HAL
open science

Simulating radial velocity observations of trappist-1 with SPIRou

Baptiste Klein, J-F Donati

To cite this version:

Baptiste Klein, J-F Donati. Simulating radial velocity observations of trappist-1 with SPIRou. Monthly Notices of the Royal Astronomical Society, 2019, 488 (4), pp.5114-5126. 10.1093/mnras/stz1953 . obspm-02399724

HAL Id: obspm-02399724

https://hal-obspm.ccsd.cnrs.fr/obspm-02399724v1

Submitted on 28 May 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **488,** 5114–5126 (2019)
Advance Access publication 2019 July 15

doi:10.1093/mnras/stz1953

Simulating radial velocity observations of trappist-1 with SPIRou

Baptiste Klein[1,2★] and J.-F. Donati[1,2]

[1] *Universite de Toulouse, UPS-OMP, IRAP, 14 Avenue E. Belin, Toulouse F-31400, France*
[2] *CNRS, IRAP/UMR 5277, Toulouse, 14 Avenue E. Belin, Toulouse F-31400, France*

Accepted 2019 July 11. Received 2019 July 8; in original form 2019 January 29

ABSTRACT

We simulate a radial velocity (RV) follow-up of the TRAPPIST-1 system, a faithful representative of M dwarfs hosting transiting Earth-sized exoplanets to be observed with SPIRou in the months to come. We generate an RV curve containing the signature of the seven transiting TRAPPIST-1 planets and a realistic stellar activity curve statistically compatible with the light curve obtained with the *K2* mission. We find a ± 5 m s^{-1} stellar activity signal comparable in amplitude with the planet signal. Using various sampling schemes and white noise levels, we create time-series from which we estimate the masses of the seven planets. We find that the precision on the mass estimates is dominated by (i) the white noise level for planets c, f, and e and (ii) the stellar activity signal for planets b, d, and h. In particular, the activity signal completely outshines the RV signatures of planets d and h that remain undetected regardless of the RV curve sampling and level of white noise in the data set. We find that an RV follow-up of TRAPPIST-1 using SPIRou alone would likely result in an insufficient coverage of the rapidly evolving activity signal of the star, especially with bright-time observations only, making statistical methods such as Gaussian Process Regression hardly capable of firmly detecting planet f and accurately recovering the mass of planet g. In contrast, we show that using bi-site observations with good longitudinal complementary would allow for a more accurate filtering of the stellar activity RV signal.

Key words: techniques: radial velocities – stars: individual: TRAPPIST-1 – stars: activity – methods: statistical.

1 INTRODUCTION

Transiting Earth-sized exoplanets are prime targets to better understand planet formation and evolution. In addition to estimating planet radii from transit depths, one can also measure planet masses through ground-based follow-ups of the host star radial velocity (RV) using high-precision velocimeters. The resulting mass-radius relations are used to constrain the planet interior structures and, beyond that, the whole paradigm of planet formation (e.g. Weiss & Marcy 2014; Zeng, Sasselov & Jacobsen 2016; Alibert & Benz 2017; Dorn et al. 2018). However, most of Earth-sized planets unveiled with the *Kepler* space telescope through transit photometry lack mass measurement, as they produce low-amplitude RV signatures ($\lesssim$1 m s^{-1}) on stars that are too faint for RV surveys in the *V* band. As a result, the mass–radius diagram of Earth-sized planets is populated by only a handful of planets with well-constrained bulk density (Santerne et al. 2018).

M dwarfs are the most promising targets to unveil Earth-like exoplanets with precise masses and radii. Beyond the fact that they largely outnumber stars with earlier spectral type in the solar neighbourhood (Henry et al. 2006), they feature smaller sizes and masses, and more compact habitable zones (HZ), making HZ terrestrial planets easier to detect as well as rosy candidates for further atmosphere characterization with the *James Webb Space Telescope* (Morley et al. 2017). Over the past few years, a growing number of attractive Earth-sized exoplanets have been discovered around M dwarfs from photometric surveys (Berta-Thompson et al. 2015; Dittmann et al. 2017; Gillon et al. 2017). This trend is expected to step up with the ongoing TESS mission (Ricker et al. 2015) as M dwarfs are known to frequently host multiple terrestrial planetary systems (e.g. Bonfils et al. 2013; Dressing & Charbonneau 2015; Gaidos et al. 2016).

The Spectro-Polarimetre Infra-Rouge (SPIRou; see Donati et al. 2018 for a complete review of the instrument and related science) is a near infrared (nIR) échelle spectropolarimeter at the Canada–France–Hawaii Telescope (CFHT) whose science observations have recently started. The combination of a resolving power of ~70 000 over the *YJHK* bands and a goal RV precision of ~1 m s^{-1} makes SPIRou ideally suited for detecting Earth-twins around M dwarfs, not least in the HZ where they typically produce an RV stellar reflex motion of ~1 m s^{-1} (Artigau et al. 2018).

★ E-mail: baptiste.klein@irap.omp.eu

© 2019 The Author(s)
Published by Oxford University Press on behalf of the Royal Astronomical Society

Downloaded from https://academic.oup.com/mnras/article/488/4/5114/5532370 by guest on 28 May 2023

The SPIRou Legacy Survey (SLS) is a 300n programme dedicated to the main science goals of SPIRou, with 75n focussing on the RV follow-up of transiting planets around M dwarfs. This component of the SLS (called the SLS-TF for SLS transit follow-up) is expected to provide accurate mass measurements for the most attractive transiting planets to be unveiled by a wide range of photometric surveys (e.g. TESS; *K2*, Howell et al. 2014; MEarth, Nutzman & Charbonneau 2008; ExTrA, Bonfils et al. 2015; TRAPPIST, Gillon et al. 2017; NGTS, Wheatley et al. 2018). SPIP,[1] the upcoming SPIRou's twin at Telescope Bernard Lyot (TBL at Pic du Midi observatory), will provide additional means to refine the observation strategies of such follow-ups.

Transit photometry provides a priori information on the orbits of the planets around host stars to be monitored within the SLS-TF that can be used to optimize the RV follow-up strategy. Various strategies of RV curve sampling have already been studied in the literature (e.g. Ford 2008; Burt et al. 2018). It turns out that uniform samplings of planet RV curves perform better than more specific strategies, e.g. aiming at orbital quadratures where RVs are expected to be largest.

