



**QUEEN'S
UNIVERSITY
BELFAST**

Centroid vetting of transiting planet candidates from the Next Generation Transit Survey

Günther, M. N., Queloz, D., Gillen, E., McCormac, J., D. R. Bayliss, D., Bouchy, F., Walker, S. R., West, R. G., Eigmüller, P., Smith, A. M. S., Armstrong, D. J., Burleigh, M., Casewell, S. L., Chaushev, A. P., Goad, M. R., Grange, A., Jackman, J., Jenkins, J. S., M. Loudon, T., ... J. Wheatley, P. (2017). Centroid vetting of transiting planet candidates from the Next Generation Transit Survey. *Monthly Notices of the Royal Astronomical Society*, 472(1), 295-307. <https://doi.org/10.1093/mnras/stx1920>

Published in:

Monthly Notices of the Royal Astronomical Society

Document Version:

Publisher's PDF, also known as Version of record

Queen's University Belfast - Research Portal:

[Link to publication record in Queen's University Belfast Research Portal](#)

Publisher rights

Copyright 2017 The Authors. This work is made available online in accordance with the publisher's policies. Please refer to any applicable terms of use of the publisher.

General rights

Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact openaccess@qub.ac.uk.

Open Access

This research has been made openly available by Queen's academics and its Open Research team. We would love to hear how access to this research benefits you. – Share your feedback with us: <http://go.qub.ac.uk/oa-feedback>

Centroid vetting of transiting planet candidates from the Next Generation Transit Survey

Maximilian N. Günther,^{1*} Didier Queloz,¹ Edward Gillen,¹ James McCormac,² Daniel Bayliss,³ Francois Bouchy,³ Simon. R. Walker,² Richard G. West,² Philipp Eigmüller,⁴ Alexis M. S. Smith,⁴ David J. Armstrong,² Matthew Burleigh,⁵ Sarah L. Casewell,⁵ Alexander P. Chaushev,⁵ Michael R. Goad,⁵ Andrew Grange,⁵ James Jackman,² James S. Jenkins,^{6,7} Tom Louden,² Maximiliano Moyano,⁸ Don Pollacco,² Katja Poppenhaeger,⁹ Heike Rauer,⁴ Liam Raynard,⁵ Andrew P. G. Thompson,⁹ Stéphane Udry,³ Christopher A. Watson⁹ and Peter J. Wheatley²

¹*Astrophysics Group, Cavendish Laboratory, J.J. Thomson Avenue, Cambridge CB3 0HE, UK*

²*Department of Physics, University of Warwick, Gibbet Hill Road, Coventry CV4 7AL, UK*

³*Observatoire de Genève, Université de Genève, 51 Ch. des Maillettes, CH-1290 Sauverny, Switzerland*

⁴*Institute of Planetary Research, German Aerospace Center, Rutherfordstrasse 2, D-12489 Berlin, Germany*

⁵*Department of Physics and Astronomy, Leicester Institute of Space and Earth Observation, University of Leicester, Leicester LE1 7RH, UK*

⁶*Departamento de Astronomía, Universidad de Chile, Casilla 36-D, Santiago 7591245, Chile*

⁷*Centro de Astrofísica y Tecnologías Afines (CATA), Casilla 36-D, Santiago 7591245, Chile*

⁸*Instituto de Astronomía, Universidad Católica del Norte, Angamos 0610, Antofagasta 1270709, Chile*

⁹*Astrophysics Research Centre, School of Mathematics and Physics, Queen's University Belfast, Belfast BT7 1NN, UK*

Accepted 2017 July 25. Received 2017 June 29; in original form 2017 May 24

ABSTRACT

The Next Generation Transit Survey (*NGTS*), operating in Paranal since 2016, is a wide-field survey to detect Neptunes and super-Earths transiting bright stars, which are suitable for precise radial velocity follow-up and characterization. Thereby, its sub-mmag photometric precision and ability to identify false positives are crucial. Particularly, variable background objects blended in the photometric aperture frequently mimic Neptune-sized transits and are costly in follow-up time. These objects can best be identified with the centroiding technique: if the photometric flux is lost off-centre during an eclipse, the flux centroid shifts towards the centre of the target star. Although this method has successfully been employed by the *Kepler* mission, it has previously not been implemented from the ground. We present a fully automated centroid vetting algorithm developed for *NGTS*, enabled by our high-precision autoguiding. Our method allows detecting centroid shifts with an average precision of 0.75 milli-pixel (mpix), and down to 0.25 mpix for specific targets, for a pixel size of 4.97 arcsec. The algorithm is now part of the *NGTS* candidate vetting pipeline and automatically employed for all detected signals. Further, we develop a joint Bayesian fitting model for all photometric and centroid data, allowing to disentangle which object (target or background) is causing the signal, and what its astrophysical parameters are. We demonstrate our method on two *NGTS* objects of interest. These achievements make *NGTS* the first ground-based wide-field transit survey ever to successfully apply the centroiding technique for automated candidate vetting, enabling the production of a robust candidate list before follow-up.

Key words: surveys – eclipses – occultations – planets and satellites: detection – binaries: eclipsing.

* E-mail: mg719@cam.ac.uk

1 INTRODUCTION

Transiting exoplanets allow determination of their radii relative to their host star. If the host star is bright enough, the planet mass is accessible through radial velocity (RV) monitoring. Together, these allow the planet density and hence bulk composition to be determined, making such planets favourable targets for detailed characterization of their atmospheric structure and composition. Gaining this insight is a key factor in the search for habitable planets. However, given the design of previous surveys, the known transiting planet population is typically faint ($V > 13$). The Next Generation Transit Survey (NGTS) (Chazelas et al. 2012; Wheatley et al. 2013, Wheatley et al., in preparation) is a wide-field survey designed to detect Neptune-sized exoplanets orbiting bright host stars. The survey commenced full operation at ESO’s Paranal Observatory in early 2016.

In addition to planets, transit-like signals in light curves can be caused by astrophysical false positives, such as eclipsing binaries, which can be incorrectly interpreted as bona fide planetary transits (see e.g. Cameron 2012). In the case of NGTS, we previously showed that the ability to identify false positives, which outnumber the planet yield by an order of magnitude, is expected to be a major factor for the survey’s scientific success (Günther et al. 2017). In particular, variable background objects (blended in the photometric aperture) can mimic Neptune-sized transits and are costly in follow-up time (Fig. 1). Such variable background objects can best be identified with the centroiding technique: if the photometric flux is lost off-centre during an eclipse, the flux centroid shifts towards the centre of the target star. Although this method has successfully been employed by the space-based *Kepler* mission (Batalha et al. 2010 and 2012, Bryson et al. 2013), it has previously not been proven feasible for ground-based surveys.

If a background star lies within the photometric aperture of a target star, the centre of flux in the aperture, ξ , is offset from the true centre of the target (Fig. 1). Assuming a symmetric point spread function, we introduce the concept of a ‘photometric centre of mass’. One can directly translate this principle from classical mechanics into the photometric scenario by replacing the term ‘mass’ with ‘flux’, denoted as $F_c(t) = F_c$ for the constant object in the aperture and $F_e(t)$ for the eclipsing object:

$$\xi(t) = \frac{F_c \mathbf{x}_c + F_e(t) \mathbf{x}_e}{F_c + F_e(t)}. \quad (1)$$

The CCD position of the two objects is denoted as \mathbf{x}_c and \mathbf{x}_e , respectively. Consequently, any change in brightness of one object leads to a shift of the centre of flux in the aperture. Any centroid shift is hence dependent on $F_e(t)$:

$$\Delta\xi(t) = \xi(t) - \xi(t=0). \quad (2)$$

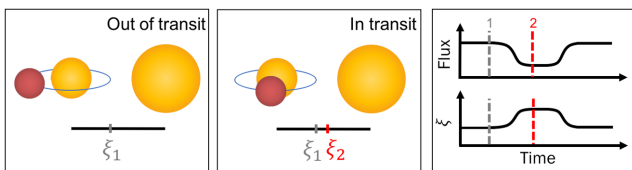


Figure 1. Sketched illustration of a centroid shift $\Delta\xi = \xi_2 - \xi_1$ correlated to a transit-like signal, which is caused by a background eclipsing binary diluted in the aperture of a constant target star. Typically, both systems are photometrically extracted as a single source, and are not visually resolvable in NGTS images. Note that if the system can be visually resolved, the direction of the shift indicates which object undergoes the eclipse.

A detailed derivation can be found in Section 3.1. Note that these equations are for two objects, yet this model allows for an arbitrary number of objects, which are described by their common centre of flux.

In the following, we distinguish four distinct cases. Thereby, we assume that the brighter object in the aperture will be identified as the ‘target’, and denote the fainter objects in the aperture as ‘blended’ or ‘background objects’:

(i) Diluted planet (*dilP*): The target hosts a transiting planet, the background objects are constant (on the respective time-scales and period of the detected signal). The planet transit signal is diluted and the measured depth is decreased. For example, a Hot Jupiter might be miss-identified as a Neptune-sized planet. Note that this means that detecting a correlation between flux and centroid data is not sufficient to disregard a planet candidate.

(ii) Diluted eclipsing binary (*dilEB*): The target is an eclipsing binary, the background objects are constant. The binary’s eclipse signal is diluted and the measured depth is decreased. If the dilution is high and/or the eclipse is shallow, this can mimic a planetary transit.

(iii) Background planet (*BP*): The target is constant, one of the background objects hosts a transiting planet. The transit signal is diluted and the measured depth is decreased. The transit depth would be decreased by >50 per cent, in most cases hindering the detection of the signal.

(iv) Background eclipsing binary (*BEB*): The target is constant, one of the background objects is an eclipsing binary. The transit signal is diluted and the measured depth is decreased. If the dilution is high and/or the transit is shallow, this can mimic a planetary transit around the target star.

