The Astronomical Journal, 158:96 (14pp), 2019 August https://doi.org/10.3847/1538-3881/ab2a7a

Empirical Predictions for the Period Distribution of Planets to Be Discovered by the *Transiting Exoplanet Survey Satellite*

Jonathan H. Jiang[1], Xuan Ji[2], Nicolas Cowan[3], Renyu Hu[1], and Zonghong Zhu[2]
[1] Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California, USA
[2] Department of Astronomy, Beijing Normal University, Beijing, People's Republic of China
[3] Department of Earth and Planetary Sciences and Department of Physics, McGill University, Montreal, Canada
Received 2019 March 3; revised 2019 June 11; accepted 2019 June 16; published 2019 August 2

Abstract

Launched in 2018 April, NASA's *Transiting Exoplanet Survey Satellite* (*TESS*) has been performing a wide-field survey for exoplanets orbiting bright stars with a goal of producing a rich database for follow-on studies. Here we present estimates of the detected exoplanet orbital periods in the 2 minute cadence mode during the *TESS* mission. For a two-transit detection criterion, the expected mean value of the most frequently detected orbital period is 5.01 days, with the most frequently detected range of 2.12–11.82 days in the region with observation of 27 days. Near the poles where the observational duration is 351 days, the expected mean orbital period is 10.93 days, with the most frequently detected range being from 3.35 to 35.65 days. For one transit, the most frequently detected orbital period is 8.17 days in the region with observation of 27 days and 11.25 days in the region near the poles. For the entire *TESS* mission containing several sectors, we estimate that the mean value of orbital period is 8.47 days for two-transit detection criterion and 10.09 days for one-transit detection criterion. If *TESS* yields a planet population substantially different from what is predicted here, the underlying planet occurrence rates are likely different between the stellar sample probed by *TESS* and that by *Kepler*.

Key words: methods: data analysis – planetary systems – surveys – telescopes

1. Introduction

The *Transiting Exoplanet Survey Satellite* (*TESS*), a NASA Explorer-class exoplanet finder mission, has been in orbit since 2018 April 18. *TESS* is monitoring 26 observational sectors, each covering about 2300 square degrees of the sky (Ricker et al. 2015; Barclay et al. 2018). For most of *TESS*'s field of view (FOV), the spacecraft will observe the selected stars in the *TESS* Input Catalog (TIC) to measure the brightness at a 2 minute cadence for 27 days. It will also obtain 30 minute cadence observations of all objects in the *TESS* fields of view, but we do not discuss the yield of this kind of observation in this paper. Over the mission design lifetime of 2 yr, *TESS* will continuously survey ∼85% of the sky for 27 days. Thus, the majority of *TESS* sensitivity will be devoted to short-period exoplanets orbiting closer to their parent stars. However, the survey methodology has been designed such that certain parts of the sky will be surveyed across multiple runs with overlapping FOVs, enabling longer durations of observations, especially at the ecliptic poles (Ricker et al. 2015). Therefore, *TESS* will have additional observational durations of 54, 81, 108, 189, and up to 351 days in different regions of the celestial sphere to allow for sensitivity to exoplanets with different orbital periods. The region surrounding the ecliptic pole will be especially visible in multiple *TESS* FOVs with observational baselines of close to a year, enabling the search for planets in Earth-like orbits.

To confirm the exoplanets detected by *TESS* and to determine their orbital period, at least two transits are required. Simulations by Sullivan et al. (2015), Barclay et al. (2018), Huang et al. (2018), and Ballard (2019) have predicted the exoplanet properties yielded by *TESS*'s two or more transit events. Sullivan et al. (2015) focused on the model for the relevant stellar and planetary populations in order to forecast the properties of the brightest transiting planet systems in the sky. These studies used a synthetic stellar population rather than a real catalog. Barclay et al. (2018) updated the simulation of *TESS*'s yield using the TIC Candidate Target List (CTL); their detection model is similar to Sullivan et al. (2015). Huang et al. (2018) improved the simulation using an empirically based simulation of multiplanet systems and updating the photometric noise model and so on. All of their simulations used the planet occurrence rates from Fressin et al. (2013) for stars with effective temperature ($T_{\rm eff}$) $\geqslant$ 4000 K, and from Dressing & Charbonneau (2015) for stars with $T_{\rm eff} <$ 4000 K. Fressin et al. (2013) determined the rate of occurrence of planets assuming initial distributions of planets as a function of planet size and orbital period provided by Howard et al. (2012). And the true planet population was obtained by simulating it in detail with no prior assumption and then adding their simulated false positives to match the list of Batalha et al. (2013), which only contains the planets detected by *Kepler* during the first 16 months after accounting for the detectability of both planets and false positives. Now we have the full sample of 4 yr of observation, which has less bias on orbital period. Dressing & Charbonneau (2015) used the *Kepler* data set, but they still have some dead zones at orbital periods of 60–100 days and radii of 0.5–1 Earth radius ($R_{\oplus}$). The occurrence rates from both Fressin et al. (2013) and Dressing & Charbonneau (2015) are based on *Kepler* data, but are limited in orbital period up to 85 and 200 days, respectively, beyond which the statistics are still too poor to provide results.

The possible yield of *TESS*'s single transiting planets has also been studied by Villanueva et al. (2019), in which they predict up to 241 single-transit events during *TESS*'s primary mission. Planet occurrence rates from Villanueva et al. (2019) were also based on the results of Fressin et al. (2013) and Dressing & Charbonneau (2015). These planet occurrence rates are only complete to periods of ∼100 days, but they extrapolate these rates to periods of >1000 days to explore the probability

of finding planets at longer periods. They assume the occurrence of the longer orbital period is equal to that of the longest complete orbital period. Note that the sample of orbital period within 503 days has a completeness >90% according to our following analysis. We can dependably predict the occurrence of the single transit for which the orbital period is within 503 days to check the result.

To summarize, previous *TESS* yield predictions are mostly based on simulated data informed by *Kepler* yields. In order to obtain the actual occurrence rate of exoplanets, it is necessary to correct the data for detection bias and incompleteness, and it always leads to the result being limited in a small range due to the incompleteness of data. As *TESS* uses the same detection method as *Kepler*, both should be subject to the same type of detection bias. Since the *Kepler* mission provided biased data for us, we can perform our analysis directly on these data. All we need to do is to compare the difference in target stars and the properties of the detector between the two missions. Then the distribution of *TESS*'s transiting planets can be derived from those of *Kepler*, through which we can prevent the uncertainties of estimating the occurrence of a planet with a given star and eliminate the extra steps, and the result will not be limited in such a small range like before.

In this study, we compute the probabilities of exoplanet orbital period distributions detected in each of *TESS*'s observational duration regions assuming a two-transit detection criterion and a one-transit detection criterion. The probability density function (PDF) of orbital period is calculated based on the observed data during the *Kepler* mission. We first review the detection bias of different exoplanet detection methods (Section 2), followed by a description of the methodology for computing the orbital period distributions of detected exoplanets (Section 3). The results, uncertainty analysis, and comparisons with previous studies are presented in Section 3.3, followed by a summary and discussion in Section 4.

2. Detection Bias

The orbital period of exoplanets is critical for understanding the formation and evolution of the star system that hosts the exoplanet (Wang & Peng 2015). It also provides important knowledge about the orbital radius, which plays a key role in determining whether or not the planet can support life (Kopparapu et al. 2013).

Using data from the NASA Exoplanet Archive (https://exoplanetarchive.ipac.caltech.edu/; "Archive" hereafter), we plot in Figure 1 the normalized histogram of exoplanets' orbital periods (P) detected by different methods: transit, radial velocity (RV), and others. The exoplanets detected using the transit method are almost all from the *Kepler* mission. As shown in Figure 1, the distribution of exoplanet orbital periods varies from the detection methods. The most detectable orbital periods using transit photometry range from 1 to 100 days, while for the RV method the most detectable are exoplanets with orbital periods of 200–2000 days. A combination of all other methods yields exoplanets with a wide range of orbital periods from a few days to a few tens of years.

The method of detection of both transit- and RV-detected exoplanets depends on the exoplanets' orbital period. We plotted the Hertzsprung–Russell (HR) diagram for their host stars, as shown in Figure 2, where the x axis shows the effective temperature of the host star in Kelvin $T_{\rm eff}$(K) provided in the Archive. The y axis shows the log base 10 luminosity

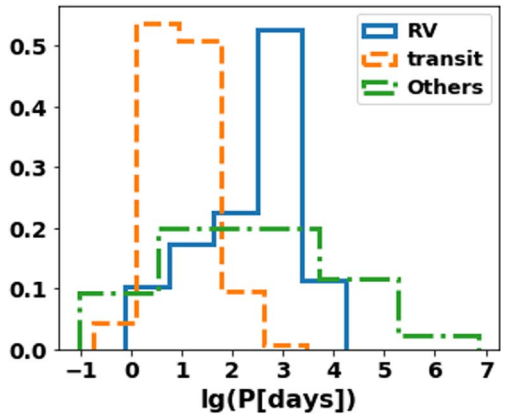

Figure 1. Normalized histogram of exoplanets' orbital periods (P) detected by different methods: transit, radial velocity (RV), and others. The x axis is base 10 logarithmic T (days), and the y axis is the probability distribution of the orbital periods normalized by the total number of samples.

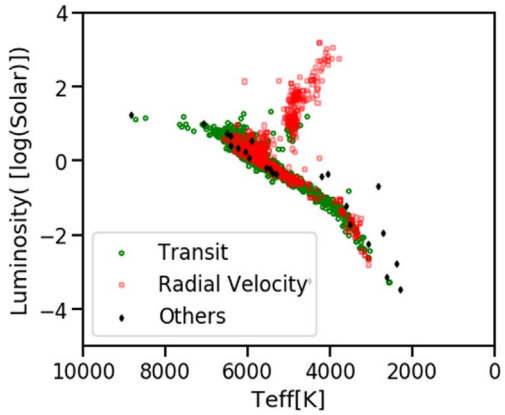

Figure 2. HR diagram of exoplanet host stars. The green circles are stars with planets detected by the transit method, the red open squares are stars with planets detected by the RV method, and the black diamonds are stars with planets detected by other methods. The gap in the middle of the main sequence for the red squares is likely due to the selection effect of radial velocity.

(L) divided by solar luminosity ($L_\odot$), which is derived using the Stefan–Boltzmann relation $L/L_\odot = (R_s/R_\odot)^2(T_{\rm eff}/T_\odot)^4$, where R_s is the radius of the host star, also available in the Archive, and $R_\odot$ is the solar radius. It can be seen that the transit method detects mainly exoplanets orbiting the main-sequence stars, while almost all exoplanets orbiting red giants are detected by RV (e.g., Jiang & Zhu 2018). This means we should reject the hypothesis that the two sets of data, transit and RV, come from the same distributions.

