



Universiteit  
Leiden  
The Netherlands

## **Discriminating between hazy and clear hot-Jupiter atmospheres with CARMENES**

Sánchez-López, A.; López-Puertas, M.; Snellen, I.A.G.; Nagel, E.; Bauer, F.F.; Pallé, E.; ... ; Zechmeister, M.

### **Citation**

Sánchez-López, A., López-Puertas, M., Snellen, I. A. G., Nagel, E., Bauer, F. F., Pallé, E., ... Zechmeister, M. (2020). Discriminating between hazy and clear hot-Jupiter atmospheres with CARMENES. *Astronomy & Astrophysics*, 643, A24. doi:10.1051/0004-6361/202038629

Version: Accepted Manuscript

License: [Leiden University Non-exclusive license](#)

Downloaded from: <https://hdl.handle.net/1887/3133910>

**Note:** To cite this publication please use the final published version (if applicable).

# Discriminating between hazy and clear hot-Jupiter atmospheres with CARMENES

A. Sánchez-López<sup>1</sup>, M. López-Puertas<sup>2</sup>, I. A. G. Snellen<sup>1</sup>, E. Nagel<sup>3</sup>, F. F. Bauer<sup>2</sup>, E. Pallé<sup>4,5</sup>, L. Tal-Or<sup>6,9</sup>, P. J. Amado<sup>2</sup>, J. A. Caballero<sup>7</sup>, S. Czesla<sup>8</sup>, L. Nortmann<sup>9</sup>, A. Reiners<sup>9</sup>, I. Ribas<sup>10,11</sup>, A. Quirrenbach<sup>12</sup>, J. Aceituno<sup>13</sup>, V. J. S. Béjar<sup>4,5</sup>, N. Casasayas-Barris<sup>4,5</sup>, Th. Henning<sup>14</sup>, K. Molaverdikhani<sup>14</sup>, D. Montes<sup>15</sup>, M. Stangret<sup>4,5</sup>, M. R. Zapatero Osorio<sup>16</sup>, and M. Zechmeister<sup>9</sup>

<sup>1</sup> Leiden Observatory, Leiden University, Postbus 9513, 2300 RA, Leiden, The Netherlands  
e-mail: alexsl@strw.leidenuniv.nl

<sup>2</sup> Instituto de Astrofísica de Andalucía (IAA-CSIC), Glorieta de la Astronomía s/n, 18008 Granada, Spain

<sup>3</sup> Thüringer Landessternwarte Tautenburg, Sternwarte 5, 07778 Tautenburg, Germany

<sup>4</sup> Instituto de Astrofísica de Canarias (IAC), Calle Vía Láctea s/n, 38200 La Laguna, Tenerife, Spain

<sup>5</sup> Departamento de Astrofísica, Universidad de La Laguna, 38026 La Laguna, Tenerife, Spain

<sup>6</sup> Department of Physics, Ariel University, Ariel 40700, Israel

<sup>7</sup> Centro de Astrobiología (CSIC-INTA), ESAC, Camino bajo del castillo s/n, 28692 Villanueva de la Cañada, Madrid, Spain

<sup>8</sup> Hamburger Sternwarte, Universität Hamburg, Gojenbergsweg 112, 21029 Hamburg, Germany

<sup>9</sup> Institut für Astrophysik, Georg-August-Universität, Friedrich-Hund-Platz 1, 37077 Göttingen, Germany

<sup>10</sup> Institut de Ciències de l'Espai (CSIC-IEEC), Campus UAB, c/ de Can Magrans s/n, 08193 Bellaterra, Barcelona, Spain

<sup>11</sup> Institut d'Estudis Espacials de Catalunya (IEEC), 08034 Barcelona, Spain

<sup>12</sup> Landessternwarte, Zentrum für Astronomie der Universität Heidelberg, Königstuhl 12, 69117 Heidelberg, Germany

<sup>13</sup> Observatorio de Calar Alto, Sierra de los Filabres, 04550 Gérgal, Almería, Spain

<sup>14</sup> Max-Planck-Institut für Astronomie, Königstuhl 17, 69117 Heidelberg, Germany

<sup>15</sup> Departamento de Física de la Tierra y Astrofísica & IPARCOS-UCM (Instituto de Física de Partículas y del Cosmos de la UCM), Facultad de Ciencias Físicas, Universidad Complutense de Madrid, 28040 Madrid, Spain

<sup>16</sup> Centro de Astrobiología (CSIC-INTA), Carretera de Ajalvir km 4, 28850 Torrejón de Ardoz, Madrid, Spain

## ABSTRACT

**Context.** Relatively large radii of some hot Jupiters observed in the ultraviolet (UV) and blue-optical are generally interpreted to be due to Rayleigh scattering by high-altitude haze particles. However, the haze composition and its production mechanisms are not fully understood, and observational information is still limited.

**Aims.** We aim to study the presence of hazes in the atmospheres of HD 209458 b and HD 189733 b with high spectral resolution spectra by analysing the strength of water vapour cross-correlation signals across the red optical and near-infrared wavelength ranges.

**Methods.** A total of seven transits of the two planets were observed with the CARMENES spectrograph at the 3.5 m Calar Alto telescope. Their Doppler-shifted signals were disentangled from the telluric and stellar contributions using the detrending algorithm SYSREM. The residual spectra were subsequently cross-correlated with water vapour templates at 0.70–0.96  $\mu\text{m}$  to measure the strength of the water vapour absorption bands.

**Results.** The optical water vapour bands were detected at  $5.2\sigma$  in HD 209458 b in one transit, whereas no evidence of them was found in four transits of HD 189733 b. Therefore, the relative strength of the optical water bands compared to those in the near-infrared were found to be larger in HD 209458 b than in HD 189733 b.

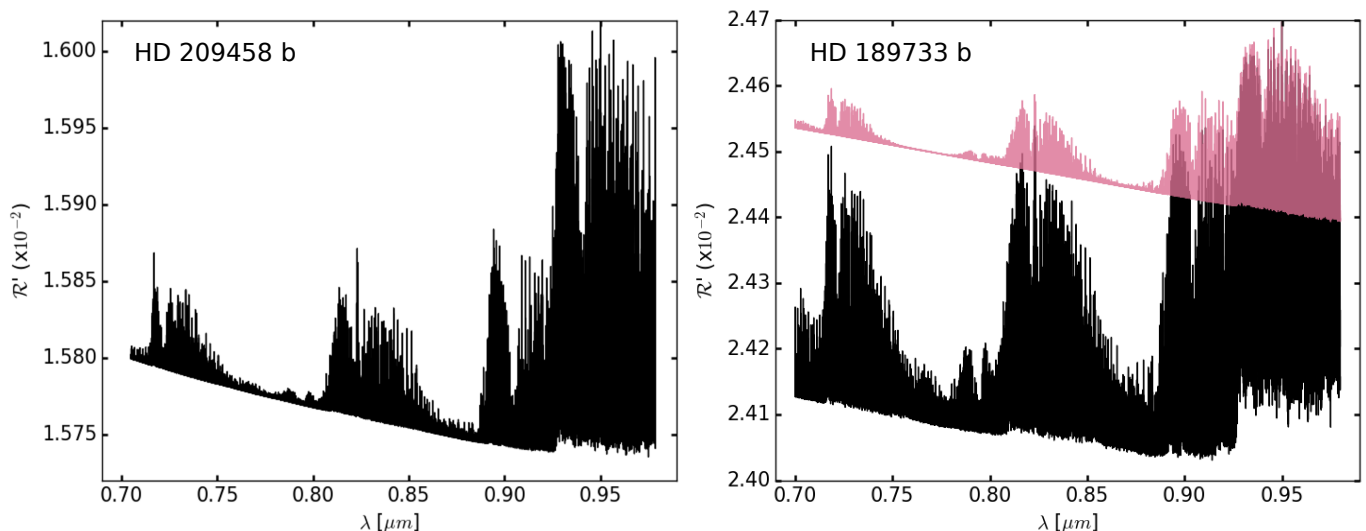
**Conclusions.** We interpret the non-detection of optical water bands in the transmission spectra of HD 189733 b, compared to the detection in HD 209458 b, to be due to the presence of high-altitude hazes in the former planet, which are largely absent in the latter. This is consistent with previous measurements with the *Hubble Space Telescope*. We show that currently available CARMENES observations of hot Jupiters can be used to investigate the presence of haze extinction in their atmospheres.

**Key words.** planets and satellites: atmospheres – planets and satellites: individual: HD 209458 b, HD 189733 b – techniques: spectroscopic – infrared: planetary systems

## 1. Introduction

Ground-based spectroscopic observations at high resolution ( $R > 40,000$ ) have flourished as a very powerful tool for the unprecedented characterisation of exoplanet atmospheres (e.g. Birkby et al. 2018; Ehrenreich et al. 2020). In recent years, a wide variety of atomic and molecular species, such as Fe I, Fe II, TiO, CO, H<sub>2</sub>O, CH<sub>4</sub>, HCN, and NH<sub>3</sub>, among others, have been successfully identified in the atmospheres of hot Jupiters using

cross-correlation techniques (Snellen et al. 2010; Birkby et al. 2013; de Kok et al. 2013; Nugroho et al. 2017; Brogi & Line 2017; Hawker et al. 2018; Yan & Henning 2018; Casasayas-Barris et al. 2018; Brogi et al. 2018; Hoeijmakers et al. 2018; Brogi & Line 2019; Alonso-Floriano et al. 2019; Sánchez-López et al. 2019; Guilluy et al. 2019; Casasayas-Barris et al. 2019). These detections provide unique information on the physical and chemical processes in hot planet atmospheres, with the aim to



**Fig. 1.** Model transmission spectra of HD 209458 b (*left*) and HD 189733 b (*right*) using the pressure-temperature profiles from Brogi & Line (2017) and Brogi et al. (2018) and the water vapour volume-mixing ratios of  $10^{-5}$  and  $10^{-4}$ , respectively. The same model for HD 189733 b with an additional haze contribution at an arbitrary level is shown in magenta to illustrate the reduction of the water signal. All models were computed at the CARMENES spectral resolution in the VIS channel ( $R = 94,600$ ) and they included the contributions from Rayleigh scattering and collision-induced absorption.

constrain their composition and ultimately their formation and evolution scenarios.

