The optical transmission spectrum of the hot Jupiter HAT-P-32b: clouds explain the absence of broad spectral features?

Published in:
Monthly Notices of the Royal Astronomical Society

Document Version:
Publisher's PDF, also known as Version of record

The optical transmission spectrum of the hot Jupiter HAT-P-32b: clouds explain the absence of broad spectral features?

N. P. Gibson,[1]* S. Aigrain,[2] J. K. Barstow,[2] T. M. Evans,[2] L. N. Fletcher[3]
and P. G. J. Irwin[3]

[1]*European Southern Observatory, Karl-Schwarzschild-Str. 2, D-85748 Garching bei München, Germany*
[2]*Department of Physics, University of Oxford, Denys Wilkinson Building, Keble Road, Oxford OX1 3RH, UK*
[3]*Atmospheric, Oceanic and Planetary Physics, University of Oxford, Clarendon Laboratory, Oxford OX1 3PU, UK*

ABSTRACT

We report Gemini-North Gemini Multi-Object Spectrograph observations of the inflated hot Jupiter HAT-P-32b during two primary transits. We simultaneously observed two comparison stars and used differential spectrophotometry to produce multiwavelength light curves. 'White' light curves and 29 'spectral' light curves were extracted for each transit and analysed to refine the system parameters and produce transmission spectra from 520 to 930 nm in $\approx$14 nm bins. The light curves contain time-varying white noise as well as time-correlated noise, and we used a Gaussian process model to fit this complex noise model. Common mode corrections derived from the white light-curve fits were applied to the spectral light curves which significantly improved our precision, reaching typical uncertainties in the transit depth of $\sim$2 $\times$ 10^{-4}, corresponding to about half a pressure scale height. The low-resolution transmission spectra are consistent with a featureless model, and we can confidently rule out broad features larger than about one scale height. The absence of Na/K wings or prominent TiO/VO features is most easily explained by grey absorption from clouds in the upper atmosphere, masking the spectral features. However, we cannot confidently rule out clear atmosphere models with low abundances ($\sim$10^{-3} solar) of TiO, VO or even metal hydrides masking the Na and K wings. A smaller scale height or ionization could also contribute to muted spectral features, but alone are unable to account for the absence of features reported here.

Key words: methods: data analysis – techniques: spectroscopic – stars: individual: HAT-P-32 – planetary systems – techniques: Gaussian processes..

1 INTRODUCTION

Remarkable progress has been achieved in understanding the diversity of extrasolar systems in our Galaxy, owing to the success of radial velocity and transit surveys. Transiting planets allow both their masses and radii to be measured, leading to bulk densities and a first-order understanding of structure and composition. The next, natural step is to probe the atmospheres of extrasolar planets using spectroscopy, and thereby understand their chemical composition, energy budget and dominant physical processes. Transiting planets allow such measurements by analysis of temporal variations in the light received from the planet and star combined, rather than spatially resolving light from the planet.

Transmission spectroscopy is a measurement of the effective size of the planet as a function of wavelength during primary transit.

The effective size of the planet is sensitive to the height at which the atmosphere becomes opaque to starlight, which in turn depends on the opacities in the atmosphere and is therefore wavelength dependent. Consequently, transmission spectroscopy enables us to probe the composition of planets' atmospheres (Seager & Sasselov 2000; Brown 2001).

Until quite recently, transmission spectroscopy was dominated by space-based observations (e.g. Charbonneau et al. 2002; Pont et al. 2008; Sing et al. 2008, 2011; Berta et al. 2012; Gibson et al. 2012b; Huitson et al. 2012; Deming et al. 2013). However, nearly all attempts to produce transmission (or emission) spectra have to deal with instrumental systematics considerably larger than the atmospheric signature from the planet. We therefore rely on simple (and often arbitrary) models of the systematics to extract spectra from the time series. Experience with the *Hubble Space Telescope* (*HST*)/Near-Infrared Camera and Multi-Object Spectrometer (NIC-MOS) camera shows we should be cautious when interpreting such data (Gibson, Pont & Aigrain 2011); indeed, repeated observations

* E-mail: ngibson@eso.org

© 2013 The Authors
Published by Oxford University Press on behalf of the Royal Astronomical Society

of the same target have confirmed the unreliability of such simple instrument models (Crouzet et al. 2012; Deming et al. 2013). Whilst other instruments have proved to be considerably more reliable (e.g. Pont et al. 2008; Sing et al. 2011; Berta et al. 2012; Deming et al. 2013), and better analysis techniques can remove the dependence on arbitrary systematics models (Gibson et al. 2012a; Waldmann 2012), we should remain cautious in overinterpreting marginal results from single epoch observations, as no simple systematics model can perfectly account for a complex instrument response. The best way to assess the reliability of spectra is to re-observe at the same wavelengths; unfortunately, given the expense of space-based observations this is usually not feasible.

Ground-based observations of transmission spectra have been rapidly advancing in recent years with the adoption of multi-object spectrographs (MOS) to perform differential spectrophotometry (Bean, Miller-Ricci Kempton & Homeier 2010), and offer a powerful and complementary alternative to space-based observations with some noteworthy advantages. First, given the relative cost of observations, it is easier to obtain multiple transit observations. Secondly, we can obtain continuous observations of transit light curves, unlike *HST* which is restricted to observe in ∼46 min blocks corresponding to half of its low Earth orbit. Whilst instrumental systematics remain a problem for ground-based observations, in principle at least, this should allow us to address the systematics more easily and therefore extract more robust spectra. Finally, and arguably most importantly, we can access wavelength regions not obtainable using current space-based instrumentation (e.g. *K* band; Bean et al. 2011). However, ground-based differential observations are always limited by the target's nearby, bright comparison stars and therefore space-based observations are likely to continue providing the strongest constraints on planetary atmospheres through observations of the brightest systems; nonetheless, ground-based observations can provide important tests on the validity of systematics models used for space-based analyses, as well as obtain reliable spectra of fainter targets.

Here we report on the use of Gemini Multi-Object Spectrograph (GMOS) on Gemini-North to observe the transmission spectrum of HAT-P-32b. The GMOS instruments on the Gemini telescopes have been used by Gibson et al. (2013), Stevenson et al. (2013) and Crossfield et al. (2013) to measure the transmission spectra of WASP-29b, WASP-12b and GJ 3470b, respectively, and demonstrated that a precision of ∼1 × 10^{-4} in transit depth is achievable. HAT-P-32b (Hartman et al. 2011) is an inflated hot Jupiter class planet with a mass and radius of $0.860 \pm 0.164 \, M_J$ and $1.789 \pm 0.025 \, R_J$. It orbits a late-F/early-G type dwarf ($V = 11.3$) with a period of 2.15 d. Whilst our first Gemini/GMOS target, WASP-29b, has a relatively small scale height and was selected as a test case for the GMOS instrument, HAT-P-32b is larger, has a much higher equilibrium temperature (∼1800 K) and a lower surface gravity, and therefore should have a considerably larger atmospheric scale height and spectral features. Thus HAT-P-32b is the first science target in our Gemini-GMOS program.

This paper is structured as follows. In Section 2 we describe the observations and data reduction and in Section 3 we present our light-curve analysis and extraction of the transmission spectra. Finally in Sections 4 and 5 we present our results and conclusions.

2 GMOS OBSERVATIONS AND DATA REDUCTION

Two transits of HAT-P-32b were observed using the 8-m Gemini-North telescope with the GMOS (Hook et al. 2004) on the nights of 2012 September 5 and 2012 October 18. Data were taken as part of program GN-2012B-Q-16, and used a similar observing strategy to Gibson et al. (2013). GMOS has an imaging field-of-view of 5.5×5.5 arcmin^2, and consists of three 2048×4608 pixel CCDs arranged side by side with small gaps in-between. We observed the target ($V = 11.3$, $R = 11.2$) and two brighter comparison stars ($V = 9.8$ and 11.0) simultaneously in multi-object mode for 5.4 and 5.8 h each night, allowing several hours either side of transit given the 3.1 h transit duration. Conditions were not photometric for the duration of either night, and the observations were degraded due to variable cloud cover. This was considerably worse for the first transit, and we discuss the implications of this later.

Observations used the R400 grism + OG515 filter with a central wavelength of 725 nm in 2×2 binning. The dispersion is 0.14 nm per (binned) pixel, giving wavelength coverage from about 510–930 nm. Similarly to Gibson et al. (2013), we read out only three regions of interest including the target and the two comparison stars to reduce the readout time to 11.5 s. For the first transit, exposure times started at 30 s and were reduced to 24 s towards the end of the observations to account for varying conditions, allowing for 482 exposures. For the second transit, owing to more stable conditions, the exposure times were kept at 25 s (except the first few exposures), resulting in 552 exposures. To minimize slit losses we created a mask with slits of 30 arcsec length and 15 arcsec width for the three stars designed using a pre-image taken with GMOS, giving seeing limited (therefore variable) resolution ranging from $R \approx 650$ to 1300 at 725 nm. Fig. 1 shows the pre-image of the field with the approximate positions of the slits marked. Immediately before and after the observations, standard calibrations were taken consisting of flat-fields and arc lamp exposures. A calibration mask was also constructed using narrower 1 arcsec slits at the same positions. Arcs were taken with the calibration mask, and flat-fields were obtained with both the science and calibration mask.
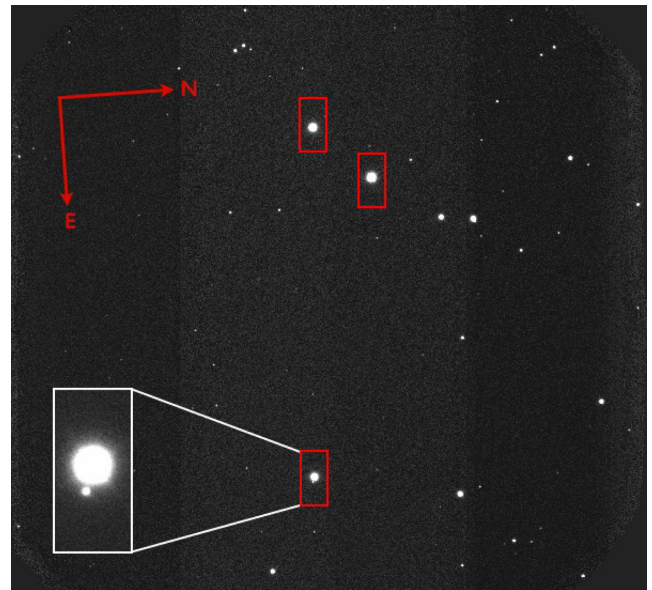
Figure 1. Mosaiced pre-image of the HAT-P-32 field taken with GMOS, acquired for the mask design. The Hα filter was used to avoid saturating the CCDs. The red boxes mark the approximate positions of the slits for the science targets, and the dispersion direction is along the *x*-axis. HAT-P-32 is located at the bottom of the field, and is the faintest of the three stars. The inset at the bottom left shows a zoomed in image of the target star, alongside the contaminant M-dwarf.

