A&A 606, A144 (2017)
DOI: 10.1051/0004-6361/201730814

Astronomy & Astrophysics

Search for water vapor in the high-resolution transmission spectrum of HD 189733b in the visible

R. Allart[1], C. Lovis[1], L. Pino[1,2], A. Wyttenbach[1], D. Ehrenreich[1], and F. Pepe[1]

[1] Observatoire Astronomique de l'Université de Genève, Université de Genève, 51 chemin des Maillettes, 1290 Versoix, Switzerland
e-mail: romain.allart@unige.ch
[2] Dipartimento di Fisica e Astronomia "Galileo Galilei", Univ. di Padova, Vicolo dell'Osservatorio 3, 35122 Padova, Italy

Received 17 March 2017 / Accepted 25 May 2017

ABSTRACT

Context. Ground-based telescopes equipped with state-of-the-art spectrographs are able to obtain high-resolution transmission and emission spectra of exoplanets that probe the structure and composition of their atmospheres. Various atomic and molecular species, such as Na, CO, H_2O have been already detected in a number of hot Jupiters. Molecular species have been observed only in the near-infrared while atomic species have been observed in the visible. In particular, the detection and abundance determination of water vapor bring important constraints to the planet formation process.

Aims. We aim to search for water vapor in the atmosphere of the exoplanet HD 189733b using a high-resolution transmission spectrum in the visible obtained with HARPS.

Methods. We used the atmospheric transmission code `Molecfit` to correct for telluric absorption features. Then we computed the high-resolution transmission spectrum of the planet using three transit datasets. We finally searched for water vapor absorption in the water band around 6500 Å using a cross-correlation technique that combines the signal of 600–900 individual lines.

Results. Telluric features are corrected to the noise level. We place a 5-σ upper limit of 100 ppm on the strength of the 6500 Å water vapor band. The 1-σ precision of 20 ppm on the transmission spectrum demonstrates that space-like sensitivity can be achieved from the ground, even for a molecule that is a strong telluric absorber.

Conclusions. This approach opens new possibilites for the detection of various atomic and molecular species with future instruments such as ESPRESSO at the VLT. Extrapolating from our results, we show that only one transit with ESPRESSO would be sufficient to detect water vapor on HD 189733b-like hot Jupiter with a cloud-free atmosphere. Upcoming near-IR spectrographs will be even more efficient and sensitive to a wider range of molecular species. Moreover, the detection of the same molecular species in different bands (e.g., visible and IR) is key to constrain the structure and composition of the atmosphere, such as the presence of Rayleigh scattering or aerosols (cloud and/or hazes).

Key words. planets and satellites: atmospheres – planets and satellites: individual: HD 189733b – methods: observational – techniques: spectroscopic

1. Introduction

Over the past two decades, the field of exoplanets has expanded on a large scale with the development of numerous ground-based and space missions to detect and determine mass and radius of exoplanets with the radial velocity and transit techniques. Two of the most studied exoplanets, HD 209458b (Charbonneau et al. 2000) and HD 189733b (Bouchy et al. 2005), have been detected both by radial velocity and transit. These two exoplanets have a bright host star and are hot Jupiters, two characteristics that make them amenable to in-depth characterization. The aim of the characterization of an exoplanet is not only to determine its basic physical parameters, but also to determine its bulk composition and the composition of its atmosphere. To do so, transit techniques are available, either using the primary eclipse (for the transmission spectrum) or the secondary eclipse (for the thermal or reflected light). The first detection of atomic species was made on the transmission spectrum of HD 209458b by Charbonneau et al. (2002), who detected the Na doublet in the visible with the STIS spectrograph on board the *Hubble* Space Telescope (HST). The first detection with a ground-based telescope was made by Redfield et al. (2008) on HD 189733b with the Na doublet detection, then followed by Snellen et al. (2008) for the same species in the atmosphere of HD 209458b. These first ground-based sodium detections were made with slit spectrographs. Previous studies have emphasised the potential of stabilised and fiber-fed spectrographs for studying exoplanet atmospheres in the optical (Vidal-Madjar et al. 2010; Arnold et al. 2014); this potential was demonstrated with HARPS for a hot gas giant by Wyttenbach et al. (2015). The presence of CO was detected in the infrared high-resolution transmission spectrum of HD 209458b and of HD 189733b by Snellen et al. (2010) and Brogi et al. (2016) using the CRIRES spectrograph at the Very Large Telescope (VLT). Deming et al. (2013) and McCullough et al. (2014) reveal the presence of H_2O in the infrared transmission spectrum of HD 209458b and HD 189733b with WFC3 instrument on HST. Today, tens of hot Jupiters have been characterized. Several studies indicate the presence of aerosols (clouds and/or hazes, e.g., Sing et al. 2016) through Rayleigh and Mie scattering signatures (e.g., Lecavelier Des Etangs et al. 2008; Sing et al. 2013; Sing et al. 2015), atmospheric evaporation (e.g., Lecavelier Des Etangs et al. 2010; Bourrier et al. 2013; Ehrenreich et al. 2015), temperature gradients within the atmosphere (e.g., Huitson et al. 2012; Wyttenbach et al. 2015; Heng et al. 2015), atmospheric circulation (e.g., Snellen et al. 2010; Kataria et al. 2016; Wyttenbach et al. 2015;

Table 1. Adopted physical and orbital parameters of HD 189733b.

Parameter	Symbol	Value	Reference
Stellar radius	R_*	$0.756 \pm 0.018\,R_\odot$	Torres et al. (2008)
Planet radius	R_p	$1.138 \pm 0.027\,R_J$	Torres et al. (2008)
White-light radius ratio	R_p/R_*	0.15617 ± 0.00011	Sing et al. (2011)
Stellar mass	M_*	$0.823 \pm 0.029\,M_\odot$	Triaud et al. (2009)
Planet mass	M_p	$1.138 \pm 0.027\,M_J$	Triaud et al. (2009)
Epoch of transit	T_0	$2\,454\,279.436714 \pm 0.000015$ BJD$_{tdb}$	Agol et al. (2010)
Duration of transit	T_{14}	0.07527 ± 0.00037 d	Triaud et al. (2009)
Orbital period	P	$2.21857567 \pm 0.00000015$ d	Torres et al. (2008)
Systemic velocity	γ	-2.2765 ± 0.0017 km s^{-1}	Boisse et al. (2009)
Semi-amplitude	K_*	200.56 ± 0.88 m s^{-1}	Boisse et al. (2009)

Louden & Wheatley 2015) and an enhanced C/O ratio (e.g., Madhusudhan et al. 2011; Moses et al. 2013; Kreidberg et al. 2015).

In this paper, we focus on HD 189733b (see Table 1 for the physical and orbital parameters used). Its host star, HD 189733, is an active, bright, metal-rich star of type K0V (V-band magnitude of 7.65). The exoplanet is supposed to be a tidally-locked hot Jupiter and exhibiting a blue color in the visible (Evans et al. 2013). Its atmosphere contains Na (Redfield et al. 2008; Huitson et al. 2012; Wyttenbach et al. 2015; Khalafinejad et al. 2017), H_2O (Birkby et al. 2013; McCullough et al. 2014; Brogi et al. 2016), CO (de Kok et al. 2013; Brogi et al. 2016) exhibits a Rayleigh scattering slope at blue wavelenghs (Pont et al. 2008; Lecavelier Des Etangs et al. 2008) likely caused by high-altitude hazes and loses hydrogen as shown by Ly-α absorption (Lecavelier Des Etangs et al. 2010; Lecavelier des Etangs et al. 2012; Bourrier et al. 2013).