Much effort has been made to the characterisation of the two most studied hot Jupiters: HD 209458 b (Charbonneau et al. 2000; Henry et al. 2000) and HD 189733 b (Bouchy et al. 2005). In particular, space observations of the former are better understood when considering partial cloud coverage (Barstow et al. 2017; MacDonald et al. 2017), whereas the latter presents a rather steep Rayleigh scattering slope produced by a strong haze extinction (Sing et al. 2016; Barstow et al. 2017). In this context, the study of  $\text{H}_2\text{O}$  cross-correlation signals, using ground-based high resolution spectrographs across different spectral intervals, can provide strong constraints on the presence of these types of atmospheric aerosols (Pino et al. 2018). This is because the strength of the  $\text{H}_2\text{O}$  rovibrational bands as well as the intensity of the haze extinction change with wavelength. Specifically, the strength of the water vapour bands decreases towards shorter wavelengths, whereas the extinction due to hazes increases (e.g. see right panel in Fig. 1).

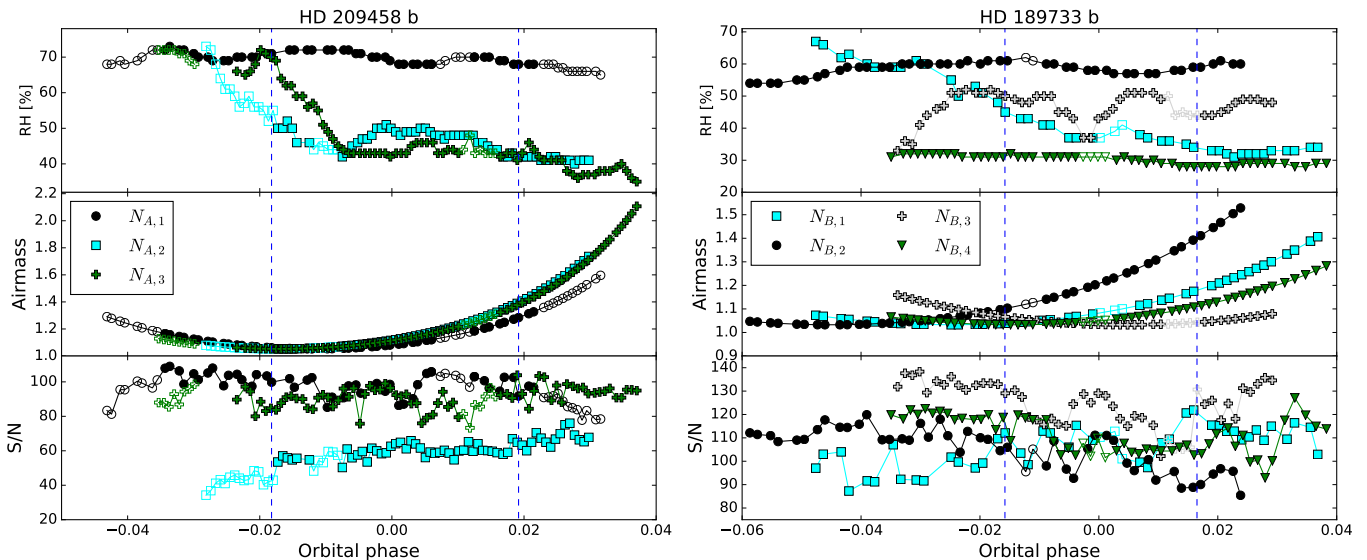
Earlier studies have investigated the presence of  $\text{H}_2\text{O}$  in HD 189733 b and HD 209458 b from the ground at near-infrared wavelengths, finding robust detections (Brogi et al. 2018; Hawker et al. 2018; Cabot et al. 2019; Alonso-Floriano et al. 2019; Sánchez-López et al. 2019). In previous work by our group, we compared the water vapour detections using the CARMENES spectrograph in HD 189733 b (Alonso-Floriano et al. 2019) and HD 209458 b at  $0.96\text{--}1.71\ \mu\text{m}$  (Sánchez-López et al. 2019). We found a strong  $\text{H}_2\text{O}$  signal using the band at  $\sim 1.0\ \mu\text{m}$  in HD 209458 b, while only weak signatures from this band were obtained for HD 189733 b. In the latter planet, however, we obtained strong  $\text{H}_2\text{O}$  signals using the  $\sim 1.15\ \mu\text{m}$  and  $\sim 1.40\ \mu\text{m}$  bands. This result already hinted at a possible muting of the shorter-wavelength signals in HD 189733 b by a strong haze extinction. However, different telluric conditions in the two nights and the variability of the telluric water vapour reported in Sánchez-López et al. (2019) prevented drawing further conclusions.

In this work, we aim to probe the atmospheres of HD 189733 b and HD 209458 b in the red optical by observing several transits of these hot Jupiters with CARMENES. We focus our attention on the  $0.70\text{--}0.96\ \mu\text{m}$  spectral region, where the strongest  $\text{H}_2\text{O}$  absorption in the optical occurs. Previous attempts at detecting water vapour in HD 189733 b in the optical were reported in Allart et al. (2017), who studied the  $\sim 0.65\ \mu\text{m}$  spectral band with the HARPS spectrograph (Mayor et al. 2003). They reported a non-detection and a  $5\sigma$  upper limit of 100 ppm on this band’s strength. Similar efforts to detect  $\text{H}_2\text{O}$  in the optical were undertaken in Esteves et al. (2017), targeting the super-Earth 55 Cancri e. The authors reported a non-detection of water vapour in the optical after combining ESPaDOns (506–795 nm, Donati 2003) and Subaru HDS (524–789 nm, Noguchi et al. 2002) data. However, no similar studies have been reported for HD 209458 b in the optical region so far.

Here, we report a detection of water vapour in HD 209458 b in the optical ( $0.70\text{--}0.96\ \mu\text{m}$ ) from one transit. In contrast, we do not detect  $\text{H}_2\text{O}$  in HD 189733 b in this spectral region even after combining four transits, which is in line with previous studies. In Sect. 2 we describe the observations and first steps of the analyses. In Sect. 3, we describe the telluric and stellar signal removal using the SYSREM algorithm and present the cross-correlation technique to extract planet atmospheric signals. In Sect. 4 we discuss the results and their implications followed by the conclusions in Sect. 5.

## 2. Observations and data reduction

We used the CARMENES spectrograph (Quirrenbach et al. 2016, 2018) to observe three transits of HD 209458 b (target A) during the nights of 16 September 2016 (hereafter  $N_{A,1}$ ), 8 November 2016 (hereafter  $N_{A,2}$ ), and 6 September 2018 (hereafter  $N_{A,3}$ ), and four transits of HD 189733 b (target B) on 8 August 2016 (hereafter  $N_{B,1}$ ), 16 September 2016 (hereafter  $N_{B,2}$ ), 30 June 2019 (hereafter  $N_{B,3}$ ), and 9 August 2019 (hereafter  $N_{B,4}$ ). CARMENES is mounted at the 3.5 m telescope at the Calar Alto Observatory (Almería, Spain). We analysed the ob-



**Fig. 2.** Relative humidity (*top panel*), air mass (*middle panel*), and mean S/N per spectra (*bottom panel*) during the transit observations of HD 209458 b (*left*) and HD 189733 b (*right*). The transits occur between the vertical dashed lines. Out-of-transit open symbols represent the spectra that are not included in the analyses because of their low S/N. Open symbols during transit represent spectra that are not included due to the overlap of telluric and planet lines when the latter has a near-zero velocity in the observer’s frame.

servations of its visible channel in the  $0.52\text{--}0.96\ \mu\text{m}$  spectral region at a resolving power of  $\mathcal{R} = 96,000$ . We restricted the analyses to the spectral orders covering the  $0.70\text{--}0.96\ \mu\text{m}$  interval, where the strongest  $\text{H}_2\text{O}$  absorption occurs (see Fig. 1). We did not include the weak  $\text{H}_2\text{O}$  band at  $\sim 0.65\ \mu\text{m}$ , because of its expected small contribution.

The relative humidity, air mass, and mean signal-to-noise ratio (S/N) per spectrum for each night are shown in Fig. 2. In each case, the raw spectra were processed using the CARMENES data reduction pipeline *caraca1* v2.20 (Zechmeister et al. 2014; Caballero et al. 2016) and, subsequently, analysed using custom Python subroutines. For  $N_{A,1}$ , eight spectra observed at the beginning of the night and another 12 recorded at the end were discarded due to their decreasing S/N, which hampered the telluric correction procedure discussed below (see out-of-transit open symbols in Fig. 2). Regarding  $N_{A,2}$ , we discarded the first 15 pre-transit spectra to avoid low ( $< 50$ ) S/N observations. The significantly lower S/N of the observations during this night was due to worse seeing and extinction conditions. Unfortunately, the observational epoch in all nights caused some in-transit spectra to be recorded at times when the velocity of the exoplanet with respect to the Earth ranged from  $-2.6$  to  $+2.6\ \text{km s}^{-1}$  (two wavelength steps around zero). This yielded a poor Doppler separation of the exo-atmospheric lines from the telluric contribution, which would result in a strong telluric contamination in the final signal. Therefore, these spectra were also discarded from the analysis (see open symbols during transit in Fig. 2).