Data were reduced using the same procedure as Gibson et al. (2013), with the standard GMOS pipeline contained in the Gemini IRAF[1]/PyRAF[2] package. First the region of interest (ROI) images were processed to be in the standard GMOS format. Basic reductions included bias subtraction and wavelength calibration. A notable difference in our GMOS-North observations compared to GMOS-South is that fringing is not present at a significant level, and consequently we were able to extract useful data over the full spectral range. We flat-fielded the data using flats taken with the science mask, although we experimented with various strategies and note this had little influence on the results.

Wavelength calibrated spectra for the three stars were extracted by simply summing in the cross-dispersion direction after sky subtraction using the GSEXTRACT routine, with an aperture (diameter) of 4 arcsec ($\approx$28 binned pixels). A few pixel columns (i.e. along the spatial direction) showing significant temporal variation were masked from the extraction. The aperture size was chosen to both optimize the rms, and to avoid a previously unknown companion star $\approx$2.8 arcsec ($\approx$20 binned pixels) from the target star. The location of this star is shown in Fig. 1. Whilst the contaminant spectrum is resolved, there is some overlap in the cross-dispersion point spread functions (PSFs). With our final extraction aperture, we estimate a few per cent of the contaminant star's flux will fall within our aperture for HAT-P-32. The contaminant star is an M-dwarf, and consequently will only affect the transmission spectrum at the reddest wavelengths. At these wavelengths the detector registers about 1/60th of the peak flux from the contaminant star compared to the target. Combined with the estimated few per cent of light that falls within the aperture we calculate that the contaminant could dilute the transit depth at the reddest wavelengths by less than 10^{-5}, far below the expected precision of our observations. For this crude calculation we assumed a Gaussian PSF, and median seeing. However, even with larger contamination we expect this would not significantly dilute our light curves, but we keep this in mind during interpretation of our results, and in comparison with published parameters of the HAT-P-32 system. More detailed correction for the contaminant star is difficult, given the varying seeing over the course of the observations, although we note that our results do not change considerably when using a larger aperture.

Examples of extracted spectra of HAT-P-32 and the two comparison stars are shown in Fig. 2. We divided the target spectrum into 29 wavelength bins (each of the three chips were divided into 10 evenly spaced bins, and the faintest channel at the blue end was excluded by the pipeline). Differential 'spectral' light curves were produced by dividing the flux of the target star by the combined flux of the comparison stars in each wavelength bin. We also produced a 'white' light curve by first integrating over all wavelengths. The raw white flux time series of the target and comparison stars are shown in Fig. 3, as well as the differential white light curves for the two transits and the differential light curves for the two comparison stars. The raw fluxes illustrate the variability of the conditions as a function of time, and the effect on the differential light curves is apparent. For the first transit we see a gradual rise in the counts as well as some short-term variations, corresponding to a gradual
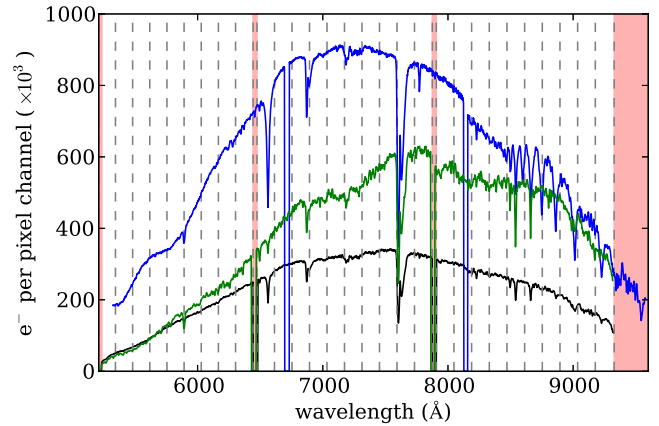
Figure 2. Example spectra extracted from a single exposure. The black line is HAT-P-32, and the blue and green lines the two comparison stars. Dashed lines define the wavelength channels, and shading marks excluded regions from the analysis at the edges and in-between detectors.

reduction of the white noise in the differential light curve. For the second transit the conditions are stable until egress where the flux suddenly drops and becomes unstable due to cloudy conditions. This corresponds to a sudden increase in the white noise at egress, marked by the vertical dashed line. This motivates our light-curve noise models, discussed in the following section. Figs 4 and 5 show the 29 raw, differential spectral light curves for the first and second transit, respectively.

3 LIGHT-CURVE ANALYSIS

3.1 White light-curve analysis

3.1.1 White noise model

We analysed the light curves in a similar way to Gibson et al. (2013). Our light-curve model is calculated using the equations of Mandel & Agol (2002), using quadratic limb darkening. We fitted for the central transit time (T_C), the system scale ($a/R_\star$), the planet-to-star radius ratio ($\rho = R_p/R_\star$), the impact parameter (b), the two quadratic limb darkening parameters (c_1, c_2) and three parameters describing a quadratic function of time for the baseline (f_{oot}, T_{grad}, $T_{grad,2}$) that we justify at the end of this section. We fixed the period at the value given in Hartman et al. (2011) of 2.150008 d. Hereafter, the transit model is denoted as $T(t, \phi)$, where t is a vector of time and ϕ is the vector of transit parameters. We denote the vector of flux measurements as f.

Given the variations in the observing conditions for both nights we cannot assume that the white noise is stationary (i.e. constant with respect to time). The calculated photon noise does not provide a realistic model of the noise (even with linear rescaling). We therefore opted to construct parametric models of the white noise for each transit; whilst not ideal, we do not expect the somewhat arbitrary choice to affect our results, due to a common mode correction that we discuss later. For the first transit, we found that an exponential decay function provided a good estimate of the noise properties:

$$\sigma_1(t_n) = \sigma_a \exp\left(-\frac{n}{\sigma_b}\right) + \sigma_c,$$
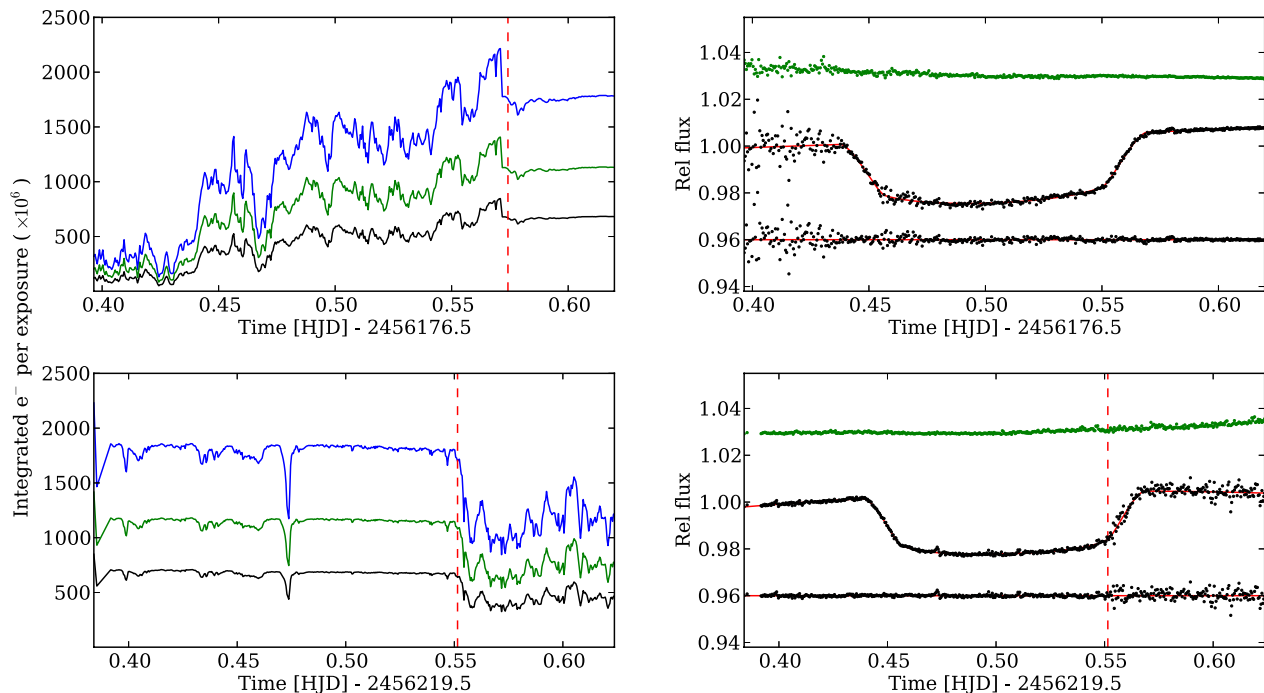
Figure 3. Left: integrated flux over all wavelengths for the target and comparison stars as a function of time, using the same colour scheme as Fig. 2. Right: differential 'white' light curves and the residuals from a best-fitting model are shown by the black dots (see Section 3), and the relative flux of the two comparison stars are shown by the green dots, offset for clarity. The top and bottom plots represent the first and second visit, respectively. For the first visit, the vertical dashed line indicates the change in exposure time. Overall, the flux gradually increases throughout the observations, resulting in a gradual increase in S/N apparent in the differential light curve. For the second visit, the dashed line represents a sudden change in sky transparency, resulting in a sudden increase in noise in the differential light curve, where the same position is marked.

and for the second transit, a simple, piecewise function was sufficient:

$$\sigma_2(t_n) = \begin{cases} \sigma_d & \text{if } n < 380, \\ \sigma_e & \text{if } n \geq 380, \end{cases}$$

where $\{\sigma_a, \sigma_b, \sigma_c, \sigma_d, \sigma_e\}$ are the parameters of the white noise models, and n indexes the time from 0. These where in part motivated by the raw flux counts plotted in Fig. 3, where we see a gradual increase in the flux for the first transit, and a sudden drop in flux for the second transit, corresponding to a sudden change of observing conditions.

We first assume that the time-varying noise followed these simple functions, but with no correlations between flux measurements. Therefore our likelihood function for each transit is

$$p(\boldsymbol{f}|\boldsymbol{t}, \phi) = \mathcal{N}\left(T(\boldsymbol{t}, \phi), \sigma_L(\boldsymbol{t})^2 \mathbf{I}\right),$$

where $\mathcal{N}(\mu, \Sigma)$ is the multivariate normal distribution with mean μ and covariance matrix Σ, $\sigma_L(\boldsymbol{t})$ are the white noise models giving the uncertainty of each data point where $L = \{1, 2\}$ and $\mathbf{I}$ is the identity matrix. This likelihood defines a diagonal covariance matrix with each element on the diagonal equal to $\sigma_L(t_n)^2$.