In this study, we used data obtained with the HARPS high-resolution spectrograph ($\lambda/\Delta\lambda \sim 115\,000$, Mayor et al. 2003) to search for water vapor in the optical transmission spectrum of HD 189733b. The high stability of HARPS allow us to optimally co-add hundreds of spectra while its high resolution allows us to resolve each individual line in the water spectrum which can be used to build a cross-correlation function (CCF) concentrating all the available water signal. So far, no detection of H_2O was ever made in the visible for an exoplanet, although weak water signatures are expected to be present.

In Sect. 2 we describe the HARPS data used in this paper. In Sect. 3, we correct these spectra from telluric features with the ESO tool `Molecfit`. In Sect. 4, we derive the transmission spectrum of the planet using the same method as Wyttenbach et al. (2015, 2017) and describe the cross-correlation function technique to study water vapor. Section 5 shows the results which are then discussed in Sect. 6. We conclude in Sect. 7.

2. HARPS observations

HD 189733 was observed with HARPS mounted on the ESO 3.6-m telescope at La Silla Observatory, Chile, in the programs 072.C-0488, 079.C-0127 (PI: Mayor) and 079.C-0828 (PI: Lecavelier des Etangs). These three programs contain four transits measured in 2006 and 2007 which are already analyzed by several authors. These observations have yielded a number of important results: the study of the Rossiter-McLaughlin effect (Triaud et al. 2009), a 10-σ detection of the Na doublet in the planet atmosphere (Wyttenbach et al. 2015), winds circulating from the day side (hotter) to the night side (cooler) (Wyttenbach et al. 2015; Louden & Wheatley 2015), an atmospheric temperature gradient of 0.2–0.4 K km^{-1} (Wyttenbach et al. 2015; Heng et al. 2015), and a tentative of detection (2.5-σ) of Rayleigh scattering (Di Gloria et al. 2015). Even the properties of the stellar surface occulted by the planet can be retrieved (Collier Cameron et al. 2010; Cegla et al. 2016).

Table 2. Observation log of the different nights showing the total number of spectra and the number of in- and out-transit spectra.

Date	7 Sep. 2006	19 Jul. 2007	28 Aug. 2007
Total spectra	20	39	40
In-transit	11	19	19
Out-of-transit	9	20	21
t_{exp} [s]	600–900	300	300

Table 2 provides the log of the observations used in this paper. We note that an additional transit was observed on 29 July 2006 but only the first half of the transit was obtained due to bad meteorological conditions. Therefore, the transmission spectrum derived from this night is much more noisy than the other ones. As a consequence, we do not take it into account in the remainder of this paper.

The data reduction applied here was made with version 3.5 of the HARPS data reduction software (DRS). Spectra were extracted order by order (total of 72), flat-fielded using calibrations obtained at the beginning of the night, deblazed and wavelength calibrated. Finally, a one-dimensional spectrum from 3800 to 6900 Å with a step of 0.01 Å in the solar system barycentric rest frame was produced.

3. Correction of telluric contamination with `Molecfit`

Figure 1 shows the red part of a HARPS spectrum and the influence of telluric lines in our optical ground-based observations. As we can see from the close-up views, the depth of some water telluric lines is ~20%. As we will see in Sect. 6.2, the depth of the planetary water lines is expected to be 30–50 ppm in the visible, which implies that the telluric correction is a crucial step to probe the transmission spectrum of an exoplanet atmosphere.

To perform this correction, we used version 1.2.0 of `Molecfit` (Smette et al. 2015; Kausch et al. 2015), an ESO tool to correct telluric features in ground-based spectra. `Molecfit` uses a line-by-line radiative transfer model (LBLRTM) to create a telluric spectrum at a very high resolution ($\lambda/\Delta\lambda \sim 4\,000\,000$). The LBLRTM needs an atmospheric profile that describes temperature, pressure, humidity and abundance of molecular species as a function of altitude for a given observatory site at one particular time and at a given airmass. To create this atmospheric profile, `Molecfit` merged an atmospheric standard profile and

