Water observed in the atmosphere of τ Boötis Ab with CARMENES/CAHA

Rebecca K. Webb [1,2]★ Siddharth Gandhi [1,2,3] Matteo Brogi,[1,2,4] Jayne L. Birkby [5] Ernst de Mooij,[6] Ignas Snellen[3] and Yapeng Zhang[3]

[1]*Department of Physics, University of Warwick, Gibbet Hill Road, Coventry CV4 7AL, UK*
[2]*Centre for Exoplanets and Habitability, University of Warwick, Gibbet Hill Road, Coventry CV4 7AL, UK*
[3]*Leiden Observatory, Leiden University, Postbus 9513, NL-2300 RA, Leiden, The Netherlands*
[4]*INAF-Osservatorio Astrofisico di Torino, Via Osservatorio 20, I-10025, Pino Torinese, Italy*
[5]*Astrophysics, University of Oxford, Denys Wilkinson Building, Keble Road, Oxford OX1 3RH, UK*
[6]*Astrophysics Research Centre, Queen's University Belfast, Belfast BT7 1NN, UK*

ABSTRACT

Characterizing the atmospheres of hot Jupiters is important in understanding the formation and migration of these exotic planets. However, there are still many open questions about the chemical and physical properties of these atmospheres. Here, we confirm the detection of water vapour in thermal emission from the non-transiting hot Jupiter τ Boötis Ab with the high-resolution NIR CARMENES spectrograph. Combining over 17 h of observations (560 spectra) and using a Bayesian cross-correlation to log-likelihood approach, we measure a systemic velocity of $V_{sys} = -11.51^{+0.59}_{-0.60}$ km s^{-1} and a radial velocity semi-amplitude of $K_P = 106.21^{+1.76}_{-1.71}$ km s^{-1} for the planet, which results in an absolute mass of $M_P = 6.24^{+0.17}_{-0.18}$ M$_J$ and an orbital inclination of $41.6^{+1.0}_{-0.9}$ degrees. Our retrieved V_{sys} shows a significant shift ($+5$ km s^{-1}) from the literature value, which could be caused by an inaccurate time of periastron. Within the explored model grid, we measure a preference for solar water abundance (VMR $= 10^{-3}$) and find no evidence for additional minor species in the atmosphere. Given the extensive orbital coverage of the data, we searched for a phase dependency in the water signal but found no strong evidence of variation with orbital phase. This detection is at odds with recent observations from SPIRou/CFHT and their tight upper limit on water abundance. We recommend further observations of the atmosphere τ Boötis Ab to try and resolve these discrepancies.

Key words: techniques: spectroscopic – planets and satellites: atmospheres – planets and satellites: fundamental parameters – planets and satellites: individual: τ Boötis Ab.

1 INTRODUCTION

Having been discovered by Butler et al. (1997), τ Boötis b was one of the first main-sequence exoplanets to be observed with precise ($\sim$ m s^{-1}) Doppler shift measurements of stellar absorption lines using high-resolution echelle spectrographs. These first generation exoplanet discoveries became known as hot Jupiters due to their large (Jupiter-like) masses and close-in orbits that are a fraction of the orbit of Mercury in the Solar system, which results in high day-side temperatures for these planets ($T_P \geq 1000$ K). Due to these extreme conditions but favourable signals, these exoplanets are also the most ideal candidates for atmospheric characterization despite accounting for $\lesssim 1$ per cent of sun-like stars (e.g. Wright et al. 2012; Petigura et al. 2018; Boley et al. 2021). τ Boötis b orbits a hot, bright F7 (Gray, Napier & Winkler 2001) star ($V = 4.49$, $H = 3.55$) (Skrutskie et al. 2006; van Belle & von Braun 2009) that is part of a binary system with an M3 dwarf com-

panion (Joy & Abt 1974). Follow-up photometric observations found that the planet was not to be transiting (Baliunas et al. 1997).

The close proximity of τ Boötis b to one of the brightest planetary host stars has made this system one of the most observed non-transiting planets over the last couple of decades. Searches for the reflected light from τ Boötis b using high-resolution instruments commenced soon after its discovery (Charbonneau et al. 1999; Collier Cameron et al. 1999) to directly detect the atmospheric composition, geometric albedo, absolute mass and the orbital inclination. Successive observations (Leigh et al. 2003; Rodler, Kürster & Henning 2010) have also been unsuccessful in detecting the reflected light from the atmosphere of τ Boötis b. Hoeijmakers, Snellen & van Terwisga (2018) produced a meta-analysis of all the data, determining an upper-limit on the planet-to-star contrast and the optical albedo of 1.5×10^{-5} and 0.12, respectively. Recent radial velocity measurements of τ Boötis A have constrained the orbital parameters of the M-dwarf companion and the planet (Justesen & Albrecht 2019), the former of which is expected to reach periastron by $\sim$2026. By this time, the M-dwarf is likely to be within 1 arcsec of the

★ E-mail: r.k.webb@warwick.ac.uk

Table 1. Observations of τ Boötis with CAMRNES. The SNR was calculated as the average SNR of order 51 ($\lambda = 1.19$–$1.21\,\mu m$) over the night. The average V_P is calculated from equation (1) assuming a planet K_P of 110 km s^{-1} and a V_{sys} of -16.9 km s^{-1} averaged over the entire night.

Night	Date	Number of spectra	Exposure time (s)	Phase range	SNR	Average V_P (km s^{-1})
1	2018-March-26	110	40	0.559–0.599	148	-76.6
2	2018-May-11	261	40	0.399–0.459	165	44.3
3	2019-March-12	161	70	0.511–0.576	222	-58.4
4	2019-March-15	165	65	0.420–0.481	246	5.74
5	2019-April-11	133	66	0.542–0.594	186	-61.4

A star, which could cause significant contamination for spectroscopic measurements of the host star and the planet.

The first direct detections from an atmosphere of a non-transiting hot Jupiter was observed from τ Boötis b in the near-infrared (NIR) (Brogi et al. 2012; Rodler, Lopez-Morales & Ribas 2012) with the high-resolution CRIRES instrument (Kaeufl et al. 2004) at the VLT. At $\sim 2.3\,\mu m$, they were able to accurately constrain the radial velocity of the planet by tracing the thermal emission of CO around superior-conjunction where the hot day-side comes into view. Later, Lockwood et al. (2014) reported the first detection of water vapour in the atmosphere of τ Boötis b in the *L* band with a 6σ detection from NIRSPEC/Keck (McLean et al. 1995). Over the past decade, the use of high-resolution spectroscopy has been optimal for measuring the chemical composition (e.g. Giacobbe et al. 2021), wind speeds (e.g. Louden & Wheatley 2015) and temperature–pressure ($T - p$) profiles (e.g. Brogi et al. 2014). More recently, developments have been made on a Bayesian framework to enable a full atmospheric retrieval from these high-resolution spectra (Brogi et al. 2017; Brogi & Line 2019; Gibson et al. 2020; Nugroho et al. 2020). This has allowed for the retrieval of precise absolute chemical abundances (Pelletier et al. 2021) and the C/O ratio (Line et al. 2021). The result from Pelletier et al. (2021) is particularly interesting in this context, as only carbon monoxide was measured in the atmosphere of τ Boötis Ab. In spite of the precise abundance derived for this species ($\log(CO) = -2.46^{+0.25}_{-0.29}$), they reported a non-detection of water vapour and placed a tight upper limit of VMR $\log(H_2O) \leq -5.66\,(3\sigma)$. Given the previous detection of water by Lockwood et al. (2014) and the inconclusive evidence from Brogi et al. (2012), τ Boötis is an interesting planet to revisit with independent data, which is what we present in this paper.

The bulk of NIR observations of hot Jupiters thus far have been to measure the thermal emission brightness variation of the planet as a function of line of sight (i.e a phase-curve). This gives a three-dimensional (3D) picture of the temperature and cloud coverage across the longitude of the planet (e.g. Knutson et al. 2007; Kreidberg et al. 2018). These phase variations leave imprints in their chemical and $T - p$ structure as a function of the orbital phase (e.g. Stevenson et al. 2017). These phase variations are expected to become more prominent for the hotter class of hot Jupiters known as the ultra-hot Jupiters (> 2000 K) (Parmentier et al. 2018) because of the extreme temperature gradient between the hot day-side and cooler night-side. This variation in the temperature structure causes thermal dissociation to occur, which plays an important role in the heat recirculation of the hottest planets (Bell & Cowan 2018). Ehrenreich et al. (2020) and Kesseli & Snellen (2021) measured the phase-resolved variation in the absorption of neutral iron in the transmission spectrum of the ultra-hot Jupiter WASP-76b, attributed to the rain-out of iron on the night-side, with high-resolution spectroscopy. However, at high resolution, phase variations in the thermal emission spectrum have remained elusive.

In this analysis, we confirm the detection of water vapour in the thermal emission with day-side observations of the hot Jupiter τ Boötis b with high-resolution spectroscopy. τ Boötis b is estimated to have day-side temperature limits between $\sim$1980 and 1670 K assuming zero and perfect heat redistribution to the night-side, respectively (Cowan & Agol 2011). We also split the spectra into pre- and post-superior conjunctions to measure any variations in the water signal as a function of phase. We outline the observations in Section 2, the data reduction and analysis in Section 4, and the results in Sections 3 and 5. Finally, we discuss the results in Section 6 with a summary of our conclusions in Section 7.

2 OBSERVATIONS

The day-side emission from the τ Boötis Ab system was observed over five nights [see also Zhang et al. (2020)] with the high-resolution spectrograph CARMENES mounted on the 3.5-m telescope at the Calar Alto Observatory (Quirrenbach et al. 2014). We used two nights of spectra taken from the Calar Alto archive (PI: J.A.Caballero and F.J.Alonso-Floriano), which were taken on 2018 March 26 and 2018 May 11 (nights 1 and 2, hereafter). We also observed (PI: M.Brogi) τ Boötis b for a further three nights on 2019 March 12, 2019 March 15, and 2019 April 11 (nights 3, 4, and 5, hereafter). CARMENES consists of separate optical and NIR spectrographs with spectral resolutions of $R = 94\,600$ and 80 400, respectively. In this analysis, we utilize the NIR spectra to probe the thermal emission from the planet with a wavelength coverage of $\lambda = 0.96$–$1.71\,\mu m$, which spans the *Y*, *J*, and *H* bands with a sampling precision of ~ 3.7 km s^{-1} pixel^{-1}. The NIR spectra are dispersed into 28 echelle orders, orders 63–36, which are imaged on two 2k × 2k Hawaii-2RG detectors.

