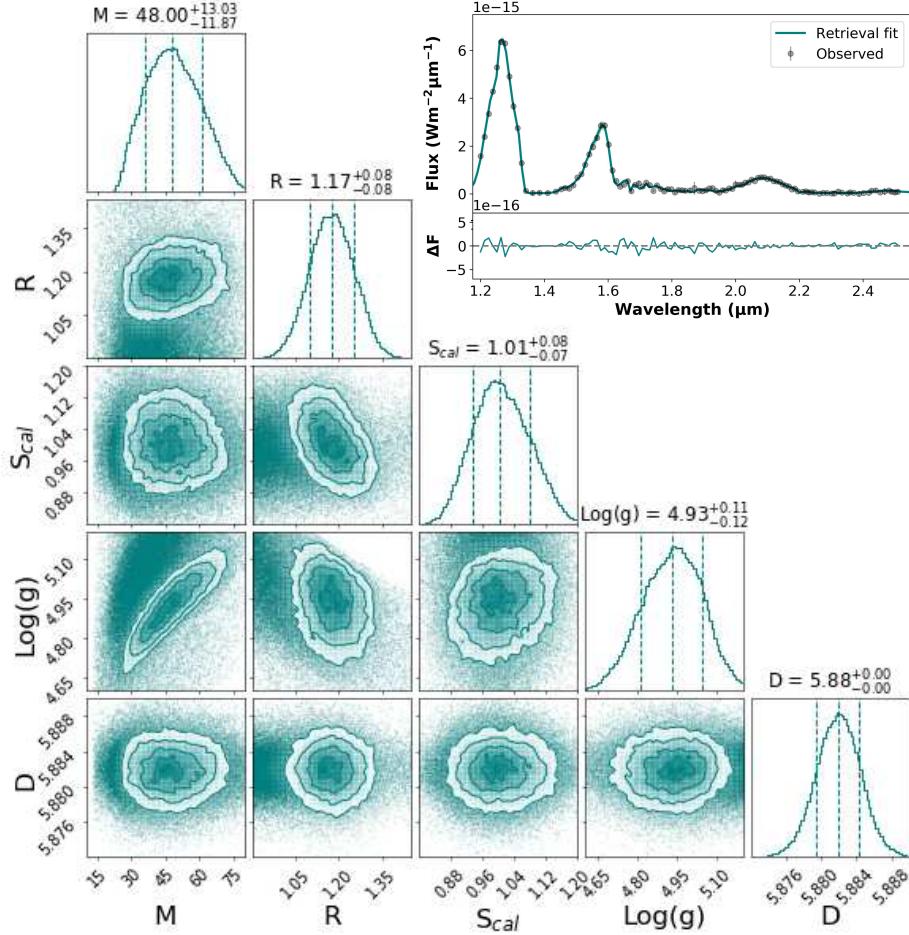


Figure 8. GJ 570D bulk parameter posterior probability distributions for the spectral fit of the 1.2–2.5 $\mu$ m data used.

## 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 *K1* and *K2* data. We present comparison posteriors in Figure 11, with the results from the individual data sets shown in Figures A1 and A2. As Samland et al. (2017) did not fit the GPI *K1* and *K2* 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 figure A3. 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

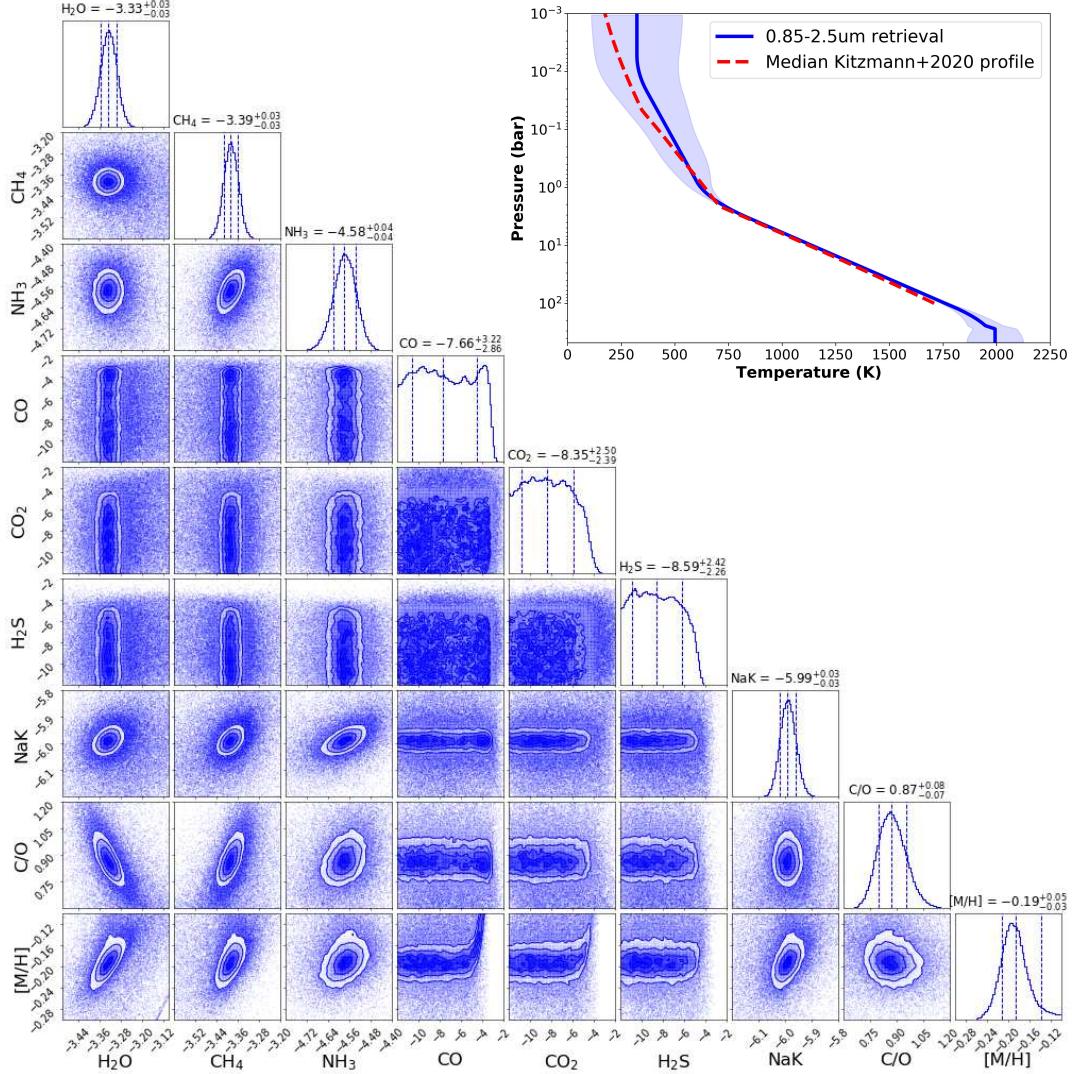
Our highest Log(Ev) posterior probability distributions for the Mass, Radius,  $S_{\text{cal}}$  and Distance, along with the inferred surface gravity are presented in Figure 11. We find that our retrieval analysis is able to produce excellent fits to 51 Eri b's observed SED (see Figure 9) while deriving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

The  $S_{\text{cal}}$  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  $S_{\text{cal}} \sim 1$  is derived, indicating a model preference for this absolute flux calibration given the priors set. All the derived  $S_{\text{cal}}$  values can be found in Figure A1, A2 and A3.

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.