The transit method is so far the most effective and sensitive method for detecting extrasolar planets, particularly from space telescopes. The limitation and biases of this method have been studied in detail by Kipping & Standford (2016). To the first order, assuming the stellar disk is of uniform brightness and neglecting any flux from the planet, the ratio of the observed relative change in flux can be approximated as $\Delta F/F = R_p^2/R_s^2$, where R_p is the radius of the exoplanet. The transit may be too faint to distinguish if R_p/R_s is too small. A planet's transit lasts only a tiny fraction of its total orbital period, so even when we observe a star with a transiting planet, it is unlikely to repeat the observation if the telescope observes the system for less than the orbital period of the planet. Even if the transit repeats, we are less likely to detect the outer exoplanets with longer orbital period

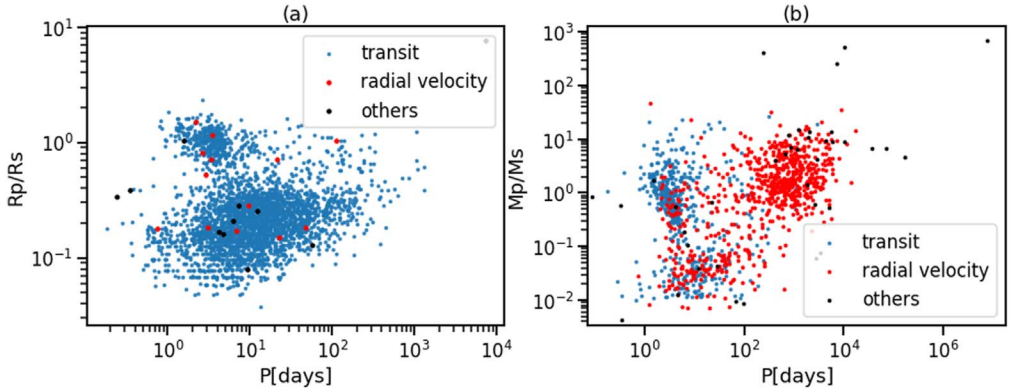

Figure 3. Distributions of confirmed exoplanets for all types of host stars in the Archive (a) in the period–radius plane and (b) in the mass–radius plane.

compared to the inner one with the same radius due to the lower signal-to-noise ratio, which is related to the time of transits.

Figure 3(a) shows R_p/R_s versus orbital period using data from the Archive. For comparison, Figure 3(b) shows M_p/M_s versus orbital period, where M_p and M_s are the radii of the exoplanet and host star, respectively. These distribution results in Figures 3(a) and (b) are similar to the simulated distribution by Sullivan et al. (2015), except that *Kepler*'s transiting exoplanets seem to be divided into two groups above and below the $R_p/R_s \sim 5$ in Figure 3(a), which is a topic of ongoing investigation (e.g., Fulton et al. 2017). Also noted is that the mass data for the transiting exoplanets are limited to those with a short period (<5 days). Only a small proportion of RV exoplanets have such a short period, which is in contradiction with the fact that the RV method is more sensitive to the closer exoplanet if the M_p/M_s ratios are equal. Due to the strong selection bias shown in Figure 3, it is difficult to obtain the accurate occurrence of the planet as a function of the planetary properties.

As discussed above, the detection bias leads to different statistics, which means it cannot be ignored or underestimated. So, it is crucial to correct the bias if we want a general conclusion. As a result, such conclusions are always within a given range because of the incompleteness of the data. If the simulation to estimate the yields of exoplanets in the future is based on those general conclusions, the result will be restricted to the range. Here we exploit the data set of transiting exoplanets to predict the distribution of planetary parameters of new transiting exoplanets in a straightforward manner.

The transiting exoplanets we used in this study are those detected by the *Kepler* telescope during 4 yr (or 1459 days) of its mission from 2009 May 13 00:01:07Z to 2013 May 11 12:16:22Z. The *Kepler* mission's specified exoplanet detection criterion is a minimum of three observed transit events (Twicken et al. 2016). Thus, most of *Kepler*'s confirmed exoplanets are those with an orbital period less than 486 days. There are some exceptions; for example, the maximum orbital period among the confirmed exoplanets in the Archive, listed under the *Kepler* Object of Interest (KOI) Catalog Q1–Q17 Data Release 25 (DR 25; Twicken et al. 2016), is from Kepler-167e, which has an orbital period of 1071 days. There is no way for *Kepler* to observe its three transits, and thus Kepler-167e must be confirmed by other means. Nevertheless, the *Kepler* data are used to compute the PDF for this study because they contain realistic observational information of exoplanets discovered using the transit method. In particular, they also contain the transit probability bias related to the stellar radius and semimajor axis, which will be discussed in the next section.

3. Methodology

3.1. Sample Selection

The sample we used to obtain the PDF of the orbital period of exoplanets detected by *Kepler* is the confirmed exoplanets listed in KOI Catalog Q1–Q17 DR 25 (Twicken et al. 2016). The *Kepler* photometer acquired data at 29.4 minute intervals, known as "long cadences." Science acquisition of Q1 data began at 2009 May 13 00:01:07Z, and acquisition of Q17 data concluded at 2013 May 11 12:16:22Z. This time period is 1459 days and contains 71,427 long-cadence intervals. Science data acquisition was interrupted periodically: monthly for data downlink, quarterly for maneuvering to a new roll orientation, and once every three days for reaction wheel desaturation (Jenkins et al. 2010). Since those interruptions are periodic and very short, we ignore them. In addition to the normal interruption, data acquisition was suspended for 11.3 days (555 long-cadence samples) in Q16, and 1145 long-cadence intervals were excluded from searches for transiting planets because of data anomalies. Only if the transit had occurred three times and one of them exactly occurred in the interrupted period—which means the orbital period is from 364.75 to 486.3 days—can we miss this exoplanet. Since exoplanets with such long orbital periods cannot be observed by *TESS* for the required two transits, we do not discuss this effect. We take those interruptions into consideration when calculating the signal-to-noise ratio (S/N).

TESS is searching for small transiting planets, which leads primarily to the selection of bright, cool dwarfs (Stassun et al. 2018). The CTL includes both dwarfs and subgiants. In order to match the CTL, we filter the confirmed exoplanets in KOI by excluding those with giant host stars. For the exoplanets that have stellar information of Yerkes spectral classification, we only choose VI (subgiants) and V (dwarfs) stars in our sample, while for those lacking spectral information, we limit the stellar surface gravity ($\log g$) from 3.8 to 5.0 (Figure 4; Allen 1976). After filters, the size of the final sample is 2141.

To match the CTL, we separate the sample into M dwarfs (<4500 K) and AFGK-type stars (>4500 K) according to the stellar effective temperature and adjust the proportion in the later analysis. The size of the subsample of M dwarfs is 218, and that of AFGK-type stars is 2023. The fitting result is shown in Figure 5.

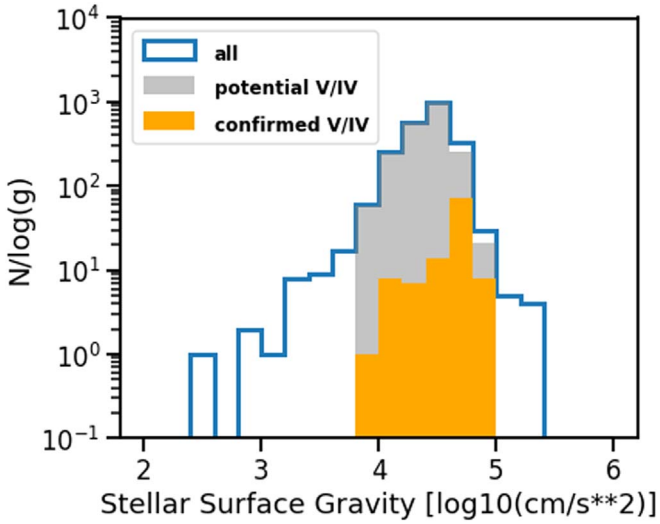

Figure 4. Histogram of log g for host stars of all confirmed exoplanets in DR 25 (blue), host stars that are confirmed dwarfs and subgiants (orange), and host stars whose surface gravity (log g) is from 3.8 to 5.0 (gray).

3.2. PDF Expression

The purpose of this study is to obtain the orbital period distribution of exoplanets discovered by *TESS* from *Kepler* data. A comparison between two theoretical PDFs of exoplanet orbital period discovered by *TESS* and *Kepler* is made; then the ratio can be calculated based on the different input catalogs and mission parameters. The ratio is a function of orbital period. The PDF of *Kepler* is obtained by fitting the data (Figure 5), and then we obtain the PDF of *TESS* by multiplying the theoretical ratio with the fitting PDF of *Kepler*.

3.2.1. Probability Expression

Orbital period is a continuous variable, which means the theoretical probability of a specific orbital period is zero, and only the integral of PDF makes sense. The probabilities mentioned in the following analysis are all obtained by taking the limit that denotes the values of PDF at a certain point. Based on Bayes's theorem, the probability that the orbital period of a detected exoplanet is P days is

$$\mathrm{Prob}(P|TESS) = \mathrm{Prob}(TESS|P) \cdot \mathrm{Prob}(P)/\mathrm{Prob}(TESS), \quad (1)$$

where *TESS* denotes the event that the exoplanet is detected by *TESS*, and $\mathrm{Prob}(TESS|P)$ can be expanded as

$$\begin{aligned}\mathrm{Prob}(TESS|P) = {} & \mathrm{Prob}(tr|P) \cdot \mathrm{Prob}(Ntrs_T|P, tr) \\ & \cdot \mathrm{Prob}(\mathrm{S/N}_T > \mathrm{S/NT}_{\min}|P, tr, Ntrs_T),\end{aligned} \quad (2)$$

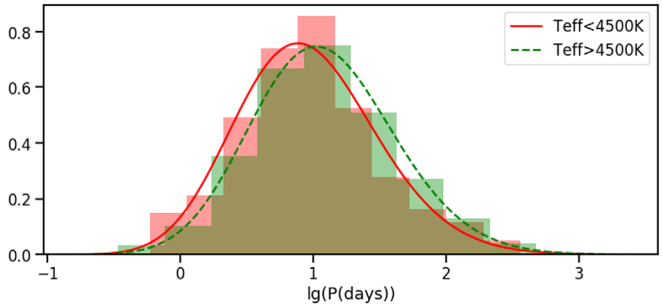

Figure 5. Normalized histogram of orbital period for each subsample. The solid and dashed lines are the fitting probability density functions for subsample 1 and subsample 2, respectively.

where $\mathrm{Prob}(tr|P)$ is the geometric probability of detecting a transit around a star for a fixed period, $\mathrm{Prob}(Ntrs_T|P, tr)$ is the probability of observing the transit(s) more than N times during the finite observing baseline of observations for *TESS* for a fixed period, given that the transit is detected, and $\mathrm{Prob}(\mathrm{S/N}_T > \mathrm{S/NT}_{\min}|P, tr, Ntrs_T)$ is the probability that the S/N of the exoplanet is higher than the threshold given that it transits at least N times over the course of the observations.

