The Transmission Spectrum of WASP-17 b From the Optical to the Near-infrared Wavelengths: Combining STIS, WFC3, and IRAC Data Sets

Arianna Saba1, Angelos Tsiaras1,2, Mario Morvan1, Alexandra Thompson1, Quentin Changeat1, Billy Edwards1,3, Andrew Jolly4, Ingo Waldmann5, and Giovanna Tinetti6
1 Department of Physics and Astronomy, University College London, Gower Street, WC1E 6BT London, UK; arianna.saba.20@ucl.ac.uk
2 INAF—Osservatorio Astrofisico di Arcetri, Largo E. Fermi 5, I-50125 Firenze, Italy
3 AIM, CEA, CNRS, Université Paris-Saclay, Université de Paris, F-91191 Gif-sur-Yvette, France
4 School of Physics, University of New South Wales, Sydney, NSW 2052, Australia
Received 2021 August 31; revised 2022 April 17; accepted 2022 April 27; published 2022 June 6

Abstract

We present the transmission spectrum of the inflated hot Jupiter WASP-17 b, observed with the STIS and WFC3 instruments aboard the Hubble Space Telescope, allowing for a continuous wavelength coverage from ~0.4 to ~1.7 μm. Observations taken with IRAC channels 1 and 2 on the Spitzer Space Telescope are also included, adding photometric measurements at 3.6 and 4.5 μm. HST spectral data were analyzed with Iraclis, a pipeline specialized in the reduction of STIS and WFC3 transit and eclipse observations. Spitzer photometric observations were reduced with the TLCD-LSTM method, utilizing recurrent neural networks. The outcome of our reduction produces incompatible results between STIS visit 1 and visit 2, which leads us to consider two scenarios for G430L. Additionally, by modeling the WFC3 data alone, we can extract atmospheric information without having to deal with the contrasting STIS data sets. We run separate retrievals on the three spectral scenarios with the aid of TauReX 3, a fully Bayesian retrieval framework. We find that, independently of the data considered, the exoplanet atmosphere displays strong water signatures and, potentially, the presence of aluminum oxide and titanium hydride. A retrieval that includes an extreme photospheric activity of the host star is the preferred model, but we recognize that such a scenario is unlikely for an F6-type star. Due to the incompleteness of all STIS spectral light curves, only further observations with this instrument would allow us to properly constrain the atmospheric limb of WASP-17 b, before the James Webb Space Telescope or Ariel will come online.

Unified Astronomy Thesaurus concepts: Exoplanet atmospheres (487); Hubble Space Telescope (761); Astronomy data reduction (1861); Bayesian statistics (1900)

1. Introduction

In recent years, the Hubble Space Telescope (HST) has been at the forefront of exoplanet atmospheric characterization. Thanks to its Wide Field Camera 3 (WFC3) and its Space Telescope Imaging Spectrograph (STIS), often combined with the InfraRed Array Camera (IRAC) aboard the Spitzer Space Telescope, the atmospheres of hot Jupiters (Sing et al. 2016; Pinhas et al. 2018; Tsiaras et al. 2018; Pluriel et al. 2020a; Skaf et al. 2020), ultrahot Jupiters (Kreidberg et al. 2018; Mikal-Evans et al. 2019; Changeat et al. 2020; Edwards et al. 2020a), super-Earths (Tsiaras et al. 2016a, 2019; Edwards et al. 2020b; Mugnai et al. 2021), sub-Neptunes (Guilluy et al. 2020; Guo et al. 2020), and hot Saturns (Anisman et al. 2020; Carone et al. 2021) have been studied in transmission, in emission, and across their phase curves (Knutson et al. 2012; Stevenson et al. 2017; Kreidberg et al. 2018; Arcangeli et al. 2019; Feng et al. 2020; Irwin et al. 2020; Changeat et al. 2021). Although small rocky planets are difficult to probe with current spectrographs owing to the limited signal-to-noise ratio (S/N) of their gaseous envelope, atmospheric studies conducted on extrasolar gas giants reported the detection of various molecules, in particular water vapor.

Among the exoplanet population of inflated hot Jupiters, WASP-17 b (Anderson et al. 2010) is one of the least dense planets discovered so far, with a radius equal to 1.932 ± 0.053 Rj and a mass 0.477 times that of Jupiter (Southworth et al. 2012). The transiting planet, confirmed to possess a retrograde orbit (Bayliss et al. 2010; Hébrard et al. 2011; Simpson et al. 2011), forms an angle of 148.5° with respect to the rotation axis of its parent star (Triaud et al. 2010). The considerable orbital obliquity of WASP-17 b cannot be explained by inward migration (Lin et al. 1996; Ida & Lin 2004) but could rather be due to an initial misalignment between the star and the protoplanetary disk (Bate et al. 2010), planet–planet scattering, or the Kozai mechanism (Nagasawa et al. 2008). These processes also induce the very small eccentricities reported in orbitally misaligned planets (Johnson et al. 2009; Winn et al. 2009; Bourrier et al. 2018); in fact, Southworth et al. (2012) were not able to find any evidence of the orbital eccentricity of WASP-17 b. The exoplanet takes 3.7 days to orbit its host star (Southworth et al. 2012), an F6-type star with V mag = 11.6 (Anderson et al. 2010).

Given the planet’s very large atmospheric scale height (Anderson et al. 2010), transmission spectroscopy studies have been extremely successful. From the ground, sodium was detected at optical wavelengths by Wood et al. (2011) and Zhou & Bayliss (2012) using the Very Large Telescope (VLT) and by Khalafinejad et al. (2018) with the Magellan Telescope. At low resolution, the spectrograph FORS2 mounted on the VLT detected potassium in the atmosphere of WASP-17 b with a 3σ significance and water absorption (Sedaghati et al. 2016). From space, data taken with HST in the optical and near-infrared and with Spitzer at 3.6 and 4.5 μm report water

Original content from this work may be used under the terms of the Creative Commons Attribution 4.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
absorption and alkali features, with little to no clouds/hazes (Sing et al. 2016). A more recent atmospheric study on WASP-17 b (Alderson et al. 2022) further confirms the presence of H$_2$O in addition to CO$_2$ absorption in the infrared. Lastly, the study conducted by Barstow et al. (2016) reports the presence of water at solar abundances in the limb of WASP-17 b, possibly accompanied by CO signatures. However, contrary to the results by Sing et al. (2016), their best atmospheric model includes scattering aerosols at relatively high altitudes, up to 1 mbar.

Here we present the spectroscopic analysis performed on transit observations taken with STIS (gratings G430L and G750L) and WFC3 (grisms G102 and G141) on board HST and with IRAC channels 1 and 2 of the Spitzer Space Telescope. The combination of the aforementioned instruments allows us to investigate the spectrum of WASP-17 b from the optical to the near-infrared wavelengths (0.4–5 μm).

2. Methodology

2.1. HST Data Processing

2.1.1. WFC3

We downloaded the HST/WFC3 raw spectroscopic and spatially scanned observations of WASP-17 b from the Mikulski Archive for Space Telescopes (MAST), as part of the HST Proposal 14918 (P.I. Hannah Wakeford). In our study we decided to not include the HST/WFC3 data set from proposal 12181, obtained in staring mode, because of its low S/N. The data would produce error bars 6–7 times bigger compared to the other WFC3 observations, obtained in spatial scanning mode.

Two transits of the exoplanet were observed with the grisms G102 and G141, which cover wavelengths in the ranges of 0.8–1.1 μm and 1.1–1.7 μm, respectively. We employed Iraclis (Tsiaras et al. 2016b), an open-source reduction pipeline for WFC3, to reduce the data and extract the light curves by performing the following steps: zero-read subtraction, reference pixels correction, nonlinearity correction, dark current subtraction, gain conversion, sky background subtraction, calibration, flat-field correction, bad-pixel and cosmic-ray correction, and light-curve extraction.