2 COMPUTATION OF THE STELLAR CENTROID TIME SERIES DATA

The 12 NGTS telescopes have a combined field of view (FOV) of almost 100 deg^2 . Each CCD is a deep depleted $2k \times 2k$ Ikon-L produced by Andor, with pixel size of $13.5 \mu\text{m}$ (4.97 arcsec). The default radius of the circular photometric aperture is 3 pixel, covering a total area on sky of 700 arcsec^2 . For all observations, the survey employs the DONUTS autoguiding algorithm developed by McCormac et al. (2013), which ensures the telescopes stay centred on target over the course of one night.

The centre of aperture per exposure, $\mathbf{x}_{\text{apt}}(t)$, is determined by a global fit to all reference stars in the FOV. Thereby, the high-precision autoguiding minimizes random scatter of these aperture positions to ~ 0.1 pixel between subsequent exposures, and a total drift of < 1 pixel over multiple hours. In theory, the centre of flux per exposure, $\mathbf{x}_{\text{flux}}(t)$, is equal to $\mathbf{x}_{\text{apt}}(t)$ in the case of isolated stars with a Gaussian point spread function and perfect alignment of the aperture mask. However, in the presence of blended objects or stray light, the two are offset from each other. This offset depends on the magnitude and position of the background object.

Both $\mathbf{x}_{\text{apt}}(t)$ and $\mathbf{x}_{\text{flux}}(t)$ are computed for each exposure in the NGTS pipeline using CASUTOOLS¹ (Irwin et al. 2004). We introduce the centroid as a relative value, ξ , relating the two:

$$\mathbf{x}_{\text{flux}}(t) = \mathbf{x}_{\text{apt}}(t) + \xi(t). \quad (3)$$

¹ <http://casu.ast.cam.ac.uk/surveys-projects/software-release>, online 2017 May 12.

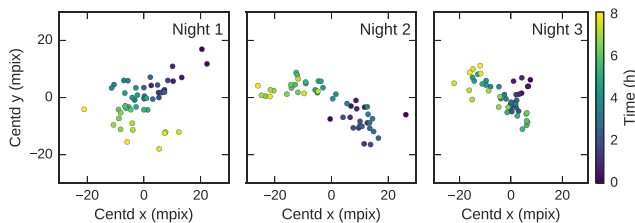


Figure 2. The centroid motion varies between different observing nights. Centroid values are shown in milli-pixel (mpix). The colour coding illustrates the time from the start of the observations on a given night.

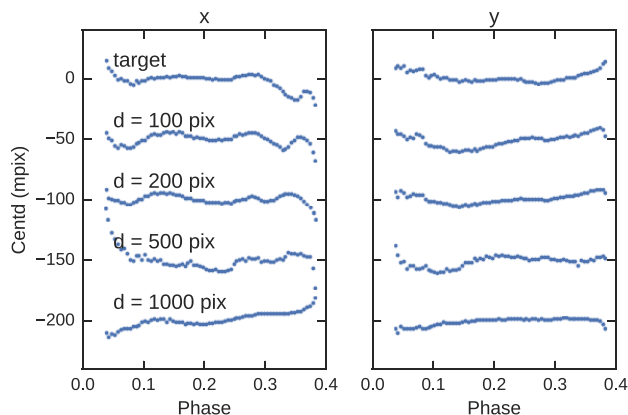


Figure 3. The similarity of the centroid motion between objects decreases with distance on the CCD. Examples show the centroid data of randomly selected objects. This data was collected over the course of 4 months and phase folded on a sidereal day period.

As a result, the centroid is automatically corrected for any global drift of the grid of apertures across the CCD, and is representing the remaining (local) residuals. Note that $\xi(t)$ is a time series containing centroid measurements for all exposures.

2.1 Pre-whitening the flux-centroid time series

NGTS typically observes each field down to an elevation of 30° above the horizon. Given the large ($\sim 3^\circ$) FOV there is a ~ 5 arcsec difference in the atmospheric refraction between the higher and lower elevation sides of the image. This difference ranges from 0.5 pixels at the zenith, to 1.75 pixels at an elevation of 30° (see equation G in Bennett 1982), with a pixel size of 4.97 arcsec. This effect is temperature and pressure dependent and acts along the parallactic angle, which rotates as the field crosses the sky. Additionally, small amounts of field rotation may occur due to residual polar misalignment of each NGTS telescope. Hence, the centroids display a systematic low-amplitude drift over the course of each night (see Fig. 2). The strength of this effect varies across the image and may differ night-to-night, but is correlated between neighbouring objects. The correlation between objects decreases as the distance between them increases (see Fig. 3). The NGTS autoguider aims to fix the global average position of the field to the same sub-pixel level, and can hence not correct for this. To address this, we instead follow a three-step approach to correct the centroid motion for each target star:

(i) *Flattening*: We compute the median centroid of each night and use it to correct the night-to-night offset, in order to combine all nights together.

(ii) *Detrending*: We use the centroid correlation between neighbouring objects to pre-whiten the centroid time series of our planet candidates. For a given target, we select N_{ref} reference stars. We perform a least-squares fit to determine which linear combination of these resembles the target’s centroid curve best, and remove this trend:

$$\xi_{\text{target,detrended}} = \xi_{\text{target,raw}} - \sum_i^{N_{\text{ref}}} c_{i,\text{raw}} \cdot \xi_{i,\text{raw}} \quad (4)$$

Here, $\xi_{i,\text{raw}}$ is the centroid data of the i th reference star, and $c_{i,\text{raw}}$ is the scale parameter that is fitted for. When selecting reference stars, we only regard objects within a certain distance d_{max} (in pixel) from the target on the CCD, based on our observations of decreasing correlation with distance (see Fig. 3). Further, we pre-select the most correlated objects to decrease the number of free fit parameters. We phase-fold the centroid time series on the transit period, exclude the in-transit data, and calculate Pearson’s correlation coefficient of the target with each selected neighbour. The N_{ref} most correlated objects are selected. Finally, to choose reference stars which are less affected by residuals of the sky background ($\lesssim 50 \text{ ADU s}^{-1}$) subtraction and to avoid saturated stars ($\gtrsim 50\,000 \text{ ADU s}^{-1}$), only objects with flux of $500\text{--}10\,000 \text{ ADU s}^{-1}$ are included.

(iii) *Sidereal day correction*: The observing pattern of ground-based surveys can lead to systematic noise on the period of a sidereal day. As our centroid detrending is applied to data that has been phase-folded on the transit period, residuals of sidereal day systematics may be present in long-period systems. To further enhance our algorithm, we phase-fold the centroid time series of the target on the mean period of a sidereal day and perform a moving average fit to correct for any remaining trends. We only consider data outside of the primary or secondary eclipses. Hence, the correction does not affect the transit signal. Generally, the effect of the sidereal day correction is marginal, as most NGTS targets are found at short periods.

2.2 Effect of the different detrending steps

Fig. 4 demonstrates the effect of our centroid detrending steps for a chosen NGTS planet candidate (see section 2.1). First, the night-to-night offsets in the time series are removed (‘flattening’). Secondly, we detrend the centroid phase-folded on the transit period, using a reference signal computed from the most correlated neighbours. This leads to a clear improvement, removing almost all systematics. Finally, a sidereal day correction is applied. Since NG 0409-1941 020057 has a period of 1.61 d (see Section 3.2.2), which is close to the sidereal day period, these systematics have mostly been removed before this stage, such that their effects are negligible. Note that long-period transits, however, will profit from this additional step. After detrending, a clear centroid signal remains at phase 0, indicating the presence of a background object in the aperture (see Section 1).

2.3 Global performance: sub-milli-pixel precision

To assess the global performance of our technique, we mimic the centroiding process for planet candidates on ~ 1200 targets from a typical NGTS field. For each star, we randomly uniformly draw a period between 0.8 and 15 d, on which we phase-fold the centroid time series. This range is based on the minimum value used for the *Kepler* planet occurrence rates in Fressin et al. (2013) and the period sensitivity of NGTS. We select stars with flux counts of $500\text{--}10\,000$

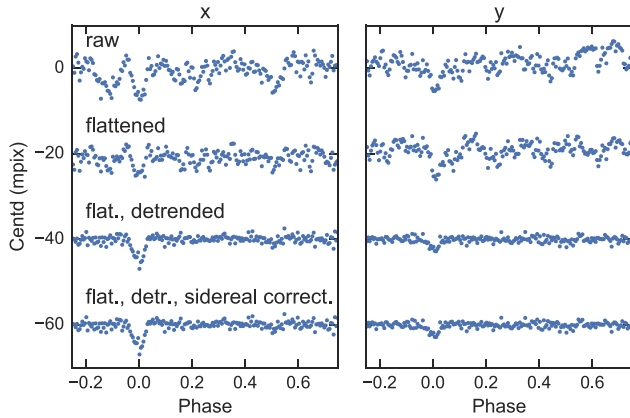


Figure 4. Improvement of the centroiding systematics with each detrending step (see Section 2.1) on the example of NG 0409-1941 020057. Shown are the centroid time series phase-folded on the period of the transit-like signal at 1.61 d. First, we correct the night-to-night offsets in the time series by subtracting the median value per night (‘flattened’). Secondly, the centroid data is detrended by a reference signal calculated from their most correlated neighbours. Lastly, a sidereal day correction is applied. Centroid data in x and y direction are shown in the left and right column of the figure, respectively. The time series are offset from each other by 20 mpix on the vertical axis for clarity.

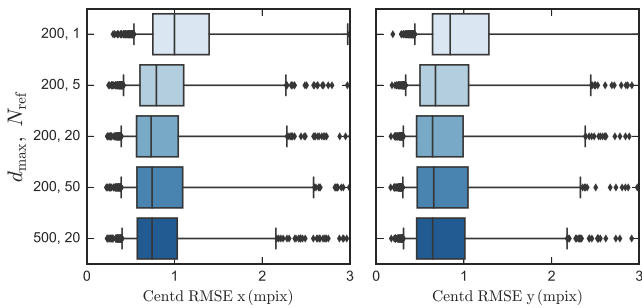


Figure 5. Comparison of different centroid detrending settings. Shown is the root mean squared error (RMSE) of the phase-folded data after detrending. Boxes represent the median, and 25th and 75th percentile of all objects. Whiskers display the 5th and 95th percentile, and outlying objects are plotted as symbols. A search radius of $d_{\max} \approx 200$ pixel and choice of $N_{\text{ref}} \approx 20$ reference objects is the best compromise between a high-precision and computational efficiency.