A time series of spectra was taken pre- ($\phi < 0.5$, nights 2 and 4) and post-superior-conjunction ($\phi > 0.5$, nights 1, 3, and 5). In Table 1, we show the observations of τ Boötis taken with CARMENES with their respective average exposure times, number of spectra obtained, and observed phase range. For nights 1 and 2, we removed the final 10 and 52 spectra, respectively, due to a rapidly decreasing SNR in the spectra, likely due to increased cloud coverage. For each exposure, one fibre was held on the target and a second fibre was placed on the sky for accurate thermal background subtraction. Each spectrum was subsequently calibrated and reduced using the CARACAL v2.10 (Caballero et al. 2016) pipeline, which performs a dark/bias correction, order tracing, a flat-relative optimal extraction, and an accurate wavelength calibration from a U–Ne lamp. We use this solution, which is known to be stable at the m s^{-1} level, for our analysis. Thus, we do not rebin the spectra on to a grid at constant resolving power, nor do we apply any correction for shifts at the sub-pixel level.

From Fig. 1, night 4 had the best observing conditions out of the five with the other nights suffering from large variability in
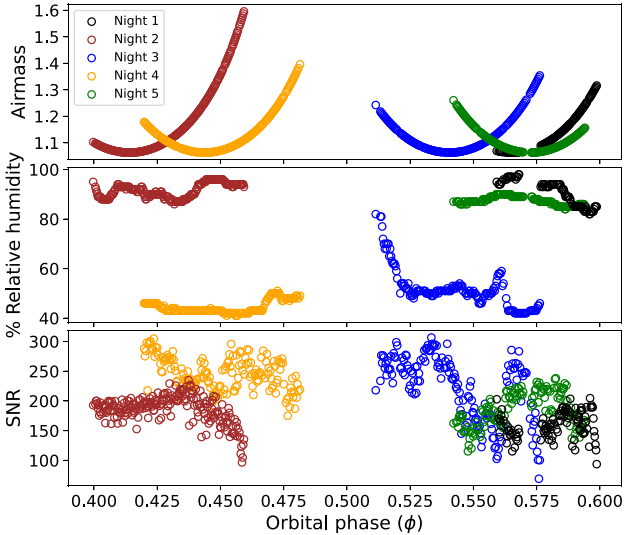
Figure 1. The airmass (top panel), relative humidity (middle panel), and the SNR (lower panel) of order 51 ($\lambda = 1.19$–$1.21\,\mu$m) as a function orbital phase for the observations of τ Boötis b of spectra taken with CARMENES.

atmospheric conditions and high humidity levels. This is reflected in the most stable signal-to-noise ratio (SNR) for all wavelengths (or orders) as shown in the bottom panel of Fig. 1. We note that while the relative humidity at the level of the telescope is not an exact proxy for the precipitable water vapour over the entire atmospheric column density, it is at least an indicator of the overall quality of the night.

3 TELLURIC REMOVAL

First, we need to remove the contaminating telluric and host stellar absorption that dominate the extracted spectra. For each order, the time-resolved spectra are set up in 2D matrices with dimensions $n_f \times n_x$, where n_f is the frame number and n_x is the number of pixels or wavelength channel, for CARMENES detectors, this is 4080. Before passing the data through the telluric pipeline, we removed orders 41–46 ($\lambda = 1.32$–$1.50\,\mu$m), 52–55 ($\lambda = 1.10$–$1.19\,\mu$m) as these were saturated with telluric lines.

We remove the variation in the observed flux as a function of time. As the telluric features are stationary over a night's observations, we can effectively treat each wavelength channel as a light curve, which can then be detrended from telluric depth variations. Since these spectra were taken at orbital phases close to superior conjunction, we are observing the maximum rate of change in the radial velocity shift of the planet ($\frac{dV_P}{dt}$), thereby minimising the losses of the planetary signature, which will be Doppler shifted across several wavelength channels over the time series. Among the various techniques to remove telluric lines, in this analysis we primarily follow the algorithm used in Brogi & Line (2019) and Webb et al. (2020). Fig. 2 shows an example of the 2D time-resolved spectra as a function of wavelength for the CARMENES echelle order 49 for night 4. Each panel shows each step in the reduction of the spectra, which we outline below:

(1) Extracted spectra from the CARACAL calibration pipeline (see Section 2) stacked in time.

(2) The variations in the light throughput are removed by normalising each spectrum (each row in Fig. 2) by the median of the spectrum.

(3) The time-averaged spectrum is fit with a second-order polynomial to each observed spectrum. This fit is thus divided through each of the spectra to remove the telluric variability as a function of wavelength.

(4) The telluric lines are further detrended in time for each wavelength channel (each column in Fig. 2). This is achieved by fitting the spectra in wavelength as a second-order polynomial with time (we use the orbital phase as our time stamp here). This function is then divided through for each wavelength channel to remove the telluric variations as a function of time.

(5) At this stage, we pass each order through a Gaussian high-pass filter (bin-size of 80 pixels) to remove the wavelength-dependent continuum fluctuations which are still present in the data.

(6) The strongest telluric residuals that remained were masked from the analysis. To determine these highly deviant channels, we calculated the ratio between the standard deviation and the relative errors output from the CARACAL pipeline and thus calculated a median value. A calculation of the difference between this ratio and the median values was then used to determine the channels that rose above the average noise level of the data. Highly deviant pixels are also removed through a $5\,\sigma$ clipping.

As described above, in step (6) we had to mask telluric residuals that were not corrected for in the telluric removal algorithm and were seen as highly deviant channels in the time-series spectra. This is necessary so as to prevent strong time-correlated telluric noise from appearing in the final cross-correlation analysis (see fig. 4 in Brogi et al. 2018). The bottom panel in Fig. 2 shows the final standard deviation in time as a function of wavelength, the lack of highly deviant wavelength channels indicates that the strongest residuals that were left in the data have now been removed. Overall orders, this resulted in a total of ∼ 9, 7, 7, 6, and 7 per cent of the pixels being masked for nights 1–5, respectively. In the case of night 2, we found particularly strong telluric residuals that still remained in the spectra despite this additional masking when correlated with a pure water model (see Fig. 3. As a result, we remove night 2 from the remainder of the analysis in order to avoid any biases in our results. This is further explained in Section 5.1. We do, however, repeat the analysis with the inclusion of night 2 in Appendix A and B.

4 DATA ANALYSIS

4.1 Cross-correlation to log-likelihood mapping

As seen in the residual spectral matrix in panel (5) in Fig. 2, the planetary signal is still hidden within the noise of the data, i.e the SNR ≪ 1 per line. To extract this signal, we cross-correlate the observed spectra with model spectra of opacity sources that may be present in the atmosphere. This amplifies the planet SNR by ∼ $\sqrt{N_{\rm lines}}$, where $N_{\rm lines}$ is the number of strong spectral lines.

We Doppler shift the model spectra into the planet rest frame via spline interpolation, and by the radial velocity of the planet in the observer's frame,

$$V_P(t) = K_P \sin\{2\pi[\phi(t) + 0.5]\} - V_{\rm bary}(t) + V_{\rm sys}, \qquad (1)$$

where K_P is the radial velocity semi-amplitude of the planet, $V_{\rm bary}$ and $V_{\rm sys}$ are the barycentric-earth radial velocity correction and the systemic velocity, respectively. The minus sign in $V_{\rm bary}$ corrects for the fact that barycentric velocities are the observer's velocity computed in the barycentre of the Solar system, and not vice-versa
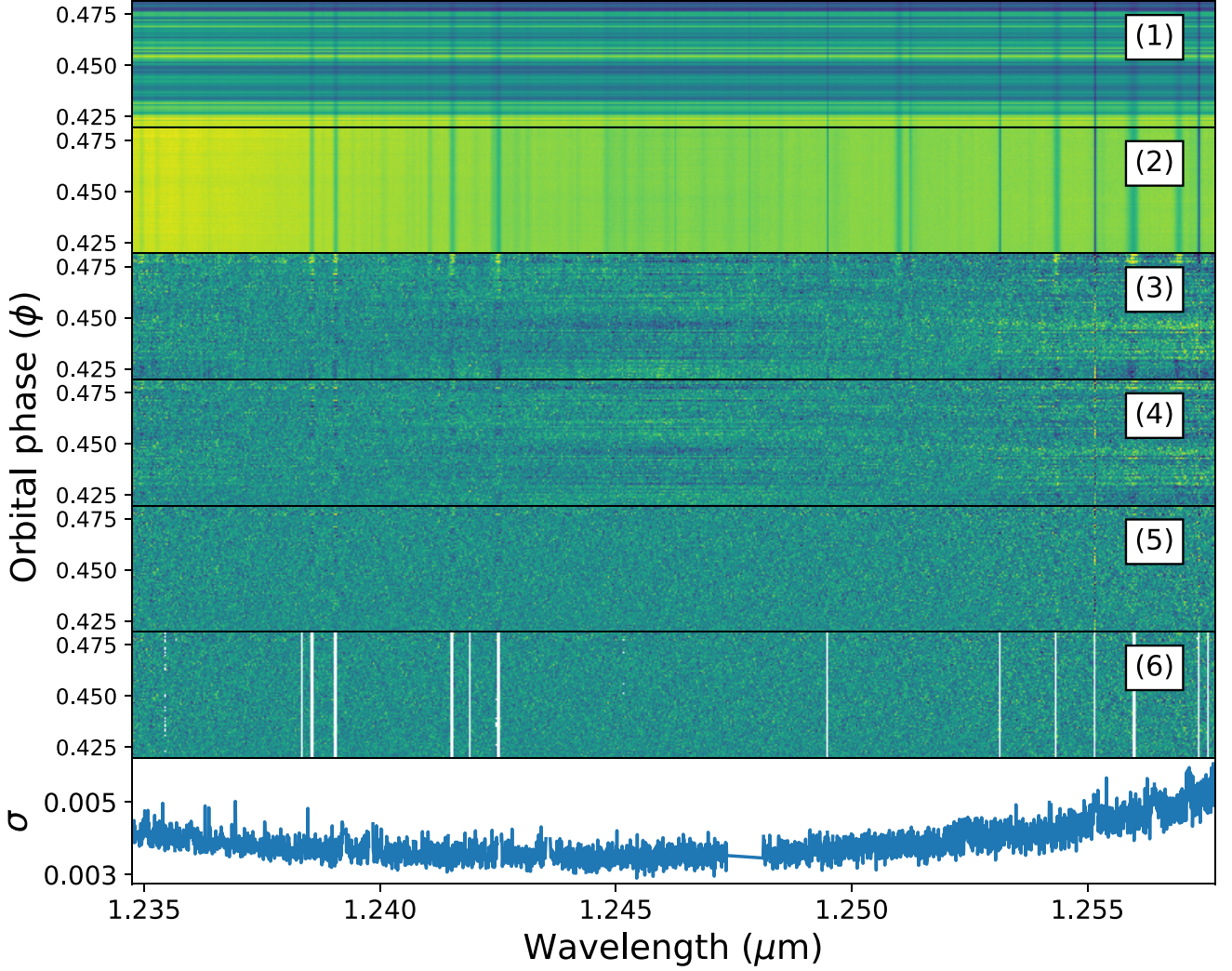
Figure 2. Example of the telluric removal sequence on order 49 of the NIR arm of CARMENES for night 4. The panel numbers indicate the telluric removal steps as described in the text. The bottom panel shows the final standard deviation in time for the spectra. The contrast is increased in panels 3–6 for better visualization of the noise structure in the spectral sequence.

