A high-contrast search for variability in HR 8799bc with VLT-SPHERE

Biller, B.A.; Apai, D.; Bonnefoy, M.; Desidera, S.; Gratton, R.; Kasper, M.; … ; Zurlo, A.

A High-Contrast Search for Variability in HR 8799bc with VLT-SPHERE

B.A. Biller,[1,2]* D. Apai,[3] M. Bonnefoy,[4] S. Desidera,[5] R. Gratton,[5] M. Kasper,[6]

M. Kenworthy,[7] A.M. Lagrange,[4] C. Lazzoni,[5,8] D. Mesa,[5] A. Vigan,[9]

K. Wagner,[3,10] J.M. Vos,[11] A. Zurlo[9,12]

[1] *SUPA, Institute for Astronomy, University of Edinburgh, Blackford Hill, Edinburgh EH9 3HJ, UK*
[2] *Centre for Exoplanet Science, University of Edinburgh, Edinburgh, UK*
[3] *Steward Observatory, The University of Arizona, Tucson, AZ 85721, USA*
[4] *Univ. Grenoble Alpes, CNRS, IPAG, F-38000 Grenoble, France*
[5] *INAF - Osservatorio Astronomico di Padova, Vicolo della Osservatorio 5, 35122, Padova, Italy*
[6] *European Southern Observatory (ESO), Karl-Schwarzschild-Str. 2,85748 Garching, Germany*
[7] *Leiden Observatory, Leiden University, PO Box 9513, 2300 RA, Leiden, The Netherlands*
[8] *Dipartimento di Fisica a Astronomia "G. Galilei", Universita' di Padova, Via Marzolo, 8, 35121 Padova, Italy*
[9] *Aix Marseille Univ, CNRS, CNES, LAM, Marseille, France*
[10] *NASA Hubble/Sagan Fellow*
[11] *Department of Astrophysics, American Museum of Natural History, Central Park West at 79th Street, New York, NY 10024, USA*
[12] *Escuela de Ingeniería Industrial, Facultad de Ciencias y Ciencias, Universidad Diego Portales, Av. Ejercito 441, Santiago, Chile*

22 January 2021

ABSTRACT

The planets HR8799bc display nearly identical colours and spectra as variable young exoplanet analogues such as VHS 1256-1257ABb and PSO J318.5-22, and are likely to be similarly variable. Here we present results from a 5-epoch SPHERE IRDIS broadband-*H* search for variability in these two planets. HR 8799b aperture photometry and HR 8799bc negative simulated planet photometry share similar trends within uncertainties. Satellite spot lightcurves share the same trends as the planet lightcurves in the August 2018 epochs, but diverge in the October 2017 epochs. We consider $\Delta(mag)_b - \Delta(mag)_c$ to trace non-shared variations between the two planets, and rule out non-shared variability in $\Delta(mag)_b - \Delta(mag)_c$ to the 10-20% level over 4-5 hours. To quantify our sensitivity to variability, we simulate variable lightcurves by inserting and retrieving a suite of simulated planets at similar radii from the star as HR 8799bc, but offset in position angle. For HR 8799b, for periods <10 hours, we are sensitive to variability with amplitude > 5%. For HR 8799c, our sensitivity is limited to variability > 25% for similar periods.

Key words: planets and satellites: atmospheres – Planetary Systems, planets and satellites: gaseous planets – Planetary Systems, infrared: planetary systems – Resolved and unresolved sources as a function of wavelength, (stars:) brown dwarfs – Stars

1 INTRODUCTION

Over the last 15 years, a small but growing cohort of young, widely-separated giant planets, such as HR 8799bcde (Marois et al. 2008, 2010) and β Pic b (Lagrange et al. 2010), have been directly imaged via their own thermal emission in the infrared (see e.g. Bowler 2016). Instruments such as SPHERE at the VLT (Beuzit et al. 2019), GPI at Gemini (Macintosh et al. 2014), SCExAO-Charis at Subaru (Jovanovic et al. 2015; Currie et al. 2019), and Gravity at the VLT (Gravity Collaboration et al. 2019) have enabled spectroscopy and deep characterization of the atmospheres of these planets (Konopacky et al. 2013; Snellen et al. 2014; Barman et al. 2015; Bonnefoy et al. 2014, 2016; Zurlo et al. 2016; De Rosa et al. 2016; Ingraham et al. 2014; Rajan et al. 2015; Lavie et al. 2017; Wang et al. 2018b; Greenbaum et al. 2018; Ruffio et al. 2019; Gravity Collaboration et al. 2019). However, most of these studies have captured only a snapshot of the planet's atmospheric state, taken over a 1-3 hour long observation, or

have combined multiple observations taken at different points in the planet's rotation, creating essentially an "integrated" spectrum. Only one directly imaged exoplanet has a measured rotation period, β Pic b, which is a fast rotator, with a 7-9 hour period (Snellen et al. 2014), assuming a similar inclination as the disc. Top-of-atmosphere asymmetric structure (e.g. from spots, patchy clouds and various other mechanisms) could cause significant changes in flux as a function of time and wavelength, as different structures rotate in and out of view. These structures may also evolve dynamically over multiple rotation periods, producing longer-term changes in planet brightness. To determine if either of these phenomena are present requires time-resolved observations of exoplanets, preferably covering multiple rotation periods. Detecting and characterizing such features, especially as a function of wavelength, will provide strong tests for atmospheric models of these planets.

Studies of field brown dwarfs, the more massive cousins of these young planets, suggest that we would expect to find significant changes in the flux of young planets with time. Quasiperiodic variability is commonly found in field L and T type brown

dwarfs, with periods from 1.5-30 hours, and amplitudes ranging from 0.01-27% (Radigan et al. 2012, 2014; Radigan 2014; Wilson et al. 2014; Metchev et al. 2015; Eriksson et al. 2019). These objects are fast rotators (3-20 hours, see e.g. Zapatero Osorio et al. 2006) with silicate clouds expected in their atmospheres (e.g. horizontal banded clouds recently detected in Luhman 16A via linear polarization by Millar-Blanchaer et al. 2020). These potentially inhomogeneous cloud features likely drive the observed variability, modulated by the fast rotation of these objects. Spectroscopic variability monitoring has shown, for instance, that the variability in high amplitude L/T transition brown dwarfs such as 2MASS J21392676+0220226 (henceforth 2M2139) and SIMP J013656.5+093347 (henceforth SIMP 0136) is best reproduced by inhomogeneous coverage of both thin and thicker cloud components (Apai et al. 2013).

Directly imaged young, giant exoplanets share similar effective temperatures and compositions as these L and T dwarfs, and likely possess comparable or even higher amplitude variability. One key physical difference between brown dwarfs and young giant exoplanets is the much lower surface gravity of the young planets compared to more massive, older brown dwarfs. Surface gravity strongly affects the placement of clouds in these atmospheres, with ramifications for the variability properties of younger, lower mass objects (Marley et al. 2012). The variability amplitudes of young, free-floating planetary mass objects appear to be significantly boosted relative to older, higher mass brown dwarf counterparts with similar spectral types. Vos et al. (2019) find that 30% of low surface-gravity L dwarfs are variable, compared to 11% of field L dwarfs. The near-IR variability detections in the L6-L7 exoplanet analogues VHS J125601.92-125723.9 (henceforth VHS 1256-1257ABb, Gauza et al. 2015; Bowler et al. 2020; Zhou et al. 2020), PSO J318.5338-22.8603 (henceforth PSO J318.5-22 ,∼ 8 M_{Jup} β Pic moving group member Liu et al. 2013; Allers et al. 2016), WISEPJ004701.06+680352.1 (henceforth W0047, AB Dor moving group, Gizis et al. 2015), and 2MASS J2244316+204343 (henceforth 2M2244, AB Dor moving group, Vos et al. 2018) have peak-to-trough amplitudes >5% (Biller et al. 2015, 2018; Vos et al. 2018; Lew et al. 2016). These are the **highest-amplitude** variability detections for **any** L spectral type object, with VHS 1256-1257ABb potentially the most variable planetary mass or substellar object yet observed (∼25% variability observed over an 8-hour HST observation, Bowler et al. 2020). These objects have spectra that are nearly identical to those of the outer two HR 8799 planets (Zurlo et al. 2016; Bonnefoy et al. 2016) and HIP 65426b (Chauvin et al. 2017). Given their close spectral match, the HR 8799 planets are potentially equally intrinsically variable (although the observable variability amplitude is likely smaller, as the HR 8799 planets are probably viewed close to pole-on, Ruffio et al. 2019; Wang et al. 2018a).

The extreme contrast difference between star and planet has hindered studies of variability in exoplanet companions such as HR 8799bcde. Planet imagers such as SPHERE at the VLT (Beuzit et al. 2019) and GPI at Gemini (Macintosh et al. 2014) now enable short cadence observations of bright giant planets – previous generations of imagers required ≥1 hour for a high S/N ratio image of such a planet. In contrast, HR 8799bcd are visible even in <1 min SPHERE *H*-band single exposures. Apai et al. (2016) conducted a pilot variability study (<2 hours on sky) of the HR 8799 planets as part of SPHERE science verification and conclusively show that variability monitoring is possible using high-contrast coronagraphic imaging on SPHERE. Here we present the first in-depth, long (>4 hour) search for variability in the HR 8799 planets.

2 PROSPECTS FOR VARIABILITY MONITORING OF THE HR 8799 SYSTEM

The star HR 8799 is a F0V member of the 42^{+6}_{-4} Myr Columba moving group (Torres et al. 2008; Zuckerman et al. 2011; Bell et al. 2015), with an estimated mass of ∼1.5 solar masses. HR 8799 is a γ Doradus pulsating variable. As in all γ Doradus stars, many pulsation modes are excited; the mode with highest amplitude has a period of ∼0.5 day (Sódor et al. 2014), unfortunately close to the potential rotation period of the planets. The optical lightcurve from the Microvariability and Oscillations in STars (MOST) space telescope from Figure 1 of Sódor et al. (2014) displays clear peaks at the ∼0.5 day period; these peaks, however, have variable heights due to constructive and destructive interference from other weaker pulsation modes. At its most variable, HR 8799 displays a peak-to-valley amplitude of ∼0.1 mag in the visible (Sódor et al. 2014). While we expect a smaller amplitude at longer wavelengths (and indeed, observations of other γ Dor variables with TESS suggest this is the case, Antoci et al. 2019), the red optical or near-IR variability of HR 8799 has not been directly measured. Thus, the fact that SPHERE enables monitoring of the multiple HR 8799 planets simultaneously as well as the ability to produce artificial "satellite spots" as photometric references (Langlois et al. 2013) crucially allow us to disentangle any intrinsic variability of the star from variability of the planets. As we do not expect the planets to have the same rotational period, phase or variability amplitude, any variability with the same period and phase shared by multiple planets would then probably either be intrinsic to the star or caused by changing observing conditions. Similarly, the satellite spot light curves should follow the trend produced by the changing conditions as well as any astrophysical variability from the star, providing a reference lightcurve that can be used to detrend single planet lightcurves. Detrending using a lightcurve from a photometric reference has been successfully implemented in numerous ground-based variability monitoring campaigns for isolated objects (Radigan et al. 2014; Vos et al. 2019), building a "calibration curve" from other stars in the field to remove variations due to changing conditions; here we will attempt a similar detrending using both the satellite spot lightcurves and the other planets as the photometric references.

The HR 8799 system also hosts a 3-component debris disk (a warm inner belt separated by the orbits of the planets from a cold planetesimal belt and an outer halo of small grains) imaged at mid-IR, submillimeter, and mm wavelengths (Su et al. 2009; Matthews et al. 2014; Booth et al. 2016; Wilner et al. 2018). The disk is viewed nearly face-on (Matthews et al. 2014). Current orbital fits suggest that the planet orbits are close to coplanar with the disk, with a small inclination of ∼27° (Ruffio et al. 2019; Wang et al. 2018a; Konopacky et al. 2016; Pueyo et al. 2015; Maire et al. 2015). If the 4 planets indeed share a similar inclination with the disk, they are likely viewed close to pole-on. This is a non-ideal viewing angle for variability detection. The brown dwarfs with the highest observed variability amplitudes are generally those which are observed at or close to equator-on (90° inclination, Vos et al. 2017), as inhomogeneous features on objects observed at lower inclinations will appear smaller due to projection effects. Thus, given the probable pole-on viewing angle of the HR 8799 planets, any variability amplitude we potentially measure is likely only a fraction of the intrinsic variability amplitude.

With the current generation of near-IR extreme adaptive optics (AO) planet-finding cameras, high-contrast ground-based variability monitoring is feasible at wavelengths of 1-4 μm. After correcting for inclination effects, field brown dwarfs (Radigan et al. 2014; Metchev et al. 2015) display a monotonic decrease in variability amplitude

across this wavelength range, with the highest amplitude variability in *J* band (~1.1-1.3 μm), decreasing somewhat through the rest of the near-IR, and with considerably lower mid-IR (3-5 μm) amplitudes; thus, SPHERE's near-IR capabilities are ideal for a first search for variability in the HR 8799 planets, which share similar atmospheric properties to these objects. In our initial July 2015 observation, we chose to observe in the SPHERE *J*23 filter, as close as possible to the expected *J*-band peak of variability for similar objects. However, as HR 8799bcde are intrinsically fainter at *J* than in *H* or *K*, we found that the planets were not detected with sufficient S/N in the July 2015 data to enable variability studies, hence for the rest of the epochs, the broadband *H* filter was chosen as the best compromise.

3 OBSERVATIONS AND DATA REDUCTION

From July 2015 to August 2018, we obtained 7 epochs of SPHERE-IRDIS variability monitoring for the HR 8799 planets (programs 095.C-0689(A), 099.C-0588(A), and 101.C-0315(A)). A summary of the observations is provided in Table 1. SPHERE-IRDIS is a dual-band imaging (DBI) camera, enabling coronagraphic observations in two wavelength bands simultaneously (Dohlen et al. 2008; Vigan et al. 2010; Beuzit et al. 2019). In July 2015, we collected one epoch of simultaneous imaging in the SPHERE dual-band *J*23 filters (*J*2: λ_c=1189.5 nm, $\Delta\lambda$=47.6 nm, *J*3: λ_c=1269.8 nm, $\Delta\lambda$=50.8 nm) and a second epoch of simultaneous imaging in the *K*12 filters (*K*1: λ_c=2102.5 nm, $\Delta\lambda$=102 nm, *K*2: λ_c=2255 nm, $\Delta\lambda$=109 nm). All other epochs were obtained using the SPHERE broadband *H* filter (λ_c=1625.5 nm, $\Delta\lambda$=291 nm) in both of IRDIS' wavelength channels[1]. HR 8799 is relatively far north for the VLT, thus we were only able to observe for 4-6 hours per epoch with airmass <2.5. Observations were taken in pupil-tracking mode, allowing the field to rotate with parallactic angle on the sky, in order to build an accurate point spread function (PSF) for quasistatic speckle suppression (see e.g. Marois et al. 2006; Lafrenière et al. 2007; Soummer et al. 2012; Amara & Quanz 2012). We used the *N_ALC_YJH_S* coronagraphic configuration (inner working angle of 0.0925") of SPHERE's apodized pupil Lyot coronagraph (Carbillet et al. 2011; Guerri et al. 2011) in all epochs except for the *K*12 observations taken on 31 July 2015, which utilized the *N_ALC_Ks* coronagraphic configuration instead (inner working angle of 0.12"). The 2015 epochs were taken in designated visitor modes for full half nights, enabling longer overall observations (albeit sometimes at higher airmasses). All other epochs were taken in service mode. Conditions on 20 August 2018 were excellent, leading the service observer to adjust the base exposure time (DIT) on the fly to reduce saturation of HR 8799. Unfortunately, the number of exposures (NEXP) taken was not also adjusted to compensate for the decreased DIT, leading to a significant loss in time coverage at this epoch. On some nights, initial exposures were taken with different DIT values or satellite spot illumination; these frames have been removed from the subsequent analysis.

To ensure a simultaneous photometric reference at all epochs, observations were conducted using the *star center* template that provides satellite spots. Adding an additional periodic modulation to the deformable mirror (above and beyond that used to correct the incoming wavefront) produces 4 satellite spots in a cross pattern on the detector (Langlois et al. 2013). These satellite spots appear relatively circular when using a narrow-band filter and elongated when

using a wide-band filter. They share the spectra of the star and can potentially be used as both photometric and astrometric references (Langlois et al. 2013; Wang et al. 2014). In the July 2015 observations, off-axis point spread function (PSF) images were also interspersed every 20-30 minutes throughout the observation to use as a photometric reference. We found that this compromised the stability of the observations, so in later epochs we only obtained off-axis PSF images at the start and finish of each 4-5 hour observing sequence.

Conditions varied considerably between epochs. Obtained *H*-band Strehl ratio, seeing, and coherence time τ_0 over each observation are plotted in Fig. 1 and wind speed and precipitable water vapor are plotted in Fig. 2. The *H*-band Strehl ratio and wind speeds are estimates from the SPARTA AO real-time computer (Beuzit et al. 2019); seeing (from the observatory Differential Image Motion Monitor measurements), coherence time τ_0 and precipitable water vapor are drawn from the raw data file headers. Most epochs had quite good conditions and AO performance, with seeing <1" and Strehl ratios between 0.8 and 0.9. The *K* band epoch taken on 31 July 2015, however, was obtained in considerably worse conditions (seeing>1") and with notably lower Strehl ratios from 0.5 to 0.7. The *J* band epoch taken on 2015-07-30 was obtained in reasonably good conditions, but, unlike other epochs, HR 8799b is not apparent in single frame images as the planet is fainter in *J* than in *H* or *K* band. Thus, the 2015 epochs either do not have sufficient photometric stability or sufficiently high signal-to-noise detection per planet to enable variability monitoring and are excluded from further variability analysis.

When the wind speed at 10 m or 30 m drops below 3 m/s, SPHERE observations can be affected by the "Low Wind Effect" (LWE) (Sauvage et al. 2016; Milli et al. 2018), causing a highly distorted PSF shape. For most of our observations, measured wind speed remained above 3 m/s, dropping below this speed for most of the observation on 2018-08-18 and a portion of the observation on 2017-10-13. We visually inspected PSF images for these epochs obtained via the SPARTA AO data handling system to determine if these observations were affected by LWE and found that this effect was mild or non-existent in our observations.

Data were reduced and aligned by the SPHERE Data Center using the standard SPHERE pipeline. For details on the SPHERE Data Center and the pipeline, please see Delorme et al. (2017) and Pavlov et al. (2008). We adopted standard values for pixel scale of 12.25 mas / pixel and true north orientation of -1.75° in these reductions (Maire et al. 2016). Fully reduced images for the full observing sequences for the October 2017 and August 2018 data are presented in Fig. 3; these were processed using the Vortex Image Processing Principle Component Analysis (VIP-PCA) pipeline (Gomez Gonzalez et al. 2017). This PCA analysis utilizes the library of frames composing the full observing sequence to build a series of principal component images (henceforth PCA modes), with the majority of the speckle pattern well-modeled with the first few components. Subtracting out an image built from these PCA modes will then remove the speckle pattern, with an increasing number of modes further improving the modeling of the speckle pattern, at the expense of subtracting out some planet light as well. For the full observing sequence, we found that removing 25 principal components provided the best balance between fully suppressing the speckle pattern without causing significant planet self-subtraction.

4 SATELLITE SPOT TIME-SERIES PHOTOMETRY

Our observations were acquired with satellite spots present in all frames. The standard SPHERE pipeline uses images taken with satel-

---

[1] We report results from IRDIS channel 1 here, as the data obtained from channel 2 is functionally identical to the channel 1 data.
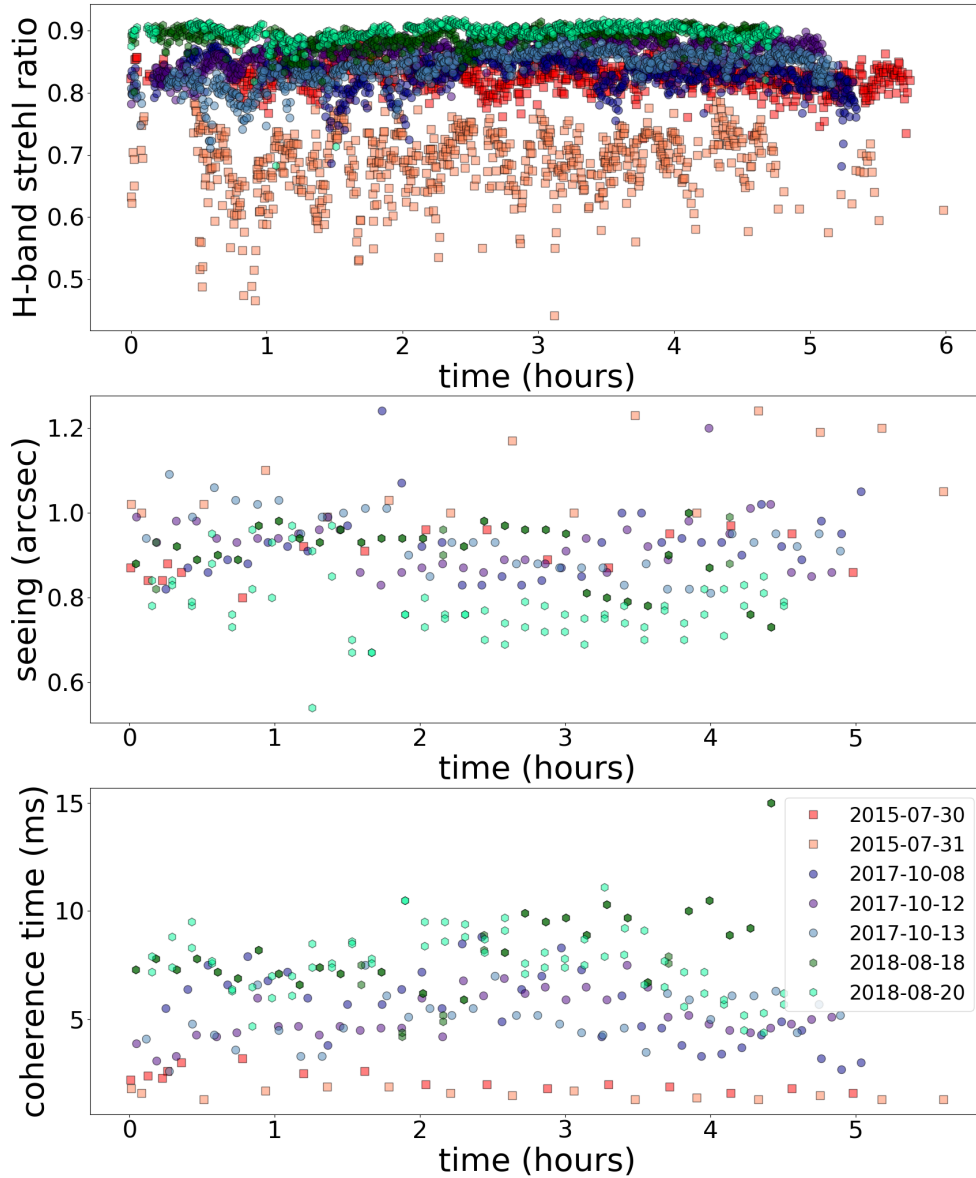
Figure 1. *H*-band Strehl ratio, seeing, and coherence time as a function of time during each of our 7 variability monitoring observations. Ranking observations in order of observing conditions: the 2015 observations were taken in the worst conditions among those probed by our variability search and are excluded from our variability analysis. The October 2017 observations were taken in very good conditions and the August 2018 observations were taken in excellent conditions.

