

# Disentangling auroral-, cloud-, and magnetic spot-driven variability in three early L dwarfs with HST/WFC3

C. O’Toole<sup>1,\*</sup>, J. M. Vos<sup>1</sup>, E. N. Nasedkin<sup>1</sup>, J. S. Pineda<sup>2</sup>, M. M. Kao<sup>3</sup>, Y. Zhou<sup>4</sup>,  
M. Schrader<sup>1</sup>, and A. M. McCarthy<sup>5,1</sup>

<sup>1</sup> School of Physics, Trinity College Dublin, The University of Dublin, Dublin 2, Ireland

<sup>2</sup> University of Colorado Boulder, Laboratory for Atmospheric and Space Physics, 3665 Discovery Drive, Boulder, CO 80303, USA

<sup>3</sup> Lowell Observatory, 1400 W Mars Hill Road, Flagstaff, AZ 86001, USA

<sup>4</sup> Department of Astronomy, University of Virginia, 530 McCormick Road, Charlottesville, VA 22904, USA

<sup>5</sup> Department of Astronomy & The Institute for Astrophysical Research, Boston University, 725 Commonwealth Avenue, Boston, MA 02215, USA

Received 3 December 2025 / Accepted 24 March 2026

## ABSTRACT

Variability monitoring provides unparalleled insights into the atmospheric processes of brown dwarfs and directly imaged exoplanets. Inhomogeneous clouds, aurorae, and magnetic spots have all been postulated as potential drivers of variability. While objects at the L/T transition have had their variability studied extensively, the variability of early L dwarfs remains an understudied region of the parameter space. We used observations from the Hubble Space Telescope in the near-infrared, using WFC3/G141 to disentangle the drivers of variability in three known variable early L dwarfs: 2MASS J1721039+334415, 2MASS J00361617+1821104, and 2MASS J19064801+4011089. We find that all three objects exhibit significant variability at all wavelengths, with white-light amplitudes of 0.53–1.41%. We find that their colour variations are brighter and bluer compared to later spectral types, except for 2MASS J19064801+4011089, which exhibits largely grey variations. We report a new period for 2MASS J1721039+334415 of  $4.9^{+0.4}_{-0.2}$  hours. We find evidence of long-term light curve stability in each object, which may indicate the presence of long-lived features on their surfaces. We created a flexible modelling framework to model three potential drivers of variability: clouds, aurorae, and magnetic spots. We fit our models to the spectral variability amplitude from 1.1–1.67  $\mu\text{m}$  of each object. We find that changing cloud properties or magnetic spots are the most likely drivers of variability in each object. Auroral models do not reproduce the variability within the HST wavelengths; however, future observations at longer wavelengths that probe higher in the atmosphere may be more sensitive to auroral effects. This work provides a foundation for future variability studies of early L dwarfs and directly imaged exoplanets to disentangle auroral, cloud, and magnetic spot driven variability.

**Key words.** Planets and satellites: atmospheres – brown dwarfs

## 1. Introduction

Brown dwarfs are often regarded as analogues to directly imaged giant exoplanets as they share many similar properties such as radii, effective temperature, thermal profiles, surface gravity, and chemistry (e.g. Faherty et al. 2016; Liu et al. 2016). The majority of brown dwarf and brown dwarf companion observations are not limited by the extreme star-to-planet contrasts of exoplanet observations, which enables more detailed analyses, particularly for time-resolved studies (e.g. Apai et al. 2013; Zhou et al. 2020a).

Variability monitoring allows the detailed exploration of extrasolar atmospheres and the dynamical processes that shape them via measurements of an object’s luminosity as it rotates. Variability studies with space-based facilities such as the Hubble Space Telescope (HST) (e.g. Apai et al. 2013; Yang et al. 2015; Bowler et al. 2020), the Spitzer Space Telescope (e.g. Metchev et al. 2015; Vos et al. 2018, 2022), and more recently the James Webb Space Telescope (JWST) (e.g. Biller et al. 2024; Chen et al. 2025; McCarthy et al. 2025; Nasedkin et al. 2025) have provided unparalleled insights into the dynamics of brown dwarf atmospheres.

Photometric variability monitoring across the optical, near-infrared (near-IR), and mid-infrared (mid-IR) has shown that dramatic variability is common among brown dwarfs (e.g. Radigan et al. 2014a; Metchev et al. 2015; Vos et al. 2022). This variability has been attributed to a number of different processes, from patchy cloud cover (e.g. Radigan et al. 2014a; Vos et al. 2023) to auroral processes (Hallinan et al. 2015) and magnetic spots (Gizis et al. 2015; Croll et al. 2016). However, disentangling the processes that drive variability requires spectroscopic monitoring.

Spectroscopic variability monitoring probes different pressures of the atmosphere simultaneously (e.g. Biller et al. 2024; McCarthy et al. 2025; Chen et al. 2025), and provides a method for probing the mechanisms that drive variability. HST has paved the way for these studies, allowing us to peer into the vertical structure of these atmospheres (Buenzli et al. 2012; Apai et al. 2013). Modelling how the variability changes with wavelength (the spectral variability amplitude) is key to understanding what processes drive the variability. Cloud-driven variability models have been the primary method used to reproduce the spectral variability amplitude for L and T dwarfs (e.g. Apai et al. 2013; Yang et al. 2015; Bowler et al. 2020). HST has shown that the spectral variability amplitude changes depending on the spectral

\* Corresponding author: cotool3@tcd.ie

type of the object. This is due to the fact that different spectral bands probe different atmospheric pressures. This is particularly evident within the water band feature: for L dwarfs, the spectral variability amplitude is characterised by high-altitude hazes and is consistent both in and out of the water feature while L/T transition objects have lower variability in the water feature. At the L/T transition, clouds are typically located at lower altitudes (higher pressures) compared to L dwarfs, resulting in lower variability in the water band (e.g. Yang et al. 2015; Lew et al. 2020a; Bowler et al. 2020). Planetary scale waves have also been proposed as a mechanism of cloud driven variability on brown dwarfs (e.g. Apai et al. 2017; Zhou et al. 2022; Fuda et al. 2024), and have been successful at constraining the driver of variability for observations spanning hundreds of rotations (Fuda et al. 2024).

The variability in early L dwarfs, <L4, has not been studied as extensively as their cooler counterparts at the L/T transition. Early L dwarfs can provide an understanding of how atmospheres transition across the stellar–substellar boundary, as they are at the M/L dwarf transition, where there is an overlap between the coolest M dwarfs and hottest L dwarfs. Variability in this regime may be caused by a number of mechanisms, such as clouds, aurorae, or magnetic spots, (e.g. Apai et al. 2013; Hallinan et al. 2015; Barnes et al. 2015; Kao et al. 2016; Zuckerman et al. 2026). While cloud driven variability has been extensively investigated with HST, (e.g. Apai et al. 2013; Yang et al. 2015; Bowler et al. 2020), auroral and magnetic spot driven variability have not.

Variability driven by the presence of magnetic spots has been hypothesised for early L dwarfs, based on observations with ground-based telescopes and the Kepler Space Telescope (Lane et al. 2007; Gizis et al. 2013; Croll et al. 2016). Magnetic spot driven variability has been observed on M dwarfs (e.g. Rockenfeller et al. 2006; Lane et al. 2007; McLean et al. 2011; Rackham et al. 2018). As early L dwarfs straddle the stellar–substellar boundary, it is plausible that they may also host magnetic spots that can drive variability. Late M dwarfs (M7–M9) commonly exhibit magnetic spot driven variability (e.g. Lane et al. 2007; McLean et al. 2011; Rackham et al. 2018).

Paudel et al. (2020) detected white light flares on nine L dwarfs, spanning spectral types from L0–L5, using observations from the K2 mission. As a result of the K2 observations spanning 4402 days, they had a long baseline to observe flaring activity. Based on results from Baylor et al. (2011); De Luca et al. (2020), and Paudel et al. (2020) found that flares in a range of stars, from spectral type G to L, share a number of properties despite the differences in their internal structures. They estimate that a  $10^{33}$  erg superflare occurs on early and mid-L dwarfs approximately once every 2.4 years. This low flaring frequency highlights why so few L dwarf flares are found in the literature as typical observing campaigns of brown dwarfs with facilities such as JWST, HST, and Spitzer are of the order of 10s–100s hours, and thus are highly unlikely to observe such a flare. Magnetic spots that produce these flares may therefore be ubiquitous in the early to mid-L dwarf regime, and may drive variability on these objects. Therefore, early L dwarfs with strong magnetic fields are among the coolest objects where magnetic spot driven variability may occur. However, magnetic spot driven variability of brown dwarfs have yet to be explored in the context of HST observations.

Thermochemical instabilities driven by diabatic processes have also been proposed as a potential driver of variability, (Tremblin et al. 2019, 2020; Oliveros-Gomez et al. 2026). These mechanisms are typically associated with the L/T transition

where CO is converted into CH<sub>4</sub> (Tremblin et al. 2019, 2020). Therefore, we can rule this method out as it is dependent on the presence of CH<sub>4</sub>, which is not expected at these spectral types.

Using spectroscopic variability observations of three early L dwarfs obtained with HST/WFC3, we seek to disentangle their drivers of variability by analysing their spectral variability amplitude. This study is structured as follows. In Section 2, we outline the properties of the sample. In Section 3, we describe the HST observations and data reduction, including the light curve extraction, Gaussian process modelling, and periodogram analysis. We illustrate the colour modulation of each object in Section 4, and highlight these objects as a complementary parameter space to previously studied brown dwarfs. In Section 5, we present the wavelength dependence of the variability amplitude for each object and we outline our modelling framework for multiple drivers of variability. We then discuss the results of our modelling in Section 6, evaluate what the most likely drivers of variability are in our objects, and compare them to their light curves in the literature. The conclusions are summarised in Section 7.

## 2. Sample properties

We chose our three L dwarf targets to characterise and test near-IR signatures of clouds, aurorae, and magnetic fields. 2MASS J00361617+1821104 (hereafter 2M0036+18) was selected as a known auroral object likely to possess strong magnetic fields, which we contrast with 2MASS J1721039+334415 (hereafter 2M1721+33), which has weak magnetic activity and known cloud behaviours. Against these two benchmarks we compare data for our third object, 2MASS J19064801+4011089 (hereafter 2M1906+40) with observed variability over 2200 rotations, which may be due to long-lived surface features. We proceed with the characteristics of individual objects as follows.

2M0036+18, is an aurorally emitting L3.5 dwarf. We show its fundamental parameters in Table 1. It emits coherent radio emission characteristic of aurora, including consistent periodic radio pulses that are highly circularly polarised with high brightness temperatures,  $P_{Radio} = 3.08 \pm 0.05$  hours (Hallinan et al. 2008). Its radio aurora has also been consistent over >10 years (Berger 2002; Hallinan et al. 2008). 2M0036+18 has shown photometric variability in the optical, near-IR, and mid-IR (Harding et al. 2013; Metchev et al. 2015; Croll et al. 2016). This makes it the ideal candidate to search for aurorally driven variability in the infrared. 2M0036+18 has an infrared photometric period of  $3.080 \pm 0.001$  hours Croll et al. (2016) and has previously had near-IR variability monitoring with the Spitzer Space Telescope Metchev et al. (2015) at 3.6 and 4.5  $\mu\text{m}$ .

2M1721+33, is an L3 dwarf whose variability is thought to be driven by clouds Metchev et al. (2015). Its fundamental parameters are shown in Table 1. It is magnetically inactive and thus unlikely to host aurorae or magnetic spots (Schmidt et al. 2015; Richey-Yowell et al. 2020). 2M1721+33 is an ideal candidate to act as a control for cloud-driven variability, against which we can contrast the variability of our other targets. This is strengthened by the fact that 2M1721+33 has a silicate feature at 10  $\mu\text{m}$ , indicative of a cloudy atmosphere (Suárez & Metchev 2022). 2M1721+33 had mid-IR variability monitoring with the Spitzer Space Telescope at 3.6 and 4.5  $\mu\text{m}$ , with a measured infrared photometric period of  $2.6 \pm 0.1$  hours (Metchev et al. 2015).

2M1906+40 is an L1 dwarf whose variability has been attributed to long-lived surface features such as magnetic spots or an inhomogeneous cloud deck (Gizis et al. 2013, 2015). Its fundamental parameters are shown in Table 1.

**Table 1.** Fundamental parameters and measured variability parameters for each object in our sample.

Object	2M1721+33	Ref	2M0036+18	Ref	2M1906+40	Ref
<b>Spectral type</b>	L3	(K03)	L3.5	(R00)	L1	(G11)
<b><math>T_{\text{eff}}</math> (K)</b>	$1802 \pm 174$	(S23)	$1869 \pm 64$	(F15)	$2201 \pm 202$	(S23)
<b>Log (g)</b>	$5.15 \pm 0.43$	(S23)	$5.21 \pm 0.17$	(F15)	$5.23 \pm 0.39$	(S23)
<b>Literature period (h)</b>	$2.6 \pm 0.1$	(M15)	$3.080 \pm 0.001$	(C16)	$8.88364 \pm 5 \times 10^{-5}$	(G15)
<b>GP period (h)</b>	$4.93^{+0.36}_{-0.22}$	(TW)	$3.13^{+0.09}_{-0.08}$	(TW)	$19.4^{+46.5}_{-10.5}$	(TW)
<b>Light curve amplitude (%)</b>	$0.64^{+0.03}_{-0.02}$	(TW)	$1.41^{+0.16}_{-0.02}$ (epoch 1)	(TW)	$0.53^{+0.03}_{-0.01}$	(TW)
			$1.36^{+0.11}_{-0.01}$ (epoch 2)	(TW)		

**Notes.** K03 = Cruz et al. (2003), R00 = Reid et al. (2000), G11 = Gizis et al. (2011), M15 = Metchev et al. (2015), F15 = Filippazzo et al. (2015), G15 = Gizis et al. (2015), C16 = Croll et al. (2016), S23 = Sanghi et al. (2023), TW = this work.

**Table 2.** Observation details for the three targets, from HST GO 15924.

Object	Orbits	Observation dates (UTC)	Observation times (UTC)	Total exposures	Time per exposure (s)
2M1906+40	6	07–08 December 2020	21:37–06:19	258	63
2M1721+33	5	16–17 October 2020	23:56–07:02	215	63
2M0036+18	5	26 July 2020	12:43–19:50	215	63
2M0036+18	5	07 December 2021	00:33–07:40	215	63

Long-term observations with the Kepler Space Telescope revealed that its light curve remained stable over 2200 rotations (Gizis et al. 2015). This light curve stability is unusual compared to light curve evolution driven by clouds on rotational timescales, (Karalidi et al. 2016; Apai et al. 2017; Vos et al. 2018). Gizis et al. (2013) suggest that the variability may be due to magnetic spots, but Gizis et al. (2015) favour a stable cloud spot model to reproduce the observed variability. However, neither auroral driven nor magnetic spot driven variability has been ruled out. The stable nature of 2M1906+40’s light curve suggests that there is a long-lived surface feature driving its variability. Gizis et al. (2015) present mid-IR monitoring at 3.6 and 4.5  $\mu\text{m}$  with the Spitzer Space Telescope. (Gizis et al. 2013, 2015) measure an optical photometric period of  $8.88364 \pm 5 \times 10^{-5}$  hours using its Kepler light curve. By analysing our sample as a whole, we can place this object in context by comparing it with others that show evidence for cloud-driven variability and auroral activity.

### 3. Observations and data reduction

#### 3.1. Observations

Each target was observed simultaneously with the HST/WFC3 instrument and the Very Large Array (VLA), (PID #15924, PI: Vos). This provided multi-wavelength coverage across both the near-IR and the radio. Analysis of the VLA observations will be reported in a future publication. HST’s WFC3/IR camera was used to observe the three targets using the G141 grism over a wavelength range of 1.1–1.67  $\mu\text{m}$ . These observations followed the observing strategy demonstrated by numerous successful examples of brown dwarf HST/WFC3 grism monitoring, (e.g. Buenzli et al. 2012; Apai et al. 2013; Yang et al. 2015).

The G141 grism has a resolving power  $\approx 130$  over the range of 1.1–1.67  $\mu\text{m}$ , which provides access to the 1.4  $\mu\text{m}$  water feature that is inaccessible from the ground. A direct image was taken at the start of each orbit using the F127M filter to determine the location of the target on the detector and to provide

the G141 grism spectrum’s wavelength zero point. Every image was captured in a  $256 \times 256$  pixel sub-array and each target was observed in staring-mode. Each image had an exposure time of 63 seconds, with 43 exposures per orbit. Further details of the observations are summarised in Table 2.