as needed here. The orbital phases are defined as,

$$\phi(t) = \frac{t - T_0}{P},$$ (2)

where t is the time of observations in HJD, T_0 and P are the phase zero-point in HJD and the orbital period in days, respectively. We note that the T_0 stated in Justesen & Albrecht (2019) is in MJD. We needed to convert T_0 into HJD using the `astropy.time` module (Astropy Collaboration et al. 2013, 2018). The T_0 was also determined in the rest frame of the host star therefore the orbital phases calculated with this solution will be in the rest frame of the star. Thus, we need to correct the orbital phases by 0.5 to be in the rest frame of τ Boötis b, as indicated in equation (1). Before cross-correlation, we must also scale the model spectra to the stellar flux of τ Boötis,

$$F_{\text{scaled}}(\lambda) = \frac{a F_P}{\pi B(\lambda, T_{\text{eff}})} \left(\frac{R_P}{R_\star} \right)^2,$$ (3)

where $B(\lambda, T_{\text{eff}})$ is the blackbody stellar flux calculated from the measured effective temperature (T_{eff}), F_P is the modelled emergent flux from the planet in W m^{-2} m^{-1}, R_P, and $R_\star$ are the estimated planetary and measured stellar radii, respectively. The scaling factor

a in equation (3) is assumed to be unity in Sections 5.1 and 5.2; however, it is allowed to vary in the MCMC analysis as described in Section 5.3. We assume a simple blackbody for the stellar flux as we do not expect any significant molecular absorption in the stellar atmosphere of τ Boötis, which has a temperature of $\sim$ 6400 K. The parameter values that were used in the analysis are summarized in Table 2. We adopt a radius of 1.2 R_J, which is approximately the radius adopted in Pelletier et al. (2021) (1.15 R_J) corrected by their retrieved scaling factor of $\sim$1.04. Since τ Boötis b is a non-transiting planet, the K_P is uncertain therefore we must test a range of velocities of the planet around the expected value of $K_P = 110$ km s^{-1} from the literature (Brogi et al. 2012; Lockwood et al. 2014). We tested a range of velocities at $0 \geq K_P \geq 200$ km s^{-1} and $-60 \geq V_{\text{sys}} \geq 60$ km s^{-1} in steps of 2.0 km s^{-1}, which is within the velocity resolution of CARMENES ($\sim$3.7 km s^{-1} in the NIR). The planet radius is also unknown for non-transiting systems; however, we can absorb any uncertainty in the radius into a scaling parameter log (a), which is then retrieved with the other atmospheric parameters (see Section 5.3).

We now add an additional step into the analysis whereby we include a cross-correlation to log-likelihood (CC-log (L)) mapping
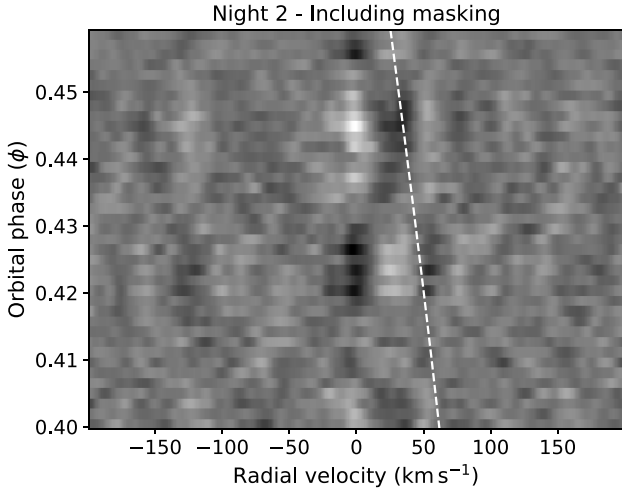
Night 2 - Including masking

Figure 3. The relative CC-log (L) values binned in phase with a pure water model with a log (VMR) $= -3.0$ with the spectra from night 2. The spectra have been co-added in $\phi = 0.0015$ wide bins. Darker shades indicate anticorrelation, whereas lighter shades indicate correlation with the water model. The white dashed line shows the expected radial velocity of τ Boötis b. The strong correlation and anticorrelation stripe in the telluric rest frame (i.e. ~ 0 km s^{-1}) suggests strong telluric absorption in these spectra.

H₂O
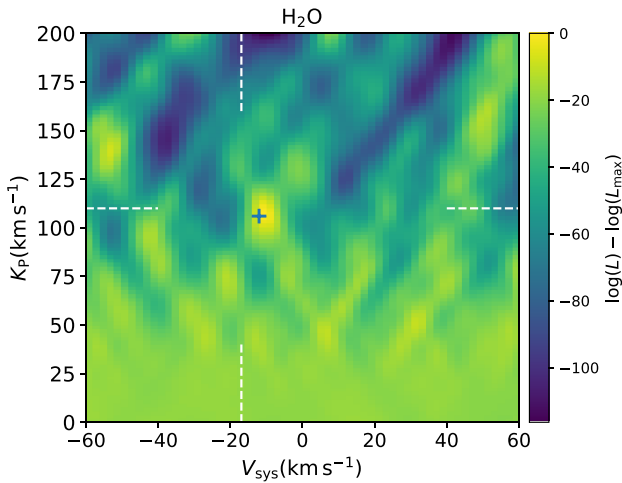
Figure 4. K_{P}–V_{sys} velocity map of the CC-log (L) mapping of the observed spectra with the best-fitting water model with a VMR $= 10^{-3}$ in $\Delta \log (L) = \log (L) - \log (L)_{\max}$. The white dashed lines indicate the expected position of the signal from τ Boötis b from the literature. The blue cross shows the location of the log $(L_{\max})$.

in order to extract statistically robust atmospheric and orbital parameters at high resolution. Since we have time-resolved spectra on a time-scale where any planet signal will be Doppler shifted, we use the mapping from Brogi & Line (2019) as described by,

$$\log(L) = -\frac{N}{2} \log \left[s_f^2 - 2R(l) + s_g^2 \right], \quad (4)$$

where the cross-covariance, $R(l)$, is related to the correlation coefficient by,

$$C(l) = \frac{R(l)}{\sqrt{s_f^2 s_g^2}}. \quad (5)$$

In equation (4), N is the number of wavelength channels used in the cross-correlation and s_f^2 and s_g^2 refer to the variance of the data and

the model, respectively. In equations (4) and (5), l represents the cross-correlation lag, which is equal to the planet's Doppler shift at the time the log (L) is evaluated. Using this form of the log (L), we are assuming the case where an additional atmospheric scaling factor a is equal to unity. This is necessary because we treat the scaling factor as a model parameter, and therefore we apply it to the model spectrum *prior* to the likelihood computation (equation 3), which is necessary to account for the effects of the analysis on the model. Using this method, we are able to directly convert the correlation values into a log-likelihood velocity map.

The log-likelihood values from equation (4) are calculated for each night, each order, and each spectrum and subsequently summed to obtain a single log-likelihood value for each model and each set of parameters,

$$\log(L)_{\mathrm{tot}} = -\sum_{k=1}^{4} \sum_{i=1}^{N_o} \sum_{j=1}^{N_s} \frac{N_{kij}}{2} \log \left[s_{f,kij}^2 - 2R_{kij}(l) + s_{g,kij}^2 \right], \quad (6)$$

where the index k denotes the observing night, i the spectral order, and j the frame number. We further note that the number of orders N_o and the number of spectra N_s vary from night to night.

4.2 Atmospheric models

To model the thermal emission from τ Boötis b, we use the line-by-line radiative transfer code GENESIS (Gandhi & Madhusudhan 2017). These model spectra were produced using the same methods as described in Hawker et al. (2018), Cabot et al. (2019), Webb et al. (2020), and Gandhi, Brogi & Webb (2020b). The spectra were generated at a wavelength range of $\lambda = 0.96$–$1.8\,\mu$m, with a wavenumber spacing of 0.01 cm^{-1} which corresponds to a resolution of $R(\lambda) = \frac{10^6}{\lambda}$, where λ is in μm. Before cross-correlation, these models were re-grid to a constant resolution (i.e. $\lambda/\Delta\lambda$) and then convolved with a Gaussian kernel with a full width at half-maximum (FWHM) equivalent to the instrumental resolution of CARMENES in the NIR ($R = 80, 400$). We also assume each model spectrum has a 1D $T - p$ profile that has been parametrized by upper (T_2, p_2) and lower (T_1, p_1) points in the atmosphere. Above ($p < p_2$) and below ($p > p_1$) these points, the atmosphere is assumed to be isothermal.