lite spots before and after the main observing sequence (in most cases, without satellite spots) as astrometric references to measure the stellar position behind the coronagraph and finely centre images, while photometric calibration is performed relative to unsaturated images of the star taken before and after the main observing sequence. Since variations in seeing, Strehl, and airmass over an observation will cause brightness fluctuations in the stellar PSF, the planets, and the

speckle pattern, establishing a photometric reference free of variability from the planets is vital for detrending and interpreting our exoplanet observations. As the star is behind the coronagraph during our observations, it is unfeasible to obtain unsaturated images of the star to use as a photometric reference. Hence, here we establish whether or not the satellite spots serve as appropriate photomet-
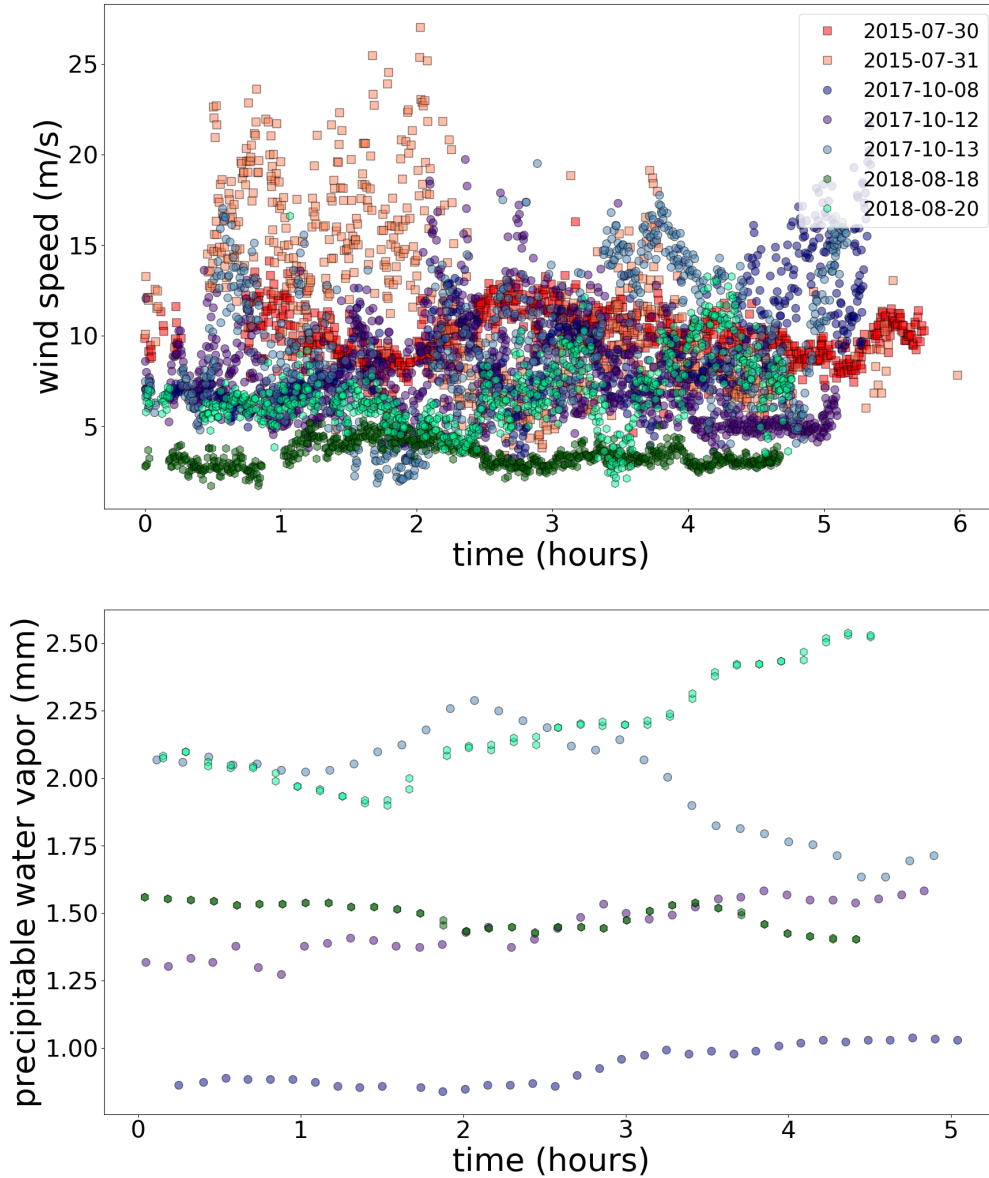
Figure 2. Windspeed and precipitable water vapor as a function of time during each of our 7 variability monitoring observations. Precipitable water vapor measurements were not available for the 2015 epochs, and thus are shown here only for the October 2017 and August 2018 observations. The windspeed measurement presented here is estimated from the SPARTA AO real-time computer. In the 2015-07-31 epoch, this windspeed value is likely an overestimate of the actual windspeed, as telescopes on Paranal cease operations at windspeeds above 18 m/s.

ric references and a proxy for measurements of unsaturated stellar images.

While satellite spots appear relatively circular when using a narrow-band filter, with the broadband H filter they appear as roughly elliptical spots. Thus, to determine satellite spot position in each frame in our time series data, we fit a 2-dimensional elliptical Gaus-

sian to each of the four satellite spots using the astropy photutils modeling package (Bradley et al. 2019). We then extracted aperture photometry using the astropy photutils aperphot package in an elliptical aperture with semi-major and semi-minor axes of 4 and 2 pixels, rotated to the correct angle to recover the flux from each satellite spot, and centered on the satellite spot positions from the 2-d
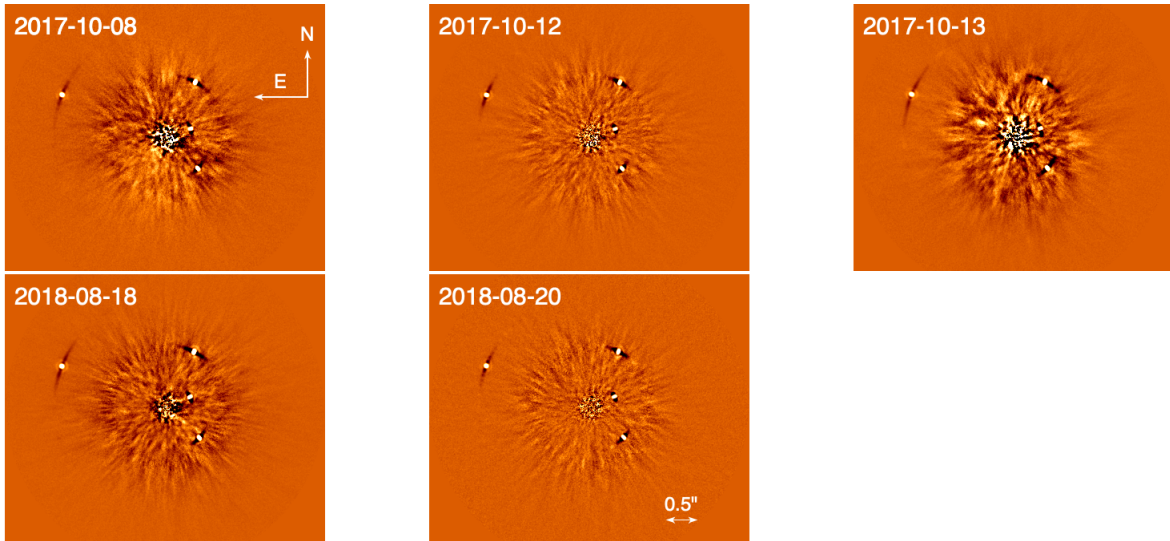
Figure 3. Full sequence reduced images for data taken in October 2017 and August 2018. Data were processed using the VIP PCA pipeline (Gomez Gonzalez et al. 2017), removing 25 principal component modes. All four planets are easily detected; the image scale runs from -10 to 10 ADU and has been adjusted to highlight the residual speckle noise remaining after subtraction of the principal modes.

elliptical Gaussian fits. We also extracted aperture photometry from "background" elliptical apertures placed 6 pixels away on either side of each satellite spot. The background as a function of time was then calculated as the average of these two background elliptical apertures. The geometry of the extraction apertures and the numbering convention for satellite spots is shown in Fig. 4. In Fig. 5, we plot photometry (normalized by the median value across the whole observation for each satellite spot) as a function of time for each satellite spot for the epoch of 2018-08-18 (taken in excellent conditions), for the various cases of: 1) the amplitude of the 2d-gaussian fit (shown as crosses), 2) elliptical aperture photometry (shown as filled circles), and 3) elliptical aperture photometry, with the background subtracted from the background elliptical apertures (shown as filled squares). We plot as well the median of the normalized satellite spot curves for each case. Photometry for the other epochs of observations can be found in Appendix A. In all cases, we find that the 2-d Gaussian fit and the elliptical aperture photometry yield similar results, as expected. We found amplitudes for the 2-d Gaussian fits of 200-300 counts for the 2017-10-08 epoch and 2000-3000 counts for the other epochs. Likely owing to better observing conditions, the satellite spot photometry for the August 2018 epochs is notably more stable and less variable than that for the October 2017 epochs.

In most cases, subtracting the background photometry did not significantly change the shape of the satellite spot light curves. Notably, however, for the 2017-10-08 epoch, where the satellite spots were under-illuminated, the satellite spot photometry seemed to follow that of the background regions and was significantly flattened and "detrended" via the background subtraction. Also, for satellite spot 4 in the 2017-10-13 epoch, background subtraction appears to have corrected a linear trend seen in this satellite spot but not in the other 3 satellite spots during this epoch. We extracted photometry from a set of 5 circular "background regions" for comparison as well, with one region centered on the core of the star PSF, and the other four spaced evenly around the image. Backgrounds extracted from each of these regions were similar as a function of time and resembled the background photometry obtained in the elliptical background regions selected close to the satellite spots. In Appendix B, we overplot the median flux in the satellite spots, measured background flux,

obtained Strehl ratio and seeing, to search for correlations in the satellite spot lightcurves as a function of ambient conditions. Other than the 2017-10-08 epoch (when, as noted previously, the satellite spots were under-illuminated), the satellite spot light curves do not appear to be strongly correlated to the Strehl ratio or anti-correlated with the background level.

4.1 Correlations between satellite spots

Following Apai et al. (2016), we calculate the degree of correlation between individual satellite spots using the Spearman's correlation test. The Spearman's correlation test ranks two sets of data in order of ascending data values, and then assesses how correlated the ranked sets of data are – this is the Spearman ρ value. If the datasets are perfectly correlated (a higher data value in one dataset always corresponds to a higher value in the other), this will produce a ρ value of 1. Perfectly anti-correlated datasets will produce a ρ value of -1. The Spearman's correlation test requires that the relationship between two variables follows a monotonic function, unlike a Pearson correlation test which specifically requires a linear relationship. We used spearmanr from from scipy.stats in python to calculate the Spearman ρ values between different normalized satellite spot lightcurves. The Spearman ρ values between different satellite spots are shown in Fig. 6.

We find that the satellite spots are mildly correlated in the October 2017 data, but at a low level and with a wide spread, with correlations ranging from 0.23 to 0.90, but generally around 0.5. This is significantly lower than the typical correlations of 0.7-0.9 found by Apai et al. (2016). In our August 2018 data, we find much smaller correlations between satellite spots, even though these data were taken in considerably better conditions, with correlations ranging from 0.03 to 0.71, with values of around 0.3-0.4 for most of the pairs of satellite spots. These low Spearman ρ values likely stem from the very small spread in satellite spot photometric values for these data, due to the excellent conditions, causing the Spearman's correlation test to be an imperfect proxy for observation quality in this case. This can be seen more clearly if we examine the correlation plots between satellite spots directly – an example of which is given in Fig. 7, and the rest
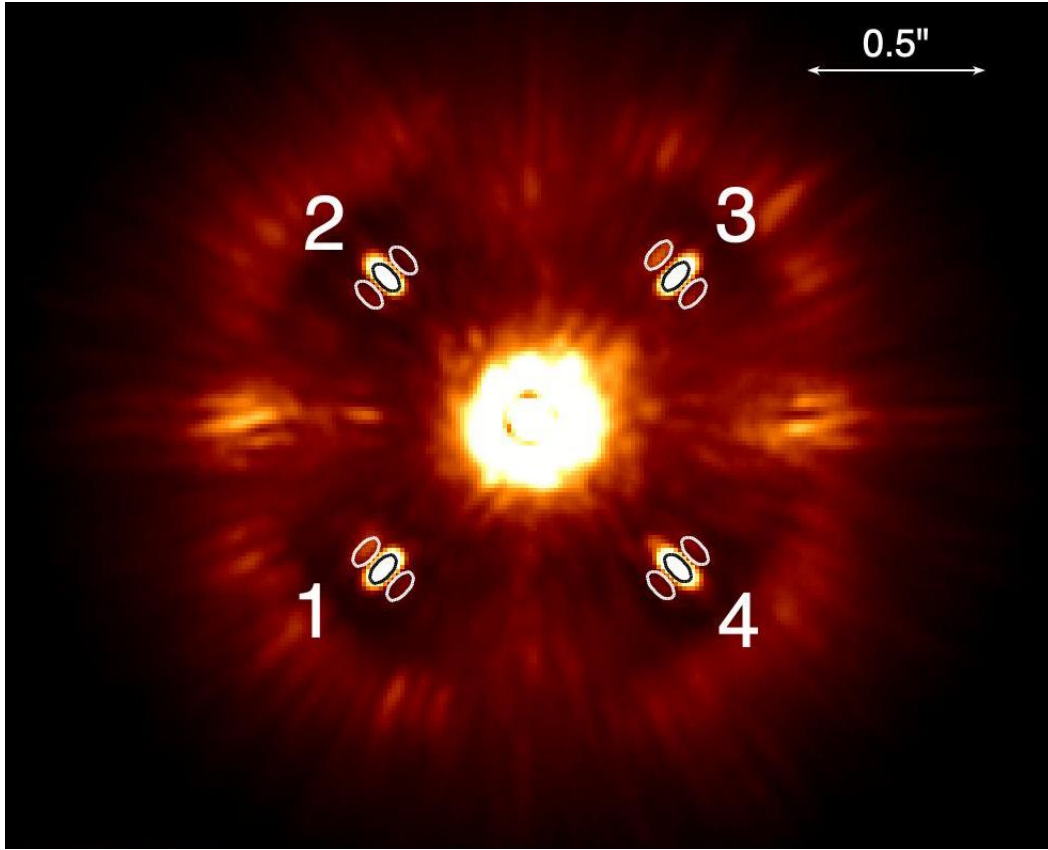
Figure 4. Satellite spots aperture photometry extraction regions and naming convention, overplotted on a single frame of data from 2018-08-18.

of which appear in Appendix C. Due to larger changes in observing conditions and poorer Strehl values overall, values for satellite spot photometry varied more over the October 2017 observations compared to the August 2018, leading to higher values of the Spearman ρ value (in other words, the greater range of values sampled allows a seemingly larger correlation between values).

Following Apai et al. (2016), we also calculate Spearman ρ values after dividing through by the mean satellite spot lightcurve. This removes shared variations between satellite spots, thus highlighting the difference in behavior between the individual satellite spots. The Spearman ρ values between different satellite spots after division by the mean are shown in Fig. 8, with the corresponding correlation plots shown in Appendix C. Apai et al. (2016) found a fairly constant pattern of no or weak anti-correlation between most of the satellite spots, with relatively strong anti-correlations between one set of directly opposite spots, and two sets of adjacent spots. We find more night-by-night diversity in our results, which are discussed in detail in Appendix C – however, given the small range of flux values covered in our observations, we do not consider these Spearman ρ values to be robust.

## 5 EXOPLANET TIME SERIES PHOTOMETRY

5.1 Achievable time resolution

To minimize self-subtraction and ensure sufficient speckle suppression, coverage of a sufficient range in parallactic angle is necessary within each temporal bin within the full observation to move each planet off of itself over the course of that segment of the observation.

We fit a 2-dimensional Gaussian to the PSF images taken before and after each temporal sequence. On all 5 nights, the FWHM of the PSF was measured to be roughly 4 pixels. Thus, ideally, we would cover at least 20 pixels of on-sky rotation ($\geq$5 PSF FWHMs) in each temporal bin to minimize planet self-subtraction. This is a quite conservative choice; the parallactic angle range we adopt for each temporal bin is essentially that of the shortest useful angular differential imaging (ADI see e.g. Marois et al. 2006; Lafrenière et al. 2007) observation for that planet-star separation. We found that speckle evolution in our datasets was sufficiently rapid that frames taken too far before or after a given frame were not useful to include in building the PSF using PCA, thus the frames selected for each temporal bin were run through the PCA analysis separately. The parallactic angle range we adopt for each temporal bin ensures that for each frame in the temporal bin, a sufficient number of frames taken far enough away in parallactic angle are available to build an accurate PSF model without significant self-subtraction. Each of our observations covered >70° of sky rotation for HR 8799, over 4-5 hours; the parallactic angle coverage necessary for 20 pixels of on-sky rotation will vary according to the planet-star separation.

For HR 8799b, we adopt 10° of rotation on sky per temporal bin as this minimum coverage; at a planet-star separation of ~1.7", 10° of sky rotation corresponds to a shift of 24 pixels on sky, sufficient to shift the planet >5 PSF-widths away from its initial position. Thus, we binned the observing sequences into 6 equal parallactic angle bins, ensuring a minimum variation in parallactic angle of 10° in each bin. For HR 8799c at a planet-star separation of ~1.0", 10° of sky rotation corresponds to a shift of 14 pixels on sky (~3.5 PSF-widths), which may not be sufficient to completely prevent self-subtraction. Thus,
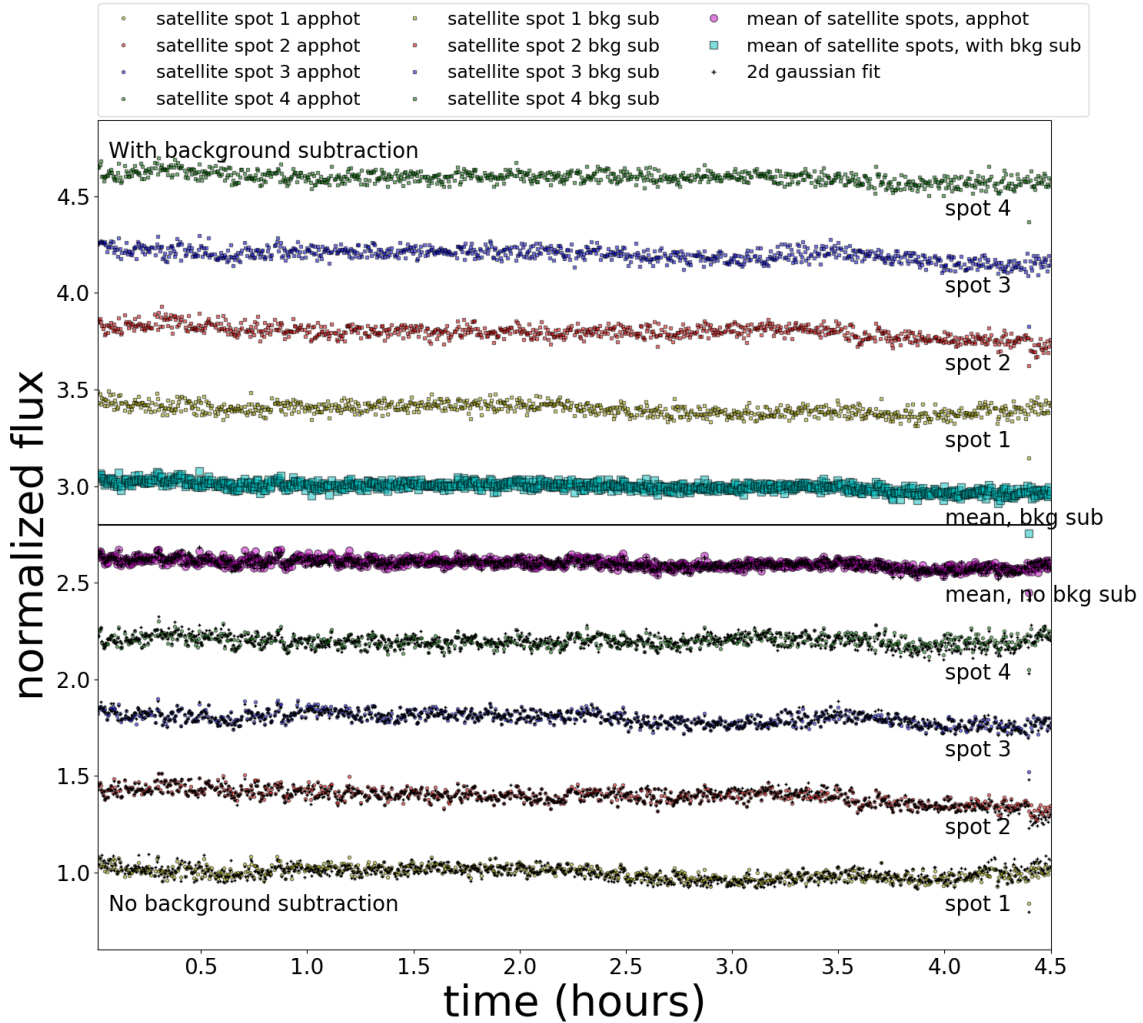
Figure 5. Satellite spot photometry for 2018-18-18 (normalized by the median value across the whole observation for each satellite spot) as a function of time for each satellite spot, with an offset in flux included between each satellite spot lightcurve for clarity. Results with background subtraction are shown on the top half of the plot, while results without background subtraction are plotted on the bottom half of the plot. The amplitudes from a 2-d gaussian fit to each satellite spot are plotted as black crosses, elliptical aperture photometry is plotted as filled circles, and elliptical aperture photometry with an additional background subtraction is plotted as filled squares. The mean of the elliptical aperture photometry for all 4 satellite spots are plotted in the middle of the figure, with magenta circles for the elliptical aperture photometry without background subtraction, and cyan square for elliptical aperture photometry with background subtraction.

for HR 8799c, we adopt 14° of rotation per sky per temporal bin, corresponding to 5 equal parallactic angle bins and a shift of ~20 pixels (~5 PSF-widths) on sky at ~1.0".

For the inner two planets, HR 8799d and HR8799e, at separations of ~0.6" and ~0.4", 10° of sky rotation corresponds to shifts of ~9 pixels and ~6 pixels on sky, insufficient to avoid significant self-subtraction and preventing an appropriate cadence for variability monitoring. In Section 5.5, we consider whether useful variability measurements can be obtained for the inner two planets if we adopt 20° of rotation on sky per temporal bin as the minimum coverage, and hence only consider a total of 3 equal parallactic angle bins per observation.