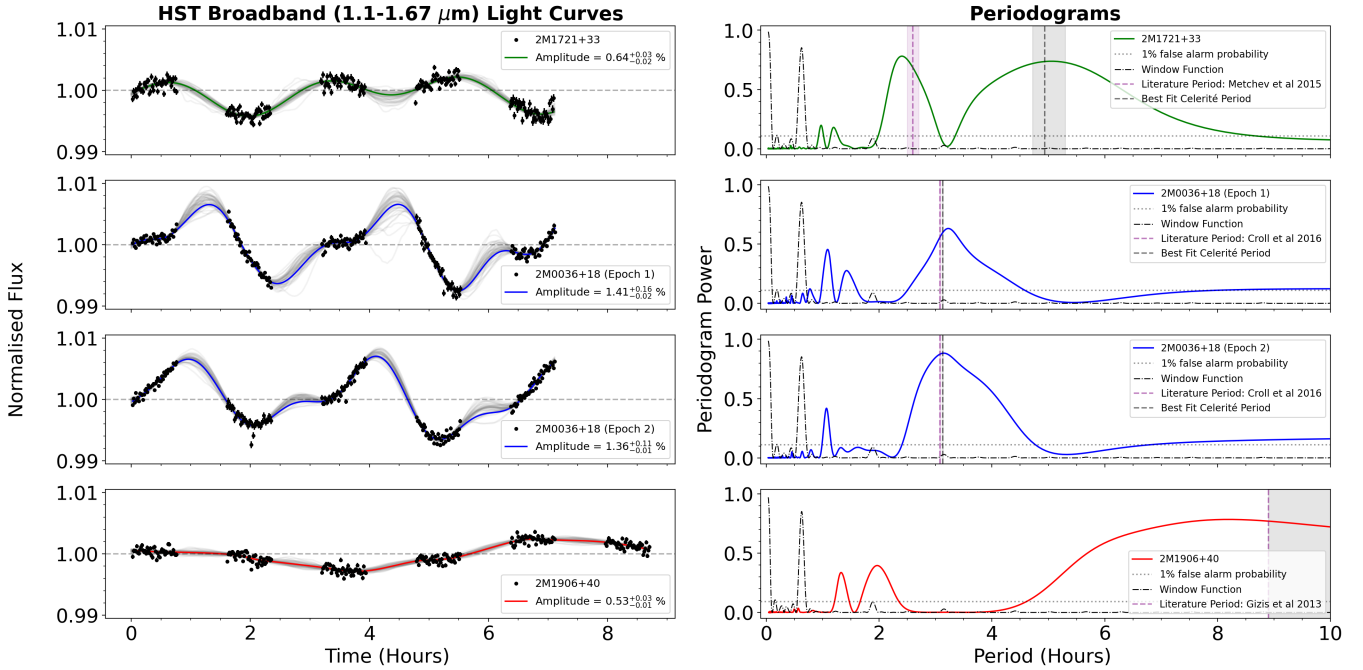
We observed 2M0036+18 and 2M1721+33 for five consecutive orbits in order to obtain full phase coverage. 2M1906+40 was observed for six consecutive orbits to obtain  $\sim 8.7$  hours of observations, slightly less than its reported period of  $8.88364 \pm 5 \times 10^{-5}$  hours, Gizis et al. (2015). We obtained two datasets for 2M0036+18 as the VLA was not available at the time of the first observation. Therefore, we conducted a repeat observation in December 2021 simultaneously with the VLA (see Table 2).

#### 3.2. Data reduction

In order to begin the spectral extraction process for our targets, the fit.fits files – produced by the calwf3 pipeline – were downloaded from the Barbara A. Mikulski Archive for Space Telescopes (MAST). We used custom python scripts and aXe (Kümmel et al. 2009) to reduce the data following best practices for HST/WFC3 (Buenzli et al. 2012; Apai et al. 2013; Yang et al. 2015).

We begin by identifying bad pixels by accessing the Data Quality extension from the fit.fits files. These arrays use flag values of 4, 16, 32, and 256 to identify bad pixels, hot pixels, unstable response pixels, and full-well saturation pixels, respectively. We corrected these pixels using bi-linear interpolation of the four neighbouring pixels in the  $x$  and  $y$  directions.

Background sky subtraction was undertaken in the same manner as that of Brammer et al. (2015). This considers both the zodiacal light background, i.e. the reflected solar continuum, and the spectral variation of the 1.083  $\mu\text{m}$  emission line in the Earth’s upper atmosphere, which varies in intensity over the course of an orbit. This algorithm is described fully in Appendix 6 of Brammer et al. (2015). Then, pixels that were flagged as cosmic rays were masked from the fit.fits files by



**Fig. 1.** *Left:* broadband G141 filter light curves of each of our observations. A sample of 100 posterior fits of each light curve are shown in grey, while the best-fit light curve is shown in colour for each observation. All four light curves are variable. *Right:* periodogram for each light curve. The periodogram for the window function of HST is shown by the dash-dotted black line in each plot. The period of each object from the literature is shown by the purple vertical line in the periodogram, while the period output from our *celerité2* models are shown by the black vertical lines. The uncertainties for the period values from the literature and from this work are represented by the shaded regions in purple and black, respectively. Our model periods agree with the literature for each observation except 2M1721+33, where we estimate a longer period of  $4.93^{+0.36}_{-0.22}$  hours.

interpolating over the neighbouring pixels. Aperture extraction widths of 5.65, 5.50, 5.50, and 5.60 pixels in diameter were used for 2M1906+40, 2M1721+33, 2M0036+18 (observation 1) and 2M0036+18 (observation 2), respectively. These widths were chosen as they provided the highest signal to noise in our spectra for each observation. We encountered wavelength calibration issues in the spectra of 2M1906+40. The assigned wavelengths were offset every second orbit, producing an artificial flux increase in the fifth orbit of the light curve. To correct this, we fixed the wavelength reference for each orbit to the position measured in the F127M image from the first orbit.

### 3.3. Light curve creation

In order to analyse the variability of our objects, we created light curves. We applied  $\sigma$ -clipping to the spectra of each object to remove any outliers, discarding any points  $>3\sigma$  away from the median spectrum of each observation. Using our spectra, we produced light curves with the *species* (Stolker et al. 2020) package for synthetic photometry, with the HST G141 filter.

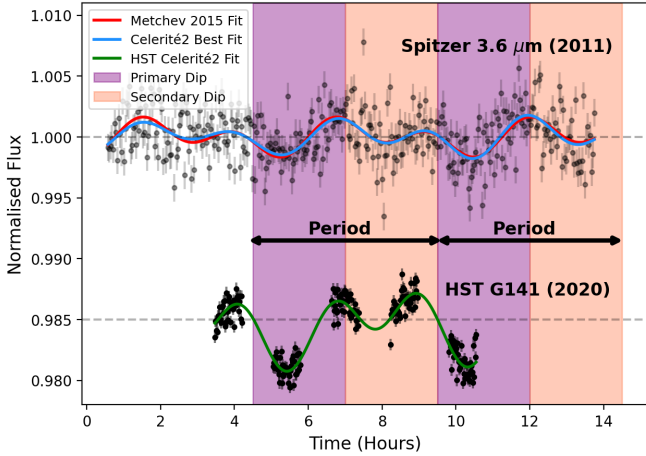
We noted a clear ramp effect in our resulting light curves. The ramp effect manifests as a tail at the start of each orbit and is most pronounced in the first orbit of each observation. This tail starts off dimmer before rapidly increasing in brightness. This effect is systematic of HST/WFC3 IR time series observations (Buenzli et al. 2012; Apai et al. 2013; Yang et al. 2015). We used the Ramp Effect Charge Trapping Eliminator (RECTE; Zhou et al. 2017) to correct for the ramp effect. RECTE models the systematic ramp effect using a physically motivated method that takes into account electron charge trapping and the fast and slow lifetimes of these trap populations. Charge carriers generated by incident photons can become trapped as they diffuse across the depletion

region of the photodiode. At the end of an exposure, the trapped charge carriers remain in the depletion region. This can lead to the ramp-shaped light curve in time series observations. The free parameters of the model are the initial fast and slow trap populations,  $E_{f,0}$ ,  $E_{s,0}$ , and their changes between orbits,  $\Delta E_{f,0}$ ,  $\Delta E_{s,0}$ . We constrained these parameters by fitting the RECTE model to the broadband G141 light curve, performing the fit separately for each observation. The final light curves for each observation are shown in Figure 1.

### 3.4. Light curve fitting

We used *celerité2* (Foreman-Mackey et al. 2017; Foreman-Mackey 2018) to model each of our light curves. *Celerité2* is an algorithm for fast and scalable Gaussian processes to determine a best-fit model to describe the variability. We used the *RotationTerm* kernel for each of our light curves. This kernel is typically used for modelling stellar rotation (Barros et al. 2020; Miles-Páez 2021) and incorporates two stochastically driven damped simple harmonic oscillating terms, one at a given period and one at half that period. We used this kernel as it has been demonstrated successfully on brown dwarf light curves in the past (Miles-Páez 2021; McCarthy et al. 2024, 2025). It is a data driven model that does not take into account the underlying physics of the variability mechanism, simply providing a fit to the data. There are five free parameters in this Gaussian process model:

- $\sigma$ : the standard deviation of the process.
- Period: the primary period of variability.
- Q0: the quality factor for the secondary oscillation.
- dQ: the difference between the quality factors of the first and second modes.



**Fig. 2.** Long-term stability of 2M1721+33 across two different wavebands. At the top is the [Metchev et al. \(2015\)](#) Spitzer 3.6  $\mu\text{m}$  light curve with their fit in red and our Gaussian process `celerité2` fit in blue applied to the Spitzer and HST light curves independently. Both models fit the data very similarly. At the bottom is the HST white light curve from our work with our Gaussian process `celerité2` overplotted (in green). Although these two light curves were observed over 9 years apart and at different wavelengths, they have a similar shape. They both exhibit a primary dip, followed by a secondary dip in each period. This may suggest long-term stability in 2M1721+33's light curve.

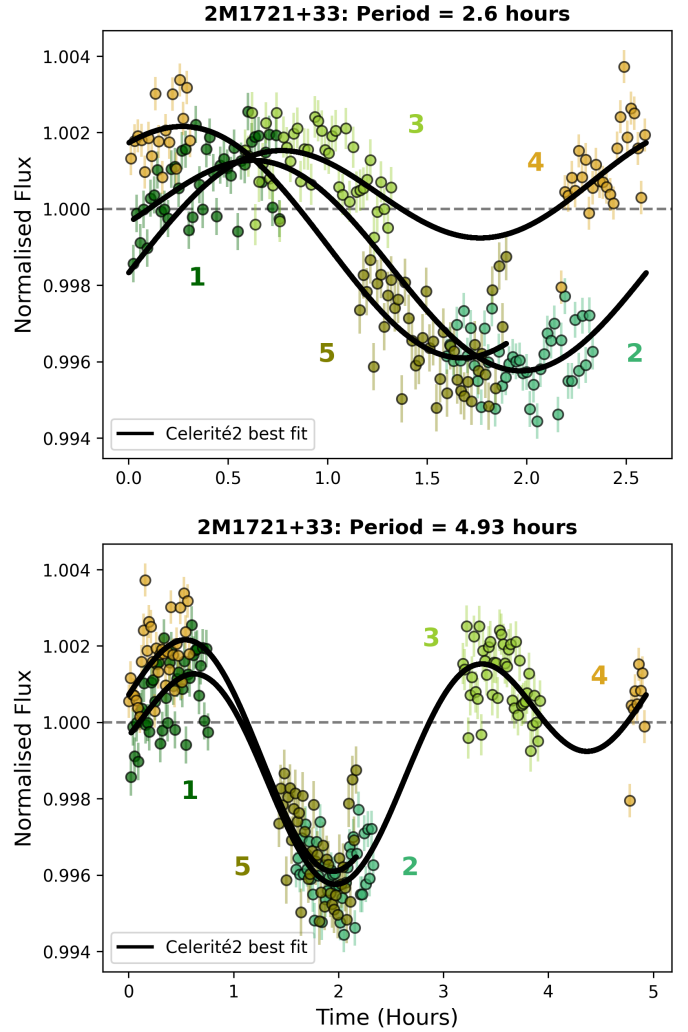
- $f$ : the fractional amplitude of the second mode compared to the primary.

We then used the Markov chain Monte Carlo (MCMC) package `emcee` ([Foreman-Mackey et al. 2013](#)) to find the best-fit model for each of the G141 white light curves, as well as the 2MASS J and 2MASS H' light curves. We used 35 walkers, with a burn in of 2000 steps followed by 10 000 steps. We verified that the chains had converged by visually inspecting the parameter traces. The parameter uncertainties were taken from the 16th and 84th percentiles of the posterior distributions.

We selected the maximum likelihood model as the best-fit light curve for each target. We show the best-fit models and a sample of 100 posterior fits for each target in the left panel of Figure 1. We calculated the amplitude of each light curve in the same manner as [Wilson et al. \(2014\)](#); [Radigan \(2014b\)](#); [Biller et al. \(2024\)](#). We define the maximum deviation for each light curve as the percent variation between the highest peak and deepest trough over the course of each HST observation. The amplitudes are shown in Table 1.

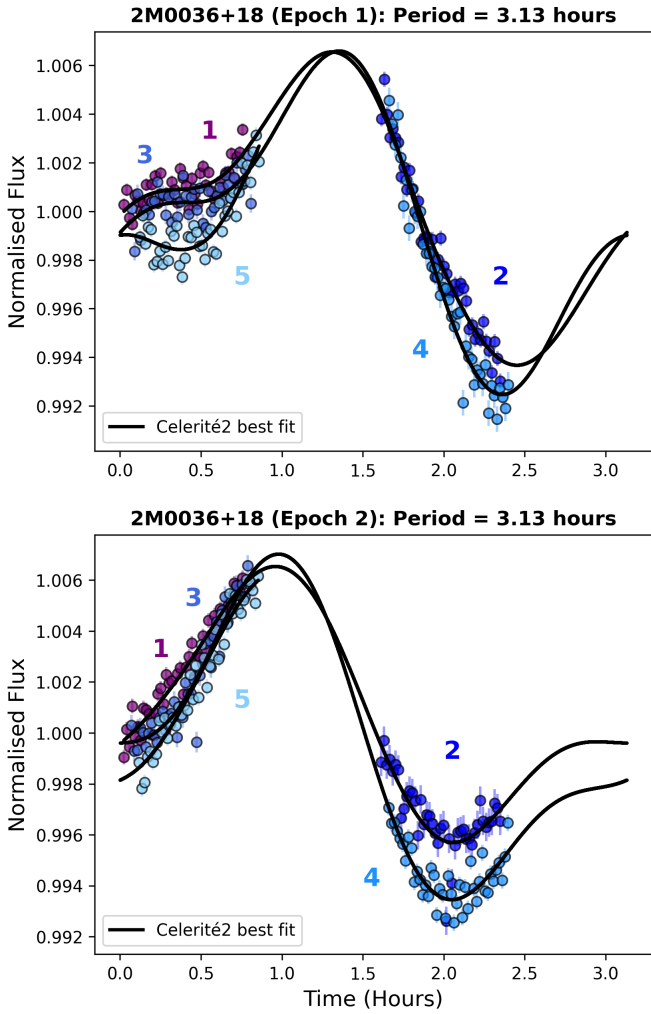
We also compute periodograms from our light curves using the astropy Lomb–Scargle function, ([Astropy Collaboration 2013](#); [Astropy Collaboration 2018](#); [Astropy Collaboration 2022](#)). In the right panel of Figure 1 we show the periodogram of each dataset with the period from our best-fit models in black, and the literature period in purple. The rotation period of 2M0036+18 from our model agrees with the rotation period estimated by [Croll et al. \(2016\)](#).

However, we find that for 2M1721+33, our model favours a period of  $4.9^{+0.4}_{-0.2}$  hours, in disagreement with the [Metchev et al. \(2015\)](#) value of  $2.6 \pm 0.1$  hours. The [Metchev et al. \(2015\)](#) results were based on Spitzer 3.6  $\mu\text{m}$  and 4.5  $\mu\text{m}$  photometric light curves of 2M1721+33 over 20 hours of observations. Their light curve model was a truncated Fourier series. They varied the number of terms,  $n$ , in their Fourier series until the residuals were consistent with random noise. To investigate the period of 2M1721+33, we used `celerité2` to fit a model to the



**Fig. 3.** *Top*: phase-folded white light curve of 2M1721+33 using the  $2.6 \pm 0.1$  hour period reported by [Metchev et al. \(2015\)](#). The photometric points from each HST orbit are colour-coded and the best-fit `celerité2` model is plotted in black. Using this period for phase-folding does not give a clean light curve. *Bottom*: phase-folded white light curve according to the period of  $4.9^{+0.4}_{-0.2}$  hours from this work. While this observation does not cover more than 1.5 rotations, the phase-folded regions (orbits 1 and 4, and orbits 2 and 5) overlap with each other in both the HST data and the `celerité2` model. This cleaner phase-folded light curve, along with Figure 2 motivated us to adopt this new period for 2M1721+33.

