Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **465,** 3379–3389 (2017)
Advance Access publication 2016 November 10

doi:10.1093/mnras/stw2908

A new yield simulator for transiting planets and false positives: application to the Next Generation Transit Survey

Maximilian N. Günther,[1]★ Didier Queloz,[1] Brice-Olivier Demory[1] and Francois Bouchy[2]

[1]*Astrophysics Group, Cavendish Laboratory, J.J. Thomson Avenue, Cambridge CB3 0HE, UK*
[2]*Observatoire de Genève, Université de Genève, 51 Ch. des Maillettes, CH-1290 Sauverny, Switzerland*

Accepted 2016 November 7. Received 2016 November 7; in original form 2016 March 16

ABSTRACT

We present a yield simulator to predict the number and characteristics of planets, false positives and false alarms in transit surveys. The simulator is based on a galactic model and the planet occurrence rates measured by the *Kepler* mission. It takes into account the observation window function and measured noise levels of the investigated survey. Additionally, it includes vetting criteria to identify false positives. We apply this simulator to the Next Generation Transit Survey (NGTS), a wide-field survey designed to detect transiting Neptune-sized exoplanets. We find that red noise is the main limitation of NGTS up to 14 mag, and that its obtained level determines the expected yield. Assuming a red noise level of 1 mmag, the simulation predicts the following for a 4-yr survey: 4 ± 3 Super-Earths, 19 ± 5 Small Neptunes, 16 ± 4 Large Neptunes, 55 ± 8 Saturn-sized planets and 150 ± 10 Jupiter-sized planets, along with 4688 ± 45 eclipsing binaries and 843 ± 75 background eclipsing binaries. We characterize the properties of these objects to enhance the early identification of false positives and discuss follow-up strategies for transiting candidates.

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

1 INTRODUCTION

Exoplanets transiting their host star give insight into their formation, bulk composition and atmospheric properties. Dedicated wide-field transit surveys, both from the ground (e.g. HAT, Bakos et al. 2002; WASP, Pollacco et al. 2006) and from space (e.g. *CoRoT*, Baglin et al. 2002; *Kepler*, Borucki et al. 2010), have discovered ∼2700 exoplanets.[1]

Transit-like shape variability in the light curve may not only be caused by planets. False alarms introduced by correlated noise may cause a time-periodicity with a pattern similar to a transit shape. In addition, false-positive transit events related to an eclipsing astrophysical object can cause a transit signal of small amplitude that may be interpreted as a planetary transit (see, e.g. Cameron 2012). Eclipsing binaries (EBs) can be very expensive in telescope time to follow up. First, low-mass companions such as brown dwarfs and very low mass stars can be of similar size as gas giant planets. Distinguishing them from planets necessitates radial velocity follow-up to measure the mass, and may be aided by measuring ellipsoidal effects or obtaining colour information during transit and eclipse. Second, EBs with grazing events lead to transit depths mimicking a planet-sized object even if the secondary is significantly larger. Another class of false positives are background eclipsing binaries (BEBs), which are faint and distant EBs that are aligned along the line of sight behind a bright target star and hence diluted. The dilution reduces the apparent transit depth on to a planet-like scale, making BEBs one of the most difficult false positives to rule out. Similar to this are triple and higher order star systems with one or more pairs of stars eclipsing, referred to as hierarchical EBs. In wide-field transit surveys, false positives can be up to two orders of magnitude more prevalent than planets (see, e.g. Almenara et al. 2009; Hartman, Bakos & Torres 2011).

Estimating the yield of a transit experiment provides a way to assess the false positive to planet ratio in detail. Brown (2003) raised awareness of the contamination impact by false positives in upcoming surveys, but most yield simulations have focused on the number of planets only. As one of the first, Brown & Latham (2008) applied false-positive models to predict the yield of the *Transiting Exoplanet Survey Satellite* (*TESS*) mission. Recently, Sullivan et al. (2015) estimated the planet yield, false-positive contamination rates and the success of ad hoc vetting methods for the *TESS* mission. In addition to enabling insight into future surveys, yield simulations

★ E-mail: mg719@cam.ac.uk
[1] http://exoplanetarchive.ipac.caltech.edu/ (2016 August 17).

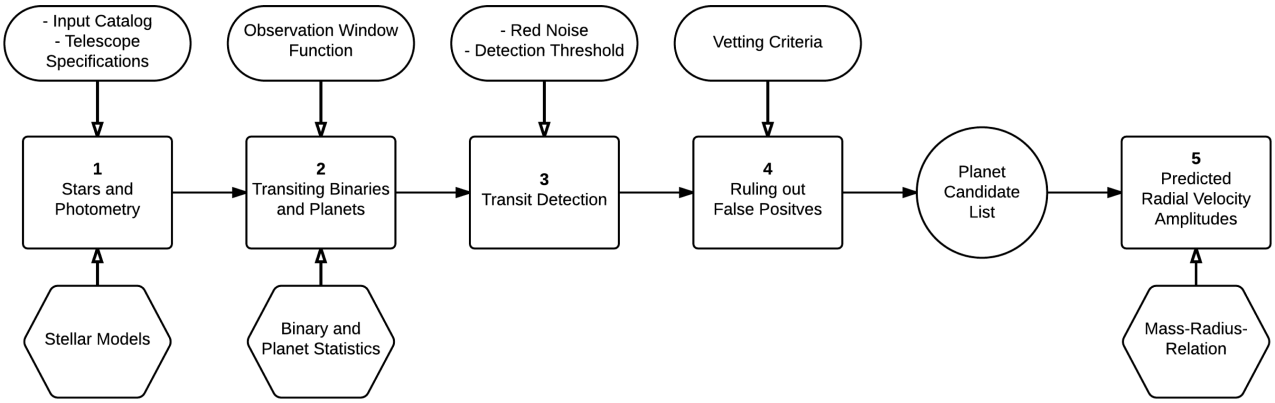

Figure 1. The simulation can be structured in five main modules, containing user input and settings (ovals), observed parameters and models (hexagons), steps of calculations (rectangles) and output files (circle). All user input is given in a parameter file when starting the simulation and is displayed here in respect to where it contributes to the simulation. The numbering refers to Sections 2.1–2.5.

can be used to evaluate how well current instruments achieve their possibilities, as well as how current observing strategies may be optimized.

Here, we develop a yield simulator with the goal of estimating the planet merit and the impact of false signals applicable to any upcoming transit survey. The simulations specifically take into account red noise, false alarms and false positives. In order to assess the impact of various observing strategies, the simulation takes as input: target list, telescope parameters, bandpass, field of view, cadence, noise models and detection criteria. To mimic the vetting processes for false positives, we implement methods examining the transit parameters (depth, shape and duration), secondary eclipses, centroid movement and the feasibility of planet follow-up and characterization.

We apply our simulator to estimate the yield of planets and false positives for the recent Next Generation Transit Survey (NGTS; Chazelas et al. 2012; Wheatley et al. 2013; Wheatley et al., in preparation). Previous ground-based facilities have limited photometric precision, for example 10–50 mmag for HAT (Bakos et al. 2002) and 3–30 mmag for WASP (Pollacco et al. 2006), and are hence more prone to detect hot Jupiters. NGTS is designed to be the first ground-based exoplanet survey to reach sub-mmag photometry. It aims at detecting transiting Neptunes with short orbital periods around small stars. The survey had its first light in early 2015 in Paranal, Chile,[2] and started its full science operation in early 2016. The facility consists of 12 independent 20 cm telescopes with a 7.4 deg^2 each; the total field of view adds up to 88.8 deg^2, similar to *Kepler*. NGTS covers a new field of this size every few months, allowing it to survey many bright stars for short-orbit planets. The sensitivity is optimized between 500 and 900 nm to maximize observation efficiency of K and early-M stars.

We organize this paper in two major parts. Part 1 describes the computational layout and the mechanisms of the simulations, which are adoptable to any transit survey. Part 2 applies the simulations to the example of NGTS. We describe our simulations and the set of priors in Section 2. In Section 3, we describe the validation process of our code using the results from *Kepler*. In Section 4, we examine the case of NGTS, the effects of different red noise levels and detection criteria, and the expected planets and false positives. We estimate the feasibility to identify false positives with NGTS' photometric data alone, and provide an outlook into the necessary follow-up facilities for planet candidates. Finally, we discuss our findings and conclude this work in Sections 5 and 6.