Opacity sources were included from the following line list databases; HITEMP, for H₂O (Rothman et al. 2010) and CH₄ (Hargreaves et al. 2020), ExoMol for HCN (Harris et al. 2006; Barber et al. 2014), NH₃ (Coles, Yurchenko & Tennyson 2019), and C₂H₂ (Chubb, Tennyson & Yurchenko 2020). We note that CARMENES has no sensitivity to CO in the *Y*, *J*, and *H* bands and therefore we do not include this molecule in the modelling. Even though there are CO lines at $\sim 1.6\,\mu$m, these are over two orders of magnitude weaker than in the *K* band (e.g. Gandhi et al. 2020b). In addition, at $1.6\,\mu$m the water opacity is stronger than for CO, thus the weaker CO lines will be significantly shielded by water lines. These models also include collisionally induced absorption from H₂-H₂ and H₂-He (Richard et al. 2012) and broadening from each opacity source (Gandhi et al. 2020a). In Table 3, we show the grids of models that were used in our analysis tested against the observed spectra. We fix the pressure points to be $p_1 = 1$ and $p_2 = 10^{-3}$ bars for all the models in the grids. For pressures lower than 10^{-3} bars, the $T - p$ profile is predicted to be isothermal from GCM and 1D modelling (e.g. Beltz et al. 2021) therefore most of the core of the lines will be formed at pressures of $> 10^{-3}$ bars. We generated a $T - p$ profile grid to explore various temperature gradients for the atmosphere with a single opacity source of water with a fixed chemical equilibrium and solar composition abundance of VMR $= 10^{-3.3}$. The choice of

Table 2. Stellar and planetary parameters and their values that were used in the analysis.

τ Boötis A	Symbol (units)	Value	Reference
Spectral type		F7	Gray et al. (2001)
H-band infrared brightness	H (mag)	3.55	Skrutskie et al. (2006)
Effective temperature	$T_{\rm eff}$ (K)	6399 ± 45	Borsa et al. (2015)
Mass	$M_\star$ (M$_\odot$)	1.35 ± 0.03	Takeda et al. (2007)
Radius	$R_\star$ (R$_\odot$)	1.42 ± 0.08	Borsa et al. (2015)
Luminosity	$L_\star$ (L$_\odot$)	3.06 ± 0.16	Borsa et al. (2015)
Distance	d (pc)	15.66 ± 0.08	Gaia Collaboration et al. (2018)
Radial velocity semi-amplitude	$K_\star$ (m s^{-1})	468.42 ± 2.09	Justesen & Albrecht (2019)
Systemic velocity	$V_{\rm sys}$ (km s^{-1})	-16.9 $\pm$ 0.3	Gaia Collaboration et al. (2018)
τ Boötis Ab			
Orbital period	P (d)	$3.31245 \pm 3 \times 10^{-6}$	Justesen & Albrecht (2019)
Radius	$R_{\rm P}$ (R$_{\rm J}$)	1.2	(*estimated*)
Phase zero-point (in the rest frame of τ Boötis A)	T_0 (MJD)	56401.879 ± 0.004	Justesen & Albrecht (2019)
Orbital separation	a (au)	$0.04869^{+0.00039}_{-0.00040}$	Rosenthal et al. (2021)
Eccentricity	e	$0.0074^{+0.0059}_{-0.0048}$	Rosenthal et al. (2021)

Table 3. The grid of models that were used in the analysis. The range of abundances tested for each species is shown, these varied in steps of 1 dex. For each species, the abundance was also allowed to drop to zero in each opacity grid to model each species individually. The exceptions to this are the water abundances, which were fixed to $10^{-3.3}$ and 10^{-3} for the water only models (top row) and the H$_2$O, HCN, and C$_2$H$_2$ grids (bottom row), respectively.

Opacity source(s)	log$_{10}$(VMR)	T_1 (K)	T_2 (K)	Number of models
H$_2$O	-3.3	[1400,1600,1800,2000]	[800,1200,1600,2000]	12
H$_2$O, CH$_4$, HCN	-3.0 to -5.0, -4.0 to -6.0, -5.0 to -7.0	1800	1200	64
H$_2$O, NH$_3$, C$_2$H$_2$	-3.0 to -5.0, -4.0 to -6.0, -5.0 to -7.0	1800	1200	64
H$_2$O, HCN, C$_2$H$_2$	-3.0, -2.0 to -7.0, -2.0 to -7.0	1800	1200	49

temperatures was guided by the range of equilibrium temperatures of the day-side of the planet, $T_{\rm eq} = 1600$–2000 K, which depends on the efficiency of day–night heat redistribution. We also generated a large opacity grid with water combined with further minor species at varying abundances, including at zero abundance. We assume that the abundance for each opacity source in all of the models have a constant VMR with pressure. For these opacity grids, we fixed the temperature gradient to $T_1 = 1800$ and $T_2 = 1200$ K. We note here that with a fixed grid of models, we are limited in our ability to constrain the atmospheric lapse rate and the molecular abundances individually due to the partial correlation between these parameters. A full atmospheric Bayesian retrieval with free parameters for the $T − p$ profile and abundances will constrain these parameters individually for which we defer to a future study.

As with the data, we also pass these models through the telluric removal pipeline as described in Section 3 prior to cross-correlation. This is to replicate the unavoidable scaling effects that occur to the atmospheric signal in the telluric removal sequence, thus avoiding potential biases in the cross-correlation analysis (see final panel in fig. 2 of Brogi & Line (2019).

5 RESULTS

5.1 Single species analysis

We first analysed the data against five single species individually at abundances of log$_{10}$VMR; H$_2$O $= −3.0$ to $−5.0$, CH$_4 = −4.0$ to $−6.0$, HCN $= −5.0$ to $−7.0$, NH$_3 = −4.0$ to $−6.0$, and C$_2$H$_2 = −5.0$ to $−7.0$. These tests were done on the combined species models in rows 2 and 3 of Table 3, with the additional

species effectively removed by setting their abundances to zero. For H$_2$O, the range of abundances used corresponds to the expected solar abundance assuming a solar C/O ratio and metallicity at thermochemical equilibrium down to sub-solar values approaching the upper-limit determined in Pelletier et al. (2021) [log$_{10}$(VMR) $<$ $−5.66$]. For the remaining species, the expected solar abundances are expected to be too low to be observable for a planet with a temperature of ~ 1800 K at a solar C/O ratio and metallicity (e.g. Madhusudhan 2012; Moses et al. 2013). However, we tested a range of enhanced abundances for CH$_4$, HCN, NH$_3$, and C$_2$H$_2$ to include potential scenarios where the atmospheric C/O ratio and metallicity are super-solar, which has been observed in recent high-resolution atmospheric studies (Giacobbe et al. 2021). For these modelling tests, we do not optimize the atmospheric scaling parameter, i.e. $a = 1$.

In Fig. 3, we show the result of correlating the reduced spectra from night 2 with a pure water model for τ Boötis b, shifted in radial velocity. The prominent alternating pattern of correlation (lighter shades) and anticorrelation features (darker shades) at zero lag radial velocity shows that the tellurics are strongly correlating with the atmospheric water models. Even though the radial velocity trail of τ Boötis b is significantly shifted from the telluric rest frame, the correlated telluric noise is overwhelming any potential signal from the planet. Therefore, the inclusion of these spectra will cause spurious telluric noise in the CC-log(L) analysis and they have subsequently been removed from the rest of the analysis (see appendix B and Fig. B2). We show the same phase-resolved correlation with a water model with all nights combined in Fig. B1 in Appendix B.

For the four nights of spectra, co-added in time and wavelength, we observe a signal for the presence of water vapour in the thermal
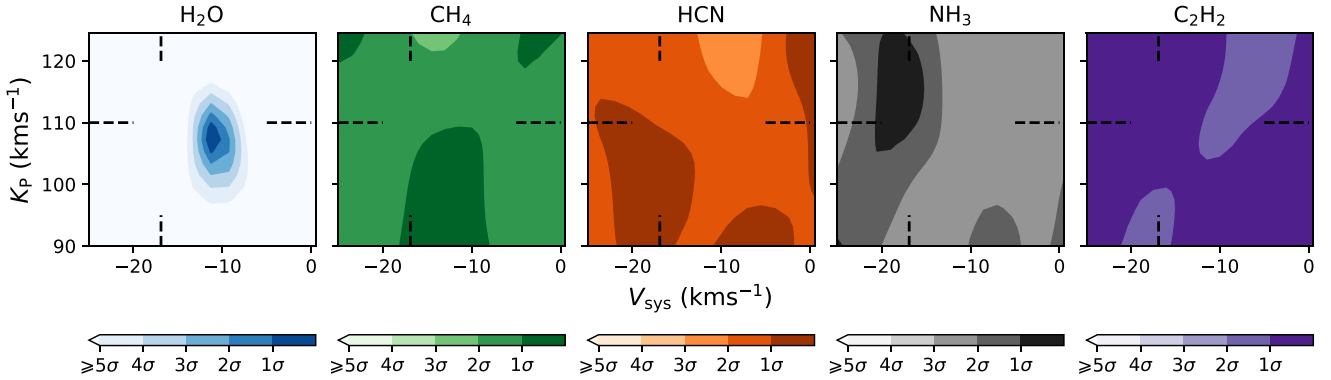
Figure 5. CC-log (L) significance K_P–V_{sys} maps of the five species at solar abundance co-added for the four nights of spectra. The filled contours indicate the areas of significance away from the peak in the log (L). From left to right, we show the maps of; H_2O, CH_4, HCN, NH_3, and C_2H_2. The H_2O map shows a zoomed-in version of the signal seen in Fig. 4. The black dashed lines indicate the location of the orbital solution given from the literature. There is a clear signal from water close to the velocity of the planet and no evidence for any other minor species.

emission of this atmosphere, which is shown in Fig. 4. We find that the water signal peaks at a planet velocity semi-amplitude of $K_P = 106.0^{+2.8}_{-2.2}$ km s^{-1}, which is consistent within 1σ from the literature values (Brogi et al. 2012; Lockwood et al. 2014; Pelletier et al. 2021). However, we find a shift in the systemic velocity of τ Boötis with a $V_{sys} = -12.0^{+1.0}_{-0.5}$ km s^{-1}, which is a $\sim +5.4$ km s^{-1} shift from the expected $V_{sys} = -16.9$ km s^{-1} (Gaia Collaboration et al. 2018; Justesen & Albrecht 2019).