**Table 5.** Summary of retrieved bulk parameters for 51 Eri b. Values from previous studies are also included for comparison. L17: [Lavie et al. \(2017\)](#). GPI all: GPI  $J, H, K1, K2$  data. SPHERE & GPI: SPHERE  $Y, J, H$  & GPI  $K1, K2$  data. SPHERE only: SPHERE  $Y, J, H$  data. U: Uniform clouds coverage. P: Patchy cloud coverage

	TP	Profile	Type	Log(Ev)	Mass (M <sub>Jup</sub> )	Radius (R <sub>Jup</sub> )	log(g)	T <sub>eff</sub> (K)	C/O	[M/H]
TW, Cloudless (GPI $J, H, K1, K2$ data)		<i>npoint</i> , flexible		5596.02	8.50 <sup>+2.86</sup> <sub>-3.44</sub>	1.09 <sup>+0.11</sup> <sub>-0.11</sub>	4.26 <sup>+0.16</sup> <sub>-0.23</sub>	769 <sup>+37</sup> <sub>-41</sub>	0.92 <sup>+0.19</sup> <sub>-0.27</sub>	-0.26 <sup>+0.66</sup> <sub>-0.18</sub>
TW, Cloudless (SPHERE $Y, J, H$ & GPI $K1, K2$ data)		<i>npoint</i> , flexible		5141.33	7.93 <sup>+3.22</sup> <sub>-3.54</sub>	1.18 <sup>+0.12</sup> <sub>-0.12</sub>	4.16 <sup>+0.18</sup> <sub>-0.26</sub>	700 <sup>+42</sup> <sub>-45</sub>	0.97 <sup>+0.09</sup> <sub>-0.20</sub>	-0.04 <sup>+0.95</sup> <sub>-0.49</sub>
TW, Cloudless (SPHERE $Y, J, H$ data)		<i>npoint</i> , flexible		2330.45	8.25 <sup>+3.01</sup> <sub>-3.33</sub>	1.31 <sup>+0.12</sup> <sub>-0.11</sub>	4.09 <sup>+0.15</sup> <sub>-0.23</sub>	909 <sup>+37</sup> <sub>-50</sub>	0.40 <sup>+0.26</sup> <sub>-0.15</sub>	-0.66 <sup>+0.14</sup> <sub>-0.11</sub>
<b>Nielsen et al. 2019</b>				2.6 <sup>+0.3</sup> <sub>-0.3</sub>	-	-	-	-	-	-
Samland et al. 2017 (PTC-uniform clouds) <sup>1</sup>				9.1 <sup>+4.9</sup> <sub>-3.3</sub>	1.11 <sup>+0.16</sup> <sub>-0.14</sub>	4.26 <sup>+0.25</sup> <sub>-0.25</sub>	760 <sup>+20</sup> <sub>-20</sub>	-	-	-
Samland et al. 2017 (PTC-patchy clouds) <sup>1</sup>				14.5 <sup>+8.7</sup> <sub>-5.6</sub>	1.11 <sup>+0.16</sup> <sub>-0.14</sub>	4.47 <sup>+0.24</sup> <sub>-0.26</sub>	757 <sup>+24</sup> <sub>-24</sub>	-	-	-
Samland et al. 2017 (PTC-clear) <sup>1</sup>				14.5 <sup>+4.7</sup> <sub>-3.1</sub>	0.40 <sup>+0.02</sup> <sub>-0.02</sub>	5.35 <sup>+0.15</sup> <sub>-0.12</sub>	982 <sup>+18</sup> <sub>-15</sub>	-	-	-
Samland et al. 2017 (Morley et al. 2012 clouds)				64.9 <sup>+19.1</sup> <sub>-15.6</sub>	1.01 <sup>+0.07</sup> <sub>-0.06</sub>	5.19 <sup>+0.10</sup> <sub>-0.11</sub>	684 <sup>+16</sup> <sub>-20</sub>	-	-	-
Rajan et al. 2017 (Iron-silicate, patchy clouds)				-	0.68 <sup>+0.13</sup> <sub>-0.14</sub>	3.25	737 <sup>+39</sup> <sub>-46</sub>	-	-	-
Rajan et al. 2017 (Sulfide, salt, uniform clouds)				-	0.90 <sup>+0.23</sup> <sub>-0.26</sub>	4.05 <sup>+0.36</sup> <sub>-0.35</sub>	605 <sup>+61</sup> <sub>-86</sub>	-	-	-
Macintosh et al. 2015 (cloud-free)				67	0.76	5.5	750	-	-	-
Macintosh et al. 2015 (partial-cloud)				2	1	3.5	700	-	-	-



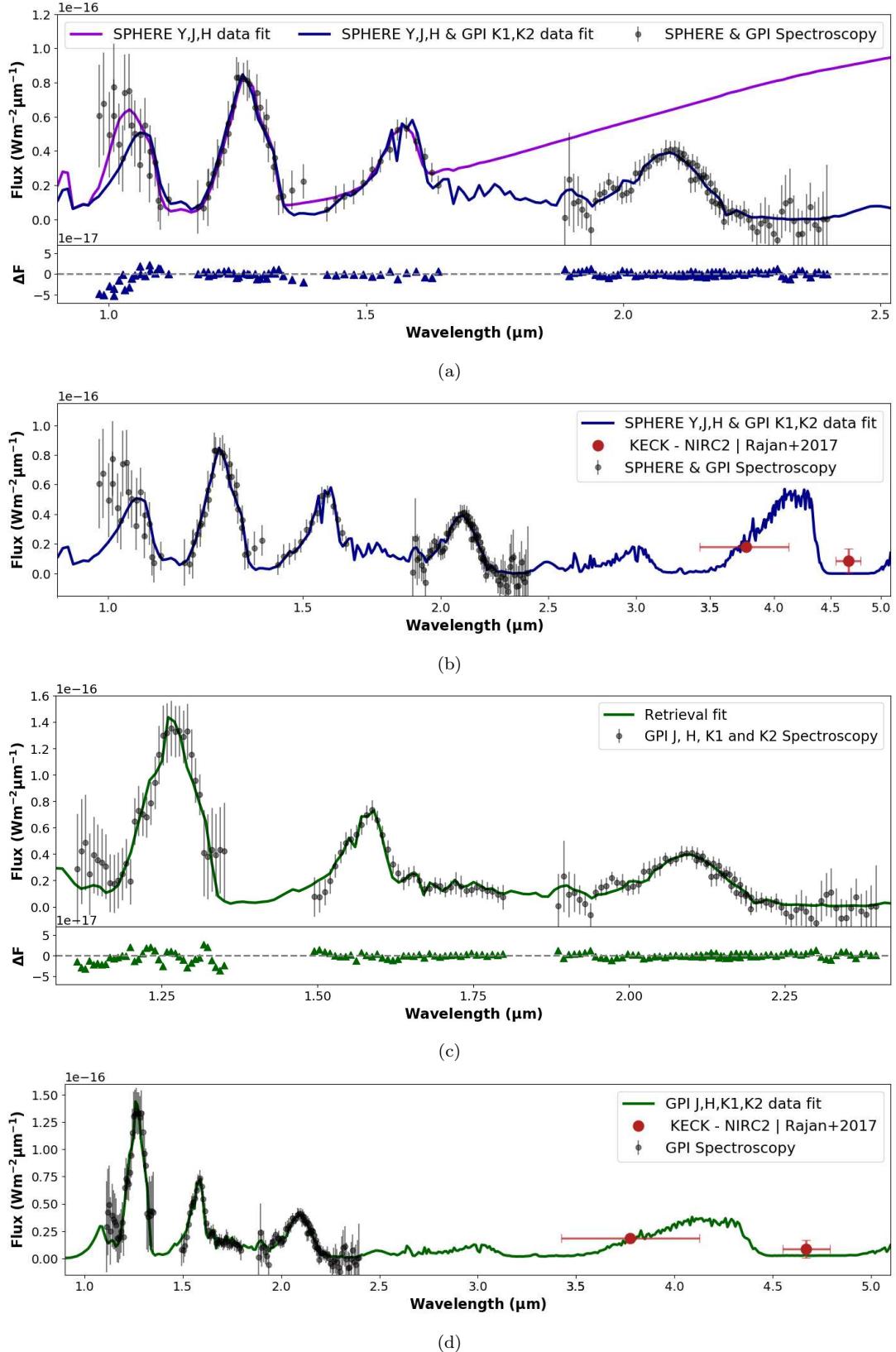
**Figure 9.** GJ 570D mixing ratio posteriors. C/O and [M/H] posteriors are inferred parameters, while all the other parameters are sampled as part of the retrieval. The retrieved temperature-pressure profile is also shown along with a comparison to the median profile retrieved in the Kitzmann et al.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