We then fitted both transits simultaneously, by multiplying the two likelihood functions and any prior distributions together to produce the joint posterior probability distribution. A Markov chain Monte Carlo (MCMC) algorithm was used to sample the posterior distribution and obtain marginalized probability distributions for each of our model parameters (see Gibson et al. 2013, and references therein for details on our implementation). We do not explicitly state priors for all of the parameters, implying uniform, improper priors,

and we restrict the limb darkening parameters, $a/R_\star$, ρ and b, to be positive using an improper prior of the form

$$p(x) = \begin{cases} 0 & \text{if } x < 0, \\ 1 & \text{if } x \geq 0. \end{cases}$$

We also restrict the sum of the limb darkening parameters to be less than 1 using the joint prior:

$$p(c1, c2) = \begin{cases} 0 & \text{if } c_1 + c_2 > 1, \\ 1 & \text{if } c_1 + c_2 \leq 1. \end{cases}$$

Four MCMC chains of length 200 000 were computed, and the first 20 per cent of each chain was discarded. The light curves and their best-fitting models are shown in Fig. 3.

Our choice of a second-order polynomial of time as the baseline function was justified by computing the Bayes factor for models with a linear and quadratic function of time, which compares the probability of the data under each model (assuming equal prior probabilities for each model) and naturally takes into account model complexity. We first calculate the Bayesian evidence for each model using an importance sampler (see e.g. Bishop 2006, for details). Importance sampling uses a proposal distribution from which it is straightforward to draw samples, and evaluates the posterior distribution at each sample point. Expectations are then estimated by weighting each sample using the proposal distribution, thus correcting for any bias due to sampling from the wrong distribution, i.e. the expectation of a function f of the probability distribution is given by

$$\mathbb{E}[f] = \frac{1}{S} \sum_{s=1}^{S} \frac{p(\phi_n)}{q(\phi_s)} f(\phi_s),$$
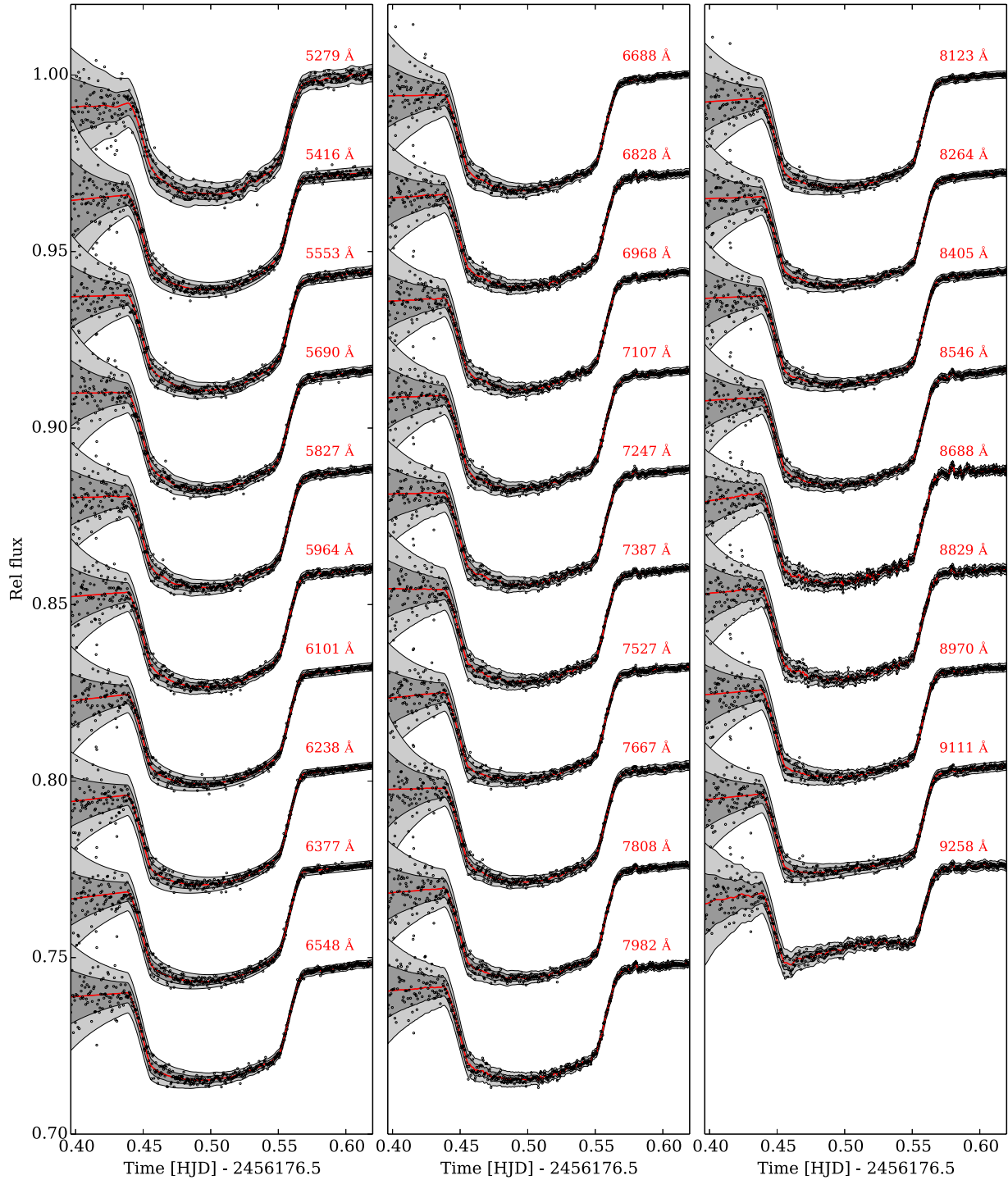
Figure 4. Raw spectral light curves for the first transit, offset for clarity. The central wavelengths for each channel are given. The red line indicates the best-fitting model from the GP fit to the transit and systematics simultaneously, and the grey shading marks the 1σ and 2σ limits of the best-fitting GP model.

where ϕ_s is the sample point, $p(\phi_s)$ and $q(\phi_s)$ are the values of the posterior and proposal distributions at ϕ_s and S is the number of samples. The success of an importance sampler depends on how closely the proposal distribution matches the true posterior. Here, we assume a Gaussian proposal distribution and estimate its mean and covariance from the MCMC distribution, and double the covariance to ensure the posterior is well sampled in the case of significant departures from a Gaussian distribution. We take 50 000 samples from the proposal, and evaluate the posterior of our model at each

sample point.[3] The Bayesian evidence (E) is simply the integral under the posterior and is calculated as

$$E = \frac{1}{S} \sum_{s=1}^{S} \frac{p(\phi_s)}{q(\phi_s)}.$$

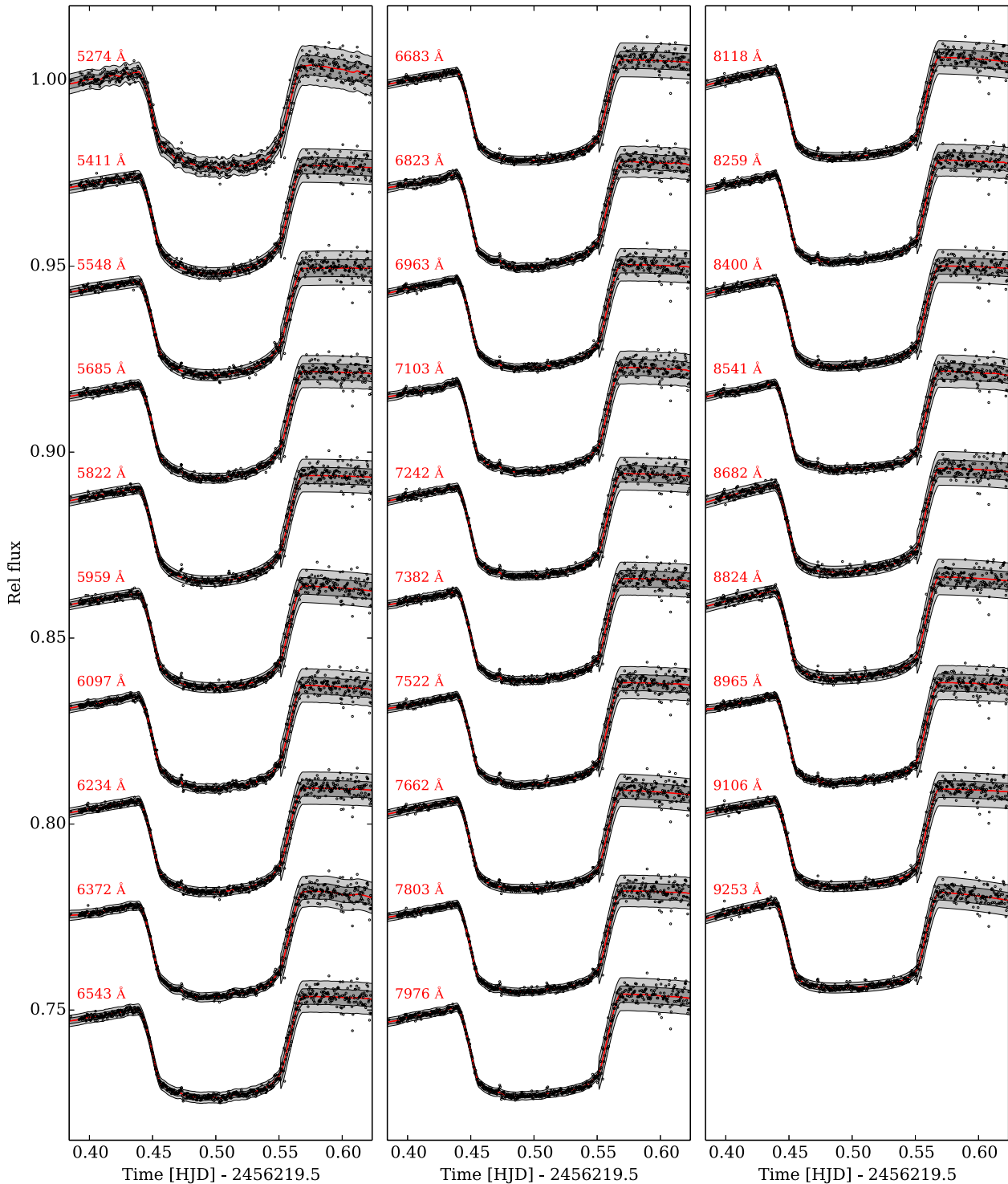[3] In practice we use the log posterior and calculate the log evidence.

Figure 5. Raw spectral light curves for the second transit, offset for clarity. Shading and models are the same as Fig. 4.

We verify our calculation by ensuring that the mean and standard deviation estimated for each model parameter are consistent with those estimated using the MCMC, and by re-running the importance sampler to ensure the evidence calculation was consistent. The Bayes factor is then simply the ratio of the Bayesian evidence for each model. We find that the Bayes factor strongly favours the quadratic model (by many orders of magnitude), and we adopt this as our model of choice. Given how strongly favoured the quadratic

model is over the linear model, we do not consider choice of improper priors for the model parameters important.