ADU s^{-1} , to avoid influence of the sky background ($\sim 50 \text{ ADU s}^{-1}$) and saturation ($> 50\,000 \text{ ADU s}^{-1}$).

We test the impact of different settings for the maximum distance on the CCD, d_{\max} , and maximum number of reference stars, N_{ref} (Fig. 5). Selecting the most correlated neighbour as the sole reference object already leads to an average milli-pixel (mpix) precision. Including more reference objects further increases this precision, yet saturates at $N_{\text{ref}} \approx 20$. Widening the search radius does not yield any improvement, as the correlation of the centroid systematics decreases with distance (see Section 2.1 and Fig. 3). We find an optimal performance for $d_{\max} \approx 200$ pixel and $N_{\text{ref}} \approx 20$. In theory, additional reference objects add additional information, but can also lead to overfitting or converging to a local minimum. In any case, pre-selecting a limited number of the most correlated reference stars is advantageous for computational efficiency.

Fig. 6 illustrates the remaining root mean squared error (RMSE) of the phase-folded centroid data after detrending. We achieve an RMSE of < 1 mpix for 73 and 75 per cent of all targets in the x and

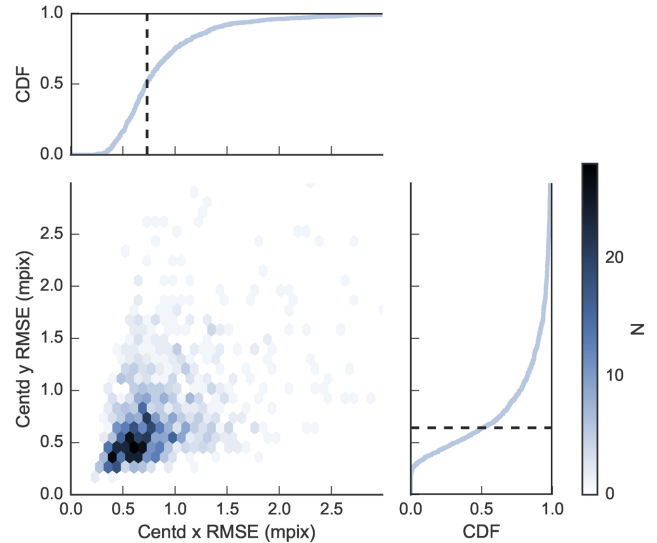


Figure 6. Achieved centroid precision for an analysis run over ~ 1200 stars from a typical *NGTS* field. Each star is given a period, which is randomly uniformly drawn between 0.8 and 15 d, and the centroid time series is phase-folded on the respective period. Shown is the root mean squared error (RMSE) of the phase-folded data in x and y direction after detrending, as well as the cumulative distribution function (CDF) in each direction.

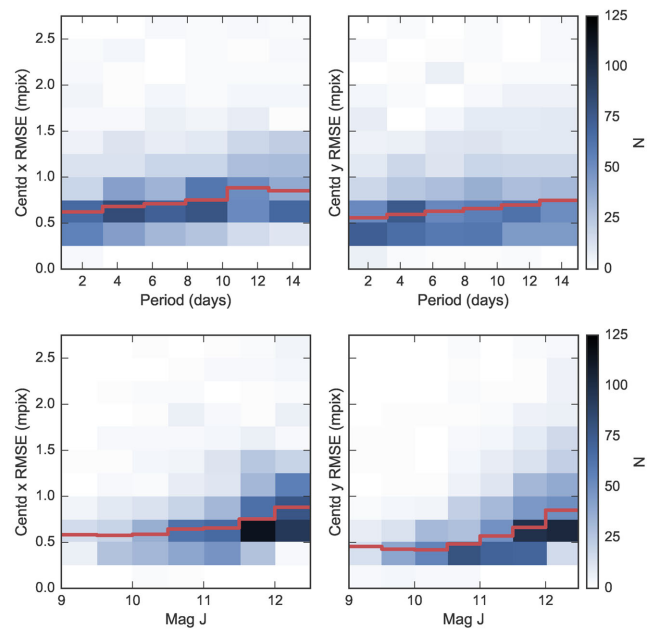


Figure 7. Dependence of the centroid precision versus period of the transit-like signal (upper panel) and the J -magnitude of the host star taken from 2MASS (lower panel), shown for ~ 1200 targets from a typical *NGTS* field. Solid red lines indicate the median centroid RMSE for each period and magnitude bin. Long-period signals show slightly higher noise in the phase-folded centroid curves. Fainter stars are more influenced by the sky background, leading to increased noise in the phase-folded centroid curves.

y direction, respectively. 61 per cent show a precision of < 1 mpix in both directions at the same time. We achieve an average centroid precision of ~ 0.75 mpix in both directions, and as low as 0.25 mpix for individual targets.

We identify an increase of the achieved centroid precision for shorter periods (Fig. 7, upper panel). This is a direct consequence

of the white noise statistics and the amount of data that can be binned up in the phase-folded centroid curve. Similarly, the centroids of fainter host stars are stronger influenced by sky background variations, leading to higher noise in the phase-folded centroid curves (Fig. 7, lower panel).

3 IDENTIFICATION OF CENTROID SHIFTS CAUSED BY BLENDED SOURCES

To identify centroid shifts caused by blended sources in the aperture of a planet candidate, we phase-fold the detrended centroid time series on the period of the respective transit feature. We compare the flux and centroid phase curves using four methods:

(i) Manual inspection of the phase-folded flux and centroid curves (see panel A in Figs 8 and 11). This allows a qualitative investigation of the noise in each time series, and the identification of any systematic features that might mimic or hide a correlation between flux and centroid.

(ii) Pearson’s correlation for a rolling window² (see panel B in Figs 8 and 11). The window size should be longer than the signal width. In praxis, we employ multiple window sizes and include all results into our further analyses. For the purpose of this paper, we demonstrate our analyses using a window size of 0.25 in phase. To start, we place this window centred on the transit and calculate Pearson’s correlation coefficient between the flux and centroid data. We then slide the window across the data, repeating the measurement for each position.

(iii) Cross-correlation (see panel C in Figs 8 and 11). We calculate Pearson’s correlation coefficient between the phase-folded flux and each centroid curve. We then shift one series of data against the other (with periodic bounds), repeating the measurement for each position. This is equivalent to the cross-correlation function widely used in astronomy, but normalized to a range of -1 to $+1$.

(iv) Hypothesis tests and manual inspection of the rain plots (see panel D in Figs 8 and 11). Rain plots were successfully used to qualitatively identify correlations between the flux and centroid time series for *Kepler* (see e.g. Batalha et al. 2010). While a single target undergoing an eclipse would ‘rain’ straight down (local centre of flux is unaffected), a blended object causes a trend (‘wind’) sideways. In our analyses for *NGTS*, we extract the in-transit data as a subset. We set the Null Hypothesis that this in-transit centroid data is distributed around the mean of the out-of-transit centroid data, i.e. around 0. We perform statistical hypotheses tests, a two-tailed *T*-test and two-tailed binomial test, and record their *p*-value. For a chosen significance level, e.g. $\alpha = 0.01$, we test if the Null Hypothesis can be rejected. The two-tailed tests are employed to verify both if the mean is significantly greater or significantly smaller than 0 (whereas a one-tailed test would only test for one direction).

3.1 Analysis of blended systems

A detailed study of the centroid signal is needed to determine which scenario causes the transit-like feature (dilP, dilEB, BP or BEB; see Section 1). At this stage in the vetting process we assume that the period has been identified correctly.

In the first step of the centroid analysis, any information on visually identified nearby objects must be used to identify blended

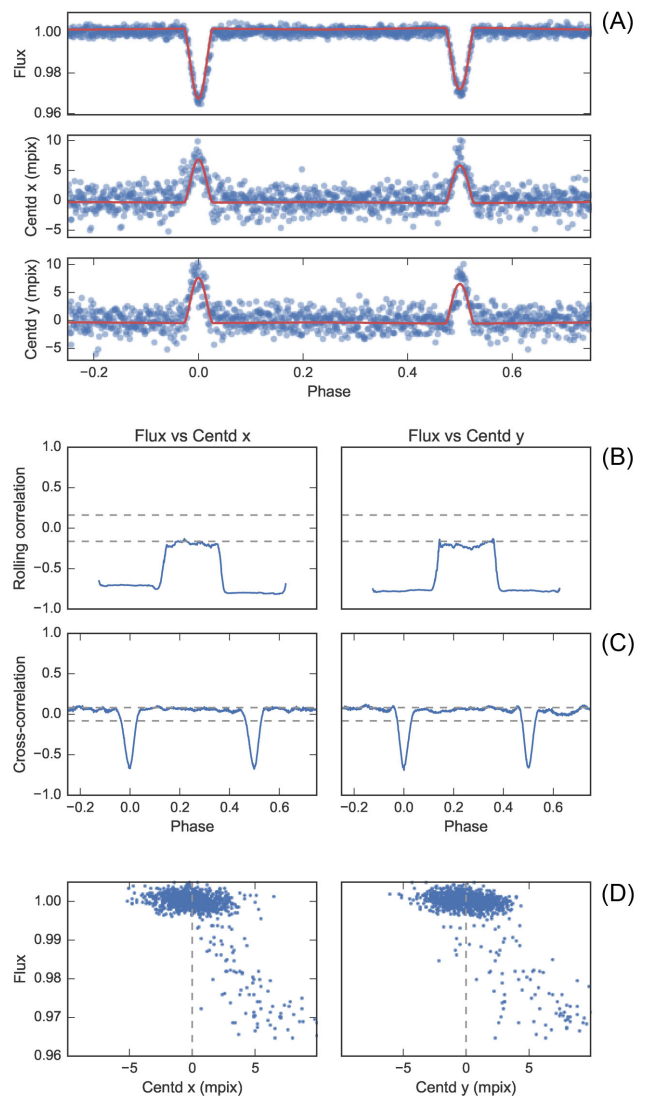


Figure 8. Identification and model fit of a centroid shift correlated to the transit-like signal in NG 0522-2518 017220. (A) The correlation is clearly visible in the manual inspection of the flux and centroid phase curves, even though the centroid shift is at the ~ 6 mpix level. Red curves represent the result of the MCMC analysis. (B) The rolling correlation analysis shows a significant correlation around phase 0. Dashed lines indicate confidence intervals of 99 per cent, calculated as $2.58/\sqrt{w/\Delta t}$ (see Fisher 1921), whereby w is the window size and Δt is the difference in phase between binned points. (C) The cross-correlation shows a significant correlation at lag 0 in phase. Dashed lines show confidence intervals of 99 per cent calculated as $2.58/\sqrt{N}$ (see Fisher 1921), whereby N is the total number of binned points. (D) Rain plots show a clear trend between flux and centroid shift.