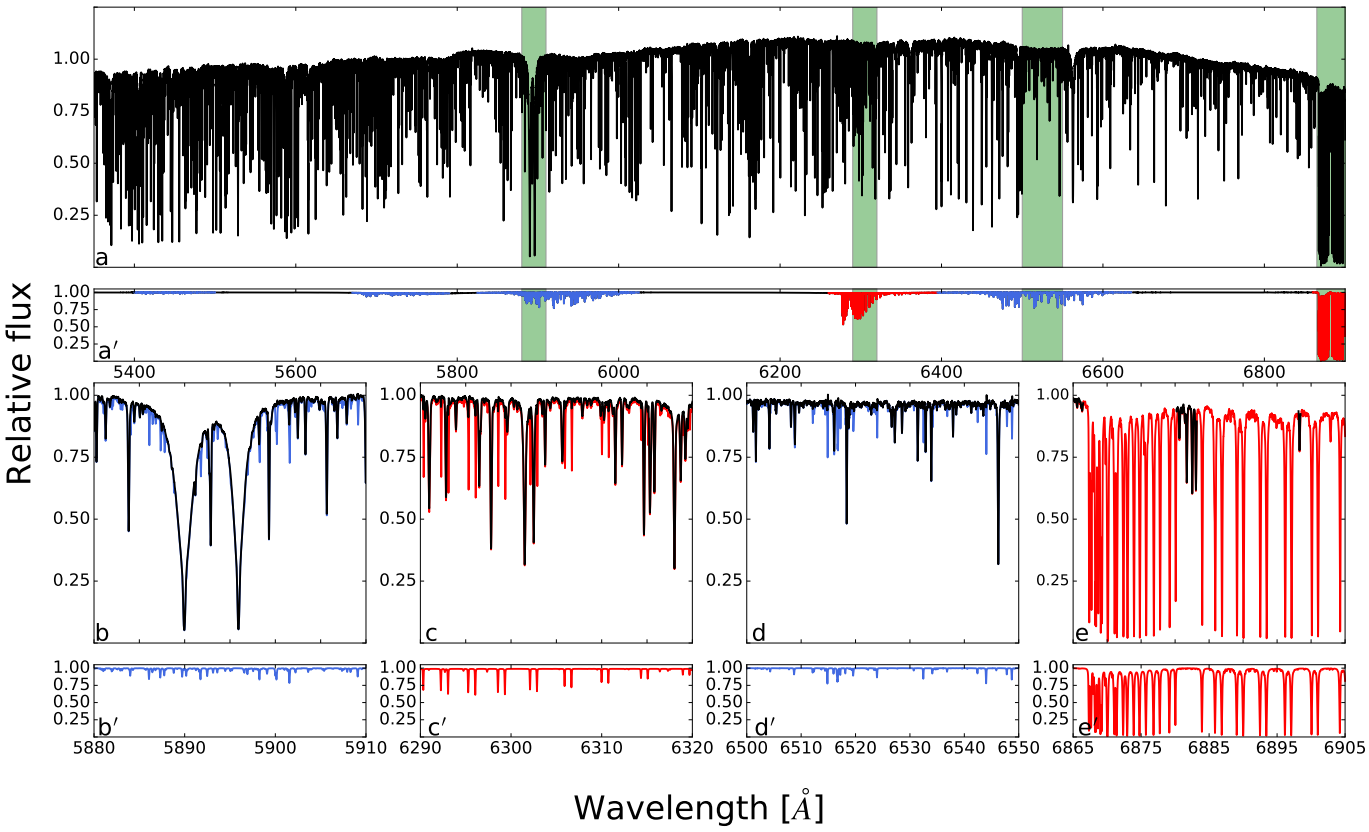

Fig. 1. Influence of telluric lines on the spectrum of HD 189733 obtained with HARPS on the red part of the CCD (5380 to 6900 Å) for the night of 19 July 2007. The *upper panel a–e* is the observed stellar spectrum with telluric contamination. *Panel a′–e′* is the best fit telluric model obtained with `Molecfit` with H_2O in blue and O_2 in red. *Panels b, b′*; *c, c′*; *d, d′*; and *e, e′* are close-up views of the green zones in *panel a, a′*. These contain the strongest telluric bands in the visible (from left to right: H_2O (203-000, 302-000, 321-000) band at ~5900 Å around the Na doublet, O_2 γ band at ~6200 Å, H_2O (311-000) band at ~6500 Å and O_2 B band at ~6900 Å).

a Global Data Assimilation System (GDAS) profile. The standard profile, which is provided by the Reference Forward Model (Remedios et al. 2001), describes pressure, temperature, molecular abundances (up to tens molecular species) as a function of altitude for a specific latitude (e.g., equatorial, mid-latitude, polar-latitude for day or night). GDAS profiles provided by the National Oceanic and Atmospheric Administration (NOAA), are dedicated to wheater forecast. They contained meteorological data set (pressure, temperature, relative humidity as a function of altitude) and are updated every three hours for specific locations. The resulting merging atmospheric profile can be described in two possible grids, a fixed grid and a natural grid. The first one describes the variation of temperature, pressure, humidity and abundance of H_2O and O_2 from 0 to 120 km with a fixed number of layers (50), while the second one is more precise with 100 to 150 layers. In this paper, we used the second grid which is more sensitive and yields a better telluric correction. Then the model spectrum is fitted to the observed spectrum by adjusting the continuum, the wavelength calibration and the instrumental resolution.

3.1. Adaptation of `Molecfit` to high-resolution visible spectra

For the first time, `Molecfit` is used on HARPS spectra, which are given in the solar system barycentric rest frame, while the modeled spectrum by `Molecfit` is given in the terrestrial rest frame. As a first step, we thus shifted the HARPS spectra into the terrestrial rest frame taking into account the Barycentric Earth Radial Velocity (BERV).

We then transformed the wavelength scale of the modeled spectrum from vacuum to air to match the observed spectrum. To optimize the correction of the telluric features, we decomposed the spectrum into about fifteen regions (part of them fall inside the green bands in Fig. 1). We chose these regions such as to have only strong lines for a single molecule (H_2O or O_2), a flat continuum, and no stellar features within them. Indeed, `Molecfit` does not model the stellar spectrum, and the fitting algorithm (Levenberg-Marquardt) is very sensitive to any stellar feature. It is thus necessary to be very rigorous on the selection of spectral regions.

We refer to Smette et al. (2015) for the detailed description of the free parameters of `Molecfit`. The goal of the fitting process is to adjust the continuum, the wavelength scale and the instrumental resolution for each fitted telluric band. The Levenberg-Marquardt χ^2 convergence criterion and the parameter convergence criterion were both set to 10^{-9}. The continuum was adjusted with a third-degree polynomial. The wavelength calibration was made with a Chebyschev second-degree polynomial. Finally, the instrumental profile was assumed to be a Gaussian with a FWHM of 4.5 pixels. Appendix A shows all the parameters used by `Molecfit`.

To correct the entire spectral range of interest, we used the Calctrans tool provided with `Molecfit` . It takes the best-fit parameters from the `Molecfit` optimization and applies them to the entire spectrum. This operation was done for every individual

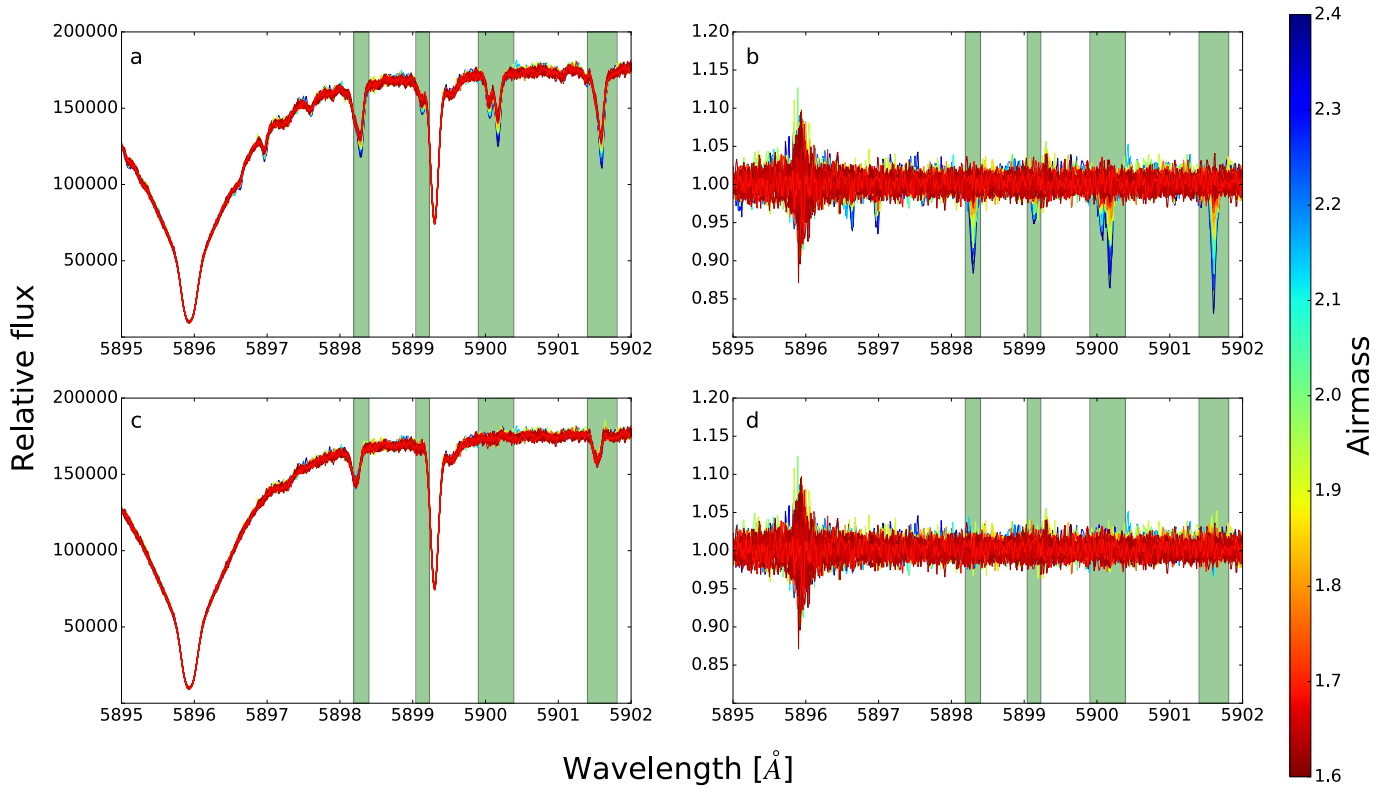

Fig. 2. Variability in the spectra over the night of 19 July 2007 for some water lines in the band at 5900 Å with the Na I D1 line on the left. The colorbar indicates the airmass. *Panel a* shows the non-corrected spectra. Green bands indicate the strongest telluric lines. *Panel b* is the variation of each non-corrected spectrum from the mean spectrum of the night. *Panel c* shows spectra corrected from the telluric lines using `Molecfit`. *Panel d* is the variation of each corrected spectrum from the mean spectrum.