The WFC3 detector induces two different time-dependent systematics to the light curve: one exponential, at the beginning of each HST orbit, and one linear throughout the visit. Iraclis corrects for these systematics by fitting a transit fit model $F(t)$, computed using Pylightcurve, on the white light curve. $F(t)$ is a function of the limb-darkening coefficients, $R_p/R_e$, and the orbital parameters $T_0, P, i, a/R_e, e, \omega$. The transit fit function is then multiplied by a normalization factor $n_0$ and a function $R(t)$ containing a linear and an exponential term as follows:

$$R(t) = (1 - r_b(t - T_0))(1 - r_{bl}e^{-\tau_b(t-t_0)}).$$ (1)

Here $t$ is time, $T_0$ is the mid-transit time, $t_0$ is the time when each orbit starts, $r_b$ is the slope of the linear term, and $(r_{bl}, r_{t0})$ are the coefficients of the exponential ramp. The spectral light curves are first fitted using Equation (1). Then, they are divided by the white light curve and further fitted via

$$n_\lambda(1 + \chi_\lambda(t - t_0))(F(\lambda, t)/F_w(t)), $$ (2)

where $n_\lambda$ is a wavelength-dependent normalization factor, $\chi_\lambda$ is the coefficient of the wavelength-dependent linear slope, $t$ is time, $T_0$ is the mid-transit time, $F(\lambda, t)$ is the wavelength-dependent transit model, and $F_w(t)$ is the best-fit model on the white light curve.

Further information on Iraclis can be found in Tsiaras et al. (2016a, 2016b, 2018).

2.1.2. STIS

We downloaded the HST/STIS raw spectroscopic observations of WASP-17 b from the Mikulski Archive for Space Telescopes (MAST), as part of the HST Proposal 12473 (P.I. David Sing).

Three transits of the exoplanet were observed with the gratings G430L (two transits) and G750L (one transit), which cover wavelengths in the ranges of 0.3–0.57 μm and 0.5–0.94 μm, respectively. For the analysis of these data sets we used again Iraclis but adapted most of the reduction steps to STIS. These steps are bias level subtraction, bias correction, image subtraction, dark image subtraction, flat-field correction, calibration, background subtraction, bad-pixel and cosmic-ray correction, and light-curve extraction. The first five steps were performed following the recipes described in the STIS Handbook (Bostroem & Profitt 2011), while the rest were implemented in the same way as for the WFC3 staring-mode observations. For the extraction we used 6-pixel-wide apertures along the cross-dispersion direction and smoothed aperture edges along the dispersion direction (the smoothing factors were 5 Å for the G430L grating and 10 Å for the G750L grating, corresponding to approximately 2 pixels in each case).

2.2. HST Light-curve Modeling

Table 1 summarizes the parameters used for the light-curve modeling. For these parameters we consulted the Exoplanet Characterisation Catalogue developed as part of the ExoClock project (Kokori et al. 2021). In all the light-curve modeling steps we assumed a circular orbit and a fixed period for the planet, while the limb-darkening effect was modeled using the Clare

| Table 1 | Stellar and Planetary Parameters for WASP-17 b Used for the Data Reduction with Iraclis and TLCD-LSTM, and the Atmospheric Modeling with TauREx 3 |
|---------|---------------------------------------------------------------------------------|
| Stellar and Planetary Parameters | Value |
| $T_e$ (K) | 6550 ± 80$^a$ |
| $R_e$ (R$_e$) | 1.572 ± 0.056$^a$ |
| $M_e$ (M$_e$) | 1.306 ± 0.026$^a$ |
| log$_{10}(\delta)$ (cm s$^{-2}$) | 4.161 ± 0.026$^a$ |
| [Fe/H] | $-0.19 ± 0.09^a$ |
| $a/R_e$ | $0.727^{+0.24}_{-0.44}$ |
| $e$ | 0 (fixed) |
| $i$ (deg) | 87.96$^{+1.34}_{-1.56}$ |
| $\omega$ | 0 (fixed) |
| $P$ (days) | $3.7354845 ± 1.9 \times 10^{-6}b$ |
| $T_0$ (BJD$_{TDB}$) | 2,454,592.8015 ± 0.0005$^b$ |
| $R_p/R_e$ | 0.12229$^{+0.00017}_{-0.00013}$ |
| $M_p$ (M$_J$) | 0.477 ± 0.033$^b$ |
| $R_p$ (R$_J$) | 1.932 ± 0.053$^b$ |

Notes.

$^a$ Anderson et al. (2011).

$^b$ Southworth et al. (2012).
4-coefficient law (Claret 2000) and the ExoTETHyS (Morello et al. 2020) package, which takes into account the stellar parameters (Table 1) and the response curve of the instrument.

At a first stage, the initial orbit of each HST visit was discarded since it presents more significant wavelength-dependent systematics compared to the subsequent orbits. We then analyzed only the white light curves (broad band) and left the semimajor axis, the inclination, the mid-transit time, and the $R_p/R_\ast$ as free parameters, together with a number of parameters for the systematics implemented in Iraclis.

To ensure consistency between the analysis of the different data sets, we estimated again the inclination and the semimajor axis of the planet using all the available data. We combined the light curves extracted from the five HST observations, together with five TESS (Ricker et al. 2014) observations. The updated parameters are reported in Table 1.

Finally, we analyzed all the white and spectral HST light curves, fixing the values for the semimajor axis and the inclination. Figures 1, 2, 3, 4, and 5 show the fitted white and spectral light curves for all the HST observations, while Tables 4 and 5 in Appendix A report the transmission data for STIS and WFC3, respectively.

2.3. Spitzer Data Analysis

Spitzer data are available on the IRSA online archive. We downloaded the WASP-17 b transit observations taken with IRAC channels 1 and 2 in 2013 (P.I. Jean–Michel Desert, program 90092). The two observations were analyzed separately by employing the light-curve detrending method named TLCD-LSTM developed by Morvan et al. (2020). This approach uses a long short-term memory network (Hochreiter & Schmidhuber 1997) to predict the transit light curve without the need for any prior assumptions on the noise or transit shape. The only assumption regarding the systematics is that the associated correlated noise can be inferred from the out-of-transit light curve and the centroid time series. First, photometric light curves are extracted by using a circular aperture around the center of light. We tested a grid of radius apertures ranging from 2 to 4 pixels, by increments of 0.25 pixels. We then selected the aperture that maximized the S/N, i.e., 3.25 pixels. We subtracted the background light computed at each time step using the median flux in each image after excluding a disk of radius 15 pixels around the source center.

The neural network was then trained to learn the temporal variability on the pre-ingress and post-egress parts of the photometric light curve. While the stellar flux is masked during transit and then predicted autoregressively, the 2D Gaussian centroid time series is used as a covariate time series during the whole observation. This allows us to leverage its correlations with the photometric light curve and infer the systematics during the transit while excluding the planetary signal. However, intermediate transit fits are performed during training on the temporary detrended light curve in order to keep track of the model’s progress and provide an additional stopping criterion. A first run is done with large margins around the expected ingress and egress times to ensure that the transits do not overlap with the training ranges. The margins are then refined after this first run to include 105% of the transit duration centered on the mid-transit time. In practice, the results are very stable and unaffected by the chosen margins provided that these remain low (under $\approx 10\%$ of the transit duration). Several architecture and learning parameters are then tested. In both cases we find that two LSTM layers of 512 units and 10% dropout trained with an Adam optimizer and decay rate of $\beta = 0.95$ provide an optimal residual noise after the transit fit while avoiding overfitting. After the model was trained for 50 epochs, the best network—i.e., with the lowest residual noise—is saved along with the corresponding detrended light curve, obtained by dividing the raw light curve by the network’s prediction. A final transit fit is performed using a Markov

---

https://github.com/ucl-exoplanets/deepARTransit
Chain Monte Carlo procedure embedded in the PyLightcurve package (Tsiaras et al. 2016b). Both during and after training, the fixed planetary parameters used are those listed in Table 1 or computed using the ExoTETHyS open package (Morello et al. 2020), while the mid-transit time and $R_p/R_*$ were left free. The raw and detrended Spitzer light curves are shown in Figure 6 (top: IRAC channel 1; bottom: IRAC channel 2), and the corresponding transmission data is available in Table 6, Appendix A. The full list of parameters used to train the network is available in Table 7, Appendix B.

2.4. Atmospheric Modeling

The spectrum resulting from our data reduction process is shown in Figure 7. Each STIS data set is indicated with a different color depending on the date on which the observation was taken. WFC3 and Spitzer data are shown in blue and purple, respectively, independently of the grism or channel employed. From the aforementioned plot we can easily notice divergent results coming from the two STIS/G430L observations, namely, those taken on 2012 June 8 and 2013 March 15. For this reason we decided to model the observed spectrum in
three separate scenarios. Case 1 includes STIS visits 2012 June 8 (visit 1) and 2013 March 19 (visit 3), WFC3 and Spitzer data, while in Case 2 we aimed to model STIS visits 2013 March 15 (visit 2) and 2013 March 19 (visit 3), WFC3 and Spitzer data sets. Additionally, we investigated the spectrum constituted by the WFC3 observations only (Case 3), thus discarding the inconsistencies of the STIS observations.