[Metchev et al. \(2015\)](#) Spitzer light curve, as seen in the top panel of Figure 2, and only applied it to the 3.6  $\mu\text{m}$  light curve as the 4.5  $\mu\text{m}$  light curve was not reported to be variable. From this model, we obtain a period of  $5.2^{+0.7}_{-1.6}$  hours for the Spitzer 3.6  $\mu\text{m}$  light curve. This period is consistent within uncertainties with the period that we obtained from the HST light curves. The shape of our periodogram for 2M1721+33 in Figure 1, replicates the shape of the periodogram from Figure 3 in [Metchev et al. \(2015\)](#), with two significant peaks at 2.6 hours and 4.9 hours. The two peaks have a very similar periodogram power, with the 2.6 hour peak being slightly higher. When taking into account both the Spitzer and HST light curve fits, our models prefer a double peaked light curve with a period of  $4.9^{+0.4}_{-0.2}$  hours. This double peaked feature is especially evident when comparing both the Spitzer and HST light curves as shown in Figure 2. Despite these



**Fig. 4.** *Top:* phase-folded white light curve of epoch 1 of 2M0036+18, using the period from this work ( $3.13^{+0.9}_{-0.8}$  hours). The best-fit *celerité2* model folds neatly as does the HST white light curve. *Bottom:* phase-folded white light curve for epoch 2 of 2M0036+18 using the same period. As with the first epoch, both the best-fit *celerité2* model and the HST white light curve overlap neatly. Despite both epochs covering slightly different phases of the light curve, the same period works well for both.

two light curves being observed over 9 years apart and at different wavelengths, there is a common shape to both light curves. They both exhibit a primary dip, followed by a secondary dip in each period. The primary and secondary dips in the light curve are seen across both light curves. This double peaked feature can be explained by the presence of two different atmospheric structures, one in either hemisphere (Vos et al. 2018). We show how the 2M1721+33 light curve looks when phase-folded according to the Metchev et al. (2015) period and the period from this work.

We found that our period for 2M1906+40,  $19.4^{+46.5}_{-10.5}$  hours, is poorly constrained. Our data do not have a sufficient time baseline to constrain the period, and therefore we assumed the period from Gizis et al. (2015) for this work. Gizis et al. (2013) and Gizis et al. (2015) both measured 2M1906+40's period to high precision,  $8.88364$  hours  $\pm 5 \times 10^{-5}$  hours, using observations from Kepler. They achieved such high precision on the rotation period due to their observations, which spanned  $>800$  days of continuous observations, from 28 June 2011 to 08 May 2013, (Programs GO 30101, GO 40004). During this time, the sinusoidal shape of

the light curve remained consistent. In comparison, we observe  $\sim 98\%$  of one period in this work.

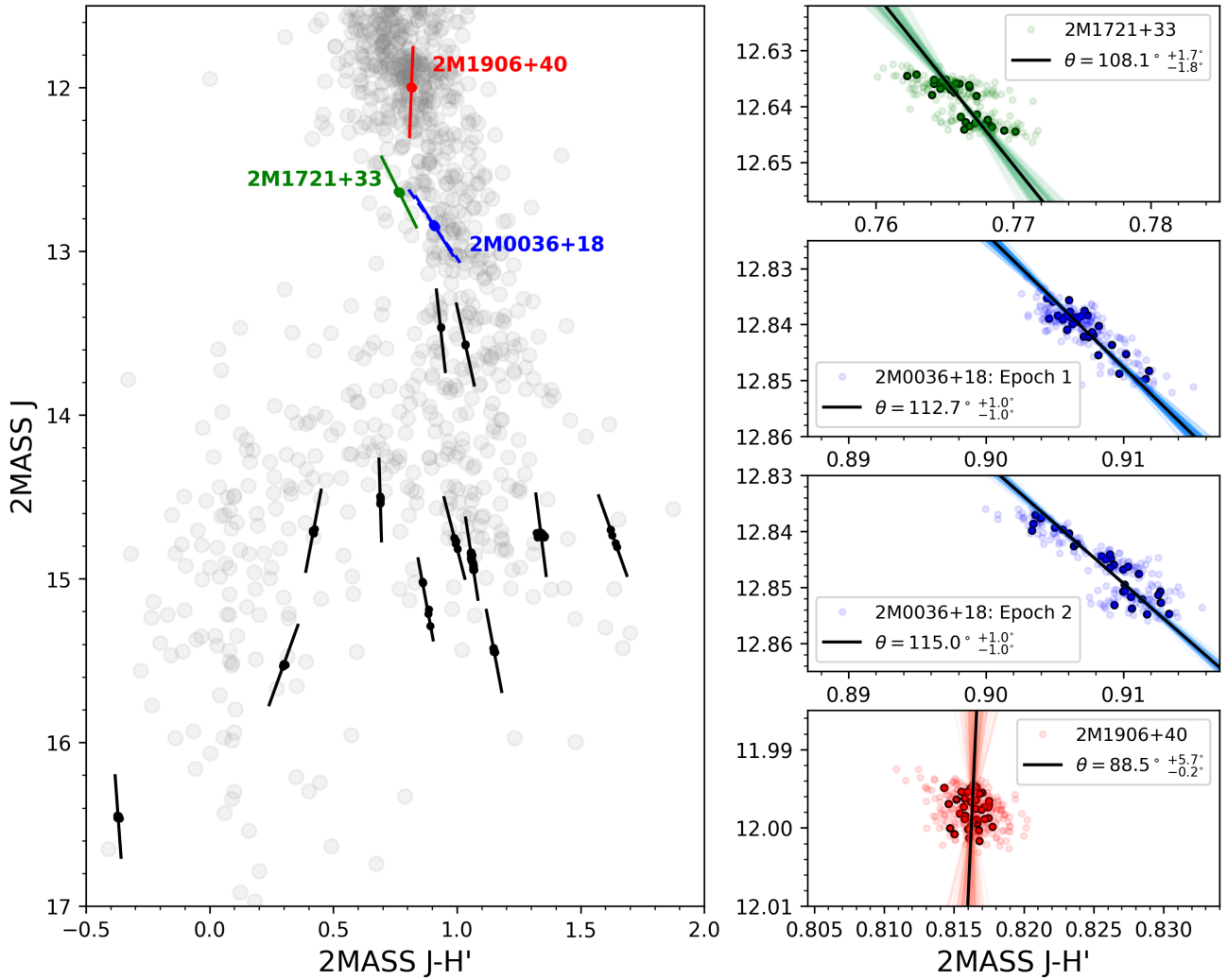
#### 4. Rotational modulations on the colour magnitude diagram

The colour magnitude diagram provides a convenient way to compare the broad observational properties of brown dwarfs at a glance. We can also investigate how the colours of brown dwarfs with rotational modulations vary by plotting their colour modulations on this diagram. By fitting a line to their J versus J-H' colour, we can measure how the colour of a brown dwarf changes as it rotates. We calculated the J and H' band magnitudes at each timestamp for each object following the procedure presented in Lew et al. (2020b). The full H band range extends past the range of the HST WFC3 instrument ( $\sim 1.48\text{--}1.82 \mu\text{m}$ ). Therefore, we define H' as a new band, which only extends to  $1.67 \mu\text{m}$ , the H-band region that is observable by WFC3. This is consistent with the method in Lew et al. (2020b). We show the J band magnitudes against the J-H' magnitudes to display the colour modulation in Figure 5.

Previously, Lew et al. (2020b) measured the J versus J-H' slopes for a sample of 12 brown dwarfs from spectral types L5-T8. These 12 brown dwarfs can be seen, plotted in black, along with our three brown dwarfs in the left panel of Figure 5. Following Lew et al. (2020b), we used orthogonal distance least squares regression (ODR), using the *scipy.odr* package to estimate the slope of the variability in colour-magnitude space as a function of rotation phase. ODR uses the orthogonal distance to define the best-fit slope of the modulations across the rotation period, accounting for uncertainties in both dimensions. In the right panel of Figure 5 we show the J versus J-H' colour modulation for each of the brown dwarfs from our observations, along with their ODR fits. We fit for  $\theta$ , the angle our line makes with the horizontal. A positive slope ( $\theta < 90^\circ$ ) suggests that the object becomes brighter and redder, while a negative slope ( $\theta > 90^\circ$ ) suggests that the object becomes brighter and bluer as it rotates. While some of the Lew et al. (2020b) objects have large J-band variations, i.e. 2M2139, all of the ODR slopes, are very steep.  $\theta$  varies from  $94^\circ < \theta < 103^\circ$  among L dwarfs, and  $76^\circ < \theta < 97^\circ$  among T dwarfs. This indicates that despite large variability amplitudes, their colour does not change significantly with phase.

Adding our three early L dwarfs to this sample allows us to probe a complementary parameter space of hotter effective temperatures. The slopes of the ODR fits to the three brown dwarfs follow an interesting trend: both 2M1721+33 and 2M0036+18 (epoch 1 and epoch 2) exhibit considerably shallower slopes than the other brown dwarfs in this sample ( $\theta = 108.1^{+1.7}_{-1.8}$ ,  $\theta = 112.7^{+1.0}_{-1.0}$ , and  $\theta = 115.0^{+1.0}_{-1.0}$ , respectively). The  $\theta$  of both observations of 2M0036+18 are similar but are slightly different. This can be attributed to the fact that each epoch observes a slightly different phase of 2M0036+18. This can be seen in the middle sub panel of Figure 5. As these two objects rotate they become brighter and bluer in J-H'. 2M1906+40 on the other hand has a drastically different slope compared to the other two objects in this sample, with  $\theta = 88.5^{+5.7}_{-0.2}$ . This results in a near vertical slope, indicative of grey variability modulations. It is also the steepest slope among all the L dwarfs between this and Lew et al. (2020b)'s sample. We note that we use these slopes as tools to help illustrate the trend in the colour modulation of each object.

In order to further justify our fits to the colour modulation, we use the approach described in McCarthy et al. (2024). We plotted the colour modulation of the 2MASS J and the 2MASS



**Fig. 5.** *Left:* colour modulation of 2M1721+33 (green), 2M0036+18 (blue), and 2M1906+40 (red) compared with the sample from Lew et al. (2020b) (black circles and lines). The grey circles represent brown dwarfs with known parallaxes across spectral types, taken from the UltracoolSheet Best et al. (2024). The slopes of 2M1721+33 and 2M0036+18 in particular are shallower in comparison to the rest of the sample, showing that as these objects rotate, they become brighter and bluer. *Right:* 2MASS J vs 2MASS J-H' magnitudes for each observation (dark circles are binned down to a 441-second cadence) along with the individual fits for each brown dwarf using orthogonal distance regression (ODR), with their  $\theta$  values and associated uncertainties. The colour modulation plot for 2M1721+33 appears as nearly two individual regions; however, this is due to a lack of phase coverage as a result of HST observations. The two observations of 2M0036+18 are plotted separately and they have a very similar slope, even though the observations were performed 16 months apart.

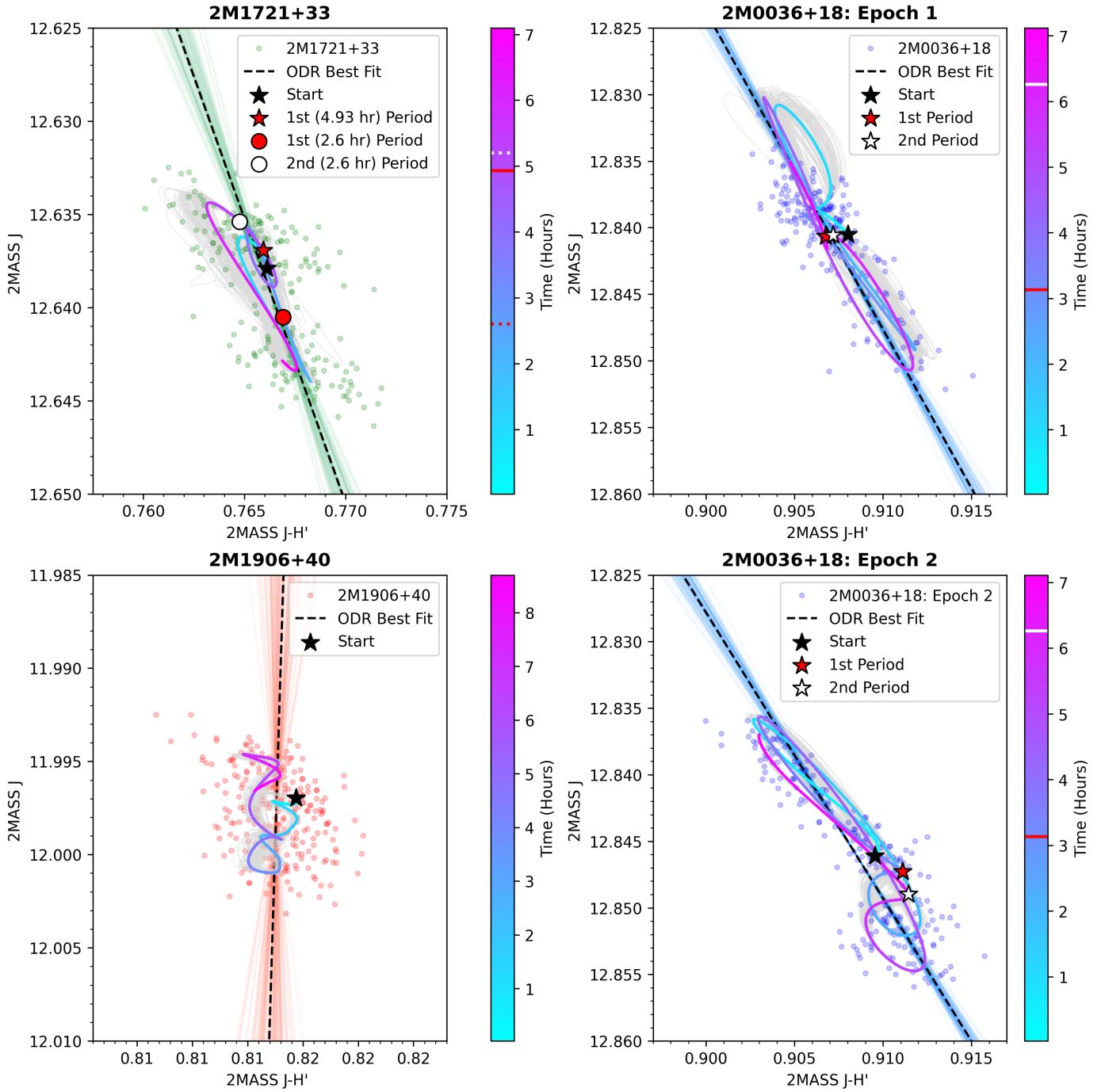
J-H' celerit 2 light curves. This is particularly useful for 2M1721+33 and 2M1906+40, where the trend in colour modulation is not immediately obvious. As illustrated in Figure 6, the general trend of the light curve modulation matches the ODR fits shown in Figure 5. The shape of 2M1721+33's colour modulation is distinctly split into an upper and lower region, as seen in Figure 5. However, this shape is due to the fact that we do not obtain full phase coverage of 2M1721+33's light curve with HST. As illustrated in Figure 3, we observe the peaks and troughs of its period, missing out on the middle portion, which results in our two regions in colour space. However, when using the interpolated light curve colour modulation in Figure 6, it is evident that our ODR fits follow the same path. Similar colour-modulation shapes have been seen in other HST variability papers (Lew et al. 2020b; Chapleski & Zhou 2026).

Lew et al. (2020b) suggest a potential trend in the colour modulation of brown dwarfs across their sample. Early T dwarfs become brighter with little to no colour change, while L dwarfs

become brighter and bluer. Our sample allows us to investigate this trend at earlier spectral types, and we find that the trend continues at these warmer temperatures. Both 2M1721+33 and 2M0036+18 become brighter and bluer than any other brown dwarf in the Lew et al. (2020b) sample. Whereas 2M1906+40 has a near vertical, positive slope that is steeper than any other L dwarf, with grey modulations. The behaviour of our three objects in Figure 5, suggests that brown dwarfs continue to become brighter and bluer as they rotate, except for 2M1906+40, which has grey modulations.