5.2 HR 8799b direct aperture photometry time-series

At ~1.7" from the star, as the widest-separated of the 4 planets, HR 8799b lies outside of the AO-corrected region, where uncorrected

atmospheric halo dominates over the static instrumental speckles. Unlike the closer-in planets, which are swimming in a sea of speckles and require a full suite of post-processing techniques for accurate photometry, deriving photometry for HR 8799b is amenable to simpler approaches. Therefore, we extracted aperture photometry directly for HR 8799b. The observing sequences were binned in 6 equal parallactic angle bins, and then we simply derotated and stacked data within each parallactic angle bin. We first fit a 2-d gaussian to the planet position in each binned data slice using photutils.modeling to determine the best planet center position and then extracted photometry in a 3-pixel radius circular aperture centered on the 2-d gaussian fit position. Since we have not removed the stellar PSF, the background varies considerably as a function of radius at the position of the planet, hence we chose to use a quite small aperture here – the 3-pixel aperture is sufficiently small to minimize variation in the background at the planet position. To correct for this radially-dependent background, we subtracted both the median of an annulus
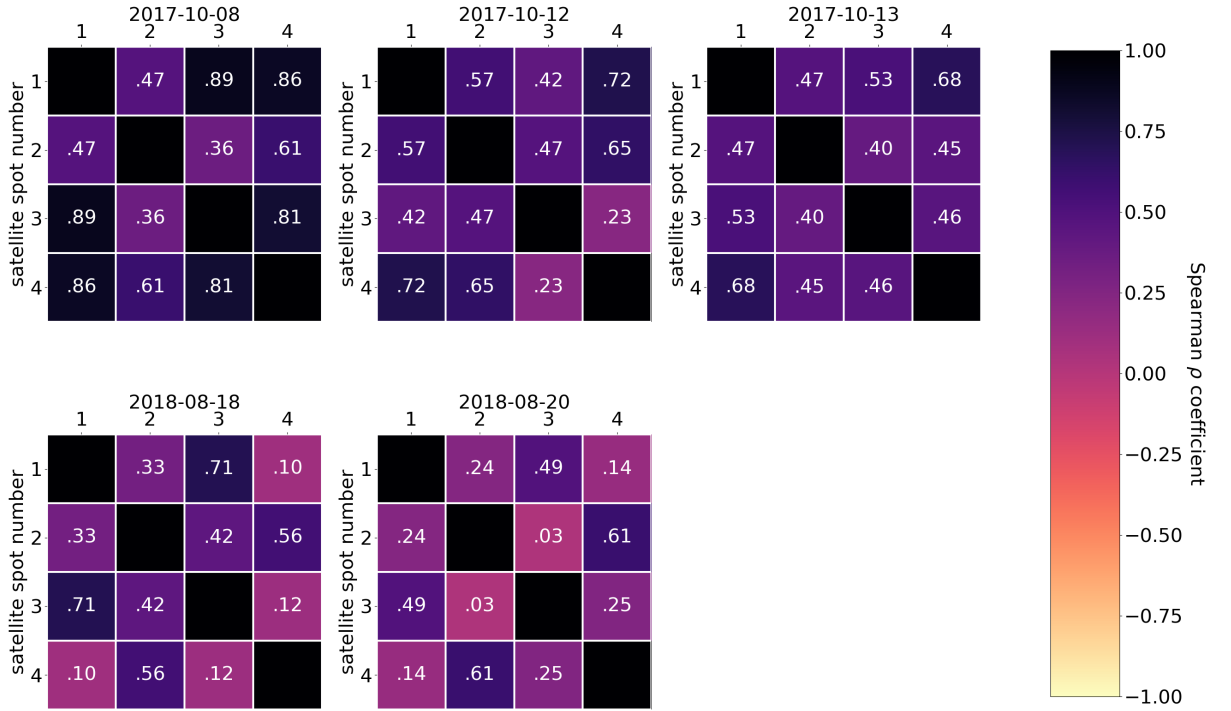
Figure 6. Spearman ρ coefficients between satellite spots, with no division by the mean lightcurve. A ρ value of 1 implies perfect correlation; a ρ value of -1 implies perfect anti-correlation. Conditions were particularly excellent for our August 2018 datasets; our satellite spot photometry thus covered a limited range of values, meaning that correlation between satellite spots is not a particularly good proxy for the quality of the observation.

(10 pixel width) at the planet radial separation to estimate the background at this position and also the median in a circular annulus aperture from 3-5 pixels around the 2-d Gaussian best fit position to remove any remaining trend in the background. Aperture photometry for HR 8799b as well as negative simulated planet photometry for HR 8799bc are presented in Figure 9.

5.3 Negative simulated planet subtraction

We must robustly remove speckle noise while preserving a photometric reference in order to determine well-calibrated time-series photometry for the HR 8799 planets. To do so, we adopt the standard technique of negative simulated planet subtraction. We first divide the observation into a number of temporal bins, which vary according to the planet considered (see Section 5.1). For each temporal bin, we then inserted negative simulated planets at the approximate positions of each of the four planets in a grid of radius, planet position angle (henceforth θ), and Δ(magnitude) into each frame. Three unsaturated point-spread function (PSF) images were taken before and after the time-series data. Planets were simulated by median combining these unsaturated PSF images and scaling them according to the $\Delta(mag)$ of the planet.

We performed the insertion in two steps:

1) Coarse Grid – in order to precisely determine the position of each planet, we ran an initial grid in radius and θ, with planet Δ(mag) kept fixed and set to literature values. After inserting the negative simulated planets, the VIP PCA pipeline (Gomez Gonzalez et al. 2017) was used to remove 10 principal modes, de-rotate, and stack data from each bin. From the de-rotated and stacked data, we then calculated the χ^2 residual for each planet at each grid point. Over the full grid, we then find the minimum χ^2 and hence the best solution for planet position. The χ^2 value for each grid point was calculated as the square of the sum of the counts in a circular region with radius of 5 pixels centered on the planet divided by the standard deviation in a circular noise annulus around the planet (5-10 pixels). For the negative simulated planet subtraction, the inhomogeneous background due to the PSF of the star is removed via the PCA analysis, so we chose a larger aperture size in comparison to the direct aperture photometry of HR 8799b described in Section 5.2, in order to obtain better statistics for evaluating the quality of our PSF subtraction. In radius, we ran the wide grid over a radial offset of -4 to +4 pixels from the initial radius guess, in steps of 0.2 pixels. In angle, we ran the wide grid over an angular offset of -2 to +2 degrees from the initial angle guess, in steps of 0.1 degrees.

2) Fine Grid – after estimating planet position with the coarse grid, we ran a finer grid in position as well as Δ(mag), centered on the best position found during the coarse grid step. From the fine-grid χ^2 grid, we then converted to likelihood (e.g. $exp(-\chi^2)$) and renormalized
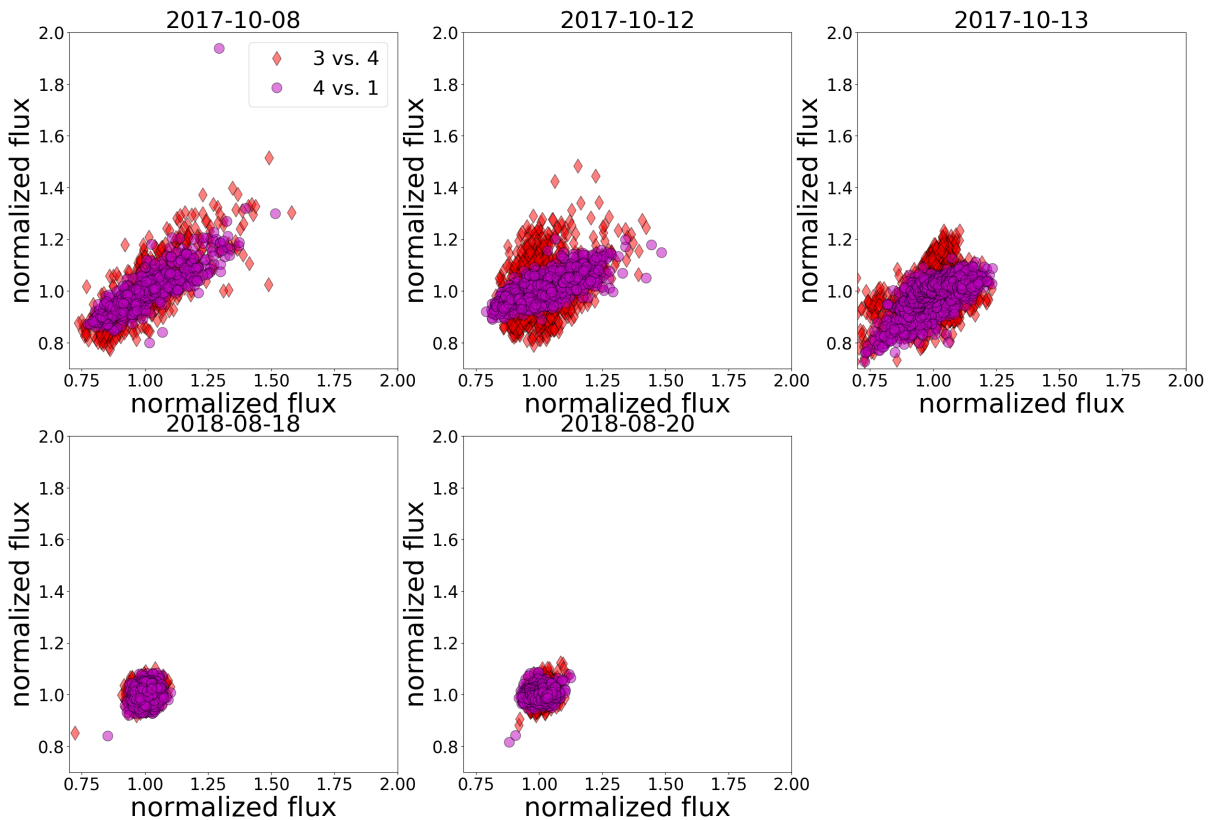
Figure 7. Correlation plots for satellite spots 3 vs. 4 and 4 vs. 1. The satellite spot photometry covers a much more limited range of values on nights with excellent conditions (August 2018) compared to nights with more moderate conditions (October 2017), leading to lower Spearman ρ values for August 2018 vs. August 2017, despite the higher quality of the August 2018 data.

the posterior probability distribution function (henceforth PDF) to 1, assuming that we had estimated planet position and $\Delta(mag)$ with sufficient accuracy to ensure that each planet was within our fine grid. We then marginalized over radius and θ, and fit a 1-d Gaussian to the 1-d posterior. The peak and width of the best-fit Gaussian was adopted as the best $\Delta(mag)$ and error on $\Delta(mag)$ respectively; in fact, this ability to robustly estimate errors is why we chose to calculate the posterior across the full fine grid, instead of using e.g. an amoeba downhill optimization procedure to select the highest probability set of parameters. In radius, we ran the fine grid over a radial offset from the best radius found by the wide grid, running from -1 to +1 pixel, in steps of 0.1 pixels. We ran the fine grid over an angular offset running from -1 to +1 degrees from the best angular position found by the wide grid, in steps of 0.1 degrees. In $\Delta(mag)$, we ran over a grid of offsets from -0.5 to 0.5 mag, in steps of 0.05 mag for the October 2017 epochs and from -0.333 to 0.333, in steps of 0.033 mag for the August 2018 epochs.

We selected 10 PCA modes as the ideal choice to best subtract speckles while preserving planet flux – we motivate this choice further in Sections 5.4 and 5.7. For HR 8799b and HR8799c, we then calculated best $\Delta(mag)$ values for 6 and 5 temporal bins respectively, evenly spaced in parallactic angle across the observation. For the night of 18 August 2018, we show a zoomed-in view of HR 8799b

or c respectively in Figures 10 and 11, before and after removing the best-fit negative simulated planet. The image scale in these figures runs from -10 to 10 ADU; while at the position of HR 8799b, speckle noise has been fully removed and noise statistics can be assumed to be Gaussian, a considerable amount of speckle noise remains at the separation of HR 8799c even after subtraction of 10 PCA modes.

Negative simulated planet photometry as a function of time (and converted to flux units) for HR 8799bc are presented in Figure 9. For each photometric measurement, to convert errors calculated in $\Delta(mag)$ to flux units, we drew 100000 samples from a Gaussian centered on the best $\Delta(mag)$ value and with σ given by the error on $\Delta(mag)$, then converted to flux units and adopted the standard deviation of the 100000 samples as our error in flux.

From Figure 9, it is clear that some nights yielded considerably better data than others – specifically, compared to the October 2017 epochs, the August 2018 epochs have much less spread in satellite spot measurements and correspondingly smaller spread in planet photometry measurements. Unsurprisingly, the two August 2018 epochs had the longest coherence times and highest Strehl ratios across all epochs (see Fig. 1). For the August 2018 epochs, the planet and satellite spots measurements generally agree. This implies that in these epochs, all or at least the bulk of the variation in planet flux we measure here is simply systematic trends as seeing, Strehl, etc., vary
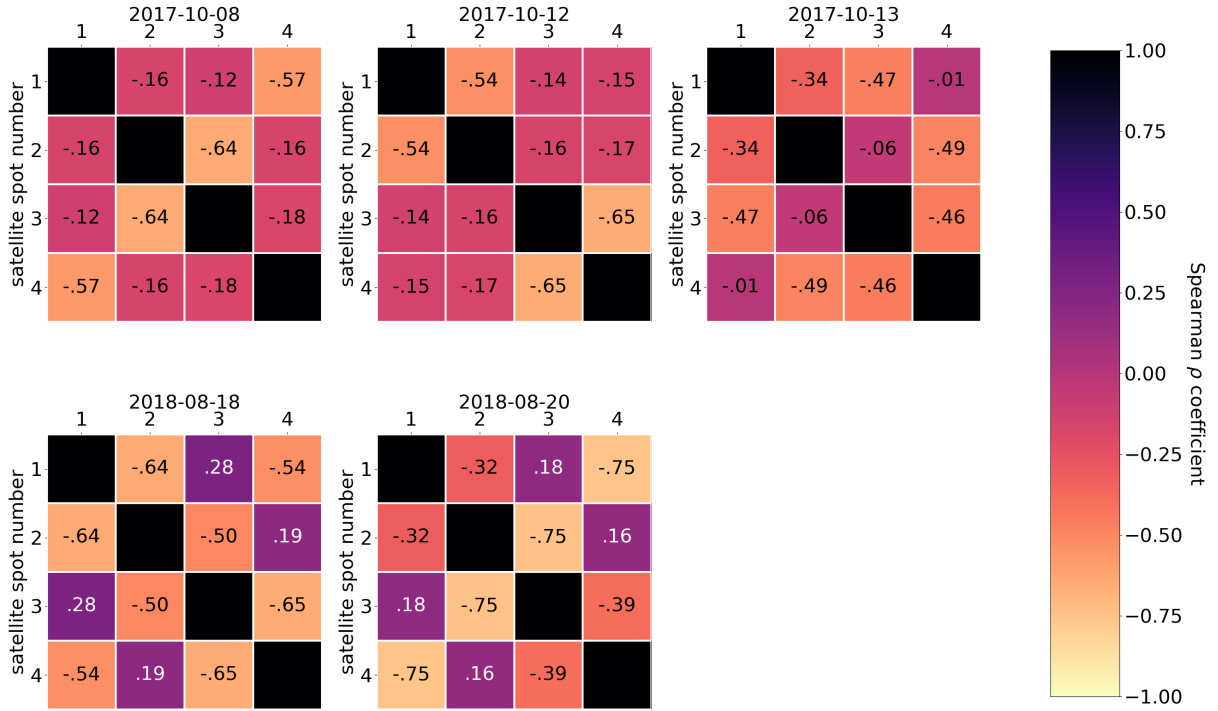
Figure 8. Spearman ρ coefficients between satellite spots, after division by the mean lightcurve (thus removing common variations). A ρ value of 1 implies perfect correlation; a ρ value of -1 implies perfect anti-correlation.

across the observation. In October 2017, however, we find significant divergence between planet and satellite spot lightcurves. Measurements for HR 8799b and HR 8799c also appear to be in agreement, except for epoch 2018-08-18, where the first and last measurements for HR 8799c are divergent from the HR 8799b and satellite spot lightcurves. However, we do not believe this is a bonafide detection of variability, as this result appears to be PCA-mode dependent and these two time bins were taken right at the beginning and end of the observation, where the airmass was at its highest and sky rotation at its lowest. We discuss this further in Sections 5.4 and 5.7.

5.4 Effect of number of PCA modes removed

Obtaining accurate photometry for planets embedded in speckle noise requires balancing the number of PCA modes utilized with the acceptable level of planet self-subtraction. Utilizing a greater number of PCA modes will more fully remove the speckle noise, but at the expense of removing flux from the planet itself. We tested the effect of using different numbers of PCA modes with the 18 August 2018 dataset (which appears by eye to be the highest-quality and most stable of the datasets) by re-running the fine-grid planet photometry pipeline with 4, 6, 10, and 16 modes. We used 6 equally spaced parallactic angle temporal bins for both b and c (as opposed to 5 for c) so we could directly compare results for the same sets of input data frames. Lightcurves are plotted in Fig. 12. Negative simulated planet photometry gives similar results for HR 8799b, independent of the

number of PCA modes used. However, we found that the negative simulated planet photometry for c varies more strongly with number of PCA modes chosen. For the most part, the retrieved photometry for HR 8799c is still consistent using a wide range of number of PCA modes, with the exception that the first and last parallactic angle bins diverge by $\sim 1\sigma$ from the trend given by the satellite spots if the number of PCA modes used is ≥ 10. However, these two points do agree with the satellite spot trend within 2σ. Given that they were taken at the extrema of the observation, at high airmasses and with a slow rate of sky rotation, they simply may not be reliable or may have true uncertainties higher than the uncertainty estimated here. In Section 5.7, we further consider the ideal number of PCA modes to adopt in this case, using simulated planet tests.

5.5 Inner Planets

Even with 4-5 hour long observations spanning >70° degrees in parallactic angle, we do not achieve sufficient sky-rotation to obtain high-sensitivity variability measurements for the innermost HR 8799 planets. Binning the observation into 3 equal parallactic angle bins will ensure $\geq 23°$ degrees of sky rotation. For the inner two planets, HR 8799d and HR8799e, at separations of ~0.6" and ~0.4", 23° of sky rotation corresponds to shifts of ~20 pixels and ~14 pixels on sky. Examples of final images after negative simulated planet removal for 2018-08-18 are presented in Figures 13. Like with HR
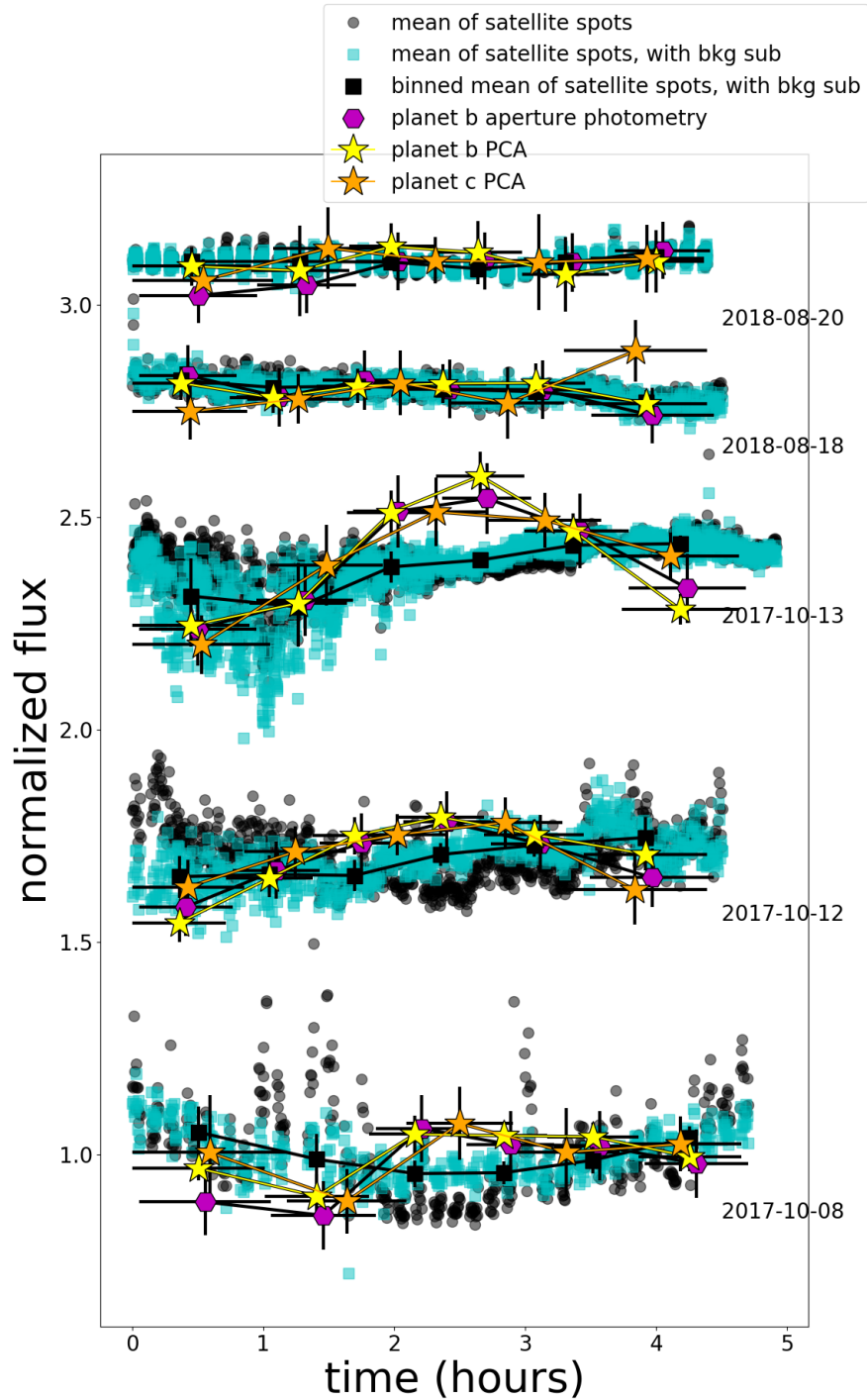
Figure 9. Satellite spot aperture photometry (with and without background subtraction), planet b aperture photometry, and planet b and c psf-fitting photometry as a function of time during the observation for October 2017 and August 2018 broadband-*H* datasets. Each lightcurve has been normalized using its mean flux value. Planet b aperture and psf-fitting photometry have been slightly offset in time to improve error bar legibility. Planets b and c almost always display the same trends as a function of time; in the August 2018 data these trends are also well-correlated with the trends seen in the satellite spot photometry.
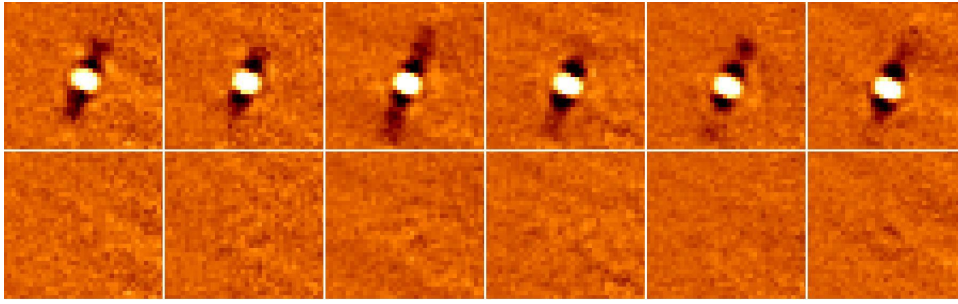
Figure 10. Best negative simulated planet subtractions for HR 8799b on 18 August 2018. The full ~4.5 hour observation has been divided into 6 equal parallactic angle bins, shown in temporal order from left to right. The image stretch runs from -10 to 10 ADU in a linear scale.
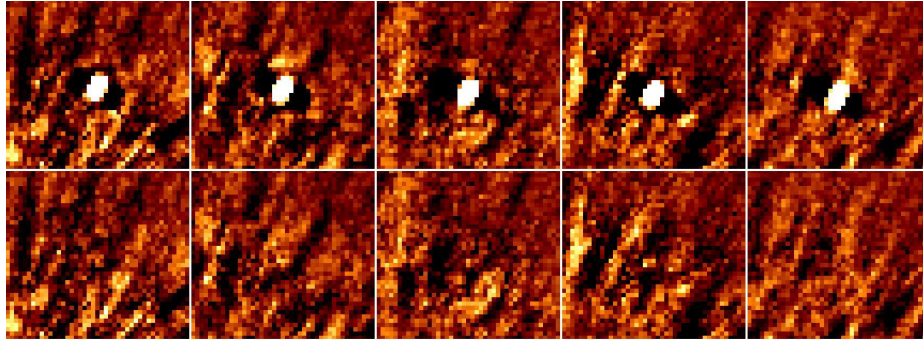
Figure 11. Best negative simulated planet subtractions for HR 8799c on 18 August 2018. The full ~4.5 hour observation has been divided into 5 equal parallactic angle bins, shown in temporal order from left to right. The image stretch runs from -10 to 10 ADU in a linear scale, with contrast adjusted to highlight the residual speckle noise after subtraction of 10 PCA modes.