In Figure 7 we compare our data with the analysis from Sing et al. (2016). We observe that our STIS visit 1, STIS visit 3, WFC3/G141, and Spitzer data agree with the previous analysis by Sing et al. (2016), finding, however, an inconsistency with the data point at 0.6 μm. Additionally, we notice that our STIS visit 2 data do not have a counterpart in the analysis by Sing et al. (2016), although their study claims to have analyzed all STIS observations available. We believe that their approach consisted in performing a weighted average of STIS visit 1 + 2, but we cannot be sure, as this information is not included in their methodology.

In each spectral case, we modeled the atmosphere at the terminator of WASP-17 b using the open-source atmospheric retrieval framework TauREx 3 (Al-Refaie et al. 2021). As is customary for hot Jupiters, we assumed this exoplanetary
and residuals we estimated the equilibrium temperature to be $T_p = 1769$ K when assuming $A = 0.1$ and $\epsilon = 0.8$. However, since we cannot ascertain the exact albedo and emissivity of the planet, we considered some extreme cases that produce a range where we expect the planetary temperature to fall. For instance, if we assume WASP-17b to possess an albedo of 0.5 and an emissivity of 0.7 (analog to Jupiter), the planetary equilibrium temperature suddenly drops to $T_p = 1579$ K. If, instead, we consider the case of a planet that reflects a tiny amount of incoming light ($A = 0.1$) and reemits some energy (possibly energy left over from the planet's formation process) with $\epsilon = 0.7$, then the equilibrium temperature increases to $T_p = 1829$ K.

As a result of the broad wavelength coverage of the combined observations and the high planetary equilibrium temperature (Sharp & Burrows 2007), we explored the presence of a variety of molecular opacities, ranging from the optical (TiO, McKemmish et al. 2019; VO, McKemmish et al. 2016; FeH, Dulick et al. 2003; Wende et al. 2010; SiH, Yurchenko et al. 2018; TiH, Gharib-Nezhad et al. 2021; AlO, Patrascu et al. 2015; K, Allard et al. 2019; Na, Allard et al. 2019) to the near-infrared and infrared absorbers (H$_2$O, Polyansky et al. 2018; CH$_4$, Yurchenko & Tennyson 2014; CO, Li et al. 2015; CO$_2$: NH$_3$, Coles et al. 2019), by employing the line lists from the ExoMol (Tennyson et al. 2016; Chubb et al. 2021), HITRAN (Rothman et al. 2010; Rothman & Gordon 2014), HITEMP (Rothman et al. 1987; Gordon et al. 2016), and NIST (Kramida 2013) databases.

In our final retrievals we considered only the presence of H$_2$O, AlO, TiH, SiH, CO, and CH$_4$ as trace gases. The molecular profile of each species was set to be constant at each atmospheric layer. Additionally, collision-induced absorption (CIA) from H$_2$–H$_2$ (Abel et al. 2011; Fletcher et al. 2018) and H$_2$–He (Abel et al. 2012), Rayleigh scattering for all molecules, and the presence of gray clouds and hazes modeled with the Lee et al. (2013) parameterization were included. We also modeled the contamination of the host star on the spectrum of WASP-17b using the formalization by Rackham et al. (2018, 2019).

We assumed a planetary atmosphere constituted by 100 layers in a plane-parallel geometry, uniformly distributed in log space between $10^{-5}$ and $10^6$ Pa. The temperature structure was modeled with an isothermal $T - p$ profile. The trace gases considered were allowed to vary freely between $10^{-12}$ and $10^{-1}$ in volume mixing ratio, while the planetary equilibrium temperature, $T_p$, could vary between 500 and 3000 K. Regarding the hazes, their top pressure $P_{\text{top}}$ ranged from $10^{-4}$ to $10^6$ Pa, the radius of the particles $R_{\text{Lee}}$ ranged between $10^{-3}$ and 1 $\mu$m, and the Mie cloud mixing ratio $\Lambda_{\text{Mie}}$ could change between $-30$ and $-4$ in log space. Furthermore, we included a layer of gray clouds, whose top pressure could vary between $10^{-1}$ and $10^6$ Pa (Robinson & Catling 2014; Kawashima & Ikoma 2018; Charnay et al. 2021). Lastly, the planetary radius bounds were set between 1 and 2 $R_J$. The prior bounds employed for each fitted parameter are reported in Table 2.

We investigated the parameter space with the nested sampling algorithm MultiNest (Feroz et al. 2009) with 750 live points and an evidence tolerance of 0.5.

2.5 Stellar Modeling

We attempted to explain the steep downward slope in the blue wavelengths of STIS visit 2 observation (see Figure 7) by modeling the effect of a combination of unocculted stellar spots and faculae on the Case 2 transmission spectrum of WASP-17b. We accounted for solar size spots and faculae to be homogeneously distributed on the stellar surface. Current stellar models and observational evidence (Mamajek & Hillenbrand 2008; Ciardi et al. 2011; Rackham et al. 2017, 2019) show that the variability of F dwarfs is $\sim 0.1\%$, with a minimum of 0.07% and a maximum of 0.36%. This is one of the lowest among all stellar spectral types, with the stellar variability being inversely proportional to the
Figure 7. Transmission spectrum of WASP-17 b, constituted by HST/STIS, HST/WFC3, and Spitzer/IRAC channel 1 and 2 data analyzed in this study compared to the data from Sing et al. (2016).

| Retrieved Parameters | Prior Bounds | Scale | Case 1 | Case 2 | Case 3 |
|----------------------|--------------|-------|--------|--------|--------|
|                      |              |       | STIS 1&3, WFC3, IRAC | STIS 2&3, WFC3, IRAC | WFC3 |
| $R_p$ ($R_J$)        | 1 : 2        | linear | 1.64 $\pm$ 0.03 | 1.59 $\pm$ 0.04 | 1.61 $\pm$ 0.04 |
| $T_{\text{spec}}$ (K) | 4000 : 5000  | linear | 4367.85 $\pm$ 296.58 | 4306.83 $\pm$ 267.22 | 4329.70 $\pm$ 278.98 |
| $T_{\text{Rad}}$ (K) | 6600 : 7000  | linear | 6741.68 $\pm$ 73.49 | 6919.28 $\pm$ 53.01 | 6914.43 $\pm$ 75.08 |
| $F_{\text{spot}}$    | 0.0 : 0.9    | linear | 0.13 $\pm$ 0.03 | 0.19 $\pm$ 0.03 | 0.20 $\pm$ 0.04 |
| $F_{\text{facc}}$    | 0.0 : 0.9    | linear | 0.51 $\pm$ 0.01 | 0.43 $\pm$ 0.01 | 0.50 $\pm$ 0.16 |
| $T$ (K)              | 500 : 3000   | linear | 1415.28 $\pm$ 194.43 | 1539.70 $\pm$ 317.21 | 1396.20 $\pm$ 231.79 |