Then Equation (1) is expanded as follows:

$$\mathrm{Prob}(P|TESS) = \frac{\mathrm{Prob}(tr|P) \cdot \mathrm{Prob}(Ntrs_T|P, tr) \cdot \mathrm{Prob}(\mathrm{S/N}_T > \mathrm{S/NT}_{\min}|P, tr, Ntrs_T) \cdot \mathrm{Prob}(P)}{\mathrm{Prob}(TESS)}. \quad (3)$$

Likewise, repeat the above analysis but for *Kepler*:

$$\mathrm{Prob}(P|Kepler) = \frac{\mathrm{Prob}(tr|P) \cdot \mathrm{Prob}(Ntrs_K|P, tr) \cdot \mathrm{Prob}(\mathrm{S/N}_K > \mathrm{S/NK}_{\min}|P, tr, Ntrs_K) \cdot \mathrm{Prob}(P)}{\mathrm{Prob}(Kepler)}, \quad (4)$$

where $Ntrs_K$ is the minimum required times of transits for the *Kepler* mission, $\mathrm{S/N}_K$ is the S/N detected by *Kepler*, and $\mathrm{S/NK}_{\min}$ is the threshold of the *Kepler* pipeline.

3.2.2. Calculation of Each Term

We assume that the planetary radius is much less than the stellar radius ($R_p \ll R_*$) and planetary mass is much smaller than stellar mass ($M_p \ll M_*$). Using Kepler's third law, the geometric probability is

$$\begin{aligned}\mathrm{Prob}(tr|P) &= \int \frac{R_*}{a} f_{R_*,a|P}(R_*, a) dR_* da \\ &= \int \left(\frac{4\pi^2}{G}\right)^{\frac{1}{3}} R_* M_*^{-\frac{1}{3}} P^{-\frac{2}{3}} f_{R_*,M_*|P}(R_*, M_*) dR_* dM_*\end{aligned} \quad (5)$$

where a is the semimajor axis. Here, $f_{R_*,a|P}$ is the joint PDF of the stellar radius and semimajor axis and $f_{R_*,M_*|P}$ is the joint PDF of the stellar radius and stellar mass. This probability is only related to the stellar properties except for the orbital period. We assume that the stellar parameters obey the same distribution in each subsample, so the intrinsic stellar parameters of *TESS*

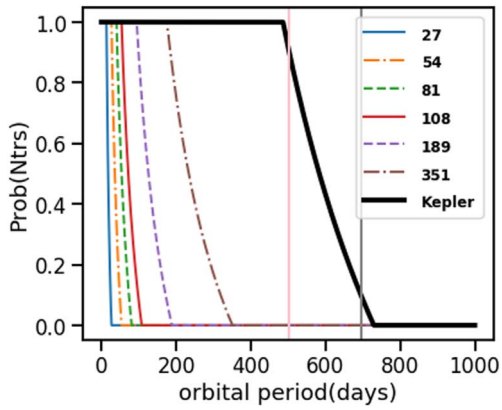

Figure 6. Probability of observing two transits (*TESS*) or three transits (*Kepler*'s detection criterion) during the finite observational duration of 27, 54, 81, 108, 189, or 351 days for *TESS* and 1459 days for *Kepler*. The gray vertical line is 694.76 days, beyond which the probability that it can be detected by *Kepler* is less than 10%, and the pink vertical line is 503.10 days, within which the probability is higher than 90%.

M dwarf targets are statistically the same as for *Kepler* M dwarf targets, even if they differ in number and distance. This integral only contains intrinsic stellar parameters, so the value does not change with the mission. Equation (5) is only the expected value; its uncertainty will be discussed in Section 4.4.

Identifying that a transit signal is due to an exoplanet requires a minimum of N times of observed transit events. The probability an exoplanet with a given orbital period (P) can be detected during a given observational duration (t days) can be estimated as follows:

$$\mathrm{Prob}(Ntrs_T|P,\ tr) = \begin{cases} 0, & t \leqslant (N-1)P \\ \dfrac{t-(N-1)\cdot P}{P}, & (N-1)P < t < NP. \\ 1, & t \geqslant N\cdot P \end{cases} \quad (6)$$

Figure 6 shows the probability that an exoplanet with a given orbital period can be discovered. *TESS* has different observation durations in different sky areas due to the multiple sectors, so the probability varies from one sky area to another.

Here we discuss the impact of the condition on the number of transits observed in the term $\mathrm{Prob}(\mathrm{S/N}_T > \mathrm{S/NT_{min}}|P,tr, Ntrs_T)$. Under the condition where the orbital period (P) is fixed, three scenarios are introduced in Equation (6). In the first scenario, the value of $\mathrm{Prob}(Ntrs_T|P,\ tr)$ is 0, so the value of this term does not matter. In the second scenario, some exoplanets for which the number of times of the transit event is $(N-1)$ will be excluded. Since every exoplanet will be excluded with equal probability, regardless of whether its S/N is higher than the threshold or not, the ratio $(\mathrm{Prob}(\mathrm{S/N}_T > \mathrm{S/NT_{min}}))$ will not change. In the third scenario, the probability that the number of transits observed exceeds N is 100%, which means there is no constraint on this term. In conclusion, the condition on the number of transits observed in $\mathrm{Prob}\ (\mathrm{S/N}_T > \mathrm{S/NT_{min}}\ |P,\ tr,\ Ntrs_T)$ has no influence on the final value of the right side of Equation (2), so we can remove it. Then the third term in Equation (3) is $\mathrm{Prob}(\mathrm{S/N}_T > \mathrm{S/NT_{min}}\ |P,\ tr)$. Likewise, the third term in Equation (4) can be rewritten as $\mathrm{Prob}(\mathrm{S/N}_K > \mathrm{S/NK_{min}}|P,\ tr)$. The detailed calculation of this term will be discussed in Section 3.3.

3.2.3. Comparison between Two Missions

In each subsample, $\mathrm{Prob}(P)$ does not change with the mission. Because $\mathrm{Prob}(tr|P)$ is a function of stellar properties, the ratio between two of them is not 1 but a distribution. Here we only consider the maximal probabilistic estimations, which means we substitute the random variable with their expected value. According to Equations (3) and (4), the PDF of orbital period of each subsample for the *TESS* mission can be expressed in terms of that of *Kepler*:

$$\mathrm{Prob}_i(P|TESS) = c_i \mathrm{Prob}_i(P|Kepler) \times \frac{\mathrm{Prob}_i(Ntrs_T|P,\ tr)\cdot \mathrm{Prob}_i(\mathrm{S/N}_T > \mathrm{S/NT_{min}}|P,\ tr)}{\mathrm{Prob}_i(Ntrs_K|P,\ tr)\cdot \mathrm{Prob}_i(\mathrm{S/N}_K > \mathrm{S/NK_{min}}|P,\ tr)} \quad i = 1,\ 2 \quad (7)$$

where $i = 1$ denotes the subsample of M dwarfs and $i = 2$ denotes the subsample of AFGK-type stars, and c_i is the normalization coefficient. Since $\mathrm{Prob}(Kepler)$ and $\mathrm{Prob}(TESS)$ in Equations (3) and (4) are constants, they can be incorporated into the normalization constants.

For *Kepler* DR 25, the required $Ntrs_K$ is 3, and we perform analysis for the two-transit criterion and one-transit criterion for the *TESS* mission.

3.3. S/N Model

The transit S/N is $\Delta F_*/\sigma_*$, where ΔF_* is the change in stellar flux during transit and σ_* is the uncertainty in stellar flux during transit.

We exclude grazing transits, for which the planet never completely obscures the star. The signal for a full transit is

$$\frac{\Delta F_*}{F_*} = \left(\frac{R_p}{R_*}\right)^2, \quad (8)$$

where we have ignored limb darkening. The Poisson (shot) noise, σ_*, in the many-photon limit is simply $\sqrt{N}$, where N is the number of photons detected in transit. The precision is the noise divided by the stellar flux, $\sigma_*/F_* = 1/\sqrt{N}$. The S/N is therefore

$$\frac{\Delta F_*}{\sigma_*} = \frac{\Delta F_*/F_*}{\sigma_*/F_*} = \left(\frac{R_p}{R_*}\right)^2\sqrt{N}. \quad (9)$$

Technically, transit photometry is a differential measurement, but we optimistically assume that the uncertainty outside the transit baseline is negligible, which is a reasonable assumption for *Kepler* or *TESS*.

Following Cowan et al. (2015), the number of in-transit photons detected over the course of a mission is

$$N = N_{trs} f_t PA\left(\frac{R_*}{r}\right)^2 \int_{\lambda_1}^{\lambda_2} \tau\pi B(\lambda,\ T_*)\left(\frac{\lambda}{hc}\right)d\lambda, \quad (10)$$

where $Ntrs$ is the time of transit events during the total observation baseline, f_t is the fraction of time the planet spends in transit, A is the collecting area of the telescope, τ is the system throughput (photon conversion efficiency = electrons out per photon in), P is the orbital period, r is the distance to the star, $\lambda_1 - \lambda_2$ is the bandpass, and $B(\lambda,\ T_*)$ is the Planck function of the star.

The fraction of time the planet spends in transit depends on both planetary and stellar parameters. The length of the chord transited by the planet is $l = 2\sqrt{R_*^2 - b^2}$, where b is the

Table 1
$k[h(\mathbb{M}_T, T_*)/h(\mathbb{M}_K, T_*)]$

	T_*	k (27 days)	k (54 days)	k (81 days)	k (108 days)	k (189 days)	k (351 days)
<4500 K	3974.41	0.0327	0.0463	0.0567	0.0655	0.0866	0.1180
>4500 K	5653.53	0.0621	0.0878	0.1075	0.1242	0.1643	0.2239

Note. k is the ratio $h(\mathbb{M}_T, T_*)/h(\mathbb{M}_K, T_*)$. It is a constant for the certain subsample and certain observing baseline for *TESS*, of which the value is shown in the table, respectively.

planet's impact parameter. The fraction of time spent in transit is therefore the ratio of the chord to the circumference of the orbit:

$$f_t = \frac{2R_*\sqrt{1-b^2}}{2\pi a} = \frac{R_*\sqrt{1-b^2}}{\pi a}, \tag{11}$$

where a is the planet's semimajor axis, and we have adopted a circular orbit and constant transverse velocity for simplicity. This introduces a nuisance variable, b, but we can marginalize over it:

$$\langle f \rangle = \frac{R_*}{\pi a}\int_0^1 \sqrt{1-b^2}\,db = \frac{R_*}{4a}. \tag{12}$$

The S/N is therefore

$$\mathrm{S/N} = \frac{R_p^2}{\sqrt{a}}\sqrt{\frac{N_{trs}AP}{4R_*r^2}\int_{\lambda_1}^{\lambda_2}\tau\pi B(\lambda, T_*)\left(\frac{\lambda}{hc}\right)d\lambda}, \tag{13}$$

where we have isolated all of the star- and mission-dependent parameters under the radical.