To test for the preference in the water abundance in the modelling, we set the abundances for the combined species in the second row of Table 3 to zero, i.e. CH_4 and HCN are set to zero but H_2O is allowed to vary. We find that an abundance of VMR$_{H_2O} = 10^{-3}$ is strongly preferred over the lower abundances of 10^{-4} and 10^{-5} by 3.8σ and 5.0σ, respectively. In Fig. 5, we show the confidence intervals in the K_P–V_{sys} correlation maps for each individual species. This shows the tight confidence intervals from the water signal that we observe in the continuous CC-log (L) map in Fig. 4 and no evidence for the presence of other minor species in this analysis. To calculate confidence intervals, we converted the CC-log (L) mapping to a chi-square distribution (with two degrees of freedom, K_P and V_{sys}) from the peak in the log-likelihood, log (L_{max}), using Wilks' theorem (Wilks 1938), $\chi^2 = -2\Delta\log(L) = -2\log(L/L_{max})$. Hence, we can determine the p-values from log (L_{max}) by halving the two-tail survival function of the χ^2 distribution. Finally, we are able to convert these p-values into σ levels by calculating the normal distribution inverse survival function.

We also analysed the water signal against varying $T - p$ profiles (see Section 4.2) with a fixed water abundance of VMR = $10^{-3.3}$. We find that a steeper temperature gradient is preferred for this atmosphere with $T_1 = 1800$ K preferred over lower temperatures by 3.2σ. With T_1 fixed to 1800 K, the upper temperature $T_2 = 800$ K is marginally preferred over higher temperatures by 1.6σ (1200 K) and 2.9σ (1600 K). We do note, however, that these are only qualitative constraints on these parameters as the atmospheric lapse rate and chemical abundances partially correlate. Recent studies with high-resolution spectroscopy observations (Line et al. 2021; Pelletier et al. 2021) have shown that by using a full Bayesian atmospheric retrieval, the lapse rate and the absolute abundances can be retrieved will little correlation between these parameters. Thus, constraining both parameters in these spectra will be possible with a full atmospheric retrieval analysis for which we defer to a future study.
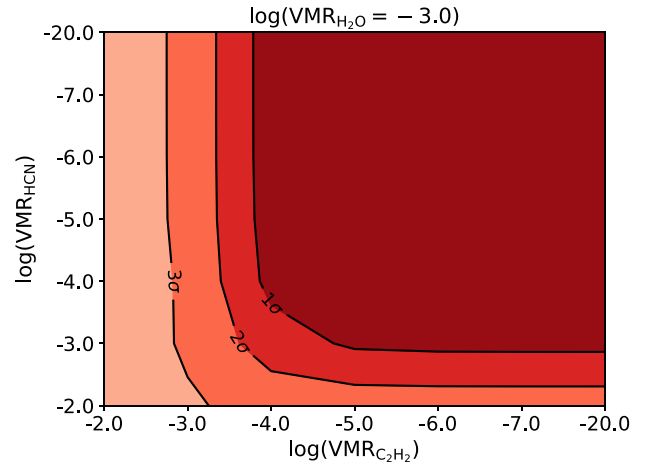

Figure 6. Abundance constraints on combining the best-fitting water abundance with additional HCN and C_2H_2 or without (abundance of -20.0). There is no preference for the addition of these additional species, and we can only place upper limits on the abundances based on this analysis.

5.2 Combined species analysis

In addition to the individual species, we also analysed a grid of models that combined the additional minor species to the best-fitting water model [log $(VMR) = -3.0$]. In initial tests with models that have solar abundances in chemical equilibrium, we saw a marginal increase $(< 1\sigma)$ in the log (L_{max}) with the addition of HCN and C_2H_2 in the modelling. Therefore, we expanded the abundance range of these species with water fixed at VMR = 10^{-3} to super-solar values (see the bottom row in Table 3) to explore whether these species significantly increase the detection significance from the pure water models. In Fig. 6, we show the abundance constraints on the grid containing HCN and C_2H_2. We find no evidence $> 1\sigma$ that these additional species in the modelling improve the detection significantly from the pure water models in this analysis. However, we can place a 3σ and 2σ upper limits on the abundance of C_2H_2 and HCN at a VMR = 10^{-3} and VMR = $10^{-2.5}$, respectively. Again, we note that these limits in abundance are dependent on the choices for the change of temperature with pressure, i.e. the atmospheric lapse rate.
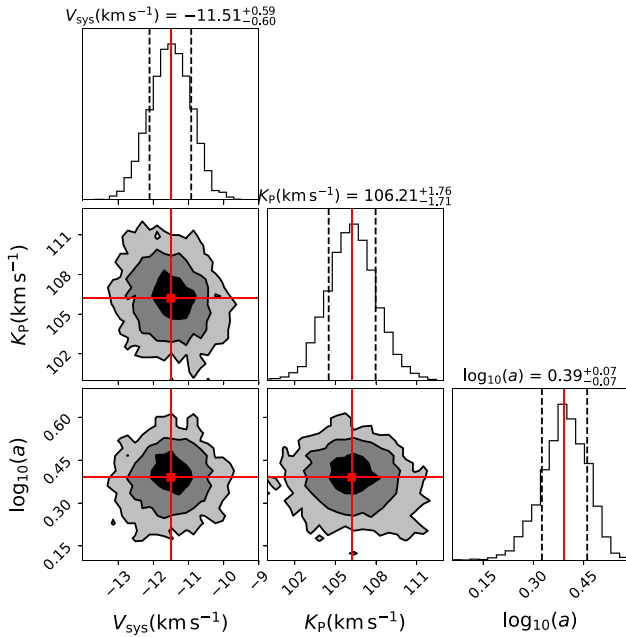
Figure 7. Posterior distributions of the orbital semi-amplitude (K_P), the systemic velocity (V_{sys}) and the logarithm of the atmospheric scaling parameter [$\log_{10}(a)$] retrieved from the best-fitting model. The median values for each parameter are given by the solid red lines on the corner plots and histograms. The black dashed lines on the histograms show the 0.16 and 0.84 quantiles. The filled-in contours show the 1, 2, and 3σ regions (darkest to lightest shades, respectively). The retrieved posteriors shows a constructive, co-added signal is retrieved from the best-fitting atmospheric model.

5.3 MCMC analysis

We performed an MCMC analysis on this data-set with the best-fitting model from the opacity model grid, i.e. a pure water model with a VMR = 10^{-3} and a T-p profile of $(T_1, p_1) = (1800\,\mathrm{K}, 1\,\mathrm{bar})$ and $(T_2, p_2) = (1200\,\mathrm{K}, 10^{-3}\,\mathrm{bars})$. For the combined four-night data-set of 560 spectra, it was too computationally expensive ($\sim$82 s per step pooled over 40 processes) to viably explore all of the model grid, and thus only the best-fitting model was used here. For this analysis, we used the emcee (Foreman-Mackey et al. 2013) package with 12 walkers and a chain-length of 1000, resulting in a total chain length of 12 000 points. We fit for a two-parameter circular orbital solution (K_P and V_{sys}) and an atmospheric scaling parameter [$\log_{10}(a)$] which allows the line strengths to vary. We allowed the MCMC to sample from uniform distributions of the prior parameters with ranges of, $-60 < V_{sys}\,(\mathrm{km\,s^{-1}}) < 60$, $0 < K_P\,(\mathrm{km\,s^{-1}}) < 200$ and $-2.0 < \log_{10}(a) < 2.0$. We initialize each Markov chain with the parameters set to the expected literature values, i.e. $K_P = 110\,\mathrm{km\,s^{-1}}$, $V_{sys} = -16\,\mathrm{km\,s^{-1}}$, and $\log_{10}(a) = 0$. The evolution of each chain is also driven by the CC-log (L) mapping as described in Section 4.1.

In Fig. 7, we show the retrieved posterior distributions from the combined set of spectra. It is clear that all of the spectra do co-add constructively to converge on to a single orbital solution from the water detection with $K_P = 106.21^{+1.76}_{-1.71}$ and $V_{sys} = -11.51^{+0.59}_{-0.60}$. The retrieved scaling factor of $\log_{10}(a) = 0.39$ suggests that the observed water lines are $\sim$2.5 × deeper compared to the best-fitting model. For these retrieved parameters, we determine a detection significance of 5.5 σ when comparing the log (L) values of the best-fitting water model and a featureless spectrum (i.e. a blackbody spectrum).

To test whether there is any phase dependence on the retrieved atmospheric parameters, we have split the data into pre- (i.e. all spectra
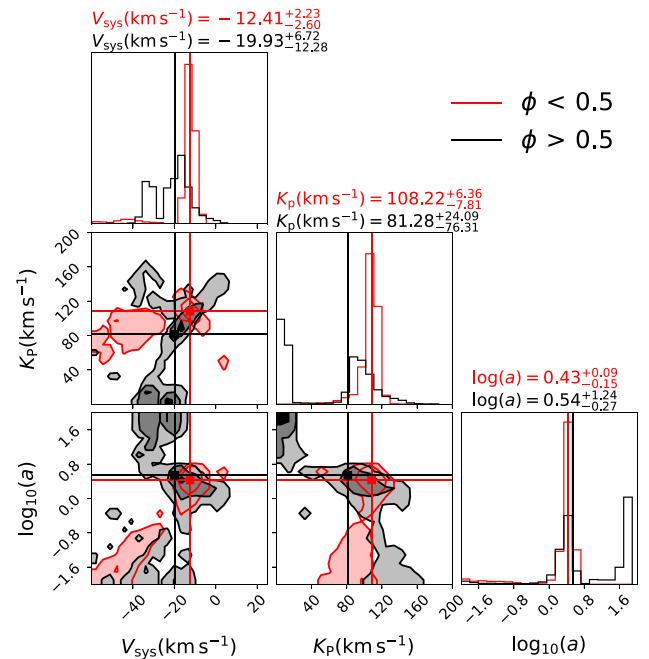

Figure 8. Same as Fig. 7 but the data-set has been split into pre- (red) and post-superior (black) conjunction. The red and black solid lines show the median values of the pre- and post-superior conjunction data-set, respectively.