| Retrieved Accounting for an Active Star | Retrieved Accounting for a Homogeneous Star |
|---------------------------------------|------------------------------------------|
| $R_p$ ($R_J$)        | 1 : 2         | linear | 1.72 $\pm$ 0.01 | 1.71 $\pm$ 0.01 | 1.71 $\pm$ 0.01 |
| $T$ (K)              | 500 : 3000   | linear | 951.13 $\pm$ 155.25 | 950.73 $\pm$ 93.36 | 900.56 $\pm$ 133.18 |
| $H_2$                 | –12 : –1     | log$_{10}$ | –3.22 $\pm$ 0.47 | –3.28 $\pm$ 0.61 | –2.99 $\pm$ 0.36 |
| $AlO$                | –12 : –1     | log$_{10}$ | –8.06 $\pm$ 0.03 | –8.14 $\pm$ 0.03 | –6.35 $\pm$ 0.03 |
| TiH                  | –12 : –1     | log$_{10}$ | –9.70 $\pm$ 1.14 | –9.71 $\pm$ 1.70 | –7.82 $\pm$ 1.59 |
| SiH                  | –12 : –1     | log$_{10}$ | –8.39 $\pm$ 2.21 | –7.03 $\pm$ 2.64 | –7.08 $\pm$ 3.05 |
| CH$_4$               | –12 : –1     | log$_{10}$ | –8.34 $\pm$ 2.24 | –7.48 $\pm$ 2.72 | –8.08 $\pm$ 3.24 |
| CO                   | –12 : –1     | log$_{10}$ | –7.09 $\pm$ 2.97 | –6.51 $\pm$ 3.15 | –7.55 $\pm$ 3.03 |
| $F_{\text{clouds}}$ (Pa) | –1 : 6      | log$_{10}$ | 4.11 $\pm$ 1.19 | 1.95 $\pm$ 0.36 | 3.98 $\pm$ 1.33 |
| $R_{\text{clouds}}$ (nm) | –3 : 1     | log$_{10}$ | –1.24 $\pm$ 1.19 | –1.25 $\pm$ 1.27 | –1.38 $\pm$ 1.43 |
| $P_{\text{Mic}}$ (Pa) | –4 : 6      | log$_{10}$ | 2.20 $\pm$ 1.57 | 2.11 $\pm$ 3.44 | 2.06 $\pm$ 3.55 |
| $\lambda_{\text{Lee}}$ | –30 : –4    | log$_{10}$ | –20.90 $\pm$ 8.77 | –20.13 $\pm$ 8.06 | –19.87 $\pm$ 8.27 |

List of the Parameters, Their Prior Bounds, the Scale Used, and the Retrieved Values (Median and Standard Deviation) for the Case 1–3 Spectra Assuming an Isothermal Atmospheric Profile.
temperature of the star (McQuillan et al. 2014). Though generally FGK stars are mostly inactive and do not contaminate the transmission spectrum of the planets orbiting them, their active counterparts do, and great care must be taken when analyzing the planetary spectra that they pollute.

According to Rackham et al. (2019), for FGK-type stars the stellar spot temperature can be calculated as

$$T_{\text{spot}} [K] = 0.418 \cdot T_{\text{phot}} + 1620$$  \hspace{1cm} (4)

and the faculae temperature as

$$T_{\text{fac}} [K] = T_{\text{phot}} + 100,$$  \hspace{1cm} (5)

with $T_{\text{phot}}$ being the temperature of the stellar photosphere, in our study equal to 6550 ± 80 K. The spectral emission density of the heterogeneous star is modeled using the BT-Settl models (Allard et al. 2013), and the stellar spectral model grid is generated using the PHOENIX library (Husser et al. 2013). In essence, the spots and faculae are modeled as separate cooler and hotter stars, respectively (A. Thompson et al. 2022, in preparation). The stellar emission densities (SEDs) of the three surface components (spots, faculae, and quiescent photosphere) are then combined in the ratio of their covering fractions to produce the observed disk-integrated stellar spectrum

$$S_{\lambda, \text{star}} = (1 - F_{\text{spot}} - F_{\text{fac}}) \cdot S_{\lambda, \text{phot}} + (F_{\text{spot}} \cdot S_{\lambda, \text{spot}}) + (F_{\text{fac}} \cdot S_{\lambda, \text{fac}}),$$  \hspace{1cm} (6)

where $S_{\lambda, \text{star}}$ is the flux of the heterogeneous star at a given wavelength; $S_{\lambda, \text{spot}}, S_{\lambda, \text{fac}},$ and $S_{\lambda, \text{phot}}$ are the spectra of the quiet photosphere, the starspots, and the faculae at the same wavelength, respectively; and $F_{\text{spot}}$ and $F_{\text{fac}}$ are the covering fractions of the spots and faculae with respect to the observed stellar disk.

Compared to the nominal transit depth, the effect of a heterogeneous star produces a wavelength-dependent offset on the transmission spectrum given by

$$\epsilon_{\lambda, t+f} = \frac{1}{1 - F_{\text{spot}} \left(1 - \frac{S_{\lambda, \text{spot}}}{S_{\lambda, \text{phot}}} \right) - F_{\text{fac}} \left(1 - \frac{S_{\lambda, \text{fac}}}{S_{\lambda, \text{phot}}} \right)},$$  \hspace{1cm} (7)

where $F_{\text{spot}}$ and $F_{\text{fac}}$ are the covering fractions of the spots and faculae and $S_{\lambda, \text{spot}}, S_{\lambda, \text{fac}},$ and $S_{\lambda, \text{phot}}$ are the emitted fluxes of the spots, faculae, and imaculate photosphere at a given wavelength, respectively. An important limitation to consider is that this simplistic method of modeling stellar active regions may be inaccurate, particularly with respect to faculae modeling, as it fails to describe any center-to-limb variations that they may exhibit or their strong dependence on magnetic field strength (Norris et al. 2017). However, given the precision of the observations in question, this simplified stellar model should be sufficient to begin to investigate any potential heterogeneity displayed by WASP-17.

In our retrievals, we considered a varying spot temperature between 4000 and 5000 K and a faculae temperature between 6600 and 7000 K. Both the spots and faculae coverage fraction bounds were set between 0.0 and 0.9 of the total stellar surface.

We tested the possibility of a variable stellar photosphere also on the spectrum in Cases 1 and 3.

### 3. Results

#### 3.1. Retrievals Accounting for an Active Star

When comparing the posteriors for each of the three cases in Figure 8, one can find common atmospheric features that are stable independently of the data sets modeled. First, the planetary temperatures retrieved are within 1σ of each other, and their upper bounds are inside the limits of the equilibrium temperature [1579 K; 1829 K] calculated in Section 2.4. Similarly, the planetary radii at 10 bars are compatible with one another.

All three spectra agree on a number of aspects: the presence of water in the solar/supersolar regime, which is consistent with the result from Pinhas et al. (2018), $\log_{10}(\text{H}_2\text{O}) = -4.04 +0.91_{-0.42}$. The reference solar water abundance was calculated under thermochemical equilibrium at a pressure of 1 bar for the equilibrium temperature of WASP-17 b, and it is equal to $\log_{10}(X_{\text{H}_2\text{O}}) = -3.3$ (Pinhas et al. 2018).

Our results suggest the presence of AlO and TiH on the order of $10^{-8}$ in volume mixing ratio, unconstrained SiH, and haze nondetection. Furthermore, the retrievals indicate a strong stellar contamination on the spectrum of WASP-17 b in the WFC3 data set too, where we do not expect stellar activity to play a big part. In all three scenarios we retrieved a spot temperature of around 4300 K, as well as spot and faculae covering fractions of the star on the order of 10%–20% and 40%–50%, respectively. Our results report a faculae temperature in the Case 1 transmission spectrum between 33 and 165 K higher than what Equation (5) predicts. On the other hand, $T_{\text{fac}}$ on Case 2 and Case 3 spectra is between 217 and 322 K higher than expected ($T_{\text{fac}} = 6919.28^{+53.01}_{-40.73}$ and $T_{\text{fac}} = 6914.43^{+55.02}_{-47.08}$). Although the study by Rackham et al. (2018) does not provide typical uncertainties on the temperature of spots and faculae, we can easily compare the retrieved values with the stellar photospheric uncertainty (~80 K) from Table 1. Hence, the $T_{\text{fac}}$ retrieved in Case 1, being less than 2σ away from the 80 K uncertainty, could be deemed acceptable compared to the faculae temperature in Cases 2 and 3, which is three to four times above the 80 K limit.

In Case 2 and Case 3 spectra we find supersolar water abundances ($\log_{10}(\text{H}_2\text{O}) = -1.93^{+0.79}_{-0.70}$ and $\log_{10}(\text{H}_2\text{O}) = -2.19^{+0.55}_{-0.69}$, respectively), while Case 1 returns a $\log_{10}(\text{H}_2\text{O}) = -3.20^{+0.99}_{-0.60}$ being 0.97× solar and 0.79 times the result reported by Barstow et al. (2016), equal to $5 \times 10^{-4}$.