spectrum in each night. The only difference in the settings used for each night consists of the choice of the fifteen spectral regions, which are optimized in each night with respect to the wavelength shift caused by the barycentric correction (Appendix B).

3.2. A first assessment of the telluric correction

`Molecfit` produces an output file which contains all parameters of the fit. We explored a range of plausible parameter values to initialize the fitting process and verified that the fit converges to the same χ^2 minimum, which it does. The obtained χ^2_r ranges from 3 to 12, for 68 to 80 parameters, and 6760 to 10 250 data points depending on the night. These relatively large values can likely be explained by the non-perfect model that is fitted to the data, in particular the non inclusion of stellar lines in the model.

Figure 2 shows all the spectra for the night of 19 July 2007 before and after the telluric correction. The scatter in the Na doublet is due to the low flux in the core while the variability of water telluric features is due to changes in the water column density (mainly due to airmass variations). As we can see with panel c, all the telluric lines are corrected to the noise level (including the telluric lines blended with stellar lines).

4. Methods

In this section, we describe how the transmission spectrum was derived and how we can search for water vapor in it with the cross-correlation technique.

4.1. Transmission spectrum

The transmission spectrum (Seager & Sasselov 2000; Brown 2001) was obtained during the transit, when the planet passes in front of the star. It is defined as the area occulted by the planet over the stellar disk as a function of wavelength and probes the highest layers of the atmosphere. We computed the transmission spectrum following the same formalism as Wyttenbach et al. (2017).

- Once the spectra are corrected from telluric features, we separated in-transit spectra from out-transit spectra. An in-transit spectrum, $f(\lambda, t_{\rm in})$, is a spectrum obtained when the planet occults a part of the stellar disk, while an out-transit spectrum, $f(\lambda, t_{\rm out})$, corresponds to the full-disk stellar spectrum.
- We corrected each spectrum for the stellar reflex motion induced by the planet, using the orbital parameters in Table 1.
- Then, we created the normalized master out-transit spectrum which is the sum of the out-transit spectra:

$$\tilde{F}_{\rm out}(\lambda) = \sum f(\lambda, t_{\rm out}). \quad (1)$$

- We normalized all the in-transit spectra to the continuum level of the master out-transit spectrum with a fourth-degree polynomial. This normalization was made from 5338 Å to 6900 Å, which corresponds to the red CCD. We applied a sigma-clipping rejection algorithm on the spectra in order to replace all the cosmic ray hits by the mean value of the other spectra at each wavelength. A normalized in-transit spectrum is noted $\tilde{f}(\lambda, t_{\rm in})$.
- Then, we computed the transmission spectrum as the sum of each individual transmission spectrum $(\tilde{f}(\lambda, t_{\rm in})/\tilde{F}_{\rm out}(\lambda))$,

Table 3. Spectral range and number of lines used for each mask. 55 226 water lines are available for this spectrale range in HITRAN and HITEMP database.

Mask	Spectral range [Å]	No. of lines used
$H_2O_{296\ K}$	6424.92–6612.53	151
$H_2O_{1300\ K}$	6415.08–6763.10	239
$H_2O_{1700\ K}$	6400.63–6794.76	873
$H_2O_{2100\ K}$	6400.93–6794.76	606
$H_2O_{2300\ K}$	6409.47–6859.60	413

after applying a Doppler shift to the planet rest frame p to compensate for the planet orbital motion during transit:

$$\tilde{\Re}(\lambda) = \sum_{t\in \mathrm{in}} \left.\frac{\tilde{f}(\lambda, t_{\mathrm{in}})}{\tilde{F}_{\mathrm{out}}(\lambda)}\right|_{\mathrm{p}} . \quad (2)$$

- Finally, we expressed the variation of the occulted area by the planet as a function of wavelength on an absolute scale, by using the known white-light radius ratio (Table 1) :

$$\frac{R_{\mathrm{p}}^2(\lambda)}{R_*^2} = 1 - \tilde{\Re}(\lambda) + \frac{R_{\mathrm{p}}^2(\lambda_{\mathrm{ref}})}{R_*^2}. \quad (3)$$

4.2. Cross-correlation function

The study of water vapor in the transmission spectrum in the visible is difficult considering the low intensity of the water bands and the noise level, even if we are able to resolve each line. To detect water vapor we need to co-add hundreds of lines to maximize the signal-to-noise ratio (S/N). To do so we used the cross-correlation function (CCF), expressed by:

$$CCF(v) = \sum_i S(\lambda_i) \cdot M(\lambda_i(1 + v/c)). \quad (4)$$

This technique (Baranne et al. 1996; Pepe et al. 2002) consists of projecting a binary mask M (with an aperture width of one pixel) on the spectrum S and to sum the transmitted flux at each wavelength. The mask contains the theoretical wavelengths of water vapor transitions, and is Doppler shifted successively to scan a radial velocity range of –80 to 80 km s^{-1} with a step of 0.82 km s^{-1}. This step corresponds to the size of one pixel on the CCD of HARPS.

In this study, we used masks at five different temperatures for the water band between 6400 and 6800 Å, including one at 296 K to check the presence of telluric residuals. The wavelength of each line was retrieved from HITRAN for the telluric mask (296 K) and from HITEMP for exoplanetary masks (1300 to 2300 K) (Rothman et al. 2009, 2010). The cross-section of each transition was scaled according to the temperature using the formulae provided in HITRAN. The number of lines in each mask was optimized to obtain the best S/N on the CCF. On the one hand, if the mask contains too few lines, even if they are strong, the noise will be large compared to the atmospheric signal. On the other hand, if too many lines are included, the noise will be lower but the weak lines will decrease the average signal. We found an optimal number of lines of 151–873 depending on temperature but note however that S/N has only a weak dependence on the exact line cut-off. Table 3 lists the different water vapor masks with their spectral range and number of lines. The mask at 296 K is used to check for the presence of telluric residuals. All the others are used to search for water vapor in the exoplanet.