**Table 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) (cloudless, *npoint* TP) retrievals.

	51 Eri b (1) <sup>1</sup>	51 Eri b (2) <sup>2</sup>	GJ 570D
log(H <sub>2</sub> O)	-3.52 <sup>+0.16</sup> <sub>-0.16</sub>	-3.50 <sup>+0.16</sup> <sub>-0.19</sub>	-3.33 <sup>+0.03</sup> <sub>-0.03</sub>
log(CH <sub>4</sub> )	-3.63 <sup>+0.12</sup> <sub>-0.13</sub>	-3.60 <sup>+0.09</sup> <sub>-0.11</sub>	-3.39 <sup>+0.03</sup> <sub>-0.03</sub>
log(NH <sub>3</sub> )	-4.85 <sup>+0.15</sup> <sub>-0.18</sub>	-4.61 <sup>+0.11</sup> <sub>-0.14</sub>	-4.58 <sup>+0.04</sup> <sub>-0.04</sub>
log(CO)	-3.32 <sup>+1.13</sup> <sub>-5.68</sub>	-5.10 <sup>+2.27</sup> <sub>-4.54</sub>	-7.66 <sup>+3.22</sup> <sub>-2.86</sub>
log(Na+K)	-9.52 <sup>+1.69</sup> <sub>-1.59</sub>	-7.65 <sup>+2.58</sup> <sub>-2.83</sub>	-5.99 <sup>+0.03</sup> <sub>-0.03</sub>
C/O	0.97 <sup>+0.09</sup> <sub>-0.20</sub>	0.92 <sup>+0.19</sup> <sub>-0.27</sub>	0.87 <sup>+0.08</sup> <sub>-0.07</sub>
[M/H]	-0.04 <sup>+0.95</sup> <sub>-0.49</sub>	-0.26 <sup>+0.66</sup> <sub>-0.18</sub>	-0.19 <sup>+0.05</sup> <sub>-0.03</sub>

<sup>1</sup> 51 Eri b (1) refers to results retrieved using SPHERE *Y*, *J*, *H* and GPI *K1*, *K2* band data.

<sup>2</sup> 51 Eri b (2) refers to results retrieved using GPI *J*, *H*, *K1* and *K2* band data



**Figure 9.** 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  $K1$  and  $K2$  data. (b) shows the SPHERE  $Y, J, H$  and GPI  $K1$  and  $K2$  data fit extrapolated to longer wavelengths, with the inclusion of KECK-NIRC2 photometry. (c) illustrates the model fit for the retrieval using the GPI  $J$ ,  $H$ ,  $K1$  and  $K2$  data. (d) shows the GPI  $J$ ,  $H$ ,  $K1$ ,  $K2$  data fit extrapolated to longer wavelengths, with the inclusion of KECK-NIRC2 photometry.

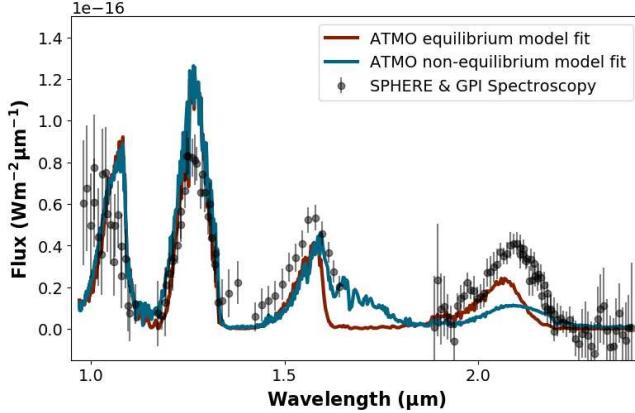


Figure 10. 51 Eri b spectral fit using ATMO

(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  $1\sigma$  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 Figure 10 and Figure A0, illustrating that these grid models are unable to explain the SED of this object or to constrain its surface gravity, radius or effective temperature, using both chemical equilibrium and chemical disequilibrium assumptions as shown in Figures A0.

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 is neglected in this instance. This resulting SED fit is, however, inaccurate when extrapolated to the  $K$  band as shown in Figure 9a.

## 5.2 Abundances, tentative ammonia detection

The highest Log(Ev) posterior distributions for the retrieved abundances are shown in Figure 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  $1\sigma$ , 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 Figure 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  $1\sigma$  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  $\sim 1\mu m$  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 observations, and at  $2.5\sigma$  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). 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 disks (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 Figure 11 and Table 6 are driven by the GPI  $K1$  and  $K2$  band data. We do not detect ammonia when only analysing the  $Y$ ,  $J$  and  $H$  band data as shown in Figure A3. Also, as noted previously, this analysis does not account for potential cross correlated noise, which could reduce the confidence of this detection.