In order to express the S/N in terms of the orbital period rather than the semimajor axis, we make use of Kepler's third law. As the value of S/N is used to calculate the probability, the expected value of *Ntrs* will be applied but not the single value of every exoplanet. In other words, we are only concerned with the statistic. Then we can plug $\langle Ntrs\rangle = t_m/P$ into the above equation, the error of which will be discussed in Section 4.4.1:

$$\mathrm{S/N} = R_p^2 P^{-\frac{1}{3}}\left(\frac{4\pi^2}{GM_*}\right)^{\frac{1}{6}}\sqrt{\frac{At_m}{4R_*r^2}\int_{\lambda_1}^{\lambda_2}\tau\pi B(\lambda, T_*)\left(\frac{\lambda}{hc}\right)d\lambda}. \tag{14}$$

We can think of the S/N as a separable function, but the stellar effective temperature in the Planck function integral makes it impossible to separate the stellar parameters from the mission parameters. Given that we determine the subsample according to effective temperature, we assume that the effective temperature obeys the same distribution for bins with different orbital periods. Thus we can use the mean value of it ($\langle T_*\rangle$) to calculate the $P(\mathrm{S/N} > \mathrm{S/N_{min}}|P, tr, Ntrs)$ and then think of the S/N as a separable function:

$$\begin{aligned}
&\mathrm{S/N} = f(\mathbb{P})g(\mathbb{S})h(\mathbb{M}, \boldsymbol{T}_*)\\
&f(\mathbb{P}) = R_p^2 P^{-\frac{1}{3}}\\
&g(\mathbb{S}) = \left(\frac{4\pi^2}{GM_*}\right)^{\frac{1}{6}}\sqrt{\frac{1}{4R_*}}\\
&h(\mathbb{M}, \boldsymbol{T}_*) = \sqrt{\frac{At_m}{r^2}\int_{\lambda_1}^{\lambda_2}\tau\pi B(\lambda, T_*)\left(\frac{\lambda}{hc}\right)d\lambda},
\end{aligned} \tag{15}$$

where $\mathbb{P} \equiv \{R_p, P\}$ are the planetary parameters, $\mathbb{S} \equiv \{M_*, R_*\}$ are the stellar parameters, and $\mathbb{M} \equiv \{A, \tau, t_m, \lambda_1, \lambda_2, r\}$ are the mission parameters. Since the distance to the star (r) is not an intrinsic property but depends on the input catalog of the mission, we take it as a mission parameter. The *TESS* and *Kepler* mission parameters are $\mathbb{M}_T$ and $\mathbb{M}_K$. Thus for a given transit event, if it is detected by *Kepler* and *TESS*, the S/N should be

$$\mathrm{S/N}_K = f(\mathbb{P})g(\mathbb{S})h(\mathbb{M}_K, T_*);\ \mathrm{S/N}_T = f(\mathbb{P})g(\mathbb{S})h(\mathbb{M}_T, T_*). \tag{16}$$

The $\mathrm{S/N}_T$ can then be expressed in terms of $\mathrm{S/N}_k$:

$$\mathrm{S/N}_T = \mathrm{S/N}_K\frac{h(\mathbb{M}_T, T_*)}{h(\mathbb{M}_K, T_*)} = k(\mathbb{M}_T, \mathbb{M}_K, T_*)\cdot\mathrm{S/N}_K, \tag{17}$$

where k is the ratio $h(\mathbb{M}_T, T_*)/h(\mathbb{M}_K, T_*)$, which is a function of $\mathbb{M}_T$, $\mathbb{M}_K$ and T_*. So, k is a constant for the certain subsample and certain observing baseline for *TESS*. The values of mission parameters and the response functions are given on the website of *Kepler* (https://keplerscience.arc.nasa.gov/the-kepler-space-telescope.html) and *TESS* (https://heasarc.gsfc.nasa.gov/docs/tess/the-tess-space-telescope.html). We use the mean value of effective temperature in each subsample as T_*, and the error range will be discussed in Section 4.4.4. Because r is a random variable, we generate it randomly under its distribution for *Kepler* and *TESS* to calculate k for 1000 times and take its mean value. The distribution of the distance to CTL stars is given on the portal (https://filtergraph.com/1371737). The results of k according to different baselines (t_m) for each subsample are listed in Table 1.

The probability that the S/N caused by an exoplanet with a given orbital period is higher than the threshold is a complex convolution including the planetary parameters and stellar parameters without the constant mission parameters. As defined above, in each subsample, the stellar parameters and planetary parameters do not change with the mission. Therefore, the complex convolution to calculate the probability is the same for both missions, and the difference between two probabilities is only due to the different mission parameters, which will be explained in more detail below.

Suppose in subsample $i (i = 1, 2)$ that the PDFs of S/N caused by the exoplanet with a fixed period for *TESS* and *Kepler* are $f_{\mathrm{S/NT}i}(\mathrm{S/N}|P, tr)$ and $f_{\mathrm{S/NK}i}(\mathrm{S/N}|P, tr)$. For a series of transit events, if they are observed by *Kepler*, the distribution of S/N will be $f_{\mathrm{S/NK}i}(\mathrm{S/N}|P, tr)$, and if they are observed by *TESS*, the distribution will be $f_{\mathrm{S/NT}i}(\mathrm{S/N}|P, tr)$. For a certain event, if it is observed by *Kepler*, the value of S/N is $\mathrm{S/N}_K$, and if it is observed by *TESS*, the value is $\mathrm{S/N}_T$, which means

$$\begin{aligned}
f_{\mathrm{S/NT}i}(\mathrm{S/N}_T|P, tr) &= f_{S/NTi}(k\cdot\mathrm{S/N}_K|P, tr)\\
&= f_{S/NKi}(\mathrm{S/N}_K|P, tr).
\end{aligned} \tag{18}$$

So the relation between two PDFs is

$$f_{\mathrm{S/NT}i}(\mathrm{S/N}|P,\,tr) = f_{\mathrm{S/NK}i}(\mathrm{S/N}/k|P,\,tr). \quad (19)$$

and the probability can be calculated as

$$\begin{aligned}
&\mathrm{Prob}_i(\mathrm{S/N}_T > \mathrm{S/NT_{min}}|P,\,tr)\\
&= \int_{\mathrm{S/NT_{min}}}^{\infty} f_{\mathrm{S/NT}i}(S/N'|P,\,tr)dS/N'\\
&= \int_{\mathrm{S/NT_{min}}}^{\infty} f_{\mathrm{S/NK}i}\left(\frac{S/N'}{k}|P,\,tr\right)dS/N'\\
&= \int_{\frac{\mathrm{S/NT_{min}}}{k}}^{\infty} k\cdot f_{\mathrm{S/NK}i}(S/N''|P,\,tr)dS/N''\\
&= k\cdot \mathrm{Prob_i}\left(\mathrm{S/N}_K > \frac{\mathrm{S/NT_{min}}}{k}|P,\,tr\right), \quad (20)
\end{aligned}$$

where we use the integral substitution. Only for a given stellar subsample and a given observing baseline can the substitution work. So the ratio of the probability is

$$\begin{aligned}
&\frac{\mathrm{Prob}_i(\mathrm{S/N}_T > \mathrm{S/NT_{min}}|P,\,tr)}{\mathrm{Prob}_i(\mathrm{S/N}_K > \mathrm{S/NK_{min}}|P,\,tr)}\\
&= k\cdot\frac{\mathrm{Prob_i}\left(\mathrm{S/N}_K > \frac{\mathrm{S/NT_{min}}}{k}|P,\,tr\right)}{\mathrm{Prob}_i(\mathrm{S/N}_K > \mathrm{S/NK_{min}}|P,\,tr)}. \quad (21)
\end{aligned}$$

According to the analysis in Section 3.2.2, adding a condition where at least three transits are observed has no influence on the right side of Equation (7) for orbital periods within 4 yr. Thus, the ratio of the S/N term can be calculated with the following substitution, in order to make use of the *Kepler* data that have the condition where every detected exoplanet has at least three transits:

$$\begin{aligned}
&\mathrm{Prob_i}\left(\mathrm{S/N}_K > \frac{\mathrm{S/NT_{min}}}{k}|P,\,tr\right)\\
&\quad\rightarrow \mathrm{Prob_i}\left(\mathrm{S/N}_K > \frac{\mathrm{S/NT_{min}}}{k}|P,\,tr,\,3trs_K\right)\\
&\mathrm{Prob}_i(\mathrm{S/N}_K > \mathrm{S/NK_{min}}|P,\,tr)\\
&\quad\rightarrow \mathrm{Prob}_i(\mathrm{S/N}_K > \mathrm{S/NK_{min}}|P,\,tr,\,3trs_K). \quad (22)
\end{aligned}$$

Then the situation for the *Kepler* mission is that the probability can be obtained from the known data set based on the law of large numbers:

$$\begin{aligned}
&\frac{\mathrm{Prob_i}\left(\mathrm{S/N}_K > \frac{\mathrm{S/NT_{min}}}{k}|P,\,tr,\,3trs_K\right)}{\mathrm{Prob}_i(\mathrm{S/N}_K > \mathrm{S/NK_{min}}|P,\,tr,\,3trs_K)}\\
&= \frac{N_{Pi}\left(\mathrm{S/N}_K > \frac{\mathrm{S/NT_{min}}}{k}\right)/N_{Pi}}{N_{Pi}(\mathrm{S/N}_K > \mathrm{S/NK_{min}})/N_{Pi}}\\
&= \frac{N_{Pi}\left(\mathrm{S/N}_K > \frac{\mathrm{S/NT_{min}}}{k}\right)}{N_{Pi}(\mathrm{S/N}_K > \mathrm{S/NK_{min}})}, \quad (23)
\end{aligned}$$

where N_{pi} is the total number of exoplanets with a given orbital period (P days), which has been detected in at least three transits. For the ones with brackets, only the exoplanets meeting the condition described in the brackets are counted.

The result can be derived from the *Kepler* data set by dividing the data into bins of different orbital period that contain a certain number of exoplanets (Tables 2, 3), and then performing linear regression or polynomial curve fitting based on the trend shown in the scatter (Figure 7). For subsample 1 ($T_{\mathrm{eff}} < 4500$ K), each bin holds 43 exoplanets, and for subsample 2 ($T_{\mathrm{eff}} > 4500$ K), each bin holds 202 exoplanets. We drop the data with the longest orbital period to make sure each bin has the same number of exoplanets. Here, we adopt the S/N threshold of 7.1 that was calculated for *Kepler* by Jenkins et al. (2010) and 7.3 for *TESS* applied by Sullivan et al. (2015). The results are shown in Table 2 for subsample 1 and Table 3 for subsample 2. If the fitting value is negative, it is taken as zero. Then we obtain the fitting analytic formula of the ratio of the terms about S/N in Equation (7), which can be calculated easily for a continuous orbital period.