[2] http://www.eso.org/public/news/eso1502/ (2016 August 17).

2 LAYOUT OF THE SIMULATIONS

The overall sketch of the simulation layout as described below is shown in Fig. 1. The simulation input contains a list of stars in the field of view as well as the instrument specifications. An example input file can be found in Appendix A. From this we calculate the collected flux for each object (Section 2.1). We first randomly assign host stars with planets, and compute the signals of transiting planets as well as eclipsing binaries (depth, duration, shape, visibility and possible dilution; Section 2.2). Second, we calculate the total noise for each observation and compute which systems would be detectable (Section 2.3). Third, we rule out false positives if detectable from photometric data (Section 2.4). Finally, we assess the feasibility to follow-up planetary signals with radial velocity instruments (Section 2.5).

2.1 Stars and photometry

The simulation considers an input catalogue with information about the multiplicity, radius, mass, effective temperature and magnitude of all stars in the field of view. The input catalogue is built using the TRILEGAL galaxy model (Girardi et al. 2005), up to $V = 23$. We keep the pre-set adjustments of TRILEGAL referring to the standard Milky Way model and simulate binaries with a fraction of 33 per cent (Raghavan et al. 2010). While this value was estimated for solar-type stars, it is also consistent with predictions for low-mass stars, given the dispersion reported in the literature (see, e.g. Duchêne & Kraus 2013, and references therein). The binary mass ratio q of the secondary and primary mass, $q = M_s/M_p$, is drawn uniformly between 0.08 and 1. To include higher order multiples, we randomly select single stars and assign them to be higher order multiples. This way, at the end the input catalogue consists of 56 per cent single stars, 33 per cent binaries is defined by modelsand 11 per cent higher order multiples (Raghavan et al. 2010).

From this input catalogue we identify a target list of stars according to magnitude and spectral type as listed in Table 1. We assume all input catalogue stars are randomly distributed across the field of view. Then we compute the photometric flux using the effective temperatures and V-band magnitudes of the input catalogue and the transmission function of the telescope. The stellar parameters are

Table 1. Settings for NGTS in the yield simulation. The target list is chosen from all FGKM stars bright enough for follow-up. The presented noise values and CCD parameters are based on observed data from both test and commissioning phases (Walker 2013, and private correspondence within the NGTS consortium).

Target list	Noise levels
$V < 15$	$N_{\rm sky}^{\rm new\ moon} = 65\ {\rm e^-\ s^{-1}\ pixel^{-1}}$
$R_* < 2\ {\rm R_\odot}$	$N_{\rm sky}^{\rm full\ moon} = 600\ {\rm e^-\ s^{-1}\ pixel^{-1}}$
$T_{\rm eff} < 10^5$ K	$N_{\rm sky}^{\rm median} = 125\ {\rm e^-\ s^{-1}\ pixel^{-1}}$
$\log g < 6.5$	$N_{\rm dark} = 0.06\ {\rm e^-\ s^{-1}\ pixel^{-1}}$
	$N_{\rm read} = 10\ {\rm e^-\ pixel^{-1}}$
	$\sigma_{\rm red} = 0$–2 mmag
CCD parameters	
$t_{\rm exp} = 10$ s	
$t_{\rm read} = 1.49$ s	

converted into photometric flux using spectrophotometric reference stars from Pickles (1998). The zero-point of the V band is defined by models of Vega from the Kurucz atlas (Kurucz 1993). The Johnson V-band model is adopted from Buser & Kurucz (1978). Any other stars lying in the photometric aperture of the target star are considered as background stars for the particular target.

2.2 Transiting binaries and planets

2.2.1 Binaries

For all input catalogue stars identified as binaries, we draw orbital periods P in days from a log-normal distribution with mean of 4.8 and a standard deviation of 2.3 (Duquennoy & Mayor 1991). The eccentricities follow a uniform distribution with a maximum eccentricity given by $e_{\rm max} = 0.4 \log P - 0.2$ (approximated from Raghavan et al. 2010). Only detached eclipsing binaries are considered using the Roche limits as criteria. Contact eclipsing binaries are evident from the light curves and can be readily ruled out. We do not consider eclipses within triple or higher order hierarchical eclipsing binaries.

We define orbit and transit parameters following Winn (2011). In a binary system, if star 1 is fully transited by star 2, we compute the transit depth as

$$\delta_1 = \frac{A}{A_1} \frac{F_1}{F_1 + F_2}, \tag{1}$$

where $A = \min(A_1, A_2)$. In the case of a grazing eclipses, we replace A with the overlapping area of two circles with radii $R_{1,2}$ and midpoint distance y.

2.2.2 Planets

The planet occurrence rates are based on the results of the *Kepler* mission as by Fressin et al. (2013) for FGK stars and by Dressing & Charbonneau (2015) for small planets around M dwarfs. While recent results made progress on long period and small planets (Burke et al. 2015), in the planet regime targeted by all-sky surveys Fressin et al. (2013) provide to date still the most complete study. The occurrence rates denote the average number of planets per star binned by planet radius and orbital period. In Fressin et al. (2013), they are discretely sampled in radius and period: Earths ($<1.25\ {\rm R_\oplus}$), Super-Earths ($<2\ {\rm R_\oplus}$), Small Neptunes ($<4\ {\rm R_\oplus}$), Large Neptunes ($<6\ {\rm R_\oplus}$) and giant planets ($<22\ {\rm R_\oplus}$), as well as 10 logarithmically spaced period ranges. We randomly assign a value for period and radius within each discrete interval. The period is drawn within each interval from a logarithmic distribution. We draw the radius within each interval from a uniform distribution, except for the last interval ($>6\ {\rm R_\oplus}$), in which we draw from a logarithmic distribution for consistency with empirical findings (see, e.g. Grether & Lineweaver 2006, and Exoplanet Archive[3]).

We assign planets to all stars, single or binary, in the input catalogue according to the occurrence rate (which may be greater than 1), except when binary systems have orbital periods shorter than 5 d. Although these short-orbit binary systems are frequent (Slawson et al. 2011) and theoretically can host circumbinary planets, these planets are likely undetectable (Muñoz & Lai 2015). These constraints lead to a smaller number of assigned planets around close-in binaries than for wide binaries or single stars. Note that there is no evidence that wide binaries affect the occurrence rate of short-orbit planets (Deacon et al. 2016). We do not consider planets around the 11 per cent of higher order multiples in our target list, as in most cases their transit signals will be diluted too much by the other stars in the system to be detectable. In cases where there is more than one planet assigned to the same star, the orbital parameters of planets are drawn completely independently. This avoids influencing the yield by setting criteria for multiplanetary systems while leading to the same statistical average over all stars of the field. For all remaining planets in binary systems we compute stability criteria following Holman & Wiegert (1999) and reject planets in unstable orbits (<1 per cent of all planets).

We investigate the impact of eccentric planetary orbits by assigning various mean eccentricities between 0 and 0.5 and find a consistent planet yield as with circular orbits. We therefore employ circular orbits for all planets considering the short-orbit sensitivity of all-sky transit surveys.

We define orbit and transit parameters following Winn (2011). For grazing geometry the overlapping area of the two objects is computed instead. If the host system is part of a binary system, we account for dilution by the other star.

2.2.3 Observation window function

For each transiting planet we estimate the number of transits that can be detected. The time of the first transit is set randomly between 0 and the orbital period P. To compute the observation window, we calculate the average visibility duration per night of each field. We implement average weather information at the telescope location and reject a certain fraction of nights to simulate bad conditions.