We normalised each spectral order by using second-order polynomial fits for their pseudo-continuum. This step provided a self-calibration that allowed us to compare flux variations between spectra at small spectral scales. Next, we masked the spectral regions where the telluric absorption was larger than 80% of the flux continuum. Unfortunately, this necessary step also prevented us from obtaining information from the pixels with the largest possible  $\text{H}_2\text{O}$  signal from the planet atmosphere. In addition, we used the methods described in Alonso-Floriano et al. (2019) and Sánchez-López et al. (2019) to mask the outliers

present in the data, which are likely produced by cosmic rays, and telluric emission lines.

### 3. Data analysis

#### 3.1. Detrending using SYSREM

With the objective of recovering the weak exo-atmospheric signal, we removed the telluric and stellar contributions that dominated the spectral matrix for each night. We used SYSREM (Tamuz et al. 2005; Mazeh et al. 2007), which is a principal component analysis algorithm that has been widely used and tested in the past for exo-atmospheric studies (Birkby et al. 2013, 2017; Nugroho et al. 2017; Hawker et al. 2018; Alonso-Floriano et al. 2019; Sánchez-López et al. 2019; Gibson et al. 2020; Stangret et al. 2020). In total, we ran 15 SYSREM iterations in each spectral order, night, and target, and we stored the resulting 2520 residual matrices for further analyses. The individual treatment of the orders allowed us to take into account their different contamination levels, caused by different telluric or stellar contributions, their different detector efficiencies, or changing instrumental responses during the night. An imperfect removal of these contributions is unavoidable regardless of the number of iterations performed because SYSREM fits the spectral matrix by working with a set of two linear and independent coefficients that change with time and wavelength, respectively. The residual spectral matrix is obtained after subtracting the best least square fit from the original spectral matrix, which presents a non-linear behaviour by itself. Nevertheless, the aforementioned studies extensively showed the suitability of this algorithm for removing telluric and stellar contributions even in the presence of high water vapour levels above the observatory.

After a certain number of SYSREM iterations, which could be different for each spectral order, the algorithm also identifies and removes a possible signal from the exo-atmosphere. In order to minimise this effect, we injected a synthetic  $\text{H}_2\text{O}$  signal at the expected velocities for the planets early in the procedure (i.e. before the telluric correction) and analysed its behaviour. We used the cross-correlation technique (see Sect. 3.2) to maximise the

recovery of the injected signal **in each spectral order and** for each of the seven nights (see Figs. A.1 and A.2). Since the water vapour bands in the 0.70–0.96  $\mu\text{m}$  spectral range are rather weak, we injected the signal at five times (5 $\times$ ) the expected strength to ensure that it was clearly recovered above the noise. The spectral orders for which the S/N of the injected signal peak was lower than three (below the dashed horizontal lines in both figures) were discarded due to the very small actual signal that we would expect to retrieve from them. For the remaining orders and for the rest of the analysis, we used the iteration that allowed us to maximise the recovery of the injected signal (marked with star symbols in Figs. A.1 and A.2).

Additionally, we visually inspected the cross-correlation results for each spectral order and each night individually in search for possible residuals that might not be detected in the previous step. For instance, a strong injected signal might still be retrieved above a strong residual, but the latter would mask a weaker, real peak. We found that some spectral orders presented noise structures around the expected planet velocities or telluric residuals in the form of large cross-correlation values around 0  $\text{km s}^{-1}$  (i.e. the Earth’s rest-frame). Hence, these spectral orders were left out of the analyses (grey curves in the figures of Appendix A).

In total, the discarded spectral orders due to both the weak recovery of injections and the presence of telluric residuals represented 39%, 46%, and 50% of the total data of HD 209458 b on  $N_{A,1}$ ,  $N_{A,2}$ , and  $N_{A,3}$ , respectively. For HD 189733 b, 50%, 25%, 32%, and 26% of the total data were excluded on  $N_{B,1}$ ,  $N_{B,2}$ ,  $N_{B,3}$ , and  $N_{B,4}$ , respectively. The highest fractions of lost spectral orders seem to correlate with the nights presenting a higher variability of the relative humidity (see Fig. 2). Therefore, a large telluric variability can preclude a successful telluric removal with SYSREM in some spectral orders, as is discussed by Sánchez-López et al. (2019).

### 3.2. Planet atmospheric signal extraction

In order to probe the planet atmospheric signal, we followed the same procedure as Sánchez-López et al. (2019). We cross-correlated the residual spectra resulting from SYSREM with high-resolution templates of the H<sub>2</sub>O absorption computed at the CARMENES spectral resolution with KOPRA (Stiller et al. 2002). In order to compare our results directly with our previous findings, we used the same atmospheric parameters for HD 209458 b and HD 189733 b as Sánchez-López et al. (2019) and Alonso-Floriano et al. (2019), respectively (see Fig. 1).

Due to the largely unknown pressure level at which the local continuum around the water vapour bands is produced, these observations, similar to most transit measurements, result in a large intrinsic uncertainty in absolute abundances (e.g. Fig. 9 in Alonso-Floriano et al. 2019). Therefore, we did not test other combinations of temperatures or H<sub>2</sub>O abundances for these atmospheres.

The cross-correlations were performed for each night and spectral order separately in a velocity interval from –200 to +200  $\text{km s}^{-1}$  (with respect to the Earth’s rest-frame) in steps of 1.3  $\text{km s}^{-1}$ , which corresponds to the mean pixel size of the instrument in the red optical range. At this stage, no planet atmospheric features could be observed in the orders individually. With the objective of enhancing a possible signal, the cross-correlations matrices of all orders were co-added to form one total cross-correlation matrix for each night. Consecutively, we shifted the matrices to the exoplanet rest-frame and co-added all in-transit spectra over time, which allowed us to obtain a 1D cross-correlation function (CCF) per night. We tested a range

of planet orbital velocities ( $K_p$ ), creating a grid from –280 to +280  $\text{km s}^{-1}$  (i.e. 561 CCFs). This allowed us to inspect a wide velocity space to check for spurious signals, indications of telluric or stellar residuals at low  $K_p$  values, and possibly strong correlated noise sources appearing in the (unphysical) negative  $K_p$  space (Birkby et al. 2017).

In order to assess the significance of the observed signals, we calculated the S/N of each CCF at each  $K_p$  by dividing each cross-correlation value by the standard deviation obtained from the rest of the CCF velocity interval, excluding a  $\pm 15.6 \text{ km s}^{-1}$  region around it. The selection of different velocity intervals can impact the assessment of the noise and, hence, we chose a wide interval (i.e.  $\pm 200 \text{ km s}^{-1}$ ) to obtain an accurate measurement of each CCF noise.