Besides, a large fraction of M dwarfs exhibits strong magnetic activity resulting in significant photometric and spectroscopic fluctuations (e.g. Basri et al. 2010; West et al. 2011). The induced RV signal tends to mimic planet signatures, making the former extremely troublesome to filter (e.g. Saar & Donahue 1997; Queloz et al. 2001; Desort et al. 2007; Dumusque et al. 2011; Borgniet, Meunier & Lagrange 2015). If not properly modelled and filtered out, this stellar activity signal is expected to strongly degrade the RV characterization of the SLS-TF targets. Moreover, it imposes additional constraints on the temporal sampling of the RV follow-up that also needs to densely cover both the rotational cycle of the host star and the typical time-scale on which it evolves. Last but not least, it makes it extremely hard to recover planets whose orbital periods lie close to the stellar rotation period or its harmonics.

In this paper, we simulate an RV follow-up of TRAPPIST-1 (2MASS J23062928-0502285), an ultracool M dwarf hosting seven transiting terrestrial planets including 3 in the HZ (Gillon et al. 2016, 2017) and possibly more non-transiting ones (Kipping 2018). Both the star and the planetary system are well constrained from photometric surveys (Luger et al. 2017; Delrez et al. 2018; Grimm et al. 2018, hereafter G18; Van Grootel et al. 2018) and the stellar light curve exhibits quasi-periodic fluctuations, indicating the presence of active regions at the surface of the star (Vida et al. 2017) that makes this system a representative of the SLS-TF targets. Moreover, TRAPPIST-1 will be observed as a member of the SLS-TF targets in order to provide Transit Timing Variations (TTV)-independent planet mass estimates, still highly required given the significant differences between TTV and RV planet mass measurements reported in the recent literature (see Mills & Mazeh 2017 and references therein).

We describe in Section 2 the simulation of a SPIRou RV follow-up of TRAPPIST-1 involving two distinct steps. First, we synthesize a TRAPPIST-1 RV curve containing both the planet signatures and a realistic stellar activity component scaled on the *K2* light curve (Luger et al. 2017). We then select the data points according to various sampling schemes to build a range of data sets with different properties. In Section 3, we detail our model to recover the planet masses from a given data set. We present our results for the different sampling configurations in Section 4 and finally discuss the results and conclude in Section 5.

[1] http://www.tbl.omp.eu/INSTRUMENTATION2/spip

Table 1. Best estimates of TRAPPIST-1 mass, radius, and effective temperature from Van Grootel et al. (2018).

Quantity	Value
$M_*/M_\odot$	0.089
$R_*/R_\odot$	0.121
$T_{\rm eff}$ (K)	2516

2 SYNTHETIC DATA SETS

We first synthesize a densely sampled RV curve containing both the planetary signature and a realistic realization of the stellar activity RV signal. We then use various sampling schemes to select observational data sets to which we add random values in order to account for instrument and photon noises. The main stellar parameters used in this paper are indicated in Table 1.

2.1 Planetary RV signature

Based on the low planet eccentricities reported in G18 and the planetary system stability discussed therein, we consider the simple case of seven planets with circular coplanar orbits assuming no planet–planet interaction. The orbital periods are set to the best estimates from Delrez et al. (2018, hereafter D18). We used Kepler's third law to derive the RV semi-amplitudes from the stellar mass measured by Van Grootel et al. (2018, see Table 1) and the last estimates of the planet masses from TTVs (see G18). Initial phases were injected to the planetary orbits using the mid-transit times from D18 with respect to their time reference. The orbital period, mass, semi-amplitude, and initial phase of the injected planets are shown in Table 2.

2.2 Synthesis of the stellar activity RV signal

TRAPPIST-1 light curve obtained as part of campaign 12 of the NASA *K2* mission exhibits quasi-periodic fluctuations of ∼10 mmag peak-to-peak (Luger et al. 2017), symptomatic of the presence of active regions at the stellar surface (Vida et al. 2017). Such stellar activity is expected to produce a significant RV signal that needs to be taken into account in a realistic way when simulating the RV follow-up of the star. By synthesizing simultaneously a photometric curve and the corresponding RV curve, we ensure that the data set we produce is broadly consistent with the statistical properties of the *K2* light curve of Trappist-1. This implies that we neglect every RV perturbation induced by the convective blueshift which is anyway expected to be strongly damped for M dwarfs compared to Sun-like stars (e.g. Meunier et al. 2017). This also implies that we ignore the Zeeman splitting when computing the line profile, as the magnetic field of TRAPPIST-1 remains poorly constrained.

We model the stellar surface by a dense grid of 20 000 cells with identical projected area when crossing the meridian. With the aim of including dark and bright inhomogeneities at the stellar surface, we assign a relative brightness factor to each of the cells. For each grid cell, we compute the local intensity line profile using the Unno-Rachkovski analytical solution of the radiative transfer equation, assuming a plane-parallel Milne-Eddington stellar atmosphere (Unno 1956). Each local line features a continuum level that is weighted

Downloaded from https://academic.oup.com/mnras/article/488/4/5114/5532370 by guest on 28 May 2023

Table 2. Key parameters of the injected RV signatures of the seven TRAPPIST-1 planets listed in a decreasing order of RV semi-amplitude. The orbital periods are taken from D18. The planetary masses are the best estimates from G18 and the RV semi-amplitudes are derived using Kepler's third law. The orbital phases are computed from the mid-transit times reported in D18 (T_0 – 2,450,000 [BJD_{TDB}]) assuming circular orbits. The bottom row shows the planet orbital phase shifts roughly corresponding to the maximum TTV predicted by G18 models (see fig. 2 therein).

Planet		b	c	g	f	e	d	h
Orbital period	(d)	1.51087637	2.42180746	12.354473	9.205585	6.099043	4.049959	18.767953
Mass	($M_\oplus$)	1.017	1.156	1.148	0.934	0.772	0.297	0.331
RV semi-amplitude	($m\,s^{-1}$)	2.838	2.757	1.590	1.427	1.353	0.597	0.399
Initial phase	(rad)	– 1.80	0.45	– 1.27	– 0.57	1.60	2.32	– 0.18
Maximum phase-shift predicted from G18	(rad)	0.006	0.007	0.018	0.028	0.021	0.011	0.006

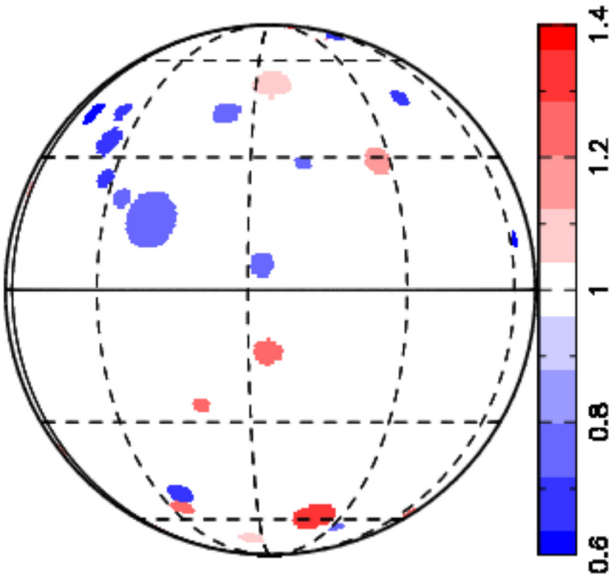


Figure 1. Illustration of the modelled stellar surface at a given time. The colour of the inhomogeneities indicates their relative brightness: 1 for the quiet photosphere (white), >1 for a bright feature (red), and <1 for a dark feature (blue).