sources. We therefore inspect the *NGTS* images and cross-match with existing catalogues including *Gaia* DR1 (Gaia Collaboration 2016a,b) and 2MASS (Skrutskie et al. 2006). If the local centre of flux shifts away from (towards) the centre of the target star, this indicates that the target star (background object) is decreasing in brightness, i.e. undergoes the eclipse (see Fig. 1). However, as discussed in Section 1, this alone does not disqualify a planet scenario. Only a detailed model of the flux and centroid time series simultaneously, taking the dilution factor into account, can guide in this question by establishing the likelihood of all astrophysical parameters.

² Known as rolling, windowed, or sliding-window correlation.

In the following, we model the aperture to contain two sources of light, one that is constant (c) and one that is eclipsing (e). This implicitly models multiple background objects as originating from a single source, located at their centre of light in the aperture. Consequently, $F_e(t)$ is the flux time series of the eclipsing object alone, while $F_c(t) = F_c$ denotes the constant object. The time series of the total flux in the aperture is

$$F_{\text{sys}}(t) = F_e(t) + F_c. \quad (5)$$

We define the dilution factor out of transit as

$$D_0 = 1 - \frac{F_e(t_0)}{F_{\text{sys}}(t_0)}, \quad (6)$$

where t_0 is a chosen time out-of-transit.

We have to distinguish between two cases: an eclipsing background object (BP or BEB; Section 3.1.1) and a constant background object (dilP or dilEB; Section 3.1.2). We consider these two cases in turn in the following two subsections.

3.1.1 Constant target, eclipsing background object

The target is constant (located at \mathbf{x}_c) and the signal comes from an eclipsing background source that is offset (located at \mathbf{x}_e). We set the origin of the coordinate system to the target's position, hence $\mathbf{x}_c = 0$. The centroid out of transit, ξ_0 , is then given following equations (1) and (6) as:

$$\xi_0 = \frac{F_c(t_0)\mathbf{x}_c + F_e(t_0)\mathbf{x}_e}{F_{\text{sys}}(t_0)} = \frac{F_e(t_0)\mathbf{x}_e}{F_{\text{sys}}(t_0)} = (1 - D_0)\mathbf{x}_e. \quad (7)$$

The centroid at any time is then given as

$$\begin{aligned} \xi(t) &= \frac{F_c(t)\mathbf{x}_c + F_e(t)\mathbf{x}_e}{F_{\text{sys}}(t)} - \xi_0 \\ &= \left(\frac{F_e(t)}{F_{\text{sys}}(t)} - 1 + D_0 \right) \mathbf{x}_e. \end{aligned} \quad (8)$$

For normalized light curves, we can express this as

$$F_e(t) = F_e^{\text{norm}}(t) \times F_e(0), \quad (9)$$

$$F_{\text{sys}}(t) = F_{\text{sys}}^{\text{norm}}(t) \times F_{\text{sys}}(0), \quad (10)$$

$$\frac{F_e(t)}{F_{\text{sys}}(t)} = \frac{F_e^{\text{norm}}(t)}{F_{\text{sys}}^{\text{norm}}(t)} \times (1 - D_0), \quad (11)$$

$$\xi(t) = \left(\frac{F_e^{\text{norm}}(t)}{F_{\text{sys}}^{\text{norm}}(t)} - 1 \right) \times (1 - D_0)\mathbf{x}_e. \quad (12)$$

3.1.2 Eclipsing target, constant background object

The target undergoes the eclipse (located at \mathbf{x}_e) and is diluted by a constant background source that is offset (located at \mathbf{x}_c). We set the origin of the coordinate system to the target's position, hence $\mathbf{x}_e = 0$. The centroid out of transit, ξ_0 , is then given following equations (1) and (6) as:

$$\xi_0 = \frac{F_c(0)\mathbf{x}_c + F_e(0)\mathbf{x}_e}{F_{\text{sys}}(0)} = \frac{F_c(0)\mathbf{x}_c}{F_{\text{sys}}(0)} = D_0\mathbf{x}_c. \quad (13)$$

The centroid at any time is then given as

$$\xi(t) = \frac{F_c(t)\mathbf{x}_c + F_e(t)\mathbf{x}_e}{F_{\text{sys}}(t)} - \xi_0 = \left(\frac{F_c(t)}{F_{\text{sys}}(t)} - D_0 \right) \mathbf{x}_c. \quad (14)$$

For normalized light curves, we can express this as

$$F_c(t) = F_c^{\text{norm}}(t) \times F_c(0) = 1 \times F_c(0), \quad (15)$$

$$F_{\text{sys}}(t) = F_{\text{sys}}^{\text{norm}}(t) \times F_{\text{sys}}(0), \quad (16)$$

$$\frac{F_c(t)}{F_{\text{sys}}(t)} = \frac{1}{F_{\text{sys}}^{\text{norm}}(t)} \times D_0, \quad (17)$$

$$\xi(t) = \left(\frac{1}{F_{\text{sys}}^{\text{norm}}(t)} - 1 \right) \times D_0\mathbf{x}_c. \quad (18)$$

3.1.3 Dependence of the blend position on transit parameters

Rearranging the previous equations allows us to express the position of the blended background object in terms of the actually measured transit depth of the system, δ_{sys} :

$$\mathbf{x}^{\text{blend}} = \pm \xi_{\text{max}} \times \left(\frac{\delta_{\text{sys}}}{1 - \delta_{\text{sys}}} \times D_0 \right)^{-1}. \quad (19)$$

The sign depends on whether the blended background object is the constant (+) or variable (−) source in the aperture.

3.1.4 Bayesian analysis

A Bayesian analysis enables us to explore complex parameter spaces and robustly estimate posterior likelihoods for each parameter. Specifically, any information on visually identified objects can be used as priors, such as the position of the blended background object.

We base our fitting model on the EB³⁴ module by Irwin et al. (2011), which incorporates the possibility of a dilution term. We establish a maximum likelihood function based on a simultaneous fit of the phase-folded flux and centroid time series following equations (12) and (18) (depending on which model we chose to fit). The free parameters in our model are the relative CCD position of the blended background object ($\mathbf{x}^{\text{blend}}$), the dilution factor D_0 , as well as the standard parameters of an EB model, namely the surface brightness ratio J , ratio of the sum of the radii over the orbital distance $(R_1 + R_2)/a$, the radius ratio R_2/R_1 and cosine of the inclination $\cos i$. We also include offset terms for the normalized flux and centroid time series, denoted F_0 , $\xi_{x,0}$ and $\xi_{y,0}$. These correct for any remaining offset after the normalization, and are usually negligible. Additionally, the error bars on the flux and centroid data are free parameters and are determined by the fit, $\sigma(F)$, $\sigma(\xi_x)$ and $\sigma(\xi_y)$. Where there is no prior information, we choose uniform priors. For systems which pass through the vetting to this stage, we expect low-brightness ratio systems that do not show significant secondary eclipses, hence uniform priors are generally expected to be uninformative.

We fix the period and epoch to the values determined by the NGTS candidate pipeline. The other parameters of the EB model (limb darkening, gravity darkening and reflectivity) are left at their standard values. Both stars are assumed to be without spots. The EB model can resemble the Mandel & Agol (2002) planet transit model if surface brightness ratio, mass ratio, light traveltime and reflectivity are zero. Hence, we can readily model all scenarios (dilP, dilEB, BP, BEB; see Section 1).

³ <https://github.com/mdwarfgeek/eb>, online 2017 February 5.

⁴ See Acknowledgements.

To find the best fit, we follow a two-step approach. First, we employ a differential evolution algorithm (Storn & Price 1997) to explore the parameter space and find a global optimization. This uses an iterative approach in which different populations of solutions are compared and the best fit is kept. As it does not rely on gradient methods, it is robust against local minima. We implement the `SCIPY` (Jones et al. 2001) distribution of the algorithm.

Secondly, we use the result of the differential evolution as the initial guess for a Markov chain Monte Carlo (MCMC) algorithm and adopt priors, to re-fine the fit and establish the likelihood of our parameters. We implement the `EMCEE` (Foreman-Mackey et al. 2013) package. Initially, the walkers are distributed following a Gaussian distribution around the initial guess with the standard deviation being 1 per cent of the given data range. We first perform several MCMC runs with 10 000 steps each, and test for convergence using the Gelman–Rubin statistic (Gelman & Rubin 1992). In the final run, we compute 5000 burn-in steps and 45 000 evaluation steps, from which we sample every n th step, whereby n is determined by the maximum of the autocorrelation time of all parameters. The Gelman–Rubin statistics for all parameters lie well below the recommended value $\hat{R} < 1.1$, suggesting convergence of the MCMC chains.