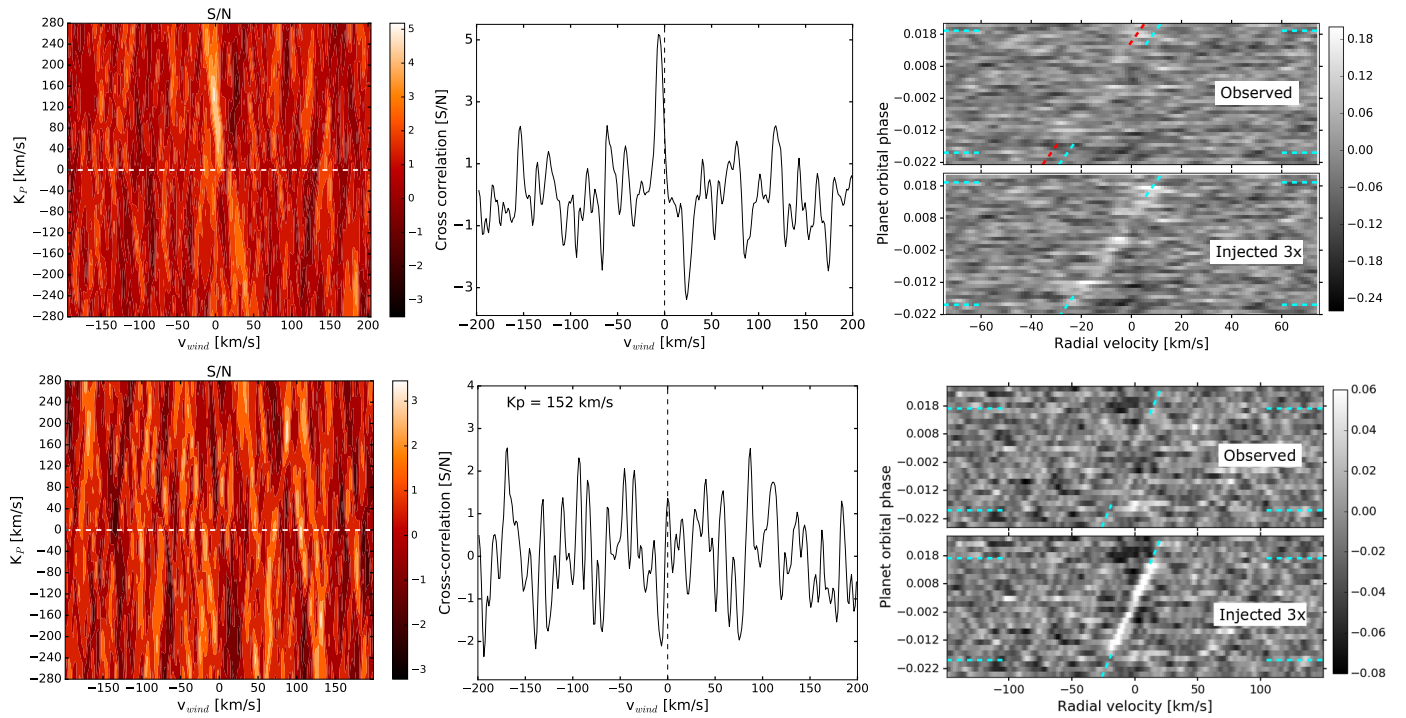
## 4. Results and discussion

### 4.1. HD 209458 b

We measured a water vapour signal on the night  $N_{A,1}$  in HD 209458 b with an S/N = 5.2 at  $K_p = 145 \pm 31 \text{ km s}^{-1}$ , which is a planet orbital velocity that is in good agreement with the widely accepted value of 145.9  $\text{km s}^{-1}$  (Brogi & Line 2017) and previous results (Snellen et al. 2010; Hawker et al. 2018; Brogi & Line 2019; Sánchez-López et al. 2019). The top row of Fig. 3 shows the S/N map with respect to the planet rest frame as well as a function of  $K_p$  (left panel) and a slice at  $K_p = 145 \text{ km s}^{-1}$ , showing the CCF peak with maximum significance (middle panel). The obtained signal follows the expected trail of the exoplanet during the observations as depicted in the cross-correlation matrix in the Earth’s rest-frame (right panel). However, it shows a significant blueshift of  $-6.5^{+3.9}_{-2.6} \text{ km s}^{-1}$ , which is consistent with what was measured in the near-infrared from the long-wavelength wing of the H<sub>2</sub>O band at  $\sim 1 \mu\text{m}$  ( $-6.5^{+2.6}_{-1.3} \text{ km s}^{-1}$ , Sánchez-López et al. 2019) and with model predictions (Rauscher & Menou 2012; Showman et al. 2013; Amundsen et al. 2016). This blueshift could be caused by global high-altitude winds blowing at the terminator from the day to the night side hemisphere of HD 209458 b.

With the objective of studying the robustness of the observed H<sub>2</sub>O signal, we repeated the cross-correlation analysis of this night by using the same number of SYSREM iterations in all the useful spectral orders, which minimises model-dependencies at the expense of a possible under- or over-correction in some of them (Alonso-Floriano et al. 2019; Stangret et al. 2020). We were able to recover a very similar 5 $\sigma$  blueshifted signal at the same planet  $K_p$  after applying five SYSREM iterations to all orders (see Fig. A.3). For a larger number of iterations, the signal was increasingly removed by the algorithm (Fig. A.4).

Furthermore, we investigated the individual contributions from the three water vapour bands probed in the 0.70–0.96  $\mu\text{m}$  spectral region. The resulting S/N maps are shown in Fig. A.5. We did not find significant signals around the expected  $K_p = 145.9 \text{ km s}^{-1}$  for any of the bands individually. The strongest contributions were observed at that  $K_p$  for the 0.82  $\mu\text{m}$  and 0.95  $\mu\text{m}$  bands, but these signals were not significant enough to claim single-band detections. Also, residual signals were observed at other velocities, making them somewhat uncertain. Therefore, the total signal discussed above did not arise from a small sample of spectral points, but rather from using a wide spectral interval that included several H<sub>2</sub>O bands. In particular, the 0.95  $\mu\text{m}$  band was already detected with the CARMENES data from its infrared channel (Sánchez-López et al. (2019) by using the long-wavelength wing of the water feature. The optical



**Fig. 3.** Cross-correlation results obtained on the night  $N_{A,1}$  for HD 209458 b (*top row*) and after combining four transits of HD 189733 b (*bottom row*). *Left column*: S/N maps of potential water vapour signals with respect to the exoplanet rest-frame (horizontal axis) and  $K_p$  (vertical axis). *Middle column*: Slice through the left panels at the  $K_p$  with the largest significance peak for HD 209458 b ( $145 \text{ km s}^{-1}$ ) and at the expected  $K_p$  of HD 189733 b. *Right column*: Cross-correlation matrices in the Earth rest-frame for the observed data and after injecting a signal at  $3\times$  the expected strength. The transit occurs between the cyan horizontal dashed lines. The cyan tilted dashed lines mark the expected velocities of the exoplanets during the observations. The tilted red dashed line in the case of HD 209458 b marks the position of the most significant signal observed.

data presented here mainly covers the short-wavelength wing, showing a weak signal with an  $S/N = 3.2$  at the expected  $K_p$ . We note that the efficiency of the optical arm of CARMENES around these long wavelengths decreases, reducing the  $S/N$  of the observations of this band.

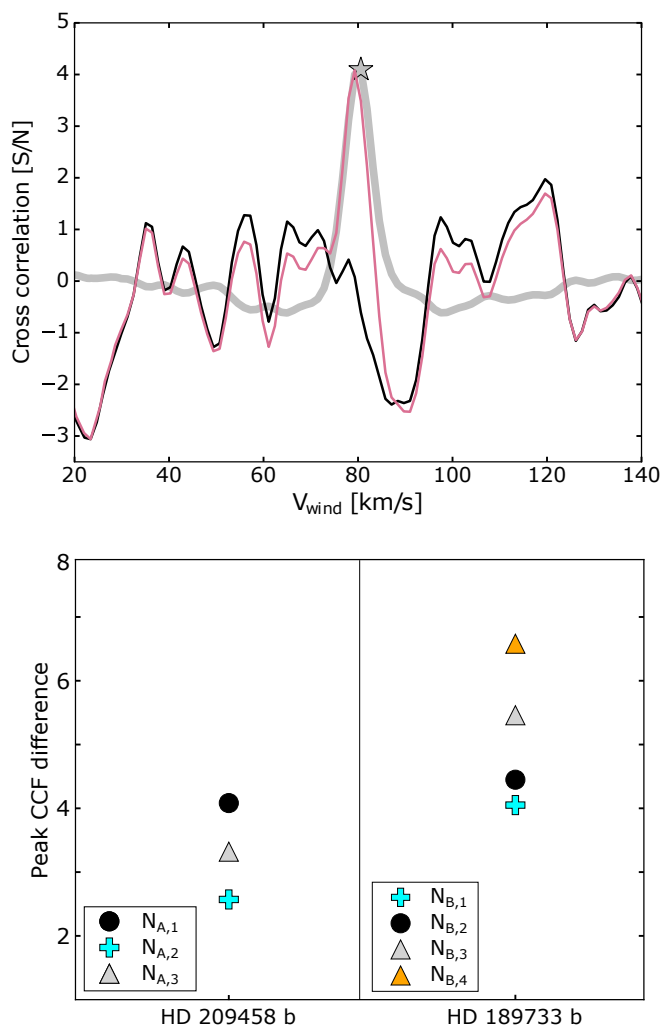
The same procedure was applied for the nights  $N_{A,2}$  and  $N_{A,3}$ . However, we did not find any evidence of planetary water absorption in these nights (see Fig. A.6). In particular, in  $N_{A,2}$ , the systematic residuals at  $0 \text{ km s}^{-1}$  (Fig. A.6, top right) caused the telluric signal at  $K_p = 0 \text{ km s}^{-1}$  to dominate the S/N map (Fig. A.6, top left). For  $N_{A,3}$ , a strong contamination at an orbital phase  $\sim 0.1$  (Fig. A.6, bottom right) produced spurious signals that peaked in the negative  $K_p$  space (Fig. A.6, bottom left).

Since we only observed a significant  $\text{H}_2\text{O}$  CCF peak in one of the three transits of HD 209458 b, we assessed the capabilities of our methodology for observing this planet’s atmosphere on the three nights. We injected signals at the expected level ( $1\times$ ) of  $\text{H}_2\text{O}$  absorption and investigated their recovery (see top panel of Fig. 4). These types of signals can be rather weak in individual orders, especially if they cover a region with low  $\text{H}_2\text{O}$  absorption. Thus, instead of using the peak value of the CCF with injection directly, we studied the difference between the CCFs with an injection and without one. This metric is robust against the contribution of potential noise sources at the velocities of the injection, which are subtracted. Furthermore, in order to avoid the influence of any potential real planetary signals or telluric residuals, we injected the model at very different velocities from those expected for the planet (i.e.  $K_p = 180 \text{ km s}^{-1}$ ,  $v_{\text{wind}} = 80.6 \text{ km s}^{-1}$ ).

We found that the injected signal was significantly better recovered on  $N_{A,1}$  with respect to  $N_{A,2}$  and  $N_{A,3}$ , in which the planet signal was at the noise level (see bottom left panel of Fig. 4). In other words, if there was  $\text{H}_2\text{O}$  optical absorption in HD 209458 b, we would most likely only be able to detect it with this methodology on  $N_{A,1}$ , which is what we observed. The reasons behind the poorer recovery of injections on  $N_{A,2}$  and  $N_{A,3}$  are unclear, although they were likely related to the poor  $S/N$  of the spectra on  $N_{A,2}$  and the unsuccessful telluric removal on both nights. The latter could be caused by a high telluric variability, inducing additional non-linear trends that hinder the performance of SYSREM. Indications of this type of variability can be inferred from the observed changes in the relative humidity during these nights (see top panel on the left side of Fig. 2). In particular, the rather rapid variability of the telluric water vapour content on  $N_{A,3}$  during the transit could also be the cause of the unsuccessful telluric removal in the reddest orders of the CARMENES near-infrared channel presented by Sánchez-López et al. (2019). Finding the exact reasons behind the different injection recoveries would require an extensive study of the telluric conditions for the different nights. However, we did not perform this type of analysis since the injection recoveries discussed above already provided us with a metric of the goodness of each data set for finding  $\text{H}_2\text{O}$  signals using the CCF technique.