by limb darkening; moreover, it is Doppler-shifted to the projected rotational velocity. The summation of all the local line profiles from the different cells weighted by the local brightness factors provides a global line profile that may be regarded as the cross-correlated function of the spectrum. We then derive the RV at the corresponding phase by fitting a Gaussian to the global profile. The relative photon flux is computed by summing the limb-darkened brightness factors of all cells. Our approach for synthesizing the stellar activity RV signal slightly differs from the one used in Boisse, Bonfils & Santos (2012) in that we can more easily include multiscale magnetic fields and differential rotation in the stellar model, as well as the Zeeman broadening and polarized Zeeman signatures of the modelled line profiles, found to be reliable proxies for diagnosing and modelling stellar magnetic activity (Haywood et al. 2016; Hébrard et al. 2016).

The synthetic star is assumed to rotate as a solid-body with a period of 3.3 d (Luger et al. 2017). The time-step for the generation of the RV and photometric data points is set at 0.01 d. Our model allows for a dynamical evolution of the active regions at the stellar surface. At each time-step, an active region is added to the model with a given probability. The position of the feature is randomly chosen at the stellar surface as active regions are likely to be found at all latitudes on late M dwarfs (Barnes et al. 2001; Barnes & Collier Cameron 2001) and, in particular, on TRAPPIST-1 (D18; Ducrot et al. 2018; Morris et al. 2018b). Consistently with Morris et al. (2018a), each newly created feature is randomly chosen to be either brighter or darker than the stellar photosphere. Its relative brightness factor with respect to the stellar photosphere is scaled from Berdyugina (2005, $\Delta T = \pm 100$ K) using Plank's law in the *K2* spectral band. The injected inhomogeneity is circular on the line of sight with a relative size S_r with respect to the whole stellar surface. The radius of each new active region is initially set to zero, increases linearly during the first third of its lifetime and decreases linearly during its remaining lifetime. After that, the feature is removed from the star. The radius of the feature at the time of its maximum size is drawn from a normal distribution centred at 0.1 per cent to ensure small active regions on the stellar surface in agreement with Rackham, Apai & Giampapa (2018) and Morris et al. (2018b). Surface features are allowed to overlap each other. As a result, our modelled star is populated by a patchwork of dark and bright small features as illustrated in Fig. 1 at any given time.

We simultaneously synthesized stellar photometric and RV curves over a 400-d period using the aforementioned time-step. We removed a 100-d burn-in period to ensure that the surface covered by active regions is more or less stable in time with the appearance of new feature roughly compensating the disappearance of old ones. We tuned both the lifetime and probability of appearance of the injected features so the simulated light curve scales with that obtained from the *K2* mission. The main attributes of the injected features on the stellar surface are listed in Table 3.

The synthetic photometric and RV activity curves are shown in Fig. 2.[2] We obtain a stellar activity signal with an amplitude ranging from ∼5 $m\,s^{-1}$ to almost 10 $m\,s^{-1}$ depending on the epoch. The lower two panels, respectively, show the mean size and surface coverage of the bright and dark features in our model. The average feature relative area, $S_r \sim 0.06$ per cent of the total stellar area, is slightly larger than the features modelled in Rackham et al. (2018) and lies close to threshold of undetected features in Morris et al. (2018b). The autocorrelation functions (ACFs) of the noiseless photometric and RV curves are shown in Fig. 3. Both time-series show 3.3-d quasi-periodic shapes symptomatic of short-lived active regions with half-life of ∼2 P_{rot}. The ACF peak at a difference of one P_{rot} is damped, though less than

[2]The modelling of the stellar activity RV curve presented here ignores both the smaller brightness contrast of surface features at nIR wavelengths on the one hand, and the impact of the magnetic fields that these features likely host on the other hand; preliminary simulations indicate that, whereas the first effect decreases the amplitude of the stellar activity RV curve, the second one amplifies it, with both effects more or less cancelling each other for magnetic field strengths of ∼2 kG. We thus conclude that despite these simplifications, our simulation is able to predict the shape and amplitude of the stellar activity RV curve within better than a factor of 2.

Downloaded from https://academic.oup.com/mnras/article/488/4/5114/5532370 by guest on 28 May 2023