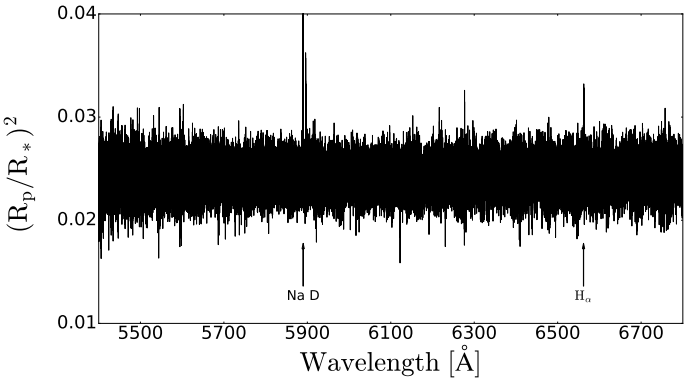

Fig. 3. The combined transmission spectrum of HD 189733b. The Na doublet is visible at ~5890 Å and the H-α line at 6562 Å.

5. Results

In this section, we present the transmission spectrum, a second verification of telluric correction and the search for water vapor.

5.1. The transmission spectrum

In order to maximize the S/N, we built a weighted mean of our three transmission spectra (for the three nights) using $w_i = 1/\sigma_i^2$ as weights, where σ_i is the standard deviation in the continuum for each night i . The transmission spectrum of HD 189733b is shown in Fig. 3 ($\sigma \sim 1500$ ppm). We confirm the excess absorption in the Na doublet at ~5890 Å detected by Wyttenbach et al. (2015). Around 6562 Å, we can see the Hα line, which varies with time (detected in two out of three transits), and is a topic of much debate in the community (Barnes et al. 2016; Cauley et al. 2015, 2016, 2017). We do not discuss it further here.

5.2. The contribution of residual telluric features to the transmission spectrum

Figure 4 shows the CCF of the transmission spectrum corrected from telluric features (in blue) for the three nights obtained with the mask at 296 K in the stellar rest frame. If the telluric correction was not sufficient, a signal would appear at the relative radial velocity of the terrestrial rest frame represented by the dashed line. The CCFs in gray are computed on the non-corrected transmission spectra and show the average telluric signature centered on the dashed line as expected. As we can see from the corrected CCFs, none of the nights shows a significant signal. The night of 19 July 2007 (panel b) shows somewhat higher dispersion in the core of the corrected CCF. Also, the non-corrected CCF shows a high dispersion in the continuum. This is due to the different mean Doppler shifts of the telluric lines between the in- and out-transit spectra (~300 m s^{-1}). Indeed, the transit occurs at the begining of the night and thus the $\tilde{F}_{\mathrm{out}}(\lambda)$ is only composed with out-transit spectra taken after the transit.

As a conclusion, `Molecfit` is capable of correcting the spectrum to the noise level even when co-adding 151 telluric lines (dispersion in the CCF continuum of about ~200 ppm).

5.3. The search for water vapor in the transmission spectrum of HD 189733b

We built the water vapor CCF from the combined transmission spectrum using the masks at 1700 K and 2100 K. These temperatures yield the lowest noise (due to the number of lines)

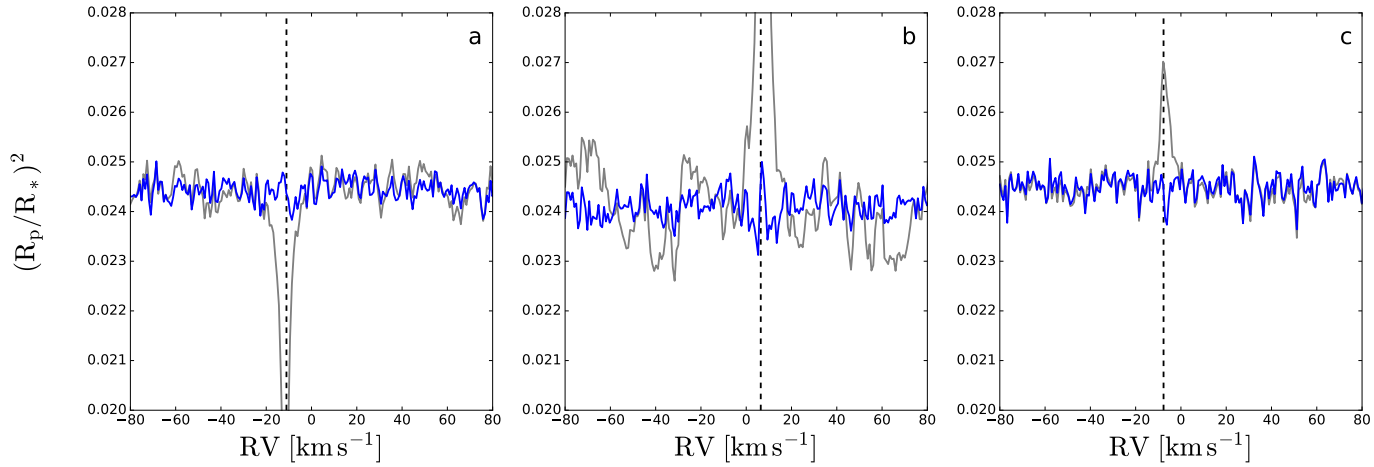

Fig. 4. CCFs of the transmission spectrum corrected from the telluric features (blue) and non-corrected (gray) in the stellar rest frame using the water vapor mask at 296 K for the three nights (*panel a*: 7 September 2006; *panel b*: 19 July 2007; and *panel c*: 28 August 2007). The dashed line shows the observer's radial velocity, where telluric residuals would be expected.

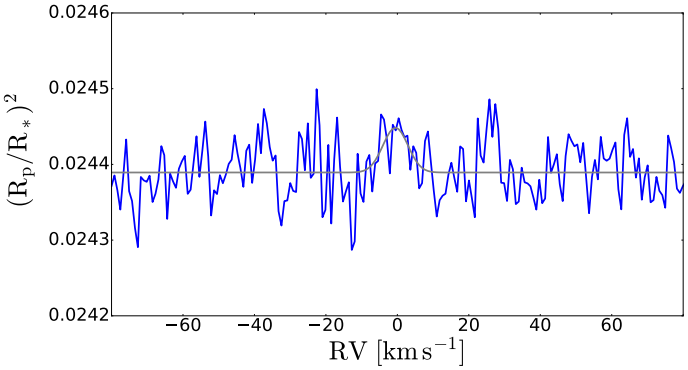

Fig. 5. Measured CCF in the planet rest frame using the water vapor mask at 1700 K (blue). The Gaussian fit is shown in gray.