#### 4.2. HD 189733 b

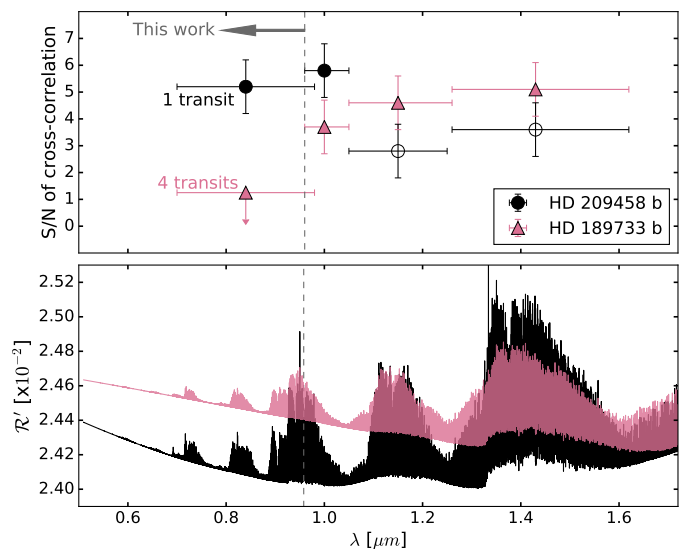
For HD 189733 b, we did not find significant water vapour signatures in any of the four transits analysed (see Fig. A.7). On  $N_{B,1}$  and  $N_{B,3}$ , the S/N maps were dominated by telluric residuals at a  $K_p$  of  $0 \text{ km s}^{-1}$  (see first and third rows in Fig. A.7). The best



**Fig. 4.** *Top:* Representative CCF obtained with an injected signal at the expected level (magenta) and without an injection (black) at a  $K_p$  of  $180 \text{ km s}^{-1}$  and  $v_{\text{wind}} = 80.6 \text{ km s}^{-1}$ . The difference between both CCFs is shown in grey and its peak is marked with a star symbol. The illustration corresponds to the results of  $N_{A,1}$ . The CCFs for the other nights have similar shapes, but different peak values. *Bottom:* Peak differences between the CCFs with and without an injection for all HD 209458 b (*left*) and HD 189733 b (*right*) data sets.

recovery of injections was found on  $N_{B,3}$  and  $N_{B,4}$  (see bottom right panel in Fig. 4), for which we did not observe any evidence of  $\text{H}_2\text{O}$  optical absorption. The maximum significance CCF at the expected  $K_p$  of  $152.5 \text{ km s}^{-1}$  was found for  $N_{B,2}$  with an S/N of 3.1. However, this night presented a significantly weaker recovery of injections, suggesting that this low-significance CCF peak was likely produced by noise. Given the similar S/N of the four transit measurements, the different observational conditions were likely the reason for the different strength of the injection recoveries.

In order to improve the detectability of a potential signal from HD 189733 b, we combined the cross-correlation matrices from the four transits weighted by their respective strength of injection recoveries (see bottom row in Fig. 3). However, it did not allow us to identify any significant signals either. We repeated the analysis by combining only  $N_{B,2}$  and  $N_{B,4}$ , which presented fewer telluric residuals in their S/N maps and also had the most stable relative humidity conditions (see Fig. 2), but we found similar results (i.e. non-detection of optical  $\text{H}_2\text{O}$  absorption).



**Fig. 5.** *Top:* S/N of the  $\text{H}_2\text{O}$  cross-correlation signal obtained for different wavelength intervals (indicated by horizontal bars) for HD 209458 b (black and open circles) and HD 189733 b (magenta triangles) with CARMENES. The  $1\sigma$  uncertainties are indicated by the vertical bars. The results in the  $0.96\text{--}1.71 \mu\text{m}$  spectral region were obtained with one transit for HD 209458 b and HD 189733 b in Sánchez-López et al. (2019). Open symbols show spectral regions for which the measurements of HD 209458 b are uncertain due to a problematic telluric removal. The larger signal towards shorter wavelengths in HD 209458 b compared to HD 189733 b is indicative of the higher atmospheric transparency of the former planet. *Bottom:* Model  $\text{H}_2\text{O}$  transmission spectra for HD 209458 b assuming a clear atmosphere (black) and for HD 189733 b including a stronger haze extinction at an arbitrary level (magenta). The spectrum of HD 209458 b is offset by  $10^{-2}$  for illustration purposes.

However, the injected models of HD 189733 b at the expected strength were well recovered in all data sets, which suggests that it would be possible to detect water vapour if the relative depth of the planetary  $\text{H}_2\text{O}$  lines were similar to those of our cloud-free model. In assuming the  $\sim 1\sigma$  CCF peak that we observed at the expected velocities (i.e.  $K_p = 152.5 \text{ km s}^{-1}$  and  $v_{\text{wind}} = 0 \text{ km s}^{-1}$ ) was a real planet atmospheric signal, we estimated that between 60 and 70 transits, observed under similar conditions to those presented here, would be required to detect the water vapour bands in the red optical at  $4\sigma$  with CARMENES.

#### 4.3. Comparison of the optical $\text{H}_2\text{O}$ analysis in both planets

The results obtained for both planets when applying the same  $\text{H}_2\text{O}$  cross-correlation analyses allowed us to compare the two different atmospheres. We were able to detect  $\text{H}_2\text{O}$  in HD 209458 b from the red optical bands, whereas we did not find any significant signals in four transit data sets of HD 189733 b over the same spectral range. However, previous studies have reported the detection of a similarly strong  $\text{H}_2\text{O}$  absorption in both planets in the near-infrared (see top panel in Fig. 5). In this context, we interpret the non-detection of  $\text{H}_2\text{O}$  optical signals in HD 189733 b to be due to the presence of atmospheric hazes, which cause a strong extinction and mute most of the water vapour molecular features (see bottom panel in Fig. 5). In contrast, we found that HD 209458 b presents a clearer atmosphere, which facilitates the identification of the weak  $\text{H}_2\text{O}$  optical signatures. Furthermore, these results are in line with space observations of these objects, which found a steep Rayleigh scattering

slope in HD 189733 b, but a more gentle one in HD 209458 b (Sing et al. 2016; Barstow et al. 2017). However, we were not able to discard the presence of other opacity contributors in the red-optical, such as TiO and VO, which might also present a contribution. This was because KOPRA does not yet include the opacities of these molecules and, in the case of the publicly available *petitRADTRANS* (Mollière et al. 2019), the linelists used are those from Plez (1998), which might not be accurate enough for high-resolution CCF studies (Merritt et al. 2020).

## 5. Conclusions

We have presented the analysis of water vapour atmospheric signals in HD 209458 b and HD 189733 b in the 0.70–0.96  $\mu\text{m}$  spectral interval, applying the cross-correlation technique to high-resolution CARMENES data of seven transits. We detected H<sub>2</sub>O in the atmosphere of HD 209458 b using one transit data set, but we were not able to reveal H<sub>2</sub>O with confidence in HD 189733 b after combining four transits. Since the latter planet exhibits a similar or even somewhat stronger H<sub>2</sub>O signal in the near-infrared bands, we attribute the relative weakness of the optical water bands in HD 189733 b to the presence of haze extinction, as already proposed for this planet from the strong UV and blue optical Rayleigh scattering signal (Sing et al. 2016; Barstow et al. 2017). This study shows that a distinct level of aerosol extinction in different exoplanets can be inferred in a similar way as proposed by Pino et al. (2018) from currently available observations of high-resolution spectrographs on 4 m-class ground-based telescopes. However, single band detections and comparisons are very challenging with our data sets due to the presence of telluric residuals, systematics, and other noise sources. This is especially important in the red optical, where the information contained in three H<sub>2</sub>O bands had to be combined to detect water vapour in one transit of HD 209458 b. Moreover, our tests using injected signals highlight the importance of telluric stability when analysing several transits of the same planet. Having similar observational conditions at Calar Alto on the different nights is a key ingredient in our capacity to disentangle planet atmospheric signals with CARMENES using transit data and cross-correlation.

High-resolution observations at higher S/N and with better observational conditions than those we analysed could allow the exploration of the aerosol content in hot-Jupiter atmospheres using the methods proposed by Pino et al. (2018). In addition, this could be attempted by performing retrievals (e.g. Brogi & Line 2017, 2019; Fisher et al. 2020), including the strength of the wavelength-dependent haze extinction as an additional parameter. For instance, a simple model for small aerosol particles could be implemented as in Sing et al. (2016) in the calculation of the templates. By using a wide spectral interval, the different strengths of H<sub>2</sub>O bands should favour atmospheric models that include a larger opacity due to hazes, hence helping us constrain their general contribution. Thus, these analyses can complement the current knowledge obtained from space observations by using a different method.