2.2.4 Dilution

Background stars in the aperture of a target star affect its extracted photometry. First, they decrease the transit depth of a planet orbiting the target star and might make the signal undetectable. Second, if the target star or any background star is an eclipsing binary, its eclipse depth will be decreased and it may appear planet-like. The dilution D for a certain source is the ratio of its stellar flux F_0 and the total flux in the aperture $\sum_i F_i$. We evaluate D for each system, with F_0 being either the flux of a single star or the whole multiple system in which the transit/eclipse occurs. The theoretical transit depth δ_0 from Sections 2.2.1 to 2.2.2 is then reduced to the measured transit

[3] http://exoplanetarchive.ipac.caltech.edu/ (2016 August 17).

depth,

$$\delta = \delta_0 \frac{F_0}{\sum_i F_i}. \quad (2)$$

2.3 Transit detection

The total noise of the extracted photometric flux can be described as a composition of uncorrelated and correlated noise, referred to as white and red noise (Pont, Zucker & Queloz 2006). White noise scales with exposure time, aperture and flux. We calculate it as the sum of individual white noise sources:

$$\sigma^2_{\rm white} = t_{\rm exp} N_\star + n_{\rm pix} \left(t_{\rm exp} N_{\rm sky} + t_{\rm exp} N_{\rm dark} + N^2_{\rm read} \right) + \sigma^2_{\rm scint}, \quad (3)$$

where $n_{\rm pix}$ is the number of pixels in the aperture and $t_{\rm exp}$ the exposure time. N_* is the photon count received from a given source and $N_{\rm sky}$ the sky background. $N_{\rm dark}$ and $N_{\rm read}$ are the counts contributed by dark and readout noise. $\sigma_{\rm scint}$ describes the scintillation noise, evaluated following Dravins (1998).

White noise averages out by the square root of the number of exposures per transit, $N_{\rm exp}$. The total noise in one transit is given by the squared quadratic sum of the binned white noise and the red noise. Red noise is composed of various sources that are not entirely known, such as weather patterns (if ground-based), correlated astrophysical and instrumental noise and software influence. We assume the driving red noise patterns are correlated on one-night time-scales, and are to first order uncorrelated over time-scales of multiple days. Therefore, red noise of measurements on different days (or different transits) averages out, and we get the total noise

$$\sigma^2_{\rm tot} = \frac{\sigma^2_{\rm white}/N_{\rm exp} + \sigma^2_{\rm red}}{N_{\rm tr}} \quad (4)$$

for a phase-folded light curve with $N_{\rm tr}$ transit events.

We require two criteria for the detection of a transit signal. First, at least three transit events must be visible, $N_{\rm tr} \geq 3$ (Section 2.2.3). Second, the signal-to-noise ratio, SNR, of the phase-folded light curve must exceed a minimum requirement. We refer to this minimum SNR as the *detection threshold*, in the following denoted by the acronym DT:

$$\mathrm{SNR} = \frac{\delta_{\rm tr/occ}}{\sigma_{\rm tot}} > \mathrm{DT}, \quad (5)$$

where $\delta_{\rm tr/occ}$ denotes the depth of the transit or occultation signal.

2.4 Ruling out false positives

Signals caused by eclipsing binaries are the most common astrophysical false positives in wide-field surveys for transiting planets (Cameron 2012, review). To identify this configuration we define criteria based on the work of the *Kepler* team (Batalha et al. 2010, 2012; Bryson et al. 2013), which can potentially be implemented in any survey's pipeline. We consider that a transit signal event originates from a false positive under the following conditions.

(i) We measure transit depths greater than a given threshold ($2R_{\rm Jup}$), assuming the radius of the host star is known.

(ii) We detect a secondary eclipse $\delta_{\rm occ}$ and clearly distinguish it from the transit signal $\delta_{\rm tra}$ if both criteria are met:

$$\frac{\delta_{\rm occ}}{\sigma_{\rm occ}} > \mathrm{DT} \quad \text{and} \quad \frac{\delta_{\rm tra} - \delta_{\rm occ}}{\sqrt{\sigma^2_{\rm tra} + \sigma^2_{\rm occ}}} > \mathrm{DT}. \quad (6)$$

As introduced in Section 2.3, DT denotes the detection threshold, and $\sigma^2_{\rm tra}$ and $\sigma^2_{\rm occ}$ the noise of the transit and occultation signals.

(iii) There are ellipsoidal variations in their light curve, a typical feature of close binaries. We use the criteria from Sullivan et al. (2015) for the simulations of the *TESS* yield and employ the model of photometric variations from Mazeh (2008), using limb darkening from Claret, Hauschildt & Witte (2012, 2013) as well as gravity darkening from Lucy (1967) to calculate the signal caused by ellipsoidal variations.

(iv) Their in-/egress time equals the transit duration and the transit depth is less than 10 per cent reduced by dilution, such that the V-shape remains clearly detectable in the light curve. Given that planet transits can be V-shaped as well, this criterion cannot be used alone (see Section 4.5).

(v) If during their eclipse the centre of flux in the aperture (centroid) shifts more than a given fraction of a pixel. This aims to identify background eclipsing binaries.

(vi) Their transit duration is significantly different from what is expected for a planet. For this purpose we calculate the theoretical transit duration of a gas giant planet ($2\,R_{\rm Jup}$) that orbits the target star with the detected period. We approximate the orbit to be circular, which is justified for short-period planets. The impact parameter b dictates the transit duration and is unknown. However, it is possible to estimate a maximum transit duration by setting $b = 0$. We identify whether the detected transit duration is greater than this maximum value.

2.5 Predicted radial velocity amplitudes

Assessing the feasibility of radial velocity follow-up for transit surveys is important to anticipate follow-up strategies. We assume a radius versus mass relationship following Weiss & Marcy (2014) for objects below $3\,\mathrm{R}_\oplus$. For planets of 3–$6\,\mathrm{R}_\oplus$ we adopt a Neptune density of $\rho_{\rm Neptune} = 1.64\,\mathrm{g\,cm^{-3}}$. For 6–$11\,\mathrm{R}_\oplus$ we adopt a Jupiter density of $\rho_{\rm Jupiter} = 1.33\,\mathrm{g\,cm^{-3}}$, and for larger planets a Jupiter mass of $M_{\rm Jup} = 1.898 \times 10^{27}$ kg as a mean value. To reflect the intrinsic diversity of planetary composition and structure, we distribute the masses following a log-normal distribution with deviation 0.5 around the mean. Finally, we estimate the radial velocity semi-amplitude using the parameters assigned to the planet systems and compare to the limits of current facilities.

3 VERIFYING THE SIMULATION ON THE EXAMPLE OF *KEPLER*

Using our simulations, we estimate the yield of *Kepler* and compare our results to the *Kepler* candidates and confirmed *Kepler* planets. We use two approaches: (1) we draw the target stars from version 10 of the Kepler Input Catalog (Brown et al. 2011), and use the TRILEGAL galaxy model to simulate background stars and distribute them randomly in the *Kepler* field of view; (2) we use solely TRILEGAL and create an ad hoc target list from all FGKM stars brighter than $V = 15$. In both cases the CCD parameters and noise levels of *Kepler* are adopted from Gilliland et al. (2011) and the *Kepler* bandpass from Koch et al. (2010).

We find that the two approaches are consistent in their yield predictions. In both cases we obtain a total of ∼3000 planets to be discovered with *Kepler*. These comprise ∼600 Earths, ∼1000 Super-Earths, ∼1000 Small Neptunes, ∼100 Large Neptunes and ∼200 giant planets. Currently there are ∼4700 objects listed as

Kepler candidates, out of which ∼2300 have so far been confirmed as planets.[4]

When comparing the statistics of *Kepler* candidates and confirmed *Kepler* planets with the simulated yield, we find a good agreement. First, the total number of simulated planets agrees with the actual findings. Second, the balance of planet types is in agreement with the statistics drawn from both the *Kepler* candidates as well as the confirmed *Kepler* planets. Hence, our yield results for *Kepler* verify our models and assumptions in the simulations. The code solely requires changing a set of priors to be used for other transit surveys. These priors contain the target list, telescope bandpass, noise levels, as well as the observation window and strategy.

4 ESTIMATING THE YIELD OF NGTS

4.1 NGTS facility, target list and background stars

NGTS is based at the European Southern Observatory's Paranal Observatory in Chile. The facility is made of 12 fully robotic 20-cm telescopes with a 7.4 deg^2 field of view each and can be operated independently. Each CCD is a deep depleted 2k × 2k Ikon-L produced by Andor, with pixel size of 13.5 µm (4.97 arcsec). The telescopes have a constant point spread function (PSF) full width at half-maximum (FWHM) of 12 µm across the field of view. More details may be found in Chazelas et al. (2012) and Wheatley et al. (2013).