Table 2
Ratios $(N_P(\mathrm{S/N}_K > (\mathrm{S/NT_{min}})/k))/(N_P(\mathrm{S/N}_K > \mathrm{S/NK_{min}}))$ in Different Bins in Subsample 1 (<4500 K)

Mean Bin Value	1.76	4.87	8.64	15.33	47.90	a	b
27 days	0.05	0.00	0.00	0.00	0.00	−0.03	0.04
54 days	0.07	0.02	0.02	0.02	0.00	−0.04	0.07
89 days	0.07	0.05	0.02	0.05	0.00	−0.04	0.08
108 days	0.07	0.07	0.07	0.07	0.00	−0.04	0.10
189 days	0.16	0.12	0.12	0.07	0.02	−0.10	0.19
351 days	0.26	0.19	0.23	0.21	0.02	−0.14	0.32

Note. There are 43 exoplanets in each bin and 10 bins in total. The mean values of each bin are listed in the top line. The last two columns show the parameters of the fitting curve: a is the coefficient of the result of linear regression, and b is the intercept.

4. Results and Discussion

4.1. Results of Two-transit and One-transit Criterion for M Dwarfs and Sun-like Stars

Based on Equation (7), we obtain the predicted distribution of orbital periods for *TESS* (Figure 8). For the two-transit criterion, the result is Figure 8(a) for subsample 1 (<4500 K) and (b) for subsample 2. For the one-transit criterion, the result is shown in Figures 8(c) and (d).

As shown in Figure 8, for exoplanets orbiting M dwarfs, the orbital period tends to be shorter compared to those with AFGK-type stars, and the peak value becomes shorter with the reduction of the observing baseline, while the trend is not that obvious for subsample 2. It indicates that the M dwarfs tend to have inner short-period exoplanets detected. But it also may result from the lack of data to estimate the probability of S/N in subsample 1.

For exoplanets orbiting stars with $T_{\mathrm{eff}} > 4500$ K, when requiring two transits, the orbital period distribution is much more sensitive to the observing baseline than when requiring at least one transit, especially for longer orbital periods. There is an obvious cutoff at the longer orbital period for subsample 2 when requiring two transits.

Although the orbital period is unconstrained when requiring only one transit, panel (c) shows the impossibility of detecting long-period exoplanets orbiting M dwarfs, which is similar to the result when requiring two transits. It is because S/N dominates the probability of detecting the long period but not the criteria of the times of detected transits for M dwarfs, which are shown in Figures 6 and 7. One reason is that M dwarfs are so faint that it is hard to confirm the single transit, due to the low S/N. The low S/N for M dwarfs also indicates the low occurrence of long-period giant exoplanets that are giant

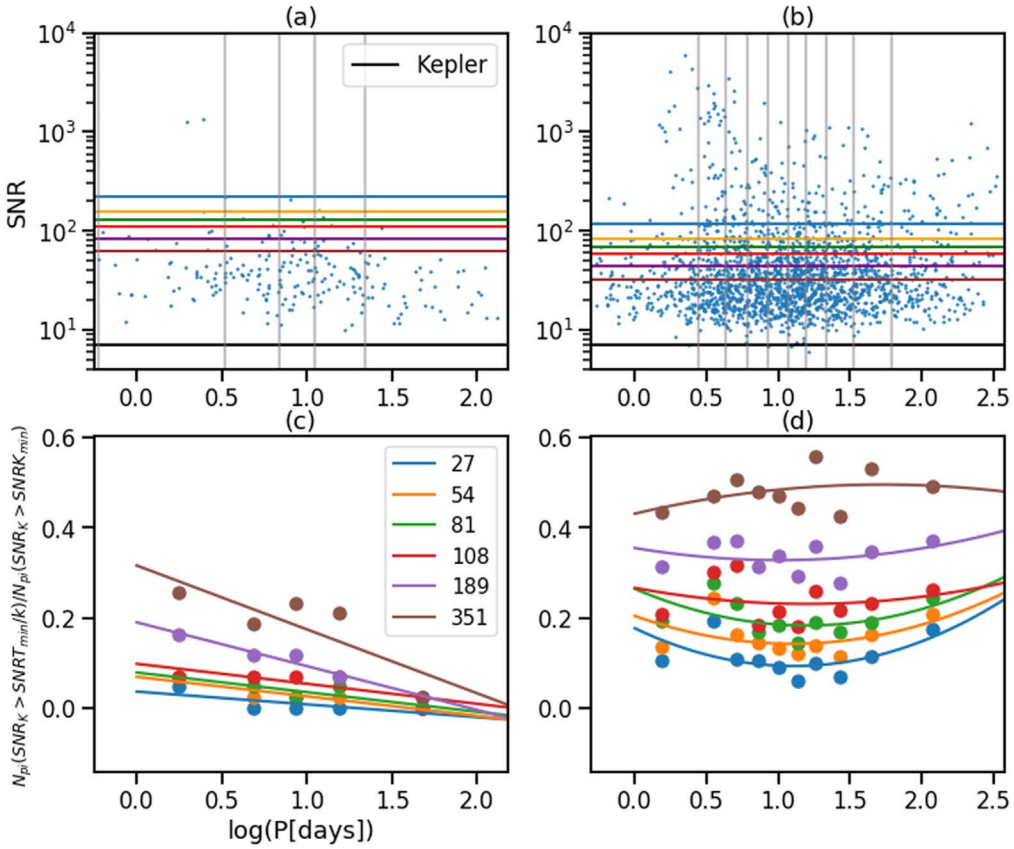

Figure 7. Top: The scatters show the orbital period and S/N of the exoplanets in the sample. The horizontal lines in different colors denote the thresholds for the different observational durations, and the legend is shown in (a). (a) Subsample 1 (<4500 K); (b) subsample 2 (>4500 K). The gray vertical line shows the boundary of each bin. There are 43 exoplanets in each bin in panel (a) and 202 exoplanets in each bin in panel (b). Bottom: (c) and (d) show the ratio of probability for each bin and the result of curve fitting.

Table 3
Ratios $(N_P(\mathrm{S/N}_K > (\mathrm{S/NT}_{\mathrm{min}})/k))/(N_P(\mathrm{S/N}_K > \mathrm{S/NK}_{\mathrm{min}}))$ in Different Bins in Subsample 2 (>4500 K)

Mean Bin Value	1.55	3.52	5.10	7.21	10.06	13.57	18.07	26.69	44.42	119.98	k_1	k_2	k_3
27 days	0.10	0.19	0.11	0.10	0.09	0.06	0.10	0.07	0.11	0.17	0.07	−0.15	0.18
54 days	0.13	0.24	0.16	0.14	0.13	0.12	0.14	0.11	0.16	0.21	0.05	−0.11	0.20
89 days	0.19	0.28	0.23	0.17	0.18	0.14	0.19	0.17	0.19	0.24	0.06	−0.14	0.26
108 days	0.21	0.30	0.32	0.18	0.21	0.18	0.26	0.22	0.23	0.26	0.02	−0.06	0.27
189 days	0.31	0.37	0.37	0.31	0.34	0.29	0.36	0.28	0.35	0.37	0.03	−0.05	0.35
351 days	0.43	0.47	0.50	0.48	0.47	0.44	0.56	0.43	0.53	0.49	−0.02	0.07	0.43

Note. There are 202 exoplanets in each bin and 20 bins in total. The mean values of each bin are listed in the top line. The last three columns show the parameters of the fitting curve: $y = k_1x^2 + k_2x + k_3$.

enough to produce the high S/N. In principle, the incompleteness of *Kepler* data could lead to this result. But according to our introduction, it will not have a significant impact.

4.2. Combination of Two Subsamples

We adopt the result of Barclay et al. (2018) that *TESS* will observe 371(n_1) M dwarfs hosting exoplanets and 922(n_2) AFGK-type stars hosting exoplanets. And we assume that the ratio of n_1 to n_2 is fixed for arbitrary sky areas. So, the PDF for the full sample is

$$\mathrm{Prob}(P|TESS) = c_T(n_1\mathrm{Prob}_1(P|TESS) + n_2\mathrm{Prob}_2(P|TESS)), \quad (24)$$

where c_T is a normalization coefficient. The PDF curves for the two-transit criterion and one-transit criterion are shown in Figure 9.

According to Figure 9, the orbital period distribution for *TESS* has an obvious cutoff at the long orbital period when at least two transits are required, especially for the short observing baseline, which indicates that there is a large incompleteness for exoplanets with long orbital periods, while there is no big difference between different observing baselines when requiring at least one transit.

4.3. Combination of Sectors

According to Barclay et al. (2018), the number of detected exoplanets of different observing baselines is 543, 273, 101, 193, 108, and 75 for 27 days, 54 days, 81 days, 108–297 days, 324 days, and 351 days. Taking them as the coefficient of each PDF for a given observing baseline and combining them, we obtain the results shown in Figure 10 and Table 6.

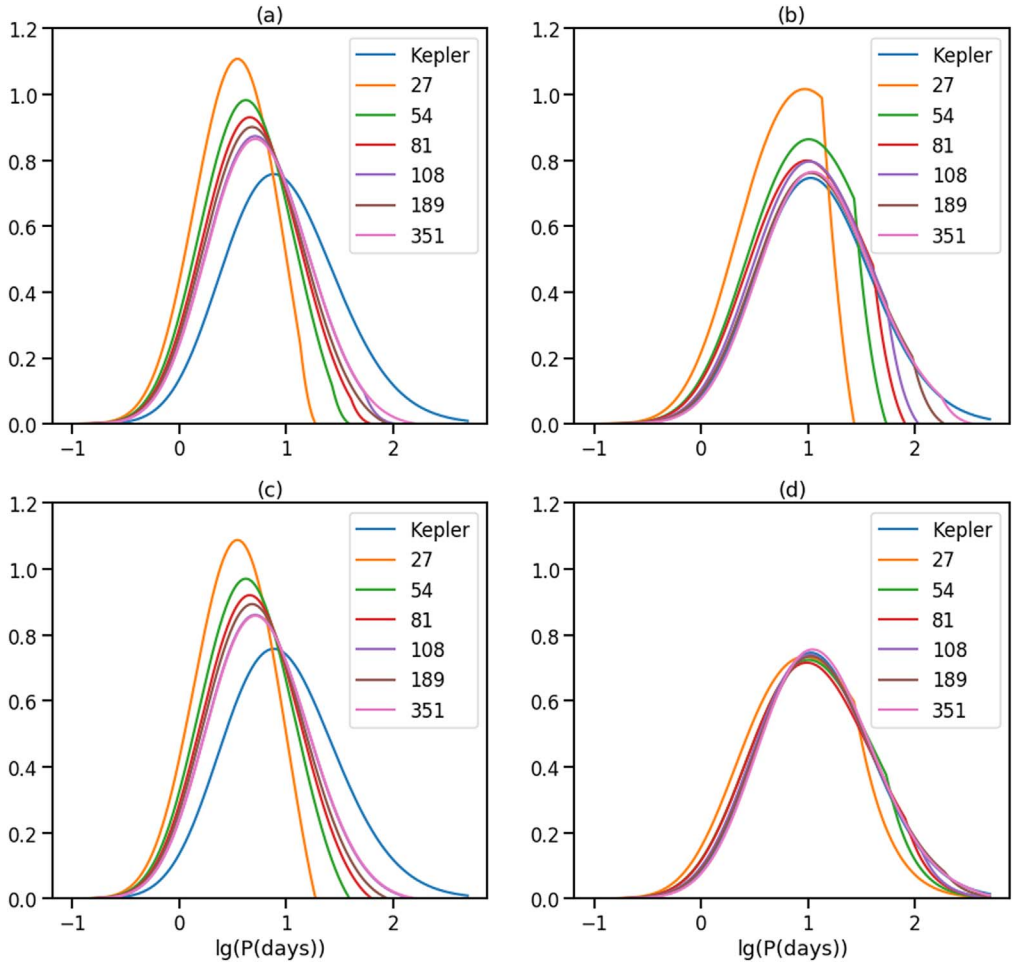

Figure 8. Top: probability density function of orbital period for different observing baselines of the *TESS* mission if at least two transits are needed: (a) the effective temperature ($T_{\rm eff}$) of the host star is <4500 K, and (b) the $T_{\rm eff}$ of the host star is >4500 K. Bottom: probability density function of orbital period for different observing baselines of the *TESS* mission if only one transit is needed: (c) the $T_{\rm eff}$ of the host star is <4500 K, and (d) the $T_{\rm eff}$ of the host star is >4500 K.