Depending on the spectral scenario considered, the AlO, TiH, and SiH abundances retrieved differ (Case 1: $\log_{10}(\text{AlO}) = -7.94^{+0.87}_{-0.81}$, $\log_{10}(\text{TiH}) = -9.70^{+1.01}_{-1.14}$, $\log_{10}(\text{SiH}) = -8.39^{+2.25}_{-2.21}$), Case 2: $\log_{10}(\text{AlO}) = -9.68^{+1.52}_{-1.42}$, $\log_{10}(\text{TiH}) = -7.91^{+1.09}_{-1.70}$, $\log_{10}(\text{SiH}) = -7.03^{+2.84}_{-1.90}$, Case 3: $\log_{10}(\text{AlO}) = -6.39^{+0.81}_{-1.41}$, $\log_{10}(\text{TiH}) = -7.82^{+1.09}_{-1.90}$, $\log_{10}(\text{SiH}) = -7.08^{+3.02}_{-3.03}$, yet they are compatible with one another being within their 1σ confidence band. According to the thermochemical equilibrium model GGChem (Woitke et al. 2018), the expected abundance of AlO and TiH at $p = 1$ bar and $T = ~1400$ K is $\sim 10^{-12}$ and $\sim 10^{-12}$ in volume mixing ratio, respectively. The theoretical estimate is in accordance with the retrieved AlO abundance in Case 2, where we find a lower bound equal to $\log_{10}(\text{AlO}) = -11.1$. However, in all other cases for both AlO and TiH, the retrieved abundances are $3\sigma$–4σ away from the expected values. On the contrary, the retrieved uncertainties on the SiH abundances are in agreement with the theoretical
value computed at $p = 1$ bar and $T \sim 1400$ K, it being $\log_{10}(\text{SiH}) = \sim -10$.

We retrieve conflicting results for the cloud deck. The Case 1 and Case 3 posteriors returned a mean cloud deck at a pressure of $\sim 1.1 \times 10^4$ Pa corresponding to $\sim 0.11$ bars. In Case 2 the cloud top pressure is higher in the atmosphere, at a pressure of 0.9 mbar. The cloud coverage found at higher altitudes could be caused by the retrieval framework trying to fit the downward slope created by the STIS visit 2 data in the optical range, which cannot be explained by any molecular absorber. The clouds cut off any molecular signature in the optical, leaving just the tip of two particular TiH and SiH signatures visible at 0.54 and 0.42 $\mu$m in the Case 2 spectrum.

Figure 10 in Appendix C shows Case 1 data fitted by the best-fit model accounting for stellar activity and the contributions of active trace gases, CIA, clouds, and hazes to the fit.

We report the retrieval results for each spectral case in Table 2 (upper part).
3.2. Retrievals Accounting for a Homogeneous Star

We modeled the three spectral scenarios previously identified by accounting for the same molecular and atmospheric parameters. However, in this instance we modeled the star as a homogeneous body. We find that the absence of stellar activity does not impact the molecular detections, but it does affect a few of the retrieved abundances (Figure 9).

As in the previous case, our data suggest the presence of water, aluminum oxide, and titanium hydride. Our retrievals return solar $\log_{10}(\text{H}_2\text{O}) = -3.22^{+0.47}_{-0.49}$ and supersolar $\log_{10}(\text{H}_2\text{O}) = -2.41^{+0.48}_{-0.56}$ and $\log_{10}(\text{H}_2\text{O}) = -1.40^{+0.19}_{-0.28}$ water concentrations when modeling the Case 1, Case 3, and Case 2 spectrum, respectively. The last two results indicate a higher abundance than what current theories predict for hot Jupiters ($\log_{10}(\text{H}_2\text{O}) = \sim -3.3$; Venot et al. 2012; Venot & Agúndez 2015; Woitke et al. 2018).

The AlO and TiH abundances retrieved from the Case 3 spectrum appear to be more compatible with those retrieved from Case 1. However, both the AlO and the TiH results obtained from each of the three cases are within their respective 1σ uncertainties. These results are not consistent with chemical

![Figure 9. Posterior distributions for the transmission spectrum of WASP-17 b accounting for a homogeneous star retrieved with different data sets. The dashed vertical lines in each histogram refer to the median value, the first quantile (lower bound error), and the third quantile (upper bound error) of each parameter. Similarly, the reported values on top of the histograms denote the parameter’s median value, the first quantile, and the third quantile. The values shown are those obtained with the spectrum constituted by STIS visit 1&3, WFC3, Spitzer (Case 1). Inset: transmission spectra with best-fit models and their 1σ uncertainties.](image-url)
equilibrium models, as Woitek et al. (2018) predict a volume mixing ratio of roughly $10^{-16}$ for AIo and $10^{-15}$ for TiH at a pressure of 1 bar and $T = \sim 900$ K. The discrepancies could be caused by chemical disequilibrium processes, e.g., photochemistry (Venot et al. 2015; Fleury et al. 2019) or quenching (Shulyak et al. 2020), occurring in the atmosphere of WASP-17 b.

SIH is still largely undetected in Case 2 and Case 3 spectra. On the other hand, exceptionally high amounts of this trace gas are detected in Case 1. We recognize that such a large contribution ($\log_10$ (SIH) $= -3.85^{+1.18}_{-2.38}$) is unlikely, as theoretical models (Woitek et al. 2018) predict a SIH volume mixing ratio of $\sim 10^{-16}$ at 1 bar and $\sim 900$ K. Moreover, all three spectra return a consistent median planetary radius, temperature, and cloud coverage, approximately equal to 1.71 $R_p$, 930 K, and $4.6 \times 10^{-5}$ Pa, respectively.

In place of stellar faculae, a Mie scattering haze could have explained the upward slope in the Case 1 spectrum. Unfortunately, this option is refused by the retrieval, which is explained the upward slope in the Case 1 spectrum.

The retrieved values for each atmospheric parameter are reported in Table 2 (bottom part), while Figure 11 in Appendix C shows Case 2 data fitted by the best-fit model and the contributions of active trace gases, CIA, clouds, and hazes to the fit.

4. Discussion

WASP-17 b has been observed, studied, and modeled intensively in the past (e.g., Anderson et al. 2011; Mandell et al. 2013; Pinhas et al. 2018; Goyal et al. 2020) by employing observations taken with WFC3/G141 (staring mode), HST/STIS, and Spitzer. Our study, complementary to the work by Alderson et al. (2022), offers a consistent analysis performed by combining and reducing in a reproducible manner all currently available WASP-17 b transmission data obtained with space-based observatories. We include all public STIS and IRAC data, plus WFC3/G141 and WFC3/G102 observations, both taken in spatial scanning mode. This large breadth of observations allows us to investigate the planet’s atmospheric spectrum spanning the wavelengths from the optical to the near-infrared.

The combination of several data sets has its pros and cons. By coupling observations taken at different wavelengths, we are able to achieve a broad wavelength coverage, paramount to obtaining more precise measurements and unambiguous molecular detection. It is extremely challenging to absolutely calibrate data obtained from different detectors, especially if each of them targets a very specific spectral range with a limited or even absent wavelength overlap (Yip et al. 2020). Moreover, each instrument creates its own distinct systematics that need to be modeled ad hoc.

The STIS detector shares a short portion of the wavelength space, about 0.14 $\mu m$, with WFC3. This wavelength overlap is useful to investigate potential calibration discrepancies between instruments. In the 0.8–0.94 $\mu m$ range, our spectrum displays six WFC3/G102 data points and one STIS/G750L point (Figure 7). The response of the STIS G750L grating shows a strong dip above 0.8 $\mu m$ (Prichard et al. 2022); hence, the last spectral bin (0.8–0.94 $\mu m$) has a flux rate $\sim 50\%$ lower than the rest of the bands, and it is possibly subject to different systematics. Because of the shape of the transmission curve, we opted to increase the size of the final spectral bin to counteract the flux loss. Our analysis returns an STIS data point at 0.87 $\mu m$ that is at least 2$\sigma$ away from the G102 data points placed between 0.8 and 0.94 $\mu m$. A similar decrease in transit depth at the same wavelengths was also recovered by Sing et al. (2016), as can be seen in Figure 7. Yet the transit depths of the light curves for G750L and G102 are within 1$\sigma$ of each other (WFC3/G102: $R_p/R_\star = 0.1218^{+0.0019}_{-0.0022}$, STIS/G750L: $R_p/R_\star = 0.1211^{+0.0014}_{-0.0013}$), confirming that the two instrumental calibrations are in overall agreement. Finally, we must note that this problematic STIS/G750L data point is not statistically significant for the spectral modeling. From Figures 8 and 9 we can observe that the TauREx 3 atmospheric model ignores this data point when computing the best-fit spectrum.