In our yield simulation we consider the situation where each telescope observes a separate neighbouring field, such that the total field of view is 88.8 deg^2 combined (hereafter called NGTS-field). We assume the survey covers three different NGTS-fields per year within 4 yr operation of the mission. In total, 12 NGTS-fields will be observed. In Paranal, in average 78 per cent of the night time is of photometric quality. A typical night lasts 10.5 h during winter and 7.5 h during summer.[5] Considering the observation duration of an NGTS-field is 4 months and detectable transit periods are less than 2 weeks, we may assume all phase-folded light curves will be randomly uniformly sampled. We select a typical NGTS-field at $l = 285°$, b $= +20°$ which is representative of the targeted stellar population. The simulation input file we employ to model NGTS can be found in Appendix A.

4.2 Red noise as the dominant limitation

We compute the white noise using equation (3) and the parameters shown in Table 1 (see also Appendix A). Based on the design of NGTS, we consider circular apertures with a 3 pixel radius. The sky noise varies strongly with lunar phase; we adopt the median value during a lunar cycle. We compute the scintillation noise using Dravins (1998) for 20-cm telescopes, an average airmass of 1.5 and location at 2400 m above sea level.

Scintillation noise is the main white noise component for a single exposure for stars brighter than $V = 11$ (Fig. 2). For fainter targets, stellar and background noise becomes driving factors, with the latter dominating at the faint end for $V > 13.5$.

Correlated noise and systematics, referred to as red noise, affect photometric measurements on time-scales comparable to the transit duration (few hours). NGTS tests in Geneva and La Palma demonstrated its ability to achieve 1 mmag sensitivity and better (Wheatley et al. 2013). The upper panel of Fig. 3 illustrates the white noise

[4] http://exoplanetarchive.ipac.caltech.edu/ (2016 August 17).
[5] https://www.eso.org/sci/facilities/paranal (2016 August 17).

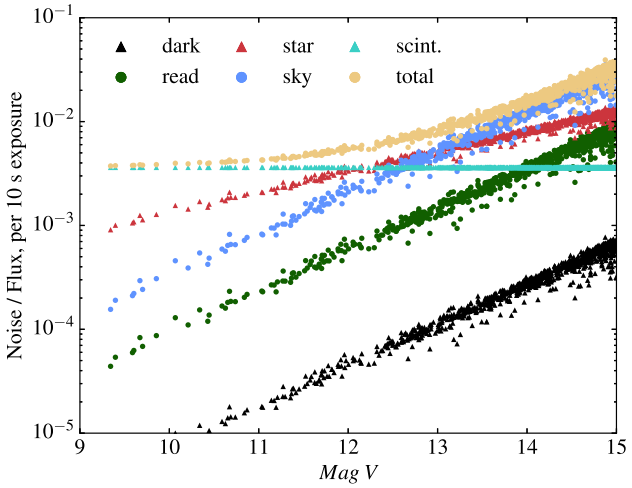

Figure 2. NGTS' white noise components. Values are calculated for each star in the simulated run time using equation (3), the stellar flux (Section 2.1) and empirical values for sky, dark and readout counts (Table 1). [A colour version of this figure is available in the online version.]

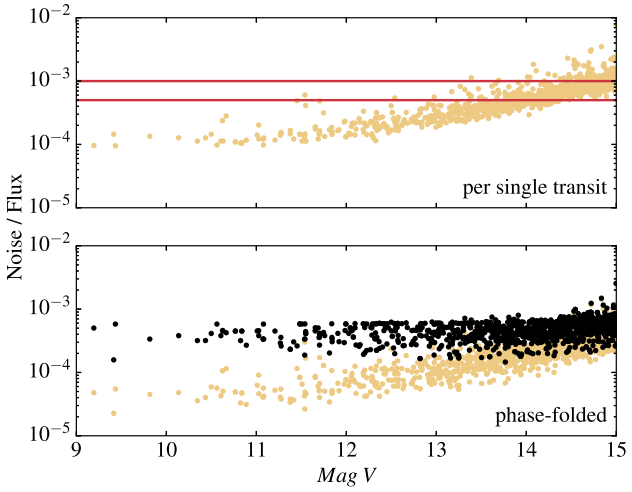

Figure 3. NGTS' white and red noise. Upper panel: white noise (yellow circles) binned over all exposures in one single transit with typical transit time-scales of 1–4 h. Red lines indicate the impact of red noise levels at 0.5 and 1 mmag. Lower panel: white noise (yellow circles) and total noise (i.e. white noise and 1 mmag red noise; black circles) for phase-folded light curves. [A colour version of this figure is available in the online version.]

binned up per single transit, $\sigma_{\rm white}/N_{\rm exp}$ (Section 2.3). Considering a single transit, a red noise level of 1 mmag dominates the total noise for objects $V < 14$, a red noise of 0.5 mmag dominates for objects $V < 13$. Note that for phase-folded light curves of bright targets a white noise model would underestimate the total noise by up to an order of magnitude compared to a model taking 1 mmag of red noise into account (Fig. 3, lower panel).

4.3 Choosing a minimum detection threshold

The choice of the detection threshold (DT), (introduced in Section 2.3), the minimum SNR required to trigger a detection, has a significant impact on the expected yield. A high value of DT leads to non-detection of small planets. On the contrary, lowering DT increases the number of false alarms. These are commonly caused by systematic errors referred to as red noise. The impact of red noise

depends on the time-scale we consider. In this section we estimate the optimal value of DT for NGTS as a function of the estimated number of false alarms.

To estimate the number of false alarms, we begin our argument with a time series of $N_{\rm tot}$ data points in which each point corresponds to an average of the measurements taken over a duration equal to the typical transit duration (1–4 h). The chance that Gaussian noise causes one data point to lie off the mean (looking like a transit) by a standard deviation more than $x\sigma$ is given by

$$p_{\rm outlier}(x) = \frac{1 - {\rm erf}(x/\sqrt{2})}{2}, \quad (7)$$

whereby erf denotes the error function. We assume there are a number of $N > 3$ outliers in the time series. In order to mimic a transit pattern, these N outliers must be distributed in a time-periodic manner. The probability for such a configuration is given by

$$p_{N\,\rm outliers} = \left(p_{\rm outlier}(x)\frac{N_{\rm tot}}{N}\right)^2 p_{\rm outlier}(x)^{N-2} \quad (N > 3). \quad (8)$$

Here, the first term $(p_{\rm outlier}(x)\frac{N_{\rm tot}}{N})$ corresponds to the probability of the first and last outliers. The time periodicity is defined by the total number of outliers and the positions of the first and last outliers in the time series. The outliers in between consequently have determined locations in the time series and the probability of this to happen is expressed by the second term of the equation. For example, if $N = 3$ the first outlier may be located anywhere in the first third of the time series ($N_{\rm tot}/3$ possible locations), while the last outlier may be located anywhere in the last third of the time series. The second outlier has to lie exactly in the middle between the first and last outliers.

For NGTS, the total error on each data point is dominated by red noise (see Section 4.2). Assuming red noise is uncorrelated on multiday time-scales and follows a Gaussian error distribution, it averages out with the number of transit events, and we can approximate $x \approx {\rm DT}/\sqrt{N}$. For example, with a detection threshold ${\rm DT} = 5$ and requiring $N = 3$ outliers, from equation (7) follows that $p_{\rm outlier}(x) \approx 1.9 \times 10^{-3}$. In our estimation for NGTS we have $t_{\rm total} \approx 885$ h. Assuming transit-like time-scales $T \approx 2$ h, this leads to $N_{\rm tot} \approx 440$. Finally, with the example of $N = 3$ outliers, we obtain from equation (8) that $p_{N=3\,\rm outliers} \approx 1.7 \times 10^4$.