## 5. Establishing and modelling the spectral variability amplitude

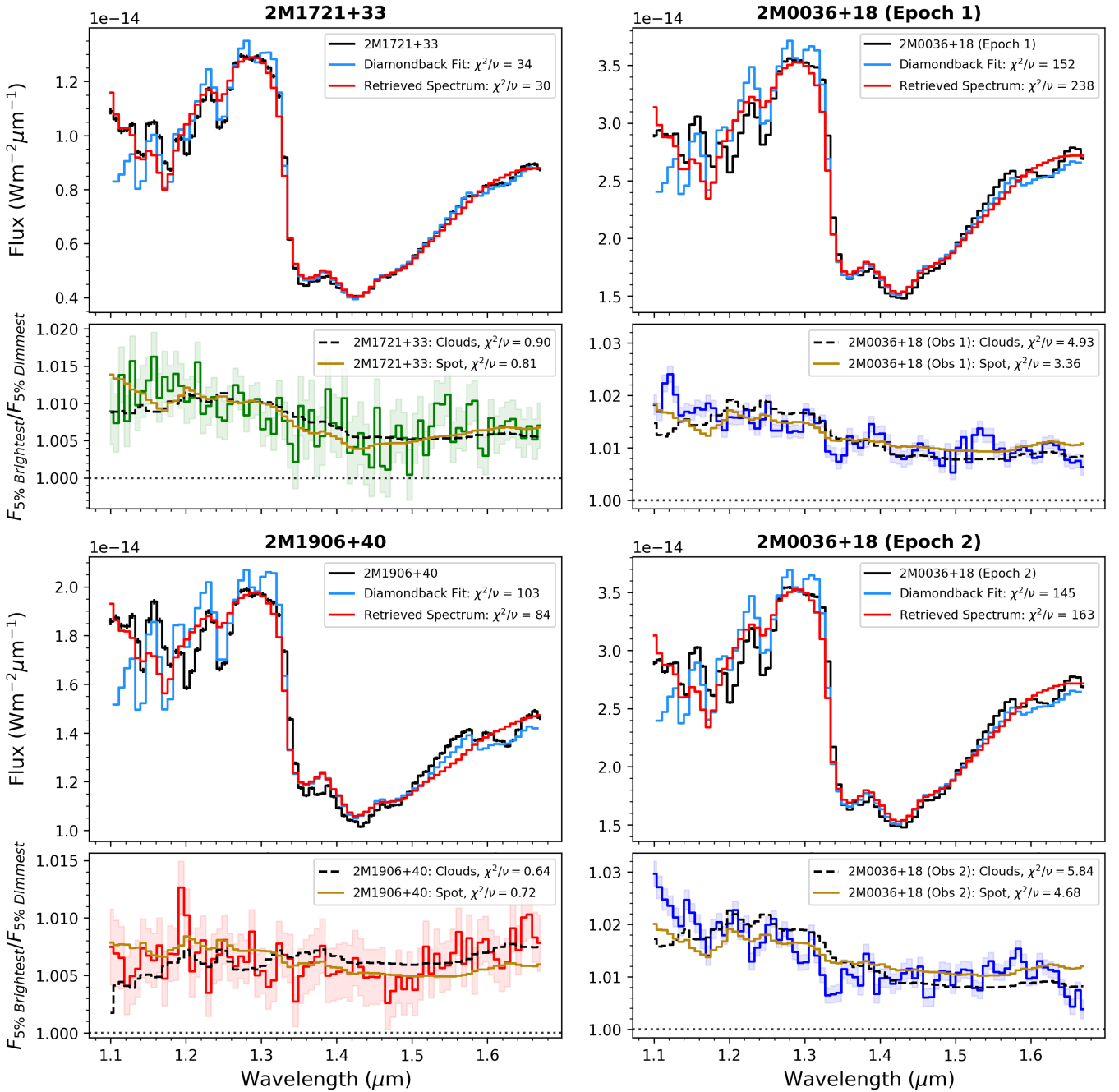
The spectral variability amplitudes (wavelength dependence of the variability) allow us to discern the main drivers of variability on brown dwarfs. Within the HST wavelengths, the shape has



**Fig. 6.** Colour modulation for each of our objects. The coloured circles represent the HST data shown in the right panels of Figure 5. The dashed and solid lines are the best-fit ODR and a selection of 100 posterior ODR fits for each object. The purple-blue solid line represents the colour modulation of the 2MASS J vs 2MASS J-H' *celerité2* light curves. This has the added benefit of interpolating over the gaps left by the HST observations. The colourbar on the right shows how the colour changes as a function of time for each object. The colour modulation of a sample of 100 posterior light curves is also shown for each object in light grey. The location of the start of the observations for each object is highlighted by a black star. A red star is used to highlight the position after one period, and a white star after two periods. The start of each period is consistent for each object with more than one rotation, further supporting our chosen periods. The ODR linear fits for each object are consistent with the path traced out by the light curve modulation.

been found to differ across spectral types, with L dwarfs typically exhibiting amplitudes that decrease smoothly as a function of wavelength (Yang et al. 2015; Manjavacas et al. 2018; Lew et al. 2020a). However, in L/T transition objects, the water band is typically less variable than the surrounding continuum (Yang et al. 2015; Zhou et al. 2018). The reason for this change is that

in the L-dwarf regime, there are spatially varying high-altitude clouds that drive the variability (Yang et al. 2015). In contrast, along the L/T transition the condensate clouds are deeper in the atmosphere, which results in a reduction of the variability amplitude within the water opacity band (Yang et al. 2015; Vos et al. 2023).



**Fig. 7.** HST spectrum (black line), Sonora Diamondback best-fit spectrum (blue line), and petitRADTRANS retrieval spectrum (red line) is in the upper panel for each plot. The bottom panels shows the spectral variability amplitude for each object (green, blue, or red lines). The cloud and spot driven spectral variability models are plotted in black and gold, respectively. For both 2M1721+33 and 2M1906+40, the cloud-driven and magnetic spot-driven models both provide adequate fits to the data. For both observations of 2M0036+18, there is a slightly more favourable  $\chi^2/\nu$  value for the spot model over the cloud model.

The spectral variability amplitude of each observation was obtained by calculating the ratio of the 5% brightest spectra to the 5% dimmest spectra in each observation. The brightest and dimmest 5% of spectra were chosen because they are representative of these regions while also being a typical amount used in the literature ( $\sim 5\text{--}10\%$ ) (Zhou et al. 2018; Lew et al. 2020b). We chose this percentage as it boosts the signal to noise as opposed to simply taking the single brightest and dimmest spectra, while also still probing a small percentage of the phase. We used the `spectres` package (Carnall 2017) to resample the wavelength points onto a common grid and interpolate the flux across these

points before calculating the brightest/dimmest ratio. The error on this ratio was calculated using MCMC bootstrapping.

The spectral variability amplitudes of our objects are shown in Figure 7. All three targets are variable at all wavelengths. It is clear that they are similar to other HST observations of brown dwarfs in the L-dwarf regime. They do not exhibit lower variability amplitudes in the water band, as is typical of L/T transition objects (Apai et al. 2013; Yang et al. 2015). The variability for 2M1721+33 and both observations of 2M0036+18 exhibit a slight slope, where the amplitude is highest at bluer wavelengths and decreases at longer wavelengths. In contrast, the variability

**Table 3.** Uniform prior ranges and results from the petitRADTRANS retrievals and the parameters for the best-fit Sonora Diamondback spectra.

Parameter	Prior	2M1721+33	2M0036+18 (epoch 1)	2M0036+18 (epoch 2)	2M1906+40
<b>petitRADTRANS</b>					
<b>Log (g)</b>	$\mathcal{U}(3.0, 5.5)$	4.24±0.04	4.27±0.01	4.27±0.01	4.96±0.06
<b>Radius (<math>R_{Jup}</math>)</b>	$\mathcal{U}(0.6, 1.4)$	0.84±0.01	0.81±0.01	0.81±0.01	0.85±0.01
<b>Cloud fraction (<math>f_{cloud}</math>)</b>	$\mathcal{U}(0, 1)$	0.73±0.01	0.83±0.01	0.83±0.01	0.82±0.01
<b><math>k_0</math> (350nm opacity)</b>	$\mathcal{U}(-5, 20)$	2.42±0.75	2.01 ±0.04	2.02 ±0.04	5.62±0.68
<b><math>\gamma</math> (Power-law exponent)</b>	$\mathcal{U}(-20, 20)$	-4.23±0.20	-4.02±0.02	-4.04±0.02	-4.01±0.02
<b>Log<sub>10</sub> CO</b>	$\mathcal{U}(-7, 0)$	-2.54±1.97	-2.60±0.03	-2.61±0.03	-1.99±0.07
<b>Log<sub>10</sub> H<sub>2</sub>O</b>	$\mathcal{U}(-7, 0)$	-3.09±0.02	-3.22±0.01	-3.22±0.01	-2.79±0.05
<b>Log<sub>10</sub> K</b>	$\mathcal{U}(-7, 0)$	-6.85±0.04	-7.00±0.01	-6.99±0.01	-7.00±0.01
<b>Log<sub>10</sub> Na</b>	$\mathcal{U}(-7, 0)$	-6.98±0.04	-7.00±0.01	-7.00±0.01	-6.98±0.03
<b>Sonora Diamondback</b>					
<b><math>T_{eff}</math> (K)</b>	–	1700	1600	1600	2000
<b>Log (g)</b>	–	5	5	5	5.5
<b>Metallicity</b>	–	+0.5	+0.5	+0.5 1.1	+0.5
<b>C/O</b>	–	1.0	1.0	1.0	1.0

**Notes.** The abundances of the species are in units of log mass mixing ratio.

of 2M1906+40 shows little wavelength dependence. 2M0036+18 exhibits the highest variability amplitude (in both observations), followed by 2M1721+33 and then 2M1906+40.

### 5.1. Modelling the spectral variability amplitude

Clouds, aurorae, and magnetic spots are three possible drivers of variability for early L dwarfs. Each of these processes has a different effect on the spectral variability amplitude. To model the spectral variability amplitude, we use a combination of radiative-convective equilibrium forward models, together with a more flexible atmospheric retrieval approach.

Multiple modelling and retrieval studies have shown that having a wide wavelength range is critical for determining the thermal profile (Burningham et al. 2021; Vos et al. 2023). As our observations cover a narrow wavelength range, we use the pressure temperature profile from the best-fit Sonora Diamondback grid models (Morley et al. 2024). In order to obtain the pressure temperature profiles, we used the Sonora Diamondback grid models to identify the best-fitting spectrum to the median spectrum of each object. Sonora Diamondback is a one-dimensional fully self-consistent model with five varying parameters: temperature, gravity, a cloud sedimentation parameter  $f_{sed}$ , metallicity, and the C/O ratio relative to solar values. We linearly interpolated these models over the grid to find the best fit using the *species* and *pymultinest* packages (Stolker et al. 2020; Buchner et al. 2014). The optimal fits, shown in Figure 7, have reduced  $\chi^2$  values of 34, 152, 145, and 103 for 2M1721+33, 2M0036+18 (epoch 1), 2M0036+18 (epoch 2), and 2M1906+40, respectively. From the interpolated best fit, we took the closest model from the grid and used its pressure temperature profile for the atmospheric retrievals. The parameters for these models are found in Table 3, and the fits themselves in Figure 7.

Clouds are expected to form in early L dwarfs (Burrows et al. 2006). We compared the condensation curve of forsterite ( $Mg_2SiO_4$ ) to the pressure temperature profile of each of our objects (taken from the Sonora Diamondback models) and found that it is expected to form in the photosphere of each of our objects (1–0.1 bar). The justification for clouds is further strengthened since both 2M1721+33 and 2M0036+18 show

prominent silicate features at 10 $\mu$ m in archival Spitzer IRS data (Cushing et al. 2006; Suárez & Metchev 2022).

We used atmospheric retrievals to model the median spectrum of each object as they provide greater flexibility for the parameters to reproduce the spectral variability amplitudes. Self-consistent grid models do not have the resolution in parameter space to capture the ~1% variability that we are trying to fit. We used the retrieval package *petitRADTRANS* (Mollière et al. 2019; Nasedkin et al. 2024) to run our analysis. We fitted the same median spectrum for each object that we used when getting the best-fit Sonora Diamondback spectrum. We fixed the pressure temperature profile to the self-consistent Diamondback profile, as the HST spectra do not have a wide enough wavelength coverage or sufficient spectral resolution to constrain it (Burningham et al. 2021; Vos et al. 2023). The distance to each object was also fixed to their estimated distances from the literature, ( $16.18 \pm 0.04$  pc,  $8.74 \pm 0.01$  pc, and  $16.76 \pm 0.03$  pc for 2M1721+33, 2M0036+18, and 2M1906+40, respectively Gaia Collaboration (2021)). The brown dwarf radius and surface gravity was allowed to vary, as well as the atmospheric mass fraction abundances of CO, H<sub>2</sub>O, K and Na (Rothman et al. 2010; Polyansky et al. 2018; Mollière et al. 2019; Allard et al. 2019). The remaining atmosphere was filled with H<sub>2</sub> and He, forming respectively 76% and 24% of the mass, similar to the composition of Jupiter (Atreya et al. 2003; Ben-Jaffel & Abbes 2015). We account for Rayleigh scattering in H<sub>2</sub> and He as well as the collision induced absorption for the H<sub>2</sub>–H<sub>2</sub> and H<sub>2</sub>–He interactions. We ran the retrievals using nested sampling, using the *pyMultiNest* package (Feroz et al. 2009, 2013; Buchner et al. 2014).

Three parameters were used to model clouds in the atmosphere. A cloud fraction ( $f_{cloud}$ ) parameter parametrises to what extent the brown dwarf’s atmosphere is covered by clouds. We also had two power-law cloud parameters. The first power-law cloud parameter is the opacity of the cloud at 350 nm,  $k_0$ , and the second is the power-law cloud exponent,  $\gamma$ . This parametrisation was chosen as it approximates the effect of clouds by introducing a power-law opacity slope to the spectrum.  $\gamma$  typically ranges from –5 to 0, with the approximated scattering increasing with increasing  $\gamma$ , as the opacity slopes also

**Table 4.** Best-fit parameters for the cloud and magnetic spot models for each brown dwarf.

Object	Retrieved $\gamma$	Model $\gamma$	Retrieved $k_0$	Model $k_0$	$\Delta A$ (%)	$\chi^2/\nu$	$\Delta T$ (K)	$\Delta A$ (%)	$\chi^2/\nu$
2M1721+33	$-4.23^{+0.15}_{-0.20}$	$-4.11^{+0.07}_{-0.09}$	$2.42^{+0.75}_{-0.43}$	$1.61^{+0.14}_{-0.19}$	$15.50^{+9.48}_{-6.81}$	0.90	-200	$2.89^{+0.07}_{-0.08}$	0.81
2M0036+18 (epoch 1)	$-4.02^{+0.01}_{-0.02}$	$-4.00^{+0.02}_{-0.03}$	$2.01^{+0.04}_{-0.03}$	$1.46^{+0.04}_{-0.04}$	$12.95^{+0.20}_{-0.27}$	4.93	-200	$2.95^{+0.02}_{-0.03}$	3.36
2M0036+18 (epoch 2)	$-4.04^{+0.02}_{-0.02}$	$-3.87^{+0.04}_{-0.06}$	$2.02^{+0.04}_{-0.04}$	$1.04^{+0.06}_{-0.10}$	$10.15^{+0.47}_{-2.65}$	5.84	-200	$3.25^{+0.03}_{-0.03}$	4.68
2M1906+40	$-4.01^{+0.01}_{-0.02}$	$-3.03^{+0.03}_{-0.03}$	$5.62^{+0.68}_{-0.49}$	$5.65^{+0.20}_{-0.22}$	$2.10^{+0.05}_{-0.06}$	0.64	-300	$1.46^{+0.04}_{-0.04}$	0.72

**Notes.**  $k_0$  is the opacity of the clouds at 350nm and  $\gamma$  is the power-law exponent of the clouds.  $\Delta A$  is the coverage fraction between the bulk atmosphere and the perturbed atmosphere (due to clouds or magnetic spots).

increase. The value  $\gamma$  characterises how non-grey the clouds are (Burningham et al. 2017), with decreasing values of  $\gamma$  resulting in increased non-greyness of the clouds and with  $\gamma = 0$  mimicking a grey cloud deck.

We tested clear, patchy, and fully cloudy atmospheres, and found that using a patchy power-law cloud fit the spectral variability amplitudes the best. A patchy cloud model, where the clouds were grey, failed to reproduce the spectral variability amplitude as accurately as the patchy power-law model. The  $\chi^2/\nu$  value for these models was at least double that of the best-fit patchy power-law cloud model that we used. We therefore favoured the patchy power-law model.

We used 400 live points and had an evidence tolerance of 0.5 in our retrievals. We calculated our models at a spectral resolving power of 180 before convolving them to the spectral resolving power of the G141 spectra,  $R \sim 130$ . They were then binned to the wavelength grid of the data to calculate the log likelihood. We set broad priors on our retrieved parameters, as shown in Table 3. We also retrieved mass fraction abundances for each of the species in our atmosphere. The best-fit retrieved spectra are shown in Figure 7.