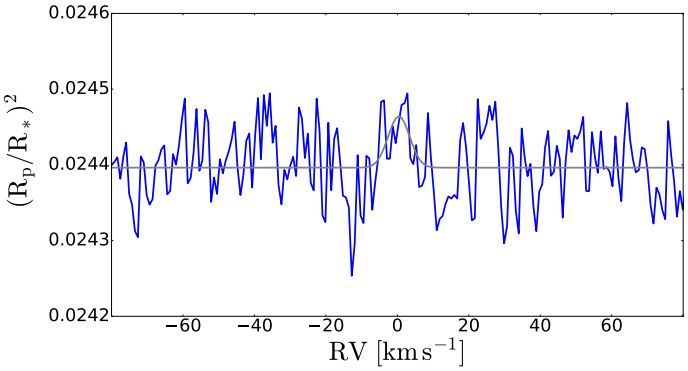

Fig. 6. Measured CCF in the planet rest frame with the water vapor mask at 2100 K (blue). The Gaussian fit is shown in gray.

compared to the masks at 1300 and 2300 K. We fit a Gaussian profile to the CCFs with free parameters for the continuum level, amplitude, FWHM and position. We also tested a Gaussian model with a FWHM fixed to 3.7 km s^{-1} (Doppler broadening and instrumental profile) but found that the fitted parameters are within 1-σ of the free FWHM parameters. Figures 5 and 6 show the CCFs for the masks at 1700 and 2100 K in the planet rest frame. The CCF at 1700 K has a dispersion in the continuum $\sigma_{\rm cont}$ of 34 ppm. $\sigma_{\rm cont}$ is defined as the standard deviation in the CCF continuum, ranging from –80 to –30 km s^{-1} and from 30 to 80 km s^{-1}. The amplitude of the fitted Gaussian is 59 ± 17 ppm, its centroid is at –0.54 ± 1.14 km s^{-1} and its FWHM is 8.2 ± 2.8 km s^{-1}. The CCF at 2100 K has a $\sigma_{\rm cont}$ of 43 ppm. The amplitude of the fitted Gaussian is 68 ± 23 ppm, its centroid is at 0.43 ± 1.16 km s^{-1} and its FWHM is 7.1 ± 2.8 km s^{-1}. These $\sigma_{\rm cont}$ values are due to white noise as they follow the square root of the number of lines (respectively 873 and 606).

In spite of the formal results and uncertainties, we do not consider this as a significant detection of H_2O because similarly-good Gaussian fits can be obtained at other radial velocity positions (see Figs. 5 and 6). Moreover, we carried out a model comparison between a Gaussian and a flat line using the Bayesian information criterion (BIC). We found that the Gaussian model is not optimal to fit these CCFs, since the straight line fit results in a ΔBIC of 4 in favor of the straight line model. We thus think that the errors obtained by the square root of the covariance matrix are unreliable, because the χ^2 minimum is poorly defined in parameter space.

6. Discussion

6.1. The telluric correction

The telluric correction is one of the most critical steps to be able to study transmission spectra with ground-based facilities. In the visible, telluric lines are dominated by H_2O and O_2. We have shown that `Molecfit` is a powerful tool which is able to correct H_2O telluric features to the noise level. However we stress here that it is necessary to be very careful in the choice of fitted spectral regions.

We might ask if the telluric correction could have removed a part of the exoplanet water vapor signal. This is unlikely because telluric features can be easily distinguished from planetary water transitions thanks to their relative Doppler shift. Indeed, the relative radial velocity between the terrestrial and the planet rest frames varies from 0.4 to –31.3 km s^{-1} for the 7 September 2006, from 16.0 to –12.8 km s^{-1} for the 19 July 2007 and from 2.4 to –27.4 km s^{-1} for the 28 August 2007. Therefore, the two line-systems overlap in only a small fraction of the observations. Moreover, `Molecfit` computes a physical model of the Earth atmosphere with essentially a single parameter controlling the depth of telluric water features. Also, the temperature

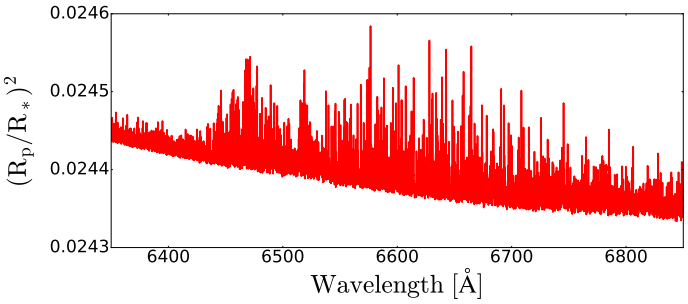

Fig. 7. Water band in the theoretical transmission spectrum at 2100 K computed with $^{\pi}\eta$. The slope is caused by the red wing of the Na doublet.

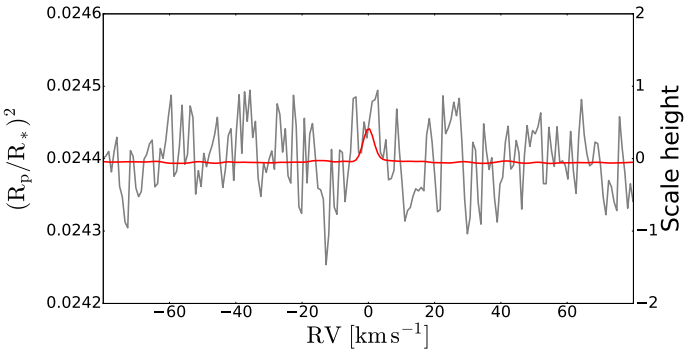

Fig. 8. Comparison between the measured CCF at 2100 K in gray and the theoretical CCF at 2100 K in red computed with $^{\pi}\eta$.

on HD 189733b is much higher than in Earth atmosphere, leading to different line intensities in the water band. Thus, even in the case of a null relative radial velocity, the telluric correction could erase only a fraction of the planetary water vapor signal.

6.2. The non-detection of water vapor

We computed a theoretical cloud-free isothermal transmission spectrum of HD 189733b at the resolution of HARPS including H, He, Na, K and H_2O with solar abundances using the $^{\pi}\eta$ tool Pino et al. (2017). $^{\pi}\eta$ is an improved version of the η code presented in Ehrenreich et al. (2006) and expanded in Ehrenreich et al. (2012) to compute transmission spectra of exoplanetary atmospheres. Its main characteristics are:

- High-resolution ($R \sim 1\,000\,000$). This is necessary to compare models to ground-based, high-resolution data.
- Broad wavelength coverage (330 nm–2 μm). This is necessary to compare models with space-borne, low- to medium-resolution data.

Figure 7 shows the 6500 Å water vapor band in the theoretical transmission spectrum at 2100 K. We can thus compute the theoretical CCF with the two same water vapor masks that we used (1700 and 2100 K). This is shown in Fig. 8. The predicted theoretical contrasts are 37 and 46 ppm respectively, for a FWHM of 3.7 and 3.8 km s^{-1}.

As we can see in Fig. 8, the theoretical contrast is of the same order as the measured dispersion in the CCF continuum which is less than one half scale height. This implies that the present transmission spectrum has a noise level too high to detect water vapor in this planet. Table 4 summarizes the noise properties and Gaussian fit parameters obtained from the data. To estimate our detection limits, we first compute the precision on the contrast

Table 4. Main characteristics of the observed and theoretical CCFs at 1700 and 2100 K.