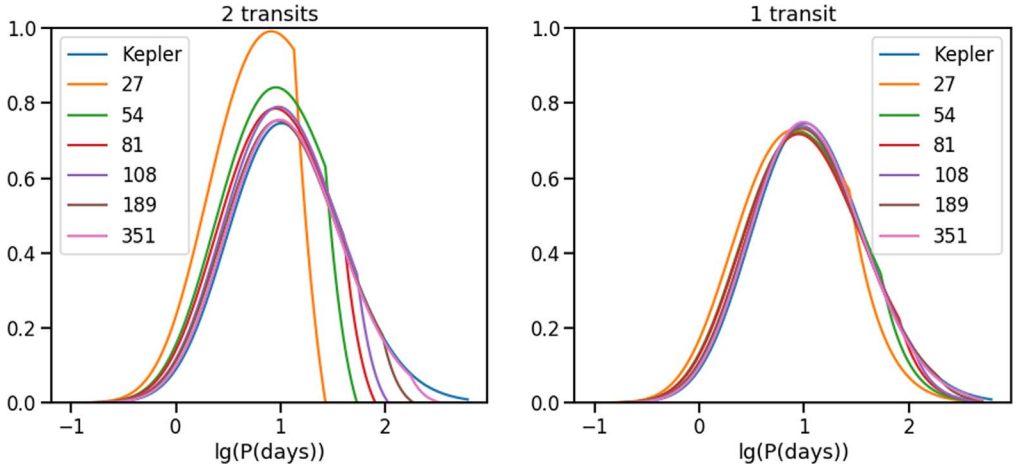

Figure 9. Left: criterion is two transits. Right: criterion is one transit.

For the two-transit criterion, the distribution has a cutoff at the orbital period longer than 20 days. This is because exoplanets with an orbital period longer than 27 days cannot be observed twice in most of the sky area. For the one-transit criterion, the orbital period distribution is similar to that of *Kepler*, though it shifts to the left a little.

4.4. Uncertainty Analysis

4.4.1. Uncertainty of Approximating Ntrs

To evaluate the influence of approximating *Ntrs*, we generate 1000 mock data points to analyze it. The distributions of every parameter contributing to S/N are shown in the first row in Figure 11.

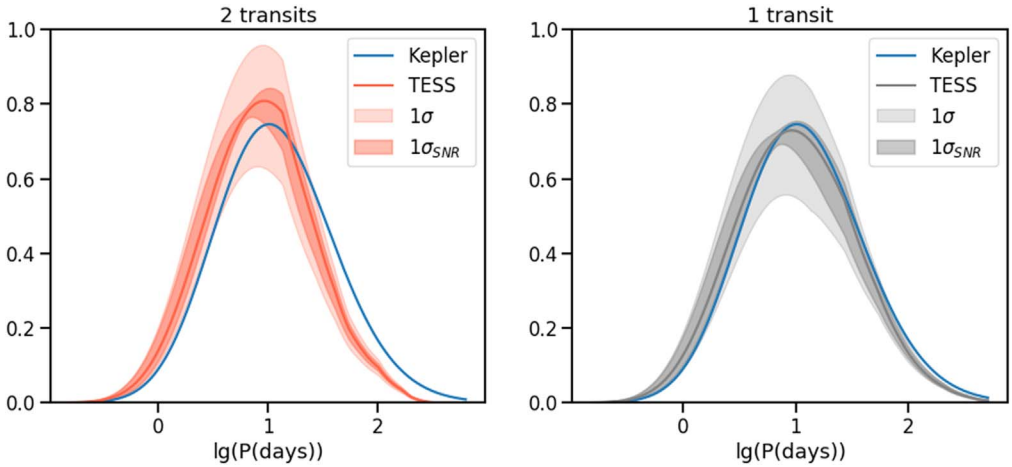

Figure 10. Left: criterion is two transits. Right: criterion is one transit. The light shadow is the possible range within one standard deviation, and the deep shadow is the possible range only considering the uncertainty caused by the term of S/N, which is discussed in Section 4.4.

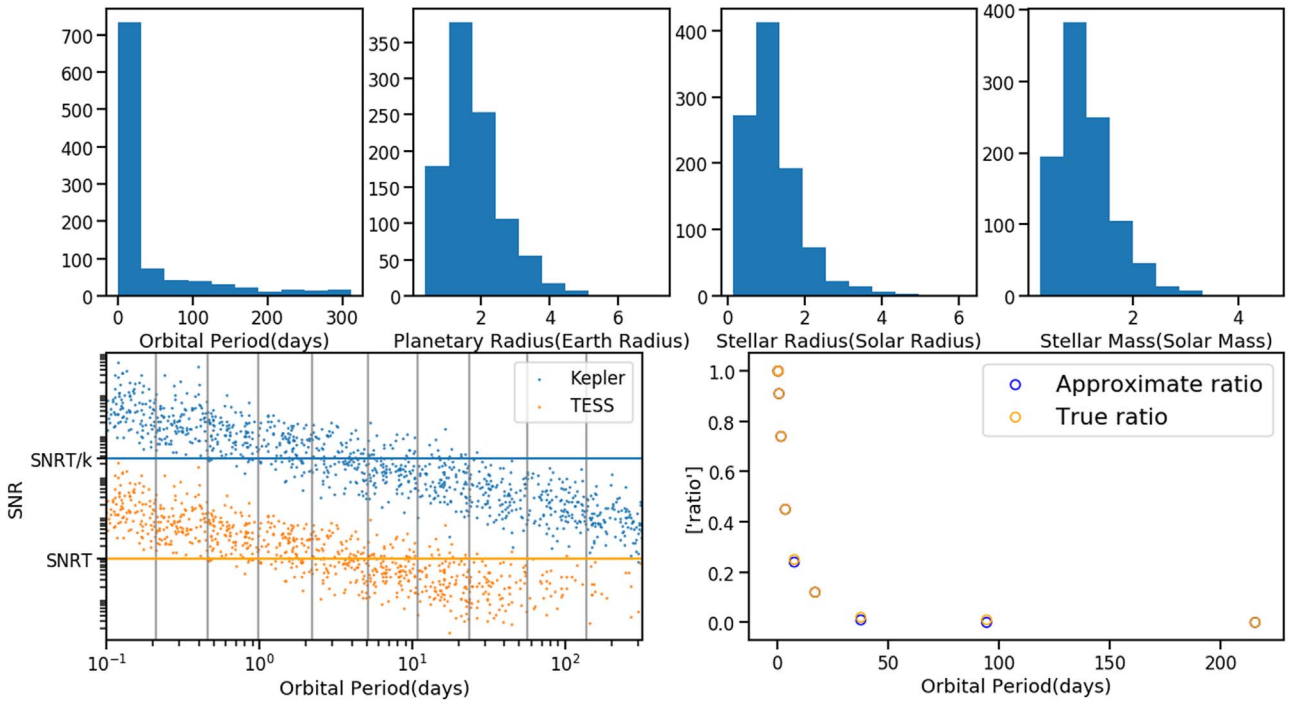

Figure 11. Top: distributions of all parameters contributing to S/N. Bottom left: The blue dots represent the signals detected by *Kepler*, and the orange dots represent the signals detected by *TESS*. The orange horizontal line is the threshold of the *TESS* mission (S/N_T), and the blue horizontal line is $S/N_T/k$ introduced in Equation (20), which is used to estimate the ratio of the terms about S/N in Equation (7). Right: The blue circles represent the ratio estimated based on *Kepler* data, and the orange circles represent the true ratio derived from the precise S/N of *TESS*. Some orange circles cover the blue circles because they have the same value.

At first, we simulated the S/N when those mock exoplanets are detected by *Kepler* using the precise formula

$$S/N = R_p^2\left(\frac{4\pi^2 P}{GM_*}\right)^{\frac{1}{6}}\sqrt{\frac{N_{trs}A}{4R_*r^2}\int_{\lambda_1}^{\lambda_2}\tau\pi B(\lambda, T_*)\left(\frac{\lambda}{hc}\right)d\lambda}, \quad (25)$$

where N_{trs} is the time of detected transit events. Unlike Equation (14) using the expected value of all exoplanets with the same orbital period, Equation (25) applies to a single detected exoplanet. The simulated signals are shown in the lower left panel in Figure 11, and N_{trs} is generated randomly under the Bernoulli distribution.

Then, we apply the same method introduced in Section 3.3 to estimate the ratio of the probability terms about S/N. The substitution of the threshold ($S/N_T/k$) is shown as the blue horizontal line in the lower left panel in Figure 11. Only exoplanets with signals above the threshold can be detectable by *TESS*. And we use this ratio of the number of exoplanets above the threshold to the number of exoplanets below the threshold to estimate the ratio of the probability terms about S/N for *TESS*, which is shown in the lower right panel in Figure 11.

In order to compare the estimate to the true ratio, we also simulated the S/N when those mock exoplanets are detected by *TESS* in the 27 day baseline using Equation (25), through which N_{trs} is generated randomly under the Bernoulli distribution. The results are shown as orange dots in the lower left panel in Figure 11, and the threshold (S/N_T) is the value of

the orange horizontal line. The results of true value are shown in the lower right panel of Figure 11 (orange circles). As we can see, the error of our estimation is small, which indicates its validity.