We chose to model the spectral variability amplitude of each object by creating two separate regions in the atmosphere. For each different variability mechanism, we created a different atmospheric region. We modelled cloud driven variability by changing the properties of the clouds in each region. We modelled aurorally driven variability by inserting a temperature inversion into the upper atmosphere of one of our regions. This temperature inversion follows findings from (Faherty et al. 2024; Nasedkin et al. 2025) where aurorae can heat the upper atmosphere. We modelled magnetic spot driven variability by reducing the temperature of the pressure temperature profile, in order to mimic the effects of a magnetic spot (Rockenfeller et al. 2006; Barnes et al. 2015; Rackham et al. 2018). Other than each specific change for the different variability mechanisms, the models were identical to the best-fit retrieval. By altering our best-fit retrieval with each of these parameters (changing the cloud properties, inserting a temperature inversion, and reducing the temperature of the pressure temperature profile), we can model the spectral variability amplitude.

We used the following equation, adapted from Lew et al. (2020a), to model the spectral variability amplitude:

$$V = 1 + \Delta A \left( \frac{F_{\text{brighter}} - F_{\text{dimmer}}}{F_{\text{min}}} \right). \quad (1)$$

Here  $\Delta A$  is the coverage fraction of the secondary region (i.e. the region with differing cloud properties, an auroral temperature inversion or a magnetic spot),  $F_{\text{brighter}} - F_{\text{dimmer}}$  is the difference between the flux of the brighter and dimmer spectra between the baseline retrieved spectrum and the secondary spectrum,  $F_{\text{min}}$

is the observed spectrum of the mean 5% dimmest spectra of the object, and  $V$  is the variability amplitude as a function of wavelength.

## 5.2. Cloud modelling

To model cloud driven variability, we assume that the variability is caused by two regions in the atmosphere with different cloud properties. The variability was modelled by changing the values of both  $k_0$ , the opacity at 350 nm, and  $\gamma$ , the power-law opacity exponent. We fit the spectral variability amplitude using Equation (1). We ran an MCMC using emcee to optimise our model and to constrain our uncertainties (Foreman-Mackey et al. 2013), and the median value of our constrained parameters for each model are shown in Table 4. We used 20 walkers and 500 steps for each observation until the model converged. As the two power-law cloud parameters are degenerate, there is a spread in the possible values for a good fit, as shown in Appendix B, Figures B.1, B.2, B.3, and B.4.

We show our model fits for cloud driven variability in Figure 7. Our model for cloud driven variability fits the spectral variability amplitude for 2M1721+33 and 2M1906+40, with  $\chi^2/\nu < 1$  for both cases, and a slightly poorer fit for 2M0036+18. The higher  $\chi^2/\nu$  for 2M0036+18 is driven by its smaller uncertainties. Our model follows the general shape of the spectral variability amplitude for all three of our objects. These models indicate that the observed spectral variability may be produced by two regions in the atmosphere that host clouds with differing properties.

## 5.3. Auroral modelling

We used a similar framework to model auroral variability. However, instead of changing the cloud property parameters, we inserted a temperature inversion in the upper atmosphere of the pressure temperature profile to create our secondary auroral region. Temperature inversions are common in our solar system, and in the case of Jupiter, the bulk of this upper atmosphere heating is driven by the redistribution of heat from hot polar auroral regions (O'Donoghue et al. 2021). Evidence of temperature inversions in brown dwarf atmospheres has recently been reported by Faherty et al. (2024), where auroral emission of methane was reproduced by inserting a temperature inversion into the pressure temperature profile. However, unlike Faherty et al. (2024), we assume that the auroral feature only exists in one region of our atmosphere. Following examples by Morley et al. (2014) and Nasedkin et al. (2025), we insert a Gaussian temperature perturbation into the pressure temperature profile. The perturbation has three parameters: the temperature of the inversion, the pressure of the peak of the inversion and the width of the temperature inversion. We insert a temperature inversion of 350 K with a peak pressure of 1 mbar and a width of one

pressure scale height (10–0.1 mbar), similar with retrieved values from Faherty et al. (2024) and Nasedkin et al. (2025). While our selection of a 350 K inversion is somewhat arbitrary, we do not have the wavelength coverage nor resolution to perform retrievals as in Faherty et al. (2024) and Nasedkin et al. (2025) to obtain a more accurate inversion value. However, the effects of the temperature of the inversion are simply more prominent with a higher temperature. We tested the effect of lower temperature inversion (100 K, 200 K) on the model and it resulted in the same shape of the model, but at a lower amplitude. 350 K helps to visualise these effects while also being similar to those in the literature.

As before, we fit the spectral variability amplitude using Equation (1). We modelled two regions of the atmosphere. The primary region is the fiducial state of the atmosphere as measured by the atmospheric retrieval, while the second region includes the auroral temperature perturbation. In Figure 8, we show the variability amplitude predicted by a 350K inversion at a range of different pressures in the atmosphere, with  $\Delta A = 0.1$ . The majority of the variability is found within the water band at  $\sim 1.3\text{--}1.45\ \mu\text{m}$ , as shown in Figure 8. When the temperature inversion is located at lower pressures, the amplitude of the variability is weaker. However, as the location of the temperature inversion moves deeper in the atmosphere, the amplitude of variability within the water band increases. This model clearly does not reproduce the spectral variability of 2M0036+18 across the full HST wavelength range.

#### 5.4. Magnetic spot modelling

Magnetic spots are regions on stellar–substellar objects where the magnetic field suppresses the emission of flux from the surface, analogous to sunspots on the Sun. They are common features on M dwarfs and are postulated to be common on L dwarfs as well (Schmidt et al. 2015). By changing the emitting flux, such spots are hypothesised to be potential drivers of variability (Croll et al. 2016).

Magnetic spot driven variability has been modelled on M dwarfs in the literature by using cooler and hotter regions in the models. These typically consist of a bulk photosphere region, with a higher  $T_{eff}$ , and a magnetic spot region with a cooler  $T_{eff}$  (Afram & Berdyugina 2015; Rackham et al. 2018). To model the spectral variability driven by magnetic spots, we altered the pressure temperature profile of our best-fit forward model. Our first region represents the bulk photospheres and is the same as in the previous models – our unperturbed retrieved spectrum for each object. Our second region mimicks the effects of magnetic spots with a Sonora Diamondback PT profile that is 100–300 K cooler for each object. This is in line with the  $\Delta T_{eff}$  values between magnetic spots and the bulk  $T_{eff}$  on mid to late M dwarfs (Rockenfeller et al. 2006; Barnes et al. 2015). Rockenfeller et al. (2006) report variability as a result of a magnetically induced cool spot on the M9 dwarf 2M1707+64, with a  $T_{eff}$  100K cooler than the  $T_{eff}$  of the bulk atmosphere.

All other parameters were kept constant. This created two regions of the atmosphere, and using our model from Equation (1), we can reproduce magnetic spot induced variability. This method of modelling the effects of magnetic spot variability is common, particularly for M dwarfs hosting exoplanet systems. For example, Afram & Berdyugina (2015) modelled the effects of starspots on M dwarfs and found that the ratio between the photosphere and the spots temperatures,  $\frac{T_{spot}}{T_{phot}} \approx 0.86$ . Using this temperature ratio relation, Rackham et al. (2018) found that

the typical covering fraction of spots on M9 dwarfs to range from  $1\% < f_{spot} < 24\%$ . We note that while this is a simple model to replicate the effects of magnetic spot induced variability, we are limited by the nature of our observations, which are low-resolution spectra covering a wavelength range of  $0.57\ \mu\text{m}$ . In order to produce a more sophisticated model (i.e. one that could Doppler image the spots and highlight the latitude and longitude of the spots) would require high-resolution spectra Roettenbacher et al. (2017).

The best-fit parameters for each observation are found on the right hand side of Table 4. We ran an MCMC to obtain the best-fit  $\Delta A$  across each of the 100–300 K cooler magnetic spots. This used the same set up as for the cloud models, (20 walkers, 500 steps). Using this MCMC we then found the optimal  $\Delta T$  and subsequently  $\Delta A$  values for each observation. We chose the  $\Delta T$  value with the lowest  $\chi^2/\nu$  value. The posterior distribution for the  $\Delta A$  value for the best-fit  $\Delta T$  for each object is shown in Figure C.1.

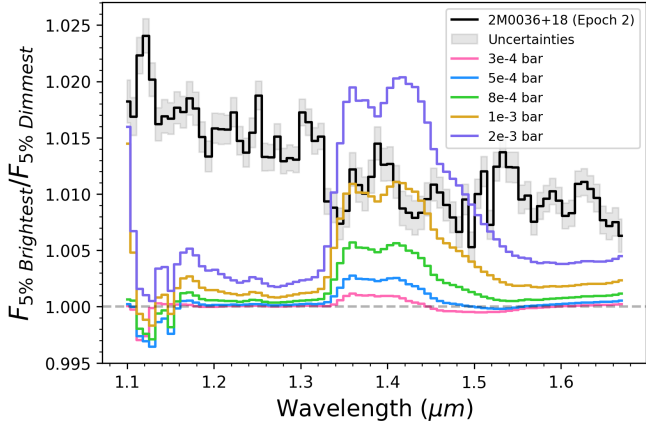
Figure 7 shows our magnetic spot driven variability model. This paper is the first to investigate the presence of magnetic spot induced variability on brown dwarfs using HST. As with our cloud driven variability, the magnetic spot variability models fit our data very well. For each observation, using a spot with  $\Delta T_{eff} = -200\ \text{K}$  provided the best-fitting model, except for 2M1906+40, where  $\Delta T_{eff} = -300\ \text{K}$  provided the best fit, with these  $\Delta T_{eff}$  values providing the lowest  $\chi^2/\nu$  value for each object. Again, these models fit 2M1721+33 and 2M1906+40 slightly better than 2M0036+18, and this can be attributed to 2M0036+18’s smaller uncertainties. These results are consistent with the M dwarf studies of Afram & Berdyugina (2015); Rackham et al. (2018). The ratio between the temperature of the photosphere and the spots are similar to the 0.86 value of Afram & Berdyugina (2015), ranging from 0.85–0.88 for each object and the coverage fractions (analogous to the  $\Delta A$  values presented in this work), also all lie between  $1\% < f_{spot} < 24\%$ . These results indicate that our model results are at least plausible as they are consistent with the literature.

## 6. Discussion

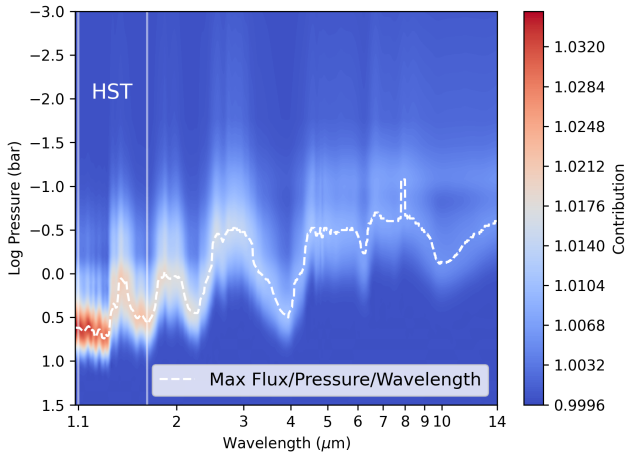
### 6.1. Comparison between the cloud, auroral and magnetic spot models

We found that our cloud and magnetic spot models fit the spectral variability amplitude the best for each of our objects. They have similar  $\chi^2/\nu$  values for each object, so from our observations we cannot rule out one model over another. Future observations at longer wavelengths are required to disentangle these two processes. Both 2M1721+33 and 2M0036+18 exhibit a  $10\ \mu\text{m}$  silicate feature in their spectra (Cushing et al. 2006; Suárez & Metchev 2022), which indicates that silicate clouds are present in their atmospheres. Variability at this  $10\ \mu\text{m}$  silicate feature has been a proposed indicator of cloud driven variability (Biller et al. 2024; McCarthy et al. 2025; Chen et al. 2025). Probing the variability at these wavelengths could determine whether the variability is primarily driven by clouds or magnetic spots. If the variability amplitude continues to decrease at longer wavelengths, that may also be indicative of magnetic spots, as seen in starspot variability on M dwarfs (Rockenfeller et al. 2006).

As shown in Figure 8, our auroral model did not fit the shape of the observed spectral variability. In Figure 9 we show the contribution function for a  $T_{eff} = 1900\ \text{K}$ ,  $\text{Log}(g) = 5$ ,  $F_{sed} = 2$ , solar metallicity and solar C/O ratio atmosphere, which closely match the fundamental parameters derived for 2M0036+18 by



**Fig. 8.** Spectral variability amplitude of 2M0036+18 (epoch 2) (black line with uncertainties in grey) with the aurora modelled by the temperature inversion method. This model is for a 350K temperature inversion at different peak pressures, with a width of 1 pressure scale height (10–0.1 mbar) and a  $\Delta A = 0.1$ . The models are shown by the coloured lines. As the peak pressure of the inversion is placed at higher pressures, lower in the atmosphere, the predicted variability amplitude in the water band increases. However, there is very little variability outside the water band, no matter the temperature inversion or peak pressure value.



**Fig. 9.** Contribution function for a  $T_{\text{eff}} = 1900\text{K}$ ,  $\log(g) = 5$ ,  $f_{\text{sed}} = 2$ , solar metallicity, and solar C/O ratio atmosphere. The pressures probed by the HST wavelengths (1.1–1.67  $\mu\text{m}$ ), highlighted by the vertical white lines, probe deeper than longer wavelengths that are accessible to JWST. Observations at longer wavelengths that probe the upper atmosphere may be more sensitive to auroral signatures.

Filippazzo et al. (2015). We use the Filippazzo et al. (2015) parameters for the contribution function instead of our own from the HST data, as their parameters were derived from analysis of the spectral energy distribution over a broader wavelength range, including SPEX-SXD (Rayner et al. 2009) and Spitzer IRS spectra (Cushing et al. 2006). The contribution shows that the majority of the HST flux (wavelengths of 1.1–1.67  $\mu\text{m}$ ) originates between 1–10 bars. In contrast, at longer wavelengths, the flux originates at shallower pressures, even probing the millibar pressures where temperature inversions have been detected, which may be more sensitive to auroral chemistry (Faherty et al. 2024). This is particularly noticeable from  $\sim 4.5$ – $5.5$   $\mu\text{m}$  and from  $\sim 7.5$ – $8.5$   $\mu\text{m}$ . Observations at these wavelengths are more sensitive to auroral effects as they probe depths that are associated with a temperature inversion (Faherty et al. 2024). We note

that the wavelengths that reproduce the variability in our auroral model in Figure 8,  $\sim 1.32$ – $1.50$   $\mu\text{m}$ , is the only region of the contribution function that probes higher in the atmosphere.

When probing the upper atmosphere at longer wavelengths, we are more sensitive to the potential effects of auroral chemistry, particularly CO as it is the main opacity source from  $\sim 4.5$ – $5.5$   $\mu\text{m}$  for early L dwarfs (Adams et al. 2023) and  $\text{H}_3\text{O}^+$  from  $\sim 7.5$ – $8.5$   $\mu\text{m}$ .  $\text{H}_3\text{O}^+$  has been postulated as a potential auroral tracer, due to  $\text{H}_3^+$ , a known auroral product in our solar system, reacting with  $\text{H}_2\text{O}$  (Helling & Casewell 2014; Gibbs & Fitzgerald 2022; Pineda et al. 2024).  $\text{H}_3\text{O}^+$  has a longer dynamical timescale than  $\text{H}_3^+$ , which may lead it to being observable (Helling & Rimmer 2019). These wavelengths are accessible with the JWST, in particular the NIRSpec G395H and MIRI MRS instruments. High-resolution observations are more sensitive to detecting auroral emission in the spectrum, as shown in Faherty et al. (2024).

## 6.2. Comparison to previous observations: Long-term stability

L/T transition objects are typically variable on rotational as well as longer-term timescales (Apai et al. 2013; Zhou et al. 2022; Biller et al. 2024; Fuda et al. 2024; Chen et al. 2025). Their light curves have been observed to evolve significantly from one rotation to another. However, this is not necessarily the case at all spectral types. There are few examples of variability remaining stable on longer timescales. These largely stem from the Spitzer variability surveys of Metchev et al. (2015) and Vos et al. (2022). These include the objects, 2M0030–14, 2M0031+57, 2M0718–64, 2M0809+44 and 2M2228–43, which span the spectral types L6–T6. 2M1047+21, a T6.5 dwarf with an infrared period of 1.74 hours also displayed stable variability for over seven rotations with Spitzer (Allers et al. 2020). These objects all displayed variability that did not evolve significantly over an observation duration of at least five rotations each. These objects all have infrared periods between 1.08–1.64 hours, except for the L6 dwarf 2M0030–14, which has a period of  $4.22 \pm 0.02$  hours. These objects typically rotate more rapidly than the three in our sample.