Mask temperature [K]	1700	2100
No. of lines	873	606
σ_{cont} [ppm]	34	43
σ_{line} [ppm]	16	20
5-σ detection limit [ppm]	79	101
Fitted Gaussian contrast [ppm]	59 ± 17	68 ± 23
Theoretical contrast [ppm]	37	46

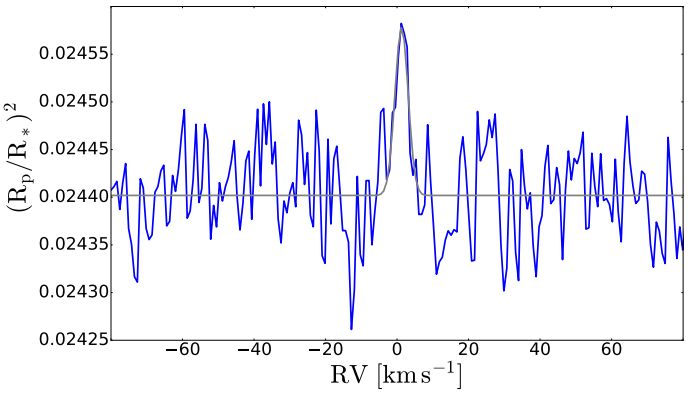

Fig. 9. Measured CCF with an injected signal scaled to an average amplitude of 100 ppm in the planet rest frame with the water vapor mask at 2100 K (blue). The Gaussian fit is shown in gray.

σ_{line} that could be achieved for a Gaussian CCF profile with a FWHM matching the theoretical value of 3.74 km s^{-1} (expressed in pixel units). It is given by:

$$\sigma_{line} = \frac{\sigma_{cont}}{\sqrt{FWHM}}, \quad (5)$$

where σ_{cont} is the measured dispersion in the continuum. The 5-σ detection limit is then simply given by 5·σ_{line}. This gives a value of about 100 ppm (see Table 4).

To further test our detection limit estimates, we injected the model of Fig. 7 scaled to an average amplitude of 100 ppm. The injection was done in the in-transit spectra before the telluric correction. We applied the same data reduction procedures as above. The Gaussian fit on the resulting CCF is shown in Fig. 9. The amplitude of the fitted Gaussian is 176 ± 29 ppm. The retrieved contrast is higher than the injected one but is still compatible with the noise level in the data (also given the uncertainty on the formal error itself, see discussion in Sect. 5.3). We also injected a 1000 ppm signal and retrieved a contrast of 1061 ± 29 ppm. This verification confirms that our technique does not remove the planetary signature.

6.3. Future prospects

At this stage, the easiest way to detect water vapor in the atmosphere of HD 189733b is by using the infrared transmission and emission spectra (McCullough et al. 2014; de Kok et al. 2013; Brogi et al. 2016). The difficulty in the detection of the H_2O molecule in the visible comes from the weak water signature in this spectral band. Nevertheless, we can improve our S/N by using more observations and/or a more favorable water band. We can ask the question of how many transits with HARPS we would need to obtain a 5-σ detection. Considering theoretical contrasts of 37 and 46 ppm, precisions of 7 and 9 ppm at 1700 K

and 2100 K must be reached. This means that we would need to observe five times more transits than used in this study, so a total of 15 transits with HARPS.

ESPRESSO (Pepe et al. 2010) is the new high-resolution spectrograph for the VLT, to be commissioned in 2017. ESPRESSO offers two major advantages compared to HARPS: first, the larger telescope diameter and improved instrument throughput will provide about six times more flux than HARPS on the 3.6-m telescope. Second, the ESPRESSO spectral coverage extends to 7800 Å, which includes the potassium doublet at 7665 and 7699 Å, as well as a water band at 7400 Å. The latter is significantly stronger than the one studied in this paper. We computed with the $^{\pi}\eta$ code an expected CCF contrast of ∼200 ppm for a cloud-free atmosphere. Given these improvements, a detection could be obtained with just one transit. A successful detection would allow us to compare the intensity of water absorption with the absorption features of the alkali doublets (sodium and potassium). By measuring their relative contrast, it is possible to infer the relative abundance of these species.

We need to keep in mind that the number of transits that we computed here is for a cloud-free atmosphere. If the atmosphere is cloudy as suggested by Lecavelier Des Etangs et al. (2008), Pont et al. (2008), Huitson et al. (2012), Pont et al. (2013), McCullough et al. (2014), water lines may be muted in these two bands. In the optical and NIR regions, scattering by aerosols is chromatic. Thus, the relative intensity of water bands at different wavelengths is indicative of the amount and type of scattering in the atmosphere. Spectral coverage is key to this characterization.

7. Conclusion

In this study, we correct telluric water lines to the noise level using the `Molecfit` tool. We reach a 1-σ precision of 20 ppm on the CCF contrast for water vapor in the atmosphere of HD 189733b. Therefore, our data would have revealed a ∼100 ppm signal at 5-σ. Furthermore, given a maximum theoretical contrast of only 46 ppm, our data are too noisy to put a meaningful constraint on the presence of water vapor in HD 189733b. Thus, we determine how many transits we would need to detect water vapor in HD 189733b-like planets with HARPS and the future ESPRESSO spectrograph for a cloud-free atmosphere. Given the increased efficiency and wavelength coverage of ESPRESSO, the stronger water band at 7400 Å could be detected in just one transit (CCF contrast of ∼200 ppm). In conclusion, ESPRESSO will not only be a terrestrial planet hunter, but also the instrument of choice for atmospheric characterization in the visible. Moreover, the detection of a given species at different wavelengths (visible and near-IR) is key to characterize atmospheric structure and composition, in particular the presence of aerosols.

Acknowledgements. This work has been carried out within the frame of the National Centre for Competence in Research "PlanetS" supported by the Swiss National Science Foundation (SNSF). The authors acknowledge the financial support of the SNSF by the grant numbers 200020_152721 and 200020_166227. We want to thank the entire atmosphere group of Geneva Observatory, the team in charge of `Molecfit` for their work and their upgrades, Jens Hoeijmaker about the discussion on the water correction and the anonymous referee for the careful reading and pertinent comments.