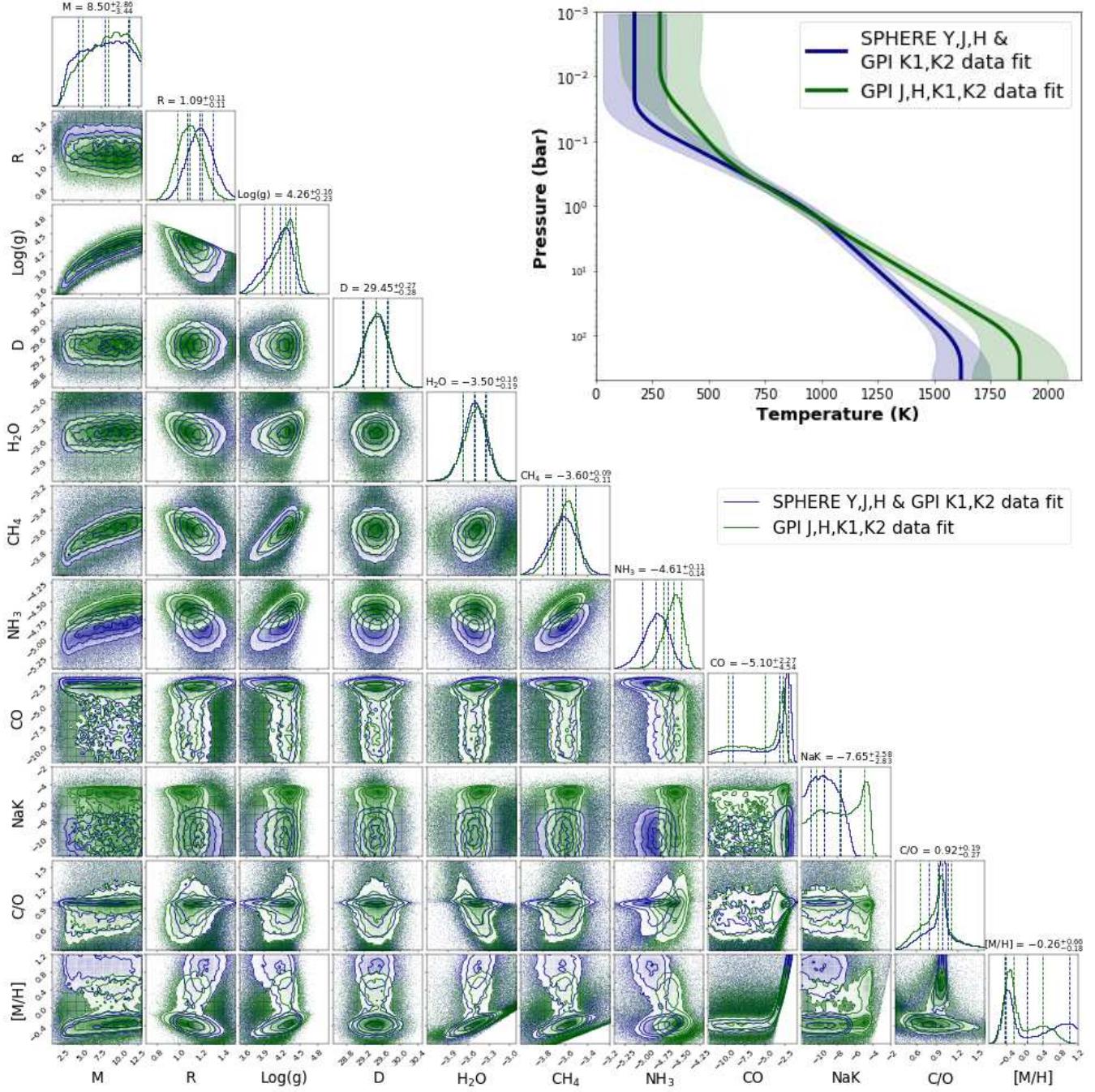
## 5.3 Temperature-Pressure profile

The derived *npoint* temperature-pressure profiles, retrieved with each data set are shown in Figure 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  $\sim 20$  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 Figure 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  $K1$  and  $K2$  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 Figure 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 Figure 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 Figure 12) and ATMO 2020 fitting (see Figure 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 Birmingham 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



**Figure 11.** 51 Eri b posteriors. Blue indicates the retrieved values from the SPHERE  $Y, J, H$  and GPI  $K1, K2$  data set. Green indicates the retrieved values from the GPI  $J, H, K1, K2$  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.

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 thermo-compositional 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 et al. 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

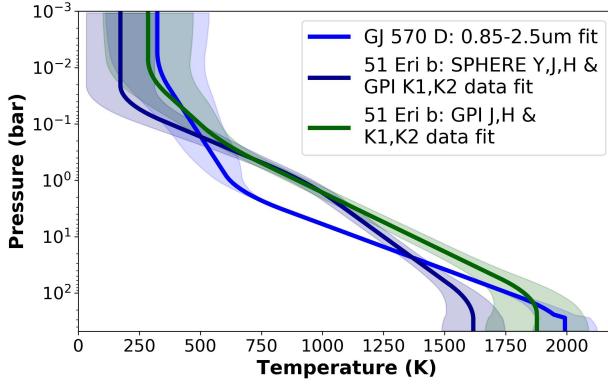


Figure 12. Retrieved 51 Eri b TP profile compared to the GJ 570D TP profile.

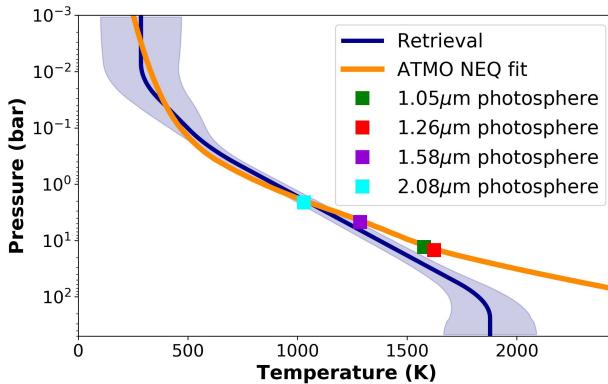


Figure 13. Temperature-pressure profile comparison between profiles derived by TauREx3 and ATMO 2020 for 51 Eri b.

star formation process. When the system is very young, the disk 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 disk. 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 timescale of core accretion is much longer than that of GI.

All our derived radii values are consistent with both the classical cold start and hot start planetary thermal evolution models from Fortney et al. (2008) (updated models from Marley et al. (2007)) as outlined in Figure 15, using age estimates from Macintosh et al. (2015) or Rajan et al. (2017). Figure 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  $\log \epsilon = 8.41$ , Oxygen  $\log \epsilon = 8.80$ ) we derive a C/O ratio of  $\sim 0.41$  for 51 Eri. Therefore, the large mass retrieved for 51 Eri b, its  $\sim 13$  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

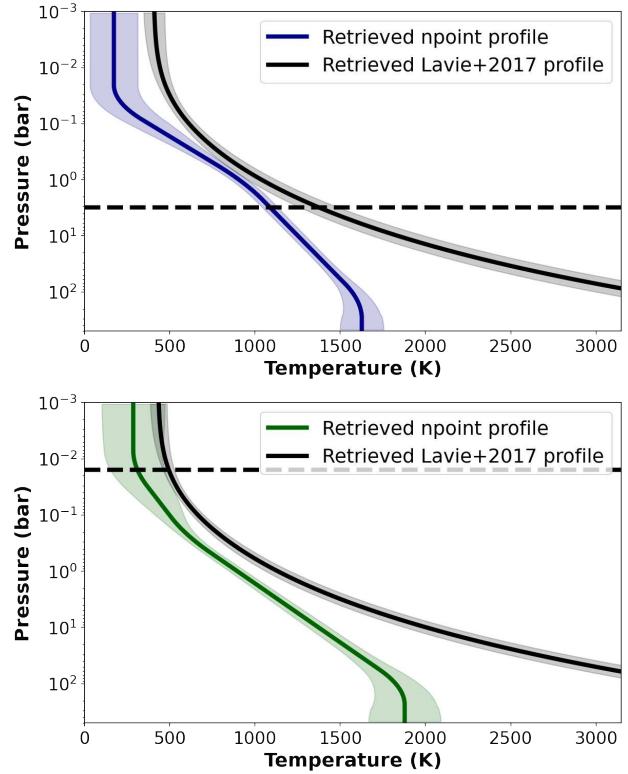


Figure 14. Retrieved temperature pressure profiles using the npoint and Lavie et al. (2017) prescriptions. Also indicated is the retrieved cloud position when using the Lavie et al. (2017) profile. Top: Using SPHERE Y, J, H and GPI K1, K2 band data. Bottom: Using GPI J, H, K1 and K2 band data.

accretion pathway would happen on a much longer timescale resulting in planetesimal enrichment (Mordasini et al. 2016), thus lowering the initial C/O ratio (Espinoza et al. 2017). However, Ille 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 vs cloudless retrievals

In order to perform a cloudy vs cloudless comparison we used the simpler but less inflexible temperature-pressure profile parameterisation 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 well documented trend of flexible temperature-pressure structures being 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 7, matching the results of Samland et al.

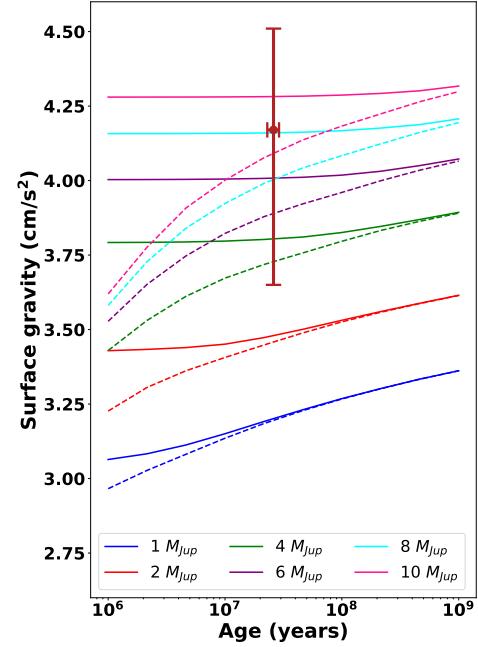
(2017) and Rajan et al. (2017). We also match Samland et al. (2017) and Rajan et al. (2017) in relation to no preference for patchy clouds vs a preference for patchy clouds, relative to the data they used when performing model fitting analysis. However, in the cases of both broadband spectroscopy combinations used in our analysis, the retrievals determined a higher or comparable  $\text{Log}(\text{Ev})$  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; Molière 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 Figure 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) has shown that longer wavelength data, particularly in the mid-IR, is crucial for accurately constraining the temperature-pressure profile when using a reasonably flexible temperature-pressure profile parameterisation. It is also very likely that the simple and flexible power law cloud parameterisations included in our analysis were not the most suitable compared to more physically motivated and condensate specific cloud modelling.