The false alarm probability for one time series of data points is then the sum over $p_{N\,\rm outliers}$ for all possible N, which happens to converge quickly with increasing N. As a false alarm can be triggered for each object in the target list, the total number of false alarms, $N_{\rm FA}$, scales with the number of objects in the observed field, $N_{\rm obj}$, and the number of covered NGTS fields, leading to

$$N_{\rm FA} = \sum_{\rm fields}\left[\left(\sum_{N=3}^{N_{\rm max}} p_{\rm N\ outliers}\right) N_{\rm obj}\right]. \quad (9)$$

We compute equation (9) for a range of typical transit-like time-scales from $T = 1$–4 h to evaluate the impact of the DT on the yield versus false alarms. With a total of 12 NGTS fields, each containing $N_{\rm obj} \sim 10^5$, ${\rm DT} = 5$ leads to a number of false alarms on the order of 10^2 among the planet candidates. This suggests a detection threshold of at least ${\rm DT} = 5$ should be used.

4.4 Expected yield and major influencing factors

4.4.1 A multitude of Neptunes and giants

We categorize the yield by object categories: eclipsing binaries (EB), background eclipsing binaries (BEB) and planet types (shown in Table 2). For the application to NGTS we subdivide the category of giant stars from Section 2.2.2 into Saturns (6–10 $R_\oplus$) and Jupiters (10–22 $R_\oplus$). The simulation over the entire survey time of 4 yr is repeated 10 times using the same parameters and input to estimate the mean values and standard deviations of the total number of expected planets and false positives (Fig. 4). We account for uncertainties in the planet and binary priors via error propagation. Statistical errors caused by re-running the TRILEGAL galaxy model for the same field are found to be negligible in comparison.

Table 2. Planet classification based on the radius.

Super-Earths	1.25–2 $R_\oplus$
Small Neptunes	2–4 $R_\oplus$
Large Neptunes	4–6 $R_\oplus$
Saturns	6–10 $R_\oplus$
Jupiters	10–22 $R_\oplus$

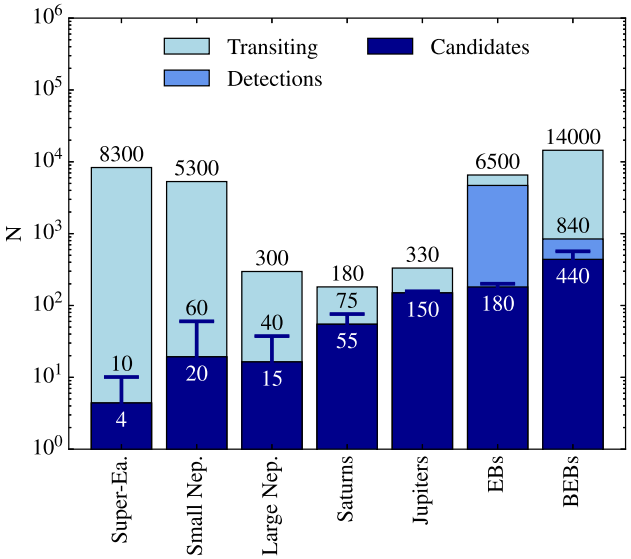

Figure 4. Expected yield for NGTS' planets (see Table 2) and false positives, i.e. eclipsing binaries (EBs) and background eclipsing binaries (BEBs). Light blue: objects undergoing a transit in the line of sight with orbital periods shorter than 20 d. Blue: objects that can be detected with a detection threshold DT = 5 for a red noise level of 1 mmag. Dark blue: planetary candidates that remain from the former group after applying the rule-out criteria for false positives described in Section 2.4. These remaining false positives in the NGTS planet candidate list need follow-up with additional instruments before they can be identified. Blue lines indicate the possible yield if a red noise level of 0.5 mmag can be reached. All values are averaged over 10 simulation runs. Uncertainties are indicated in Table 3. [A colour version of this figure is available in the online version.]

NGTS' combined fields over 4 yr contain tens of thousands of planets and binary systems that transit in the line of sight (Fig. 4). After accounting for visibility, noise and detection criteria, only a small fraction triggers a detectable signal. A detection threshold ${\rm DT} = 5$ for a 1 mmag red noise leads to the detection of ∼35 small and large Neptunes as well as ∼200 of giant planets.

4.4.2 Impact of red noise and detection criteria

The impact of the red noise level and DT is significant for Neptune-sized planets (Table 3). If we decrease the red noise from 1 to 0.5 mmag, the sensitivity for small planets increases by a factor of 3 (see also Fig. 4). Omitting red noise leads to ∼250–300 additional

Table 3. Expected yield for NGTS for different assumptions of the red noise. The table shows all planet classes, false positives before the vetting process described in Section 2.4 (EB before, BEB before) and false positives that remain unidentified after the vetting process (EB after, BEB after). Shown values are the mean and standard deviation averaged over 10 simulation runs, including any uncertainties on the priors.

	2 mmag	1 mmag	0.5 mmag	0 mmag
Super-Earths	1 ± 1	4 ± 3	10 ± 3	28 ± 3
Small Neptunes	5 ± 2	19 ± 5	60 ± 10	229 ± 20
Large Neptunes	4 ± 1	16 ± 4	38 ± 4	69 ± 6
Saturns	28 ± 5	55 ± 8	76 ± 10	86 ± 10
Jupiters	129 ± 9	150 ± 10	158 ± 11	161 ± 10
EB before	4719 ± 45	4688 ± 45	4708 ± 46	4719 ± 45
BEB before	1070 ± 88	843 ± 75	972 ± 83	1070 ± 88
EB after	211 ± 12	181 ± 12	201 ± 12	211 ± 12
BEB after	665 ± 51	439 ± 37	568 ± 47	665 ± 51

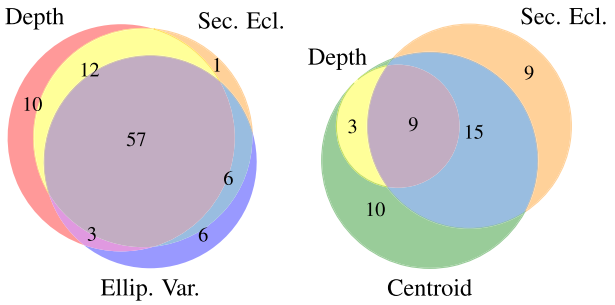

Figure 5. Percentage of false positives identified with different rule-out methods for EBs (left) and BEBs (right). Overlapping regions are designated by colour combination. The relative error on all values is ≤10 per cent. [A colour version of this figure is available in the online version.]

small planets. Increased red noise levels lead to a loss of most small planets.

Maintaining the assumption of a 1 mmag red noise and reducing the detection threshold from DT = 5 to 3 leads to a comparable increase as dividing the red noise by 2, but increases the number of false alarms by several magnitudes as discussed in Section 4.3. In contrast, increasing the detection threshold from DT = 5 to 7 leads to the loss of small planets comparable to doubling the red noise.

For the purpose of the following sections, we assume a detection threshold DT = 5 and, confirm with the design goal of NGTS, a 1 mmag red noise unless otherwise stated.

4.5 False positives and false negatives

NGTS finds a significant number of false positives, consisting of ∼4700 EBs and ∼850 of BEBs (Table 3, Fig. 4). To copy the screening process to identify them, we use the series of criteria described in Section 2.4. We only vet target stars for ellipsoidal variations, assuming light-curve features of diluted background stars are undetectable with NGTS. The centroiding sensitivity for NGTS is set to 1/100 pixel.