4.4.2. Discussion on Different Detection Models

In our model introduced in Section 3, the S/N includes all transit events during the observation. The exoplanet can be confirmed as long as its stacked S/N is above the threshold. In fact, the first single-transit event also needs to exceed a threshold so we can confirm it with subsequent multiple signals. Therefore, some exoplanets with lower single S/N would be excluded in this way, though the total S/N is higher than the threshold because of the multiple stacking. We ignore this threshold in this paper.

4.4.3. Uncertainty of Stellar Parameters

As the stellar parameters obey the same distributions for a given subsample, Prob($tr|P$) should not change with the missions, as Equation (5) indicates. But the integral in Equation (5) is only the expected value. In fact, this term is a variable drawn from a random distribution. Although the terms of two missions are drawn from the same distribution, their values may differ, leading to the dispersion of the ratio of the terms for two missions, which means that the ratio is not 1. Here we use different subscripts to distinguish these two variables. If so, Equation (7) should be rewritten as follows, and then the uncertainty of the probability density can be obtained by the propagation of uncertainty:

$$\begin{aligned}&\mathrm{Prob}_i(P|TESS)\\ &= c_i \mathrm{Prob}_i(P|Kepler)\cdot\frac{\mathrm{Prob}_a(tr|P)}{\mathrm{Prob}_b(tr|P)}\cdot\frac{\mathrm{Prob}_i(Ntrs_T|P,\,tr)}{\mathrm{Prob}_i(Ntrs_K|P,\,tr)}\\ &\cdot\frac{\mathrm{Prob}_i(\mathrm{S/N}_T>\mathrm{S/NT}_{min}|P,\,tr)}{\mathrm{Prob}_i(\mathrm{S/N}_K>\mathrm{S/NK}_{min}|P,\,tr)},\ i=1,\,2\end{aligned} \tag{26}$$

where the second term (geometric probability term) and the fourth term (S/N term) include the random parameters.

According to Equation (5), the uncertainty of geometric probability should be propagated by the uncertainties of the stellar parameters (R_*, M_*). Here, we regard this term as a single variable and obtain the statistics directly (Figure 12, Table 7). According to the propagation of uncertainty, the uncertainty of the ratio is

$$\sigma_{\mathrm{Prob}(tr|P)/\mathrm{Prob}(tr|P)}=\sqrt{2\cdot(\sigma_{\mathrm{Prob}(tr|P)}/\mathrm{Prob}(tr|P))^2}. \tag{27}$$

The values are shown in Table 7.

4.4.4. Uncertainty of the S/N Model

As for the uncertainty of the S/N term, the standard deviation of k, or σ_k, could only be derived from the deviations of effective temperature and distance to the stars. The relative error of k caused by the variance of effective temperature is

$$\frac{k_{\max}(\mathbb{M}_T,\,\mathbb{M}_K,\,T')-k_{\min}(\mathbb{M}_T,\,\mathbb{M}_K,\,T'')}{k(\mathbb{M}_T,\,\mathbb{M}_K,\,T_*)}, \tag{28}$$

where T' is the effective temperature that maximizes the k, and T'' minimizes k. For example, the relative error of k for 27 days of observation is 1.5×10^{-6} for M dwarfs with effective temperature ranging from 1 to 4500 K and 5.2×10^{-9} for Sun-like stars with effective temperature ranging from 4500 to 10,000 K, respectively. The relative errors for other observational baselines are of the same order of magnitude for each subsample. These variances lead to little influence on the fitting results shown in Figure 7, so we ignore the uncertainty of the final result caused by them.

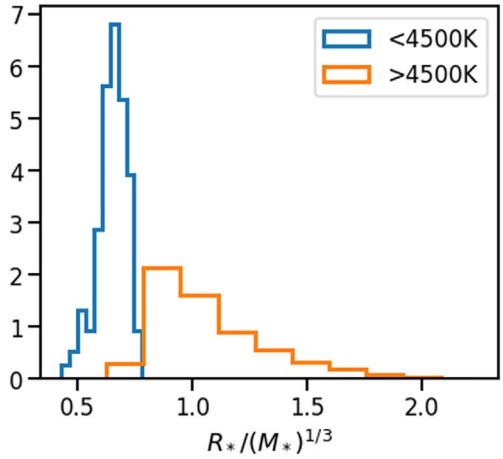

Figure 12. Distribution of the random variable term in Prob($tr|P$): $\mathrm{Prob}(tr|P)=(4\pi^2/G)^{1/3}R_*M_*^{-1/3}P^{-2/3}$. We plot the histogram of $R_*M_*^{-1/3}$ to illustrate its distribution and obtain its standard deviation to calculate the uncertainty of the ratio by the propagation of uncertainty (Equation (27)).

As Section 3.3 introduced, in order to deal with the uncertainty caused by the distance, we generate r randomly under its distribution for *Kepler* and *TESS* to calculate k and take the following steps using it 1000 times. The 1σ range appears in Figure 10 as deep shadows for the two-transit criterion and one-transit criterion, respectively.

Taking all kinds of uncertainties into consideration, the uncertainty of the PDF can be obtained by the propagation of the above uncertainties, and the results are shown in Figure 10 as light shadows.

4.5. Comparison with Previous Studies

In order to compare our results with previous studies in the literature, we generated 1000 simulated data points according to the PDF given above by rejection sampling, which is shown in Figure 13.

For the two-transit criterion, we compare our results to the newly available *TESS* alerts data (https://tess.mit.edu/alerts/) and the results from Sullivan et al. (2015) and Barclay et al. (2018). Since the available *TESS* data are only from two sectors, we simulated data according to the PDF of an orbital period for a 27 day observing baseline. Compared to the *TESS* alerts, which are from preliminary data and likely contain false positives, our simulation results show good agreement with the *TESS* preliminary data. However, our simulation shows that *TESS* will detect slightly more longer-period exoplanets, and the peak is also slightly wider than that of alerts.

Our simulated sample also contains a larger proportion of longer-period exoplanets compared to both Sullivan et al. (2015) and Barclay et al. (2018). The cause might be that the occurrence rates from both Fressin et al. (2013) and Dressing & Charbonneau (2015), which were used in the studies of Sullivan et al. and Barclay et al., are limited in orbital period to

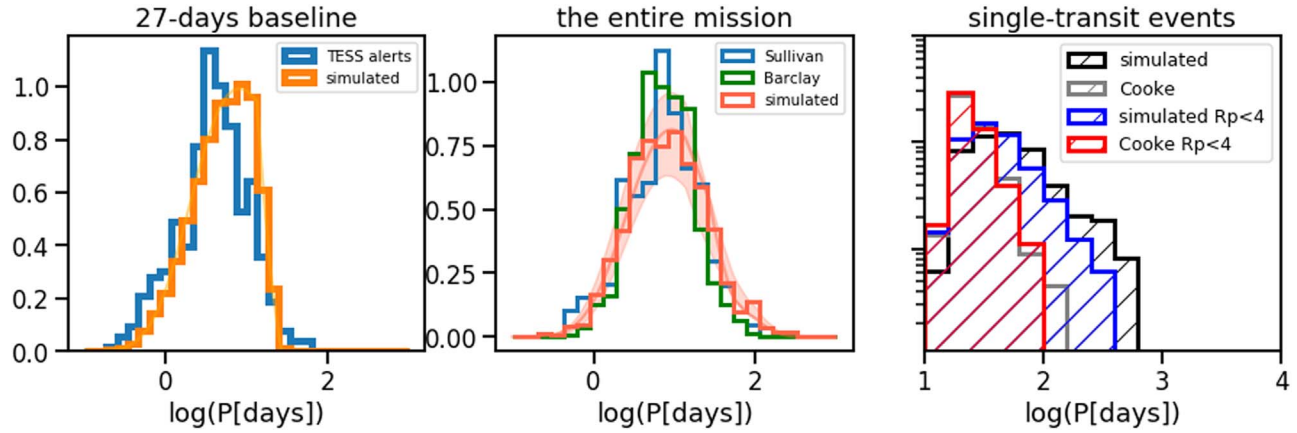

Figure 13. Left panel: The blue profile is the histogram of *TESS* alerts, and the orange profile is the histogram of the simulated data under the probability density function of orbital period for a 27 day observing baseline. Middle panel: The blue profile is the histogram of simulated data from Sullivan et al. (2015), the green profile is the histogram of the simulated data from Barclay et al. (2018), and the pink profile is the histogram of the simulated data under the probability density function of orbital period for the entire *TESS* mission. The shadow is the possible range within one standard deviation. Right: The orbital period histogram of simulated single-transit events is in black. The histograms of single-transit events of sub-Neptunes are in blue and red, representing our results and the results of Cooke et al. (2018). The length and range of the bins are the same as those of Figure 1 from Villanueva et al. (2019).

Table 4
Mean Orbital Period and Detection Range within 1σ of *TESS*'s Most Frequently Detected Orbital Period of Exoplanets in Each Duration of Observations for Each Subsample, Assuming the Detection Criteria are Two Transits and One Transit

Duration of Observation (days)	27	54	81	108	189	351	*Kepler*
MP (days)-sub1-2 transits	3.16	4.10	4.67	5.52	5.07	5.67	9.27
1σ (days)-sub1-2 transits	1.44–6.93	1.69–9.95	1.82–12.00	2.01–15.19	1.89–13.56	2.01–16.03	2.73–31.52
MP (days)-sub2-2 transits	5.17	7.57	8.59	9.71	11.05	12.16	⋯
1σ (days)-sub2-2 transits	2.20–12.14	2.89–19.83	3.05–24.23	3.38–27.95	3.51–34.29	3.80–38.97	⋯
MP (days)-sub1-1 transit	3.12	4.05	4.62	5.56	5.00	5.60	12.23
1σ (days)-sub1-1 transit	1.41–6.93	1.65–9.95	1.78–12.01	1.96–15.73	1.86–13.47	1.97–15.93	3.54–42.20
MP (days)-sub2-1 transit	8.56	10.44	10.75	11.38	12.02	12.49	⋯
1σ (days)-sub2-1 transit	2.68–27.33	3.18–34.26	3.18–36.37	3.43–37.75	3.53–40.89	3.75–41.65	⋯

Note. MP: Mean orbital period. For comparison, *Kepler*'s most frequently detected exoplanet orbital period is also shown under *Kepler*'s detection criteria of three transits.

0.5–85 and 0.5–200 days, respectively. For the longer observing baseline, their input data are not complete.

For the one-transit criterion, we extract the single-transit events from the entire yields shown in Figure 10 and generate 241 data points in the histogram in Figure 13, in order to compare it to the result shown in Figure 1 from Villanueva et al. (2019). They show the same trend, though the peak is not as obvious as the result of Villanueva et al. (2019). They assumed that the occurrence of the longer orbital period is equal to the occurrence of the longest complete orbital period. The planet occurrence rates are only complete to periods of 60 days for AFGK-type stars and 30 days for M dwarfs because they are derived from planetary radius and orbital period. In this study, we focus on the orbital period, so the distribution of orbital period up to 695 days can be derived from the *Kepler* data set, and the distribution of *Kepler* shows the low occurrence of the longer orbital period. The results of Cooke et al. (2018) are far more narrow than both results discussed above. Cooke et al. (2018) and Villanueva et al. (2019) applied the same input catalog and used the same occurrence rate, but with different approaches. It can be seen that the detection model of Cooke et al. (2018) is more stringent. In this study, we apply the S/N threshold of 7.3 for both multiple transits and a single transit, although it may be optimistic.