3.2 Case studies

In the following, we employ our centroiding technique on the example of two case studies representing different scenarios. First, we consider NG 0522-2518 017220, which is an eclipsing binary slightly diluted by a visually resolved blend. Secondly, we apply our analyses to NG 0409-1941 020057, which has no prior visual information, yet can be identified as a strongly diluted BEB from a joint MCMC fit of photometric flux and centroid data.

3.2.1 NG 0522-2518 017220

The target NG 0522-2518 017220 was first detected with a period of ~ 0.83 d. Its sinusoidal out-of-eclipse (OOE) variation unveiled that the true signal originates from a primary and secondary eclipse with a period of ~ 1.67 d, which have comparable depths of ~ 3 per cent and widths of ~ 2.6 h. With $G = 13.6$ in *Gaia* DR1 ($J = 12.6$ and $K = 12.2$ in 2MASS), the object is bright and well-suited for follow-up and potential characterization. It is located at RA = $05^{\text{h}}23^{\text{m}}31.6^{\text{s}}$, Dec. = $-25^{\text{d}}08^{\text{m}}48.4^{\text{s}}$ and has been identified as 2MASS 05233161-2508484 and GAIA 2957881682551005056.

However, the centroid time series shows clear shifts of 5–8 mpix in x and y direction, respectively, for both the primary and secondary eclipse (see Fig. 8). All correlation and hypothesis tests confirm a statistically significant centroid shift correlated to the transit signal (see Table 1). On visual inspection, we identify a neighbouring object at ~ 18 arcsec separation, which has a similar brightness at $G = 14.0$ ($J = 12.7$, $K = 12.2$) and partly blends into the target’s photometric aperture (see Fig. 9). The centroid shifts into the positive x and y direction, which combined with the visual information suggests that the target is undergoing the eclipse, while the flux from the blended background object is constant (see Fig. 1). A direct comparison of the detrended and phase-folded light curves shows that the eclipses are only visible for NG 0522-2518 017220. In the light curve of the neighbouring object, the signals are diluted beyond the noise level, and hence not detectable (Fig. 10). This verifies the conclusions drawn from the centroid analysis.

Using the *Gaia* DR1 magnitude and models for the NGTS point spread function and bandpass, we calculate a dilution factor of

Table 1. Statistical identification of a centroid shift correlated to the transit-like signals in NG 0522-2518 017220 and NG 0409-1941 020057. The table displays the signal-to-noise ratio (SNR) of the rolling correlation and cross-correlation analyses, displayed in Figs 8 and 11. Further it lists the resulting p -values from a T -test and binomial test of the in-transit centroid data, testing the Null Hypothesis that the centroid is distributed around the mean of the out-of-transit data, i.e. around 0.

	x	y
<i>NG 0522-2518 017220</i>		
SNR roll. corr.	41.43	43.01
SNR cross-corr.	53.28	26.55
p -value T -test	4.01×10^{-32}	5.69×10^{-26}
p -value binomial test	6.62×10^{-24}	9.96×10^{-18}
<i>NG 0409-1941 020057</i>		
SNR roll. corr.	51.66	44.06
SNR cross-corr.	15.75	12.28
p -value T -test	7.92×10^{-26}	4.34×10^{-16}
p -value binomial test	8.67×10^{-19}	5.65×10^{-12}

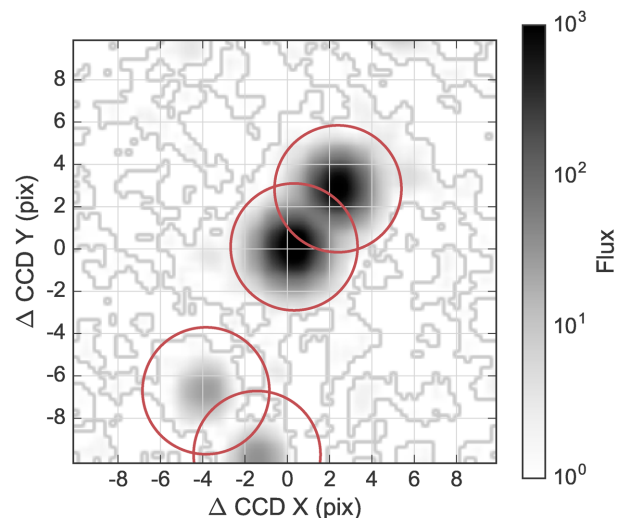


Figure 9. Visual inspection of the NGTS sky images for NG 0522-2518 017220 uncovers that a neighbouring object is blending into the photometric aperture of the target (centred in the image). The size of one NGTS pixel measures 4.97 arcsec. Red circles illustrate the photometric aperture radius of 3 pixels.

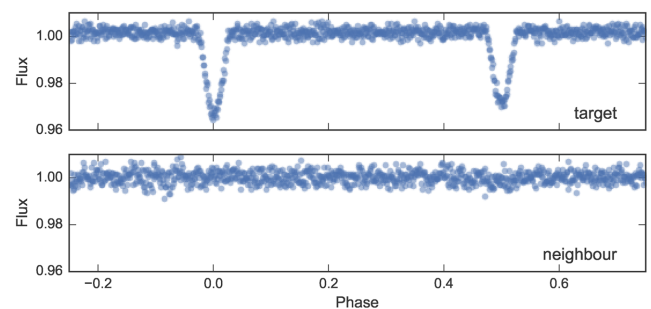


Figure 10. Comparison of the phase-folded light curves for NG 0522-2518 017220 and its blending neighbour. While the primary and secondary eclipse signals are clearly visible in the target, they cannot be identified in the blending neighbour. This verifies the outcome of the centroid analysis: the target is undergoing the eclipse, while the neighbouring object is constant.

$D_0 = 0.13 \pm 0.02$ for NG 0522-2518 017220. We further compute the centre of flux of the third light in the aperture to be at $\mathbf{x}_{\text{backg.obj.}} = (1.9 \pm 0.2, 1.7 \pm 0.2)$ pixel. This information on D_0 and $\mathbf{x}_{\text{backg.obj.}}$ is used as Gaussian priors on these parameters in our MCMC model fit.

The object shows significant OOE modulation on a ~ 1 per cent level, which appears to be sinusoidal and in phase with the eclipse signal. We first analyse the time evolution of the OOE variation by dividing the light curve in equal sections in time, and compare the variability between these sections. We find that the OOE modulation significantly changes over the 175 d observing span. We further compute Lomb Scargle periodograms for the OOE data. We identify multiple periods, which we can relate to the orbital period of the system and systematics on a sidereal day period. With an orbital period of ~ 1.67 d, it can be assumed that the orbit is circular and that the binary components are tidally locked. Hence, we conclude that the OOE modulations may result from a combination of (1) stellar spots on either or both bodies, (2) a difference in the reflection indices of the two bodies and (3) systematics introduced by the third light in the aperture.

We consequently remove the OOE variation from the flux and centroid time series using Gaussian Process Regression, before analysing the time series with our MCMC model. We employ a combination of a Matern 3/2 kernel function, a linear kernel and a white noise kernel to model the OOE data of each time series individually. We then extrapolate the model and evaluate it for the eclipse data, resulting in the detrended light curve and centroid curves shown in Figs 8A and 10.

The results of our MCMC model fit are shown in Fig. 8 and Table A1. Fig. A1 shows the resulting posterior distributions. We find a high inclination ($i = 80.14 \pm 0.34^\circ$) of the system and comparable surface brightness ratio ($J = 0.722 \pm 0.010$) of the two components. The radius ratio of the two bodies is estimated to be $R_2/R_1 = 0.2414 \pm 0.0044$.

The example of NG 0522-2518 017220 illustrates how the centroiding technique can aid us in identifying which system in the aperture shows the eclipsing signal. In particular, any visual information on the blend can be input as priors to refine the model fit and establish the physical parameters of the eclipsing system.

3.2.2 NG 0409-1941 020057

The transit-like signal of NG 0409-1941 020057 was initially detected with a period of ~ 0.8 d and width of ~ 3.4 h. However, with additional *NGTS* photometry the true period could be established as ~ 1.61 d, with a ~ 4 per cent primary eclipse and a ~ 0.5 per cent secondary eclipse. The object is well suited for follow-up at $G = 13.3$, $J = 12.1$ and $K = 11.7$. It is identified as 2MASS 04104778-2031575 and GAIA 5091012688012721664, and is located at RA = $04^{\text{h}}10^{\text{m}}47.8^{\text{s}}$, Dec. = $-20^{\text{d}}31^{\text{m}}57.5^{\text{s}}$. The signal, however, is significantly correlated with ~ 5 and ~ 2 mpix centroid shifts in x and y , respectively (see Fig. 11, Table 1).

The source is listed as a single source in various catalogues, including 2MASS and *Gaia* DR1. We investigate the archival images, but find no definite indication of two separate sources, despite a marginal ellipticity of the 2MASS point spread function. The latest release of *Gaia* DR1, however, is incomplete below ~ 4 arcsec separation.⁵ We employ this incompleteness in our MCMC model as an upper limit for uniform priors on the relative CCD position

⁵ See e.g. <https://www.cosmos.esa.int/web/gaia/dr1>, online 2017 May 12.