Long-term stability in the light curve structure suggests that long-lived features are responsible for the variability (Gizis et al. 2015; Croll et al. 2016). In order to assess whether our objects show long-term light curve evolution or display long-term stability, we compared our results to light curves published in the literature. In our unique case of two sets of observations for 2M0036+18, we compare the two epochs directly. For 2M1721+33 we compare our observations with Spitzer 3.6  $\mu\text{m}$  observations from Metchev et al. (2015). For 2M1906+40, we discuss the similarity between our observations and the Kepler light curves from Gizis et al. (2013).

We analysed our unique case of two HST observations of 2M0036+18. Despite both observations being just over 16 months apart (Table 2), both the light curves (Figure 1) and spectral variability amplitude (Figure 7) appear very similar. The observations cover slightly different phases of rotation, as can be seen in the phase-folded light curves in Figure 4. The spectral variability amplitude of both observations are very similar too, as shown in Figure 7. The second observation has a slightly higher spectral variability amplitude from 1.1–1.25  $\mu\text{m}$ , but this may be attributed to the fact that this epoch probes the peak of the light curve amplitude more than the first epoch. Due to gaps in our observations caused by HST’s low Earth orbit, we did not achieve full phase coverage with either observation of 2M0036+18. Therefore, when calculating the spectral variability

amplitude, the brightest and dimmest 5% of spectra originate from different phases of 2M0036+18 in our two epochs. The light curve amplitude of both observations are also consistent with each other within uncertainties,  $1.41^{+0.16}_{-0.02}\%$  and  $1.36^{+0.11}_{-0.01}\%$  for the first and second epochs, respectively. Croll et al. (2016) also found that the light curves of 2M0036+18 were stable across observations spanning 122 days.

2M1721+33 was previously observed with Spitzer using the 3.6  $\mu\text{m}$  filter (Metchev et al. 2015). We show the 3.6  $\mu\text{m}$  light curve in Figure 2, along with the Metchev et al. (2015) fit to the light curve and our *celerité*2 Gaussian process fit for this Spitzer light curve. As discussed in Section 3.4, the 2M1721+33 light curve contains two troughs, one deeper than the other. This shape is consistent across both epochs and wavelengths; however, the amplitudes differ. The Spitzer 3.6  $\mu\text{m}$  light curve has an amplitude of  $0.33 \pm 0.07\%$ , while the HST light curve has an amplitude of  $0.53^{+0.03}_{-0.01}\%$ . Yet, we do not expect the amplitude or the light curve shape to be the same at different wavelengths as they probe different depths in the atmosphere (Vos et al. 2020). The Spitzer observations were taken in 2011, nine years before the HST observations from this program. Much like 2M1906+40, this consistency in the light curve nine years apart suggests long-lived surface features. In this case, potentially two features causing a double trough shape in the light curve. We note that since these observations were taken at different wavelengths, they probe different pressure levels (e.g. Buenzli et al. 2012, 2014; McCarthy et al. 2025). However, it is interesting to note the shared similarities across both time and wavelength in the light curve shape.

Although our observations of 2M1906+40 covered less than one rotation, it was observed by the Kepler Space Telescope for 887 days (Gizis et al. 2013, 2015). The Kepler light curve showed a sinusoidal shape over the entire observation duration. This shape is consistent across both the Kepler and HST observations, taking place from 2011–2013 and 2020, respectively. 2M1906+40 was also observed with the Spitzer 3.6  $\mu\text{m}$  and 4.5  $\mu\text{m}$  channels for 20 hours on 17–18 October 2013 (Gizis et al. 2015). This study found that the 3.6  $\mu\text{m}$  light curve also exhibits stable sinusoidal variability, although variability was not detected at 4.5  $\mu\text{m}$ . The amplitude of the light curves in each wavelength range are different, however. The optical Kepler light curve had an amplitude of 1.5% (Gizis et al. 2015) the HST light curve from this work has an amplitude of  $0.64^{+0.03}_{-0.02}\%$  and the Spitzer 3.6  $\mu\text{m}$  light curve has an amplitude of 1.1%. As with 2M1721+33, we note that none of these amplitudes were measured at the same epoch, nor probe the same wavelength range, yet all share a common light curve shape.

The fact that all three of our objects display stable light curves across many years implies that there may be a different regime of variability in the early L dwarfs. L/T transition objects such as SIMP 0136, VHS 1256 b, and WISE 1049AB have variability that evolves from one rotation to the next (Zhou et al. 2022; Biller et al. 2024; McCarthy et al. 2025). These objects have been observed across multiple epochs with both HST and JWST and they each show changes in their overall light curve amplitude as well as the shapes of their light curves. This light curve evolution may be due to dynamical evolution of clouds, carbon chemistry or hotspots on their surfaces (McCarthy et al. 2025). In contrast, the early L dwarfs presented in this work appear somewhat stable over timescales of years.

As all three of our objects have shown evidence for little light curve evolution over many years, this suggests that long-lived features may be present on all three objects. Our results show that the variability is likely caused by clouds of different properties

or a magnetic spot. Future observations with telescopes such as JWST, across both the near and mid-infrared, would provide insight into whether clouds, magnetic spots, and aurorae drive variability across the full infrared spectrum. This would allow light curve comparison across multiple epochs in order to further analyse the light curve evolution of these objects. An observation with NIRSpec PRISM, for example, would obtain a light curve from 1–5  $\mu\text{m}$  and provide comparisons with the HST light curves from 1.1–1.67  $\mu\text{m}$  as well as the Spitzer 3.6  $\mu\text{m}$  and 4.5  $\mu\text{m}$  light curves.

### 6.3. Implications for directly imaged exoplanets

While spectroscopic variability observations can be conducted on a large number of free floating brown dwarfs (e.g. Buenzli et al. 2014; Radigan et al. 2014a; Metchev et al. 2015; Vos et al. 2022), variability studies of directly imaged exoplanets have been limited by the presence of a bright host star. Apai et al. (2016) and Biller et al. (2021) searched for variability in the HR8799 system of exoplanets with the Very Large Telescope’s SPHERE instrument; however, neither study was able to confidently detect variability. This illustrates the challenge of variability observations from even the best ground-based telescopes. There have also been attempts to observe variability with space-based telescopes. Zhou et al. (2019) observed three planetary mass companions with HST, but only found tentative detections of photometric variability. Zhou et al. (2020b) observed the directly imaged exoplanet HD 106906b with HST in the F127M, F139M and F153M filters. Tentative variability was detected with in the F127M light curve but the other two filters detected no variability. These results illustrate the difficulty of obtaining photometric variability of directly imaged exoplanets.

Recent observations with JWST have revealed variability in the directly imaged planetary mass companion 2M1207 b (Adams et al. 2025). Other ongoing programs, such as Zhou et al. (2024), will search for variability in directly imaged exoplanets such as  $\beta$  pic b with JWST.

Brown dwarfs can be considered as analogues for directly imaged exoplanets. There are a number of directly imaged exoplanets that share similar temperatures and spectral types to the early L dwarfs presented in this paper, such as  $\beta$  pic b, RXS 1609 b, and AB Pic b (Lagrange et al. 2009; Bowler et al. 2013; Palma-Bifani et al. 2023). At these temperatures the effects of clouds, aurorae and magnetic spots may all be important. By studying the variability of early L dwarfs, we can establish the dominant processes in extrasolar atmospheres at these temperatures without the need for high-contrast imaging. Early L dwarfs such as 2M1721+33, 2M0036+18, and 2M1906+40 are analogues for directly imaged exoplanets such as those listed above.

Our results indicate that early L dwarfs are variable and exhibit high-variability amplitudes. Clouds and magnetic spots are the most plausible variability mechanisms for early L dwarfs from 1.1–1.67  $\mu\text{m}$ . Therefore, it is likely that directly imaged exoplanets with similar effective temperatures and spectral types are variable. Our work shows that this variability is likely driven by condensate clouds and/or magnetic spots at near-IR wavelengths (1.1–1.67  $\mu\text{m}$ ), but auroral variability may become important at longer wavelengths. Our spectroscopic variability results help put photometric variability of directly imaged exoplanets into context. Future high-contrast variability searches with JWST, VLT, and ELT are likely to reveal variability in directly imaged exoplanets with early L spectral types.

## 7. Conclusions

In this work, we presented HST/WFC3 variability observations of three early L dwarfs, 2M1721+33, 2M0036+18, and 2M1906+40. We used a modelling framework that combines self-consistent forward models and atmospheric retrievals to model the spectral variability amplitude of each object for three potential cases: cloud-driven variability, auroral variability, and magnetic spot-driven variability. We were able to assess the likely drivers of variability in each object:

1. All three brown dwarfs show significant variability across all wavelengths, with white light curve amplitudes of  $0.64^{+0.03}_{-0.02}\%$ ,  $1.41^{+0.16}_{-0.02}\%$ , (epoch 1),  $1.36^{+0.11}_{-0.01}\%$  (epoch 2), and  $0.53^{+0.03}_{-0.01}\%$ , for 2M1721+33, 2M0036+18, and 2M1906+40, respectively;
2. We report a new period of  $4.93^{+0.36}_{-0.22}$  hours for 2M1721+33 from our Gaussian process fit to our new HST and archival Spitzer light curves. This resulted in this program not capturing a fully phase-resolved light curve. Our measured periods for 2M0036+18 and 2M1906+40 agree with the literature values;
3. The colour modulation of our objects extend the trend discussed in [Lew et al. \(2020b\)](#), where early L dwarfs become brighter and bluer as they rotate. However, this changes for the L1 target 2M1906+40, which has near grey modulations;
4. We developed a modelling framework by combining self-consistent forward models and atmospheric retrievals to reproduce the spectral variability amplitude of brown dwarfs for variability that is driven by changing clouds, aurorae, and magnetic spots. In the future, this can be applied to other objects with observations at different wavelengths to model their spectral variability amplitude;
5. Our analysis found that either clouds or magnetic spots are the plausible drivers of variability in the early L dwarfs 2M1721+33, 2M0036+18, and 2M1906+40. Aurora can be ruled out as a driver of variability at the measured wavelengths of 1.1–1.67  $\mu\text{m}$ . However, future observations at longer wavelengths are likely more sensitive to potential auroral variability as they probe shallower atmospheric depths;
6. All three of our objects have shown evidence of long-term stability of their light curves. In the case of long-term light curve stability, the variability is likely driven by a long-lived feature. This may be a cloudy spot or region, or a stellar-like magnetic spot. Future observations across the near-infrared and mid-infrared are required in order to disentangle this further.

This work presented a method to disentangle the drivers of variability in early L-dwarf atmospheres based on their spectral variability behaviour. Future studies with instruments such as NIRSpec G395H and MIRI MRS on board JWST would allow us to further characterise these processes. At longer wavelengths, we can search for variability in the 10  $\mu\text{m}$  silicate feature to disentangle whether the long-lived features on these objects are cloudy or magnetic in nature. Searches for aurora on these worlds would benefit from observations that probe higher in the atmosphere, at longer wavelengths, from 4.5–5.5  $\mu\text{m}$  and 7.5–8.5  $\mu\text{m}$ , to detect species that may be affected by auroral heating. This work has provided a flexible modelling framework that can be applied to any L, T, or Y dwarf to determine what the likely drivers of variability are across different wavelengths and resolutions. This provides an insight into the likely variable nature of directly imaged exoplanets that will be observed with the next generation of telescopes.

*Acknowledgements.* We would like to thank the referee for their helpful comments which helped improve the quality of this paper. We would like to thank Ben Lew for their guidance on fitting the colour modulation using the ODR method. We would like to thank Caroline Morley for their guidance with the Sonora Diamondback contribution function. CO'T, JMV and EN acknowledge support from a Royal Society – Research Ireland University Research Fellowship (URF/1/221932, RF/ERE/221108). MS acknowledges support from Trinity College Dublin via a Trinity Research Doctoral Award. This research is based on observations made with the NASA/ESA Hubble Space Telescope obtained from the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5–26555. These observations are associated with program 15924.