Table 4. Retrieved posterior values of τ Boötis b from a MCMC analysis with the best-fitting atmospheric model.

Data-set	V_{sys} (km s^{-1})	K_P (km s^{-1})	$\log_{10}(a)$
$\phi < 0.5$	$-12.41^{+2.23}_{-2.60}$	$108.22^{+6.36}_{-7.81}$	$0.43^{+0.09}_{-0.15}$
$\phi > 0.5$	$-19.93^{+6.72}_{-12.28}$	$81.28^{+24.09}_{-76.31}$	$0.54^{+1.24}_{-0.27}$
Combined	$-11.51^{+0.59}_{-0.60}$	$106.21^{+1.76}_{-1.71}$	0.39 ± 0.07

$\phi < 0.5$, i.e. night 4 only) and post- the superior conjunction (all spectra $\phi > 0.5$). For our data-set with only the four nights used, the pre- and post-superior conjunction spectra include a phase coverage of $\phi = 0.420$–0.481 and 0.511–0.599, respectively. In Fig. 8, we show a corner plot of the posteriors from the MCMC using only the pre-superior conjunction data in red, and the post-superior conjunction data in black. In Table 4, we show the retrieved parameters from the two data-sets. It can be seen that the spectra at pre-superior conjunction provide a much more convincing detection of water with tighter constraints retrieved from the posterior distributions, which gives a detection significance of water at 4.6 σ for night 4 alone. The retrieved posteriors for the two data-sets do, however, show some overlap at the 2 σ level at the expected radial velocity of τ Boötis b suggesting a weak agreement between the two data-sets. However, neither data-set converges to a single solution with the post-superior conjunction posteriors, in particular, showing a double peak in the distribution. This indicates that the detection of water is weak in the individual nights and that we need to co-add the full data-set of 560 spectra in order to converge on to a single solution from the water detection. The failure of convergence for both these data-sets means that we are unable to constrain the individual atmospheric scaling factors therefore we find no evidence for a phase dependence on the observed water signature from τ Boötis b.

6 DISCUSSION

Using the CARMENES high-resolution instrument, we detect a signature of water in absorption in the day-side spectrum from the non-transiting planet τ Boötis b. Co-adding all five nights of spectra, we determine an orbital solution with a $K_P = 106.21^{+1.76}_{-1.71}\,\mathrm{km\,s^{-1}}$ and a $V_{sys} = -11.51^{+0.59}_{-0.60}\,\mathrm{km\,s^{-1}}$. Using a stellar mass of $1.35 \pm 0.03\,\mathrm{M_\odot}$ (Takeda et al. 2007) and a radial velocity of $468.42 \pm 2.09\,\mathrm{m\,s^{-1}}$ (Justesen & Albrecht 2019), we derive a planetary mass of $M_P = 6.24^{+0.17}_{-0.18}\,\mathrm{M_J}$. Furthermore, we are able to derive an inclination of $i = 41.6^{+1.0}_{-0.9}$ degrees, determined from a planet separation of $0.04869^{+0.00039}_{-0.00040}$ au (Rosenthal et al. 2021), which is consistent with those determined in Lockwood et al. (2014) and Pelletier et al. (2021). Despite having $\sim 3 \times$ weaker constraints on K_P, we are able to match the constraints on the inclination and the mass of the planet due to the recent improvement on the constraints on the measured semimajor axis (Rosenthal et al. 2021) than what was used in Pelletier et al. (2021). This highlights the importance of regular improvements to the system parameters on planets with large monitoring radial velocity surveys, such as in Rosenthal et al. (2021), which are essential for retrieving precise orbital parameters in high-resolution atmospheric studies.

We find that the best-fitting model requires a water abundance of VMR $= 10^{-3}$, which is consistent with solar, and a non-inverted $T - p$ profile of 1800-1200 K over a pressure range of $1\text{-}10^{-3}$ bars. We find no evidence for the presence of any further minor species from this analysis. We also split the spectra into a pre- and post-superior conjunction and we find that there is only weak evidence for a detection of water in the post-dayside spectra, this is likely due to the poorer observing conditions for those nights (see Section 2) and perhaps an overall weak signal from the atmosphere. This may also hint at a phase dependence on the water signal from τ Boötis b, however, we find no evidence for this in these spectra. For this model, we also retrieve a scaling factor of $\log_{10}(a) = 0.39$ indicating that the water depths are $\sim 2.5 \times$ deeper than the modelled spectral lines. As this planet is at an $\sim 41.6°$ inclination, we are viewing a mixture of the hot day-side and the cooler night-side of the atmosphere. Due to the absence of external irradiation from the host star on the night-side, it is expected that the lapse rate is steeper compared to the day-side, resulting in steeper spectral lines (e.g. de Kok et al. 2014). As water is expected to be formed on both sides (Madhusudhan 2012), it is likely that our scaling factor is trying to compensate for the differences in line depths over the two-temperature regions.

6.1 The orbital parameters of non-transiting planets

Assuming a circular orbit for τ Boötis b, we retrieve a systemic velocity of $-11.51^{+0.59}_{-0.60}\,\mathrm{km\,s^{-1}}$, which is significantly shifted from the literature value of $-16.9 \pm 0.3\,\mathrm{km\,s^{-1}}$ (Gaia Collaboration et al. 2018). However, these observations were taken a few years apart from those from the Gaia data release 2 survey therefore according to the radial velocity solution of τ Boötis A in Justesen & Albrecht (2019), it is likely that the systemic velocity has shifted further to over $-17\,\mathrm{km\,s^{-1}}$. This is due to the systems M-dwarf companion, τ Boötis B, approaching periastron, which will impart an additional radial velocity shift of τ Boötis A from the planetary signal [see fig. 3 in Justesen & Albrecht (2019)]. The analysis by Pelletier et al. (2021) also recovers a significantly shifted velocity of $V_{sys} = -15.4 \pm 0.2\,\mathrm{km\,s^{-1}}$. This shift in our V_{sys} can be partially explained by assuming an eccentric solution from Borsa et al. (2015) ($e = 0.011 \pm 0.006$ and $\omega = 113.4° \pm 32.2°$), as used in Pelletier et al. (2021). If we used this eccentric solution on the phases calculated

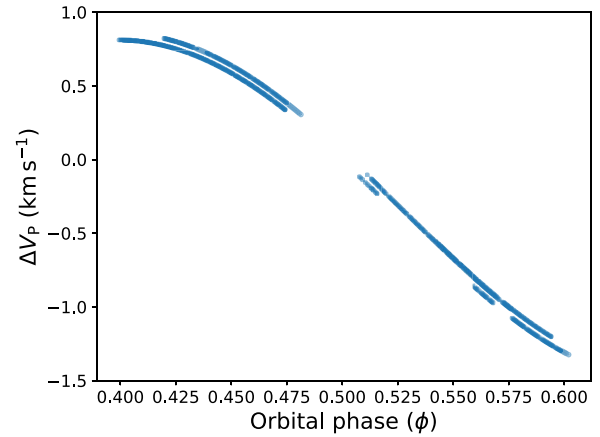**Figure 9.** Difference in the radial velocity of τ Boötis b in $\mathrm{km\,s^{-1}}$ between the eccentric and circular orbital solutions. The magnitude of these velocity shifts does not explain the $\sim 7\,\mathrm{km\,s^{-1}}$ shift we retrieve from the systemic velocity.

using the updated T_0 and the orbital period from Justesen & Albrecht (2019), this would result in a shift in the velocity by $\sim -3\,\mathrm{km\,s^{-1}}$. If we do adopt the full orbital solution from Borsa et al. (2015) (including their retrieved T_0 and orbital period), then we get at most a shift in the planetary velocity by $\pm 1.5\,\mathrm{km\,s^{-1}}$, as shown in Fig. 9, this also does not resolve the discrepancy in the retrieved systemic velocity. However, the time of periastron (T_0) for the eccentric solution obtained in Borsa et al. (2015) is highly uncertain (± 0.3 in BJD) and could therefore result in an even greater shift of several $\mathrm{km\,s^{-1}}$ for the systemic velocity. Due to this uncertainty in time of periastron from the eccentric orbital solution, we do not implement this solution into our analysis and instead adopted the circular orbital solution from Justesen & Albrecht (2019). However, this does show that even a relatively small eccentricity can lead to significantly shifted planetary velocities of up to several $\mathrm{km\,s^{-1}}$ if the eccentric solution is highly uncertain. This was also true in the high-resolution characterization of the non-transiting planet 51 Pegasi b (Birkby et al. 2017), which needed to invoke a large shift in the time of periastron by $\Delta T_0 = 0.07$ in days for the circular orbital solution to match the observed water signal to that of the observed systemic velocity. This highlights the importance of retrieving and regularly updating precise orbital solutions with long-period radial velocity surveys in order to accurately constrain the planetary velocities with follow-up high-resolution atmospheric characterization, particularly for non-transiting systems. Further observations of τ Boötis b are therefore necessary to try to explain these apparent discrepancies in the characterization of this system.

6.2 Comparison with previous analyses of the atmosphere of τ Boötis b

Our analysis confirms the detection of water in the L band by Lockwood et al. (2014) who used the NIRSPEC instrument at the Keck Observatory. Curiously, our water detection is in stark contrast to the results in Pelletier et al. (2021) who find only a 3σ upper limit on the water abundance at a VMR $= 10^{-5.66}$ with a full atmospheric retrieval. This analysis detects the presence of water at a VMR $= 10^{-3}$, which is preferred over a VMR of 10^{-4} and 10^{-5} by 3.8σ and 5.0σ, respectively, with a non-inverted $T - p$ profile of 1800–1200 K over $1\text{-}10^{-3}$ bars. However, we do emphasize that the molecular abundance and $T - p$ profile in our analysis will

be partially correlated, and therefore a full atmospheric retrieval is needed on these spectra to give an accurate comparison of the water abundance with Pelletier et al. (2021).