## 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 parameterisations.

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 selec-



**Figure 15.** 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. Purple highlighted region indicates the age of 51 Eri b, as stated in Rajan et al. (2017). The error bar indicates the retrieved 2 sigma confidence boundary for surface gravity from the SPHERE  $Y$ ,  $J$ ,  $H$  and GPI  $K1$ ,  $K2$  data set with the age estimate from Rajan et al. (2017).

tion, 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.

- The profiles of Allard et al. (2016) and Allard et al. (2019) are only considered valid up to a  $\text{H}_2$  density of  $10^{21} \text{ cm}^{-3}$ . 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 Figure 7. 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.
- The line cores of Na and K, based on the data of Allard et al. (2016) and Allard et al. (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.

**Table 7.** Comparison of 51 Eri b retrieval Log Evidences when employing different model setups. L17: [Lavie et al. \(2017\)](#). GPI all: GPI  $J, H, K1, K2$  data. SPHERE & GPI: SPHERE  $Y, J, H$  & GPI  $K1, K2$  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)	<i>npoint</i> , 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)	<i>npoint</i> , 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

- 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](#)) databases, 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 Figure 7. The larger number of non-resonance lines in the Kurucz database 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.
- 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
- 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](#); [Zalesky et al. 2019](#); [Schneider et al. 2015](#)).

We note there are now several studies that have used the updated broadening coefficients from [Allard et al. \(2016\)](#) and [Allard et al. \(2019\)](#) 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 modeling the  $0.85 - 1.2 \mu\text{m}$  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  $>1.1 \mu\text{m}$ . 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 shap-

ing 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) and Line et al. (2017) studies.

The atmospheric similarities between the bona fide exoplanet 51 Eri b and late-T field brown dwarfs extends 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 et al. 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 (Brogi & Line 2019; Hoeijmakers et al. 2018).

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, M. et al. (2020) when combining

observations of Beta Pic b. Crucially, such flexibly appears employable when using a high quality data set, as in the case of the GRAVITY Beta Pic b data used in the Nowak, M. 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 didn't 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 becomes 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 parameterised 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 parameterisation, using the cloud parameterisation 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-model 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 et al. 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 setups along with self-consistent modelling, as performed in this study, when characterising self-luminous objects. Ideally, when further data becomes 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 $\mu$ m data, while part 2 retrieved the chemical profile of the atmosphere using the 0.85-2.5 $\mu$ m 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 comparing our results to previous studies that used grid modelling. The main results of our 51 Eri b retrieval analysis are:

- 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 un-modelled clouds with this behaviour also being seen in Burningham et al. (2017) and Mollière et al. (2020).

- We confirm and constrain the presence of H<sub>2</sub>O and CH<sub>4</sub>.
- We find tentative evidence of NH<sub>3</sub> in the atmosphere of 51 Eri b, to a  $\sim$ 2.7 sigma confidence. Further observations are required to confirm this.
- We retrieve a super-solar C/O ratio, and a solar consistent [M/H] for 51 Eri b.
- Our surface gravity values are consistent with both classical hot-start and cold-start planetary thermal evolution models from Fortney et al. (2008).
- We demonstrate the importance of the *K*-band observations for constraining the effective temperature and temperature-pressure profile.
- 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.
- 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 un-modelled cloud structure in the case of 51 Eri b's retrieval by adopting a more isothermal gradient.
- 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<sub>4</sub> chemical transition (Tremblin et al. 2016, 2019) in the atmosphere of 51 Eri b.
- 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. 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 vs cloud-free and patchy clouds vs uniform clouds, can be biased by model setup and which data are employed. This points to the need for diverse and rigours testing of retrieval model setups and data combinations, particularly when analysing high-contrast 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 acknowledgements postgraduate funding from the UK's Science and Technology Facilities Council (STFC). 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, The Journal of Physical Chemistry A, 115, 6805

Abel M., Frommhold L., Li X., Hunt K. L., 2012, The Journal of chemical physics, 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., Molliere P., 2019, *A&A*, 628, A120

Amundsen D. S., Baraffe I., Tremblin P., Manners J., Hayek W., Mayne N. J., Acreman D. M., 2014, *Astronomy and Astrophysics*, 564, A59

Amundsen D. S., Tremblin P., Manners J., Baraffe I., Mayne N. J., 2017, *Astronomy and Astrophysics*, 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, *Mon. Not. R. Astron. Soc.*, 460, 4063

Barstow J. K., Heng K., 2020, arXiv e-prints, p. [arXiv:2003.14311](https://arxiv.org/abs/2003.14311)

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, arXiv e-prints, p. [arXiv:1504.07655](https://arxiv.org/abs/1504.07655)

Bernath P. F., 2020, *J. Quant. Spectrosc. Radiat. Transf.*, 240, 106687

Beuzit J.-L., et al., 2008, in *Proc. SPIE*. p. 701418, doi:10.1117/12.790120

Beuzit J. L., et al., 2019, *A&A*, 631, A155

Biller B. A., Bonnefoy M., 2018, Exoplanet Atmosphere Measurements from Direct Imaging. p. 101, doi:10.1007/978-3-319-55333-7\_101

Bodenheimer P., 1974, *Icarus*, 23, 319

Bonnefoy M., et al., 2016, *A&A*, 587, A58

Boss A. P., 1997, *Science*, 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 *Proceedings of the SPIE*, Volume 9147, id. 914721 18 pp. (2014).. p. 914721, doi:10.1117/12.2056468

Brandt T. D., Dupuy T. J., Bowler B. P., Bardalez Gagliuffi D. C., Faherty J., Mirek Brandt G., Michalik D., 2019, arXiv e-prints, p. [arXiv:1910.01652](https://arxiv.org/abs/1910.01652)

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 *Astronomical Society of India Conference Series*. pp 7–16 ([arXiv:1406.4887](https://arxiv.org/abs/1406.4887))

Burgasser A. J., Splat Development Team 2017, in *Astronomical Society of India Conference Series*. pp 7–12 ([arXiv:1707.00062](https://arxiv.org/abs/1707.00062))

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., Marley M. S., Line M. R., Lupu R., Visscher C., Morley C. V., Saumon D., Freedman R., 2017, *MNRAS*, 470, 1177

Burningham B., et al., 2021, *MNRAS*,

Burrows A., Sharp C. M., 1999, *ApJ*, 512, 843

Burrows A., Volobuyev M., 2003, *ApJ*, 583, 985

Burrows A., et al., 1997, *ApJ*, 491, 856

Chabrier G., Baraffe I., Leconte J., Gallardo J., Barman T., 2009, in Stempels E., ed., *American Institute of Physics Conference Series* Vol. 1094, 15th Cambridge Workshop on Cool Stars, Stellar Systems, and the Sun. pp 102–111 ([arXiv:0810.5085](https://arxiv.org/abs/0810.5085)), doi:10.1063/1.3099078

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*

Coles P. A., Yurchenko S. N., Tennyson J., 2019, *Mon. Not. R. Astron. Soc.*, 490, 4638

Crossfield I. J. M., et al., 2014, *nat*, 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, *Astronomy and Astrophysics*, 594, A69

Dupuy T. J., 2018, The Coolest Sample of Brown Dwarf Dynamical Masses, HST Proposal

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

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*. p. 607 ([arXiv:astro-ph/0603179](https://arxiv.org/abs/astro-ph/0603179))

Edwards B., et al., 2020, arXiv e-prints, p. [arXiv:2005.02374](https://arxiv.org/abs/2005.02374)