To further justify our model choice we used the Bayesian information criterion (BIC), which uses the maximum likelihood estimate and adds a complexity penalty for each degree of freedom in our model (Schwarz 1978), and is frequently used in the exoplanet literature (e.g. Gibson et al. 2010; Sing et al. 2011; Crossfield et al. 2013). Again, this strongly favours a second-order polynomial in

time. This is not surprising, given that the second-order trend can be seen by eye, particularly in the second light curve. We do not consider models where we allow a second-order trend in one light curve but not the other, as this could result in more strongly weighting one transit over the other if one model is significantly more flexible. We also checked for correlations between the baseline function and the airmass, but found no obvious correlations, indicating that the trend is not a simple function of airmass. This is supported by finding a consistent trend in the spectroscopic light curves, as we would expect this to be less pronounced over narrow wavelength regions if it is the result of colour variations between the target and comparison stars (see Section 3.2 and Figs 8 and 9).

After our white noise model fits, correlated noise is clearly present in the light-curve residuals, and must be accounted for in the fitting procedure. Given the non-stationary nature of the white noise, we are unable to use all of the methods employed in Gibson et al. (2013), as the time-averaging method (Pont, Zucker & Queloz 2006) and wavelet method (Carter & Winn 2009) are difficult to apply and interpret for variable white noise.

3.1.2 Gaussian process model

To fit for time-correlated noise simultaneously with our time-dependent white noise model, we employ the flexibility of the Gaussian process (GP) model introduced in Gibson et al. (2012a), and further applied in Gibson et al. (2012b). Application to Gemini/GMOS light curves is also discussed in Gibson et al. (2013). We apply similar methods here and refer the reader to these papers for details.

A GP is a collection of random variables, any finite subset of which have a joint Gaussian distribution. In practice, the only difference from our white noise model is that we allow for a non-diagonal covariance matrix with the elements set via a kernel function.[4] It is fully specified by its mean and covariance:

$$p(\boldsymbol{f}|\boldsymbol{t}, \phi, \theta) = \mathcal{N}\left(T(\boldsymbol{t}, \phi), \Sigma(\boldsymbol{t}, \theta)\right).$$

Here, $T(\boldsymbol{t}, \phi)$ is the transit function (or mean function), and $\Sigma(\boldsymbol{t}, \theta)$ is the kernel (or covariance function) with parameters θ:

$$\Sigma_{nm} = k(t_n, t_m|\theta).$$

As in Gibson et al. (2013) we use the Matérn 3/2 kernel function to model time correlations in the data, given by

$$k(t_n, t_m|\theta) = \xi^2(1 + \sqrt{3}\eta\Delta t)\exp(-\sqrt{3}\eta\Delta t) + \delta_{nm}\sigma_L(t_n)^2,$$

where ξ is a hyperparameter that specifies the maximum covariance, $\Delta t = |t_n - t_m|$ is the time difference, η is the inverse characteristic length scale and δ is the Kronecker delta. Again, we add the white noise term, $\sigma_L(t_n)^2$, to the diagonal of the covariance matrix to account for the non-stationary white noise. Whilst the GP could in principle model the quadratic baseline discussed in the previous section, we choose to keep this as part of the mean function as it likely has a separate origin than the higher frequency correlated noise remaining in the white noise residuals, and it is clearly supported by the data.

We defined a hyperprior for the inverse length scale, η. This took the form of a Gamma distribution with shape parameter unity, given by

$$p(\eta) = \begin{cases} 0 & \text{if } \eta < 0, \\ \frac{1}{l}\exp\left(-\eta/l_\eta\right) & \text{if } \eta \geq 0, \end{cases}$$

where l_η is the length scale of the hyperprior. We set the length scale to 200 as in Gibson et al. (2013). For ξ, instead of applying a Gamma prior, we fitted for $\log \xi$. This is equivalent to placing a prior on ξ:

$$p(\xi) = \frac{1}{\xi}$$

for positive ξ, and is a natural choice for a scale parameter. These priors were not intended to influence the results of the inference, but rather to ease convergence of the MCMC chains when the parameters are poorly constrained, and to discourage the GP model from fitting very high frequency systematics. Indeed, when the parameters are strongly constrained by the data, the likelihood should dominate the posterior distribution, and the (hyper-)priors should have negligible influence.

Inference is performed in the same way as for the white noise model. In this context hyperparameters are treated in exactly the same way as parameters of our transit model. The likelihoods and priors are combined, and we sample from the joint posterior probability distribution using MCMC to fit both transits along with their systematics models simultaneously. Note that each transit has separate values (and hyperpriors) for ξ and η. The white light curves along with their best-fitting models are shown in Fig. 6, as well as the best-fitting systematics models projected without the transit function. The GP captures the time-correlated systematics as well as the non-stationary white noise. Fig. 7 shows the normalized residuals after fitting our GP model, in other words the residuals after dividing through by our inferred error function, $\sigma_L(\boldsymbol{t})$. These should be (and are) distributed with zero mean and variance of 1, thus validating our choice of error function.

Our GP model has significant advantages over parametric models, in that we can avoid having to specify an unknown function in closed form to describe the systematics, but rather place a probability distribution over a broad and complex class of plausible functions with only a few parameters. We then marginalize out our ignorance of these functions, analogous to marginalizing out model parameters in 'normal' Bayesian inference. Furthermore, a GP is intrinsically Bayesian, and finds the simplest model that explains our data, automatically invoking Occam's Razor. Our choice of kernel is justified in Gibson et al. (2013), but we emphasize that any sensible choice of kernel is superior to assuming a parametric model.

3.2 Spectral light-curve analysis

Our raw, differential spectral light curves consist of 29 wavelength channels for each of the two transits. These are shown in Figs 4 and 5. We proceed by fitting the light curves with the same GP model described for the white light curves. The system parameters T_C, $a/R_\star$ and b are now fixed at their best-fitting values, as we are interested in the conditional distribution of ρ for each wavelength channel to obtain our transmission spectrum; therefore, anything that should be the same for each wavelength channel is fixed. This leaves ρ, the limb darkening parameters, the normalization parameters and the kernel hyperparameters (i.e.systematics parameters) to freely vary for each light-curve fit. We fitted the wavelength channels for the

---

[4] This means the probability distribution is defined for all possible inputs, rather than just for the specific times of our observations. It is this property that allows GPs to make predictions for any arbitrary values of inputs, and so defines a GP as a probability distribution over functions, distinguishing it from a simple multivariate Gaussian distribution.
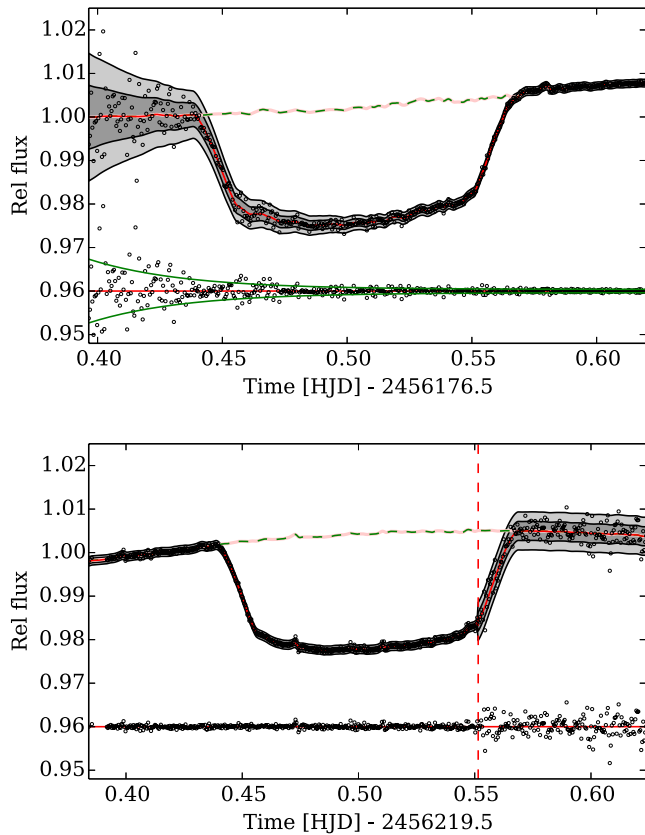
Figure 6. White light curves for the two transits, fitted with a GP model, along with their residuals from the best-fitting model, offset for clarity. The solid line is the best-fitting model, and the dashed green line is the systematics model without the transit function (but with the normalization function). The grey regions represent the 1σ and 2σ predictions of the GP model. The solid green line for the residuals of the first transit is the best-fitting white noise model, and the vertical dashed line for the second transit marks the change-point in the white noise model.
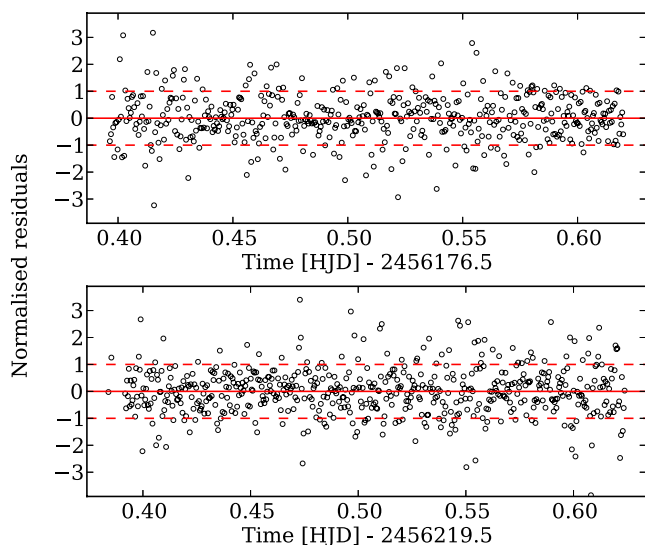

Figure 7. Normalized residuals from the GP models, i.e. the residuals from the best-fitting model divided through by their best-fitting uncertainties. These should be distributed around zero with a variance of 1 if the noise model aptly characterizes the noise.

two transits independently. The best-fitting models are shown in Figs 4 and 5, along with the best-fitting GP noise models.

After fitting the wavelength channels and inspecting the residuals, it was clear that the spectral light curves for each transit showed common mode systematics. The same effect was seen for GMOS-South data in Gibson et al. (2013), and we corrected for it in the same way. Each spectral light curve was divided through by the systematics model, which consisted of the normalization function plus the systematics model from the GP; in other words the dashed lines in Fig. 6. The systematics are clearly a multiplicative (sensitivity) effect, as the common mode systematics have the same amplitude in the *normalized* light curves, despite each channel having a large variation in raw flux. We note that the GP systematics function was additive rather than multiplicative, but this effect is negligible providing the GP does not model large amplitude, long-term trends, and is why we preferred keeping the second-order polynomial as the baseline function, rather than letting the GP account for this. Prior to fitting, we normalized by a constant value from the median of the out-of-transit flux.

Further to this common mode correction, we noticed that the residuals from the spectral light curves were correlated with residuals from other wavelength channels. We therefore implement another common mode correction, where we subtract the residuals from the corresponding white light-curve fits from each of the wavelength channels. This was implemented after dividing through by the normalization function plus systematics model, and after setting the out-of-transit flux to 1. The origin of this effect is probably related to spatial variation in the atmospheric throughput due to clouds that cannot be accounted for by the comparison stars. Applying such common mode corrections is valid when extracting transmission spectra, as we are only interested in the differential transit depths. Indeed, we are conditioning on a common systematics model just as we condition on common physical parameters.