8799c, significant speckle noise was still present after removing 10 PCA modes.

Lightcurves for HR 8799de are presented in Fig. 14, along with satellite spot photometry. Lightcurves for HR 8799de have been normalized to the middle temporal bin out of the three bins and are consistent within errors with satellite spot trends. However, the errors on the planet photometry are considerably larger than those on the satellite spot photometry, preventing any fine determination of variability relative to the satellite spots. The photometry measurements for HR 8799de are roughly consistent with each other to within 0.25 mag, thus we can rule out variability at the >30% level on 4-5 hour-long timescales.

5.6 Constraints on b-c colors – detrending b with c

We note that the lightcurves of HR 8799bc are almost always similar to each other, but in the October 2017 epochs, diverge significantly from the satellite spot lightcurves. Thus, while we initially considered detrending the planet photometry using the satellite spot photometry, it is clear that the satellite spot lightcurves do not provide an appropriate photometric reference for the planet photometry at all epochs. Instead, we follow the approach of Apai et al. (2016) and calculated $\Delta(mag)_b - \Delta(mag)_c$ – essentially detrending the b lightcurve using that of c, or vice versa. This traces non-shared variations between the two planets. We considered five temporal bins here, to assure 20 pixels of movement on the sky for HR 8799b and c, with difference lightcurves presented in Fig. 15. $\Delta(mag)_b - \Delta(mag)_c$ measured by Apai et al. (2016) is overplotted as a purple horizontal line. Within errors, our results are in agreement with those from Apai et al. (2016). The uncertainty on the measurement is dominated by the uncertainty

in our measurement of $\Delta(mag)_c$. Depending on epoch, we rule out non-shared variability in $\Delta(mag)_b - \Delta(mag)_c$ to the 10-20% level over 4-5 hours.

5.7 Simulated variability results

To quantify our sensitivity to variability in the 2018-08-18 epoch (the best quality data of our search), we simulate variable and non-variable lightcurves by inserting and retrieving a suite of simulated planets at a similar radius from the star as HR 8799bc, but offset in position angle from the true planet position angle. Thus, we sample a similar region of the speckle pattern, but place each simulated planet sufficiently far from its real counterpart that they do not interfere with each other during the PCA analysis. In the simulated lightcurves, we estimate systematic noise sources to first order (e.g. varying Strehl ratio, seeing, etc) expected in the data by scaling the simulated planet flux by the normalized satellite spot aperture photometry (averaged across all 4 satellite spots) after background subtraction. While this is not a rigorous attempt to model all potential noise sources, it is sufficient to set constraints on the amplitudes and periods of variable signals that we can potentially detect. Note the satellite spot aperture photometry is only an appropriate noise model for the August 2018 epochs (and not for the October 2017 epochs), as in August 2018, planet photometry and satellite spot photometry largely followed similar trends. We then simulate either a constant planet flux or a sinusoidal variability signal. Since the positions of the simulated planets are known a priori, we do not run the full 3-d planet retrieval grid, but only consider a 1-d grid of Δmag at the a priori planet position.

We first consider the effect of using different numbers of PCA
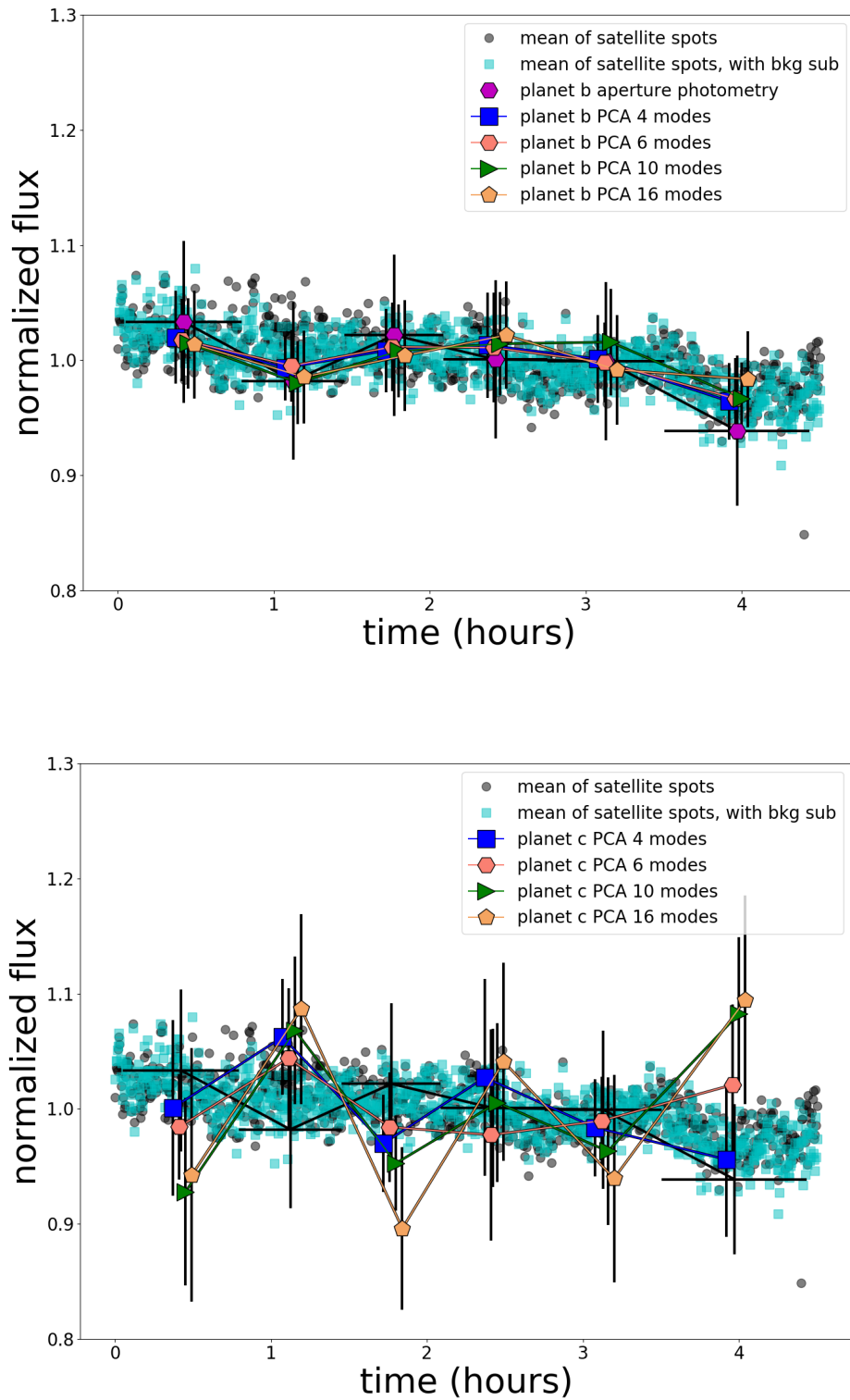
Figure 12. The effect of utilizing different numbers of PCA modes on HR8799b and c lightcurves for 18 August 2018. The lightcurve for planet b is consistent independent of the number of modes used. The retrieved photometry for HR 8799c is largely self-consistent using a wide range of number of PCA modes, with the exception that the first and last parallactic angle bins diverge by ~1σ from the trend given by the satellite spots if the number of PCA modes used is ≥10. However, these two time bins were taken at the extrema of the observation, where the airmass was at its highest.
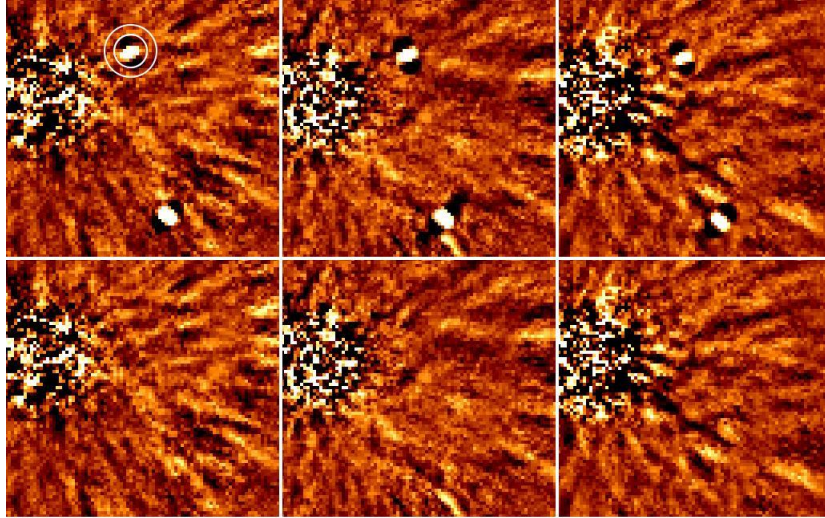
Figure 13. Best negative simulated planet subtractions for HR 8799d and HR 8799e on 18 August 2018. The full ~4.5 hour observation has been divided into 3 equal parallactic angle bins, shown in temporal order from left to right. The image stretch runs from -10 to 10 ADU in a linear scale, with contrast adjusted to highlight the residual spectral noise after subtraction of 10 PCA modes. Apertures used for calculating planet flux and background flux in the negative simulated planet algorithm are shown as a white circle and annulus respectively in the image for the first parallactic angle bin.

modes when retrieving simulated planets. Retrieved lightcurves (using 4, 10, or 16 PCA modes) for simulated constant-flux versions of HR 8799bc are presented in Fig. 16. The simulated HR 8799b lightcurve is robustly retrieved using 4, 10, or 16 modes, but the 4 and 16 mode retrievals look considerably noisier than the 10 mode retrieval, further justifying our adoption of 10 modes as standard. The retrieved lightcurves for HR 8799c are noisier than those for HR 8799b. Using 4 or 10 modes, the HR 8799c simulated lightcurves match the satellite spot lightcurve within errors. The 16 mode lightcurve is extremely noisy and is clearly oversubtracting the planet flux, resulting in significantly higher uncertainties. In all cases, the final photometric point in the HR 8799c simulated lightcurve is brighter than the satellite spot lightcurve at the same bin; this mirrors what is found in the actual HR 8799c lightcurve. At this radius from the star and this temporal point in the observations, we suspect that speckle residuals remain that are not accounted for in the satellite spot lightcurves. In other words, the satellite spot lightcurve does not serve as an appropriate photometric reference for this particular data point.

In Fig. 17, we present simulated lightcurves for HR 8799bc with a period of 8.6 hours, similar to the period of the variable free-floating planetary mass object PSO J318.5-22 (Biller et al. 2018), a phase of 40°, and with variability amplitudes of 0%, 5%, 10%, and 20%. As before, we use 6 and 5 equal parallactic angle bins for HR 8799b and c respectively, to ensure at least 5 FWHMs of motion on the sky in each parallactic angle bin. Variability is clearly apparent in the simulated HR 8799b lightcurves with amplitudes ≥10%. Variability detection is much more difficult for the noisier HR 8799c lightcurves, notably stymied by the high final photometric point.

To estimate our sensitivity to variability across a wide range of different amplitudes, periods and phases, we ran a complete grid of simulated planet lightcurves using 5 equal parallactic angle temporal bins, with periods from 2 to 20 hours (in steps of 2 hours), amplitudes from 3% to 30% (in steps of 3%), and phases from 0° to 330° (in steps of 30°). We then detrend each simulated HR8799bc lightcurve by dividing through by the mean satellite spot lightcurve, with errors on the detrended lightcurve given as the sum in quadrature of the errors on the simulated planet and satellite spot lightcurves. For

each detrended simulated lightcurve at a specific value of period, amplitude, and phase, we set up a metric for the detection of variability in the detrended lightcurve as follows: If (maximum(flux) - minimum(flux)) / maximum(error) >2, we consider variability to be detected (e.g. a 2-σ variability detection). This metric was tested by eye for a significant number of simulated lightcurves, and generally agrees with visual assessment. To build the detection map, we assign each combination of period, amplitude, and phase the value 1 if variability is detected, and 0 if variability is not detected. We then sum along the phase axis and divide by the number of phases simulated to produce a fractional detection map as a function of period and amplitude. Detection maps for HR 8799bc in the 2018-08-18 epoch are presented in Fig. 18. For HR 8799b, for periods <10 hours, we are generally quite sensitive to variability with amplitude > 5%. For HR 8799c, our sensitivity is much lower, limited to variability > 25% for similar periods.

6 EPOCH-BY-EPOCH RESULTS

In this section, we present epoch-by-epoch astrometry and photometry, and compare to literature values. For this purpose, we used only the data frames from the hour approaching meridian crossing and the hour after meridian crossing, thus covering the time period during each observation with the most field rotation and best airmass. We again ran our full negative simulated planet code to retrieve the best Δ(mag), radius, and position angle value for each planet.

Astrometric results are plotted compared to literature astrometry in the left panel of Fig. 19. A zoomed-in plot on our epochs of astrometry is presented in right panel of Fig. 19 and the epoch-by-epoch astrometry for each planet is presented in Tab. 2. To convert the uncertainties derived from the negative simulated planet code from radius and position angle values to ΔRA and ΔDEC, we drew 100000 samples from Gaussians centered on the best radius and position angle values and with σ given from the respective errors on from the negative simulated planet subtraction, converted these values to ΔRA and ΔDEC, then adopted the mean and standard deviations of these distributions as the corresponding ΔRA, ΔDEC
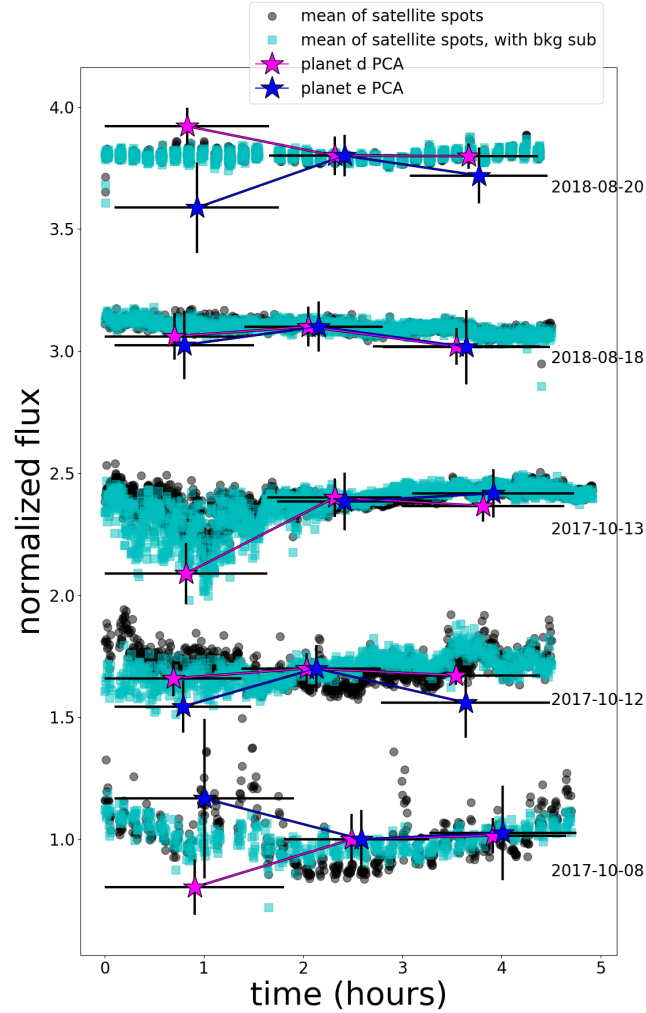
Figure 14. Satellite spot aperture photometry (with and without background subtraction), planet b aperture photometry, and planet d and e psf-fitting photometry as a function of time during the observation for October 2017 and August 2018 broadband-*H* datasets. Planets d and e are consistent within errors with the satellite spot photometry, however, our uncertainties are much higher for HR 8799de relative to HR 8799b, as this region of our images still show considerable speckle residuals even after removing 10 PCA modes.

values and 1-σ uncertainties. In this case, we conservatively adopt the 2-σ values for uncertainties, which gives us the uncertainty on the relative planet position σ_{PP} at this epoch. However, we must consider further astrometric uncertainties in our full error budget. Following Zurlo et al. (2016), we adopt a plate scale uncertainty σ_{PS} of 2 mas and star center position σ_{SC} uncertainty of 1.2 mas; the full uncertainty is then given as the sum in quadrature of σ_{PP}, σ_{PS}, and σ_{SC}. Our astrometric points are consistent with the already-observed orbital motion of these planets, however, we defer orbital fitting of these points and further epochs of SPHERE orbital monitoring to Zurlo et al. in prep.

Our epoch-by-epoch values of the absolute *H* magnitude of each planet are compared to literature photometry in the *H* band in Figs. 20 to 22 and presented in tabular form in Table 3. To convert from Δmag to absolute magnitude in *H*, we adopt a parallax of 24.2175±0.0881 mas from Gaia DR2 (Gaia Collaboration et al. 2018) and an *H*-band

magnitude of 5.280±0.018 mag from the 2MASS all-sky survey (Skrutskie et al. 2006). Since the star is a known variable with an amplitude of up to 0.1 mag in the optical, we have added a conservative error of 0.1 mag on the magnitude of the star, in quadrature with the error we derive from our negative simulated planet subtraction photometry. The large errors endemic to converting from Δ(mag) to absolute mag prevent a meaningful epoch-to-epoch comparison in absolute H, however, our photometry is constant to within error bars both between our own epochs and with *H*2 values from Zurlo et al. (2016) (a somewhat different band), H-band values for all planets from Skemer et al. (2012), and HR 8799b and c values from Oppenheimer et al. (2013). We see significant divergence from Marois et al. (2008) and HR 8799d and e values from Oppenheimer et al. (2013). Because of the inherently large uncertainties in converting from Δ(mag) to absolute mag, Apai et al. (2016) instead opt to consider $\Delta(mag)_b$ - $\Delta(mag)_c$, etc., in other words, the relative difference
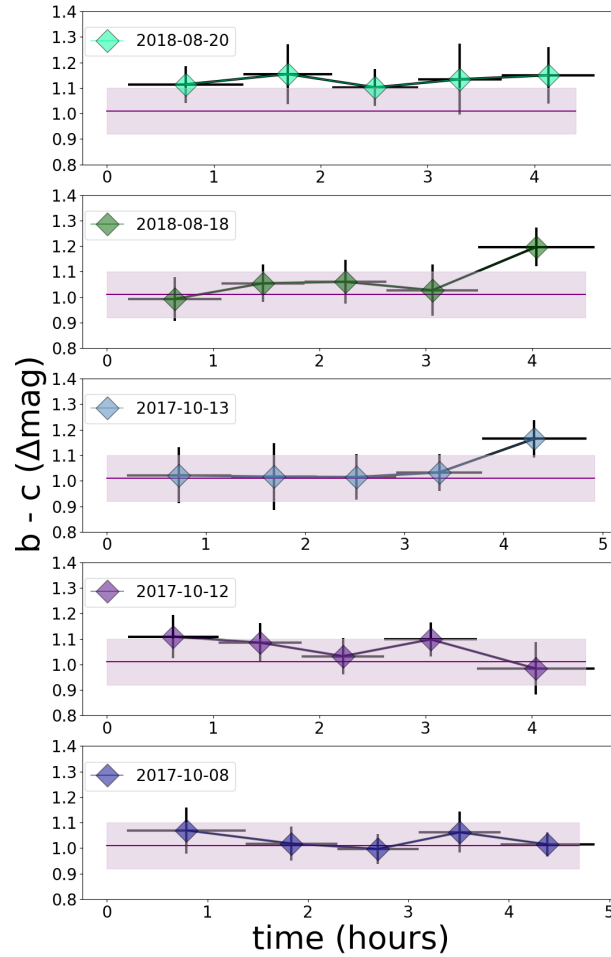
Figure 15. Δ(mag) of HR 8799b - Δ(mag) of HR 8799c as a function of time during each broadband-*H* observation, compared to the results from Apai et al. (2016) (purple line and shaded rectangle). Δ(mag) of HR 8799b - Δ(mag) is consistent with being constant within errors.

in contrast between different planets in the system. We plot our measurements of this value in comparison to those from Apai et al. (2016) in Fig. 23 and present our measurements in tabular form in Table 4; our values all agree within 1-σ with those from Apai et al. (2016), with no compelling evidence of epoch-to-epoch variability.

7 DISCUSSION

7.1 Use of satellite spots as photometric references

As noted in Section 2, HR 8799 is a γ Doradus pulsating variable, with a dominant periodicity of ~0.5 day, and a peak-to-valley amplitude of up to ~0.1 mag in the visible (Sódor et al. 2014). SPHERE crucially enables monitoring of all planets as well as the satellite spots simultaneously, thus allowing us to disentangle any intrinsic variability of the star from variability of the planets. As we do not expect the planets to have the same rotational period, phase or variability amplitude, any variability with the same period shared by

multiple planets would then probably be intrinsic to the star. Thus, our expectation was that the satellite spot light curves should follow the trend produced by the changing atmospheric conditions as well as any astrophysical variability from the star, as would the lightcurve of a non-variable exoplanet. This did not prove to be universally true. The lightcurves for HR 8799bc closely follow each other, suggesting that the variations observed are driven by the observing conditions or intrinsic variability of the star, and do not indicate variability in the planets. However, while the lightcurves for HR8799bc closely follow that of the mean satellite spot lightcurve in the August 2018 data, there are significant deviations between the satellite spot and planet lightcurves in the October 2017 data. This means that the satellite spots do not always provide appropriate photometric references for detrending the broadband *H* exoplanet lightcurves; this seems to be correlated to the observing conditions and may be due to the fact that the planet spectra are similar to each other, but much redder than the spectrum of the star (which is shared by the satellite spots). To test whether this is the case, it would be interesting to observe a binary
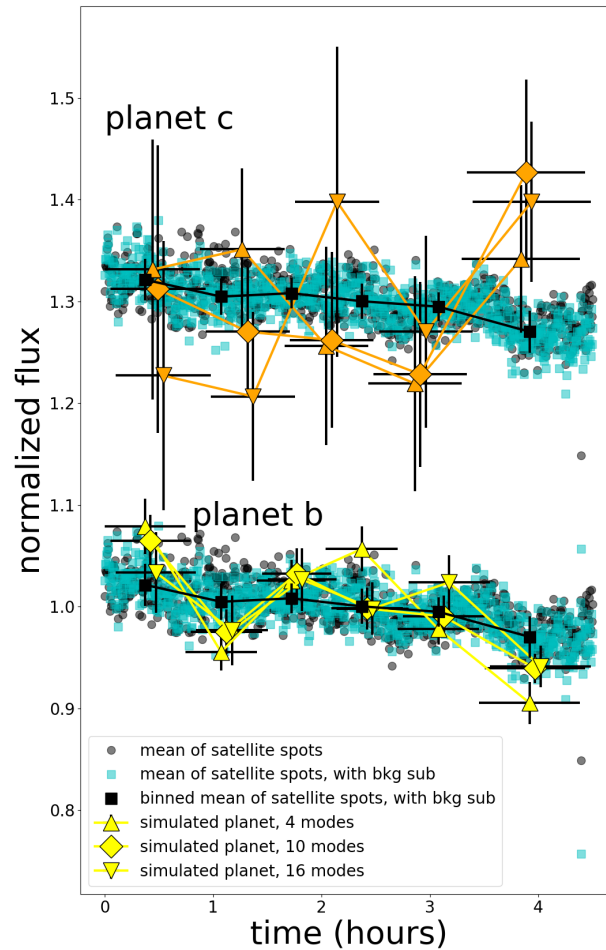
Figure 16. The effect of utilizing different numbers of PCA modes on simulated constant-flux HR8799b and c lightcurves for 18 August 2018 (noise simulated using the satellite spot lightcurve, inserted at the same radius but differing PA compared to the actual planets). The planet b simulated lightcurve is consistent independent of the number of modes used. The retrieved photometry for HR 8799c is much noisier, with the 10 mode lightcurve providing the best trade-off between preserving planet flux and suppressing speckles. The final photometric point in HR 8799c simulated lightcurve is considerably brighter than the satellite spot lightcurve at the same bin using 4, 10, or 16; this mirrors what is seen in the actual HR 8799c lightcurve. At this radius from the star and this temporal point in the observations, speckle residuals remain that affect our negative simulated planet subtraction, but are not accounted for in the satellite spot lightcurves.

companion with a more similar (e.g. bluer) spectrum to its host, along with the satellite spots. If the binary companion displays the same trends as the satellite spots in a range of different observing conditions, the divergence between planet and satellite spot lightcurves we measure here would likely be due to the red colours of the planets relative to the star. In this case, a narrow band or integral field spectroscopic search for variability would not have to contend with this particular issue, but would rather be limited by the brightness of the planet in a narrow bandpass.