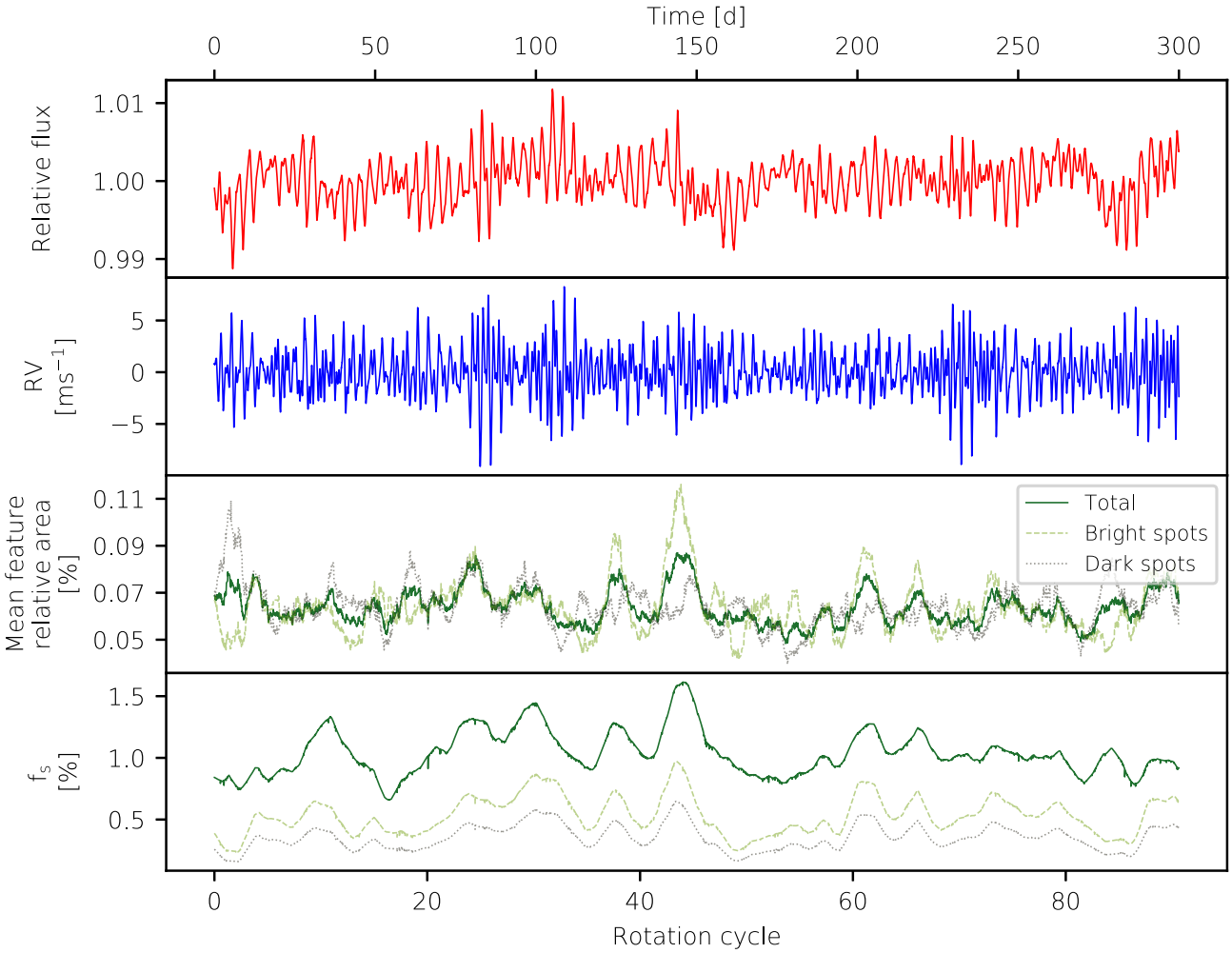


Figure 2. Synthetic RV and light curves of TRAPPIST-1. *Top:* relative brightness curve with respect to the quiet photosphere; *second*: stellar activity RV curve; *third:* mean feature relative area with respect to the visible disc; *bottom*: f_s accounting for total surface covered by bright and/or dark features with respect to the whole stellar surface. In the two lower panels, the quantities relative to dark features are plotted in grey dotted lines, those relative to bright features are plotted is light-green dashed lines and the dark-green solid lines refer to both dark and bright features.

Table 3. Parameters of the features generated at the stellar surface for synthesizing stellar photometric and RV curves. The Jeffreys's distribution is defined in Gregory (2007).

Parameter	Value	Unit
Probability of new feature at each time-step	0.04	–
Feature lifetime	$\mathcal{N}(5.0, 0.5)$	($P_{\rm rot}$)
Relative brightness dark spot	Jeffreys(0.6,0.8)	–
Relative brightness bright spot	Jeffreys(1.1,1.3)	–
Latitude of the feature	$\mathcal{U}(-90, 90)$	(deg)
$\log_{10}$ maximum relative size	$\mathcal{N}(-3.0, 0.3)$	–
Model for spot size evolution	Linear	–

the one in the original *K2* light curve (e.g. fig. 6 in Morris et al. 2018a), partly as a result of the noise not being included yet.

2.3 Scheduling of the observations

The synthetic TRAPPIST-1 RV curve (i.e. planet and stellar activity RV signals) is shown in Fig. 4. The lower two panels, respectively, display the stellar activity and planet RV curves whose levels of fluctuation are similar. We simulated TRAPPIST-1 RV follow-ups using the various sampling schemes listed in Table 4. Each sampling strategy provides a set of epochs at which we select the data points from the densely sampled RV curve containing both the planetary and stellar RV curves. All data sets are built from the same 90-d visibility window to ensure that all simulations feature the same realization of the stellar activity signal.

More specifically, sampling scheme A leads to an evenly sampled data set with a sampling frequency of 2 d^{-1}, i.e. greater than the Nyquist frequency of the planet signal. Schemes B and C assume that TRAPPIST-1 observations are carried out from CFHT and TBL, when the star is visible at an airmass of $\lesssim$1.5. To account for stochastic weather conditions, the observations are achieved with a probability of 0.85 and 0.5 at CFHT and TBL, respectively. Scheme C restricts CFHT observations to bright time periods ($\sim$15 d centred on full-Moon periods), whereas scheme B allows observations in bright and dark time periods, making scheme C more consistent with typical observational constraints at CFHT. Finally, scheme D is similar to scheme B except that we assume that the TBL is not available, leaving us with only one telescope for following-up TRAPPIST-1, this time with no constrain on the moon phase.

Downloaded from https://academic.oup.com/mnras/article/488/4/5114/5532370 by guest on 28 May 2023

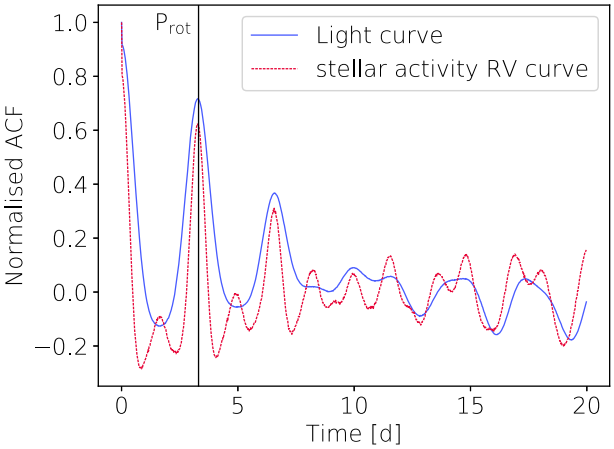


Figure 3. Normalized ACFs of the noise-free stellar photometric (blue solid line) and RV curves (red dashed line). Trappist-1 rotation period is indicated through the black vertical line.

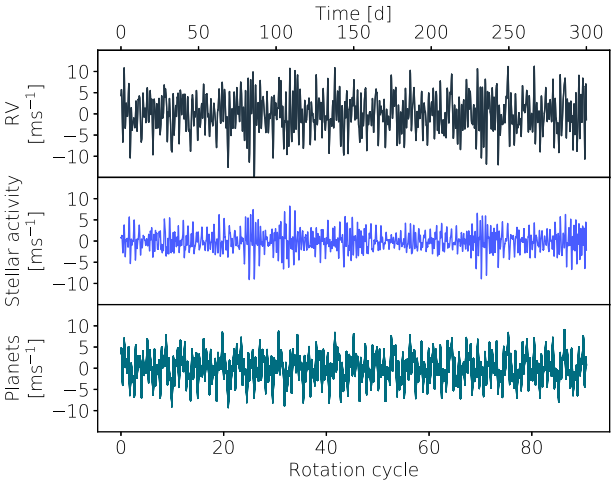


Figure 4. *Top:* Synthetic TRAPPIST-1 RV curve containing both stellar activity signal and planetary signatures. *Middle:* stellar activity RV signal shown in Fig. 2. *Bottom:* Planetary signature containing the RV signal of the seven planets.