References

Agol, E., Cowan, N. B., Knutson, H. A., et al. 2010, ApJ, 721, 1861
Arnold, L., Ehrenreich, D., Vidal-Madjar, A., et al. 2014, A&A, 564, A58
Baranne, A., Queloz, D., Mayor, M., et al. 1996, A&AS, 119, 373
Barnes, J. R., Haswell, C. A., Staab, D., & Anglada-Escudé, G. 2016, MNRAS, 462, 1012
Birkby, J. L., de Kok, R. J., Brogi, M., et al. 2013, MNRAS, 436, L35
Boisse, I., Moutou, C., Vidal-Madjar, A., et al. 2009, IAU Symp., 253, 462
Bouchy, F., Udry, S., Mayor, M., et al. 2005, A&A, 444, 5
Bourrier, V., Lecavelier des Etangs, A., Dupuy, H., et al. 2013, A&A, 551, A63
Brogi, M., de Kok, R. J., Albrecht, S., et al. 2016, ApJ, 817, 106
Brown, T. M. 2001, ApJ, 553, 1006
Cauley, P. W., Redfield, S., Jensen, A. G., et al. 2015, ApJ, 810, 13
Cauley, P. W., Redfield, S., Jensen, A. G., & Barman, T. 2016, ApJ, 152, 20
Cauley, P. W., Redfield, S., & Jensen, A. G. 2017, AJ, 153, 185
Cegla, H. M., Lovis, C., Bourrier, V., et al. 2016, A&A, 588, A127
Charbonneau, D., Brown, T. M., Latham, D. W., & Mayor, M. 2000, ApJ, 529, L45
Charbonneau, D., Brown, T. M., Noyes, R. W., & Gilliland, R. L. 2002, ApJ, 568, 377
Collier Cameron, A., Bruce, V. A., Miller, G. R. M., Triaud, A. H. M. J., & Queloz, D. 2010, MNRAS, 403, 151
de Kok, R. J., Brogi, M., Snellen, I. A. G., et al. 2013, A&A, 554, A82
Deming, D., Wilkins, A., McCullough, P., et al. 2013, ApJ, 774, 95
Di Gloria, E., Snellen, I. A. G., & Albrecht, S. 2015, A&A, 580, A84
Ehrenreich, D., Tinetti, G., Lecavelier Des Etangs, A., Vidal-Madjar, A., & Selsis, F. 2006, A&A, 448, 379
Ehrenreich, D., Vidal-Madjar, A., Widemann, T., et al. 2012, A&A, 537, L2
Ehrenreich, D., Bourrier, V., Wheatley, P. J., et al. 2015, Nature, 522, 459
Evans, T. M., Pont, F., Sing, D. K., et al. 2013, ApJ, 772, L16
Heng, K., Wyttenbach, A., Lavie, B., et al. 2015, ApJ, 803, L9
Huitson, C. M., Sing, D. K., Vidal-Madjar, A., et al. 2012, MNRAS, 422, 2477
Kataria, T., Sing, D. K., Lewis, N. K., et al. 2016, ApJ, 821, 9
Kausch, W., Noll, S., Smette, A., et al. 2015, A&A, 576, A78
Khalafinejad, S., von Essen, C., Hoeijmakers, H. J., et al. 2017, A&A, 598, A131
Kreidberg, L., Line, M. R., Bean, J. L., et al. 2015, ApJ, 814, 66
Lecavelier Des Etangs, A., Pont, F., Vidal-Madjar, A., & Sing, D. 2008, A&A, 481, L83
Lecavelier Des Etangs, A., Ehrenreich, D., Vidal-Madjar, A., et al. 2010, A&A, 514, A72
Lecavelier des Etangs, A., Bourrier, V., Wheatley, P. J., et al. 2012, A&A, 543, L4
Louden, T., & Wheatley, P. J. 2015, ApJ, 814, L24
Madhusudhan, N., Harrington, J., Stevenson, K. B., et al. 2011, Nature, 469, 64
Mayor, M., Pepe, F., Queloz, D., et al. 2003, The Messenger, 114, 20
McCullough, P. R., Crouzet, N., Deming, D., & Madhusudhan, N. 2014, ApJ, 791, 55
Moses, J. I., Madhusudhan, N., Visscher, C., & Freedman, R. S. 2013, ApJ, 763, 25
Pepe, F., Mayor, M., Galland, F., et al. 2002, A&A, 388, 632
Pepe, F. A., Cristiani, S., Rebolo Lopez, R., et al. 2010, in Proc. SPIE, 7735, 77350F
Pino, L., Ehrenreich, D., Wyttenbach, A., et al. 2017, A&A, accepted [arXiv:1709.09678]
Pont, F., Knutson, H., Gilliland, R. L., Moutou, C., & Charbonneau, D. 2008, MNRAS, 385, 109
Pont, F., Sing, D. K., Gibson, N. P., et al. 2013, MNRAS, 432, 2917
Redfield, S., Endl, M., Cochran, W. D., & Koesterke, L. 2008, ApJ, 673, L87
Remedios, J. J., Gille, J. C., & Barnett, J. J. 2001, Adv. Space Res., 27, 1479
Rothman, L. S., Gordon, I. E., Barbe, A., et al. 2009, J. Quant. Spectr. Rad. Trans., 110, 533
Rothman, L. S., Gordon, I. E., Barber, R. J., et al. 2010, J. Quant. Spectr. Rad. Trans., 111, 2139
Seager, S., & Sasselov, D. D. 2000, ApJ, 537, 916
Sing, D. K., Pont, F., Aigrain, S., et al. 2011, MNRAS, 416, 1443
Sing, D. K., Lecavelier des Etangs, A., Fortney, J. J., et al. 2013, MNRAS, 436, 2956
Sing, D. K., Wakeford, H. R., Showman, A. P., et al. 2015, MNRAS, 446, 2428
Sing, D. K., Fortney, J. J., Nikolov, N., et al. 2016, Nature, 529, 59
Smette, A., Sana, H., Noll, S., et al. 2015, A&A, 576, A77
Snellen, I. A. G., Albrecht, S., de Mooij, E. J. W., & Le Poole, R. S. 2008, A&A, 487, 357
Snellen, I. A. G., de Kok, R. J., de Mooij, E. J. W., & Albrecht, S. 2010, Nature, 465, 1049
Torres, G., Winn, J. N., & Holman, M. J. 2008, ApJ, 677, 1324
Triaud, A. H. M. J., Queloz, D., Bouchy, F., et al. 2009, A&A, 506, 377
Vidal-Madjar, A., Arnold, L., Ehrenreich, D., et al. 2010, A&A, 523, A57
Wyttenbach, A., Ehrenreich, D., Lovis, C., Udry, S., & Pepe, F. 2015, A&A, 577, A62
Wyttenbach, A., Lovis, C., Ehrenreich, D., et al. 2017, A&A, 602, A36

Appendix A

Table A.1. Initial parameters of `Molecfit` for every night.

Initial parameters	Values	Comments
ftol	10^{-9}	χ^2 Convergence criterion
xtol	10^{-9}	Parameter convergence criterion
Molecules	H_2O, O_2	
$n_{\rm cont}$	3	Degree of polynom for the continuum
a_0	2000	Constant of polynom for the continuum
n_λ	2	Chebyschev degree for wavelength calibration
b_0	0	Constant Chebyschev for wavelength calibration
$\omega_{\rm Gaussian}$	4.5	FWHM in pixel
Kernel size	15	
Pixel scale	0.16	
Slit width	$1''$	
MIPAS profile	equ	Equatorial profile
Atmospheric profile	0	Natural profile
PWV	–1	No value taken into account

Appendix B

Table B.1. Fitted regions with `Molecfit` for the three nights.

Fitted regions [μm]		
7 Sep. 2006	19 Jul. 2007	28 Aug. 2007
0.592100–0.592354	0.592040–0.592345	0.592100–0.592253
0.592500–0.592927	0.592393–0.592906	0.592500–0.592703
0.594676–0.594912	0.594530–0.594890	0.594716–0.594798
0.596814–0.597700	0.596781–0.597663	0.596916–0.597095
0.627961–0.628145	0.627882–0.628203	0.597258–0.597700
0.628313–0.628410	0.628270–0.628405	0.628034–0.628111
0.628510–0.628671	0.628470–0.628655	0.628307–0.628390
0.628861–0.629231	0.628800–0.629232	0.628510–0.628671
0.629614–0.629818	0.629578–0.629793	0.629064–0.629231
0.647405–0.647726	0.647370–0.647711	0.629614–0.629818
0.647964–0.648327	0.647790–0.648328	0.647465–0.647726
0.648530–0.649333	0.648465–0.649300	0.648078–0.648233
0.651200–0.652005	0.651100–0.651964	0.649058–0.649333
0.686800–0.688240	0.686850–0.688216	0.651626–0.651717
0.688550–0.691500	0.688509–0.691500	0.651869–0.652005
		0.686684–0.687365
		0.687521–0.688240
		0.688539–0.691236