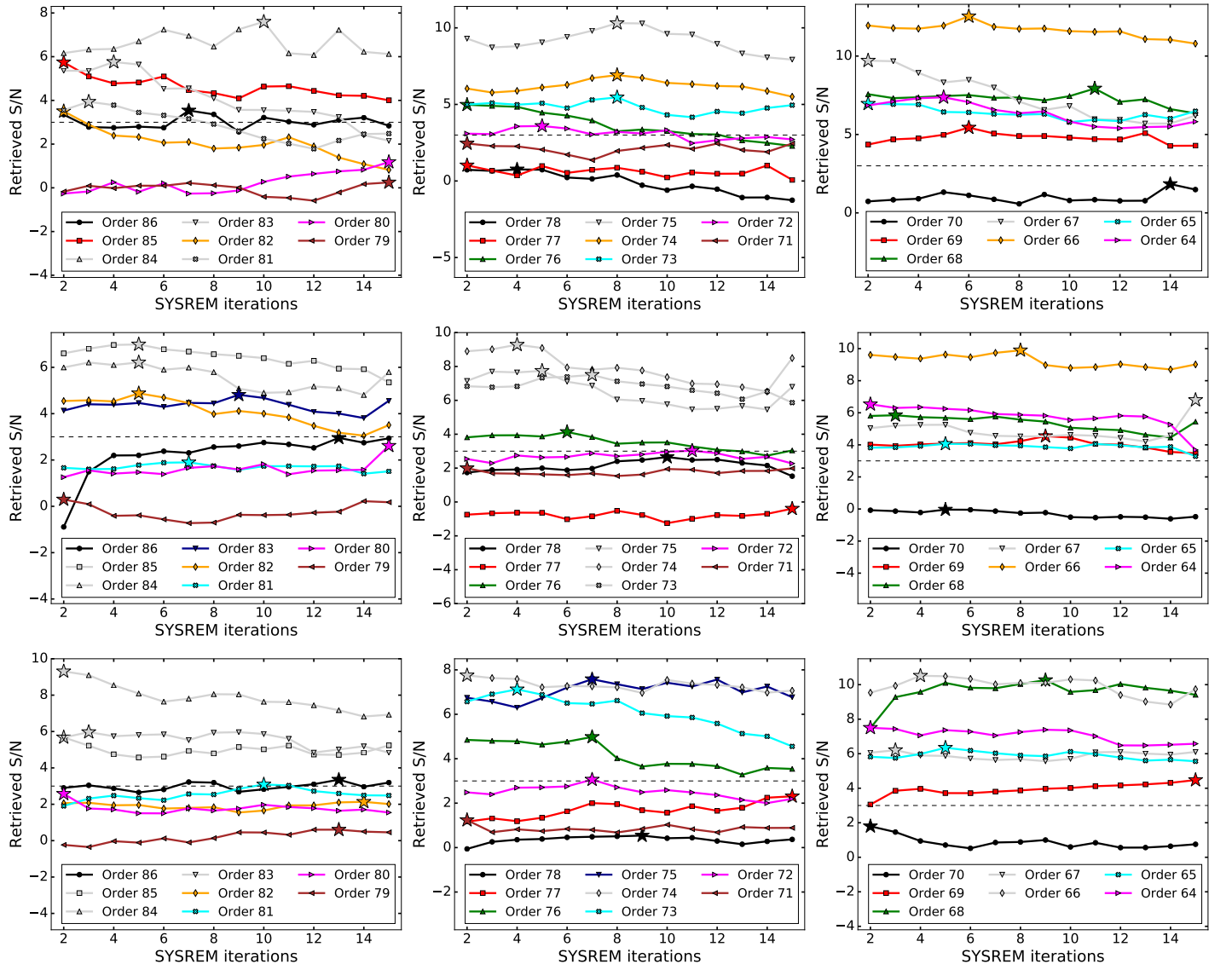
*Acknowledgements.* A.S.L. and I.S. acknowledge funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation program under grant agreement No 694513. CARMENES is an instrument for the Centro Astronómico Hispano-Alemán de Calar Alto (CAHA, Almería, Spain). CARMENES is funded by the German Max-Planck-Gesellschaft (MPG), the Spanish Consejo Superior de Investigaciones Científicas (CSIC), the European Union through FEDER/ERF FICTS-2011-02 funds, and the members of the CARMENES Consortium (Max-Planck-Institut für Astronomie, Instituto de Astrofísica de Andalucía, Landessternwarte Königstuhl, Institut de Ciències de l’Espai, Insitut für Astrophysik Göttingen, Universi-

dad Complutense de Madrid, Thüringer Landessternwarte Tautenburg, Instituto de Astrofísica de Canarias, Hamburger Sternwarte, Centro de Astrobiología and Centro Astronómico Hispano-Alemán), with additional contributions by the the Spanish Ministerios de Ciencia e Innovación and of Economía y Competitividad, the Fondo Europeo de Desarrollo Regional (FEDER/ERF), the Agencia estatal de investigación, the Fondo Social Europeo under grants AYA2011-30 147-C03-01, -02, and -03, AYA2012- 39612-C03-01, ESP2013-48391-C4-1-R, ESP2014-54062-R, ESP 2016-76076- R, ESP2016-80435-C2-2-R, ESP2017-87143-R, PGC2018-098153-B-C31, BES-2015-073500, and BES- 2015-074542, the German Science Foundation through the Major Research Instrumentation Programme and DFG Research Unit FOR2544 “Blue Planets around Red Stars”, the Klaus Tschira Stiftung, the states of Baden-Württemberg and Niedersachsen, and by the Junta de Andalucía. IAA authors acknowledge financial support from the State Agency for Research of the Spanish MCIU through the “Center of Excellence Severo Ochoa” award SEV-2017-0709. Based on observations collected at the Centro Astronómico Hispano-Alemán (CAHA) at Calar Alto, operated jointly by Junta de Andalucía and Consejo Superior de Investigaciones Científicas (IAA-CSIC). We thank the anonymous referee for the very useful comments.