On the other hand, between 3 and 5 $\mu m$ lie the Spitzer data points, whose transit depths cannot be compared to any additional data set. We shifted the two Spitzer data points 100 ppm up and 100 ppm down from their original position to see whether the results would be considerably impacted. In either case the retrieved mean molecular abundances and other atmospheric features, such as radius and temperature, fluctuate at the order of the fifth decimal figure from the initial values, implying an accurate data reduction process and stability of the results.

An additional issue to take into account when combining data sets taken months or even years apart is atmospheric variability. We must remember that planets are dynamic bodies, varying in space and time. Hence, mechanisms like star–planet interaction and global circulation processes such as polar vortexes and zonal jets can produce very strong temperature and chemical differences (Cho et al. 2003, 2021). It might not be surprising that the same planet can display contrasting spectral features at the same wavelengths, such as the two opposite spectral trends we presented between $\sim 0.4$ and $\sim 0.6$ $\mu m$. In light of this, we must stress that the publicly available STIS observations used in this study lack the post-egress part of the planetary transit (Figures 3, 4, 5). Missing light-curve data strongly impact the transit fit model, which then could potentially affect the derived transit depth, leading to discordant spectral features among different observations, as highlighted in Figure 7.

Due to the discrepancies in the data, we decided to model the contrasting STIS data sets separately: in Case 1 we considered a spectrum that includes STIS visit 1&3, WFC3, and IRAC; in Case 2 we modeled the spectrum with STIS visit 2&3, WFC3, and IRAC. Additionally, we employed the WFC3 observations alone (Case 3) to check how the results would change when we do not include the problematic STIS data sets.

To explain the unusually strong downward slope in the optical range displayed by the Case 2 spectrum (Figure 8, inset), an active star was included in the retrieval setup. For consistency, we applied the same stellar model to Case 1 and Case 3 spectra. Potentially, STIS visits 2&3, being taken just a few days apart, were impacted by a stellar event, causing the spectrum to be best fitted by a model containing high spot and faculae covering fractions. STIS visit 1 could have also been affected by a similar event, given that its best model favors stellar activity in place of aerosols.
The Astronomical Journal, 164:2 (19pp), 2022 July

Theoretically, a faculae covering fraction between 40% and 50% is unlikely in an F6-type star like WASP-17, unless we are dealing with a rare and extreme event. Moreover, we recognize that the posterior distributions between the faculae temperature and their covering fraction are degenerate (see Figure 8). This is because the retrieval is unable to distinguish between two scenarios: (1) a smaller covering fraction of hotter faculae and (2) a larger covering fraction of cooler faculae. Both scenarios result in similar stellar disk-integrated SEDs and thus introduce similar contamination features to the transmission spectrum.

There are no previous studies on the activity of the star, so we are unable to compare our results with published research. However, Khalafinejad et al. (2018) mention that their attempt to identify a sodium feature in the spectrum of WASP-17 b could have been impacted by an in-transit activity of the host star, which decreased the planetary Na signal. Hence, stellar contamination cannot be excluded as an additional cause for a decreasing transit depth in the optical regime. Continuous photometric monitoring of WASP-17 is necessary to determine accurately the activity levels of the star.

When the star is treated as a homogeneous body, the retrieval of the Case 1 spectrum results in a extremely high SiH volume mixing ratio $\log_{10}(\text{SiH}) = -3.85^{+1.18}_{-2.38}$. This result is unlikely, given that chemical equilibrium models predict a SiH abundance on the order of $10^{-16}$ at 900 K (Woitke et al. 2018). Perhaps such high SiH abundance is due to a missing molecule that is not accounted for. Either this missing gas absorbs at the same wavelengths of SiH, or it has similar absorption features. Even if we wanted to investigate this further, we do not possess all the possible molecular cross sections, and it is not feasible to test all the known molecules that absorb in the optical region. However, by inspecting the SiH results retrieved from the Case 2 and Case 3 spectra, we confirm that SiH is largely unconstrained and remains undetected.

Overall, we notice that when we consider an inactive star, the retrievals struggle to fit the data points at the end of the WFC3 wavelength coverage at $\sim1.6$–$1.8$ $\mu$m and those at $0.4$–$0.55$ $\mu$m, especially visit 2 (see Figure 9, inset). This statement is supported by the log evidence of the models: when stellar activity is included, the log(E) increases by a minimum of 4$\sigma$ (compared to the same model without stellar activity).

Furthermore, independently of the data sets modeled, our results suggest the presence of the same trace gases ($\text{H}_2\text{O}, \text{AlO}, \text{TiH}$) as when accounting for a heterogeneous star. At this stage we do not possess enough evidence to discard a particular STIS data set in favor of another one, but the information content derived by WFC3 data alone is enough to propose the existence of those species in the atmosphere of WASP-17 b.

During our preliminary retrievals, we tested different models containing a variety of trace gases on the spectrum in Case 1, but ultimately we found that including $\text{H}_2\text{O}$, AlO, TiH, SiH, clouds, hazes, and an active star would lead to the best results. Adding CH$_4$ and CO in the latter model did not substantially affect the Bayesian evidence, the two models being 0.26$\sigma$ away from each other.

On the other hand, the spectrum containing STIS visit 2 proved trickier to model. Out of the five models tested on these data, models (1), (3), (4) and (5) in Table 3 (middle section) were within less than 1$\sigma$ of each other, and model (2) containing $\text{H}_2\text{O}$, AlO, TiH, SiH, CH$_4$, CO, clouds, and hazes was 5.64$\sigma$ away from model (1), which includes the same chemical species plus stellar variability. Ultimately, we used the best-fit model for the Case 1 spectral scenario also in the Case 2 data set and on the WFC3 data alone, to make a valid comparison between the scenarios in terms of molecular abundances. The retrievals conducted exclusively on the WFC3 data also favor a model containing $\text{H}_2\text{O}$, AlO, TiH, SiH, clouds, hazes, stellar spots, and faculae.

Our results suggest that WASP-17 b possesses mostly a clear atmosphere, free from a thick layer of clouds and aerosols. Sing et al. (2016) report the atmosphere of this hot Jupiter to be the clearest among the 10 exoplanets in their study, presenting distinct water features. Pinhas et al. (2018) find similar results, their spectrum being fitted best by a model containing mostly a cloudless atmosphere with a cloud/haze fraction of 0.2 and a cloud top pressure of 0.1 mbar. On the other hand, Barstow et al. (2016) report the presence of scattering aerosols at relatively high altitudes, but not so strong as to reduce the size of the other molecular features. Mandell et al. (2013) also claim that their data are best fitted by a model that accounts for hazes or clouds.

We were unable to identify the alkali lines (Na, K) reported by Sing et al. (2016). Their data point at $\sim$0.6 $\mu$m has an uncertainty three times larger than the other ones, and it is not statistically significant to demonstrate the presence of sodium in the atmosphere of WASP-17 b.

Furthermore, during our exploratory retrievals we did not find evidence of CO$_2$, which, on the contrary, is detected by Alderson et al. (2022). Due to the sparse nature of the Spitzer data points and the influence that additional molecules can impart in the spectrum of an exoplanet at the wavelengths probed by IRAC, it is challenging to place precise constraints on the abundance of carbon-bearing molecules.

Regardless of whether we consider an active star or not, the retrieved planetary temperature is always lower than the calculated equilibrium temperature, as transmission studies often find. Discrepancies between the retrieved and the calculated temperatures have been analyzed by several authors (e.g., Caldas et al. 2019; MacDonald et al. 2020; Pluriel et al. 2020b; Changeat et al. 2021) and found to arise from a combination of aspects: the adoption of 1D atmospheric models to describe the different molecular content of the morning and evening terminators, the use of simplified models to characterize the planets’ 3D processes, and the adoption of incorrect assumptions regarding the albedo and the emissivity when calculating the planet’s equilibrium temperature. Hence, retrieving a lower temperature than calculated should not be surprising.
Table 3
Bayesian Log-difference between Different Models Depending on the Observations Fed to the Retrievals