However, the source of the divergence may instead be that the satellite spots are artificially introduced and do not fully share properties of imaged astrophysical objects. The satellite spots for SPHERE are generated by introducing a waffle pattern onto the SPHERE deformable mirror, producing 4 artificial sources in a cross-pattern

(Langlois et al. 2013). In contrast, GPI uses a square grid in the pupil plane of the telescope to produce a similarly shaped cross-pattern of 4 first-order diffraction spots (Sivaramakrishnan & Oppenheimer 2006; Wang et al. 2014). It would be interesting to investigate whether these different methods to produce artificial references cause differences in the behaviours of these references; note however, that Figure 7 of Wang et al. (2014) seems to show similar correlations / anticorrelations between different satellite spots as we found in Section 4.1 and Appendix C. Unlike a real companion or stellar image, satellite spots in SPHERE and GPI are not incoherent with stellar residuals. The coherent satellite spots suffer from pinning effects (interference between the satellite spots and PSF residuals) and provide only reduced accuracy for photometry and astrometry. Recently, Sahoo et al. (2020) has experimented with introducing fiducial incoherent faint
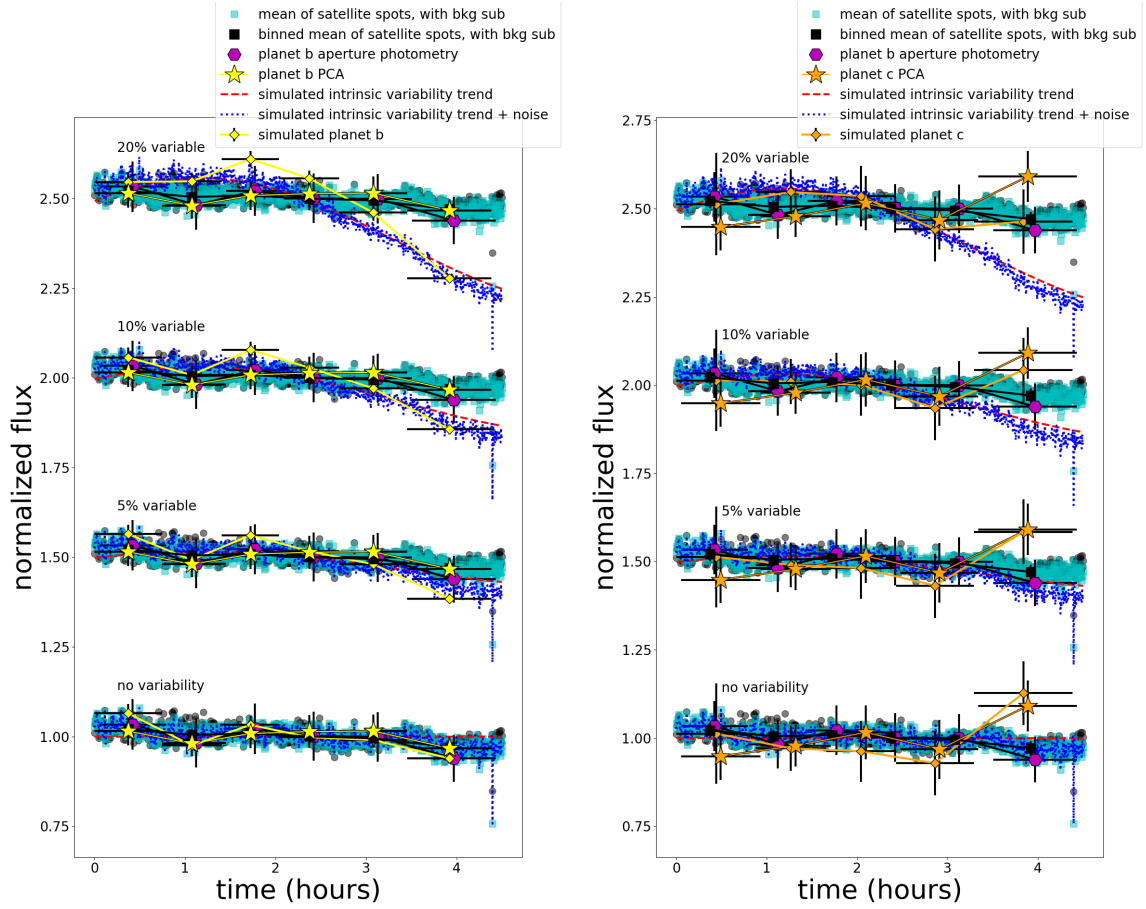
Figure 17. Simulated lightcurves for HR 8799bc with sinusoidal variability with a period of 8.6 hours, and amplitudes of 5%, 10%, and 20%. Variability is apparent by eye in the simulated HR 8799b lightcurves with amplitudes $\geq 10\%$, but there is no clear detection of variability for the simulated HR 8799c lightcurves even for the 20% amplitude curve.
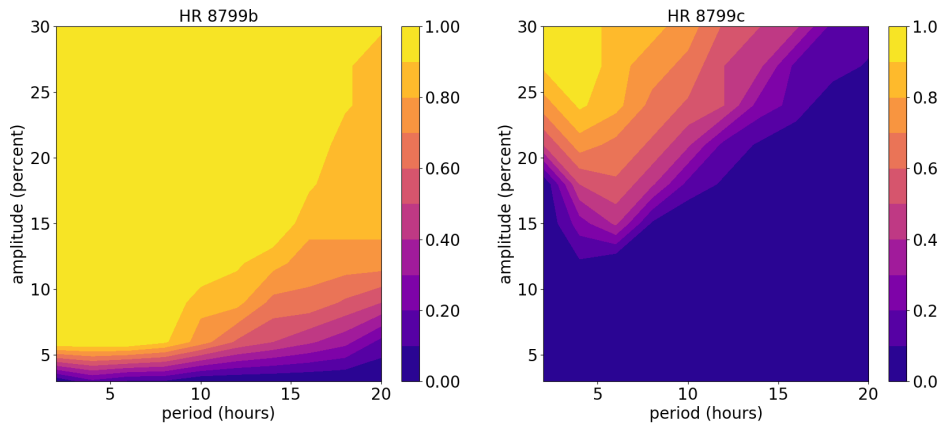
Figure 18. Simulated planet fractional detection maps for HR 8799bc in our 2018-08-18 data. The color scale is set by the fraction of simulated lightcurves detected at each combination of period and amplitude. For HR 8799b, for periods <10 hours, we are generally quite sensitive to variability with amplitude > 5%. For HR 8799c, our sensitivity is much lower, limited to variability > 25% for similar periods.
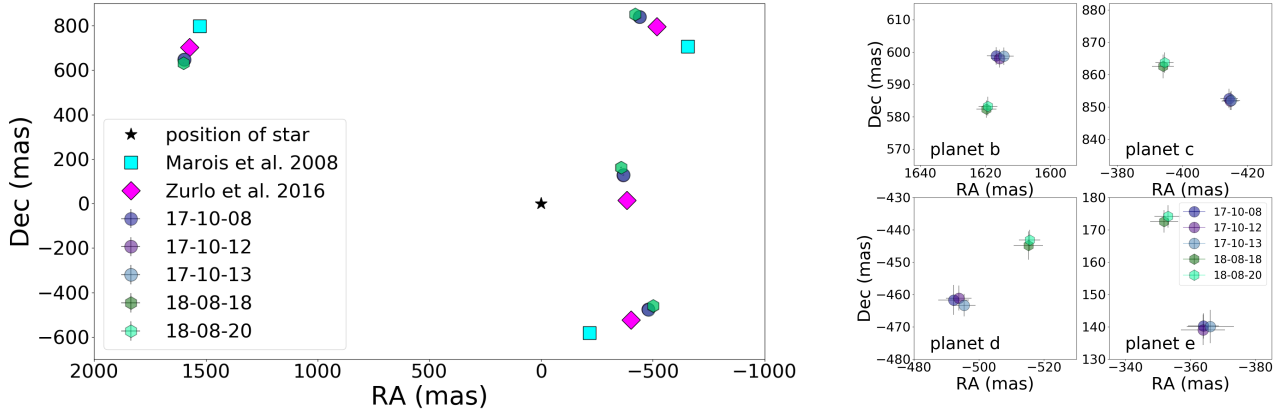
Figure 19. *Left:* Astrometric comparison between our epochs, Marois et al. (2008), and Zurlo et al. (2016). The error bars are smaller than the plot symbols; all four planets show significant on-sky motion even between our 2017 and 2018 epochs. *Right*: Planet-by-planet view of astrometry from our October 2017 and August 2018 epochs.

copies of the host star in the image plane of Subaru ScEXAO and then alternating the pattern of these copies between exposures. This produces an incoherent photometric reference and may solve some of the issues of using satellite spots as photometric references encountered in this study.

7.2 Sensitivity to variability / comparison with known variable objects

HR 8799bcde share very similar spectra and colours to highly variable mid-to-late-L variable planetary mass objects such as VHS 1256-1257ABb (Bowler et al. 2020; Zhou et al. 2020) and PSO J318.5-22 (Liu et al. 2013; Biller et al. 2015, 2018). Although variability amplitude can vary between epochs for these objects, these objects are always highly variable in the near-IR, and particularly in J-band – VHS 1256-1257ABb has been observed to be variable by 24.7% over ~8 hours at 1.27 μm (Bowler et al. 2020) and PSO J318.5-22 has been observed to be 7-10% variable over ~4 hours in the J_S band (Biller et al. 2015). These two objects are observed at high inclinations; Zhou et al. (2020) find that the $v\sin i$ and period of VHS 1256-1257ABb is most consistent with an equator-on viewing inclination and Biller et al. (2018) measure an inclination for PSO J318.5-22 of 56.2±8.1°. Field brown dwarfs (Radigan et al. 2014; Metchev et al. 2015) display the highest amplitude variability in J band (~1.1-1.3 μm), decreasing somewhat through the rest of the near-IR, and with considerably lower mid-IR (3-5 μm) amplitudes; VHS 1256-1257ABb shows a similar trend to the field brown dwarfs at all wavelengths, with the mid-IR amplitude less than half the near-IR amplitude (Zhou et al. 2020), while PSO J318.5-22 is a rare exception to this rule, with nearly equal near-IR vs. mid-IR amplitudes.[2] As the HR 8799bcde planets share comparable atmospheric properties to these objects, they may also have high intrinsic variability amplitudes in the near-IR.

In our best epoch on 18 August 2018, for periods less than 10 hours, we were sensitive to 5-10% variability in HR 8799b and to ≥25% variability in HR 8799c in the broadband H filter. However,

---

[2] PSO J318.5-22 is also notably one of few brown dwarfs or exoplanet analogues with simultaneously measured near-IR and mid-IR amplitudes.
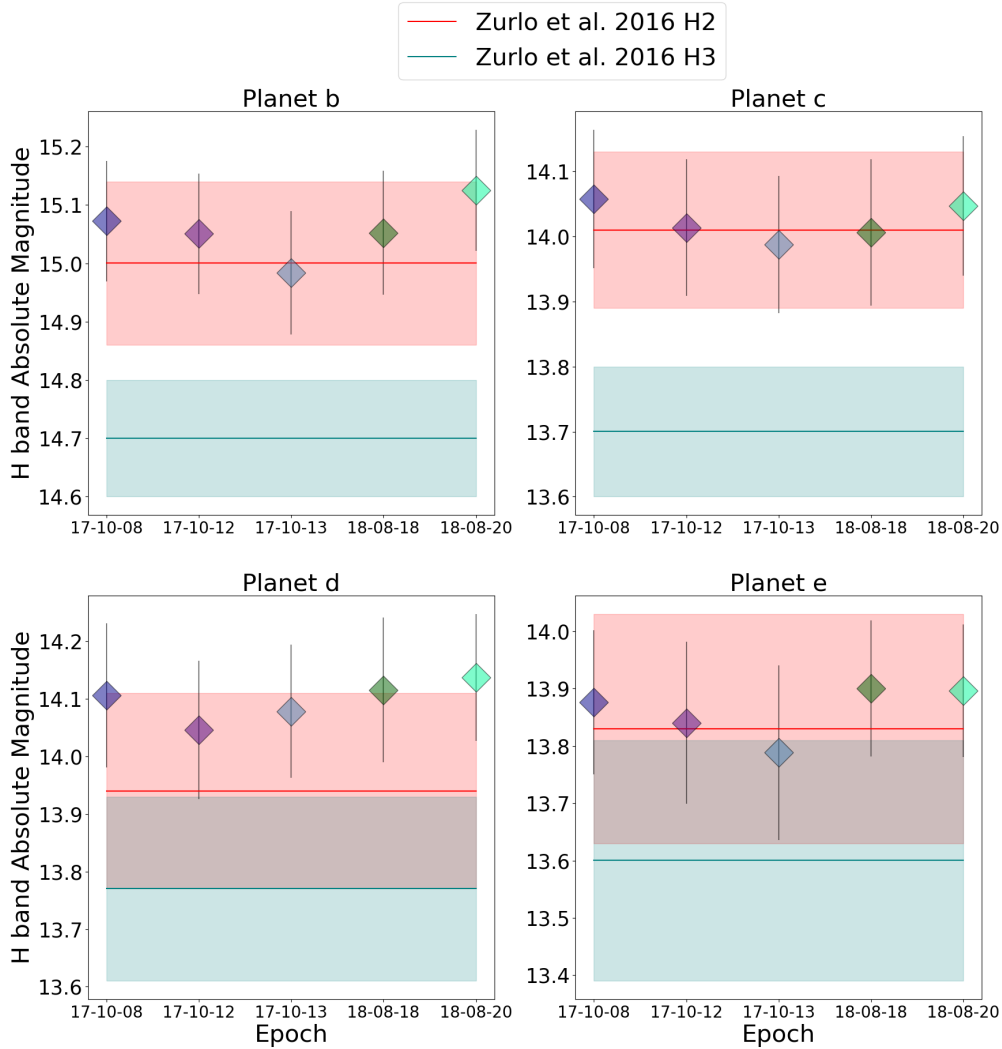
the measured variability amplitude will be less than the intrinsic variability amplitude due to geometric dilution effects and attenuation of flux within the line-of-sight through the planet's own atmosphere. Vos et al. (2017) estimate the contributions of geometric effects and atmospheric attenuation as:

$$A = A_O \sin i - \kappa \frac{dx}{\sin i} \qquad (1)$$

where A is the measured variability amplitude, A_O is the intrinsic variability amplitude, i is the inclination of the planet, $\frac{dx}{\sin i}$ is the path length of the light through the atmosphere of the planet, and κ is the degree of attenuation through the atmosphere per unit length (see also Fig. 4 in Vos et al. 2017). The first term gives the geometric dilution due to projection. The HR 8799 disk is inclined by ~27° (based both on disk studies and orbital fits for the individual planets, see e.g. Matthews et al. 2014; Ruffio et al. 2019; Wang et al. 2018a; Konopacky et al. 2016; Pueyo et al. 2015; Maire et al. 2015). For brown dwarfs with measured inclination angles, Vos et al. (2017, 2020) find a relationship between inclination angle and $(J - K)_{2MASS}$ colour anomaly (object $J - K$ - average $J - K$ for other objects with the same spectral type and gravity class). They find that objects viewed at high inclination (equator-on) are redder than the average object of the same spectral type and objects viewed at low inclination (pole-on) are comparatively bluer. Assuming that HR 8799b shares the inclination of the disk (but noting that the inclination of the planet has not been measured), adopting an L7 spectral type, and calculating colour anomaly using photometry from Marois et al. (2008); Currie et al. (2011) converted from MKO into 2MASS filters using the conversion from Stephens & Leggett (2004), we plot HR 8799b alongside field and low-gravity brown dwarfs on the same inclination vs. colour anomaly plot in Fig. 24. The colour anomaly is consistent with a wide range of possible inclinations for this planet.

However, assuming the planets do share the inclination of the disk and ignoring additional attenuation from the atmospheres of the planets themselves implies that the intrinsic variability amplitude for the HR 8799 planets will be at least a factor of 2.2 greater than any measured variability amplitude. Thus, if the planets are observed nearly pole-on, we are sensitive to **intrinsic** variability amplitudes of 11-22% for HR 8799b, and >55% for HR 8799c. Assuming that VHS 1256-1257ABb and PSO J318.5-22 are observed nearly equator-

Figure 20. Our broadband-*H* epoch-by-epoch photometry compared to SPHERE *H*2 and *H*3 photometry from Zurlo et al. (2016).

on (hence measured variability amplitude is nearly the same as the intrinsic variability amplitude), we could potentially have detected variability if HR 8799b was a VHS 1256-1257ABb analogue, but a PSO J318.5-22 analogue would have remained undetected (as we found from our simulated planet tests). Taking atmospheric attenuation into account as well (second term in Equation 1) further reduces the measured variability amplitude as a function of decreasing inclination. This term is sensitive to the wavelength being monitored; we expect light at near-IR wavelengths to be produced deeper in a given object than light at mid-IR wavelengths, thus, near-IR light will traverse a longer path length and will be more attenuated as a function of inclination.

While the observed photometric variability for brown dwarfs and planetary mass objects to date is consistent with inhomogeneous top-of-atmosphere features (e.g. quasiperiodic variability that evolves on timescales of a few rotational periods), a transiting companion to an exoplanet would provide another potential source of variability. The geometry of the HR 8799 system is not ideal for this scenario – assuming the planet and the companion to the planet are coplanar

and have inclinations similar to that of the system, observable transits would be unlikely for this face-on system. However, a companion to one of the HR 8799 planets with an orbit inclined perpendicularly to the inclination of the system could potentially produce transits in this system. Lazzoni et al. in prep simulate transits due to companions to known planets for a range of planetary systems, focusing primarily on the edge-on β Pic system. For HR 8799b, simulating an ensemble of satellites around HR 8799 b with masses in the range $[0.001,0.39] \times M_{Jup}$ and semi-major axis in the range $[0.001, R_h]$, where R_h is the Hill radius of the planet yields numerous transit events with depths greater than 10%. However, the transit durations are often quite long – up to tens of hours, with only a minority of simulated transits shorter than 10 hours. Thus, it is likely that ground-based observations with durations <10 hours, such as those reported here, would simply be too short to realistically detect such a transit.
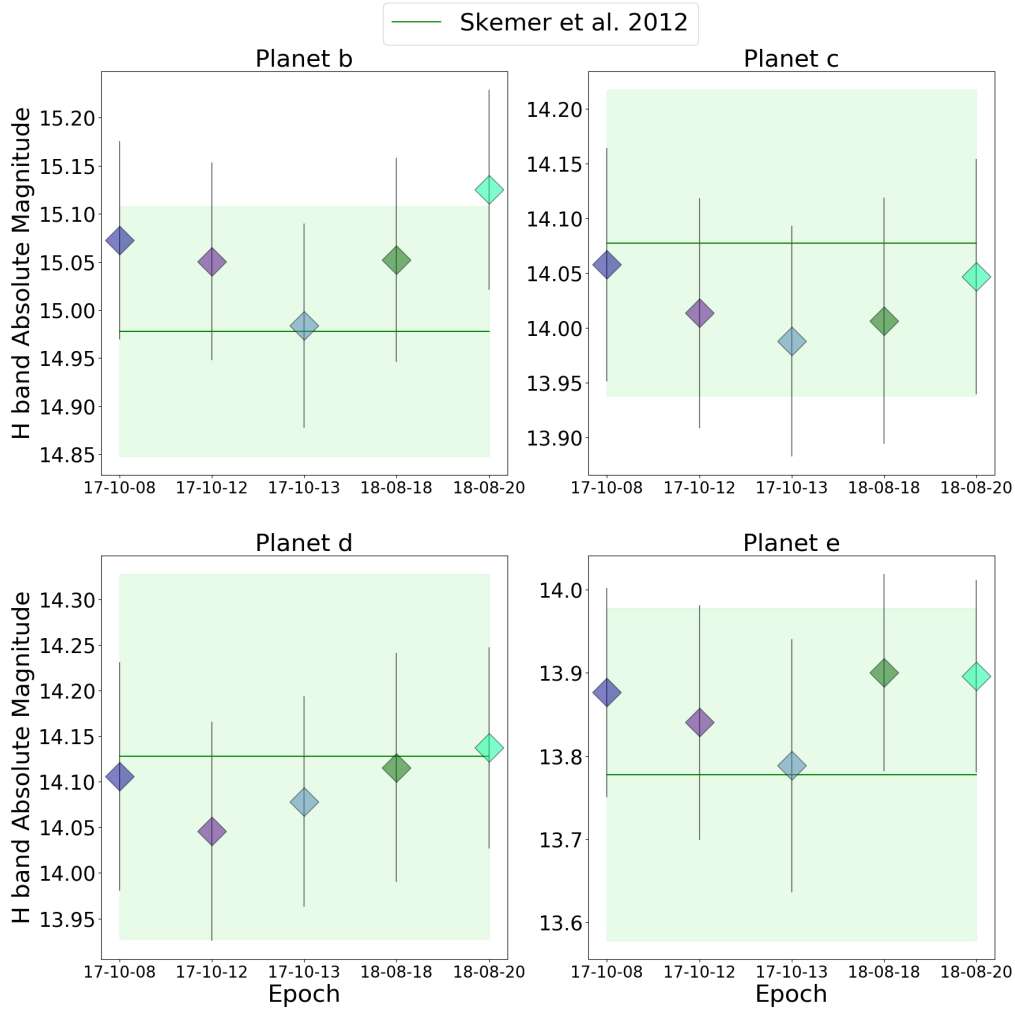
Figure 21. Our broadband-*H* epoch-by-epoch photometry compared to LBT *H*-band photometry from Skemer et al. (2012).

8 SUMMARY AND CONCLUSIONS

Here we summarize the main results of this study:

(i) The satellite spots were well-illuminated (∼2000 counts at peak) in all epochs except 2017-10-08 (∼200 counts at peak). For these well-illuminated epochs, background subtraction did not noticeably affect the mean satellite spot lightcurve, as the background level was so much lower than the satellite spot peak. For 2017-10-08, subtracting the background noticeably detrended the mean satellite spot lightcurve.

(ii) The degree of excursion (maximum vs. minimum satellite spot lightcurve value) varies strongly for different epochs of observations – the August 2018 epochs are considerably more stable and have noticeably smaller ranges of values in the satellite spot lightcurves than the October 2017 epochs. This leads to overall better sensitivity to variability and improved ability to detrend planet lightcurves.

(iii) In contrast to Apai et al. (2016), we do not find a night-by-night stable pattern of correlations / anti-correlations between satellite spots. In particular, for our best datasets (August 2018), we

find low levels of correlation / anti-correlation between satellite spots, because the satellite spot photometry is very stable on those nights and do not cover a wide range of values (in other words, the satellite spot lightcurves are quite flat in these epochs, so do not appear very correlated at all).