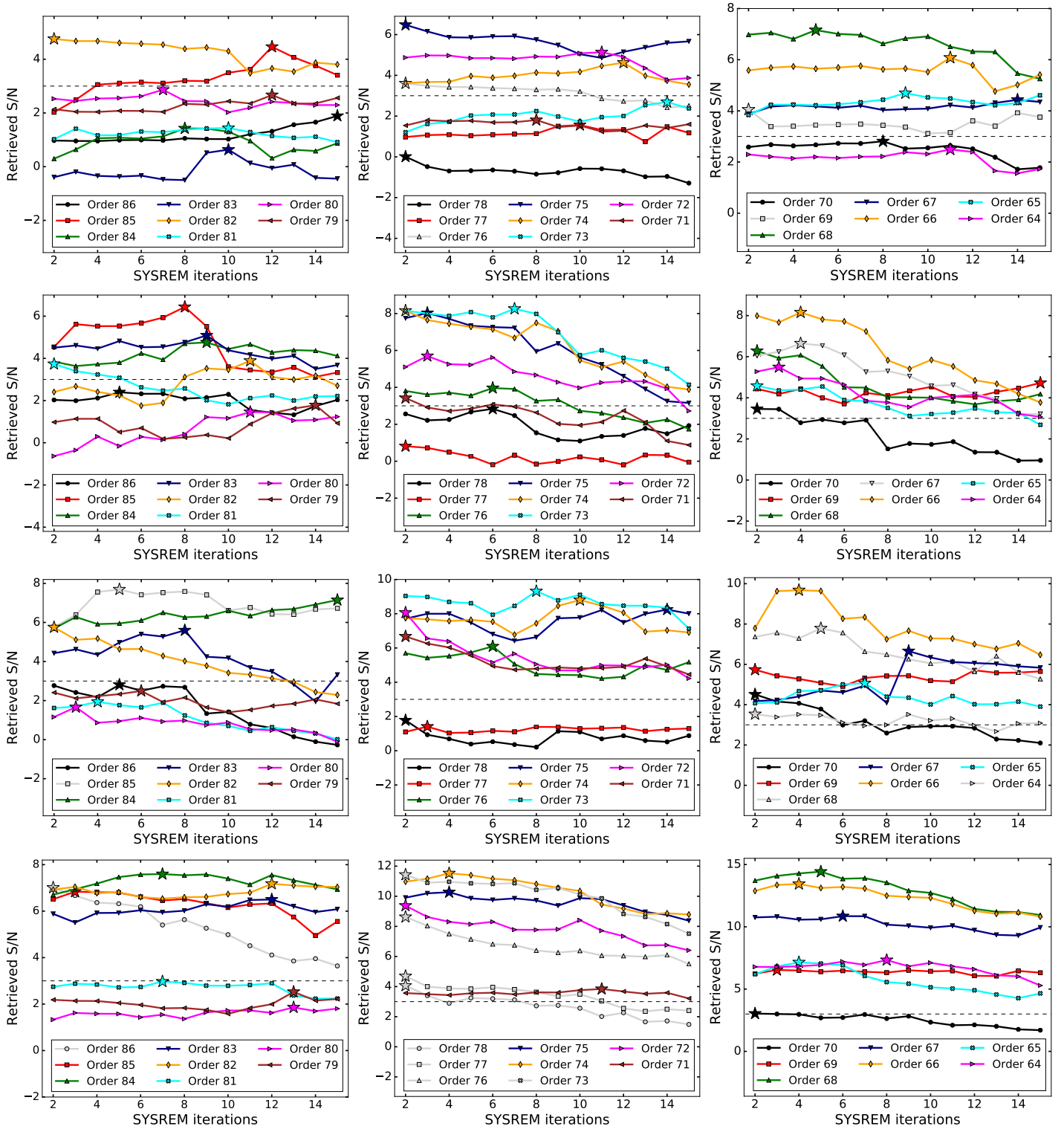
## References

- Amundsen, D. S., Mayne, N. J., Baraffe, I., et al. 2016, *A&A*, 595, A36  
 Allart, R., Lovis, C., Pino, L., Wyttenbach, A., Ehrenreich, D., & Pepe, F. 2017, *A&A*, 606, A144  
 Alonso-Floriano, F. J., Sánchez-López, A., Snellen, I. A. G., López-Puertas, M., et al. 2019, *A&A*, 621, A74  
 Barstow, J. K., Aigrain, S., Irwin, P. G., & Sing, D., K. 2017, *ApJ*, 834, 50  
 Bouchy, F., Udry, S., Mayor, M., Moutou, C., et al. 2005, *A&A*, 444, 1  
 Brogi, M., Line, M., Bean, J., Désert, J.-M., & Schwarz, H. 2017, *ApJ*, 839, L2  
 Brogi, M., Giacobbe, P., Guilluy, G., et al. 2018, *A&A*, 615, A16  
 Brogi, M., & Line, M. 2019, *AJ*, 157, 114  
 Birkby, J. L., de Kok, R. J., Brogi, M., et al. 2013, *MNRAS*, 436, L35  
 Birkby, J. L., de Kok, R. J., Brogi, M., Schwarz, H., & Snellen, I. A. G. 2017, *AJ*, 153, 138  
 Birkby, J. L., 2018, *Handbook of Exoplanets*, 1485–1508, Springer  
 Caballero, J. A., Guàrdia, J., López del Fresno, M., et al. 2016, *Proc. SPIE*, 9910, 0E  
 Cabot, S. H. C., Madhusudhan, N., Hawker, G. A., Gandhi, S., 2019, *MNRAS*, 482, 4422  
 Casasayas-Barris, N., Pallé, E., Yan, F., Chen, G., et al. 2018, *A&A*, 616, A151  
 Casasayas-Barris, N., Pallé, E., Yan, F., Chen, G., et al. 2019, *A&A*, 628, A9  
 Charbonneau, D., Brown, T. M., Latham, D. W., & Mayor, M. 2000, *ApJ*, 529, 45  
 de Kok, R. J., Brogi, M., Snellen, I. A. G., et al. 2013, *A&A*, 554, A82  
 Donati, J. F. 2003, *ASP Conf. Ser.* 307, *Solar Polarization* ed J. Trujillo-Bueno and J. S. Almeida (San Francisco, CA: ASP) 41  
 Ehrenreich, D., Lovis, C., Allart, R., et al. 2020, *Nature*, 580(7805), 597-601.  
 Esteves, L. J., De Mooij, E. J., Jayawardhana, R., Watson, C., & De Kok, R. 2017, *AJ*, 153, 268  
 Fisher, C., Højimakers, H. J., Kitzmann, D., Márquez-Neila, P., et al. 2020, *AJ*, 159, 192  
 Fortney, J. J., Lodders, K., Marley, M. S., & Freedman, R. S. 2008, *ApJ*, 678(2), 1419  
 Gibson, N. P., Merritt, S., Nugroho, S. K., Cubillos, P. E., de Mooij, E. J., et al. 2020, *MNRAS*, 493, 2215  
 Guilluy, G., Sozzetti, A., Brogi, M., Bonomo, A. S., et al. 2019, *A&A*, 625, A107  
 Hawker, G. A., Madhusudhan, N., Cabot, S. H. C., & Gandhi, S. 2018, *ApJ*, 863, L11  
 Henry, G. W., Marcy, G. W., Butler, P. R., & Vogt, S. S. 2000, *ApJ*, 529, 41  
 Højimakers, H. J., Ehrenreich, D., Heng, K., et al. 2018, *Nature*, 560, 453  
 Mazeh, T., Tamuz, O., & Zucker, S. 2007, *Transiting Extrasolar Planets Workshop*, 366, 119  
 MacDonald, R. J., & Madhusudhan, N. 2017, *MNRAS*, 469, 1979  
 Mayor, M., Pepe, F., Queloz, D., Bouchy, F., et al. 2003, *The Messenger*, 114, 20  
 Merritt, S. R., Gibson, N. P., Nugroho, S. K., de Mooij, E. J., et al. 2020, *A&A*, 636, A117  
 Molaverdikhani, K., Henning, T., & Mollière, P. 2019, *ApJ*, 883, 194  
 Mollière, P., Wardenier, J. P., van Boekel, R., Henning, T., et al. 2019, *A&A*, 627, A67  
 Noguchi, K., Aoki, W., Kawanamoto, S., Ando, H., et al. 2002, *PASJ*, 54, 855  
 Nugroho, S. K., Kawahara, H., Masuda, K., et al. 2017, *AJ*, 154, 221  
 Pino, L., Ehrenreich, D., Allart, R., et al. 2018, *A&A*, 619, A3  
 Plez, B. 1998, *A&A*, 337, 495  
 Quirrenbach, A., Amado, P. J., Caballero, J. A., et al. 2016, *Proc. SPIE*, 9908, 12  
 Quirrenbach, A., Amado, P. J., Ribas, I., et al. 2018, *Proc. SPIE*, 10702, 0W  
 Rauscher, E., & Menou, K., 2012, *ApJ*, 750, 96

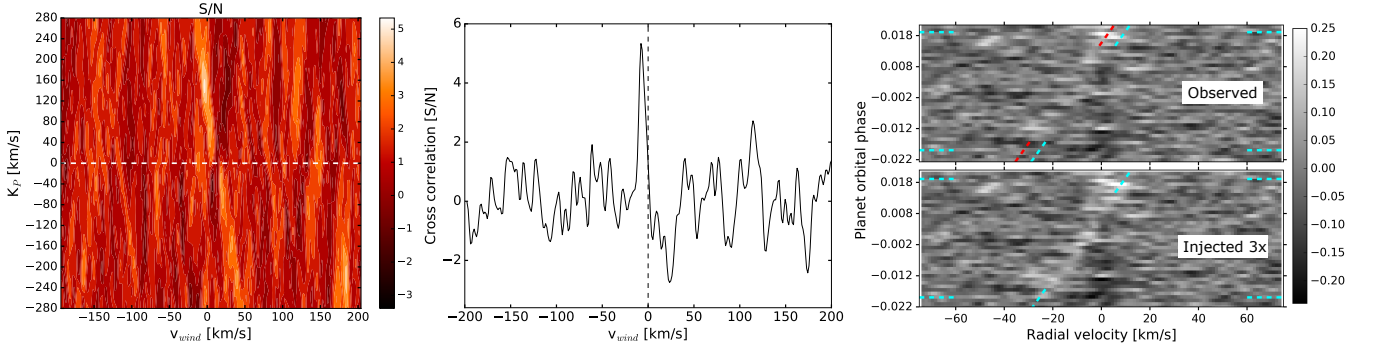
- Sánchez-López, A., Alonso-Floriano, F. J., López-Puertas, M., Snellen, I. A. G., et al. 2019, A&A, 630, A53
- Showman, A. P., Fortney, J. J., Lewis, N. K., & Shabram, M. 2013, ApJ, 762, 1
- Sing, D. K., Fortney, J. J., Nikolov, N., et al. 2016, Nature, 529, 59
- Snellen, I. A. G., de Kok, R. J., de Mooij, E. J. W., & Albrecht, S. 2010, Nature, 465, 1049
- Stangret, M., Casasayas-Barris, N., Pallé, E., Yan, F., Sánchez-López, A., & López-Puertas, M. 2020, A&A, 638, A26.
- Stiller, G. P., von Clarmann, T., Funke, B., et al. 2002, Journal of Quantitative Spectroscopy and Radiative Transfer, 72, 249
- Tamuz, O., Mazeh, T., & Zucker, S. 2005, MNRAS, 356, 1466
- Yan, F., & Henning, T. 2018, Nature Astronomy, 2, 714
- Zechmeister, M., Anglada-Escudé, G., & Reiners, A. 2014, A&A, 561, A59

**Appendix A: Supplementary figures**


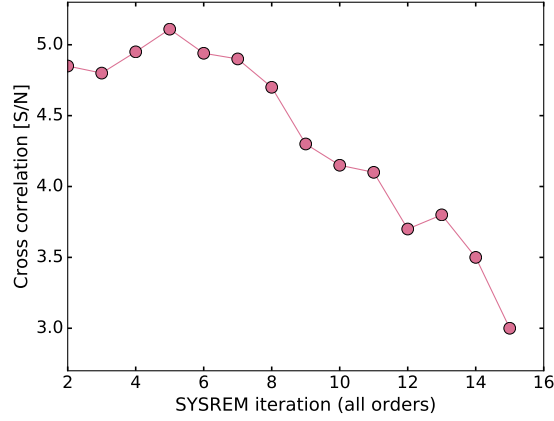
**Fig. A.1.** Evolution of the S/N of the retrieved injected signal with subsequent SYSREM iterations for HD 209458 b on the nights of  $N_{A,1}$  (top),  $N_{A,2}$  (middle), and  $N_{A,3}$  (bottom) for the H<sub>2</sub>O bands at  $\sim 0.72 \mu\text{m}$  (left),  $\sim 0.82 \mu\text{m}$  (central), and  $\sim 0.95 \mu\text{m}$  (right). The model signals were injected at 5 $\times$  the expected strength. The horizontal dashed lines mark the S/N = 3 level. The star symbols mark the iteration in which the recovery of the injected signal is maximised.



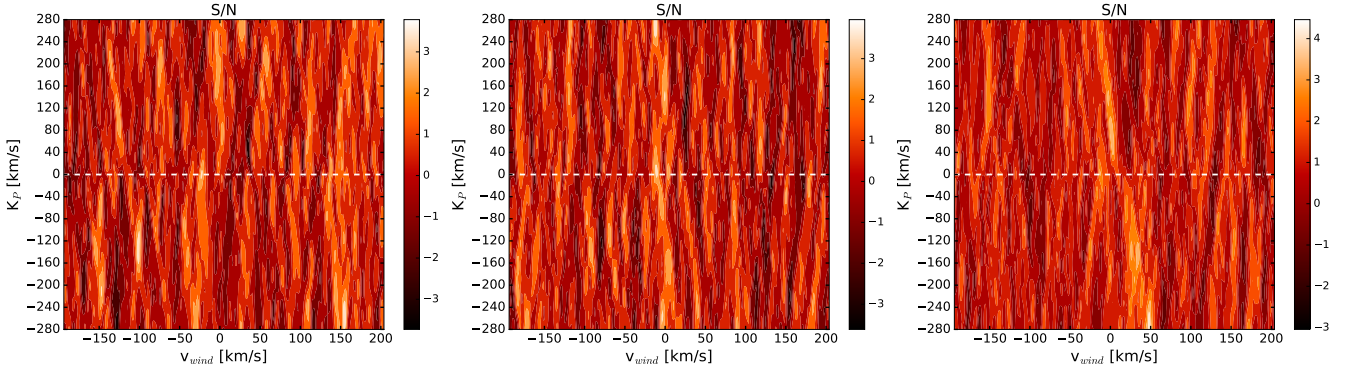
**Fig. A.2.** Same as Fig. A.1, but for HD 189733 b on the nights of  $N_{B,1}$  (first row),  $N_{B,2}$  (second row),  $N_{B,3}$  (third row), and  $N_{B,4}$  (fourth row).