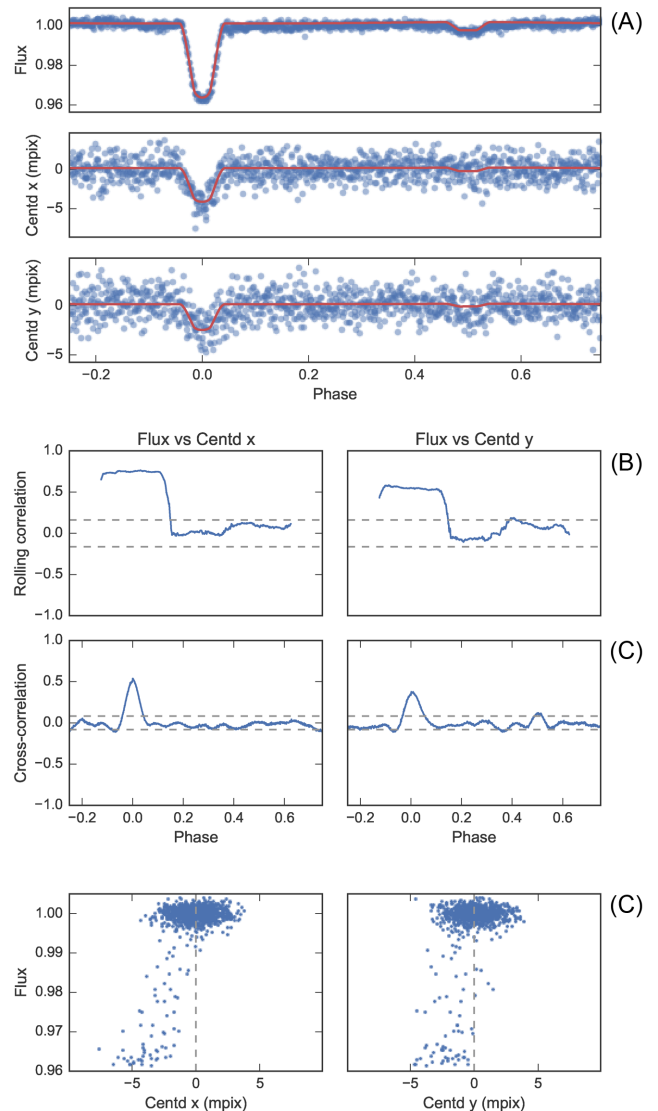


Figure 11. Identification and model fit of a centroid shift correlated to the transit signal in NG 0409-1941 020057. See caption of Fig. 8.

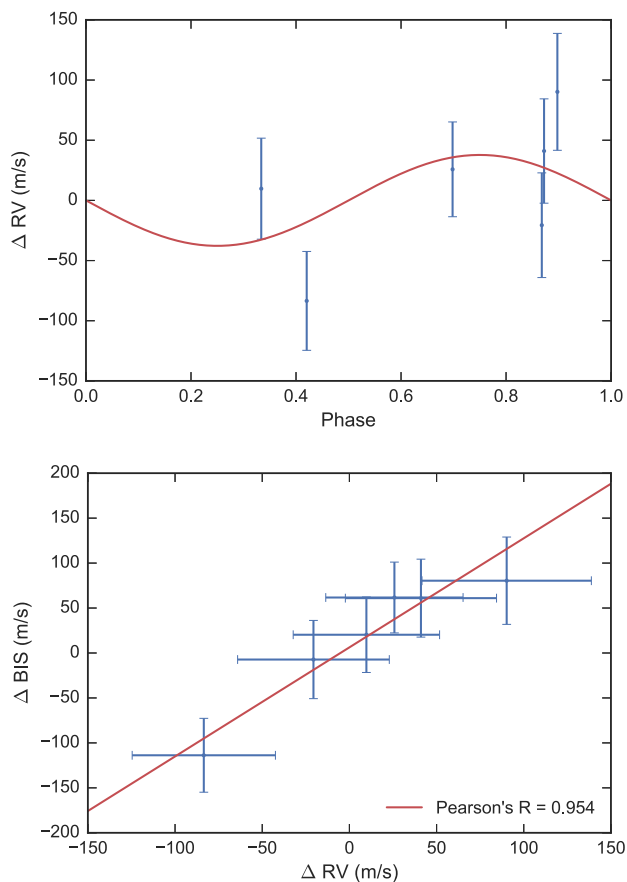
of the background object. Due to the short orbital period and the clear secondary eclipse at phase 0.5, we restrict our MCMC model to circular orbits.

The results of our MCMC analysis can be seen in Figs 11 and A2, and are summarized in Table A1. The object undergoing the eclipse is highly diluted with $D_0 = 0.849^{+0.010}_{-0.015}$, indicating that the signal originates from a blended background object. This background object would hence show an undiluted transit depth of $\delta_e = 24.5^{+2.0}_{-2.6}$ per cent, and has a surface brightness ratio of $J = 0.1061 \pm 0.0065$ and radius ratio of $0.462^{+0.018}_{-0.022}$.

Before the correct period had been established and our centroiding analysis had been performed, six reconnaissance RV measurements were taken using the *Coralie* spectrograph (Queloz et al. 2001) on the Swiss 1.2 m telescope at La Silla Observatory, Chile. These measurements are set out in Table 2. The RV signal for this system, when phase-folded at the true period, shows an in-phase variation of approximately 50 m s^{-1} (see Fig. 12). Such a signal is consistent with what may be expected for a typical hot Jupiter. However the bisectors of the RV cross-correlation functions show a significant correlation with the RV amplitude (see Fig. 12) This

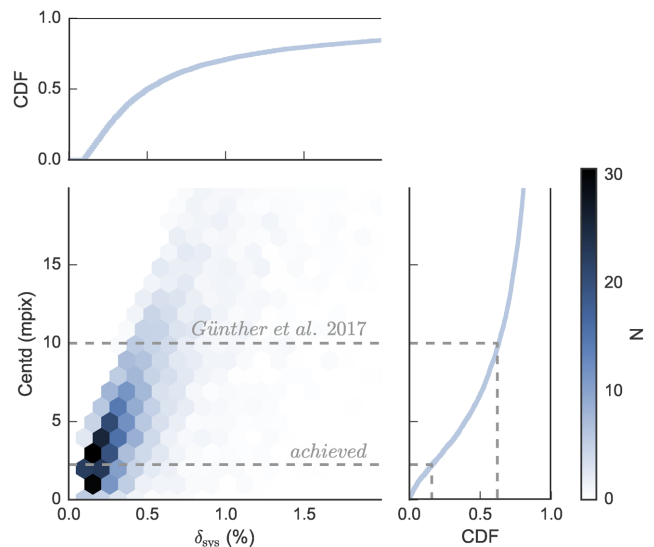
Table 2. *Coralie* radial velocities of NG 0409-1941 020057.

BJD (−2400000)	RV (km s ^{−1})	RV error (km s ^{−1})	BIS (km s ^{−1})
57605.906507	103.93259	0.04857	0.08040
57613.901203	103.88342	0.04335	0.06101
57630.852688	103.75888	0.04108	−0.11377
57632.906217	103.86820	0.03938	0.06171
57634.786375	103.82177	0.04350	−0.00728
57638.748983	103.85213	0.04198	0.02030


Figure 12. *Coralie* radial velocity (RV) measurements of NG 0409-1941 020057. Upper panel: RV signal phase folded on the photometric period and epoch. The solid line indicates the best-fitting Keplerian solution assuming a circular orbit and the photometric period and epoch. Lower panel: RV bisector spans against the measured RV signal. The strong correlation (Pearson's $R = 0.954$) indicates the detected RV variations are due to a diluted spectrum shifting at large amplitudes, such as a background eclipsing binary.

indicates that the variation seen for this target is due to a blended star that is spectroscopically contaminating the cross-correlation function and is moving in phase with the photometric period.

Note that the acceptance of the *Coralie* fibre is ~ 2 arcsec. The results from our MCMC model predict the blended background object to be offset from the target by 0.653 ± 0.040 arcsec in x and 0.396 ± 0.034 arcsec in y , further supporting the hypothesis that the *Coralie* signal is affected by the blend. The evidence provided by both the centroid vetting and the RV data are in agreement and suggest a highly diluted BEB scenario.


Figure 13. The impact of the centroiding algorithm for automated planet candidate vetting in *NGTS*, based on the simulated yield for 4 yr of survey operation. Shown is the total number of BEBs triggering a signal as a function of the measured (diluted) transit depth of the system, δ_{sys} , and the caused centroid shift, $\Delta\xi$, as well as the cumulative distribution function (CDF) of each parameter. Assuming that all centroid signals $>3\sigma$ above the noise level of 0.75 mpix can be detected, the achieved centroiding detection threshold exceeds our previous assumptions by an order of magnitude. This will allow to directly identify ~ 80 per cent of all BEBs.

4 DISCUSSION

4.1 Impact of the centroiding technique on *NGTS* candidate vetting

NGTS is the first ground-based wide-field transit survey to employ the centroid technique for automated and routine candidate vetting. The presented algorithm is already part of the *NGTS* pipeline employed for all detected transit-like signals. We achieve an average centroid precision of 0.75 mpix for all candidates, and as low as 0.25 mpix for individual objects. This precision depends on the photometric data quality as a direct consequence of equations (12) and (18). We therefore expect to observe an increase of the centroiding utility with higher signal-to-noise ratios (SNRs) of the detected transit-like signal.

The obtained centroid precision exceeds the previous assumptions in our yield estimations by an order of magnitude (Günther et al. 2017). The yield simulator is based on a galactic model and the planet occurrence rates estimated by the *Kepler* mission. It considers the measured noise levels and observation window function of *NGTS*, and simulates vetting criteria to identify false positives. We previously assumed all centroid shifts >10 mpix could be detected, and predicted that ~ 38 per cent of all BEBs can be identified with *NGTS* alone. Considering the achieved centroid precision of 0.75 mpix, we update our yield simulator and repeat this analysis. We assume that all centroid signals $>3\sigma$ above the noise level can be detected. This is motivated by the detection of a ~ 2 mpix centroid shift for NG 0409-1941 020057 (Section 3.2.2 Fig. 11). We estimate that this allows to directly identify ~ 80 per cent of all BEBs (see Fig. 13).