Cooke et al. (2018) also gave the predicted orbital distribution of sub-Neptunes detected by *TESS*, which is listed in Table 4 from Cooke et al. (2018). In order to compare our results to that, we resample the *Kepler* data by excluding those exoplanets with radius greater than 4 Earth radii and repeat the above steps introduced in Section 3 to obtain the PDF of orbital periods of sub-Neptunes. Then we generate 241 data points and plot the histogram for comparison (Figure 13). The lower yields of sub-Neptunes with longer orbital period indicate their lower occurrence rate. This trend can also be seen in the results of Cooke et al., while the range of our results is wider, which has been discussed above.

5. Summary and Discussion

The orbital period detectable by *TESS* varies in different regions of the celestial sphere, due to different observation durations. Transit surveys are strongly biased toward short periods, especially for *TESS*, whose observation period is very short in most of the celestial sphere. Our mathematical analysis finds (Table 5) that, for the two-transit criterion, the expected mean value of the most frequently detected orbital period is 5.01 days with a most detected range of 2.12–11.82 days in the region with the observation of 27 days. Near the poles where the observational duration is 351 days, the expected mean

Table 5
Mean Orbital Period and Detection Range within 1σ of *TESS*'s Most Frequently Detected Orbital Period of Exoplanets in Each Duration of Observations, Assuming the Detection Criteria are Two Transits and One Transit

Duration of Observation (days)	27	54	81	108	189	351	*Kepler*
MP (days)-2 transits	5.01	7.06	8.00	9.21	10.07	10.93	11.89
1σ (days)-2 transits	2.12–11.82	2.64–18.91	2.77–23.10	3.17–26.76	3.19–31.81	3.35–35.65	3.45–41.04
MP (days)-1 transit	8.17	9.74	10.12	10.96	10.96	11.25	⋯
1σ (days)-1 transit	2.55–26.21	2.95–32.19	2.98–34.31	3.21–35.62	3.20–37.57	3.34–37.91	⋯

Note. MP: Mean orbital period. For comparison, *Kepler*'s most frequently detected exoplanet orbital period is also shown under *Kepler*'s detection criterion of three transits.

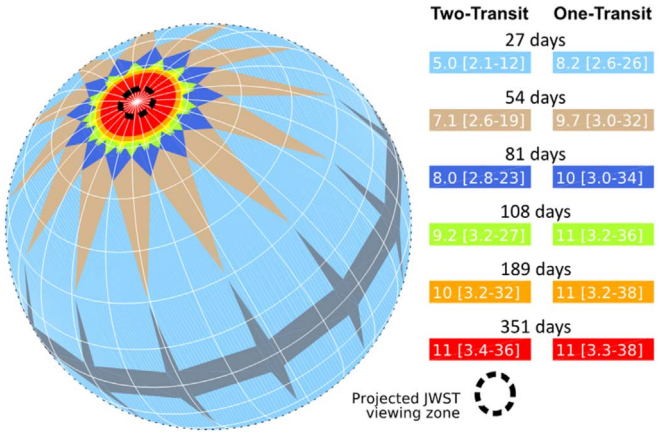

Figure 14. Duration of observations on the celestial sphere by *TESS*, distinguished by color, reproduced after Ricker et al. (2015). The mean orbital period and detection range of *TESS*'s most frequently detected orbital period of exoplanets for two-transit and one-transit criteria are shown as text inside the color bars.

orbital period is 10.93 days with a most detected range from 3.35 to 35.65 days. For one transit, the most frequently detected orbital period is 8.17 days in the region with the observation of 27 days and 11.25 days near the poles. These results are illustrated in Figure 14. For the entire *TESS* mission containing several sectors (Table 6), the mean value of orbital period is 8.47 days for two transits and 10.09 days for one transit. With such orbital periods, the majority of exoplanets detected by *TESS* will be closer to their host stars than Mercury is to our Sun, whose orbital period is 88 days.

The major difference between this study and the previous study is leveraging some of the similarities in biases between *TESS* and *Kepler* to cut out many of the steps taken to make a mock catalog. Admittedly, this study involves some approximations, including treating each of the stellar subsamples as homogeneous and simplifying the S/N model.

Unlike the previous studies that simulate the detected exoplanet with several properties, we only focus on the orbital period. Previous simulations (e.g., Sullivan et al. 2015; Barclay et al. 2018; Cooke et al. 2018; Huang et al. 2018, and Villanueva et al. 2019) require the general conclusion of the occurrence of the exoplanets. To do that, they might need to know about the distribution of different properties and the correlation among them. Actually, it is incredibly hard to recover the data, due to selection bias, let alone to obtain the general conclusion. Therefore, they have to ignore some relations and simplify the occurrence model of exoplanets to perform the simulation by generating some parameters from a random uniform distribution. If we perform the estimate

Table 6
Mean Orbital Period and Detection Range within 1σ of *TESS*'s Most Frequently Detected Orbital Period of Exoplanets, Assuming the Detection Criterion Is Two Transits and One Transit

Duration of Observation (days)	*TESS*	*Kepler*
MP (days)-2 transits	8.47	11.89
1σ (days)-2 transits	2.75–26.12	3.45–41.04
MP (days)-1 transit	10.09	⋯
1σ (days)-1 transit	2.99–34.08	⋯

Note. MP: mean orbital period. For comparison, *Kepler*'s most frequently detected exoplanet orbital period is also shown under *Kepler*'s detection criterion of three transits.

Table 7
Standard Deviation of Random Variable Terms

Statistic	<4500 K	>4500 K
$\frac{\sigma_{\mathrm{Prob}(tr\|P)}}{\mathrm{Prob}(tr \mid P)}$	0.0648	0.2571
$\frac{\sigma_{\mathrm{Prob}(tr\|P)}/\mathrm{Prob}(tr\|P)}{\mathrm{Prob}(tr \mid P)\ /\ \mathrm{Prob}(tr \mid P)}$	0.1407	0.3380

directly on the data, we can take those parameters and their relation into consideration.

We could further predict the distribution of other planetary properties. For example, the planetary radius affects the S/N, so we could estimate the radius distribution by dividing the radius sample into several subsamples based on orbital period and then perform an analysis on how the S/N makes a difference in the distribution of an individual radius. This could help us to predict the yields of habitable exoplanets.

The significance of this work lies not only in its ability to predict but also in its utility as a tool to perform an analysis between two data sets. For example, we could predict the planetary radius based on the prediction of the orbital period distribution that agrees with the real data. If the result is far different from the real yields, this would suggest that there is no dependence between radius and orbital period.

The major reason why we choose to predict the orbital period distribution rather than other properties is that the major difference between *TESS* and *Kepler* is the duration of observation. It makes the biggest difference in orbital period, which means the changes in the distribution of orbital period between the two missions should be the most notable, making it easier to do follow-on research.

Furthermore, in this study, a vital assumption is that the stellar parameters and planetary parameters that determine the occurrence of exoplanets obey the same distribution for M

dwarfs and AFGK-type stars, respectively; this similarity is often applied in the literature. If the yield of *TESS* is very different from our prediction, the assumption should be questioned. If so, we could change the assumption and perform an analysis to see which parameters can dominate the exoplanet occurrence based on the methodology of this study.

This work was supported by an exoplanet study initiative at the Jet Propulsion Laboratory, California Institute of Technology, under contract with NASA. X.J. thanks the Joint Institute for Regional Earth System Science and Engineering, University of California, Los Angeles, for hosting her summer research via a student exchange program. We thank Junyang Long of Chinese University of Hong Kong for helpful comments on the statistical method. Funding for the *TESS* mission is provided by NASA's Science Mission Directorate, and we acknowledge the use of *TESS* Alert data, which is currently in a beta test phase, from pipelines at the *TESS* Science Office and at the *TESS* Science Processing Operations Center.

Data availability: The exoplanet data used for this study can be downloaded at the NASA Exoplanet Archive (https://exoplanetarchive.ipac.caltech.edu/) and the *TESS* Alerts (https://tess.mit.edu/alerts/). The simulated data generated as the result of this study will be available for two years and can be downloaded from the third-party repository at https://zenodo.org/record/3245557. Please contact the corresponding author for any questions at Jonathan.H.Jiang@jpl.nasa.gov.

ORCID iDs

Jonathan H. Jiang https://orcid.org/0000-0002-5929-8951
Nicolas Cowan https://orcid.org/0000-0001-6129-5699
Zonghong Zhu https://orcid.org/0000-0002-3567-6743

References

Allen, C. W. 1976, Astrophysical Quantities (3rd ed.; London: Athlone)
Ballard, S. 2019, ApJ, 157, 113
Barclay, T., Pepper, J., & Quintana, E. V. 2018, ApJS, 239, 2
Batalha, N. M., Rowe, J. F., Bryson, S. T., et al. 2013, ApJS, 204, 24
Cooke, B. F., Pollacco, D., West, R., McCormac, J., & Wheatley, P. J. 2018, A&A, 619, A175
Cowan, N. B., Greene, T., Angerhausen, D., et al. 2015, PASP, 127, 311
Dressing, C. D., & Charbonneau, D. 2015, ApJ, 807, 45
Fressin, F., Torres, G., & Charbonneau, D. 2013, ApJ, 766, 81
Fulton, B. J., Petigura, E. A., Howard, A. W., et al. 2017, AJ, 154, 109
Howard, A. W., Marcy, G. W., Bryson, S. T., et al. 2012, ApJS, 201, 15
Huang, C. X., Shporer, A., Dragomir, D., et al. 2018, AJ, submitted
Jenkins, J. M., Caldwell, D. A., Chandrasekaran, H., et al. 2010, ApJL, 713, L87
Jiang, J. H., & Zhu, S. 2018, RNAAS, 2, 185
Kipping, D. D., & Standford, E. 2016, MNRAS, 463, 1323
Kopparapu, R. K., Ramirez, R., Kasting, J. F., et al. 2013, ApJ, 765, 131
Ricker, G. R., Winn, J. N., Vanderspek, R., et al. 2015, JATIS, 1, 014003
Stassun, G. K., Oelkers, R. J., Pepper, J., et al. 2018, AJ, 156, 102
Sullivan, P. W., Winn, J. N., Berta-Thompson, Z. K., et al. 2015, ApJ, 809, 77
Twicken, J. D., Jenkins, J. M., Seader, S. E., et al. 2016, AJ, 152, 158
Villanueva, S., Jr., Dragomir, D., & Gaudi, B. S. 2019, ApJ, 157, 84
Wang, C.-Y., & Peng, Y. 2015, ConPh, 56, 209