The most efficient criteria are depth, secondary eclipses and ellipsoidal variations for EBs, as well as centroiding and secondary eclipses for BEBs (Fig. 5). Overall we estimate 96 per cent of EBs and 48 per cent of BEBs can be identified by the vetting process. These values include vetting for the transit duration, which can be used to identify 5 per cent of false positives, including 2 per cent that cannot be detected with another method. V-shaped transits can

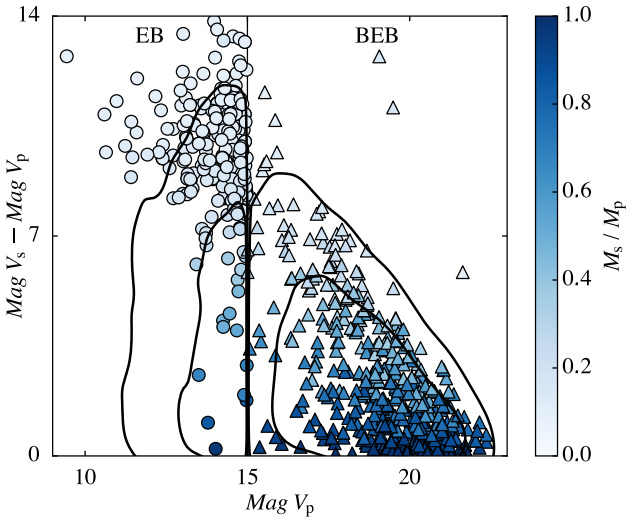

Figure 6. EBs (circles) and BEBs (triangles) that are not identified by the light-curve vetting process and need follow-up observations, shown against the magnitudes of the primary and the secondary. The colour coding displays the binary mass ratio M_s/M_p of the secondary to the primary. Contour lines illustrate the original distribution of false positives before the vetting process. [A colour version of this figure is available in the online version.]

be detected for 54 per cent of EBs, including 5 per cent that cannot be ruled out with other methods. Here, we do not use the V-shape alone to reject an object, as planets can cause V-shaped transits as well.

While individual criteria can already give a hint towards possible false positives, at least two criteria can be met in parallel for about three quarters of all EBs and one quarter of all BEBs. All criteria are effective for a wide range of EBs, but are less applicable for systems with faint secondaries. Faint BEBs are strongly diluted, which decreases the ability to identify them with any method.

The search for false positives can lead to false negatives. Giant planets undergoing grazing eclipses can still trigger detectable NGTS signals. We estimate 10 per cent of all planet detections show a distinguishable V-shape. The transit depth is a reliable measure assuming the stellar properties of the target star are known well. Secondary eclipses, ellipsoidal variations and centroid shifts for planets are not expected to be detectable with NGTS.

4.6 Characteristics of NGTS candidates

4.6.1 Distinguishing remaining false positives from planets

After the candidate vetting process described above, the majority of false positives will be identified but the undetected ones may still outnumber the planet candidates. EBs that remain undetected in the candidate list are expected to consist of binaries with low-mass companions, such as M stars or brown dwarf secondaries (Fig. 6). Small M stars or brown dwarfs are the same size as gas giants (Fortney, Baraffe & Militzer 2011) and hence pollute the sample of giant planets but scarcely affect the population of smaller planets (Fig. 7). These systems cannot be ruled out based on a transit light curve, but need radial velocity follow-up and mass measurements.

Remaining BEBs can show more variety due to different degrees of dilution of their transit signals (Fig. 6). They can mimic planetary radii over a large range but especially pollute the sample of planets in the Neptune-sized regime (Fig. 7). In addition, this is the regime where statistical false alarms will pollute the sample most.

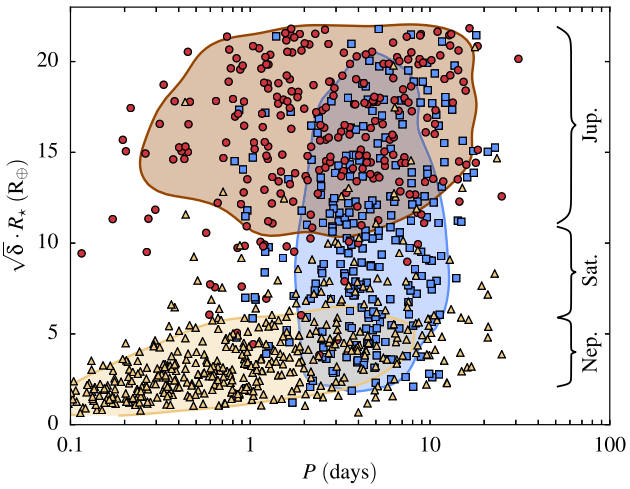

Figure 7. Objects in the expected NGTS candidate list after the photometric vetting process. Shown is the measured apparent radius without correction for dilution or grazing eclipses. Planets (blue squares), EBs (red circles) and BEBs (orange triangles) occupy different regions of the parameters space in apparent transit radius and orbital period. Symbols denote the objects in one simulation run for a red noise of 1 mmag; 1σ contours are calculated over 10 simulation runs. [A colour version of this figure is available in the online version.]

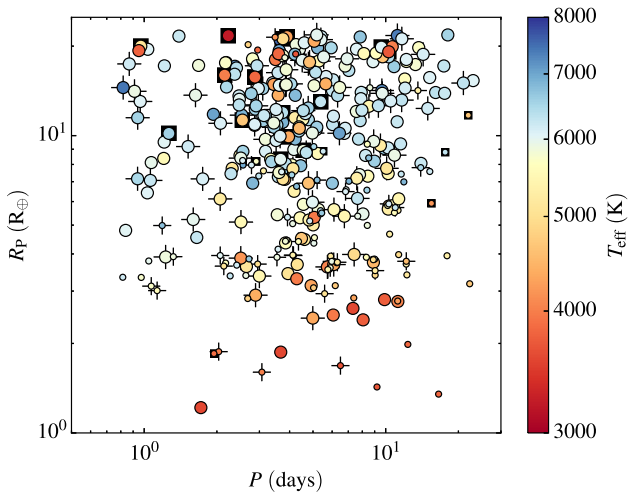

Figure 8. Distribution of detected NGTS planets in period and radius, colour coded in dependence of the host star's effective temperature. Symbol sizes represent detections for red noise levels of 0.5 mmag (small) and 1 mmag (large). An underlying cross (square) marks a planet orbiting the primary (secondary) of a binary system. [A colour version of this figure is available in the online version.]

Fig. 7 further illustrates that it will be difficult to filter out planets below 2-d orbital periods. The planet occurrence rates suggest planets are unlikely to have orbital periods of less than 1–2 d (Fressin et al. 2013; Dressing & Charbonneau 2015). Many binary systems, in contrast, are known to orbit on time-scales of only a few hours, including detached systems (see, e.g. Norton et al. 2011; Soszyński et al. 2015).

4.6.2 *Expected planet properties*

97 ± 1 per cent of all predicted NGTS planets are found at orbital periods shorter than 2 weeks, and 60 ± 3 per cent orbit at less than 5 d (Fig. 8). Decreasing the red noise by half leads to an increase of NGTS' sensitivity to detect small planets with longer orbital periods. K stars are found to be typical hosts for NGTS' smallest planets, such as Small Neptunes and Super-Earths (Figs 8 and 9). Large Neptunes can be detected around G stars. Fig. 9 further illustrates the sensitivity cut-off of NGTS as a function of planet over stellar radius. The paucity of giant planets detected around small stars is a direct consequence of the planet occurrence rates estimated from the results of the *Kepler* mission (Fressin et al. 2013; Dressing & Charbonneau 2015).

We estimate that 57 ± 17 per cent of planets detected with NGTS orbit single stars and 34 ± 12 per cent (8 ± 5 per cent) orbit the primary (secondary) stars of binary systems. Most of the planets detected in binary systems are Jupiters and inflated giants, as only large planets are still detectable given the strong dilution of the transit signal by the light of the binary companion. Dilution decreases the transit depth by 20 per cent or more for 21 per cent of Jupiters, 11 per cent of Saturns and 8 per cent of smaller planets. Circumbinary planets are not detected due to the limited sensitivity of NGTS for long-period transiting planets.

20 ± 5 planets are detected with NGTS around background stars. These consist of systems with a faint target star and a small background star orbited by an inflated giant planet.

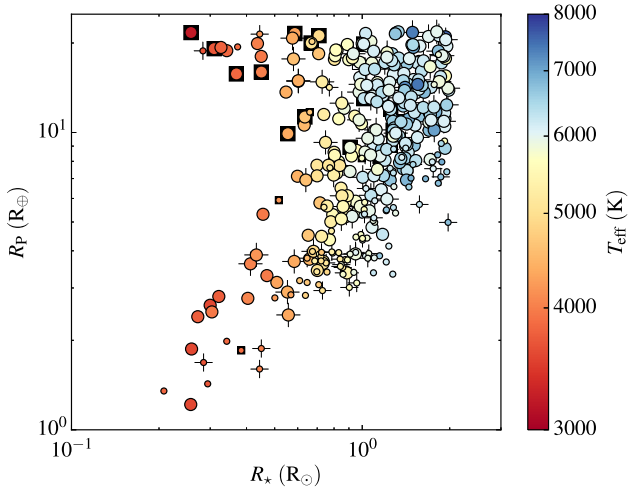

Figure 9. Radius of detected NGTS planets versus radius of their host star, colour coded by the host star's effective temperature. Symbol sizes represent detections for red noise levels of 0.5 mmag (small) and 1 mmag (large). An underlying cross (square) marks a planet orbiting the primary (secondary) of a binary system. [A colour version of this figure is available in the online version.]