Table 4. List of the observational sampling strategies considered in this study. Columns 1–3 show, respectively, the name of the configuration, the size of the data set and the sampling frequency. Additional details on the selection scheme are added in column (4).

Scheme	$N_{\rm pts}$	$f_{\rm s}$ (d^{-1})	Comments
A	180	2	Even sampling
B	120	∼2	CFHT (bright and dark times) and TBL
C	86	∼2	CFHT (bright time only) and TBL
D	76	∼1	CFHT only (bright and dark times)

2.4 White noise

We account for the instrumental and photon noises by adding random values drawn from a Gaussian distribution to each newly created data set. For each sampling scheme listed in Table 4, we consider random noise levels of 1 and 2 m s^{-1} rms in order to quantify the losses in the precision of the mass estimates when the noise level increases. At 1 m s^{-1}, we assume the photon noise to be the dominant contribution of the white noise, implying thus that other sources of noise such as the instrument itself or the correction for tellurics in the spectra produce a contribution much smaller than 1 m s^{-1} which is optimistic and less realistic than the 2 m s^{-1} case. In what follows, each sampling scheme from Table 4 will be indexed by the rms of the injected white noise expressed in m s^{-1}. For example, case A_1 designates a data set sampled with scheme A with a white noise of 1 m s^{-1} rms.

3 FITTING THE RV TIME-SERIES

We fit the synthetic RV time-series using the following model

$$V_r(t) = V_p(t) + V_j(t) + \epsilon(t), \tag{1}$$

where $V_p(t)$ and $V_j(t)$, respectively, account for the planetary signature and stellar activity signal and $\epsilon(t) \sim \mathcal{N}(0, \sigma^2(t))$, where the error on the data point at time t, $\sigma(t)$, is assumed to be known. Furthermore, we assume that systematics, quantified with simultaneous observations of RV standards with SPIRou, are already corrected.

3.1 Fit of the planetary signal only

As a sanity check, we first consider the ideal case of a data set containing only the planet signatures embedded in white noise, which can thus be modelled according to $V_r(t) = V_p(t) + \epsilon(t)$. We search for N_p planets with circular orbits assuming no planet–planet interactions (see the justifications in Section 4). Therefore, we fit the RV planet signatures using a linear combination of sinusoids such that

$$V_p(t) = \sum_{n=1}^{N_p} \alpha_n \cos\left(\frac{2\pi}{P_n}t\right) + \beta_n \sin\left(\frac{2\pi}{P_n}t\right), \tag{2}$$

where $\alpha_n = K_n \cos\phi_n$ and $\beta_n = -K_n \sin\phi_n$, with K_n and ϕ_n the semi-amplitude and orbital phase of the nth planet. The planet orbital periods, P_n, are frozen to their best estimates from D18. This results in a linear $2N_p$-parameter model $\boldsymbol{V_p} = \mathbf{X}\boldsymbol{\omega}$, where $\boldsymbol{\omega} = (\alpha_1, \beta_1, .., \alpha_{N_p}, \beta_{N_p})$ and $\mathbf{X}$ is a $(N, 2N_p)$ array, N being the number of data points. Assuming a Gaussian white noise, the parameter posterior density function, $p(\boldsymbol{\omega}|\boldsymbol{V_r})$, can be analytically derived, leading to

$$p(\boldsymbol{\omega}|\boldsymbol{V_r}) = \mathrm{N}\left(\mathbf{A}^{-1}\boldsymbol{b}, \mathbf{A}^{-1}\right) \tag{3}$$

where

$$\begin{cases} \mathbf{A} = \mathbf{X}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathbf{X} \\ \boldsymbol{b} = \mathbf{X}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{V_r} \end{cases} \tag{4}$$

where the covariance matrix of the white noise, $\boldsymbol{\Sigma}$, is such that $\Sigma_{ij} = \sigma(t_i)^2\delta_{ij}$, δ being the Kronecker delta.

We used equation (3) to derive maximum a posteriori estimates as well as 1σ uncertainties on the planet masses and orbital phases for the four sampling schemes listed in Table 4 at white noise levels of 1 and 2 m s^{-1} rms. In each case, the estimation is carried out on 50 signals with different white noise realizations from which we average the best planet parameters and error bars.

In this ideal test case, we are able to recover the masses at a $>10\sigma$-precision level for planets b, c, g, f, and e for all sampling schemes and white noise levels. The precision of the mass estimates for planets d and h systematically lies above 5σ except for cases C_2 and D_2 for which the mass of planet h is estimated at a precision $\gtrsim 3\sigma$.

Downloaded from https://academic.oup.com/mnras/article/488/4/5114/5532370 by guest on 28 May 2023

Table 5. Prior density probabilities used for the MCMC sampling for the interpolation of the stellar activity.

Parameter	Prior
$\ln\theta_1$ [ln(m s^{-1})]	$\mathcal{U}(-1, 10)$
$\ln\theta_2$ [ln(d)]	$\mathcal{U}(0, 5)$
θ_3 [$P_{\rm rot}$]	$\mathcal{U}(0.9, 1.1)$
θ_4	$\mathcal{U}(0.1, 4.0)$

3.2 Fitting the stellar activity RV signal

We now fit a signal that contains no planet, but only the stellar activity signal and white noise. In this purpose, we use Gaussian Process Regression (GPR; Rasmussen & Williams 2006) to model the stellar activity RV signal. The shape of the ACF of the synthetic data set (see Fig. 3) suggests that a quasi-periodic covariance kernel will be adapted to reconstruct the synthetic stellar activity RV curve (Haywood et al. 2014, hereafter H14):

$$k(t_i, t_j) = \theta_1^2 \exp\left[-\frac{(t_i - t_j)^2}{\theta_2^2} - \frac{\sin^2\frac{\pi(t_i - t_j)}{\theta_3}}{\theta_4^2}\right], \quad (5)$$

where t_i and t_j are the times associated with observations i and j. This covariance function depends on four hyperparameters called θ_1 to θ_4; θ_1 is the amplitude of the Gaussian Process (GP) that scales with the amplitude of the stellar activity signal, θ_2 is the time-scale for the evolution of the active regions, θ_3 is the period of the GP, whereas θ_4 is called the smoothing parameter that controls the number of high-frequency structures that can be included into the fit. In what follows, the vector containing the four hyperparameters is called $\boldsymbol{\theta}$.