(iv) Ensuring 5×FWHM sky rotation at the exoplanet radius to optimize speckle suppression but minimize self-subtraction, our ∼4.5 hour observations allow 6 equal parallactic angle temporal bins for HR 8799b, 5 such bins for HR 8799c, but only 3 for HR 8799de.

(v) Planet photometry / astrometry for HR 8799bc was retrieved using negative simulated planet subtraction, inserting scaled unsaturated images of the star at the Δmag, position of each planet in each raw data frame, for 6 temporal bins for b and 5 temporal bins for c. As HR 8799b is sufficiently far from the core of the speckle pattern, we were able to also retrieve direct aperture photometry for this planet. In all epochs, HR 8799b aperture photometry and HR 8799bc negative simulated planet photometry share similar trends within uncertainties, except for the final temporal bin on 2018-08-18, where HR 8799c appears bright compared to HR 8799b. This is
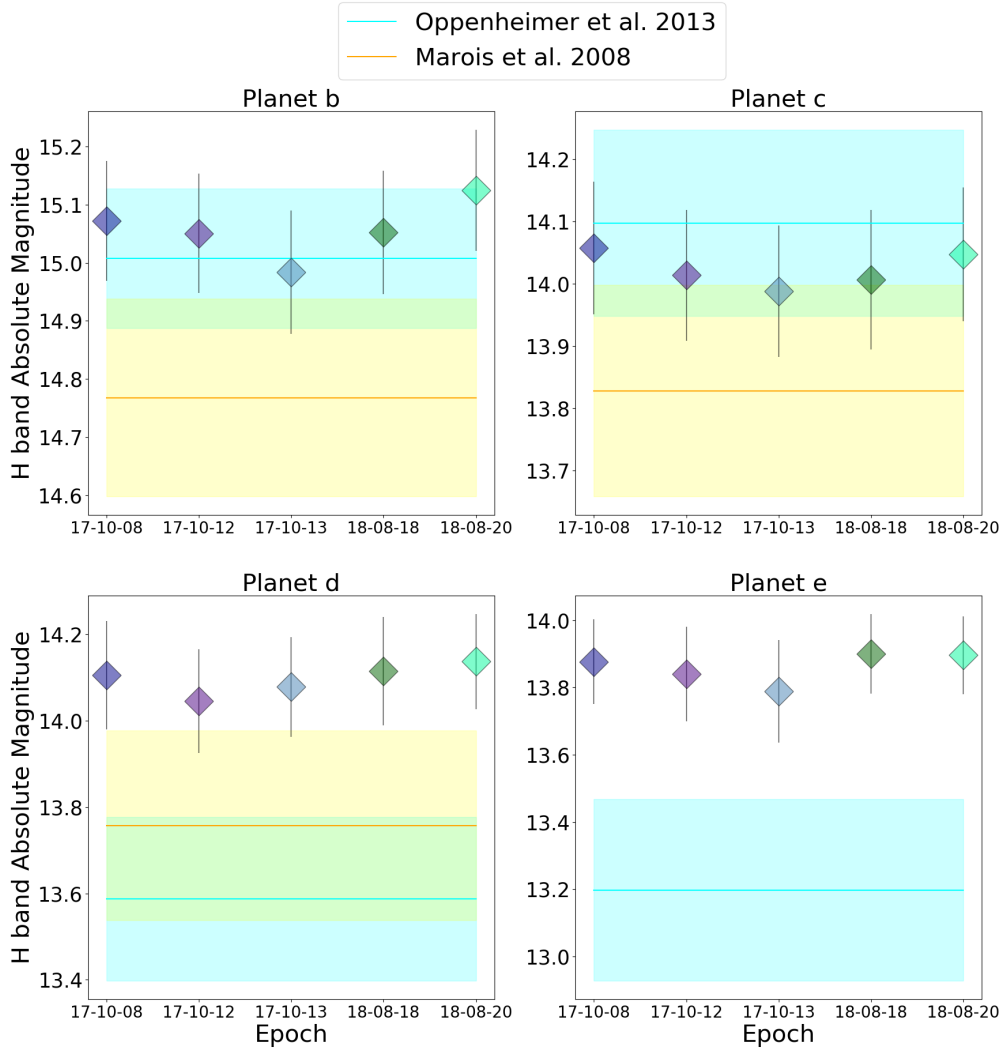
Figure 22. Our broadband-*H* epoch-by-epoch photometry compared to Palomar *H*-band photometry from Oppenheimer et al. (2013) and Keck *H*-band photometry from Marois et al. (2008).

likely due to uncorrected speckle noise at the radius of HR 8799c in this temporal bin, and is recreated in simulated planet injection tests using the satellite spots as the noise model. The fact that HR8799bc generally show the same trends means that they are appropriate to use to detrend each other, to search for non-shared variations.

(vi) Satellite spot lightcurves share the same trends as the planet lightcurves in the August 2018 epochs, however, the planets diverge significantly from the lightcurve trends in the October 2017 epochs. The divergence may be due to the much redder colors of the planets relative to the satellite spots (which should share the spectrum of the star). This means that the satellite spot lightcurves are only appropriate to detrend planet lightcurves for the August 2018 epochs, not the October 2017 epochs.

(vii) We found that the number of PCA modes removed had no real effect on our ability to retrieve photometry for HR 8799b, but HR 8799c is significantly affected by the number of modes subtracted, especially in the first and last temporal bin (also when conditions were worst / airmasses were highest). For HR 8799c, removing 16 modes

oversubtracts light from the exoplanet, but with 4 modes removed, obvious speckle noise remains in the residual images, thus we chose to subtract 10 modes as the best compromise.

(viii) Considering results for HR 8799de using 3 temporal bins, these planets appear to follow satellite spot trends to within error bars, but too much speckle noise remains at these radii to meaningfully constrain variability to high precision.

(ix) Since the satellite spot lightcurves are not appropriate to detrend planet lightcurves at all epochs, we consider $\Delta(mag)_b - \Delta(mag)_c$ as a function of time during each observation – essentially detrending the b lightcurve using that of c, or vice versa. This traces non-shared variations between the two planets. We used five temporal bins, to assure 20 pixels of movement on the sky for HR 8799b and c, with difference lightcurves presented in Fig. 15. Within errors, our results are in agreement with those from Apai et al. (2016). The uncertainty on the measurement is dominated by the uncertainty in our measurement of $\Delta(mag)_c$. Depending on epoch, we rule out
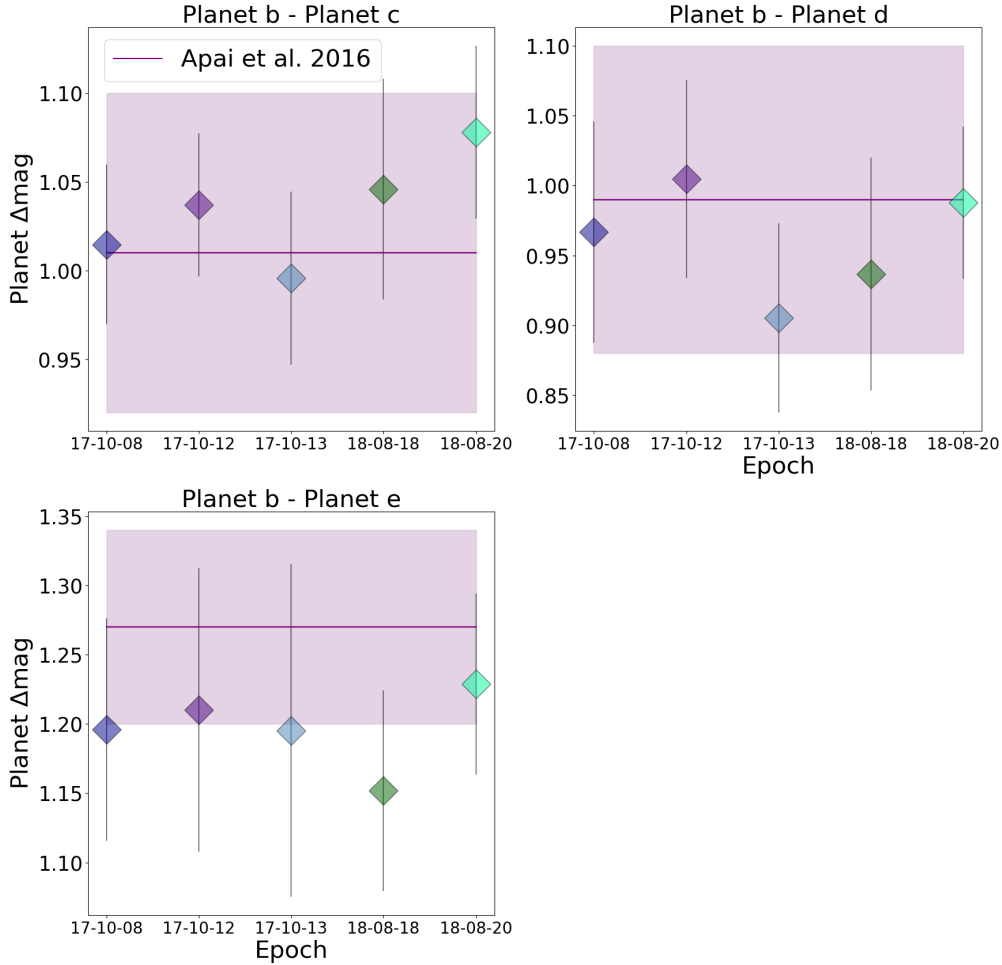
Figure 23. Our broadband-H epoch-by-epoch photometry compared to Apai et al. (2016).

non-shared variability in $\Delta(mag)_b - \Delta(mag)_c$ to the 10-20% level over 4-5 hours.

(x) To fully quantify our sensitivity to variability in the 2018-08-18 epoch (the best quality data of our search), we simulate variable and non-variable lightcurves by inserting and retrieving a suite of simulated planets at a similar radius from the star as HR 8799bc, but offset in position angle from the true planet position angle. Simulated planets are assumed to have sinusoidal lightcurves, with noise estimated using the mean satellite spot lightcurve. We ran a complete grid of simulated planet lightcurves with periods from 2 to 20 hours, amplitudes from 3% to 30%, and phases from 0° to 330°. For HR 8799b, for periods <10 hours, we are generally quite sensitive to variability with amplitude > 5%. For HR 8799c, our sensitivity is much lower, limited to variability > 25% for similar periods. Taking into account that these planets are likely observed nearly pole-on (inclination <27°) and correcting for geometric dilution, these measured limits on variability amplitudes correspond to sensitivity to **intrinsic** variability amplitudes of 11-22% for HR 8799b, and >55% for HR 8799c for periods of <10 hours, in the broadband H filter.

(xi) We derive epoch-by-epoch astrometry and photometry, and compare to literature values, using the data frames from the hour approaching meridian crossing and the hour after meridian crossing, thus covering the time period during each observation with the most field rotation and best airmass. We again ran our full negative simulated planet code to retrieve the best Δmag, radius, and position angle value for each planet. Our epoch-by-epoch photometry agrees within uncertainties to literature values which employ similar negative simulated planet subtraction techniques. Unfortunately, the uncertainties inherent in converting from relative photometry within an epoch to Δ(mag) for an entire epoch (i.e. variability of the star, error on the brightness of the star) prevent an epoch-by-epoch search for variability.

While 8-m telescopes with extreme AO systems have the sensitivity to detect HR 8799bc with high S/N in a reasonably small amount of observing time, for the inner planets, the varying speckle noise floor effectively precluded variability studies at the required accuracy for a plausible detection. This changing speckle background is driven
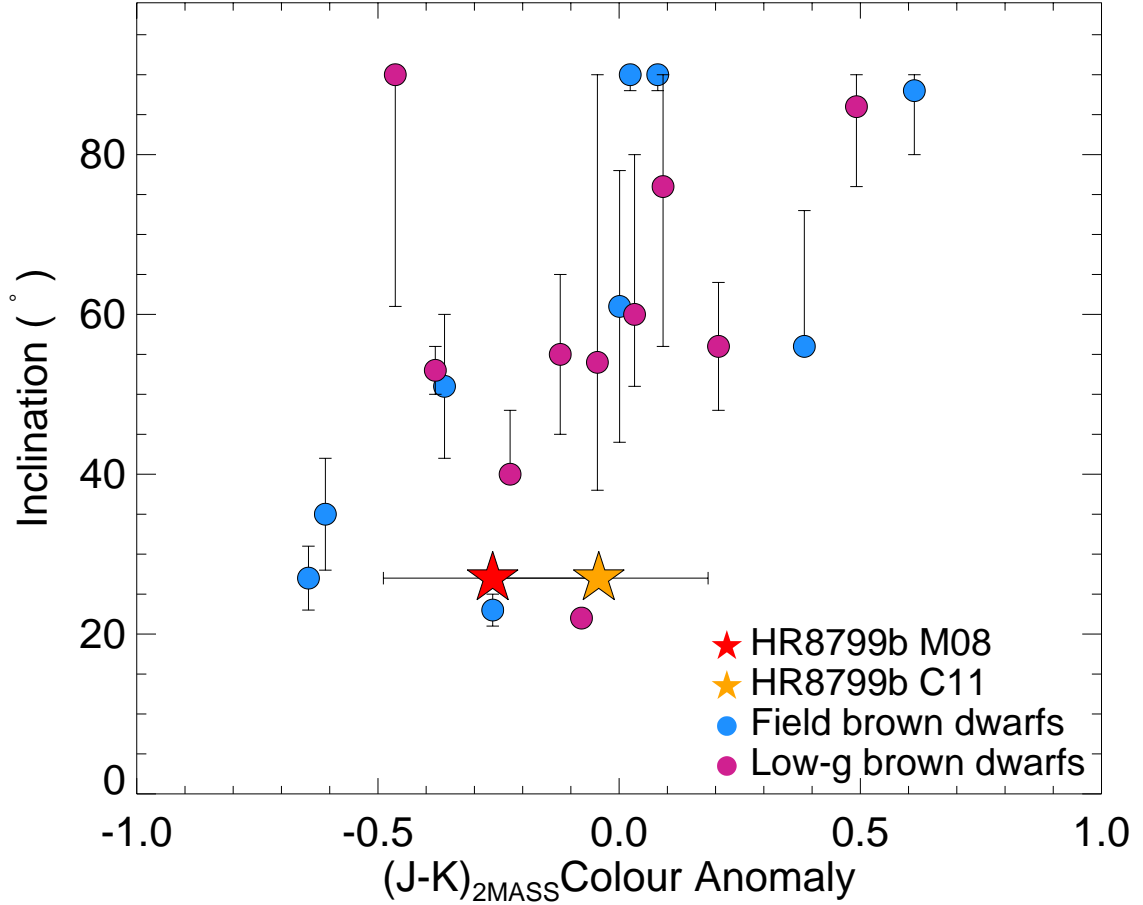
Figure 24. Inclination vs. colour anomaly for HR 8799b as well as field and low surface gravity brown dwarfs from Vos et al. (2020). HR 8799b has been plotted assuming it shares the same 27° inclination as the HR 8799 disk(Matthews et al. 2014; Ruffio et al. 2019; Wang et al. 2018a; Konopacky et al. 2016; Pueyo et al. 2015; Maire et al. 2015), however, the inclination of this planet has not been directly measured. Adopting a spectral type of L7, we calculate the colour anomaly (object $J − K$ - average $J − K$ for other objects with the same spectral type and gravity class) for HR 8799b using photometry from Marois et al. (2008); Currie et al. (2011) converted into the 2MASS filter set using the conversion from Stephens & Leggett (2004). The colour anomaly for HR 8799b is consistent with a wide range of possible inclinations for this planet, including a nearly pole-on 27° inclination.

by the fundamental instability of working from the ground; we must contend with light that has already traversed the atmosphere of the Earth and is only imperfectly corrected by our adaptive optics system. The limits of ground-based studies are also apparent in variability studies of field brown dwarfs and isolated planetary mass objects; ground-based variability monitoring is only sensitive to variations of >2-3% (Radigan et al. 2014; Vos et al. 2019), whereas space-based monitoring with telescopes such as HST and Spitzer enable sensitivity down to variations of >0.1-0.5% (Metchev et al. 2015; Apai et al. 2013). However, existing space platforms such as HST do not have coronagraphs that can attain the contrasts or small inner working angles necessary for similar studies of close-companions such as the HR 8799 planets. The upcoming James Webb Space Telescope (JWST) combines both the stability of a space platform with high-contrast imaging capability from a suite of coronagraphs and will enable much higher sensitivity to variability in these planets than we were able to achieve even in excellent conditions with VLT-SPHERE; thus, in the next few years, JWST will likely transform

our understanding of the variability of young, directly-imaged giant exoplanet atmospheres.

REFERENCES

Allers K. N., Gallimore J. F., Liu M. C., Dupuy T. J., 2016, Astrophys. J, 819, 133

Amara A., Quanz S. P., 2012, MNRAS, 427, 948

Antoci V., et al., 2019, MNRAS, 490, 4040

Apai D., Radigan J., Buenzli E., Burrows A., Reid I. N., Jayawardhana R., 2013, Astrophys. J, 768, 121

Apai D., et al., 2016, Astrophys. J, 820, 40

Astropy Collaboration et al., 2013, A&A, 558, A33

Astropy Collaboration et al., 2018, aj, 156, 123

Barman T. S., Konopacky Q. M., Macintosh B., Marois C., 2015, ApJ, 804, 61

Bell C. P. M., Mamajek E. E., Naylor T., 2015, MNRAS, 454, 593

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

Biller B. A., et al., 2015, ApJ, 813, L23

Biller B. A., et al., 2018, AJ, 155, 95

Bonnefoy M., et al., 2014, Astron. & Astrophys., 567, L9

Bonnefoy M., et al., 2016, Astron. & Astrophys., 587, A58

Booth M., et al., 2016, MNRAS, 460, L10

Bowler B. P., 2016, PASP, 128, 102001

Bowler B. P., Zhou Y., Morley C. V., Kataria T., Bryan M. L., Benneke B., Batygin K., 2020, ApJ, 893, L30

Bradley L., et al., 2019, astropy/photutils: v0.6, doi:10.5281/zenodo.2533376, https://doi.org/10.5281/zenodo.2533376

Carbillet M., et al., 2011, Experimental Astronomy, 30, 39

Chauvin G., et al., 2017, A&A, 605, L9

Currie T., et al., 2011, ApJ, 729, 128

Currie T., et al., 2019, arXiv e-prints, p. arXiv:1909.10522

De Rosa R. J., et al., 2016, Astrophys. J, 824, 121

Delorme P., et al., 2017, in Reylé C., Di Matteo P., Herpin F., Lagadec E., Lançon A., Meliani Z., Royer F., eds, SF2A-2017: Proceedings of the Annual meeting of the French Society of Astronomy and Astrophysics. pp 347–361 (arXiv:1712.06948)

Dohlen K., et al., 2008, in Ground-based and Airborne Instrumentation for Astronomy II. p. 70143L, doi:10.1117/12.789786

Eriksson S. C., Janson M., Calissendorff P., 2019, A&A, 629, A145

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

Gauza B., Béjar V. J. S., Pérez-Garrido A., Zapatero Osorio M. R., Lodieu N., Rebolo R., Pallé E., Nowak G., 2015, ApJ, 804, 96

Gizis J. E., et al., 2015, Astrophys. J, 813, 104

Gomez Gonzalez C. A., et al., 2017, AJ, 154, 7

Gravity Collaboration et al., 2019, A&A, 623, L11

Greenbaum A. Z., et al., 2018, AJ, 155, 226

Guerri G., et al., 2011, Experimental Astronomy, 30, 59

Ingraham P., et al., 2014, Astrophys. J. Lett., 794, L15

Jovanovic N., et al., 2015, PASP, 127, 890

Konopacky Q. M., Barman T. S., Macintosh B. A., Marois C., 2013, Science, 339, 1398

Konopacky Q. M., Marois C., Macintosh B. A., Galicher R., Barman T. S., Metchev S. A., Zuckerman B., 2016, AJ, 152, 28

Lafrenière D., Marois C., Doyon R., Nadeau D., Artigau É., 2007, ApJ, 660, 770

Lagrange A.-M., et al., 2010, Science, 329, 57

Langlois M., Vigan A., Moutou C., Sauvage J.-F., Dohlen K., Costille A., Mouillet D., Le Mignant D., 2013, in Esposito S., Fini L., eds, Proceedings of the Third AO4ELT Conference. p. 63, doi:10.12839/AO4ELT3.13317

Lavie B., et al., 2017, AJ, 154, 91

Lew B. W. P., et al., 2016, Astrophys. J. Lett., 829, L32

Liu M. C., et al., 2013, Astrophys. J. Lett., 777, L20

Macintosh B., et al., 2014, Proceedings of the National Academy of Science, 111, 12661

Maire A. L., et al., 2015, A&A, 576, A133

Maire A.-L., et al., 2016, in Ground-based and Airborne Instrumentation for Astronomy VI. p. 990834 (arXiv:1609.06681), doi:10.1117/12.2233013

Marley M. S., Saumon D., Cushing M., Ackerman A. S., Fortney J. J., Freedman R., 2012, Astrophys. J, 754, 135

Marois C., Lafrenière D., Doyon R., Macintosh B., Nadeau D., 2006, ApJ, 641, 556

Marois C., Macintosh B., Barman T., Zuckerman B., Song I., Patience J., Lafrenière D., Doyon R., 2008, Science, 322, 1348

Marois C., Zuckerman B., Konopacky Q. M., Macintosh B., Barman T., 2010, Nature, 468, 1080

Matthews B., Kennedy G., Sibthorpe B., Booth M., Wyatt M., Broekhoven-Fiene H., Macintosh B., Marois C., 2014, ApJ, 780, 97

Metchev S. A., et al., 2015, Astrophys. J, 799, 154

Millar-Blanchaer M. A., et al., 2020, ApJ, 894, 42

Milli J., et al., 2018, in Adaptive Optics Systems VI. p. 107032A (arXiv:1806.05370), doi:10.1117/12.2311499

Oppenheimer B. R., et al., 2013, ApJ, 768, 24

Pavlov A., Möller-Nilsson O., Feldt M., Henning T., Beuzit J.-L., Mouillet D., 2008, in Advanced Software and Control for Astronomy II. p. 701939, doi:10.1117/12.789110

Pueyo L., et al., 2015, ApJ, 803, 31

Radigan J., 2014, Astrophys. J, 797, 120

Radigan J., Jayawardhana R., Lafrenière D., Artigau É., Marley M., Saumon D., 2012, Astrophys. J, 750, 105

Radigan J., Lafrenière D., Jayawardhana R., Artigau E., 2014, Astrophys. J, 793, 75

Rajan A., et al., 2015, Mon. Not. R. Astron. Soc., 448, 3775

Ruffio J.-B., et al., 2019, AJ, 158, 200

Sahoo A., Guyon O., Lozi J., Chilcote J., Jovanovic N., Brandt T., Groff T., Martinache F., 2020, AJ, 159, 250

Sauvage J.-F., et al., 2016, in Adaptive Optics Systems V. p. 990916, doi:10.1117/12.223243559

Sivaramakrishnan A., Oppenheimer B. R., 2006, ApJ, 647, 620

Skemer A. J., et al., 2012, Astrophys. J, 753, 14

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

Snellen I. A. G., Brandl B. R., de Kok R. J., Brogi M., Birkby J., Schwarz H., 2014, Nature, 509, 63

Sódor Á., et al., 2014, A&A, 568, A106

Soummer R., Pueyo L., Larkin J., 2012, ApJ, 755, L28

Stephens D. C., Leggett S. K., 2004, PASP, 116, 9

Su K. Y. L., et al., 2009, ApJ, 705, 314

Torres C. A. O., Quast G. R., Melo C. H. F., Sterzik M. F., 2008, Young Nearby Loose Associations. p. 757

Vigan A., et al., 2010, in Ground-based and Airborne Instrumentation for Astronomy III. p. 77352X, doi:10.1117/12.856701

---

[3] http://www.astropy.org

Vos J. M., Allers K. N., Biller B. A., 2017, ApJ, 842, 78

Vos J. M., Allers K. N., Biller B. A., Liu M. C., Dupuy T. J., Gallimore J. F., Adenuga I. J., Best W. M. J., 2018, MNRAS, 474, 1041

Vos J. M., et al., 2019, MNRAS, 483, 480

Vos J. M., et al., 2020, AJ, 160, 38

Wang J. J., et al., 2014, in Ground-based and Airborne Instrumentation for Astronomy V. p. 914755 (arXiv:1407.2308), doi:10.1117/12.2055753

Wang J. J., et al., 2018a, AJ, 156, 192

Wang J., Mawet D., Fortney J. J., Hood C., Morley C. V., Benneke B., 2018b, AJ, 156, 272

Wilner D. J., MacGregor M. A., Andrews S. M., Hughes A. M., Matthews B., Su K., 2018, ApJ, 855, 56

Wilson P. A., Rajan A., Patience J., 2014, Astron. & Astrophys., 566, A111

Zapatero Osorio M. R., Martín E. L., Bouy H., Tata R., Deshpande R., Wainscoat R. J., 2006, Astrophys. J, 647, 1405

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

Zuckerman B., Rhee J. H., Song I., Bessell M. S., 2011, ApJ, 732, 61

Zurlo A., et al., 2016, Astron. & Astrophys., 587, A57

APPENDIX A: SATELLITE SPOT PHOTOMETRY

We present here satellite spot lightcurves for the epochs of 2017-10-07, 2017-10-12, 2017-10-13, and 2018-08-20 (Figs. A1 through A4). The satellite spot photometry for the August 2018 epochs is notably more stable and less variable than that for the October 2017 epochs; note the smaller plot range used for the August 2018 epochs.