These corrections enabled us to reach precision of $\approx$1.4–2.5 times the calculated Poisson noise, where for each spectral light curve, we take the median value of the ratio between the fitted white noise model and calculated photon limit. The subtraction of the residuals typically improved the (median) noise estimate by a factor of $\approx$1.6–1.8. For the clean sections of the white light curves, we reach precision of $\approx$10 times the photon limit, and considerably higher for the noisier parts of the white light curves (as much as 40–50 times). This clearly shows there is a significant non-white noise component in the differential white light curve, which is even more important when observing during variable weather conditions, and further justifies our use of a common mode correction. Note that this does not take into account any remaining correlated noise in the light curves.

Finally, we noticed that some of the light-curve fits were skewed by outliers in the spectral light curves. We decided to cull outlying points more than 4σ from the predictive GP distribution from each individual fit, and repeated our fits. In practice this resulted in a few points (≤ 4) being removed from each light curve.

The light curves after implementing these common mode corrections are shown in Figs 8 and 9. Most of the large-scale systematics are removed, and also variation in noise due to the variable conditions is considerably reduced. For completeness, we still fit all light curves using the same noise models, as the exponential decay is still present in the first transit. Whilst the sudden change in noise properties is still present to some degree in the second transit, the noise is much closer to stationary, and our model is therefore much less dependent on the assumed noise model. Furthermore, not all of the time-correlated noise is common mode, and we employ the
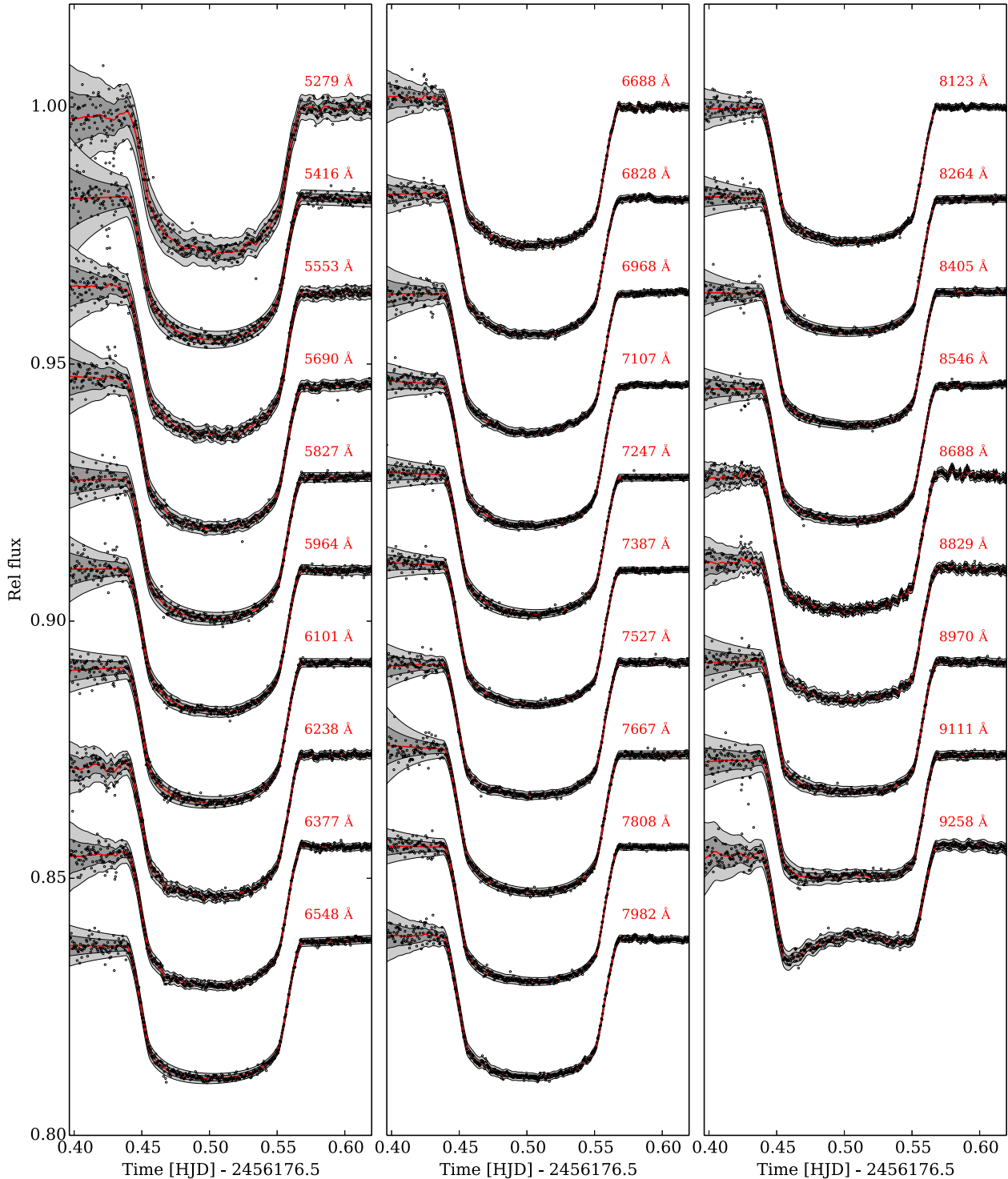
Figure 8. Spectral light curves for the first transit, after common mode corrections are applied, and offset for clarity. The central wavelengths for each channel are given. The red line indicates the best-fitting model from the GP fit to the transit and systematics simultaneously, and the grey shading marks the 1σ and 2σ limits of the best-fitting GP model.

same GP model to account for any remaining systematics in our fits. For all fits, we ran four MCMC chains of length 100 000 and checked for convergence as before. As in Gibson et al. (2013), for some runs not all parameters fully converged, with the GR statistic a few per cent form unity. This is mostly related to the relation between ξ and η parameters, where η is only constrained by the prior

when ξ is close to zero, and also the more complex noise model and baseline used for these data. However, multiple runs have verified that our transmission spectra are not significantly altered, and this does not affect our final results.

For convenience, we now summarize the steps taken to produce our final transmission spectra.
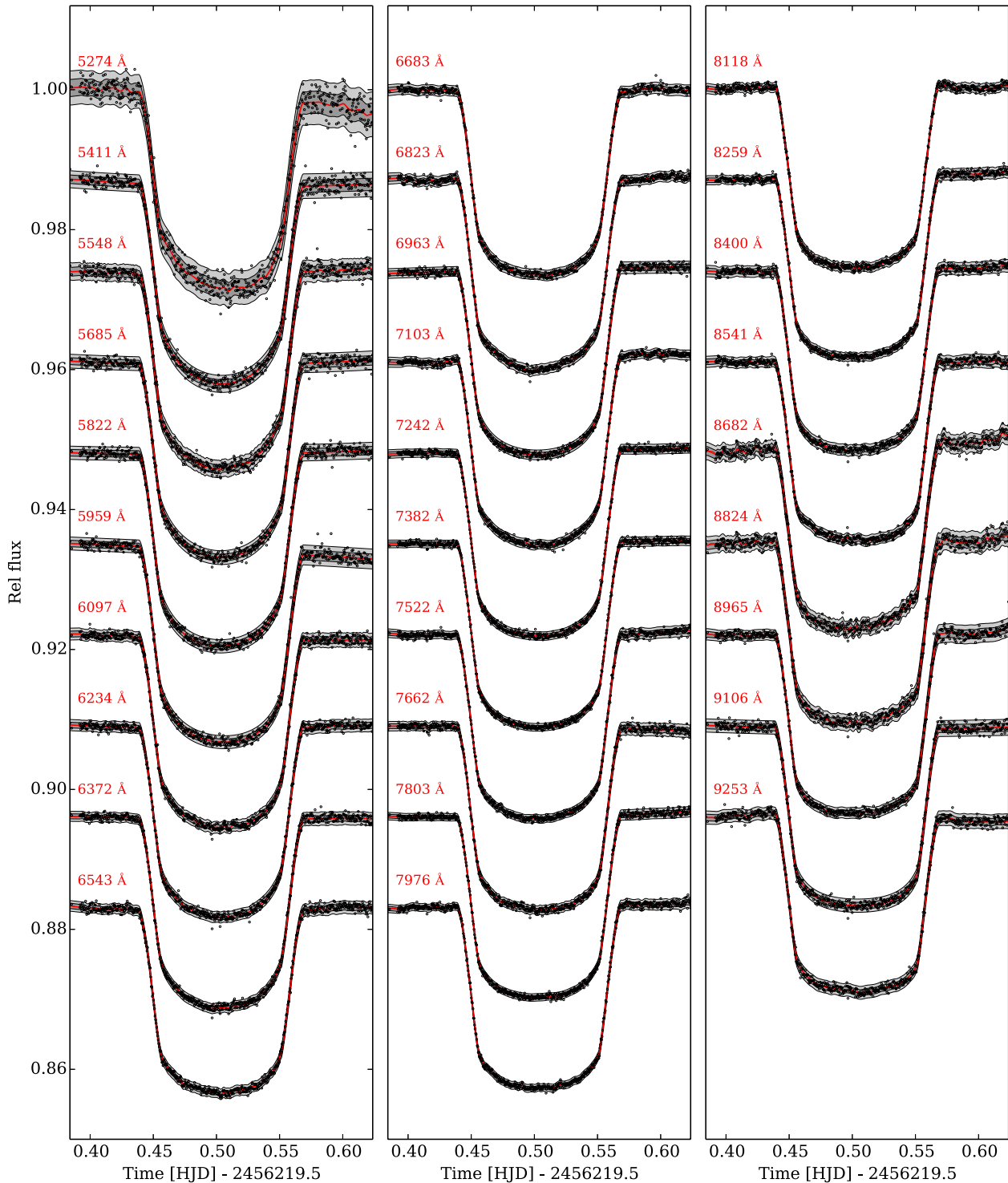
Figure 9. Spectral light curves for the second transit, after common mode corrections are applied, and offset for clarity. Shading and models are the same as Fig. 8. The offsets between the light curves have been reduced for greater clarity.

(i) The spectral time series for the target and comparison stars are extracted from the raw data, and used to produce differential light curves integrated over all wavelengths (the 'white' light curves), and for 29 spectral bins from 520 to 930 nm and ≈14 nm wide.

(ii) The white light curves for both visits are fitted simultaneously to derive the physical parameters of the HAT-P-32 system and the

GP noise model, which includes correlated noise and time-varying white noise.

(iii) To account for common mode systematics, each spectral channel is divided through by the best-fitting systematics model from the corresponding white light curve. The systematics models are shown by the dashed lines in Fig. 6.

(iv) After normalization, the residuals from the corresponding white light curve are subtracted from each spectral channel, to account for correlated noise between the channels. The effects of the last two steps are to change the light curves from Figs 4 and 5 to Figs 8 and 9.