We used GPR to independently model photometric and RV time-series of the stellar activity signal evenly sampled with scheme A. White noises of 1 m s^{-1} and 1 mmag rms were added, respectively, to the RV and photometric time-series so that the two data sets have similar signal-to-noise ratios. The 4 GP hyperparameters are jointly estimated by maximizing the posterior density function of the hyperparameters, $p(\boldsymbol{\theta}|\boldsymbol{V_r})$, sampled using a Bayesian Markov Chain Monte Carlo (MCMC). We adopted the non-informative priors listed in Table 5 to ensure a physically realistic fit of the stellar activity signal. The period of the GP is expected to lie close to the stellar rotation period (Angus et al. 2018), explaining thus the narrow prior distribution adopted for θ_3. The likelihood of the GP fit to the stellar activity RV signal, $\mathcal{L} = p(\boldsymbol{V_r}|\boldsymbol{\theta})$, is such that

$$2\ln\mathcal{L} = -N\ln 2\pi - \ln(|\mathbf{K} + \boldsymbol{\Sigma}|) - \boldsymbol{V_r}^T(\mathbf{K} + \boldsymbol{\Sigma})^{-1}\boldsymbol{V_r} \quad (6)$$

where $\mathbf{K}$ is the GP covariance matrix. Concretely, we use the EMCEE affine invariant sampler (Foreman-Mackey et al. 2013, 100 walkers, 2000 iterations) to sample the posterior density. After removing a burn-in period of ∼10 autocorrelation times from the chain, we compute the median and uncertainties on each parameter from the posterior distribution.[3]

This process is repeated for 50 photometric and RV time-series with different white noise realizations, resulting in residuals of about 0.6 m s^{-1} and 0.8 mmag rms for RV and photometric time-series, respectively. Table 6 shows the average over the 50 time-series of the best and the median values of the hyperparameters. We find similar GP periods and time-scales for both photometric and RV time-series, consistent with the injected rotation period and the one derived from the ACF of Fig. 3. We note that θ_4 is roughly twice as large in photometry as in RV, reflecting the fact that a given feature at the stellar surface produces a signature in the RV curve that evolves twice as fast as its counterpart in the light curve. The interpolation of photometric and RV time-series is shown in Fig. 5 for a given realization of white noise.

Following the approach of Aigrain, Pont & Zucker (2012), we computed the first derivative of the synthetic noise-free light curve and found a linear correlation coefficient of −0.89 with the stellar activity RV curve, indicating a net anticorrelation between the two signals. We fitted the evenly sampled stellar activity RV curve from Fig. 5 with the first derivative of the stellar light curve, evaluated at the same time values, using a linear model. This resulted in residuals of 1.4 m s^{-1} rms, i.e. larger than that obtained when the stellar activity signal is modelled with GP (0.56 m s^{-1} rms).

3.3 Simultaneous fit of stellar activity RV signal and planetary curves

We now model the synthetic data sets described in Section 2 using equation (1), including all model components (planet, stellar activity, and white noise). The goal is to simultaneously filter the stellar activity signal with a GP directly trained in the data set and estimate the planetary semi-amplitudes and orbital phases. This time, our model contains $2N_p$ linear planet parameters and four non-linear GP hyperparameters. We use a Bayesian MCMC to sample the posterior density marginalized over the planet parameters:

$$p(\boldsymbol{\theta}|\boldsymbol{V_r}) \propto \pi(\boldsymbol{\theta}) \int_{\boldsymbol{\omega}} p(\boldsymbol{V_r}|\boldsymbol{\theta}, \boldsymbol{\omega})\pi(\boldsymbol{\omega})\mathrm{d}\boldsymbol{\omega}, \quad (7)$$

where $\boldsymbol{\theta}$ and $\boldsymbol{\omega}$ are assumed independent, and where $\pi(\boldsymbol{\theta})$ and $\pi(\boldsymbol{\omega})$ stand for the prior density on $\boldsymbol{\theta}$ (listed in Table 5) and the linearized planetary amplitudes, respectively. We analytically derive equation (7), assuming an infinitely broad Gaussian prior density for $\boldsymbol{\omega}$ that is similar as a non-informative uniform prior density [see section 2.7.1 in Rasmussen & Williams (2006) for a detailed generic demonstration and expression for the likelihood]. After removing a burn-in period of a few autocorrelation times from the chain, we compute the set of GP hyperparameters, $\boldsymbol{\theta}_{\rm best}$, that maximizes the posterior density as well as the median and error bars on the hyperparameters from the posterior samples. At a given posterior sample $\boldsymbol{\theta}$, the posterior distribution for the planet parameters is such that

$$p(\boldsymbol{\omega}|\boldsymbol{V_r}, \boldsymbol{\theta}) = \mathcal{N}\left(\mathbf{A}(\boldsymbol{\theta})^{-1}\boldsymbol{b}(\boldsymbol{\theta}), \mathbf{A}(\boldsymbol{\theta})^{-1}\right), \quad (8)$$

where

$$\begin{cases} \mathbf{A}(\boldsymbol{\theta}) = \mathbf{X}^T[\boldsymbol{\Sigma} + \mathbf{K}(\boldsymbol{\theta})]^{-1}\mathbf{X} \\ \boldsymbol{b}(\boldsymbol{\theta}) = \mathbf{X}^T[\boldsymbol{\Sigma} + \mathbf{K}(\boldsymbol{\theta})]^{-1}\boldsymbol{V_r} \end{cases}. \quad (9)$$

We use equation (8) at $\boldsymbol{\theta} = \boldsymbol{\theta}_{\rm best}$ to estimate the planet parameters as well as their 1σ error bars. Note that these error bars are significantly larger than the dispersion of the best planet parameters obtained by computing equation (8) for all the samples of the joint posterior distribution. The planetary masses and their error bars are finally derived from the estimates of the semi-amplitudes using Kepler's third law. We stress that this procedure is in principle applicable to any planetary system with circular orbits of known periods (and phases).

[3]The 1σ error bars are given by $0.5(x_2 - x_1)$, where x_1 and x_2, respectively, correspond to the 16th and 84th percentiles of the posterior distributions of each parameter.

Downloaded from https://academic.oup.com/mnras/article/488/4/5114/5532370 by guest on 28 May 2023

Table 6. Best and median estimates of the four hyperparameters of the GP when fitting the light and stellar activity RV curves without any planetary component. Note that the values indicated in this table are averaged over 50 data sets with different white noise realizations of 1 m s^{-1} and 1 mmag for the RV and photometric curves, respectively.