APPENDIX B: EFFECT OF CONDITIONS ON SATELLITE SPOT PHOTOMETRY

To search for correlations in the satellite spot lightcurves as a function of ambient conditions, we overplot the median flux in the satellite spots, measured background flux, obtained Strehl ratio and seeing for the 2017-10-12 (average conditions) and 2018-08-18 (excellent conditions) epochs in Figs. B1 through B5. All satellite spot and background lightcurves have been normalized to themselves; the background levels are generally quite a bit lower than the flux in the satellite spots, but here we consider variations relative to the normalization of the component in question. As expected, background level is anti-correlated with the Strehl ratio; as Strehl ratio decreases, more starlight is lost from the core of the PSF and instead appears as a higher background level. Seeing is also roughly anti-correlated with Strehl ratio, commensurate with the expectation of poorer adaptive optics correction during poorer seeing conditions. Other than the 2017-10-08 epoch (when, as noted previously, the satellite spots were under-illuminated), the satellite spot light curves do not appear to be obviously anti-correlated with the background level. However, this may be because the AO residual background scales as (1 - Strehl ratio), whereas the satellite spots should scale directly with the Strehl ratio. Thus, for instance, a drop in Strehl ratio from 90% to 80% would have a factor of 2 effect on the AO residual background, but only a 10% effect on the satellite spot brightness.

APPENDIX C: SATELLITE SPOT CORRELATION PLOTS

We present additional satellite spot correlation plots. Correlation plots without the mean lightcurve removed are shown in Fig. C1 and C2. Plots with the mean lightcurve removed are shown in Figures C3 through C5. The goal of measuring satellite spot photometry is to eventually remove trends due to non-astrophysical variability (i.e. changes in AO correction, seeing, flexure) as well as astrophysical trends from the star. Thus, we measure two different types of trends in our satellite spot data: how much the satellite spots follow a monotonic function between each other (i.e. smooth trends in time due to changes in observing conditions) – and how much inherent spread is left in the satellite spots after removing common variations. Larger trends with time in the satellite spot data shared between spots follow major changes in Strehl, seeing, coherence time, etc – better data, such as those from August 2018 show smaller trends over the observation compared to data from October 2017, as conditions were inherently more stable. Plotting satellite spots against each other then mostly highlights the inherent scatter in the data, as for the best datasets, we covered only a small range of different star spot values. This produces a smaller Spearman ρ value, even though the data is of better quality.

On 2018-10-18, and 2018-08-20, we find the strongest anti-correlations between adjacent satellite spots 1-4 and 2-3, weaker anti-correlations between satellite spots 1-2 and 3-4, and no or slight positive correlation between the two opposite sets of satellite spots. On 2017-10-08, we find that adjacent satellite spots 1-4 and 2-3 show relatively strong anti-correlation (but not much correlation between other adjacent satellite spots). On 2017-10-12, we find anti-correlation between adjacent spots 1-2 and 3-4. On 2017-10-13, adjacent spots 1-2 and 3-4 are somewhat anti-correlated, as well as opposite spots 1-3 and 2-4. These anti-correlations are directly visible in the satellite spot lightcurves in most cases, especially in the August 2018 data. These results differ from Apai et al. (2016), who found a fairly constant pattern of no or weak anti-correlation between most of the satellite spots, with relatively strong anti-correlations between one set of directly opposite spots, and two sets of adjacent spots. However, given the small range of flux values covered in our observations, we do not consider these Spearman ρ values to be robust. Additionally, in this study we have considered observations taken only in a fairly narrow range of observing conditions – thus, we have not considered Spearman ρ for a sufficient range of different conditions to justify strong conclusions regarding satellite spot behavior. It would be interesting to consider correlation plots for all archival SPHERE-IRDIS sequences using satellite spots throughout the sequence, to search for overarching trends as a function of observing conditions.

This paper has been typeset from a TeX/LaTeX file prepared by the author.

Date (UTC)	MJD at start of observation	filter set	obs. length (hours)	field rotation (degrees)	seeing (arcsec)	airmass	median Strehl	DIT (s)	NEXP
2015-07-30	57233.2178	J2/J3	4.99	83.4	0.96	1.43 - 2.04	0.83	16	960
2015-07-31	57234.1999	K1/K2	5.96	90.9	1.30	1.43 - 2.54	0.69	16	1121
2017-10-08	58033.9927	BB-H	4.70	76.3	0.64	1.43 - 2.12	0.83	16	476
2017-10-12	58038.0299	BB-H	4.50	75.8	0.60	1.43 - 2.03	0.87	16	882
2017-10-13	58038.9936	BB-H	4.91	80.0	0.67	1.43 - 1.90	0.85	8	1856
2018-08-18	58348.1795	BB-H	4.50	73.9	0.65	1.43 - 2.02	0.89	16	896
2018-08-20	58350.1367	BB-H	4.39	72.6	0.35	1.43 - 1.99	0.90	4	896

Table 1. Log of IRDIS Observations

Epoch	Planet	ΔRA (mas)	ΔDec (mas)
17-10-08	HR8799b	1616.72±2.81	598.78±2.74
17-10-08	HR8799c	-414.39±2.72	852.59±3.01
17-10-08	HR8799d	-492.29±4.72	-461.70±4.61
17-10-08	HR8799e	-363.98±4.74	140.31±3.91
17-10-12	HR8799b	1615.61±2.73	597.93±2.77
17-10-12	HR8799c	-414.72±2.60	851.83±2.83
17-10-12	HR8799d	-493.82±3.92	-461.12±3.85
17-10-12	HR8799e	-363.87±6.79	139.06±4.73
17-10-13	HR8799b	1614.33±3.01	598.67±2.74
17-10-13	HR8799c	-415.03±2.69	852.03±2.94
17-10-13	HR8799d	-495.43±3.51	-463.32±3.45
17-10-13	HR8799e	-366.11±7.27	140.13±5.12
18-08-18	HR8799b	1619.67±3.05	582.29±2.67
18-08-18	HR8799c	-394.05±3.31	862.52±3.72
18-08-18	HR8799d	-515.20±4.53	-444.86±4.35
18-08-18	HR8799e	-351.91±4.29	172.60±3.51
18-08-20	HR8799b	1619.21±2.87	583.23±2.90
18-08-20	HR8799c	-394.53±2.79	863.73±3.15
18-08-20	HR8799d	-515.60±3.27	-443.16±3.15
18-08-20	HR8799e	-353.08±4.13	174.17±3.47

Table 2. Epoch-by-Epoch Astrometry

Epoch	Planet	$H_{abs}(mag)$
17-10-08	HR8799b	15.07±0.10
17-10-08	HR8799c	14.06±0.11
17-10-08	HR8799d	14.11±0.13
17-10-08	HR8799e	13.88±0.13
17-10-12	HR8799b	15.05±0.10
17-10-12	HR8799c	14.01±0.11
17-10-12	HR8799d	14.05±0.12
17-10-12	HR8799e	13.84±0.14
17-10-13	HR8799b	14.98±0.11
17-10-13	HR8799c	13.99±0.11
17-10-13	HR8799d	14.08±0.12
17-10-13	HR8799e	13.79±0.15
18-08-18	HR8799b	15.05±0.11
18-08-18	HR8799c	14.01±0.11
18-08-18	HR8799d	14.12±0.13
18-08-18	HR8799e	13.90±0.12
18-08-20	HR8799b	15.12±0.10
18-08-20	HR8799c	14.05±0.11
18-08-20	HR8799d	14.14±0.11
18-08-20	HR8799e	13.90±0.12

Table 3. Epoch-by-Epoch Photometry

Epoch	b -c (mag)	b - d (mag)	b - e (mag)
17-10-08	1.01±0.16	0.97± 0.18	1.20±0.18
17-10-12	1.04±0.16	1.00± 0.17	1.21±0.19
17-10-13	1.00±0.19	0.91± 0.20	1.20±0.22
18-08-18	1.05±0.19	0.94± 0.20	1.15±0.20
18-08-20	1.08±0.17	0.99± 0.18	1.23±0.18

Table 4. Epoch-by-Epoch subtracted Photometry
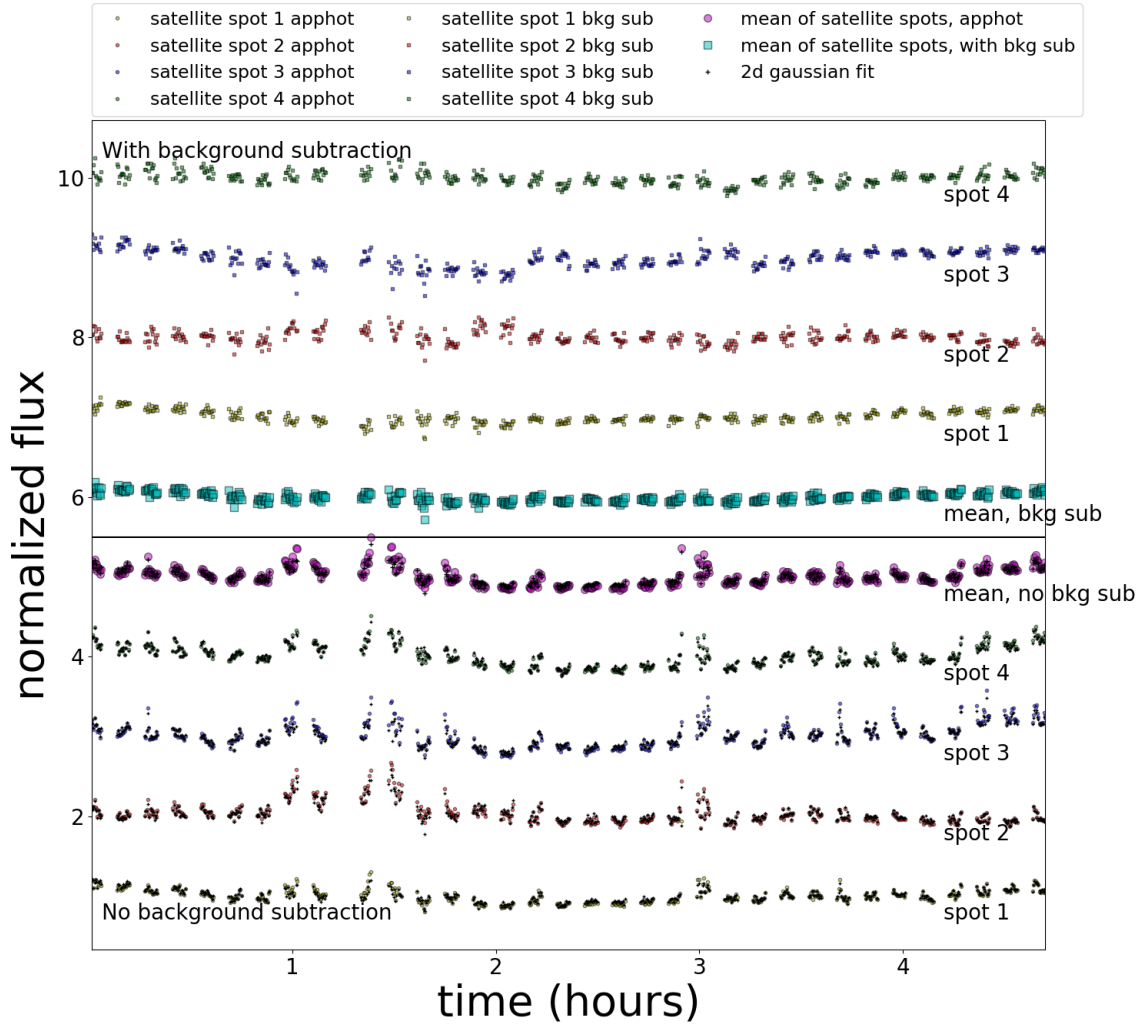
Figure A1. Satellite spot photometry for 2017-10-07 (normalized by the median value across the whole observation for each satellite spot) as a function of time for each satellite spot, with an offset in flux included between each satellite spot lightcurve for clarity. Results with background subtraction are shown on the top half of the plot, while results without background subtraction are plotted on the bottom half of the plot. The amplitudes from a 2-d gaussian fit to each satellite spot are plotted as black crosses, elliptical aperture photometry is plotted as filled circles, and elliptical aperture photometry with an additional background subtraction is plotted as filled squares. The mean of the elliptical aperture photometry for all 4 satellite spots are plotted in the middle of the figure, with magenta circles for the elliptical aperture photometry without background subtraction, and cyan square for elliptical aperture photometry with background subtraction.