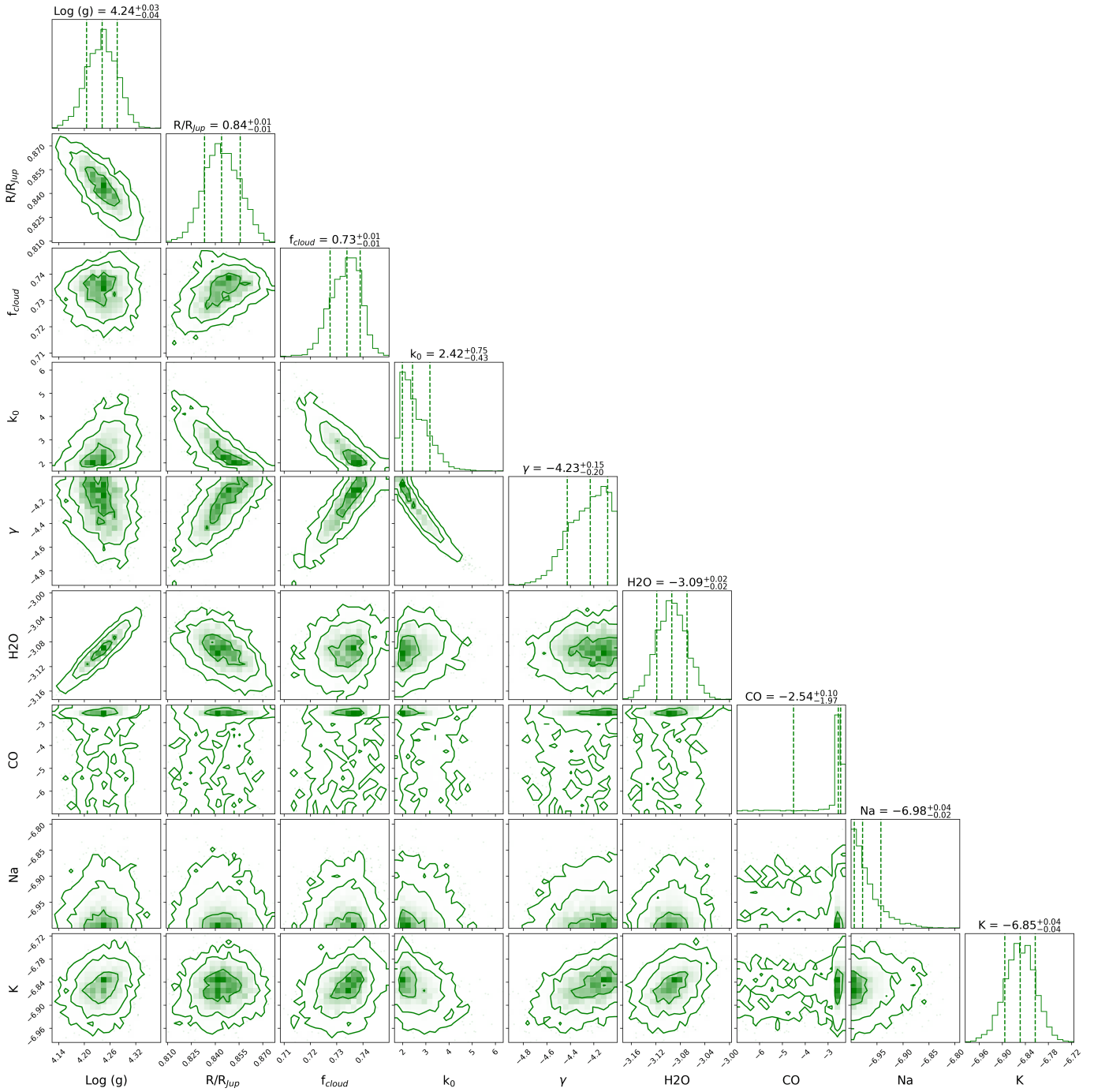
## References

- Adams, A. D., Meyer, M. R., Howe, A. R., et al. 2023, *AJ*, **166**, 192  
 Adams, A. D., Zhou, Y., Marleau, G.-D., et al. 2025, *AJ*, **170**, 289  
 Afram, N., & Berdyugina, S. V. 2015, *A&A*, **576**, A34  
 Allard, N. F., Spiegelman, F., Leininger, T., & Molliere, P. 2019, *A&A*, **628**, A120  
 Allers, K. N., Vos, J. M., Biller, B. A., & Williams, P. K. G. 2020, *Science*, **368**, 169  
 Apai, D., Radigan, J., Buenzli, E., et al. 2013, *ApJ*, **768**, 121  
 Apai, D., Kasper, M., Skemer, A., et al. 2016, *ApJ*, **820**, 40  
 Apai, D., Karalidi, T., Marley, M. S., et al. 2017, *Science*, **357**, 683  
 Astropy Collaboration (Robitaille, T. P., et al.) 2013, *A&A*, **558**, A33  
 Astropy Collaboration (Price-Whelan, A. M., et al.) 2018, *AJ*, **156**, 123  
 Astropy Collaboration (Price-Whelan, A. M., et al.) 2022, *ApJ*, **935**, 167  
 Atreya, S. K., Mahaffy, P. R., Niemann, H. B., Wong, M. H., & Owen, T. C. 2003, *Planet. Space Sci.*, **51**, 105  
 Barnes, J. R., Jeffers, S. V., Jones, H. R. A., et al. 2015, *ApJ*, **812**, 42  
 Barros, S. C. C., Demangeon, O., Díaz, R. F., et al. 2020, *A&A*, **634**, A75  
 Baylor, R. N., Cassak, P. A., Christe, S., et al. 2011, *ApJ*, **736**, 75  
 Ben-Jaffel, L., & Abbes, I. 2015, in *Journal of Physics Conference Series*, 577, *Journal of Physics Conference Series* (IOP), 012003  
 Berger, E. 2002, *ApJ*, **572**, 503  
 Best, W. M. J., Dupuy, T. J., Liu, M. C., et al. 2024, *The UltracoolSheet: Photometry, Astrometry, Spectroscopy, and Multiplicity for 4000+ Ultracool Dwarfs and Imaged Exoplanets (2.0.1)*  
 Biller, B. A., Apai, D., Bonnefoy, M., et al. 2021, *MNRAS*, **503**, 743  
 Biller, B. A., Vos, J. M., Zhou, Y., et al. 2024, *MNRAS*, **532**, 2207  
 Bowler, B. P., Liu, M. C., Shkolnik, E. L., & Dupuy, T. J. 2013, *ApJ*, **774**, 55  
 Bowler, B. P., Zhou, Y., Morley, C. V., et al. 2020, *ApJ*, **893**, L30  
 Brammer, G., Ryan, R., & Pirzkal, N. 2015, *Source-dependent master sky images for the WFC3/IR grisms*, Instrument Science Report WFC3 2015-17, 18  
 Buchner, J., Georgakakis, A., Nandra, K., et al. 2014, *A&A*, **564**, A125  
 Buenzli, E., Apai, D., Morley, C. V., et al. 2012, *ApJ*, **760**, L31  
 Buenzli, E., Apai, D., Radigan, J., Reid, I. N., & Flateau, D. 2014, *ApJ*, **782**, 77  
 Burningham, B., Marley, M. S., Line, M. R., et al. 2017, *MNRAS*, **470**, 1177  
 Burningham, B., Faherty, J. K., Gonzales, E. C., et al. 2021, *MNRAS*, **506**, 1944  
 Burrows, A., Sudarsky, D., & Hubeny, I. 2006, *ApJ*, **640**, 1063  
 Carnall, A. C. 2017, *SpectRes: A Fast Spectral Resampling Tool in Python*, [arXiv:1705.05165](https://arxiv.org/abs/1705.05165) [astro-ph.IM]  
 Chapleski, M. F., & Zhou, Y. 2026, *ApJ*, **997**, 152  
 Chen, X., Biller, B. A., Tan, X., et al. 2025, *MNRAS*, **539**, 3758  
 Croll, B., Muirhead, P. S., Han, E., et al. 2016, *arXiv e-prints* [[arXiv:1609.03586](https://arxiv.org/abs/1609.03586)]  
 Cruz, K. L., Reid, I. N., Liebert, J., Kirkpatrick, J. D., & Lowrance, P. J. 2003, *AJ*, **126**, 2421  
 Cushing, M. C., Roellig, T. L., Marley, M. S., et al. 2006, *ApJ*, **648**, 614  
 De Luca, A., Stelzer, B., Burgasser, A. J., et al. 2020, *A&A*, **634**, L13  
 Faherty, J. K., Riedel, A. R., Cruz, K. L., et al. 2016, *ApJS*, **225**, 10  
 Faherty, J. K., Burningham, B., Gagné, J., et al. 2024, *Nature*, **628**, 511  
 Feroz, F., Hobson, M. P., & Bridges, M. 2009, *MNRAS*, **398**, 1601  
 Feroz, F., Hobson, M. P., Cameron, E., & Pettitt, A. N. 2013, *Open J. Astrophys.*, **2**, 10  
 Filippazzo, J. C., Rice, E. L., Faherty, J., et al. 2015, *ApJ*, **810**, 158  
 Foreman-Mackey, D. 2018, *RNAAS*, **2**, 31  
 Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, *PASP*, **125**, 306  
 Foreman-Mackey, D., Agol, E., Ambikasaran, S., & Angus, R. 2017, *AJ*, **154**, 220  
 Fuda, N., Apai, D., Nardiello, D., et al. 2024, *ApJ*, **965**, 182  
 Gaia Collaboration (Brown, A. G. A., et al.) 2021, *A&A*, **649**, A1  
 Gibbs, A., & Fitzgerald, M. P. 2022, *AJ*, **164**, 63  
 Gizis, J. E., Troup, N. W., & Burgasser, A. J. 2011, *ApJ*, **736**, L34

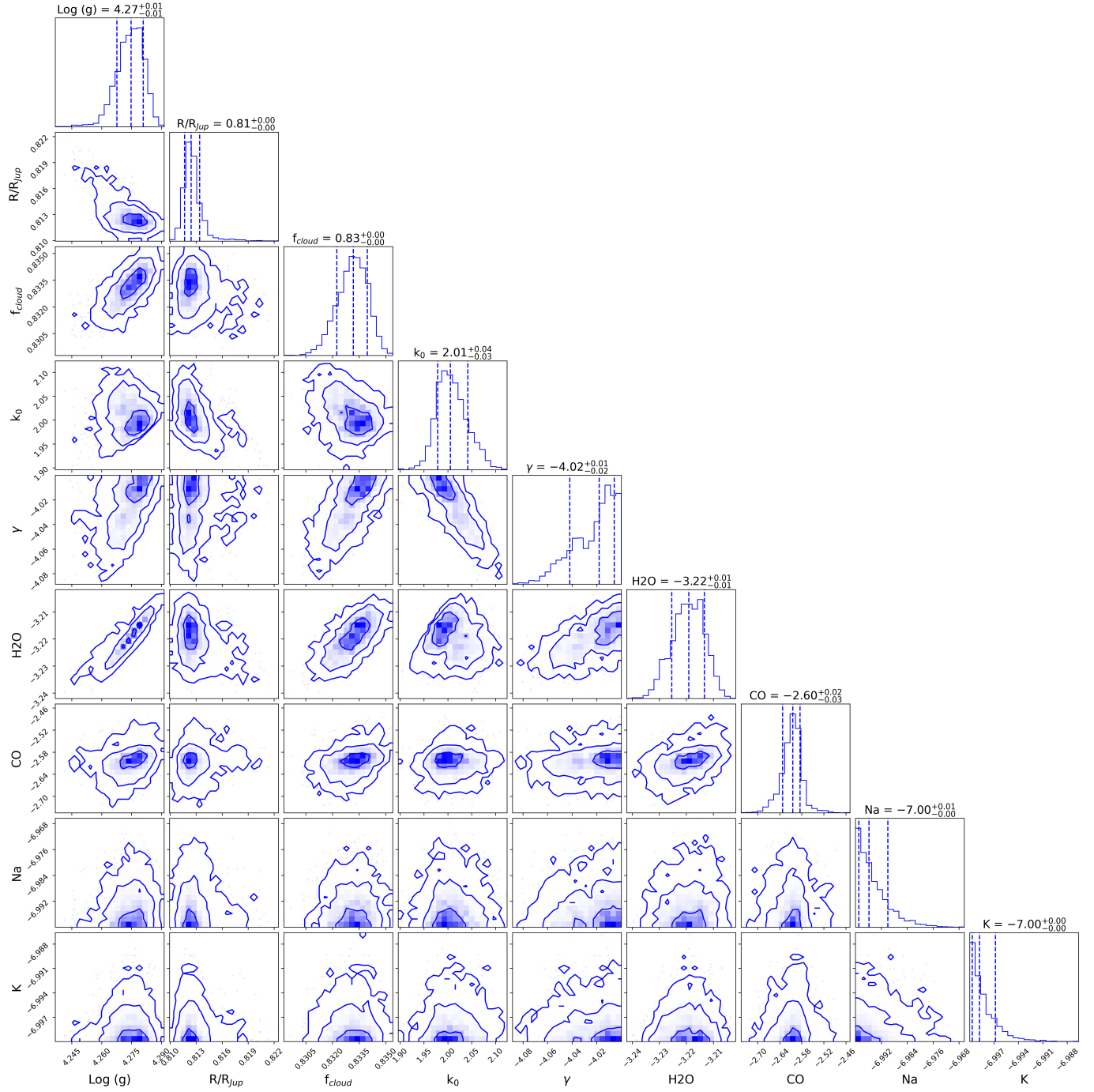
- Gizis, J. E., Burgasser, A. J., Berger, E., et al. 2013, *AJ*, 779
- Gizis, J. E., Dettman, K. G., Burgasser, A. J., et al. 2015, *ApJ*, 813, 104
- Hallinan, G., Antonova, A., Doyle, J. G., et al. 2008, *ApJ*, 684, 644
- Hallinan, G., Littlefair, S. P., Cotter, G., et al. 2015, *Nature*, 523, 568
- Harding, L. K., Hallinan, G., Boyle, R. P., et al. 2013, *ApJ*, 779, 101
- Helling, C., & Casewell, S. 2014, *A&A Rev.*, 22, 80
- Helling, C., & Rimmer, P. B. 2019, *Philos. Trans. Roy. Soc. Lond. A*, 377, 20180398
- Kao, M. M., Hallinan, G., Pineda, J. S., et al. 2016, *ApJ*, 818, 24
- Karalidi, T., Apai, D., Marley, M. S., & Buzzi, E. 2015, *ApJ*, 825, 90
- Kümmel, M., Walsh, J. R., Pirzkal, N., Kuntschner, H., & Pasquali, A. 2009, *PASP*, 121, 59
- Lagrange, A. M., Gratadour, D., Chauvin, G., et al. 2009, *A&A*, 493, L21
- Lane, C., Hallinan, G., Zavala, R. T., et al. 2007, *ApJ*, 668, L163
- Liu, M. C., Dupuy, T. J., & Allers, K. N. 2016, *ApJ*, 833, 96
- Lew, B. W. P., Apai, D., Marley, M., et al. 2020a, *ApJ*, 903, 15
- Lew, B. W. P., Apai, D., Zhou, Y., et al. 2020b, *AJ*, 159, 125
- Manjavacas, E., Apai, D., Zhou, Y., et al. 2018, *AJ*, 155, 11
- McCarthy, A. M., Muirhead, P. S., Tamburo, P., et al. 2024, *ApJ*, 965, 83
- McCarthy, A. M., Vos, J. M., Muirhead, P. S., et al. 2025, *ApJ*, 981, L22
- McLean, M., Berger, E., Irwin, J., Forbrich, J., & Reiners, A. 2011, *ApJ*, 741, 27
- Metchev, S. A., Heinze, A., Apai, D., et al. 2015, *ApJ*, 799, 154
- Miles-Páez, P. A. 2021, *A&A*, 651, L7
- Mollière, P., Wardenier, J. P., van Boekel, R., et al. 2019, *A&A*, 627, A67
- Morley, C. V., Marley, M. S., Fortney, J. J., & Lupu, R. 2014, *ApJ*, 789
- Morley, C. V., Mukherjee, S., Marley, M. S., et al. 2024, *ApJ*, 975, 59
- Nasedkin, E., Mollière, P., & Blain, D. 2024, *J. Open Source Softw.*, 9, 5875
- Nasedkin, E., Schrader, M., Vos, J. M., et al. 2025, *A&A*, 702, A1
- O'Donoghue, J., Moore, L., Bhakyaipabul, T., et al. 2021, *Nature*, 596, 54
- Oliveros-Gomez, N., Manjavacas, E., Karalidi, T., et al. 2026, *ApJ*, 997, 136
- Palma-Bifani, P., Chauvin, G., Bonnefoy, M., et al. 2023, *A&A*, 670, A90
- Paudel, R. R., Gizis, J. E., Mullan, D. J., et al. 2020, *MNRAS*, 494, 5751
- Pineda, J. S., Hallinan, G., Desert, J.-M., & Harding, L. K. 2024, *ApJ*, 966, 58
- Polyansky, O. L., Kyuberis, A. A., Zobov, N. F., et al. 2018, *MNRAS*, 480, 2597
- Rackham, B. V., Apai, D., & Giampapa, M. S. 2018, *ApJ*, 853, 122
- Radigan, J. 2014b, *ApJ*, 797, 120
- Radigan, J., Lafrenière, D., Jayawardhana, R., & Artigau, E. 2014a, *ApJ*, 793, 75
- Rayner, J. T., Cushing, M. C., & Vacca, W. D. 2009, *ApJS*, 185, 289
- Reid, I. N., Kirkpatrick, J. D., Gizis, J. E., et al. 2000, *AJ*, 119, 369
- Richey-Yowell, T., Kao, M. M., Pineda, J. S., Shkolnik, E. L., & Hallinan, G. 2020, *ApJ*, 903, 74
- Rockefeller, B., Bailer-Jones, C. A. L., & Mundt, R. 2006, *A&A*, 448, 1111
- Roettenbacher, R. M., Monnier, J. D., Korhonen, H., et al. 2017, *ApJ*, 849, 120
- Rothman, L. S., Gordon, I. E., Barber, R. J., et al. 2010, *J. Quant. Spec. Radiat. Transf.*, 111, 2139
- Sanghi, A., Liu, M. C., Best, W. M. J., et al. 2023, *ApJ*, 959, 63
- Schmidt, S. J., Hawley, S. L., West, A. A., et al. 2015, *AJ*, 149, 158
- Stolker, T., Quanz, S. P., Todorov, K. O., et al. 2020, *A&A*, 635, A182
- Suárez, G., & Metchev, S. 2022, *MNRAS*, 513, 5701
- Tremblin, P., Padioleau, T., Phillips, M. W., et al. 2019, *ApJ*, 876, 144
- Tremblin, P., Phillips, M. W., Emery, A., et al. 2020, *A&A*, 643, A23
- Vos, J. M., Allers, K. N., Biller, B. A., et al. 2018, *MNRAS*, 474, 1041
- Vos, J. M., Biller, B. A., Allers, K. N., et al. 2020, *AJ*, 160, 38
- Vos, J. M., Faherty, J. K., Gagné, J., et al. 2022, *ApJ*, 924, 68
- Vos, J. M., Burningham, B., Faherty, J. K., et al. 2023, *ApJ*, 944, 138
- Wilson, P. A., Rajan, A., & Patience, J. 2014, *A&A*, 566, A111
- Yang, H., Apai, D., Marley, M. S., et al. 2015, *ApJ*, 798, L13
- Zhou, Y., Apai, D., Lew, B. W. P., & Schneider, G. H. 2017, *AJ*, 153, 243
- Zhou, Y., Apai, D., Metchev, S., et al. 2018, *AJ*, 155, 132
- Zhou, Y., Apai, D., Lew, B. W. P., et al. 2019, *AJ*, 157, 128
- Zhou, Y., Bowler, B. P., Morley, C. V., et al. 2020a, *AJ*, 160, 77
- Zhou, Y., Apai, D., Bedin, L. R., et al. 2020b, *AJ*, 159, 140
- Zhou, Y., Bowler, B. P., Apai, D., et al. 2022, *AJ*, 164, 239
- Zhou, Y., Apai, D., Balmer, W., et al. 2024, *From Day to Season: Constraining the Rotation Period and Obliquity of Beta Pic b with Time-resolved High-contrast Imaging*, JWST Proposal. Cycle 3, ID. #4758
- Zuckerman, A., Pineda, J. S., Brain, D., Mang, J., & Morley, C. V. 2026, *ApJ*, 999, 26