| Setup | Log Evidence | Sigma (Discarded by) | Retrieved Temperature (K) |
|-------|--------------|----------------------|---------------------------|
| Case 1: STIS Visits 1 and 3, WFC3, Spitzer | 395.73 | 1415.28 | 2230.00 |
| (1) H₂O, AIO, TiH, SiH, CH₄, CO, clouds, hazes, stellar activity | | | |
| (2) H₂O, AIO, TiH, SiH, CH₄, CO, clouds, hazes | 389.31 | 4.01 w.r.t. (1) | 951.13 | 159.25 |
| (3) H₂O, clouds, stellar activity | 395.02 | 1.84 w.r.t. (1) | 1429.50 | 248.74 |
| (4) H₂O, AIO, TiH, SiH, clouds, hazes, stellar activity | 395.90 | <1 w.r.t. (1) | 1397.57 | 382.33 |
| (5) H₂O, TiO, VO, FeH, clouds, stellar activity | 393.60 | 2.6 w.r.t. (1) | 1854.96 | 495.05 |
| Case 2: STIS Visits 2 and 3, WFC3, Spitzer | 401.41 | 1539.70 | 211.29 |
| (1) H₂O, AIO, TiH, SiH, CH₄, CO, clouds, hazes, stellar activity | | | |
| (2) H₂O, AIO, TiH, SiH, CH₄, CO, clouds, hazes | 387.38 | 5.64 w.r.t. (1) | 950.73 | 93.36 |
| (3) H₂O, clouds, stellar activity | 401.68 | <1 w.r.t. (1) | 1357.63 | 324.71 |
| (4) H₂O, AIO, TiH, SiH, clouds, hazes, stellar activity | 401.14 | <1 w.r.t. (1) | 1435.55 | 256.29 |
| (5) H₂O, TiO, VO, FeH, clouds, stellar activity | 401.42 | <1 w.r.t. (1) | 1550.71 | 306.26 |
| Case 3: WFC3 | 285.37 | 1396.20 | 251.34 |
| (1) H₂O, AIO, TiH, SiH, CH₄, CO, clouds, hazes, stellar activity | | | |
| (2) H₂O, AIO, TiH, SiH, CH₄, CO, clouds, hazes | 278.77 | 4.05 w.r.t. (1) | 900.56 | 153.18 |
| (3) H₂O, clouds, stellar activity | 284.59 | 1.88 w.r.t. (1) | 1374.81 | 328.09 |
| (4) H₂O, AIO, TiH, SiH, clouds, hazes, stellar activity | 285.44 | <1 w.r.t. (1) | 1356.01 | 256.69 |
| (5) H₂O, TiO, VO, FeH, clouds, stellar activity | 284.15 | 2.15 w.r.t. (1) | 1475.35 | 309.41 |

5. Conclusion

Remote-sensing studies are the only technique through which we can unveil the atmospheric characteristics of planets outside our solar system. In recent years HST has pioneered the spectroscopic investigation of exoplanets. In this study we presented the transmission spectrum of WASP-17 b observed with the STIS and the WFC3 mounted on HST and the Infrared Camera Array aboard the Spitzer Space Telescope. The Hubble spectroscopic data were reduced with Iraclis, a specialized STIS and WFC3 data reduction routine. Photometric data obtained by Spitzer were instead reduced and analyzed with the TLCD-LSTM method by means of a long short-term memory network. Given the discordant results obtained for two STIS observations that cover the same wavelength range, we decided to model the spectrum of WASP-17 b in three separate cases: (a) including STIS visit 1&3, WFC3, IRAC (Case 1); (b) including STIS visit 2&3, WFC3, IRAC (Case 2); (c) WFC3 only (Case 3). The fully Bayesian retrieval framework TauREx 3 was then employed to find the model that can best explain each spectral scenario. Retrieval results find that the best-fit model is constituted by H₂O, AIO, TiH, SiH, clouds, hazes, and stellar activity in Cases 1 and 3. The best-fit model for Case 2 includes CH₄ and CO too; however, their addition only improved the Bayesian evidence by 0.26σ. We acknowledge that the extreme stellar activity identified on the F6-type host star is unlikely; therefore, we check whether a model containing the same trace gases but including a homogeneous star is able to describe all three spectra. Independently of the activity of the star, our analysis indicates the presence of water in the solar/supersolar regime and possible traces of AIO and TiH in the atmosphere of WASP-17 b. The atmosphere of the inflated hot Jupiter appears to be free from both an optically thick layer of gray clouds and hazes. Given the incompleteness of the STIS light curves, we reckon that additional HST/STIS observations will help to improve our knowledge about the limb of this gas giant. In fact, the remarkably large scale height of WASP-17 b makes the planet a perfect target for further atmospheric characterization, both from the ground and by current and next-generation space-based observatories such as the James Webb Space Telescope (Greene et al. 2016) and Ariel (Tinetti et al. 2018).

We thank our anonymous referee for the insightful comments that have improved the quality of our work.

Data: This work is based on observations with the NASA/ESA Hubble Space Telescope, obtained at the Space Telescope Science Institute (STScI), operated by AURA, Inc. The publicly available HST observations presented here were taken as part of proposals 12473 and 14918, led by David Sing and Hannah Wakeford, respectively. These were obtained from the Hubble Archive, which is part of the Mikulski Archive for Space Telescopes. This work is also based in part on observations made with the Spitzer Space Telescope, which is operated by the Jet Propulsion Laboratory, California Institute of Technology, under a contract with NASA. The publicly available Spitzer data were taken as part of program 90902 led by Jean–Michel Desert. This paper also uses data collected by the TESS mission, which are publicly available from the Mikulski Archive for Space Telescopes (MAST) and produced by the Science Processing Operations Center (SPOC) at NASA Ames Research Center (Jenkins et al. 2016). Funding for the TESS mission is provided by NASA’s Science Mission directorate. We are thankful to those who operate this archive, the public nature of which increases scientific productivity and accessibility (Peek et al. 2019).

Funding: This project has received funding from the European Research Council under the European Unions Horizon 2020 research and innovation program (grant agreement No. 758892, ExoAI) and the European Unions Horizon 2020 COMPET program (grant agreement No. 776403, Exoplanets A). Furthermore, we acknowledge funding by the UK Space Agency and Science and Technology Funding Council grants ST/K502406/1, ST/P000282/1, ST/
P002153/1, ST/ S002634/1, ST/T001836/1, ST/V003380/1, and ST/W00254X/1. B.E. is a Laureate of the Paris Region fellowship program, which is supported by the Ile-de-France Region and has received funding under the Horizon 2020 innovation framework program and the Marie Sklodowska-Curie grant agreement No. 945298.

Software: Iraclis (Tsiaras et al. 2016a), TauREx 3 (Al-Refaie et al. 2021), PyLightcurve (Tsiaras et al. 2016b), Lightkurve (Lightkurve Collaboration et al. 2018), ExoTETHyS (Morello et al. 2020), Astropy (Price-Whelan et al. 2018), h5py (Collette 2013), emcee (Foreman-Mackey et al. 2013), Matplotlib (Hunter 2007), Multinest (Feroz et al. 2009), PyMultinest (Buchner et al. 2014), Pandas (McKinney & Pandas Development Team 2011), Numpy (Oliphant 2006), SciPy (Virtanen et al. 2020), corner (Foreman-Mackey 2016).

Appendix A
Spectral Data

The transit depth and associated errors for the STIS, WFC3 and IRAC data sets are shown in Tables 4, 5, and 6, respectively.

Table 4
Reduced and Fitted Spectral Data from the Raw HST/STIS Transmission Data Using Iraclis