(v) To produce our transmission spectra, we extract the posterior distribution of the planet-to-star radius ratio for each wavelength channel with the same GP model used for the corresponding white light curves. This is conditioned on the common physical parameters and common mode systematics, but marginalized over individual systematics models.

4 RESULTS AND DISCUSSION

4.1 White light-curve analysis

Results for the simultaneous GP fit to the white light curves are shown in Table 1. The fitted system parameters are consistent with those reported in Hartman et al. (2011). In general, our uncertainties are slightly larger, given the noisy ingress and egress for the first and second transits, respectively, and that we were fitting for a systematics model and limb darkening parameters. Our fitted limb darkening parameters are broadly consistent with the i band values reported in Hartman et al. (2011), and the R and I filters given in the tables of Claret (2004) using stellar parameters from the discovery paper, although we do not attempt a detailed calculation for the effective passband or the spectroscopic light curves. We note that the transit depth (and therefore planet radius) reported in Hartman et al. (2011) is perhaps diluted by the contaminant star. This will dilute the transit depth by a factor of $(1 + q)$, where q is the ratio of the flux from the target and contaminant star in a particular passband. The transit parameters were inferred from i, z and g bands, and we caution that the HAT-P-32b's radius could be marginally larger than that reported in Hartman et al. (2011), depending on the apertures used for the photometry and the relative contribution to the final parameters of each of the light curves.

Using the distributions from the MCMC chains, we calculate further system parameters for HAT-P-32, also reported in Table 1. Where the distributions were not available, we generated draws from normal distributions from the values reported in the literature and combined them with those inferred from the MCMC chains. This included the stellar mass and radius ($M_\star$ and $R_\star$) and the planetary mass (M_p). We calculated values for the transit duration (T_{14}), inclination (i), planet radius (R_p), planet density (ρ_p), log surface gravity ($\log g_p$) and the equilibrium temperature (T_p). Again, these distributions are consistent with the results of Hartman et al. (2011).

4.2 Transit ephemeris

We calculated a new ephemeris for HAT-P-32b using our two transit times, and the ephemeris reported in Hartman et al. (2011). We fitted a straight line of the form

$$T_C(E) = T_C(0) + P E$$

to the transit times, where P is the period and E is the epoch. We selected $E = 0$ to correspond to the centre of mass as closely as possible (weighted as $1/\sigma_{T_C}^2$), resulting in epochs of -243, 574 and 594 for the discovery epoch and our two light curves. In order to account for prior information on the period (as we use the ephemeris rather than the individual transit times from Hartman et al. 2011), we use a Gaussian prior with a mean and standard deviation as

Table 1. Fitted parameters for HAT-P-32b from the white light curves, given for the fully marginalized Matérn 3/2 GP model, as well as parameters derived from the posterior distribution.

Parameter	Value	Unit
Global parameters		
P^a	2.150008 (fixed)	–
$a/R_\star$	$6.091^{+0.036}_{-0.047}$	–
$R_p/R_\star$	$0.1515^{+0.0012}_{-0.0012}$	–
b	$0.093^{+0.071}_{-0.065}$	–
c_1	$0.279^{+0.070}_{-0.074}$	–
c_2	$0.254^{+0.123}_{-0.122}$	–
Transit 1		
T_C	$245\,6177.0031611^{+0.0002472}_{-0.0002410}$	HJD$_{UTC}$
f_{oot}	$1.0023^{+0.0007}_{-0.0006}$	–
T_{grad}	$0.0014^{+0.0002}_{-0.0002}$	–
$T_{grad,2}$	$0.00023^{+0.00013}_{-0.00018}$	–
$\log \xi$	$-3.19^{+0.26}_{-0.15}$	–
η	$65.04^{+31.49}_{-30.24}$	–
σ_a	$0.00691^{+0.00083}_{-0.00069}$	–
σ_b	$78.4^{+8.3}_{-7.6}$	–
σ_c	$0.00038^{+0.00004}_{-0.00005}$	–
Transit 2		
T_C	$245\,6220.0036343^{+0.0001864}_{-0.0001848}$	HJD$_{UTC}$
f_{oot}	$1.0044^{+0.0006}_{-0.0005}$	–
T_{grad}	$0.0010^{+0.0001}_{-0.0001}$	–
$T_{grad,2}$	$-0.00043^{+0.00008}_{-0.00010}$	–
$\log \xi$	$-3.41^{+0.30}_{-0.21}$	–
η	$42.27^{+36.55}_{-21.63}$	–
σ_d	$0.00048^{+0.00002}_{-0.00002}$	–
σ_e	$0.00215^{+0.00012}_{-0.00011}$	–
Ephemeris		
T_0	$245\,4942.898449 \pm 0.000077$	HJD
P	2.1500085 ± 0.0000002	d
Derived parameters		
T_{14}	$0.12959^{+0.00059}_{-0.00057}$	d
i	$89.12^{+0.61}_{-0.68}$	°
$M_\star^b$	$1.160^{+0.041}_{-0.041}$	M$_\odot$
$R_\star^b$	$1.219^{+0.016}_{-0.016}$	R$_\odot$
M_p^b	$0.860^{+0.164}_{-0.164}$	M_J
R_p	$1.796^{+0.028}_{-0.027}$	R_J
ρ_p	$0.18^{+0.04}_{-0.04}$	g cm^{-3}
$\log g_p$	$2.82^{+0.08}_{-0.09}$	(cgs)
T_p	1779^{+26}_{-26}	K

^{a}Fixed at the value of Hartman et al. (2011) for model fitting.
^{b}Adopted from Hartman et al. (2011).

reported in Hartman et al. (2011). This has negligible influence on the resulting ephemeris, which is dominated by the long baseline between the discovery observations and our Gemini light curves. The updated ephemeris is reported in Table 1.
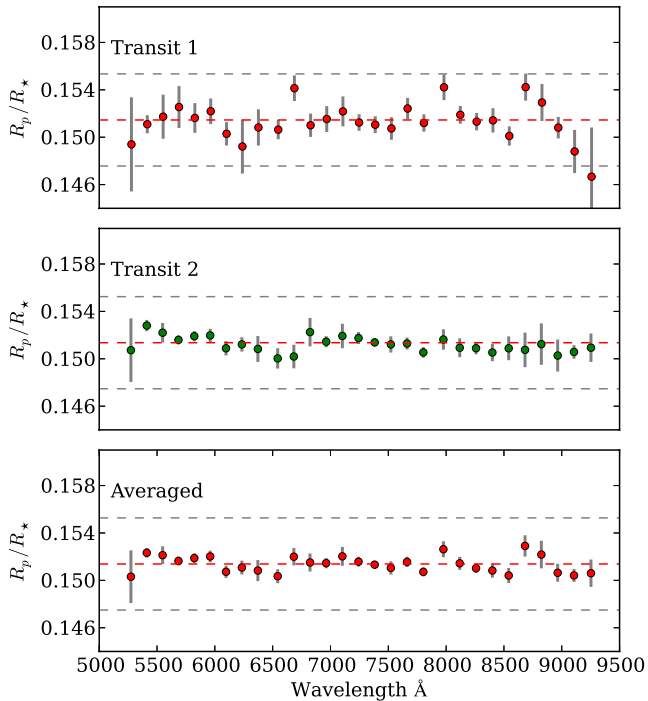
Figure 10. Transmission spectra of HAT-P-32b for the first transit, second transit and a weighted average of both. The dashed lines are the weighted mean of each spectrum, plus and minus three atmospheric scale heights.

4.3 Transmission spectrum

The transmission spectra are shown in Fig. 10 for the first and second transits, and for a combined spectrum derived from a weighted average of the two transits. These results are reported in Table 2, along with the fitted limb darkening coefficients. We ignore the small ∼5 Å shifts in the central wavelengths when combining the spectra. The dashed horizontal lines correspond to the weighted average of the spectra, plus and minus three pressure scale heights of the atmosphere, where 1 scale height is ≈1100 km when assuming the equilibrium temperature. The most striking property of the spectra is how flat they are, and they are all marginally consistent with a flat spectrum, giving reduced χ^2 of 1.23, 1.03 and 1.29 for the two transits and the combined spectrum, respectively. The slightly larger reduced χ^2 for the combined spectrum perhaps indicates that the two individual spectra are marginally inconsistent. On closer inspection, several deviant points at about 670, 800 and 870 nm in the first spectrum are mainly responsible. The limb darkening coefficients for the two runs are largely consistent, vary quite smoothly from wavelength to wavelength, and are broadly consistent with the white light-curve values at the central wavelengths, thus are not responsible for these discrepancies.[5] The disagreement at ≈670 and 800 nm is perhaps related to stars lying at the edges of the GMOS detectors.

To be conservative, given that the first transit has the larger systematics, larger uncertainties and the more convoluted noise model, we choose to focus the rest of our interpretation on the second visit's transmission spectrum, as it is easier to trust the inference using this simpler noise model. We note that this will not significantly change

our conclusions. Another GP could in principle be used to construct models for the time-varying white noise and marginalize out our ignorance of its parametric form, but we choose not to proceed with a more complex model here, which would complicate the inference further. We are cautious not to over interpret any small 'features' in the spectra such as deviant points in the first transit, or low-amplitude (≲1 scale height) variations, as these vary between the two visits, and there may still be correlations between neighbouring points in the spectra, given the nature of common mode systematics. However, we can immediately rule out the presence of broad, spectral features larger than about one atmospheric scale height, and it is this constraint from which we infer information about the atmosphere of HAT-P-32b.

Fig.11 shows the transmission spectrum of the second visit, again with the weighted average and plus and minus three scale heights indicated by the dashed lines. Given the degeneracy involved in extracting information from a flat spectrum, we do not attempt a full retrieval of the atmospheric parameters. Rather, we compare our transmission spectrum to forward models produced using the NEMESIS retrieval tool (Irwin et al. 2008), a radiative transfer code originally developed to investigate the atmospheres of Solar system planets, and recently adapted for exoplanet emission and transmission spectra (e.g. Lee, Fletcher & Irwin 2012; Barstow et al. 2013a,b). These models are plotted alongside the data in Fig. 11, and we consider the effects of Na, K, TiO, VO and H_2O in a H_2-dominated atmosphere, as these are the main atomic and molecular features we expect to see at this temperature and wavelength range. The red line is a model with solar abundances of Na, K, TiO, VO and H_2O. The large features are mainly due to TiO and VO, and clearly, if TiO and VO are present in the atmosphere, they have significantly lower mixing ratios. The blue line shows the same model now without TiO and VO, leaving large spectral features from Na and K wings that we can confidently rule out with our observations. However, it is possible that with small amounts of TiO and VO, we can significantly mask the Na and K features, and obtain a spectrum that is reasonably consistent with our data. This is shown by the green model, which is the same as the red model, only with 10^{-3} the abundances of TiO and VO. Our flat spectrum could also be produced by a clear atmosphere with only H_2O features, shown by the grey line, although we consider this explanation unlikely as we expect Na and K to be present in the atmosphere from both models and observations of hot Jupiters (e.g. Fortney et al. 2008; Huitson et al. 2012; Sing et al. 2012), and require significantly lower temperatures for them to condense out of the atmosphere (Burrows, Marley & Sharp 2000). Another obvious explanation for our data would be the presence of clouds in the upper atmosphere, which act as a grey absorber and mask potential spectral features. The most extreme effect of clouds would be to create a completely flat spectrum, consistent with the dashed red line, although intermediate scenarios are possible, where the clouds dilute some of the spectral features. We do not attempt to create a more detailed physical cloud model.