**Appendix A: Atmospheric retrieval corner plots**

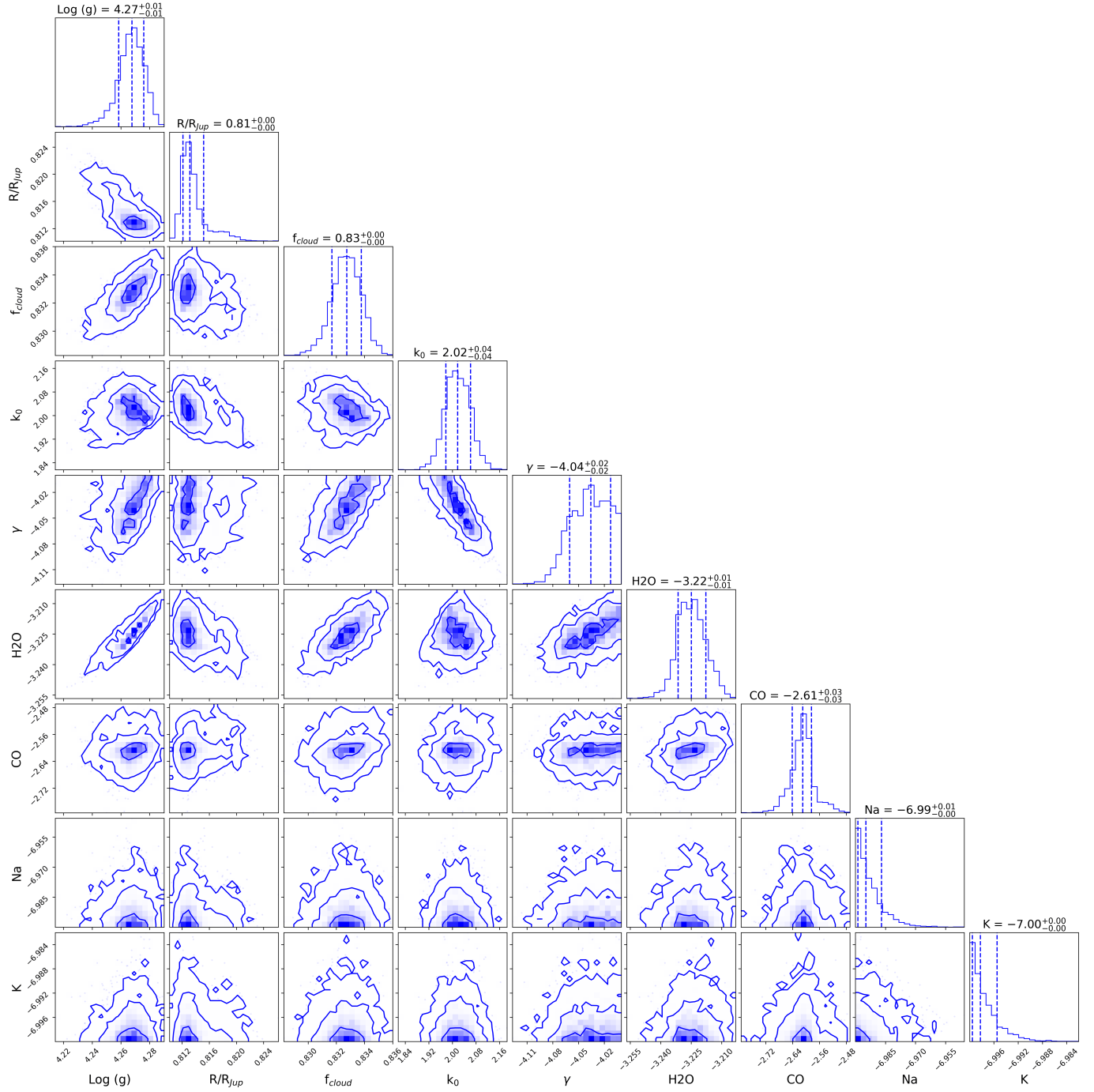
This section of the Appendix displays the corner plots for each of the petitRADTRANS atmospheric retrievals for each observation.



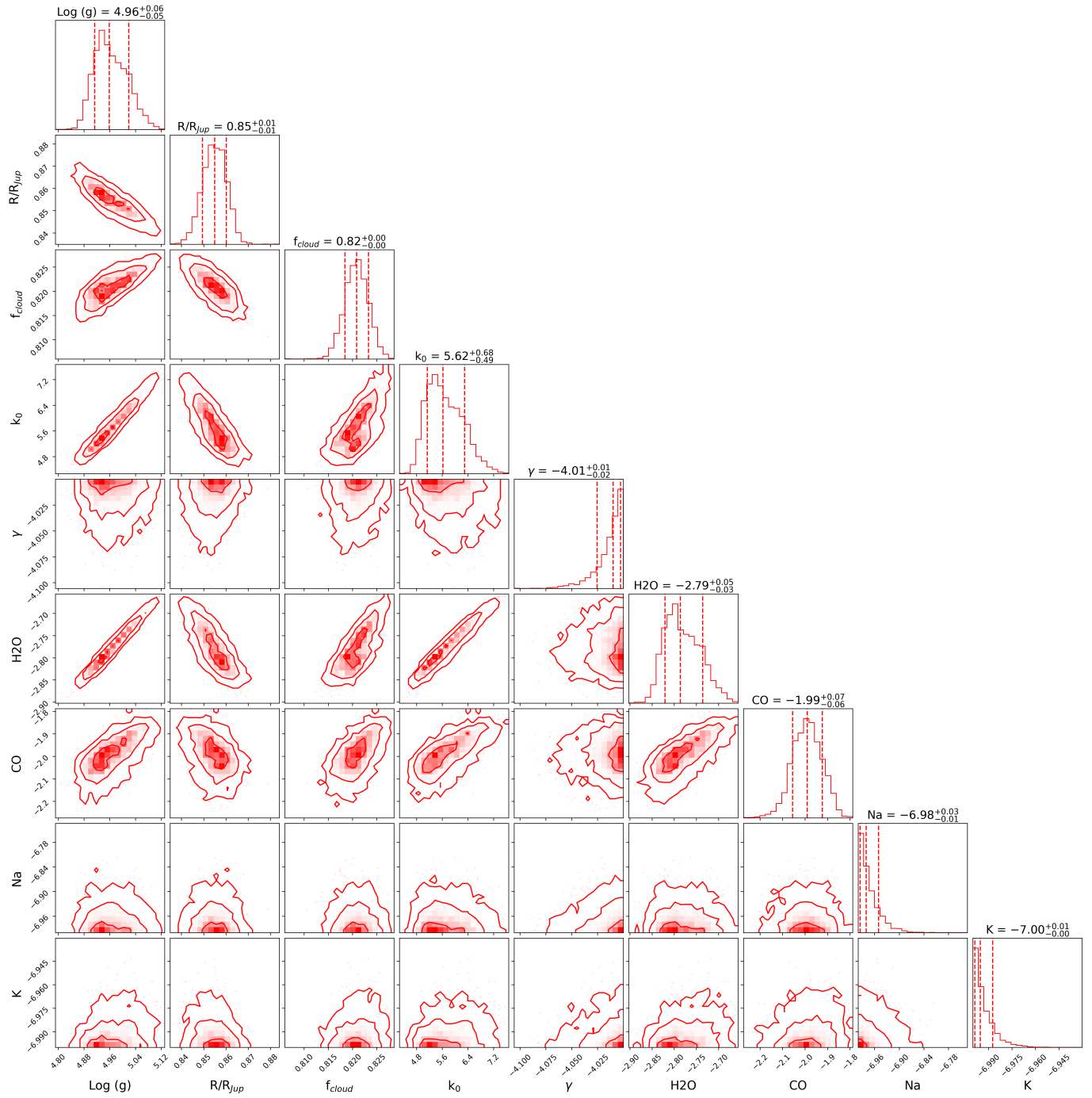
**Fig. A.1.** Posterior distributions for the petitRADTRANS atmospheric retrieval of 2M1721+33.



**Fig. A.2.** Posterior distributions for the petitRADTRANS atmospheric retrieval of 2M0036+18 (epoch 1).



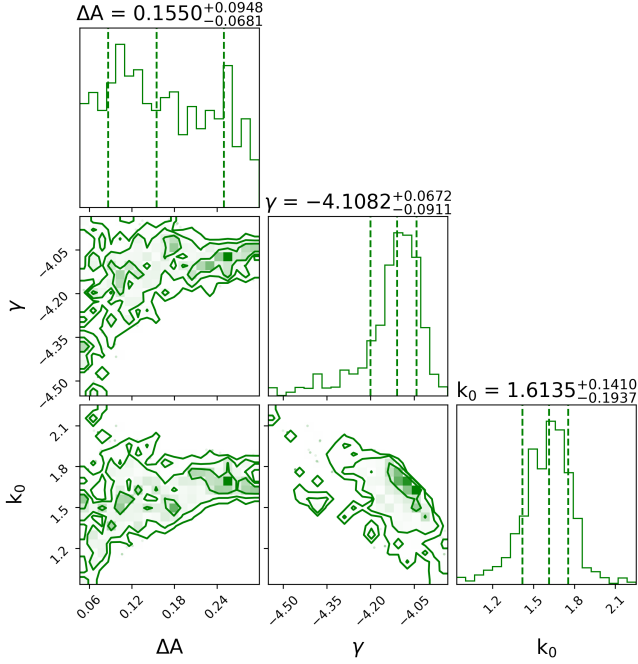
**Fig. A.3.** Posterior distributions for the atmospheric retrieval of 2M0036+18 (epoch 2).



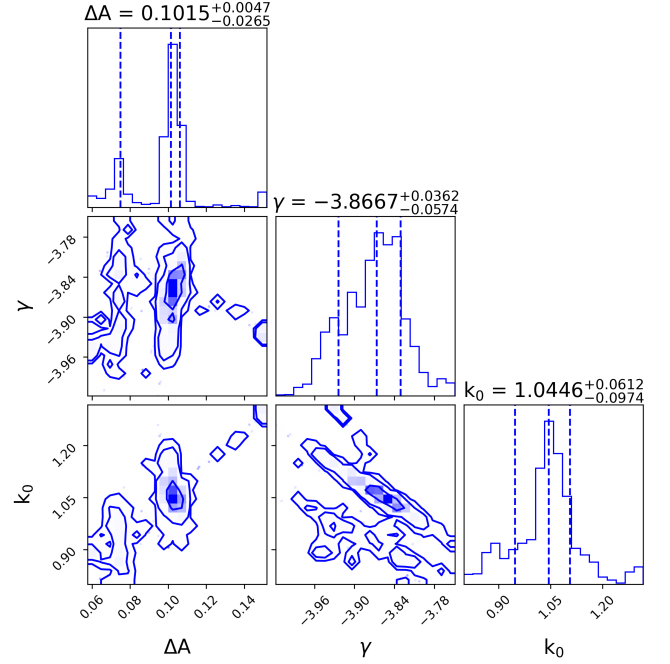
**Fig. A.4.** Posterior distributions for the petitRADTRANS atmospheric retrieval of 2M1906+40.

### Appendix B: Cloud variability model corner plots

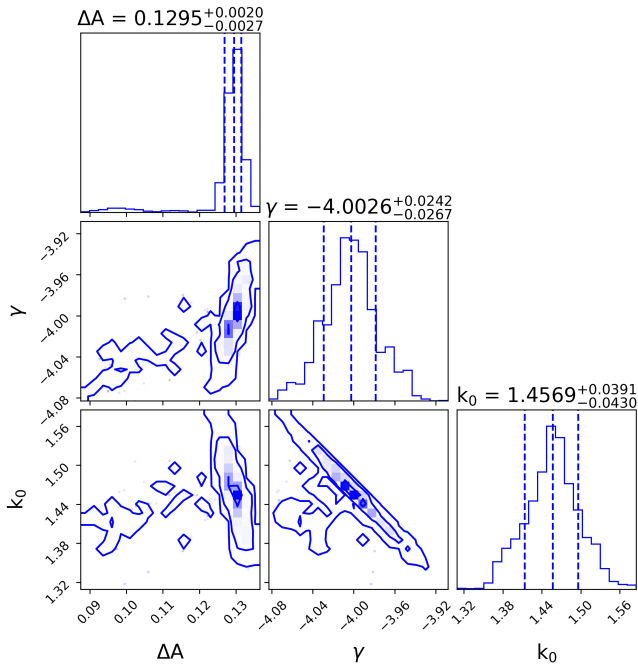
This section of the Appendix displays the corner plots for the cloud driven variability models. We show the posteriors for the  $\Delta A$ ,  $\gamma$  and  $k_0$  values for each observation.



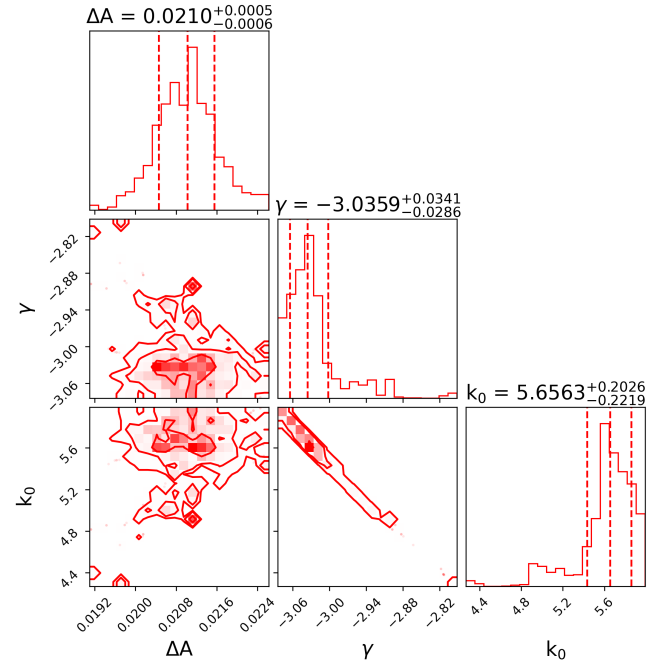
**Fig. B.1.** Posterior distributions of the cloud model parameters  $\Delta A$ ,  $\gamma$ , and  $k_0$  for 2M1721+33.



**Fig. B.3.** Posterior distributions of the cloud model parameters  $\Delta A$ ,  $\gamma$ , and  $k_0$  for 2M0036+18 (epoch 2).



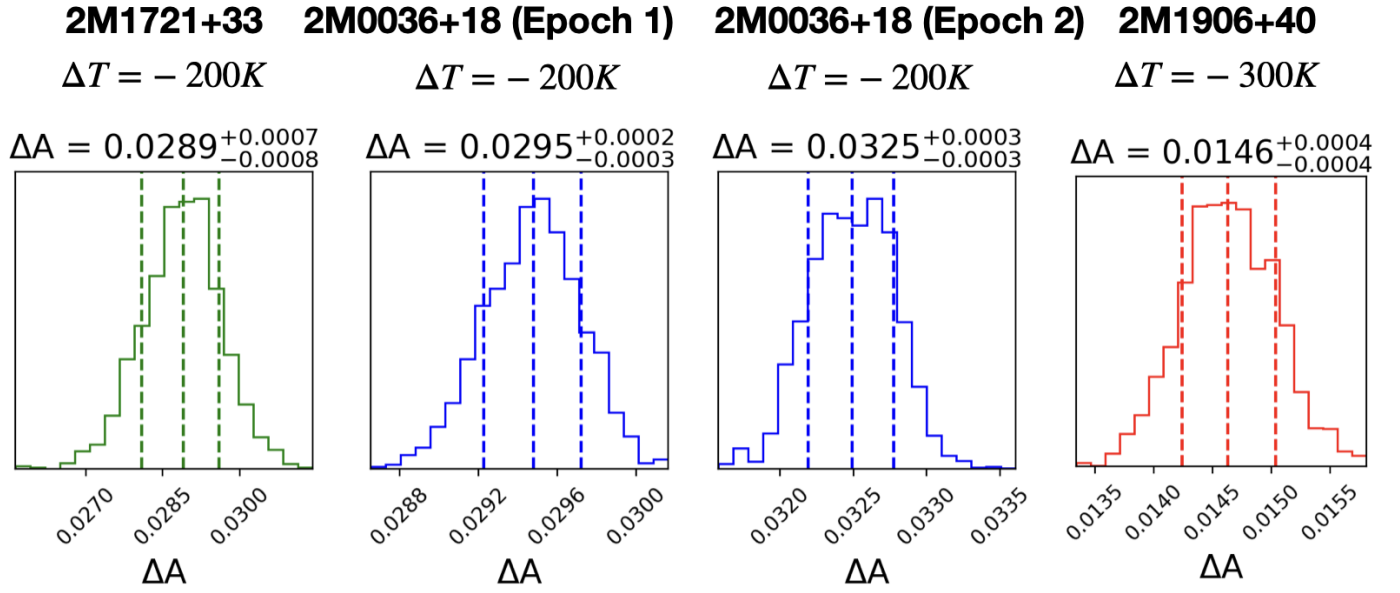
**Fig. B.2.** Posterior distributions of the cloud model parameters  $\Delta A$ ,  $\gamma$ , and  $k_0$  for 2M0036+18 (epoch 1).



**Fig. B.4.** Posterior distributions of the cloud model parameters  $\Delta A$ ,  $\gamma$ , and  $k_0$  for 2M1906+40.

### Appendix C: Magnetic spot variability posterior distribution Plots

This section of the Appendix displays the posterior distributions for the spot driven variability models. We show the posteriors for the  $\Delta A$  value for each observation for the best-fit  $\Delta T$ .



**Fig. C.1.** Posterior distributions of the best-fit magnetic spot model,  $\Delta A$  parameter, for each object.