| Wavelength (μm) | Transit Depth (%) | Error (%) | Bandwidth (μm) | Instrument | Grating |
|----------------|------------------|-----------|----------------|------------|---------|
| 0.385          | 1.593            | 0.045     | 0.050          | HST STIS   | G430L   |
| 0.420          | 1.529            | 0.030     | 0.020          | HST STIS   | G430L   |
| 0.440          | 1.525            | 0.037     | 0.020          | HST STIS   | G430L   |
| 0.460          | 1.537            | 0.028     | 0.020          | HST STIS   | G430L   |
| 0.480          | 1.484            | 0.027     | 0.020          | HST STIS   | G430L   |
| 0.500          | 1.473            | 0.029     | 0.020          | HST STIS   | G430L   |
| 0.520          | 1.525            | 0.027     | 0.020          | HST STIS   | G430L   |
| 0.540          | 1.534            | 0.036     | 0.020          | HST STIS   | G430L   |
| 0.560          | 1.472            | 0.041     | 0.020          | HST STIS   | G430L   |
| 0.385          | 1.391            | 0.032     | 0.050          | HST STIS   | G430L   |
| 0.420          | 1.444            | 0.046     | 0.020          | HST STIS   | G430L   |
| 0.440          | 1.417            | 0.028     | 0.020          | HST STIS   | G430L   |
| 0.460          | 1.407            | 0.026     | 0.020          | HST STIS   | G430L   |
| 0.480          | 1.400            | 0.027     | 0.020          | HST STIS   | G430L   |
| 0.500          | 1.431            | 0.026     | 0.020          | HST STIS   | G430L   |
| 0.520          | 1.444            | 0.027     | 0.020          | HST STIS   | G430L   |
| 0.540          | 1.466            | 0.033     | 0.020          | HST STIS   | G430L   |
| 0.560          | 1.486            | 0.037     | 0.020          | HST STIS   | G430L   |
| 0.570          | 1.479            | 0.021     | 0.040          | HST STIS   | G750L   |
| 0.605          | 1.456            | 0.036     | 0.030          | HST STIS   | G750L   |
| 0.635          | 1.464            | 0.034     | 0.030          | HST STIS   | G750L   |
| 0.665          | 1.440            | 0.027     | 0.030          | HST STIS   | G750L   |
| 0.700          | 1.488            | 0.022     | 0.040          | HST STIS   | G750L   |
| 0.760          | 1.497            | 0.033     | 0.080          | HST STIS   | G750L   |
| 0.870          | 1.394            | 0.031     | 0.140          | HST STIS   | G750L   |
Table 5
Reduced and Fitted Spectral Data from the Raw HST/WFC3 Transmission Data Using Iraclis

| Wavelength (μm) | Transit Depth (%) | Error (%) | Bandwidth (μm) | Instrument | Grism |
|-----------------|-------------------|-----------|----------------|------------|-------|
| 0.813           | 1.506             | 0.019     | 0.025          | HST WFC3   | G102  |
| 0.838           | 1.492             | 0.012     | 0.025          | HST WFC3   | G102  |
| 0.863           | 1.483             | 0.009     | 0.025          | HST WFC3   | G102  |
| 0.888           | 1.479             | 0.011     | 0.025          | HST WFC3   | G102  |
| 0.913           | 1.459             | 0.007     | 0.025          | HST WFC3   | G102  |
| 0.938           | 1.504             | 0.009     | 0.025          | HST WFC3   | G102  |
| 0.963           | 1.502             | 0.010     | 0.025          | HST WFC3   | G102  |
| 0.988           | 1.497             | 0.012     | 0.025          | HST WFC3   | G102  |
| 1.013           | 1.484             | 0.009     | 0.025          | HST WFC3   | G102  |
| 1.038           | 1.474             | 0.009     | 0.025          | HST WFC3   | G102  |
| 1.063           | 1.474             | 0.008     | 0.025          | HST WFC3   | G102  |
| 1.088           | 1.480             | 0.010     | 0.025          | HST WFC3   | G102  |
| 1.113           | 1.493             | 0.009     | 0.025          | HST WFC3   | G102  |
| 1.138           | 1.471             | 0.020     | 0.025          | HST WFC3   | G102  |
| 1.126           | 1.519             | 0.011     | 0.022          | HST WFC3   | G141  |
| 1.148           | 1.507             | 0.012     | 0.021          | HST WFC3   | G141  |
| 1.169           | 1.491             | 0.011     | 0.021          | HST WFC3   | G141  |
| 1.189           | 1.488             | 0.009     | 0.020          | HST WFC3   | G141  |
| 1.208           | 1.489             | 0.011     | 0.019          | HST WFC3   | G141  |
| 1.228           | 1.500             | 0.008     | 0.019          | HST WFC3   | G141  |
| 1.246           | 1.476             | 0.009     | 0.019          | HST WFC3   | G141  |
| 1.266           | 1.469             | 0.011     | 0.019          | HST WFC3   | G141  |
| 1.285           | 1.457             | 0.013     | 0.019          | HST WFC3   | G141  |
| 1.304           | 1.503             | 0.010     | 0.019          | HST WFC3   | G141  |
| 1.323           | 1.497             | 0.009     | 0.019          | HST WFC3   | G141  |
| 1.341           | 1.514             | 0.010     | 0.019          | HST WFC3   | G141  |
| 1.361           | 1.537             | 0.014     | 0.019          | HST WFC3   | G141  |
| 1.380           | 1.522             | 0.012     | 0.020          | HST WFC3   | G141  |
| 1.400           | 1.516             | 0.011     | 0.020          | HST WFC3   | G141  |
| 1.420           | 1.517             | 0.011     | 0.020          | HST WFC3   | G141  |
| 1.441           | 1.518             | 0.011     | 0.021          | HST WFC3   | G141  |
| 1.462           | 1.527             | 0.012     | 0.021          | HST WFC3   | G141  |
| 1.483           | 1.526             | 0.009     | 0.022          | HST WFC3   | G141  |
| 1.505           | 1.493             | 0.011     | 0.022          | HST WFC3   | G141  |
| 1.528           | 1.500             | 0.009     | 0.023          | HST WFC3   | G141  |
| 1.552           | 1.488             | 0.011     | 0.024          | HST WFC3   | G141  |
| 1.576           | 1.463             | 0.010     | 0.025          | HST WFC3   | G141  |
| 1.602           | 1.459             | 0.011     | 0.026          | HST WFC3   | G141  |
| 1.629           | 1.436             | 0.010     | 0.028          | HST WFC3   | G141  |

Table 6
Reduced and Fitted Spectral Data from the Raw Spitzer/IRAC Transmission Data Using TLCD-LSTM

| Wavelength (μm) | Transit Depth (%) | Error (%) | Bandwidth (μm) | Instrument |
|-----------------|-------------------|-----------|----------------|------------|
| 3.560           | 1.500             | 0.032     | 0.380          | Spitzer IRAC |
| 4.500           | 1.521             | 0.018     | 0.560          | Spitzer IRAC |

Appendix B
Network Parameters

Table 7 shows the LSTM parameters used for detrending IRAC observations.
|                          | IRAC Channel 1 | IRAC Channel 2 |
|--------------------------|----------------|----------------|
| Program                  | 90092          | 90092          |
| Aorkey                   | 47040000       | 47039488       |
| Radius (pixels)          | 3.25           | 3.25           |
| Centroiding method       | 2D Gaussian    | 2D Gaussian    |
| Start of test set (BMJD TDB) | 56,422.595     | 56,426.330     |
| End of test set (BMJD TDB) | 56,422.783     | 56,426.518     |
| No. layers               | 2              | 2              |
| No. units                | 512            | 512            |
| Dropout rate             | 10%            | 10%            |
| Adam learning rate       | 0.005          | 0.005          |
| $\beta$ decay rate       | 0.95           | 0.95           |
Appendix C
Retrieval Contribution Functions

Case 1 (Figure 10) and Case 2 (Figure 11) spectra fitted by the best-fit model when accounting for an active and a homogenous star, respectively. The plots show the best-fit model (in black) and the contribution of CIA, active trace gases, clouds, and hazes to the final model.

Figure 10. STIS visit 1&3, WFC3, and Spitzer data (Case 1) fitted by the best-fit model (black line) that includes an active star together with the contributions from CIA, active trace gases, clouds, and hazes. Stellar activity induces wavelength variations at the base of the model, which is why the gray cloud and Mie cloud contributions are not flat at the bottom.

Figure 11. STIS visit 2&3, WFC3, and Spitzer data (Case 2) fitted by the best-fit model (black line) that includes a homogeneous star together with the contributions from CIA, active trace gases, clouds, and hazes.
Venot, O., Hébrard, E., Agúndez, M., et al. 2012, *A&A*, 546, A43
Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, *Nature methods*, 17, 261
Wende, S., Reiners, A., Seifahrt, A., & Bernath, P. 2010, *A&A*, 523, A58
Winn, J. N., Johnson, J. A., Fabrycky, D., et al. 2009, *ApJ*, 700, 302
Woitke, P., Helling, C., Hunter, G., et al. 2018, *A&A*, 614, A1

Wood, P. L., Maxted, P. F. L., Smalley, B., & Iro, N. 2011, *MNRAS*, 412, 2376
Yip, K. H., Changeat, Q., Edwards, B., et al. 2020, *AJ*, 161, 4
Yurchenko, S. N., Sinden, F., Lodi, L., et al. 2018, *MNRAS*, 473, 5324
Yurchenko, S. N., & Tennyson, J. 2014, *MNRAS*, 440, 1649
Zhou, G., & Bayliss, D. D. R. 2012, *MNRAS*, 426, 2483