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, arXiv e-prints, p. [arXiv:1306.2144](https://arxiv.org/abs/1306.2144)

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, *The Astrophysical Journal Supplement Series*, **235**, 24

Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, *PASP*, **125**, 306

Fortney J. J., 2018, preprint, ([arXiv:1804.08149](https://arxiv.org/abs/1804.08149))

Fortney J. J., Marley M. S., Saumon D., Lodders K., 2008, *ApJ*, **683**, 1104

Gaia Collaboration et al., 2016, *A&A*, **595**, A1

Gaia Collaboration et al., 2018, *A&A*, **616**, A1

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, arXiv e-prints, p. [arXiv:2010.01224](https://arxiv.org/abs/2010.01224)

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

Goyal J. M., et al., 2020, *MNRAS*,

Gravity Collaboration et al., 2017, *A&A*, **602**, A94

Gravity Collaboration et al., 2020, *A&A*, **633**, A110

Guillot T., 2010, *A&A*, **520**, A27

Gustafsson B., Edvardsson B., Eriksson K., Jørgensen U. G., Nordlund Å., Plez B., 2008, *Astronomy and Astrophysics*, **486**, 951

Harrington J., et al., 2016, in *American Astronomical Society Meeting Abstracts #227*. p. 212.01

Helling C., 2019, *Annual Review of Earth and Planetary Sciences*, **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, arXiv e-prints, p. [arXiv:1008.4686](https://arxiv.org/abs/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. Spectrosc. Radiative Transfer*, **109**, 1136

Kass R. E., Raftery A. E., 1995, *Journal of the American Statistical Association*, **90**, 773

Kirkpatrick J. D., 2005, *ARA&A*, **43**, 195

Kirkpatrick J. D., et al., 2012, *ApJ*, **753**, 156

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, in *Ground-based and Airborne Instrumentation for Astronomy VII*. p. 1070211, doi:10.1117/12.2311836

Kramida A., Ralchenko Y., Reader J., 2013, *NIST Atomic Spectra Database – Version 5*, doi:10.18434/T4W30F

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, arXiv e-prints, p. [arXiv:2009.10714](https://arxiv.org/abs/2009.10714)

Li G., Gordon I. E., Rothman L. S., Tan Y., Hu S.-M., Kassi S., Campargue A., Medvedev E. S., 2015, *Astrophys. J. Suppl.*, **216**, 15

Line M. R., et al., 2013, *ApJ*, **775**, 137

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

Line M. R., et al., 2017, *ApJ*, **848**, 83

Lissauer J. J., Stevenson D. J., 2007, in Reipurth B., Jewitt D., Keil K., eds, *Protostars and Planets V*. p. 591

Liu M. C., Leggett S. K., Chiu K., 2007, *ApJ*, **660**, 1507

Liu M. C., et al., 2013, *ApJ*, **777**, L20

Lodders K., Fegley B. J., 2006, *Chemistry of Low Mass Substellar Objects*. p. 1, doi:10.1007/3-540-30313-8\_1

Lozi J., et al., 2018, in *Adaptive Optics Systems VI*. p. 1070359 ([arXiv:1809.08301](https://arxiv.org/abs/1809.08301)), doi:10.1117/12.2314282

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, *Proceedings of the National Academy of Science*, **111**, 12661

Macintosh B., et al., 2015, *Science*, **350**, 64

Madhusudhan N., 2018, *Atmospheric Retrieval of Exoplanets*. p. 104, doi:10.1007/978-3-319-55333-7\_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, *Mon. Not. R. Astron. Soc.*, **463**, 771

McKemmish L. K., Masseron T., Hoeijmakers H. J., Pérez-Mesa V., Grimm S. L., Yurchenko S. N., Tennyson J., 2019, *Mon. Not. R. Astron. Soc.*, **488**, 2836

McLean I. S., Sprayberry D., 2003, in Iye M., Moorwood A. F. M., eds, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 4841, Instrument Design and Performance for Optical/Infrared Ground-based Telescopes*. pp 1–6, doi:10.1117/12.461785

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, arXiv e-prints, p. [arXiv:2006.09394](https://arxiv.org/abs/2006.09394)

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, [https://exoplanetarchive.ipac.caltech.edu/docs/counts\\_detail.html](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 Astronomy*, 20, 516

Peach G., Yurchenko S., Chubb K., Baraffe I., Phillips M. W., Tremblin P., 2020, *Contributions of the Astronomical Observatory Skalnaté Pleso*, 50, 193

Pezzato J., et al., 2019, arXiv e-prints, p. arXiv:1909.06487

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, *Mon. Not. R. Astron. Soc.*, **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. Spectrosc. Radiative Transfer*, **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, *Astrophysical Journal Supplement* v.99, 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, *The Astrophysical Journal*, **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, *Nature Astronomy*, **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, *Astronomy and Astrophysics Review*, **21**, 63

Tremaine S., et al., 2002, *ApJ*, **574**, 740

Tremblin P., Amundsen D. S., Chabrier G., Baraffe I., Drummond B., Hinkley S., Mourier P., Venot O., 2016, *ApJ*, **817**, L19

Tremblin P., et al., 2017, *ApJ*, **850**, 46

Tremblin P., et al., 2019, *ApJ*, **876**, 144

Trotta R., 2008, *Contemporary Physics*, **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, *Nature Astronomy*, **3**, 1086

Udry S., et al., 2014, preprint, (arXiv:1412.1048)

Vigan A., et al., 2017, *A&A*, **603**, A3

Vigan A., et al., 2020, arXiv e-prints, p. arXiv:2007.06573

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, arXiv e-prints, p. arXiv:2007.02810

Xuan J. W., et al., 2022, arXiv e-prints, p. arXiv:2208.01657

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, *Mon. Not. R. Astron. Soc.*, p. submitted

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., Wenjun Chang 2020, arXiv e-prints, p. arXiv:2004.09513

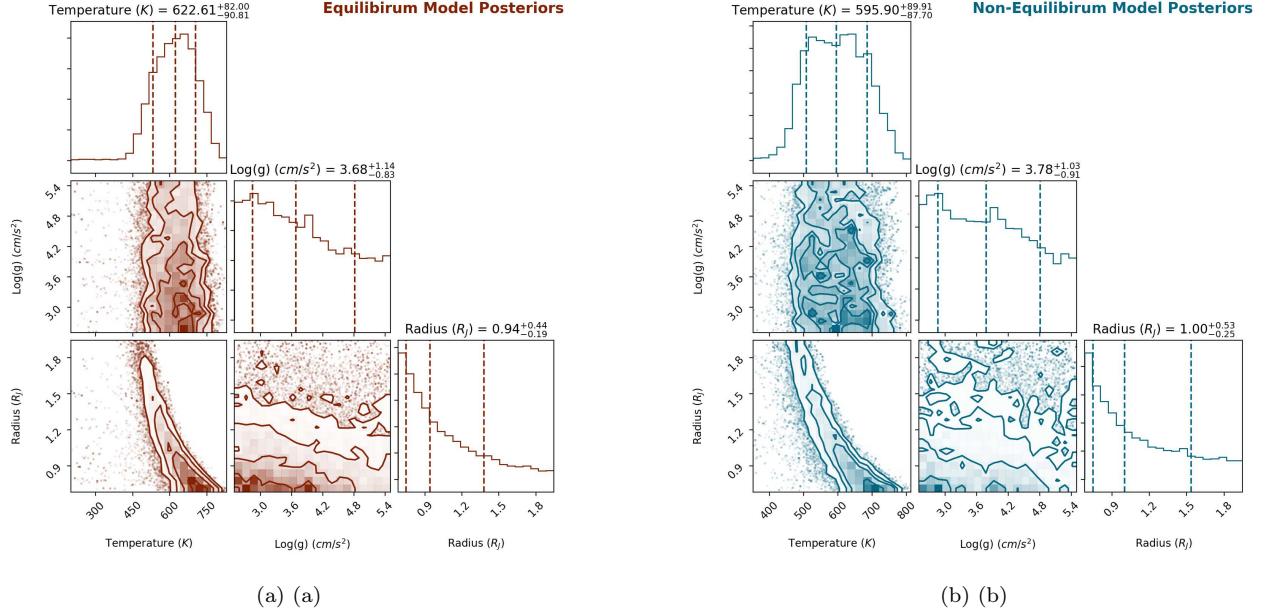
Zhou Y., Bowler B. P., Morley C. V., Apai D., Kataria T., Bryan M. L., Benneke B., 2020, *AJ*, **160**, 77