The analysis by Pelletier et al. (2021) observed the day-side thermal emission of τ Boötis with the SPIRou ($R = 70\,000$) instrument over five nights of data spanning a similar phase coverage to this analysis. Due to the wider wavelength coverage of SPIRou ($\lambda = 0.95$–$2.50\,\mu$m) compared with CARMENES, it was expected that SPIRou should have observed the day-side emission of CO and H_2O simultaneously; however, only CO was detected in these spectra. As CARMENES does not cover the strong 2-0 R-branch absorption feature at $\sim 2.3\,\mu$m, we cannot observe these two molecules simultaneously. Brogi et al. (2012) also only detected the absorption features from CO from τ Boötis b, however, their observations were taken with a narrow wavelength range to cover the $2.3\,\mu$m feature therefore it is likely that the strong CO lines obscured the H_2O lines to be observable. It could also be the case that Pelletier et al. (2021) suffer from the same sort of behaviour as the strong CO lines mask the weaker water absorption features in the atmosphere of τ Boötis b, although this explanation is perhaps unsatisfactory as they deduce that the spectra are sensitive to the presence of water from τ Boötis b with their injection and retrieval tests. We are only able to detect a convincing signal from water our fourth night, which has a phase coverage of ~ 0.42–0.48. The corresponding night in Pelletier et al. (2021) that covered the same phase range suffered from poorer seeing and a slightly lower SNR than the rest of their nights. If it is the case that there is some phase dependence and the signal is far stronger within this phase range, then this could explain the absence of water in their analysis. However, if there is a constant abundance of water across the orbit, then it should be observable in the spectra from Pelletier et al. (2021). The final explanation for the discrepancies in the water detection could be due to differing line lists used in each analysis. In this analysis, we use the HITEMP water opacities calculated using the BT2 line list (Barber et al. 2006), as also used in the analysis by Pelletier et al. (2021) therefore we rule out the possibility that line lists are the cause of the discrepancies between our two analyses.

7 CONCLUSIONS

With ~ 21 h of observations over five nights at high resolution with the CARMENES spectrograph, we unambiguously detect the presence of absorption features from water vapour through thermal emission from the atmosphere of τ Boötis b. We searched for, HCN, CH_4, NH_3, and C_2H_2 but found no evidence for these minor species. Using a grid of models, we found that these spectra prefer a high abundance of water (VMR $= 10^{-3}$), which is significantly preferred over models with lower abundances by $\leqslant 3.8\,\sigma$. However, a full atmospheric retrieval would be needed to provide confidence intervals on the retrieved abundance for water. On individual nights, we find that the predominant signal from water originates from night 4. In effect this means that we were able to detect the signature for water in only ~ 5 h of observation in a phase coverage of $\phi = 0.42$–0.48. We find no strong evidence for any phase variability in the water signal over the phase coverage of our observations in our analysis when we split the data into pre- and post-dayside observations. We retrieve an atmospheric scaling factor of $\log(a) = 0.39$, which suggests the model is underestimating the depth of the water lines by $2.5\times$; however, this value could be dominated by the mixing of day- and night-side emissions from the atmosphere due to the $\sim 41°$ inclination of the planet. This analysis is in agreement with the *L*-band detection from Lockwood et al. (2014) but is strongly in disagreement with

the more recent analysis from Pelletier et al. (2021), which finds no evidence for water in the atmosphere of τ Boötis b.

Retrieving an accurate abundance is crucial if we are to understand the C/O ratio (e.g. Madhusudhan 2012), metallicity (e.g. Moses et al. 2013) and the physical structure (e.g. Seager & Sasselov 1998) of hot Jupiter atmospheres. It is likely that further analysis of the atmosphere of τ Boötis b is needed in order to resolve some of the discrepancies that remain over the detection of water. It is also likely that a full atmospheric retrieval is necessary in order to delve deeper into the water features that we have detected from the emergent spectra from τ Boötis b in the *Y*, *J*, and *H* bands.

DATA AVAILABILITY

The data used in this article are available at http://caha.sdc.cab.int a-csic.es/calto/ with project ID's: 271375, 272952, 294943, 295318 ,and 296169.

REFERENCES

Astropy Collaboration et al., 2013, A&A, 558, A33
Astropy Collaboration et al., 2018, AJ, 156, 123

Baliunas S. L., Henry G. W., Donahue R. A., Fekel F. C., Soon W. H., 1997, ApJ, 474, L119

Barber R. J., Strange J. K., Hill C., Polyansky O. L., Mellau G. C., Yurchenko S. N., Tennyson J., 2014, MNRAS, 437, 1828

Barber R. J., Tennyson J., Harris G. J., Tolchenov R. N., 2006, MNRAS, 368, 1087

Bell T. J., Cowan N. B., 2018, ApJ, 857, L20

Beltz H., Rauscher E., Brogi M., Kempton E. M. R., 2021, AJ, 161, 1

Birkby J. L., de Kok R. J., Brogi M., Schwarz H., Snellen I. A. G., 2017, AJ, 153, 138

Boley K. M., Wang J., Zinn J. C., Collins K. A., Collins K. I., Gan T., Li T. S., 2021, AJ, 162, 85

Borsa F. et al., 2015, A&A, 578, A64

Brogi M., de Kok R. J., Birkby J. L., Schwarz H., Snellen I. A. G., 2014, A&A, 565, A124

Brogi M., Giacobbe P., Guilluy G., de Kok R. J., Sozzetti A., Mancini L., Bonomo A. S., 2018, A&A, 615, A16

Brogi M., Line M. R., 2019, AJ, 157, 114

Brogi M., Line M., Bean J., Désert J. M., Schwarz H., 2017, ApJ, 839, L2

Brogi M., Snellen I. A. G., de Kok R. J., Albrecht S., Birkby J., de Mooij E. J. W., 2012, Nature, 486, 502

Butler R. P., Marcy G. W., Williams E., Hauser H., Shirts P., 1997, ApJ, 474, L115

Caballero J. A. et al., 2016, Proc. SPIE, 9910, 99100E

Cabot S. H. C., Madhusudhan N., Hawker G. A., Gandhi S., 2019, MNRAS, 482, 4422

Charbonneau D., Noyes R. W., Korzennik S. G., Nisenson P., Jha S., Vogt S. S., Kibrick R. I., 1999, ApJ, 522, L145

Chubb K. L., Tennyson J., Yurchenko S. N., 2020, MNRAS, 493, 1531

Coles P. A., Yurchenko S. N., Tennyson J., 2019, MNRAS, 490, 4638

Collier Cameron A., Horne K., Penny A., James D., 1999, Nature, 402, 751

Cowan N. B., Agol E., 2011, ApJ, 729, 54

de Kok R. J., Birkby J., Brogi M., Schwarz H., Albrecht S., de Mooij E. J. W., Snellen I. A. G., 2014, A&A, 561, A150

Ehrenreich D. et al., 2020, Nature, 580, 597

Foreman-Mackey D., 2016, J. Open Source Softw., 1, 24

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

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

Gandhi S. et al., 2020a, MNRAS, 495, 224

Gandhi S., Brogi M., Webb R. K., 2020b, MNRAS, 498, 194

Gandhi S., Madhusudhan N., 2017, MNRAS, 472, 2334

Giacobbe P. et al., 2021, Nature, 592, 205

Gibson N. P. et al., 2020, MNRAS, 493, 2215

Gray R. O., Napier M. G., Winkler L. I., 2001, AJ, 121, 2148

Hargreaves R. J., Gordon I. E., Rey M., Nikitin A. V., Tyuterev V. G., Kochanov R. V., Rothman L. S., 2020, ApJS, 247, 55

Harris C. R. et al., 2020, Nature, 585, 357

Harris G. J., Tennyson J., Kaminsky B. M., Pavlenko Y. V., Jones H. R. A., 2006, MNRAS, 367, 400

Hawker G. A., Madhusudhan N., Cabot S. H. C., Gandhi S., 2018, ApJ, 863, L11

Hoeijmakers H. J., Snellen I. A. G., van Terwisga S. E., 2018, A&A, 610, A47

Hunter J. D., 2007, Comput. Sci. Eng., 9, 90

Joy A. H., Abt H. A., 1974, ApJS, 28, 1

Justesen A. B., Albrecht S., 2019, A&A, 625, A59

Kaeufl H.-U. et al., 2004, Proc. SPIE, 5492, 1218

Kesseli A. Y., Snellen I. A. G., 2021, ApJ, 908, L17

Knutson H. A. et al., 2007, Nature, 447, 183

Kreidberg L. et al., 2018, AJ, 156, 17

Leigh C., Collier Cameron A., Horne K., Penny A., James D., 2003, MNRAS, 344, 1271

Line M. R. et al., 2021, Nature, 598, 580

Lockwood A. C., Johnson J. A., Bender C. F., Carr J. S., Barman T., Richert A. J. W., Blake G. A., 2014, ApJ, 783, L29

Louden T., Wheatley P. J., 2015, ApJ, 814, L24

Madhusudhan N., 2012, ApJ, 758, 36

McLean I. S. et al.,1995, Proc. SPIE, 2475, 350

Moses J. I. et al., 2013, ApJ, 777, 34

Nugroho S. K., Gibson N. P., de Mooij E. J. W., Herman M. K., Watson C. A., Kawahara H., Merritt S. R., 2020, ApJ, 898, L31

Parmentier V. et al., 2018, A&A, 617, A110

Pelletier S. et al., 2021, AJ, 162, 73

Petigura E. A. et al., 2018, AJ, 155, 89

Quirrenbach A. et al., 2014, Proc. SPIE, 9147, 531

Richard C. et al., 2012, J. Quant. Spectrosc. Radiat. Transfer, 113, 1276

Rodler F., Kürster M., Henning T., 2010, A&A, 514, A23

Rodler F., Lopez-Morales M., Ribas I., 2012, ApJ, 753, L25

Rosenthal L. J. et al., 2021, ApJS, 255, 8

Rothman L. S. et al., 2010, J. Quant. Spectrosc. Radiat. Transfer, 111, 2139

Seager S., Sasselov D. D., 1998, ApJ, 502, L157

Skrutskie M. F. et al., 2006, AJ, 131, 1163

Stevenson K. B. et al., 2017, AJ, 153, 68

Takeda G., Ford E. B., Sills A., Rasio F. A., Fischer D. A., Valenti J. A., 2007, ApJS, 168, 297

van Belle G. T., von Braun K., 2009, ApJ, 694, 1085

Virtanen P. et al., 2020, Nat. Methods, 17, 261

Webb R. K., Brogi M., Gandhi S., Line M. R., Birkby J. L., Chubb K. L., Snellen I. A. G., Yurchenko S. N., 2020, MNRAS, 494, 108