To supplement our transmission spectra, we also attempted a targeted search for narrow Na and K cores using a similar strategy to previous studies (e.g. Charbonneau et al. 2002; Huitson et al. 2012; Sing et al. 2012). These features originate at low pressures and may be visible above any hypothetical cloud deck or TiO/VO absorption, and have been observed in several hot Jupiters (e.g. Huitson et al. 2012; Sing et al. 2012). Light curves from narrow spectral regions of width 15–60 Å centred on the Na and K cores were extracted and fitted in the same way as the broad-band spectral light curves, and we applied the same common mode corrections. As the spectral regions

---

[5] In some cases the limb darkening parameters are better constrained than the white light-curve fits; this is simply a result of fixing common transit parameters.

Table 2. Transmission spectra of HAT-P-32b for the GP fits to the common mode corrected light curves, including individual fits to the two transits, and a weighted average of the separate fits. The fitted limb darkening parameters are also included.

Wavelength (Å)	Transit 1			Transit 2			Combined
	c_1	c_2	$R_p/R_\star$	c_1	c_2	$R_p/R_\star$	$R_p/R_\star$
5276	0.45 ± 0.15	0.33 ± 0.20	0.14939 ± 0.00397	0.41 ± 0.11	0.25 ± 0.15	0.15073 ± 0.00268	0.15031 ± 0.00222
5413	0.40 ± 0.04	0.36 ± 0.07	0.15110 ± 0.00077	0.45 ± 0.03	0.21 ± 0.04	0.15280 ± 0.00048	0.15233 ± 0.00040
5551	0.41 ± 0.09	0.26 ± 0.13	0.15173 ± 0.00186	0.42 ± 0.05	0.21 ± 0.08	0.15220 ± 0.00082	0.15212 ± 0.00075
5688	0.39 ± 0.08	0.24 ± 0.12	0.15255 ± 0.00177	0.44 ± 0.02	0.19 ± 0.04	0.15160 ± 0.00038	0.15164 ± 0.00038
5825	0.34 ± 0.05	0.31 ± 0.07	0.15162 ± 0.00125	0.38 ± 0.02	0.27 ± 0.04	0.15191 ± 0.00042	0.15188 ± 0.00040
5962	0.36 ± 0.06	0.31 ± 0.10	0.15219 ± 0.00108	0.33 ± 0.03	0.35 ± 0.04	0.15197 ± 0.00056	0.15202 ± 0.00049
6099	0.30 ± 0.05	0.38 ± 0.08	0.15029 ± 0.00099	0.34 ± 0.04	0.30 ± 0.06	0.15089 ± 0.00060	0.15072 ± 0.00051
6236	0.29 ± 0.12	0.41 ± 0.18	0.14922 ± 0.00228	0.34 ± 0.04	0.29 ± 0.05	0.15121 ± 0.00060	0.15108 ± 0.00058
6374	0.32 ± 0.09	0.32 ± 0.14	0.15082 ± 0.00153	0.35 ± 0.05	0.26 ± 0.07	0.15083 ± 0.00109	0.15083 ± 0.00089
6546	0.24 ± 0.05	0.35 ± 0.08	0.15064 ± 0.00081	0.27 ± 0.05	0.30 ± 0.08	0.15003 ± 0.00086	0.15035 ± 0.00059
6686	0.35 ± 0.06	0.18 ± 0.09	0.15413 ± 0.00108	0.29 ± 0.05	0.30 ± 0.06	0.15019 ± 0.00099	0.15200 ± 0.00073
6825	0.33 ± 0.06	0.33 ± 0.09	0.15101 ± 0.00098	0.29 ± 0.07	0.29 ± 0.09	0.15225 ± 0.00120	0.15151 ± 0.00076
6965	0.31 ± 0.07	0.28 ± 0.10	0.15153 ± 0.00109	0.28 ± 0.03	0.31 ± 0.05	0.15144 ± 0.00047	0.15145 ± 0.00043
7105	0.22 ± 0.07	0.43 ± 0.10	0.15218 ± 0.00126	0.30 ± 0.06	0.24 ± 0.08	0.15193 ± 0.00103	0.15203 ± 0.00080
7245	0.29 ± 0.05	0.31 ± 0.08	0.15124 ± 0.00070	0.25 ± 0.03	0.33 ± 0.04	0.15175 ± 0.00051	0.15157 ± 0.00041
7385	0.25 ± 0.05	0.35 ± 0.07	0.15105 ± 0.00071	0.24 ± 0.02	0.34 ± 0.03	0.15138 ± 0.00034	0.15132 ± 0.00030
7525	0.22 ± 0.06	0.32 ± 0.08	0.15074 ± 0.00095	0.25 ± 0.04	0.31 ± 0.06	0.15121 ± 0.00068	0.15105 ± 0.00056
7664	0.34 ± 0.04	0.27 ± 0.06	0.15242 ± 0.00089	0.24 ± 0.03	0.33 ± 0.05	0.15129 ± 0.00050	0.15156 ± 0.00044
7806	0.23 ± 0.04	0.34 ± 0.06	0.15120 ± 0.00073	0.21 ± 0.03	0.36 ± 0.04	0.15054 ± 0.00045	0.15072 ± 0.00038
7979	0.31 ± 0.06	0.20 ± 0.09	0.15421 ± 0.00108	0.20 ± 0.06	0.35 ± 0.08	0.15163 ± 0.00086	0.15263 ± 0.00067
8120	0.30 ± 0.05	0.19 ± 0.07	0.15189 ± 0.00075	0.20 ± 0.05	0.37 ± 0.08	0.15093 ± 0.00080	0.15144 ± 0.00055
8261	0.24 ± 0.04	0.26 ± 0.07	0.15131 ± 0.00074	0.21 ± 0.03	0.33 ± 0.06	0.15089 ± 0.00051	0.15103 ± 0.00042
8403	0.21 ± 0.05	0.31 ± 0.07	0.15142 ± 0.00103	0.20 ± 0.05	0.34 ± 0.07	0.15053 ± 0.00073	0.15083 ± 0.00059
8544	0.24 ± 0.05	0.32 ± 0.07	0.15010 ± 0.00081	0.20 ± 0.05	0.31 ± 0.07	0.15088 ± 0.00101	0.15041 ± 0.00063
8685	0.26 ± 0.07	0.22 ± 0.11	0.15422 ± 0.00114	0.21 ± 0.09	0.33 ± 0.13	0.15075 ± 0.00146	0.15291 ± 0.00090
8826	0.24 ± 0.09	0.26 ± 0.14	0.15293 ± 0.00156	0.20 ± 0.09	0.33 ± 0.12	0.15124 ± 0.00175	0.15218 ± 0.00116
8967	0.16 ± 0.05	0.36 ± 0.08	0.15081 ± 0.00091	0.14 ± 0.07	0.38 ± 0.09	0.15027 ± 0.00135	0.15064 ± 0.00076
9109	0.05 ± 0.05	0.44 ± 0.08	0.14880 ± 0.00182	0.17 ± 0.03	0.36 ± 0.05	0.15058 ± 0.00057	0.15042 ± 0.00054
9255	0.10 ± 0.08	0.22 ± 0.10	0.14667 ± 0.00415	0.14 ± 0.06	0.38 ± 0.08	0.15093 ± 0.00120	0.15060 ± 0.00116
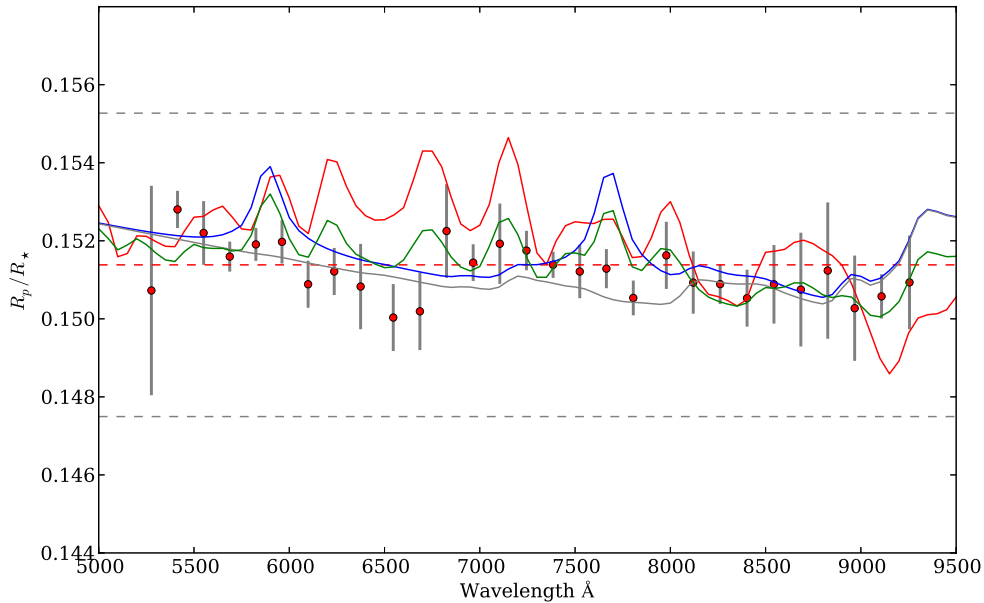
Figure 11. Transmission spectrum of HAT-P-32b for the second transit. The dashed horizontal lines represent the weighted mean and plus and minus three scale heights, where one scale height ≈1100 km. The solid lines represent various forward models produced using the NEMESIS code (see Section 4.3). The red line is a model with solar abundances of Na, K, TiO, VO and H_2O in a H_2-dominated atmosphere. The blue is the same without TiO and VO. The green is a Na, K, TiO, VO and H_2O model but with substellar ($\sim 10^{-3}$ solar) abundances of TiO and VO. Finally, the grey model is with only H_2O.

narrowed, the resulting increase in the photon noise resulted in the GP systematics models being poorly constrained. Furthermore, the systematics are considerably larger with a spectral width of 15 Å. Unsurprisingly, 140 Å bins centred on the Na and K features resulted in a planet-to-star radius consistent with a flat spectrum and with similar uncertainties to the surrounding data points. Consequently, we were unable to detect or place meaningful constraints on the presence of Na and K cores.