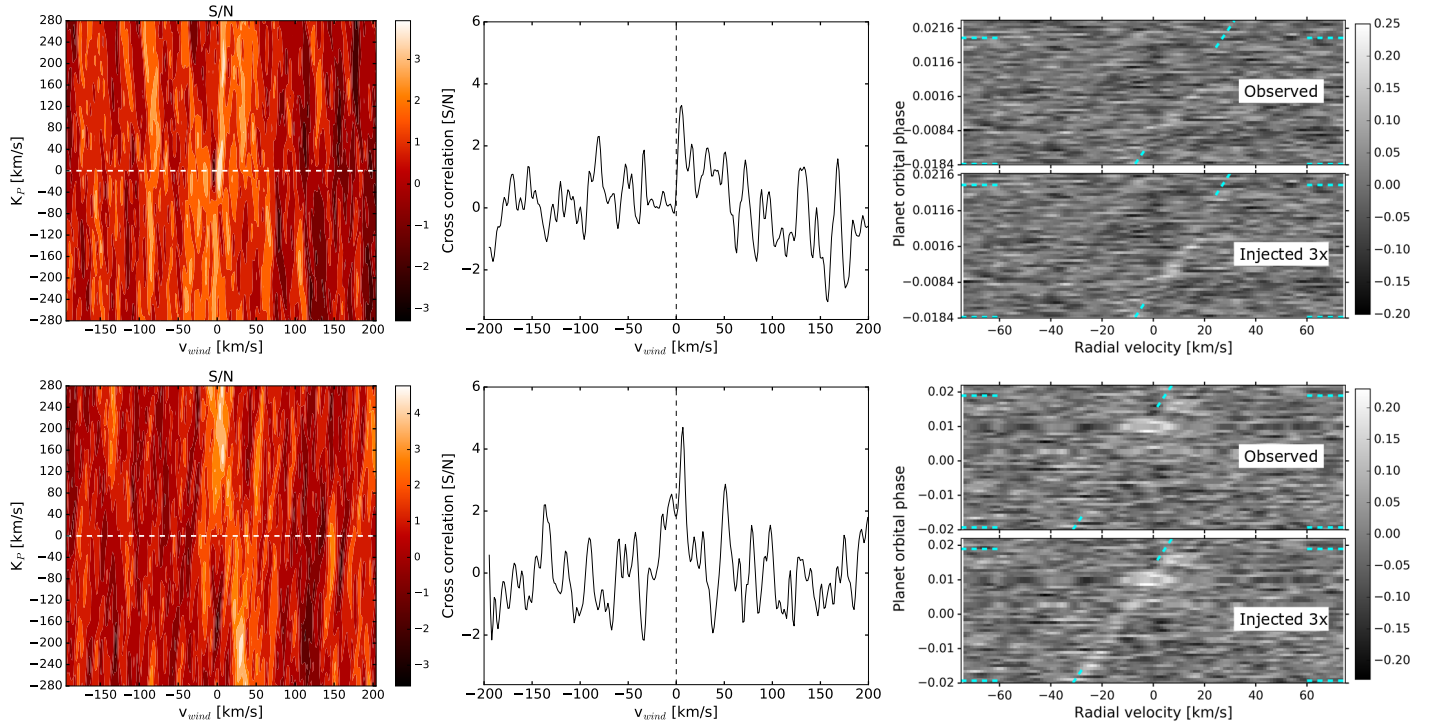
**Fig. A.3.** Same as the top row of Fig. 3, but for HD 209458 b when using five SYSREM iterations for all spectral orders.



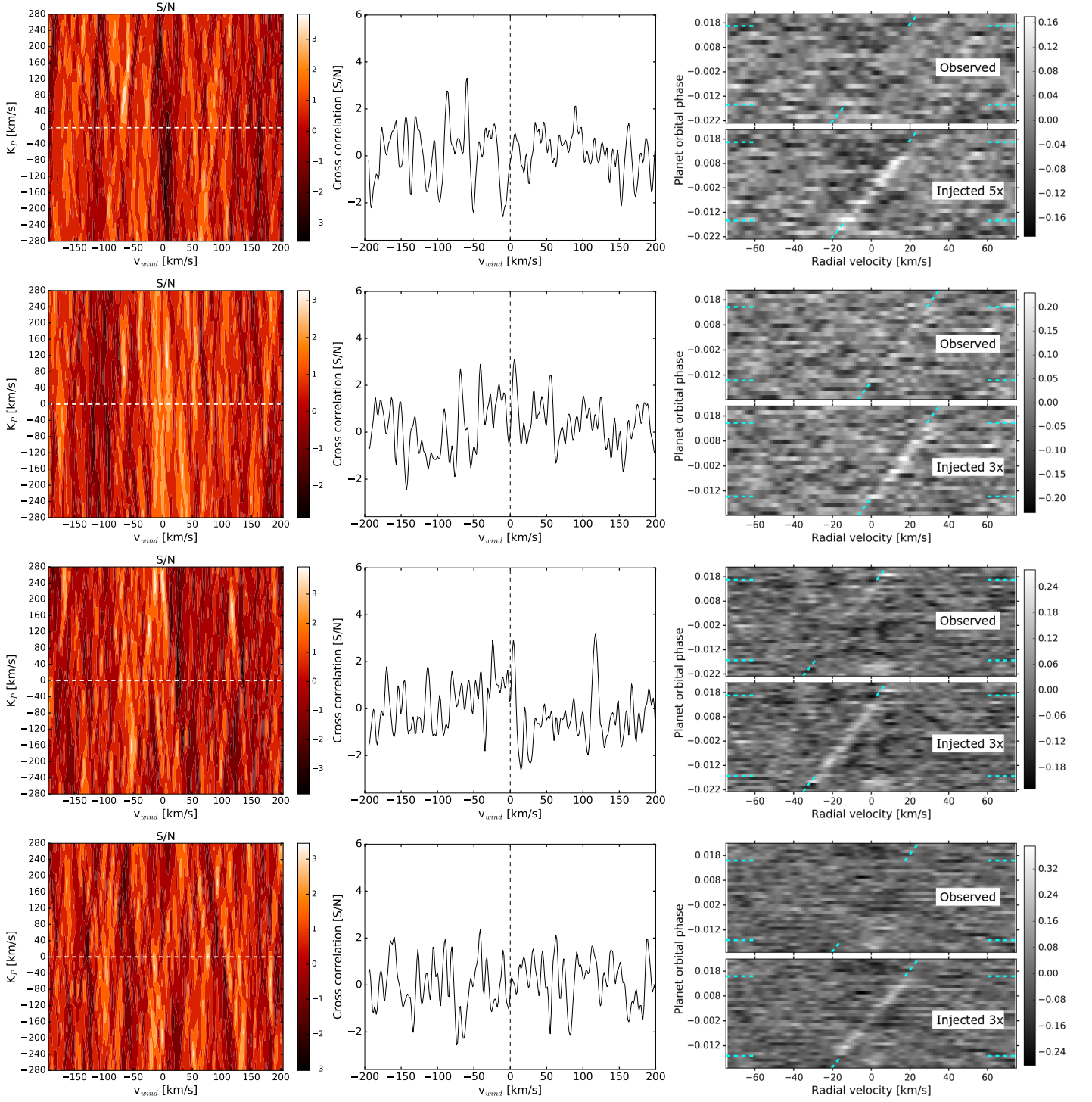
**Fig. A.4.** S/N of the CCF peak obtained in  $N_{A,1}$  for HD 209458 b when cross-correlating the residual matrix after each SYSREM iteration with the  $\text{H}_2\text{O}$  absorption model. The values correspond to the highest S/N at a  $K_p$  in the range of 130 to 160  $\text{km s}^{-1}$  and  $v_{\text{wind}}$ , ranging from  $-10$  to  $+10$   $\text{km s}^{-1}$ . The first iteration is omitted since it is mostly dominated by telluric residuals.



**Fig. A.5.** S/N map for HD 209458 b for potential water signals with respect to the exoplanet rest-frame (horizontal axis) and  $K_p$  (vertical axis) for the  $\sim 0.72$   $\mu\text{m}$  band (left panel), the  $\sim 0.82$   $\mu\text{m}$  band (middle panel), and the  $\sim 0.95$   $\mu\text{m}$  band (right panel).



**Fig. A.6.** Left column: S/N maps for HD 209458 b on the nights  $N_{A,2}$  (top) and  $N_{A,3}$  (bottom) for potential water signals with respect to the exoplanet rest-frame (horizontal axis) and  $K_p$  (vertical axis). The horizontal dashed lines mark the  $K_p = 0 \text{ km s}^{-1}$  value. Middle column: Slice through the left panels at the expected  $K_p$  for HD 209458 b ( $146 \text{ km s}^{-1}$ ). Right column: Cross-correlation matrices in the Earth rest-frame. The transits occur between the horizontal dashed lines. The tilted dashed lines trace the expected planet velocities during the observations.



**Fig. A.7.** Same as A.6, but for HD 189733 b ( $K_p = 152 \text{ km s}^{-1}$ ) on the nights  $N_{B,1}$  (first row),  $N_{B,2}$  (second row),  $N_{B,3}$  (third row), and  $N_{B,4}$  (fourth row).