4.6.3 *Radial velocity follow-up and characterization of NGTS planets*

We estimate the planetary masses and radial velocity (RV) signals following Section 2.5. The bulk of detected planets lies in magnitude between Corot 7b and GJ 1412b (Fig. 10). Instruments like Coralie (Queloz et al. 2000) are important for vetting false positives, and enable mass measurement of 43 ± 2 per cent of predicted Jupiter-sized planets. HARPS (Mayor et al. 2003) can confirm 99 ± 0 per cent of all predicted Jupiter-sized planets, 92 ± 1 per cent of Saturn-sized planets and 51 ± 2 per cent of Large Neptunes. ESPRESSO (Pepe et al. 2014) will reach the sensitivity to measure RV signals of all NGTS planets. It is worth mentioning that most stars hosting small planets are K dwarfs (Fig. 9) and are brighter in

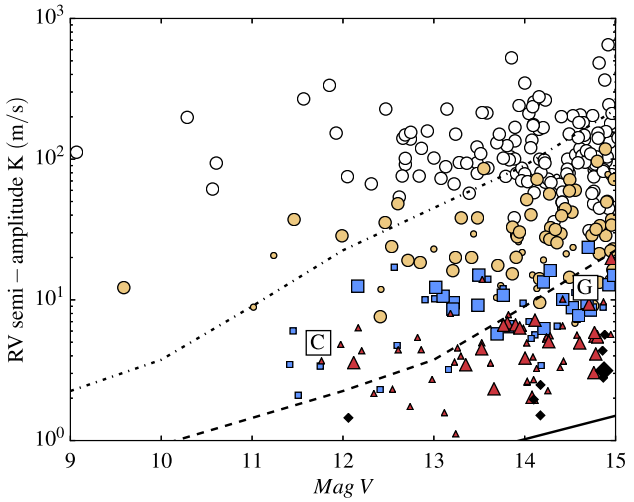

Figure 10. Estimated RV signals of NGTS planets compared with sensitivity of current and future RV instruments: Coralie (dotted line), HARPS (dashed line) and ESPRESSO (solid line). Planet symbols denote Jupiters (white circle), Saturns (yellow circle), Large Neptunes (blue squares), Small Neptunes (red triangles) and Super-Earths (black diamonds). Symbol sizes represent detections for red noise levels of 0.5 mmag (small) and 1 mmag (large). The sensitivity curve for each instrument corresponds to the criterion $2K > 3\sigma_{\rm RV}$. This reflects the minimum signal that can be detected with 3σ for a 1 h exposure considering the instrument's sensitivity, implying at least 10 independent RV measurements per target. The ESPRESSO sensitivity is taken from Pepe et al. (2014), values for HARPS and Coralie are established for 1 h exposures on non-rotating and non-active stars compiled from published results. The planets Corot 7b (C) and GJ 1412b (G), for which masses were measured with HARPS, are shown for comparison. [A colour version of this figure is available in the online version.]

the infrared. Future characterization of these objects may hence be easier in that wavelength.

5 DISCUSSION

5.1 Red noise limitation, detection criteria and planet merit

Red noise is a limiting factor in searching for small planets around bright stars. According to our simulation, at least a 1 mmag precision is essential to detect Neptune-sized planets (see Table 3). In this work we assume major components of the red noise are only short-time (nightly) correlated and average out over a whole season. This assumption is based on early NGTS test data (private communication within the NGTS consortium), and extrapolated from WASP results (Pollacco et al. 2006) and recent studies with EulerCam at La Silla (Lendl et al. 2013). Early results from NGTS data suggest this assumption is realistic (Weathley et al., in preparation).

Red noise is the dominant factor for bright objects (see Fig. 3). Decreasing it by half increases the overall number of small planets and allows to find more of them around bright stars; however, it does not significantly change the number of giant planets, as the detection of giants is limited by the observation window function rather than the noise threshold.

The detection efficiency and false alarm rate depend on the chosen detection criteria, as well as on the final version of detrending and light-curve fitting algorithms. Assuming false alarms can well be ruled out by visual inspection, a 100:1 false alarm ratio is a realistic compromise between a high planet yield and practicality. In the simulation we settled for a DT of 5σ. In systems where a transit has already been found (e.g. by the *TESS* survey) or other planets are known, additional transit signals with lower SNR may be considered.

5.2 Selecting the target list

The number of transiting planets scales with the number of stars in the field of view and hence decreases with distance from the galactic plane. Conversely, there will be more background stars, leading to increased dilution and therefore a decrease of the transit depth as well as an increased number of false positives. Specifically, the yield of EBs, BEBs and giant planets scales with the crowdedness of the field, as their signals are still detectable even in a crowded environment. In contrast, Neptunes and smaller planets show a saturation behaviour since dilution limits the detection of small planets in crowded fields. Sparse fields lead to a higher detection efficiency for small planets but a decreased number of targets. The results suggest these two factors average out for Neptunes and smaller planets. It would therefore be beneficial to stay distant from the galactic plane in order to optimize for a higher yield of Neptunes and Super-Earths while reducing the contamination by false positives.

5.3 Follow-up and characterization of NGTS candidates

We estimate for NGTS that the majority of EBs and BEBs can be identified without necessitating further follow-up measurements. EBs that cannot be identified by the vetting process usually have a low-mass secondary such as an M dwarf or brown dwarf. They can mimic planets with sizes of Jupiter or greater. It will require RV follow-up to determine the mass of the companion and reject a planet hypothesis. While some of the remaining BEBs can have low-mass companions and hence mimic various planetary signals, the majority consists of binary companions with comparable mass. These systems cause very deep transit signals, but are diluted on to a planetary scale. Follow-up of these systems can be achieved with high-precision multicolour photometry to investigate the colour dependence of the transit/eclipse depth.

Here, we focused on EBs and BEBs, but it can be assumed hierarchical eclipsing binaries lead to similar numbers as BEBs, as shown for example in the yield simulations for *TESS* by Sullivan et al. (2015). Additionally considering $\sim 10^2$ false alarms, we expect 97 per cent of all initially detected NGTS transit signals are caused by false positives and false alarms. After the candidate vetting, we expect to remain with 82 per cent of NGTS planet candidates being caused by false positives that need to be identified by follow-up. In comparison, *CoRoT*'s initial detections contained 98 per cent of false positives (Almenara et al. 2009). After the vetting process the follow-up candidates included 88 per cent of false positives. Existing ground-based surveys like WASP and HAT are even more limited in detecting false positives in their photometric data. In the RV and photometric vetting of HAT candidates, a typical frequency of 95 per cent of false positives was found (Latham et al. 2009; Hartman et al. 2011). In their estimations for *TESS*' full-frame images mode, Sullivan et al. (2015) find a contamination by false positives of 97 per cent in the detected signals and 81 per cent after the ad hoc vetting process.

Dilution by background stars can lead to an underestimation of planetary radii around target stars. Especially planets in binary systems are more difficult to detect and may appear smaller, as the light from the binary companion decreases the transit signal further. Hence, gas giants may be misclassified as Neptunes or Super-Earths. Here, we do not treat these diluted planets as false negatives, but

the findings raise awareness of the importance of follow-up measurements to resolve multiple star systems and blended objects.

Precise knowledge of neighbouring objects as well as target star radii is crucial for the vetting and characterization of planets. Current results from the *Gaia*-ESO survey (Gilmore et al. 2012; Randich, Gilmore & Gaia-ESO Consortium 2013) provide astrometric information on more than two million stars, enabling the screening of NGTS targets for nearby background stars. Upcoming data releases will provide precise parallax measurements, enhancing the precision on spectroscopic properties of target stars.