This leaves us with two plausible models for our data: (1) clouds in the upper atmosphere, masking atomic and molecular species resulting in a flat spectrum, and (2) a cloud free atmosphere with Na and K, and with TiO and VO at substellar levels ($\sim 10^{-3}$ solar), represented by the green line in Fig. 11. Clouds and/or haze are emerging as a significant feature of many hot Jupiter atmospheres. For example, they are required to explain the broad transmission spectrum of HD 189733b (Pont et al. 2008, 2013; Gibson et al. 2012b; Sing et al. 2012), and probably to account for diluted molecular features in XO-1b and HD 209458b at infrared wavelengths (Deming et al. 2013). On the other hand, the presence of TiO and VO has been suggested as a possible explanation for temperature inversions in hot Jupiter atmospheres (e.g. Fortney et al. 2008), and it has been detected in the atmosphere of HD 209458b, albeit tentatively (Désert et al. 2008). Only small ($\sim 10^{-3}$ solar) abundances of TiO are necessary to explain this potential TiO feature; however, we can rule out a HD 209458b-like atmosphere for HAT-P-32b, given the lack of Na wings. Understanding the abundance of TiO/VO in the atmosphere of hot Jupiters is a complex problem, involving 3D circulation, day/night cold traps and gravitational settling of haze particles (Parmentier, Showman & Lian 2013). Indeed, Parmentier et al. (2013) predict for HD 209458b that if TiO condenses into particles of size larger than about 1 μm on the planet's night side, it is plausible that it could be completely depleted from the atmosphere. As TiO and VO begin to condense at temperatures of about 2000 K (e.g. Lodders 2010; Parmentier et al. 2013), it seems unlikely that TiO/VO will be present at the terminator of HAT-P-32b. Another natural explanation for an underabundance of TiO and VO is a high C/O ratio (Madhusudhan 2012), or low metallicity, in which case metal hydrides could play a prominent role (Sharp & Burrows 2007). We do not construct models including metal hydrides, but assume that their effects would be broadly similar to the metal oxides but consider their presence less plausible.

Several other factors we have yet to consider could contribute to a lack of prominent Na/K features. These includes ionization which could reduce the abundance of neutral Na/K and therefore the size of the features. This is likely to occur only at high altitudes in the atmosphere, and therefore might reduce the amplitude of the line cores, but cannot easily explain the lack of pressure broadened wings which originate deeper in the atmosphere (e.g. Fortney et al. 2003). Another possibility is that the temperature of the atmosphere at the terminator could be lower than estimated from the equilibrium temperature, therefore reducing the scale height and therefore size of spectral features. However, whilst this could result in diminished Na/K features, this would require an extreme departure from the predicted temperature to explain the lack of observed features alone. Of course, the effects we have discussed are not mutually exclusive, and a combination of them could contribute to a flat transmission spectrum.

We did not attempt to establish which of these atmospheric models provide a better explanation for our observed transmission spectrum in a quantitative manner. To do so robustly would require a full Bayesian analysis; a simple least-squares or maximum likelihood (minimum χ^2) approach would not be satisfactory, as the different families of models being considered have very different levels of complexity. This in turn would require us to chose a sensible parametrization for each family of models, and suitable priors for each parameter. How to do this is far from obvious, yet the priors would have a major effect on the results, as the flat spectrum means that the data contain relatively little information on the values of many of the parameters. Furthermore, to perform such a model comparison we would have to check for residual correlated errors on the radius ratio measurements obtained from neighbouring wavelength bins, which is also difficult to ascertain. Therefore, we do not perform a statistical model comparison, but rather deduce the relative plausibility in the light of what is currently known about hot Jupiters. Whilst clouds/hazes have been conclusively detected in hot Jupiter atmospheres, TiO or VO has yet to be observed conclusively. Furthermore, the TiO/VO explanation requires an unknown depletion mechanism to remove significant amounts of Ti and V from the atmosphere so that broad TiO/VO features are not detected in the spectrum, yet leave enough to mask Na and K wings, and also with high enough temperatures at the terminator so that gaseous TiO and VO are present. We therefore consider this explanation to be more 'fine-tuned', as would a similar explanation involving metal hydrides. This leaves clouds/haze as the more plausible explanation for our spectrum out of those considered, although more precise optical spectra or information at other wavelengths would be required to break this degeneracy.

5 CONCLUSIONS

We have presented Gemini-North GMOS observations of the hot Jupiter, HAT-P-32b, during two primary transits. Transmission spectra were extracted for both transits at wavelengths of 520–930 nm in ≈14 nm bins, using several comparison stars to perform differential spectrophotometry.

Using the white light curves from the two visits, we derived the system parameters and updated the ephemeris. We found our results to be broadly consistent with the discovery paper (Hartman et al. 2011), but with larger uncertainties, owing to the variable weather conditions during observations and the increased systematics resulting from extracting light curves from spectra. A previously unknown M-dwarf contaminant star ≈2.8 arcsec from HAT-P-32 was detected in the pre-imaging, which may have diluted the discovery light curves, depending on the apertures used in the data reduction. We therefore caution that some of the system parameters in the discovery paper could be systematically biased; in particular, the planet radius could be marginally larger.

The transmission spectra were extracted using the GP model of Gibson et al. (2012a) to simultaneously model the transit light curve, instrumental systematics and time-varying white noise. The flexibility of our GP model proved particularly important for this data set. Common mode systematics were identified in the spectral light curves, and we used the GP systematics model and residuals from the white light curves to correct for these and significantly improve the precision of our transmission spectra.

The resulting data are consistent with a flat spectrum, and we can rule out broad features larger than about one pressure scale height of the atmosphere. There are two plausible explanations for the absence of the large Na, K, TiO and VO or metal hydride features that we would expect from stellar abundances in HAT-P-32b's atmosphere. The first is the presence of clouds in the upper atmosphere, with large enough particle sizes to act as a grey absorber. The second is that small amounts of TiO/VO (or metal hydrides) could mask the wings of the Na and K lines sufficiently and still be

consistent with no large-scale variations over about a scale height. The first scenario is more plausible, given that the second requires high enough temperatures at the terminator for gaseous TiO and VO to be present, but with an unknown mechanism required to deplete Ti and V so that large TiO and VO features remain undetected in our spectra, yet leave enough to mask Na and K wings. Furthermore, the cloud scenario naturally explains a flat spectrum, whereas alternative explanations can easily explain the lack of Na/K features, but need to be somewhat 'fine-tuned' in order to show no other detectable, broad features in the optical transmission spectrum. Of course, a combination of both clouds and TiO/VO is also possible, as are contributions from metal hydrides, ionization, and a smaller scale height if the temperature at the terminator region significantly departs from the equilibrium temperature. In the absence of stronger constraints on the atmosphere of HAT-P-32b, a robust retrieval of detailed information is infeasible. Despite achieving relatively high precision in the transmission spectra, this case highlights the need for broad wavelength coverage to fully understand the complexity of hot Jupiter atmospheres.

ACKNOWLEDGEMENTS

REFERENCES

Barstow J. K., Aigrain S., Irwin P. G. J., Bowles N., Fletcher L. N., Lee J.-M., 2013a, MNRAS, 430, 1188
Barstow J. K., Aigrain S., Irwin P. G. J., Fletcher L. N., Lee J.-M., 2013b, MNRAS, 434, 2616
Bean J. L., Miller-Ricci Kempton E., Homeier D., 2010, Nat, 468, 669
Bean J. L. et al., 2011, ApJ, 743, 92
Berta Z. K. et al., 2012, ApJ, 747, 35
Bishop C. M., 2006, Pattern Recognition and Machine Learning. Springer-Verlag, Berlin
Brown T. M., 2001, ApJ, 553, 1006
Burrows A., Marley M. S., Sharp C. M., 2000, ApJ, 531, 438
Carter J. A., Winn J. N., 2009, ApJ, 704, 51
Charbonneau D., Brown T. M., Noyes R. W., Gilliland R. L., 2002, ApJ, 568, 377
Claret A., 2004, A&A, 428, 1001
Crossfield I. J. M., Barman T., Hansen B. M. S., Howard A. W., 2013, A&A, preprint (arXiv:1308.6580)
Crouzet N., McCullough P. R., Burke C., Long D., 2012, ApJ, 761, 7
Deming D. et al., 2013, ApJ, 774, 95
Désert J.-M., Vidal-Madjar A., Lecavelier Des Etangs A., Sing D., Ehrenreich D., Hébrard G., Ferlet R., 2008, A&A, 492, 585
Fortney J. J., Sudarsky D., Hubeny I., Cooper C. S., Hubbard W. B., Burrows A., Lunine J. I., 2003, ApJ, 589, 615
Fortney J. J., Lodders K., Marley M. S., Freedman R. S., 2008, ApJ, 678, 1419
Gibson N. P. et al., 2010, MNRAS, 404, L114
Gibson N. P., Pont F., Aigrain S., 2011, MNRAS, 411, 2199
Gibson N. P., Aigrain S., Roberts S., Evans T. M., Osborne M., Pont F., 2012a, MNRAS, 419, 2683
Gibson N. P. et al., 2012b, MNRAS, 422, 753
Gibson N. P., Aigrain S., Barstow J. K., Evans T. M., Fletcher L. N., Irwin P. G. J., 2013, MNRAS, 428, 3680
Hartman J. D. et al., 2011, ApJ, 742, 59
Hook I. M., Jørgensen I., Allington-Smith J. R., Davies R. L., Metcalfe N., Murowinski R. G., Crampton D., 2004, PASP, 116, 425
Huitson C. M., Sing D. K., Vidal-Madjar A., Ballester G. E., Lecavelier Des Etangs A., Désert J.-M., Pont F., 2012, MNRAS, 422, 2477
Irwin P. G. J. et al., 2008, J. Quant. Spectrosc. Radiat. Transfer, 109, 1136
Lee J.-M., Fletcher L. N., Irwin P. G. J., 2012, MNRAS, 420, 170
Lodders K., 2010, in Barnes R., ed., Formation and Evolution of Exoplanets. Wiley-VCH, Berlin, p. 157
Madhusudhan N., 2012, ApJ, 758, 36
Mandel K., Agol E., 2002, ApJ, 580, L171
Parmentier V., Showman A. P., Lian Y., 2013, A&A, preprint (arXiv:1301.4522)
Pont F., Zucker S., Queloz D., 2006, MNRAS, 373, 231
Pont F., Knutson H., Gilliland R. L., Moutou C., Charbonneau D., 2008, MNRAS, 385, 109
Pont F., Sing D. K., Gibson N. P., Aigrain S., Henry G., Husnoo N., 2013, MNRAS, 432, 2917
Schwarz G., 1978, Ann. Stat., 6, 461
Seager S., Sasselov D. D., 2000, ApJ, 537, 916
Sharp C. M., Burrows A., 2007, ApJS, 168, 140
Sing D. K., Vidal-Madjar A., Désert J.-M., Lecavelier Des Etangs A., Ballester G., 2008, ApJ, 686, 658
Sing D. K. et al., 2011, MNRAS, 416, 1443
Sing D. K. et al., 2012, MNRAS, 426, 1663
Stevenson K. B., Bean J. L., Seifahrt A., Desert J.-M., Madhusudhan N., Bergmann M., Kreidberg L., Homeier D., 2013, ApJ, preprint (arXiv:1305.1670)
Waldmann I. P., 2012, ApJ, 747, 12

This paper has been typeset from a TEX/LATEX file prepared by the author.