Zingales T., Waldmann I. P., 2018, preprint, p. arXiv:1806.02906 (arXiv:1806.02906)

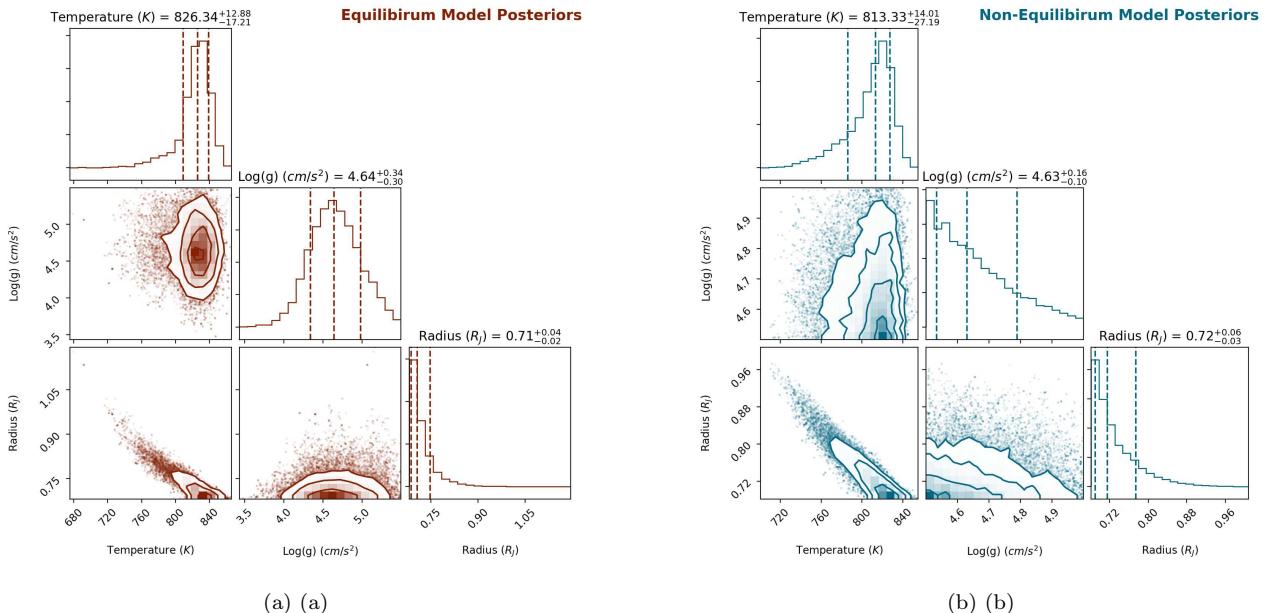
## APPENDIX A: APPENDIX FIGURES

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 datasets.

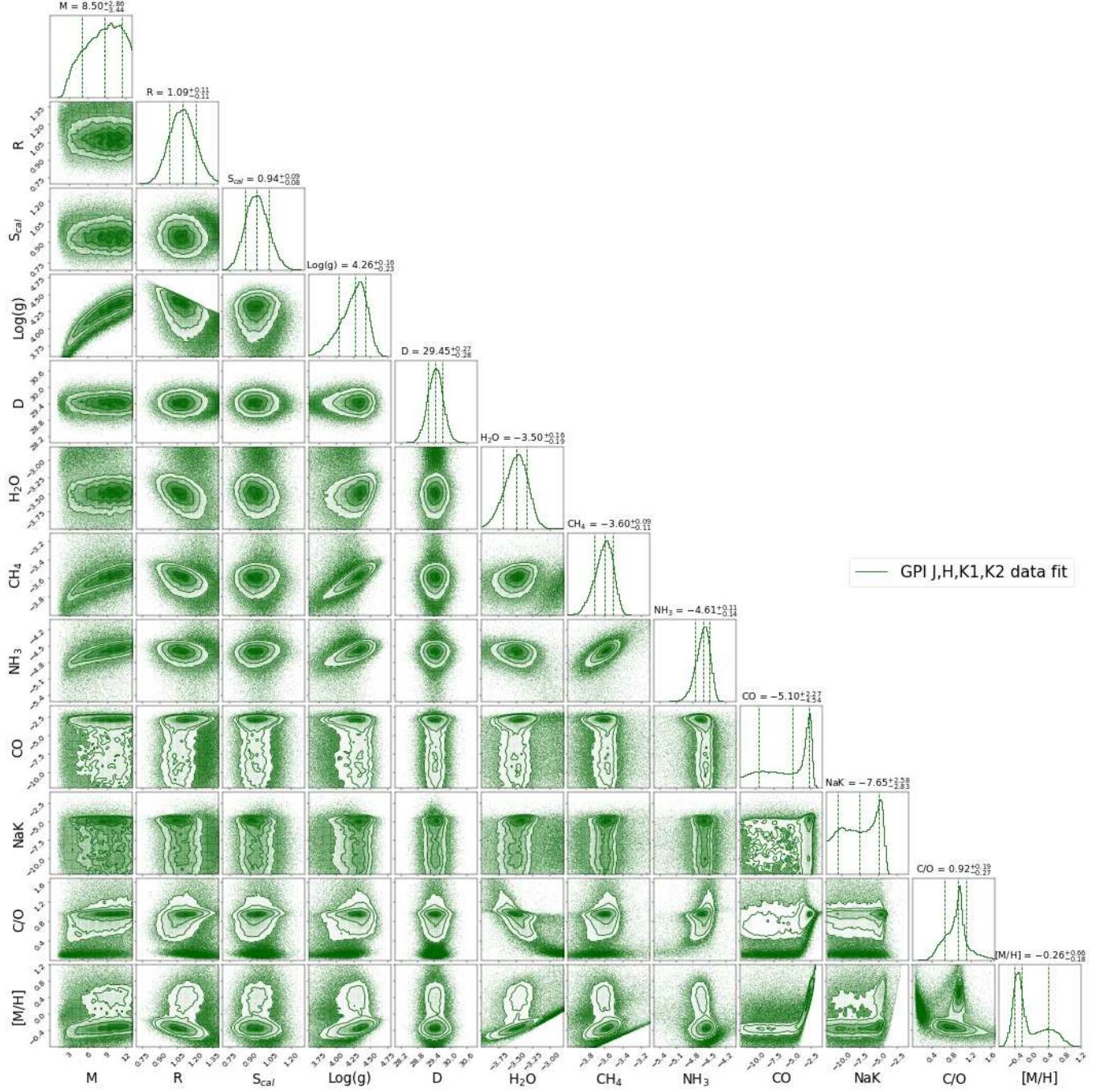
This paper has been typeset from a *T<sub>E</sub>X/L<sub>A</sub>T<sub>E</sub>X* file prepared by the author.