Wilks S. S., 1938, Anna. Math. Stat., 9, 60

Wright J. T., Marcy G. W., Howard A. W., Johnson J. A., Morton T. D., Fischer D. A., 2012, AJ, 753, 160

Zhang Y., Snellen I. A. G., Mollière P., Alonso-Floriano F. J., Webb R. K., Brogi M., Wyttenbach A., 2020, A&A, 641, A161

APPENDIX A: MCMC ANALYSIS ON THE INDIVIDUAL NIGHTS

It is clear from Fig. 8 that each night's set of spectra does not contribute equally to the overall water signal from τ Boötis b in Fig. A1. Here, we repeat the analysis from Section 5.3 but instead of combining the nights together, we run an MCMC on the individual nights to assess the contribution from each night. As the pre-superior
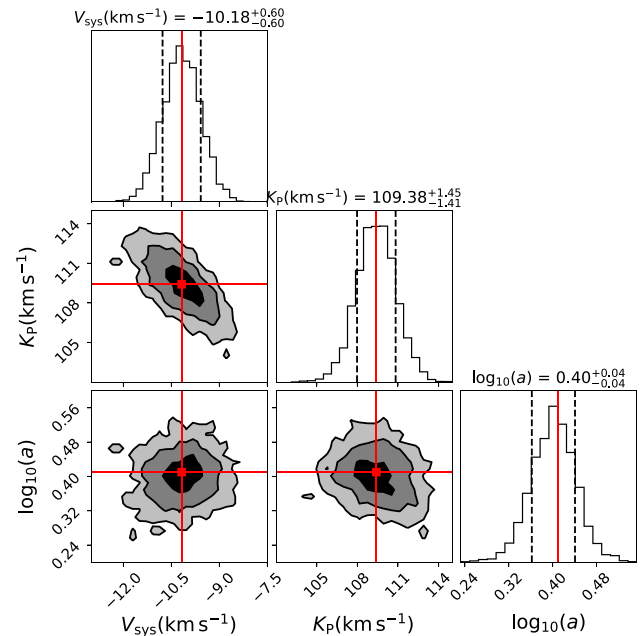

Figure A1. Same as Figs 7 and 8 but with only spectra from night 4 with the best-fitting combined water model. There is a clear signal from τ Boötis b with the retrieved parameters stated above each histogram.

Table A1. The median values retrieved from an MCMC analysis for the orbital solution and an atmospheric scaling parameter for each night in the analysis.

Night	Retrieved parameters (median)		
	V_{sys} (km s^{-1})	K_P (km s^{-1})	$\log(a)$
1	$-15.57^{+21.82}_{-2.70}$	$50.69^{+1.19}_{-0.55}$	$1.71^{+0.06}_{-0.08}$
2	21.44 ± 0.19	$50.03^{+0.05}_{-0.02}$	1.80 ± 0.01
3	$-10.56^{+10.86}_{-12.85}$	$102.25^{+37.92}_{-43.29}$	$0.27^{+0.26}_{-0.46}$
4	$-12.41^{+2.23}_{-2.60}$	$108.22^{+6.36}_{-7.81}$	$0.43^{+0.09}_{-0.15}$
5	$-15.56^{+5.67}_{-13.73}$	$110.78^{+8.00}_{-13.95}$	$1.36^{+0.60}_{-1.68}$

conjunction data only included night 4, the retrieved parameters will be the same as shown in Fig. 8.

In Table A1, we show the median values retrieved from an MCMC analysis on each night. As described in Section 5.3, the MCMC is set-up with 12 individual walkers, each with a chain length of 1000. All of the nights failed to converge to a single solution; however, night 4 performed the best, with the main peak in the posterior distribution at the expected radial velocity of the planet. Nights 1 and 2 performed worse than the other nights as those chains settled on to the lower limit of the K_P prior. This is likely an indication that those spectra suffer from some residual tellurics despite the lack of visible highly deviant wavelength columns. Night 5 shows a potentially weak signal for water at the orbital solution of τ Boötis b, however, the MCMC did not converge to a single solution with a particularly wide posterior on the scaling factor $\log_{10}(a)$.

The lack of a water detection in every night except for night 4 is not one of surprise given the variable observing conditions between each night at the Calar Alto site. Given the variable SNR for nights 1, 2, 3, and 5, it is highly likely that these suffered from cloudy conditions intermittently throughout these nights. The variable SNR for nights 1 and 2 combined with the high humidity, which consistently exceeded >85 per cent during the nights, is likely the reason why residual tellurics overpower the weak water signal in this MCMC analysis. Although we can only detect the water signal from the night 4 spectra, it is nevertheless the case that with the addition of nights together helps the water detection by tightening the confidence intervals on the retrieved parameters, which is seen in Figs 8 and A1.

APPENDIX B: CROSS-CORRELATION TO LOG-LIKELIHOOD ANALYSIS INCLUDING ALL NIGHTS

As shown in the main data analysis (see Fig. 3), night 2 suffers from strong telluric residuals despite the telluric removal steps and the additional masking of highly deviant spectral channels. However, here, we show how the inclusion of the observations from night 2 (261 additional spectra) affects the results of the analysis from the four other nights.

Fig. B1 shows all spectra correlated with a pure water spectrum as a function of phase and radial velocity shift from the telluric rest frame. The coloured dashed lines indicate the expected radial velocity trails of τ Boötis b during those observations. In Fig. B2, we show the $K_P - V_{sys}$ velocity map with the inclusion of night 2. It is clear that the telluric noise from these spectra overwhelms the planet's water signal, as seen in Fig. 4, with the retrieved $\log(L_{max})$ (shown as the
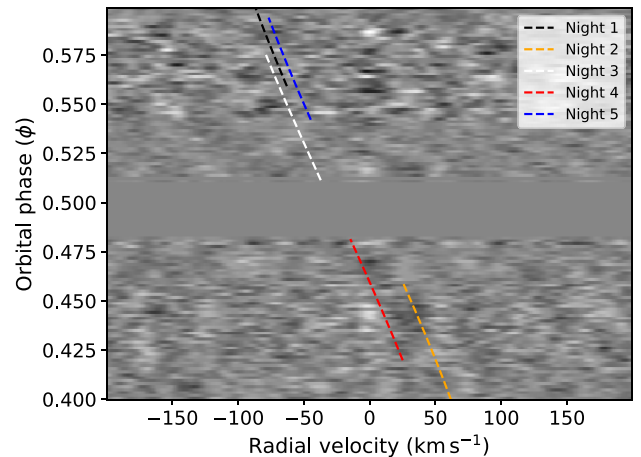

Figure B1. Same as Fig. 3 but with all nights included binned in phase with a resolution of $\Delta\phi = 0.0015$. The black, orange, white, red, and blue dashed lines show the expected radial velocity of τ Boötis b for nights 1–5, respectively.
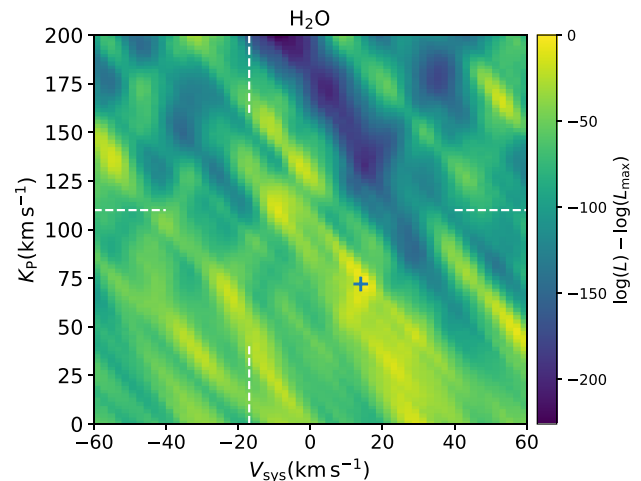

Figure B2. Same as Fig. 4 but with the inclusion of night 2. The white dashed lines show the expected position of τ Boötis b. The blue cross shows the position of the maximum $\log(L)$, which no longer appears at the position of the detection in Fig. 4.

blue cross) shifting beyond the expected orbital velocity of the planet. However, as shown in Fig. A1, the MCMC still converges on to a local maximum at the expected orbital solution of the planet. Although this is a case of the MCMC algorithm converging on to a local maximum before exploring the wider parameter space, it nevertheless shows that there is still a detectable signal of water from the atmosphere in the local vicinity of expected orbital solution. If we use the retrieved parameters from the MCMC analysis, we find a marginal preference of 1.1σ for the inclusion of HCN in the atmospheric models at an abundance of VMR $= 10^{-4}$. In Fig. B3, we show the updated abundance constraints on HCN and C_2H_2 for the inclusion of night 2 spectra. With these results, we can lower the 3σ upper limits on these species at log(VMR) ≈ -2.0 and ≈ -4.0 for HCN and C_2H_2, respectively.
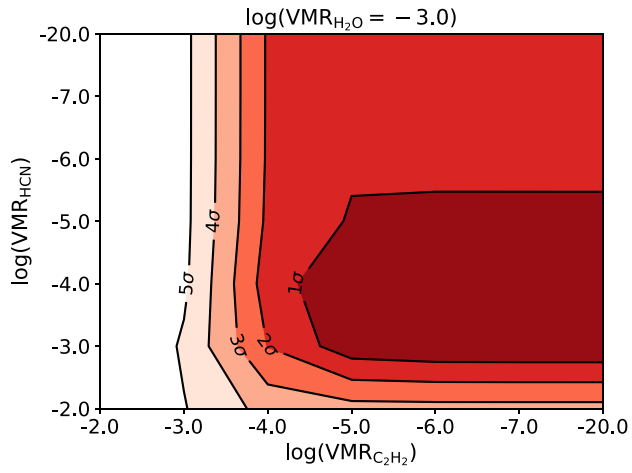
Figure B3. Same as Fig. 6 but with the addition of night 2. With the inclusion of these spectra, we see a marginal preference for the addition of HCN at a VMR $= 10^{-4}$ in the best-fitting atmospheric model.

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