6 CONCLUSION

We developed a comprehensive simulation to investigate the impact of observing strategies, target fields and noise properties on the planet and false-positive yields of transit survey programs. We considered the NGTS facility, and showed that the yield is strongly dependent on the red noise level and DT.

According to our simulation, we show NGTS will fulfil its design purpose by finding ∼240–320 close-in planets and providing a new sample of ∼40–110 characterizable Neptune-sized and smaller planets for the anticipated 4-yr survey.

ACKNOWLEDGEMENTS

We thank Peter Weathley and Simon Walker for introducing us to their previous estimations of the NGTS yield and the NGTS data, as well as Richard West, James McCormac and Daniel Bayliss for useful discussions about NGTS. We are grateful to Ed Gillen for helpful comments on the manuscript. We further thank Courtney Dressing for providing a fine-grid version of the occurrence rate of small planets around small stars. This research has made use of the NASA Exoplanet Archive, which is operated by the California Institute of Technology, under contract with the National Aeronautics and Space Administration under the Exoplanet Exploration Program. Figs 2, 3, 7 and 10 were created using the cubehelix colour scheme provided in Green (2011). MNG is supported by the UK Science and Technology Facilities Council (STFC) as well as the Isaac Newton Studentship.

REFERENCES

Almenara J. M. et al., 2009, A&A, 506, 337
Baglin A., Auvergne M., Barge P., Buey J.-T., Catala C., Michel E., Weiss W., COROT Team, 2002, in Battrick B., Favata F., Roxburgh I. W., Galadi D., eds, ESA Special Publication Vol. 485, Stellar Structure and Habitable Planet Finding. ESA, Noordwijk, p. 17
Bakos G. Á., Lázár J., Papp I., Sári P., Green E. M., 2002, PASP, 114, 974
Batalha N. M., Kepler Team, 2012, Am. Astron. Soc. Meeting Abstr., 220, 306.01
Batalha N. M. et al., 2010, ApJ, 713, L103
Borucki W. J. et al., 2010, Science, 327, 977
Brown T. M., 2003, ApJ, 593, L125
Brown T. M., Latham D. W., 2008, preprint (arXiv:0812.1305)
Brown T. M., Latham D. W., Everett M. E., Esquerdo G. A., 2011, AJ, 142, 112
Bryson S. T. et al., 2013, PASP, 125, 889
Burke C. J. et al., 2015, ApJ, 809, 8
Buser R., Kurucz R., 1978, A&A, 70, 555
Cameron A. C., 2012, Nature, 492, 48
Chazelas B. et al., 2012, Proc. SPIE, 8444, 84440E
Claret A., Hauschildt P. H., Witte S., 2012, A&A, 546, A14
Claret A., Hauschildt P. H., Witte S., 2013, A&A, 552, A16
Deacon N. R. et al., 2016, MNRAS, 455, 4212
Dravins D. L. L. M. E. Y. A. T., 1998, PASP, 110, 610
Dressing C. D., Charbonneau D., 2015, ApJ, 807, 45
Duchêne G., Kraus A., 2013, ARA&A, 51, 269
Duquennoy A., Mayor M., 1991, A&A, 248, 485
Fortney J., Baraffe I., Militzer B., 2011, Exoplanets. Univ. Arizona Press, Tuscan, AZ
Fressin F. et al., 2013, ApJ, 766, 81
Gilliland R. L. et al., 2011, ApJS, 197, 6
Gilmore G. et al., 2012, Messenger, 147, 25
Girardi L., Groenewegen M. A. T., Hatziminaoglou E., da Costa L., 2005, A&A, 436, 895
Green D. A., 2011, Bull. Astron. Soc. India, 39, 289
Grether D., Lineweaver C. H., 2006, ApJ, 640, 1051
Hartman J. D., Bakos G. Á., Torres G., 2011, European Phys. J. Web Conf., 11, 02002
Holman M. J., Wiegert P. A., 1999, AJ, 117, 621
Koch D. G. et al., 2010, ApJ, 713, L79
Kurucz R., 1993, ATLAS9 Stellar Atmosphere Programs and 2 km/s grid, Smithsonian Astrophysical Observatory. CD-ROM No. 13
Latham D. W. et al., 2009, ApJ, 704, 1107
Lendl M., Gillon M., Queloz D., Alonso R., Fumel A., Jehin E., Naef D., 2013, A&A, 552, A2
Lucy L. B., 1967, Z. Astrophys., 65, 89
Mayor M. et al., 2003, Messenger, 114, 20
Mazeh T., 2008, in Goupil M.-J., Zahn J.-P., eds, EAS Publ. Ser. Vol. 29, Observational Evidence for Tidal Interaction in Close Binary Systems. p. 1
Muñoz D. J., Lai D., 2015, Proc. Natl. Acad. Sci. USA, 112, 9264
Norton A. J. et al., 2011, A&A, 528, A90
Pepe F. et al., 2014, Astron. Nachr., 335, 8
Pickles A. J., 1998, PASP, 110, 863
Pollacco D. L. et al., 2006, PASP, 118, 1407
Pont F., Zucker S., Queloz D., 2006, MNRAS, 373, 231
Queloz D. et al., 2000, A&A, 354, 99
Raghavan D. et al., 2010, ApJS, 190, 1
Randich S., Gilmore G., Gaia-ESO Consortium, 2013, Messenger, 154, 47
Slawson R. W. et al., 2011, AJ, 142, 160
Soszyński I. et al., 2015, Acta Astron., 65, 39
Sullivan P. W. et al., 2015, ApJ, 809, 77
Walker S., 2013, PhD thesis, Univ. Warwick
Weiss L. M., Marcy G. W., 2014, ApJ, 783, L6
Wheatley P. J. et al., 2013, European Phys. J. Web Conf., 47, 13002
Winn J. N., 2011, Exoplanets. Univ. Arizona Press, Tuscan, AZ

APPENDIX A: INPUT FILE FOR THE YIELD SIMULATION ON THE EXAMPLE OF NGTS

```
########################################################################
# Target list criteria
########################################################################
15.            #MagV_target = 15.                #upper limit on MagV for target stars
22.5           #magV_limit = 22.5                #upper limit on MagV for background stars
2.             #Rs_limit = 2.                    #upper limit on stellar radius of target stars (in Rsun)
10000.         #Teff_limit = 10000.              #upper limit on effective temperature of target stars (in K)
6.5            #logg_limit = 6.5                 #upper limit on logg of target stars
########################################################################
# CCD resolution
########################################################################
50331648.      #pixel_number = 12.   * 2048.**2  #number of pixels in total; 12 cameras with 2048^2 pixel
4194304.       #pixel_per_ccd = 2048*2048        #number of pixels per CCD; 1 camera with 2048^2 pixel
########################################################################
# Observing strategy
########################################################################
88.8           #FoV = 12. * 7.4.                 #FoV that is observed (in sq.deg.)
122.           #FoV_duration = 365./3.           #time spent per FoV (in days)
12.            #N_FoVs = 4.*3.                   #number of FoVs surveyed in total (4 years * 3 fields per year)
7.             #hours_per_night = 7.             #average number of hours observed per night
########################################################################
# Instrument parameters, noise, aperture
########################################################################
10.            #exposure = 10.                   #exposure (in s)
1.49           #readout = 1.49                   #CCD readout time (in s)
0.2            #D_tel = 0.2                      #telescope aperture (in m)
1.5            #airmass = 1.5                    #average airmass
2400.          #h_tel = 2400.                    #height of the telescope above sea level (in m)
125.           #noise_sky = 125.                 #sky noise (in e- per s per pixel)
0.06           #noise_dark = 0.06                #dark noise (in e- per s per pixel)
10.            #noise_readout = 10.              #readout noise (in e- per pixel)
28.2743338823  #N_apert_pixel = np.pi*(3.**2.)   #number of pixels in aperture (in average)
########################################################################
# Detection criteria
########################################################################
5.             #detection_threshold = 5.         #detection threshold DT (planet detected if signal-to-noise > DT)
0.001          #red_noise = 1./1000.             #as fraction (1/1000 ~ 1 mmag)
3.             #num_transits_threshold = 3.      #minimum number of visible transits required for detection
```

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