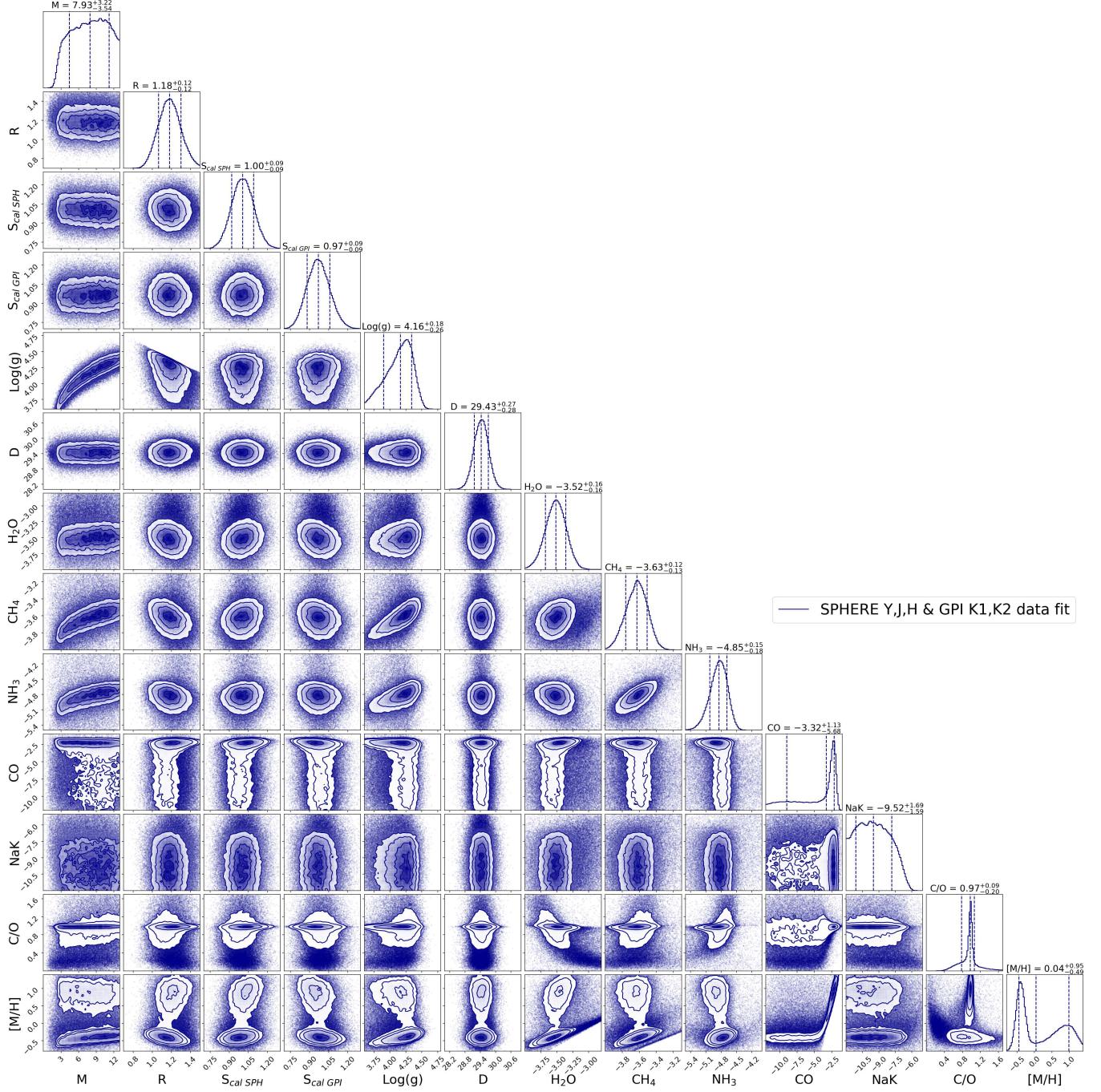
**Figure A0.** ATMO posteriors plots for 51 Eri b. Left: Equilibrium chemistry model. Right: Non-equilibrium chemistry model.



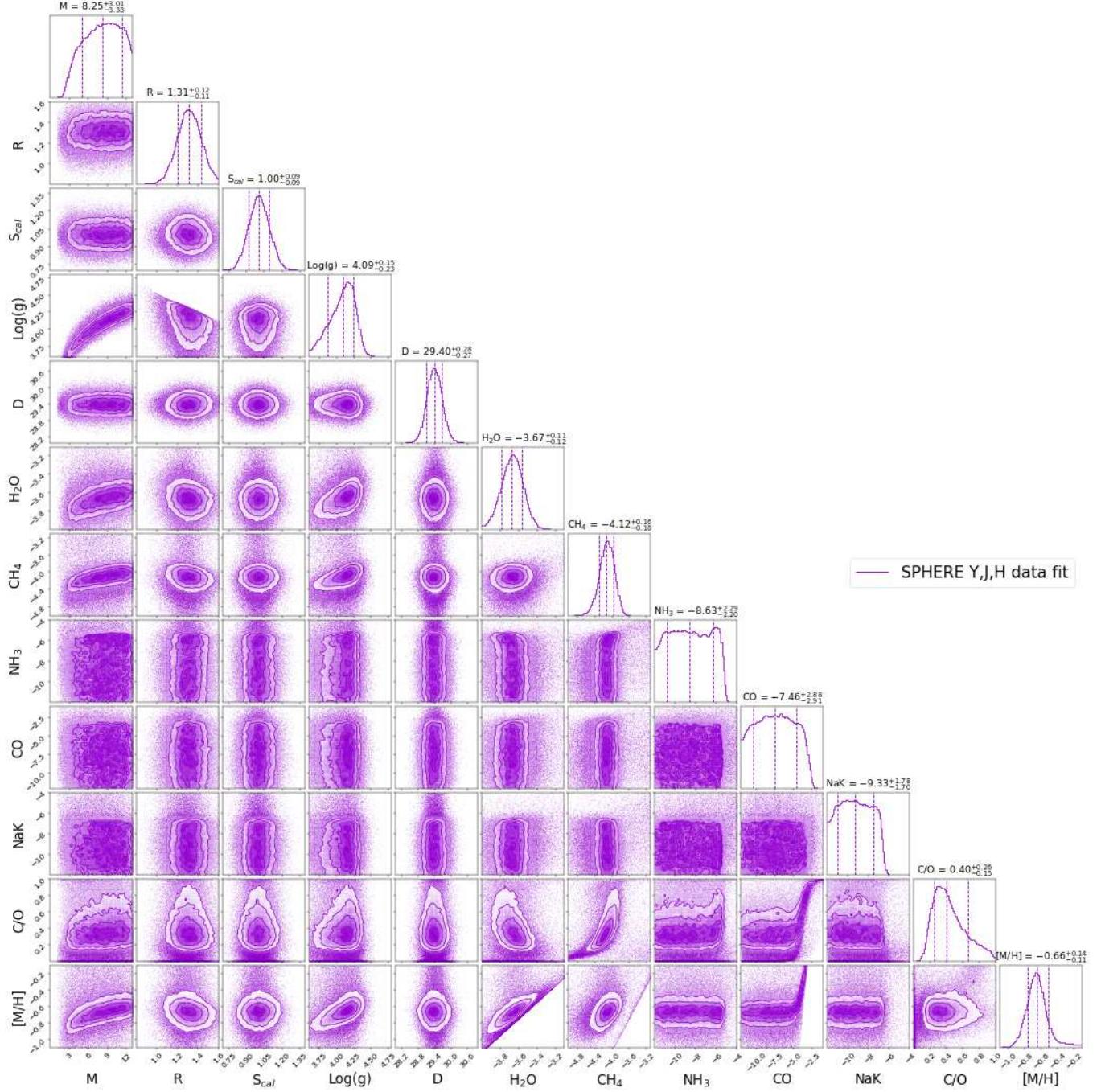
**Figure A0.** ATMO posteriors plots for GJ 570D. Left: Equilibrium chemistry model. Right: Non-equilibrium chemistry model.



**Figure A1.** 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 A2.** 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 A3.** 51 Eri b posteriors for  $Y$ ,  $J$  and  $H$  band data.  $\text{Log}(g)$ ,  $\text{C}/\text{O}$  and  $[\text{M}/\text{H}]$  posteriors are inferred parameters, while all the other parameters are sample as part of the retrieval.