To verify these estimations we test our implementation on eclipsing systems from the latest *NGTS* pipeline run. We restrict this comparison to a sample of obvious astrophysical signals with eclipse

depths >2 per cent to avoid the influence of spurious signals. Note that this sample is mostly comprised of undiluted eclipsing binaries due to their high occurrence rates. We find that 16 ± 8 per cent of this sample show a significant correlation between their photometric flux and centroid data. The given confidence interval is the standard error of the mean of a sample of binomial random variables. In comparison, from our simulations we would expect to identify a centroid shift for 12 ± 2 per cent of this sample. These findings are consistent and highlight the expected success of the centroid algorithm for the automated candidate vetting pipeline.

4.2 Investigated candidates

In Sections 3.2.1 and 3.2.2, we demonstrated our method in two case studies. First, visual information on NG 0522-2518 017220 allowed us to identify which blended object is undergoing the eclipse. We identify OOE modulation, which is likely due to either (a) star spots, (b) the reflection effect or (c) systematics from the blended contaminant, or a combination of these effects. The eclipsing system is shown to be a grazing low-mass binary, likely consisting of a *K* star primary and *M* star or brown dwarf secondary.

Second, an analysis of NG 0409-1941 020057 reveals that its signal in fact originates from a highly diluted source, and thus suggests a deep BEB (undiluted depth of $24.5^{+2.0}_{-2.6}$ per cent) as a cause. RV data previously collected with *Coralie* shows a correlation in the bisectors of the RV cross-correlation function, which supports the results of our centroiding method.

For the latter candidate, the blended objects are estimated to be <1 arcsec separated, such that no existing catalogue has resolved the system. This demonstrates the effectiveness of the centroid technique, as it predicts the relative position of blended background objects, which will eventually be confirmable with the resolution of upcoming results from *Gaia*.

The parameter space was thoroughly explored using differential evolution algorithms and multiple MCMC runs from different starting positions. Additionally, care was taken to reach convergence of all walkers and to not fall into local minima. However, systematics in the flux and centroid time series may still be present after the detrending procedure, and could potentially restrict the exploration of the parameter space. This could lead to the underestimation of MCMC posterior likelihood distributions. Hence, a future refinement of the presented work could be the implementation of a joint MCMC model directly incorporating Gaussian Process Regression (see e.g. Pepper et al. 2017; Gillen et al. 2017).

5 CONCLUSION

We developed a comprehensive framework to extract and detrend flux centroid information from *NGTS* data. The introduced algorithms are part of an automated vetting pipeline for all *NGTS* candidates. We achieve an average precision of 0.75 mpix on the phase-folded centroids over an entire field, and 0.25 mpix for specific targets. This enables the identification of systems that are currently too close (<4 arcsec) to be resolved with any photometric or astrometric all-sky survey. Case studies of *NGTS* candidates illustrate that different scenarios can lead to a centroid motion, yet our robust MCMC fitting procedure is able to determine the true origin of a given transit-like signal. In total, we estimate to be able to rule out ~ 80 per cent of all blended variable background objects with *NGTS* data alone. These systems would otherwise have to be followed-up with higher resolution imaging, or may otherwise potentially be

miss-identified as planet candidates. While the centroiding technique has previously been employed for the space-based *Kepler* mission, this is the first time it has been implemented in a ground-based wide-field photometric survey.

ACKNOWLEDGEMENTS

This research is based on data collected under the *NGTS* project at the ESO La Silla Paranal Observatory. *NGTS* is operated with support from a UK Science and Technology Facilities Council (STFC) research grant (ST/M001962/1). This work has further made use of data from the European Space Agency (ESA) mission *Gaia* (<https://www.cosmos.esa.int/gaia>), processed by the *Gaia* Data Processing and Analysis Consortium (DPAC, <https://www.cosmos.esa.int/web/gaia/dpac/consortium>). Funding for the DPAC has been provided by national institutions, in particular the institutions participating in the *Gaia* Multilateral Agreement. Moreover, this publication makes use of data products from the Two Micron All Sky Survey, which is a joint project of the University of Massachusetts and the Infrared Processing and Analysis Center/California Institute of Technology, funded by the National Aeronautics and Space Administration and the National Science Foundation. We also make use of the open-source PYTHON packages NUMPY (van der Walt, Colbert & Varoquaux 2011), SCIPY (Jones et al. 2001), MATPLOTLIB (Hunter 2007), PANDAS (McKinney 2010), EMCEE (Foreman-Mackey et al. 2013), CORNER (Foreman-Mackey 2016) and EB (Irwin et al. 2011). The latter is based on the previous JKTEBOP (Southworth, Maxted & Smalley 2004a; Southworth et al. 2004b) and EBOP codes (Popper & Etzel 1981) and models by Etzel (1981), Mandel & Agol (2002), Binnendijk (1974a,b) and Milne (1926). DJA is funded under UK Science and Technology Facilities Council (STFC) consolidated grant reference ST/P000495/1. MNG is supported by the UK Science and STFC award reference 1490409 as well as the Isaac Newton Studentship.

REFERENCES

- Batalha N. M., Kepler Team 2012, American Astronomical Society Meeting Abstracts #220, p. 306.01
 Batalha N. M. et al., 2010, *ApJ*, 713, L103
 Bennett G. G., 1982, *J. Navig.*, 35, 255
 Binnendijk L., 1974a, *Vistas Astron.*, 16, 61
 Binnendijk L., 1974b, *Vistas Astron.*, 16, 85
 Bryson S. T. et al., 2013, *PASP*, 125, 889
 Cameron A. C., 2012, *Nature*, 492, 48
 Chazelas B. et al., 2012, in Stepp L. M., Gilmozzi R., Halleds H. J., eds, SPIE Conf. Ser., Vol. 8444, Ground-based and Airborne Telescopes IV. SPIE, Bellingham, p. 84440E
 Etzel P. B., 1981, in Carling E. B., Kopal Z., eds, Photometric and Spectroscopic Binary Systems. Reidel, Dordrecht, p. 111
 Fisher R. A., 1921, *Metron*, 1, 3
 Foreman-Mackey D., 2016, *J. Open Source Softw.*, 24, 2
 Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, *PASP*, 125, 306
 Fressin F. et al., 2013, *ApJ*, 766, 81
 Gaia Collaboration, 2016a, *A&A*, 595, A1
 Gaia Collaboration, 2016b, *A&A*, 595, A2
 Gelman A., Rubin D. B., 1992, *Stat. Sci.*, 7, 457
 Gillen E., Hillenbrand L. A., David T. J., Aigrain S., Rebull L., Stauffer J., Cody A. M., Queloz D., 2017, *ApJ*, preprint ([arXiv:1706.03084](https://arxiv.org/abs/1706.03084))
 Günther M. N., Queloz D., Demory B.-O., Bouchy F., 2017, *MNRAS*, 465, 3379
 Hunter J. D., 2007, *Comput. Sci. Eng.*, 9, 90

- Irwin M. J. et al., 2004, in Quinn P. J., Bridger A., eds, Proc. SPIE Vol. 5493, Optimizing Scientific Return for Astronomy through Information Technologies. SPIE, Bellingham, p. 411
- Irwin J. M. et al., 2011, ApJ, 742, 123
- Jones E. et al., 2001, SciPy: Open Source Scientific Tools for Python, <http://www.scipy.org/>
- Mandel K., Agol E., 2002, ApJ, 580, L171
- McCormac J., Pollacco D., Skillen I., Faedi F., Todd I., Watson C. A., 2013, PASP, 125, 548
- McKinney W., 2010, in van der Walt S., Millman J., eds, Proceedings of the 9th Python in Science Conference, AQR Capital Management, pandas: a python data analysis library, p 51
- Milne E. A., 1926, MNRAS, 87, 43
- Pepper J. et al., 2017, AJ, 153, 177
- Popper D. M., Etzel P. B., 1981, AJ, 86, 102
- Queloz D. et al., 2001, The Messenger, 105, 1
- Skrutskie M. F. et al., 2006, AJ, 131, 1163
- Southworth J., Maxted P. F. L., Smalley B., 2004a, MNRAS, 351, 1277
- Southworth J., Zucker S., Maxted P. F. L., Smalley B., 2004b, MNRAS, 355, 986
- Storn R., Price K., 1997, J. Global Optim., 11, 341
- van der Walt S., Colbert S. C., Varoquaux G., 2011, Comput. Sci. Eng., 13, 22
- Wheatley P. J. et al., 2013, in EPJ Web Conf., 47, 13002

APPENDIX

Table A1. Parameters of the blended eclipsing systems NG 0522-2518 017220 and NG 0409-1941 020057.