Figure A2. Satellite spot photometry for 2017-10-12, plot axes and labels are the same as in Fig. A1.
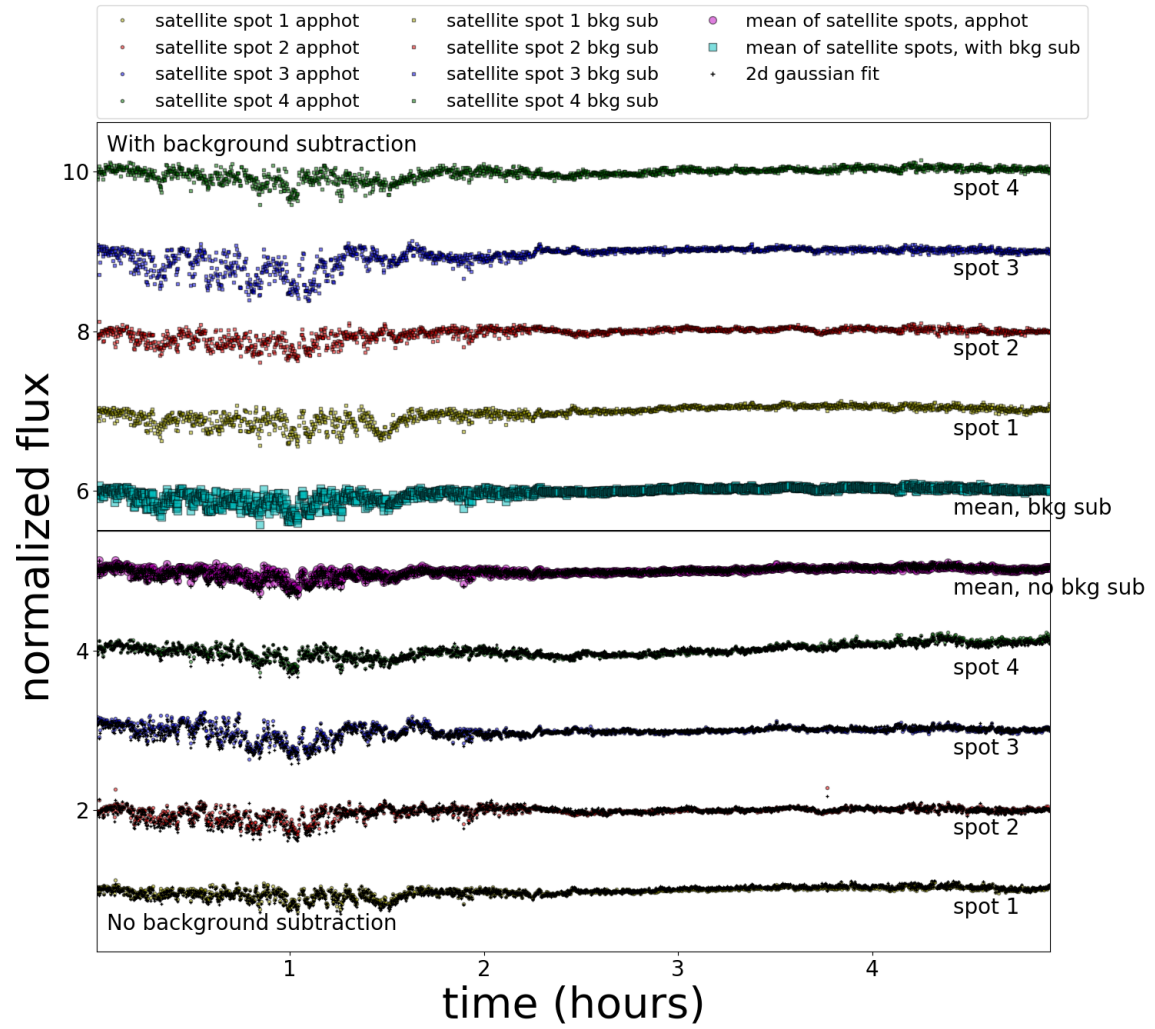
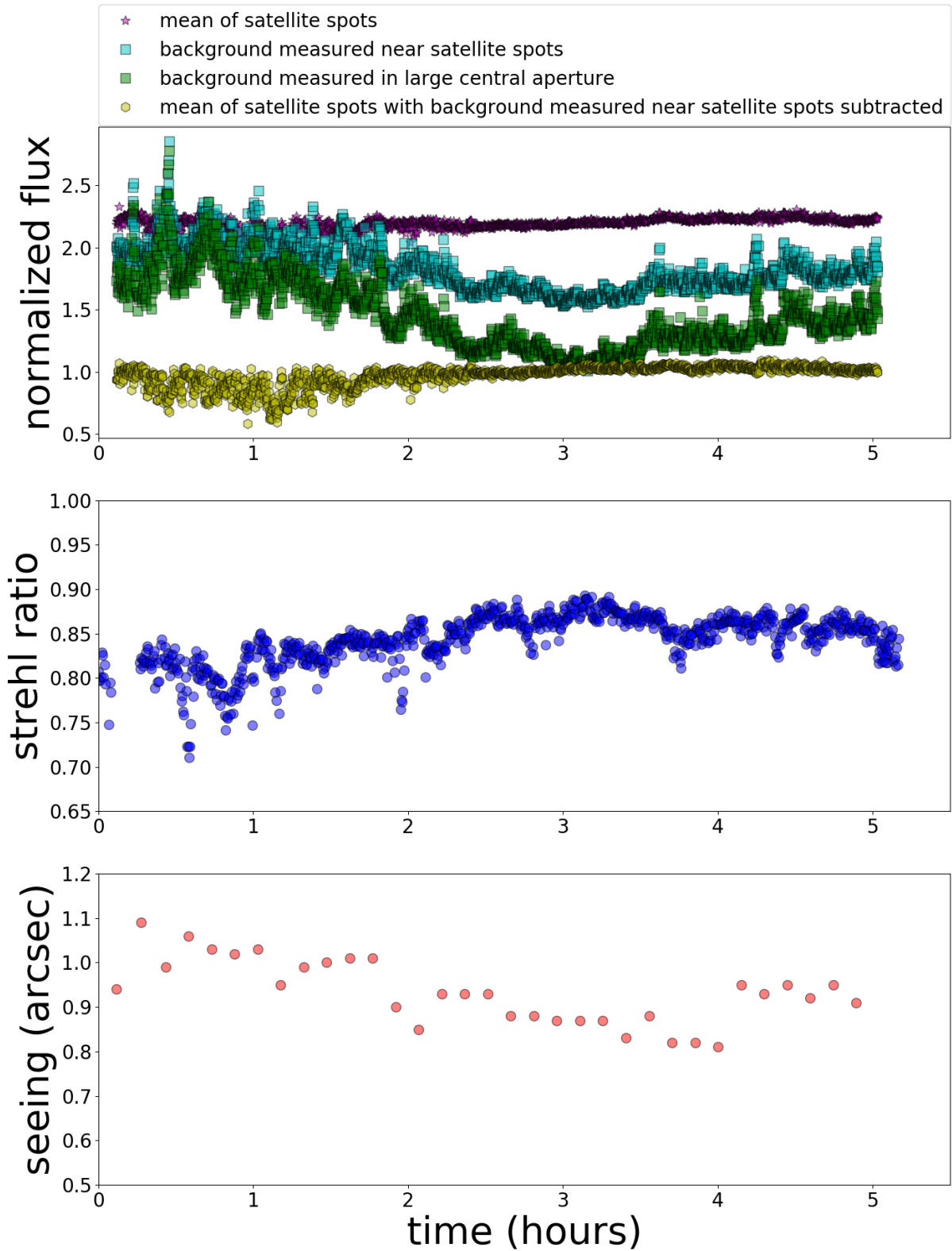**Figure A3.** Satellite spot photometry for 2017-10-13, plot axes and labels are the same as in Fig. A1.
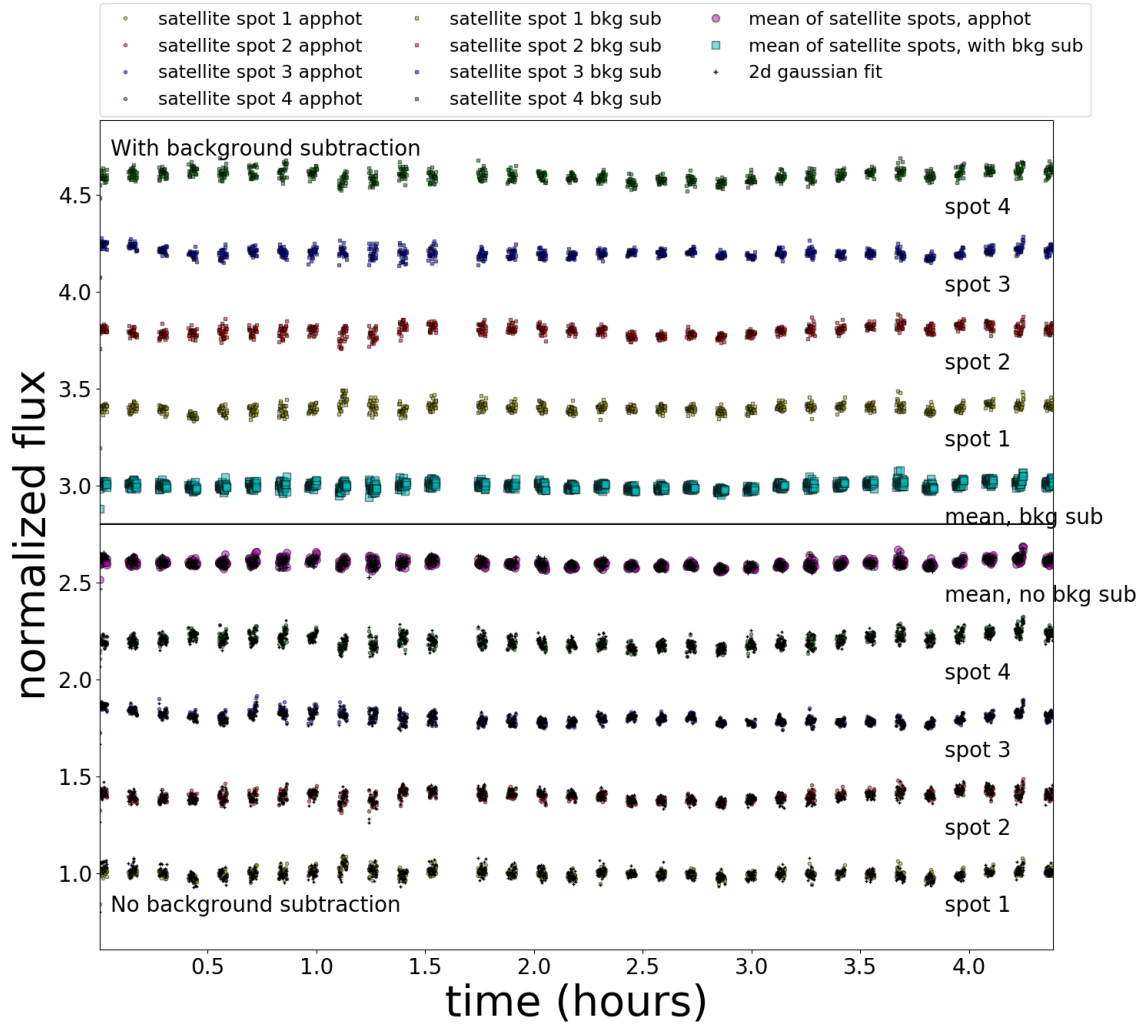
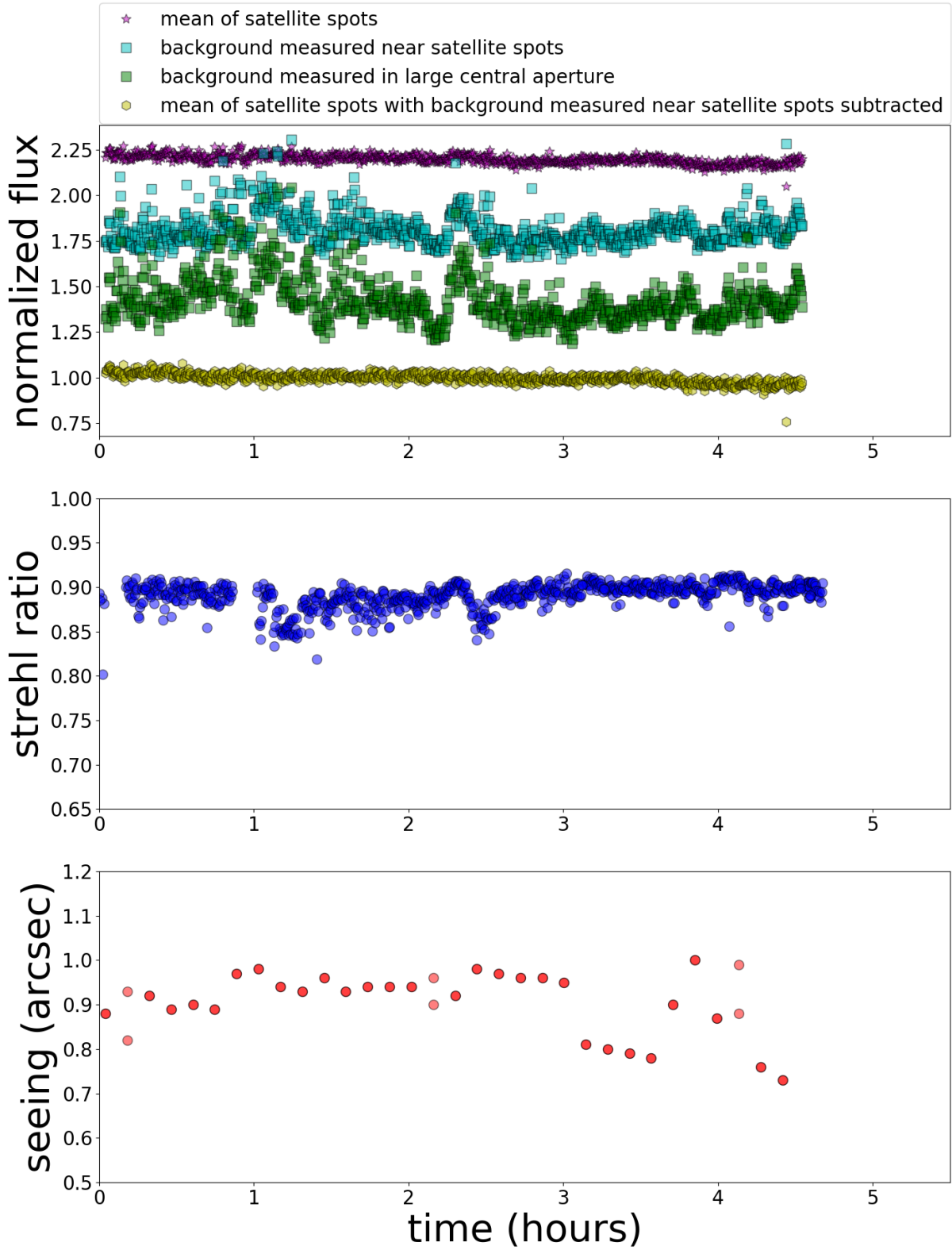**Figure A4.** Satellite spot photometry for 2018-08-20, plot axes and labels are the same as in Fig. A1.
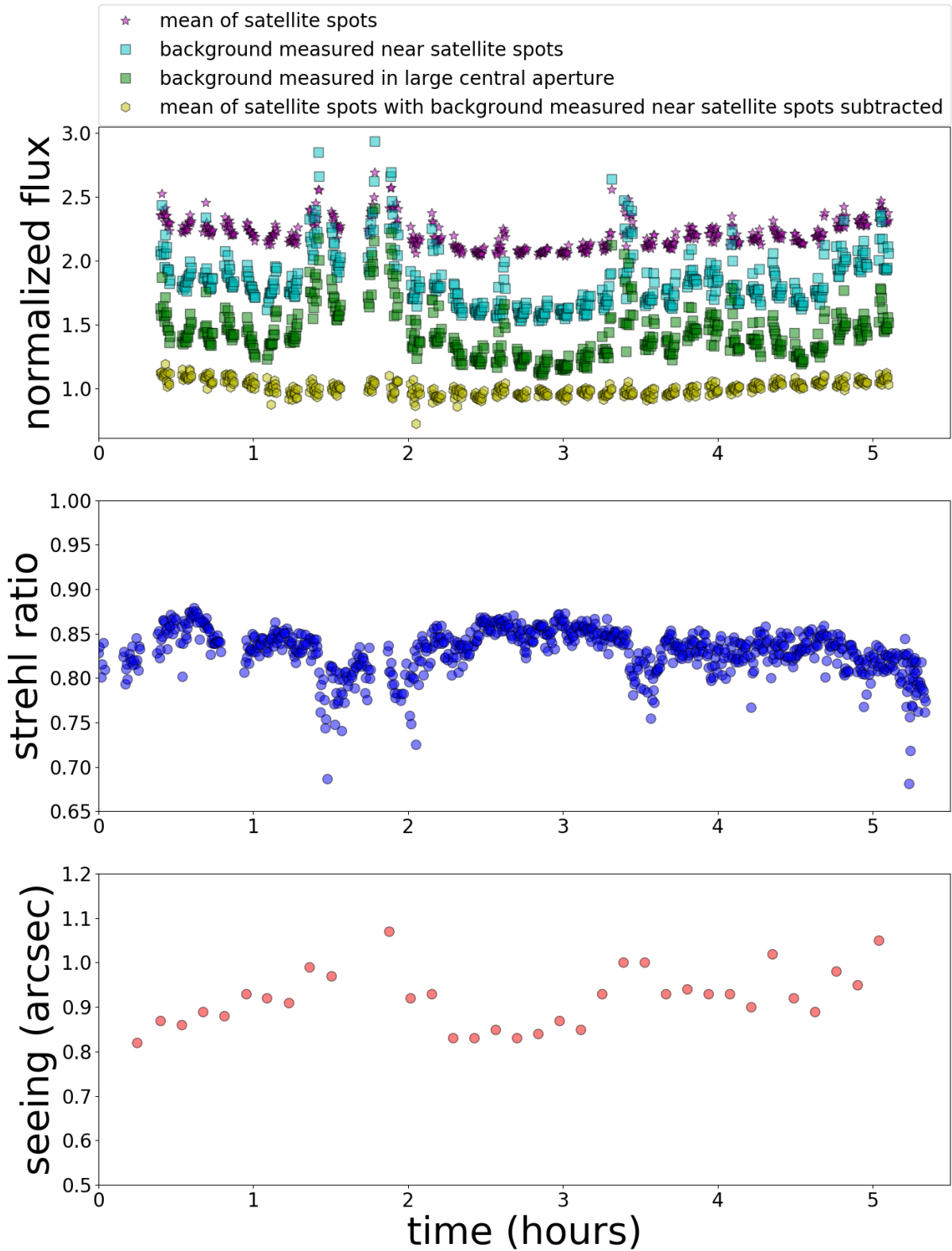
Figure B1. Mean satellite spot photometry, background photometry, strehl ratio, and seeing vs. time for 2017-10-08.
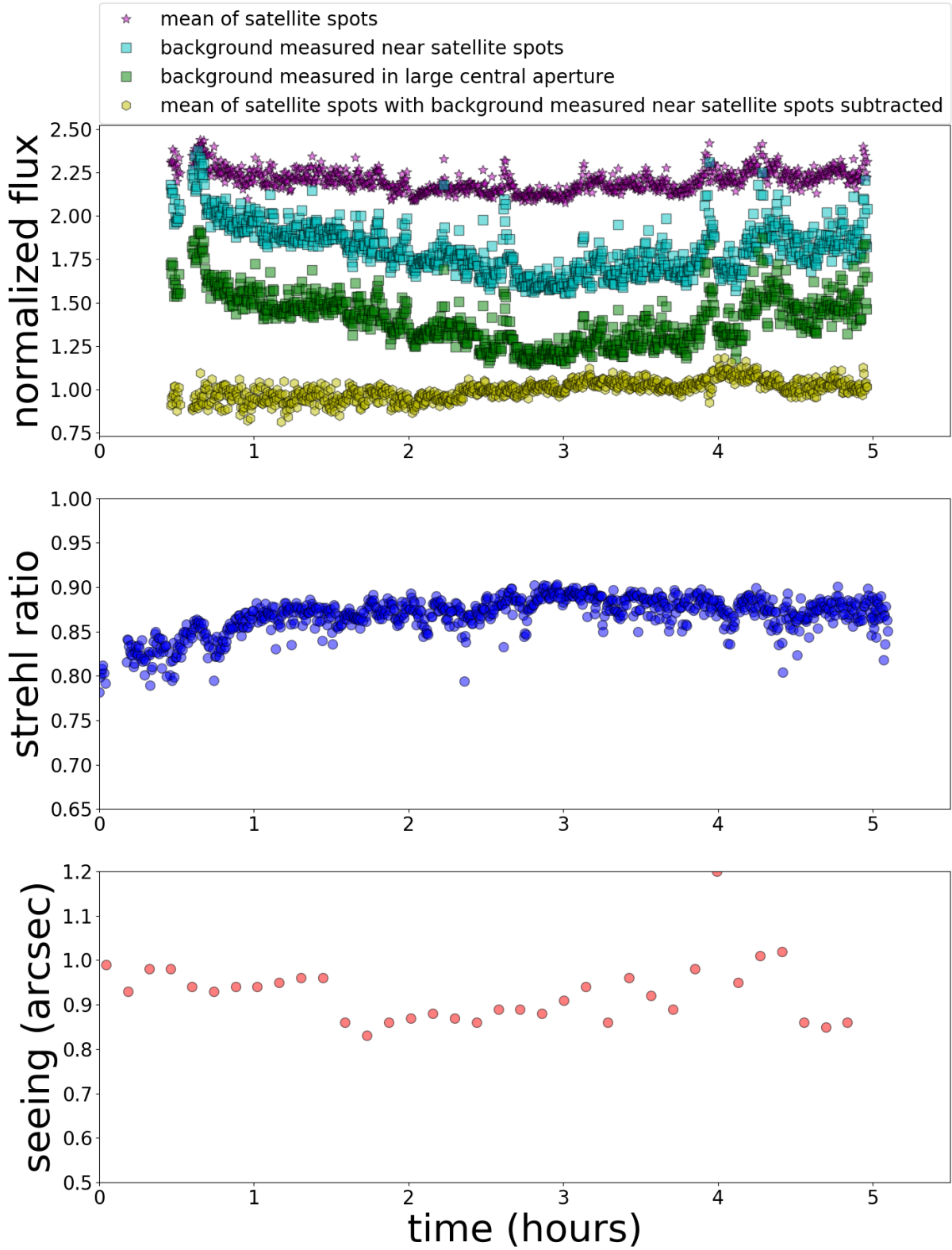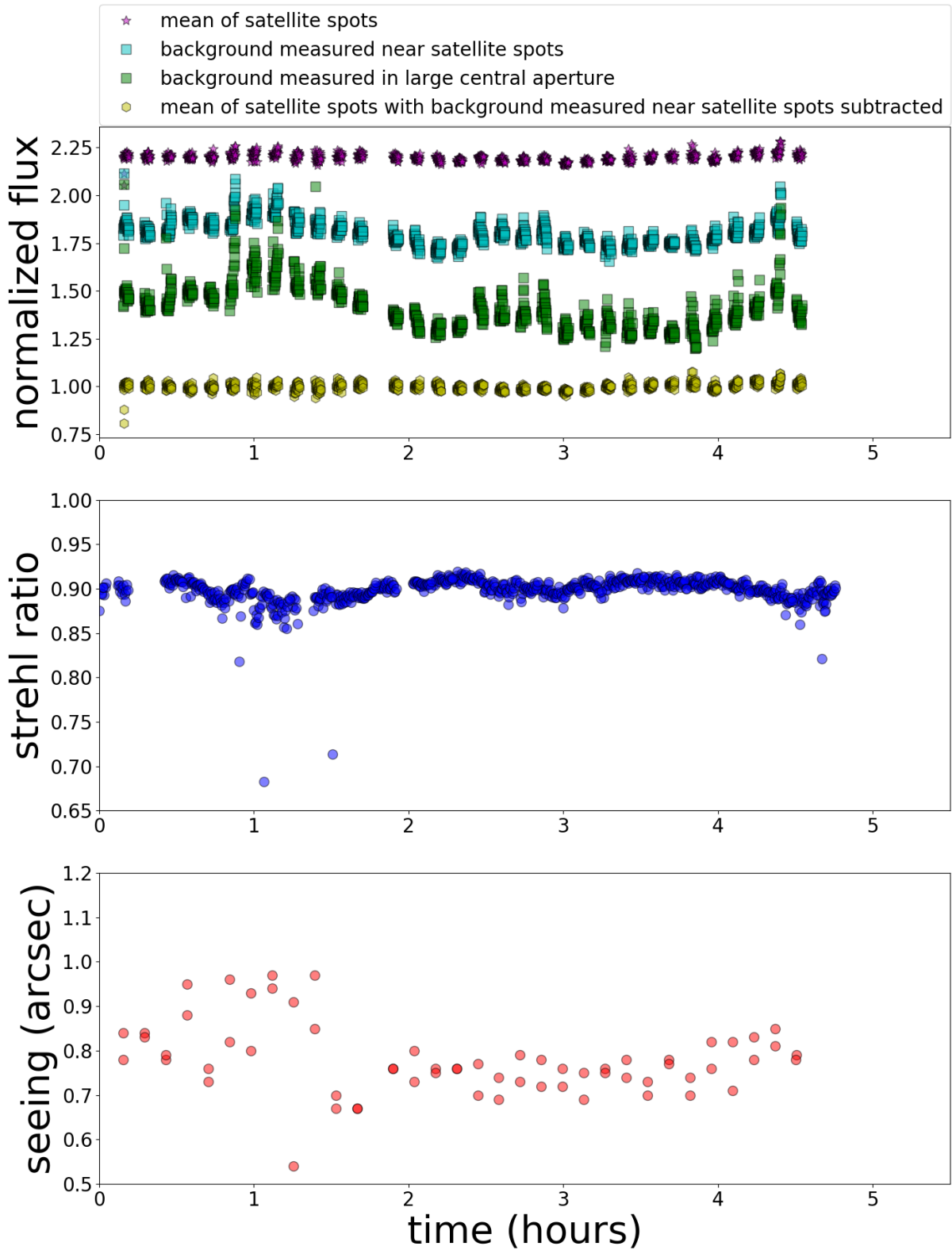
Figure B2. Mean satellite spot photometry, background photometry, strehl ratio, and seeing vs. time for 2017-10-12.
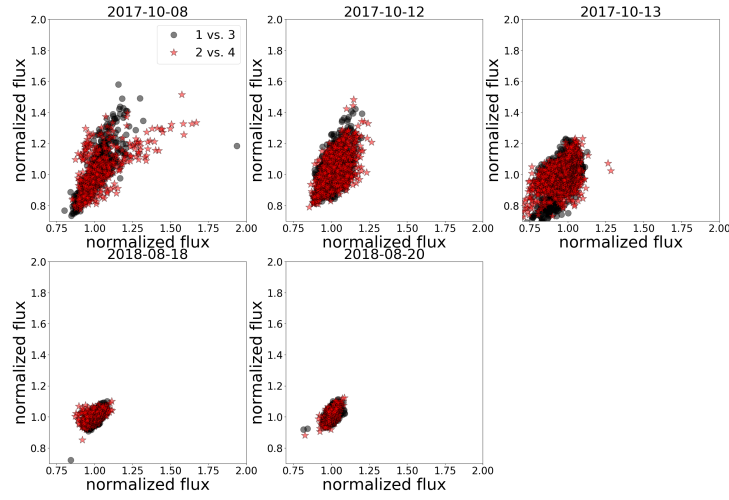
Figure B3. Mean satellite spot photometry, background photometry, strehl ratio, and seeing vs. time for 2017-10-13.

Figure B4. Mean satellite spot photometry, background photometry, strehl ratio, and seeing vs. time for 2018-08-18.

Figure B5. Mean satellite spot photometry, background photometry, strehl ratio, and seeing vs. time for 2018-08-20.

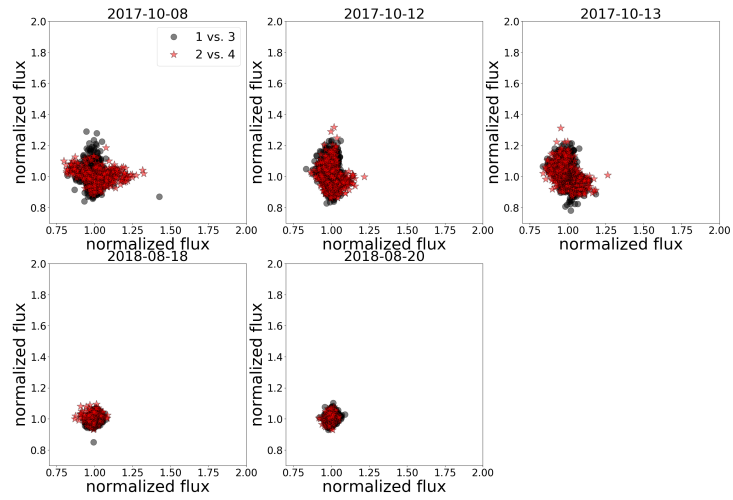**Figure C1.** Correlation plots for satellite spots across from each other.
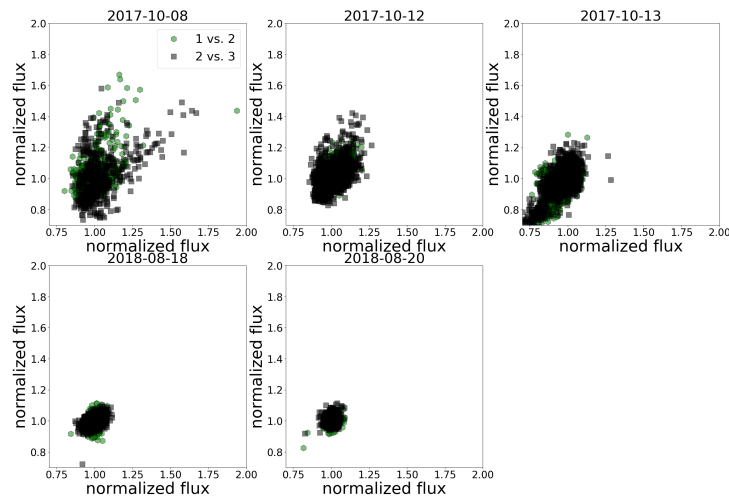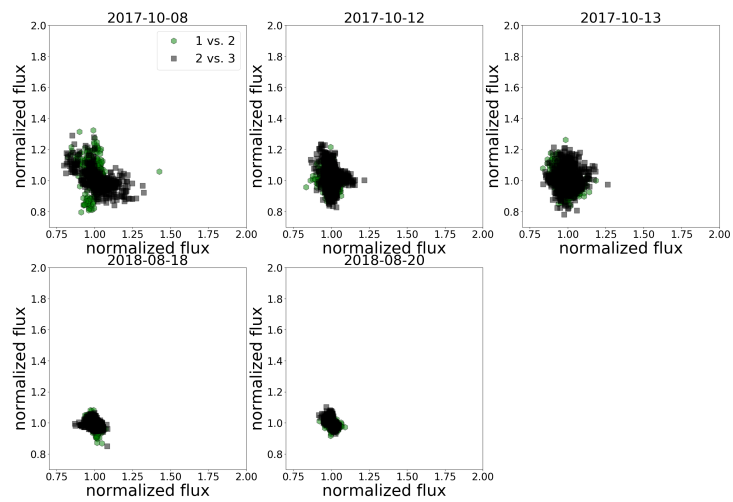

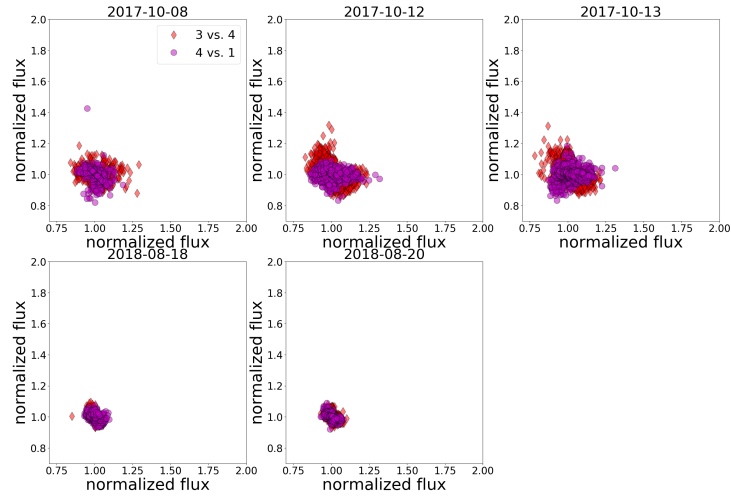**Figure C2.** Correlation plots for satellite spots on the sides.

Figure C3. Correlation plots for satellite spots across from each other, mean satellite spot lightcurve removed.

Figure C4. Correlation plots for satellite spots on the sides, mean satellite spot lightcurve removed.

Figure C5. Correlation plots for satellite spots on the sides, mean satellite spot lightcurve removed.