Parameter	Light curve		RV	
	Best	Median	Best	Median
θ_1	0.33 per cent	0.34 ± 0.05 per cent	2.6 m s^{-1}	2.6 ± 0.3 m s^{-1}
θ_2 (d)	7.2	7.2 ± 0.8	7.9	7.8 ± 0.8
θ_3 (d)	3.30	3.30 ± 0.04	3.30	3.30 ± 0.02
θ_4	0.65	0.67 ± 0.09	0.37	0.37 ± 0.03

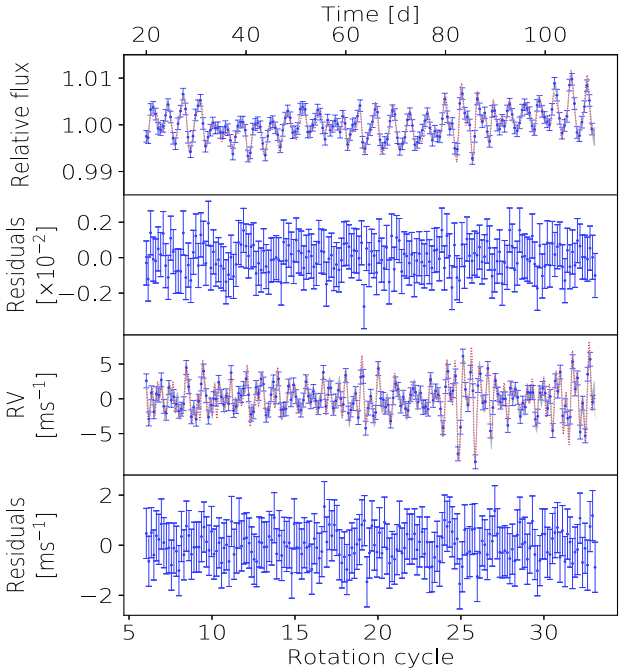


Figure 5. From top to bottom: Photometric time-series, residuals after the GP interpolation of the photometric time-series, RV time-series and residuals after the GPR. The red dotted-lines are the injected signals and the blue dots with error bars are the GP prediction at the observation times. The grey error bands are the $\pm 1\sigma$ predictions of the GP.

3.4 Model comparison and evidence estimation

Which of the TRAPPIST-1 planets are detected with the algorithm described above? Although apparently intuitive, this issue requires the use of a robust criterion to be addressed in a reliable way. In our case, we carried out a Bayesian comparison of models searching, respectively, for N_p and $N_p + 1$ planets, where N_p varies from 0 to 6. Assuming that we have no prior information on the number of searched planets, two models are compared by computing the so-called Bayes factor, i.e. the ratio of the Marginal Likelihoods (MLs) of the two models. The ML being mathematically intractable in our case, we used the method of Chib & Jeliazkov (2001, hereafter CJ01) to approximate its value from the posterior density obtained when sampling the parameter space of each model (see also H14 for more details on the implementation of the method).

For each data set, we compute the ML of models searching for 0–7 planets, where the searched planets are added by decreasing order of RV semi-amplitude (i.e. TRAPPIST-1 b, c, g, f, e, d, and h). We then derive the Bayes factor comparing every model containing at least one planet with the model searching for one fewer planet. According to Jeffreys (1961), a Bayes factor greater than 150 (∼5 in log) is associated with a fair detection of the planet added to the model. The evidence in favour of the planet will be regarded as strong if the Bayes factor lies within 150–20 (∼5–3 in log), positive if it falls within the range 20–3 (∼3–1 in log) and inconclusive otherwise.

4 RESULTS

4.1 Fitting simulated data

We used the model described in Section 3 to recover TRAPPIST-1 planet masses for the sampling schemes listed in Table 4 and for white noise levels of 1 and 2 m s^{-1} rms. Our first goal being to assess the effects of the stellar activity signal and sampling strategies on the quality of the mass estimates, we start by systematically searching for seven planets in the data sets. For each sampling scheme and white noise level, we fit 50 time-series, each with different white noise realisations of equal rms. The best estimates of the planet masses and GP hyperparameters with error bars are obtained by taking the median of the distributions. The outcome of the MCMC samplings and the best estimates for the planet masses are shown in Tables 7 and 8, respectively.

Case A_1, i.e. evenly sampled data with 1 m s^{-1} white noise, provides the best fit of the signal with an average reduced χ^2, χ^2_r (i.e. χ^2/N, where N is the number of data points), of ∼0.3. We find hyperparameters consistent with those obtained when fitting the stellar activity signal alone (see Table 6), indicating a good interpolation of the stellar activity signal by the GP. Compared to the mass estimates obtained with the TTVs, our mass estimates are significantly more precise for planets b and c, and show comparable precision for planets g, f, and e. However, planets d and h whose masses were derived with $>10\sigma$ precision in the absence of stellar activity in case A_1, now show largely degraded mass estimates, with a precision level below 3σ. The error bars on the planet orbital phase estimates are ∼3 to ∼140 times larger than the phase-shift induced by the maximum TTVs predicted from G18 models (see Table 2), justifying thus to have ignored the dynamical interactions between the planets when modelling the data. An example of the fit to our A_1 data set is shown in Fig. 6, where we displayed, in particular, the reconstructed activity and planetary RV curves as well as the residuals of the fit to the synthetic data. The corresponding posterior distribution of the four GP hyperparameters is shown in Fig. 7. In the Appendix, we show examples of reconstructed RV times series as well as posterior densities for all sampling schemes with 1 and 2 m s^{-1} rms white noise.

When we increase the white noise level from 1 to 2 m s^{-1} (i.e. case A_2), we note that the error bars on the hyperparameter estimates are roughly doubled except for the amplitude of the GP whose precision is only slightly degraded. Compared to case A_1, the error bars on the mass estimates of planets c, f, and e are multiplied by

Downloaded from https://academic.oup.com/mnras/article/488/4/5114/5532370 by guest on 28 May 2023

Table 7. Results of the interpolation of the synthetic data sets for the sampling schemes listed in Table 4 and for white noise levels of 1 and 2 m s^{-1}. Column 1 indicates the sampling case indexed by the rms of the injected white noise expressed in m s^{-1}, column 2 is the number of searched planets and columns 3 to 6 are the medians and 1σ error bars of the GP hyperparameters averaged over ∼50 input data sets with different realizations of white noise. We show the logarithm of the maximum likelihood, ML and Bayes factor for each model compared to the model with one less planet in columns 7–9. The median reduced χ^2 of the fit, χ^2_r (i.e. χ^2/N, where N is the number of data points), is shown in column 10. The hyperparameters listed without error bars were frozen to the indicated value during the MCMC sampling to ensure the convergence of the algorithm.