	NG 0522-2518 017220	NG 0409-1941 020057
<i>Catalogue values</i>		
Coordinates	RA = 05 ^h 23 ^m 31 ^s .6 Dec. = −25 ^d 08 ^m 48 ^s .4	RA = 04 ^h 10 ^m 47 ^s .8 DEC = −20 ^d 31 ^m 57 ^s .5
2MASS ID	2MASS 05233161-2508484	2MASS 04104778-2031575
<i>Gaia</i> ID	GAIA 2957881682551005056	GAIA 5091012688012721664
Magnitudes	$G = 13.6, J = 12.6, K = 12.2$	$G = 13.3, J = 12.1$ and $K = 11.7$
Colour	$J - K = 0.4$	$J - K = 0.4$
<i>Fitted parameters</i>		
Δx	Relative CCD x position of the blend in pixel	0.1312 ± 0.0071
Δy	Relative CCD y position of the blend in pixel	0.0798 ± 0.0064
F_0	Offset in normalized flux	$(9.68 \pm 0.50) \times 10^{-4}$
$\xi_{x,0}$	Offset in centroid in x in pixel	$(1.93 \pm 0.47) \times 10^{-4}$
$\xi_{y,0}$	Offset in centroid in y in pixel	$(1.61 \pm 0.44) \times 10^{-4}$
D	Dilution	$0.849^{+0.010}_{-0.015}$
J	surface brightness ratio	0.1061 ± 0.0065
$(R_1 + R_2)/a$	Sum of radii over semimajor axis	0.2634 ± 0.0031
R_2/R_1	Ratio of radii	$0.462^{+0.018}_{-0.022}$
$\cos i$	Cosine of the inclination	$0.022^{+0.021}_{-0.015}$
P	Period (s)	138 852.0 ± 1.4
T_0	Epoch (s)	59 237 250 ± 34
e	Eccentricity	0.0 (fixed)
ω	Argument of periastron in degree	0.0 (fixed)
$\sigma(F)$	Error on normalized flux	$(7.805 \pm 0.037) \times 10^{-3}$
$\sigma(\xi_x)$	Error on centroid in x	$(8.383 \pm 0.041) \times 10^{-3}$
$\sigma(\xi_y)$	Error on centroid in y	$(8.468 \pm 0.041) \times 10^{-3}$
<i>Derived parameters</i>		
i	Inclination (deg)	$88.61^{+0.86}_{-1.26}$
R_1/a	Radius of the primary over semimajor axis	$0.1800^{+0.0037}_{-0.0020}$
R_2/a	Radius of the secondary over semimajor axis	$0.0830^{+0.0030}_{-0.0033}$
T_2	Midpoint of secondary eclipse (s)	59 306 676 ± 36
$T_{\text{dur},1}$	Duration of primary eclipse (s)	281 100 ± 3600
$T_{\text{dur},2}$	Duration of secondary eclipse (s)	281 100 ± 3600
$\delta_{1,\text{dil}}$	Diluted depth of the primary eclipse	$3.628^{+0.032}_{-0.047}$
$\delta_{2,\text{dil}}$	Diluted depth of the secondary eclipse	0.238 ± 0.023
$\delta_{1,\text{undil}}$	Undiluted depth of the primary eclipse	$24.5^{+2.0}_{-2.6}$
$\delta_{2,\text{undil}}$	Undiluted depth of the secondary eclipse	2.10 ± 0.26
Δx_{sky}	Relative sky position of the blend (arcsec)	0.653 ± 0.040
Δy_{sky}	Relative sky position of the blend (arcsec)	0.396 ± 0.034

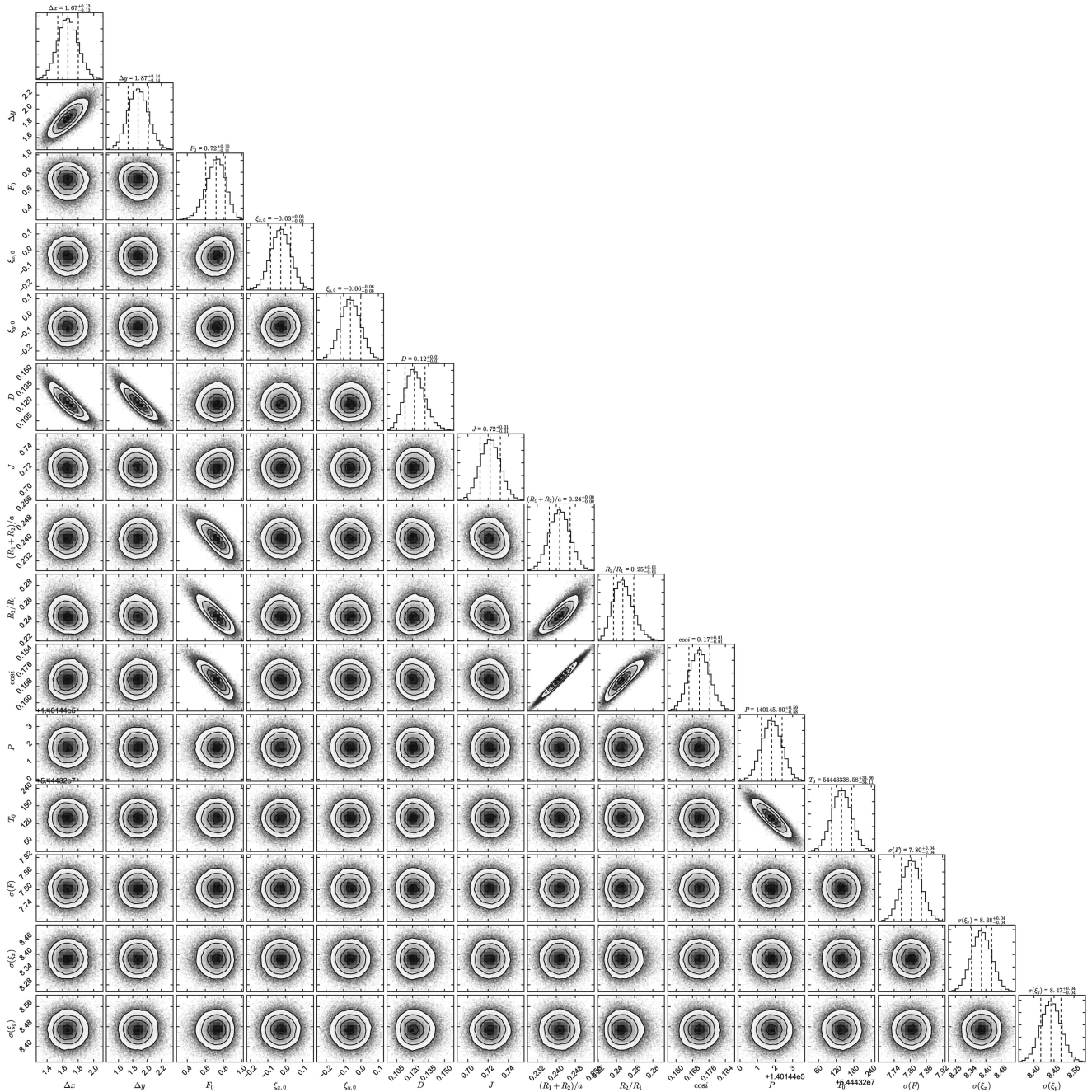


Figure A1. Posterior likelihood distributions for all parameters of the MCMC fit to NG 0522-2518 017220. A description of the model and all parameters can be found in Section 3.1.4.

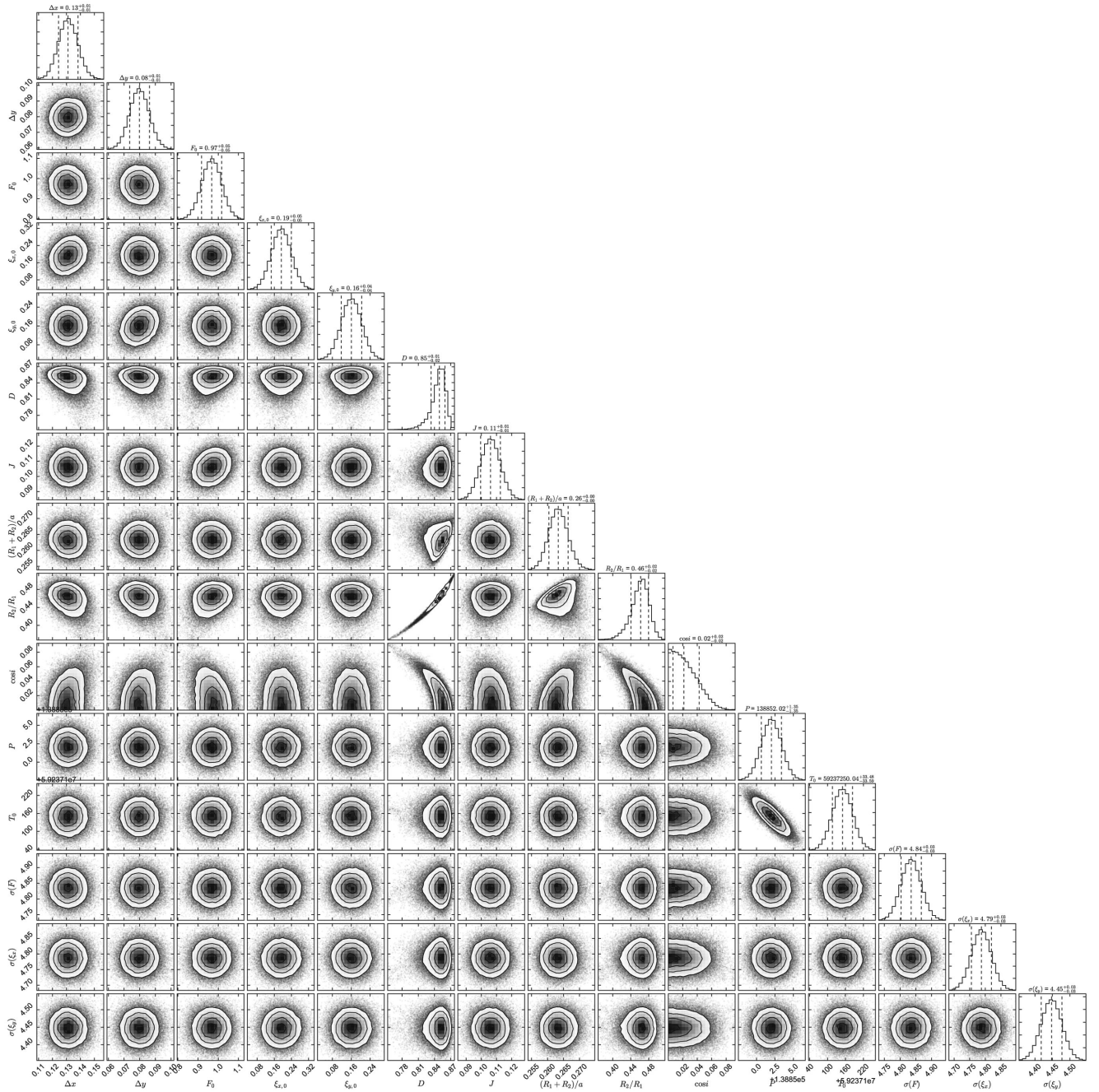


Figure A2. Posterior likelihood distributions for all parameters of the MCMC fit to NG 0409-1941 020057. A description of the model and all parameters can be found in Section 3.1.4.

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