Case	N_p	θ_1 (ms^{-1})	θ_2 (d)	θ_3 (d)	θ_4	ln($\mathcal{L}_{\max}$)	ln(ML)	ln Bayes factor (N_p;N_p-1)	χ^2_r
A_1	3	2.94 ± 0.22	23.51 ± 12.20	3.25 ± 0.04	0.10 ± 0.07	−439.7 ± 3.2	−439.1 ± 3.9	–	0.1
	4	2.77 ± 0.23	7.98 ± 0.74	3.27 ± 0.01	0.24 ± 0.02	−421.3 ± 3.5	−421.8 ± 3.4	17.6 ± 3.2	0.2
	5	2.69 ± 0.27	7.53 ± 0.76	3.30 ± 0.02	0.38 ± 0.04	−393.9 ± 4.1	−394.6 ± 4.1	27.0 ± 2.1	0.3
	6	2.64 ± 0.26	7.70 ± 0.80	3.30 ± 0.02	0.37 ± 0.04	−391.8 ± 4.1	−392.6 ± 4.1	2.2 ± 0.8	0.3
	7	2.64 ± 0.27	7.70 ± 0.81	3.30 ± 0.02	0.37 ± 0.04	−391.5 ± 4.1	−392.2 ± 4.1	0.5 ± 0.4	0.3
A_2	3	2.89 ± 0.27	20.16 ± 12.60	3.26 ± 0.09	0.12 ± 0.09	−469.1 ± 5.2	−469.6 ± 5.6	–	0.4
	4	2.73 ± 0.28	9.04 ± 2.79	3.28 ± 0.05	0.23 ± 0.06	−459.0 ± 5.5	−459.8 ± 5.7	10.3 ± 3.0	0.4
	5	2.61 ± 0.30	7.64 ± 1.74	3.29 ± 0.05	0.34 ± 0.06	−445.9 ± 4.9	−446.8 ± 4.9	13.0 ± 2.8	0.5
	6	2.58 ± 0.31	7.78 ± 1.77	3.28 ± 0.04	0.34 ± 0.06	−444.0 ± 4.8	−444.9 ± 4.8	2.2 ± 1.1	0.5
	7	2.59 ± 0.31	7.82 ± 1.80	3.28 ± 0.04	0.34 ± 0.06	−443.3 ± 4.6	−444.2 ± 4.6	1.0 ± 0.7	0.5
B_1	3	3.09 ± 0.27	9.13 ± 3.95	3.28 ± 0.06	0.16 ± 0.09	−298.7 ± 2.9	−299.1 ± 2.9	–	0.1
	4	2.82 ± 0.26	8.50 ± 1.84	3.26 ± 0.03	0.20 ± 0.04	−286.6 ± 3.0	−287.4 ± 3.0	11.8 ± 1.5	0.2
	5	2.76 ± 0.32	8.42 ± 1.17	3.29 ± 0.03	0.37 ± 0.05	−268.9 ± 3.6	−269.7 ± 3.6	18.0 ± 2.2	0.2
	6	2.76 ± 0.32	8.66 ± 1.24	3.28 ± 0.03	0.37 ± 0.05	−267.6 ± 3.4	−268.5 ± 3.4	1.4 ± 0.8	0.2
	7	2.73 ± 0.32	8.96 ± 1.35	3.28 ± 0.03	0.37 ± 0.05	−265.7 ± 3.5	−266.5 ± 3.5	2.0 ± 0.9	0.2
B_2	3	3.08 ± 0.34	11.15 ± 6.42	3.34 ± 0.11	0.16 ± 0.13	−316.2 ± 3.5	−317.0 ± 3.4	–	0.3
	4	2.83 ± 0.34	8.59 ± 3.36	3.26 ± 0.08	0.23 ± 0.09	−307.2 ± 3.3	−308.2 ± 3.3	8.9 ± 2.0	0.4
	5	2.74 ± 0.37	8.21 ± 2.22	3.27 ± 0.06	0.35 ± 0.08	−298.3 ± 3.1	−299.3 ± 3.2	9.1 ± 2.5	0.4
	6	2.74 ± 0.38	8.53 ± 2.35	3.27 ± 0.05	0.35 ± 0.08	−297.0 ± 3.0	−297.9 ± 3.0	1.6 ± 0.9	0.4
	7	2.72 ± 0.38	8.81 ± 2.53	3.27 ± 0.05	0.35 ± 0.08	−295.3 ± 3.0	−296.2 ± 3.0	1.9 ± 0.9	0.4
C_1	3	3.32 ± 0.33	9.39 ± 4.42	3.29 ± 0.04	0.16 ± 0.09	−220.3 ± 2.1	−220.8 ± 2.1	–	0.1
	4	3.16 ± 0.34	9.01 ± 3.17	3.28 ± 0.03	0.22 ± 0.06	−213.1 ± 2.4	−213.9 ± 2.4	6.9 ± 1.0	0.1
	5	2.96 ± 0.36	8.10 ± 1.91	3.31 ± 0.04	0.36 ± 0.06	−201.6 ± 2.2	−202.3 ± 2.2	11.8 ± 1.5	0.1
	6	2.90 ± 0.36	8.86 ± 2.22	3.30 ± 0.04	0.36 ± 0.06	−198.6 ± 2.3	−199.4 ± 2.2	3.0 ± 0.8	0.2
	7	2.90 ± 0.36	8.74 ± 2.20	3.29 ± 0.04	0.36 ± 0.06	−197.8 ± 2.1	−198.5 ± 2.1	0.9 ± 0.4	0.2
C_2	3	3.35 ± 0.42	11.94 ± 7.15	3.29 ± 0.11	0.17 ± 0.10	−230.2 ± 2.5	−231.1 ± 2.4	–	0.3
	4	3.18 ± 0.43	10.35 ± 5.48	3.28 ± 0.09	0.24 ± 0.09	−224.4 ± 2.0	−225.3 ± 2.0	5.9 ± 1.2	0.3
	5	2.91 ± 0.43	8.79 ± 3.90	3.29 ± 0.10	0.34 ± 0.10	−216.4 ± 2.8	−217.3 ± 2.9	8.0 ± 1.8	0.4
	6	2.86 ± 0.44	9.80 ± 4.59	3.26 ± 0.10	0.34 ± 0.10	−213.6 ± 2.5	−214.5 ± 2.5	3.0 ± 1.4	0.4
	7	2.88 ± 0.46	10.91 ± 5.51	3.24 ± 0.10	0.33 ± 0.11	−212.3 ± 2.7	−213.2 ± 2.8	1.4 ± 0.6	0.4
D_1	3	3.25 ± 0.34	7.9	3.43 ± 0.04	0.37	−190.1 ± 2.3	−191.2 ± 2.3	–	0.1
	4	3.01 ± 0.33	7.9	3.42 ± 0.04	0.37	−184.2 ± 2.3	−185.3 ± 2.3	5.9 ± 0.8	0.1
	5	2.59 ± 0.31	7.9	3.31 ± 0.05	0.37	−173.3 ± 2.5	−174.4 ± 2.5	10.9 ± 1.4	0.2
	6	2.60 ± 0.32	7.9	3.31 ± 0.05	0.37	−172.3 ± 2.5	−173.4 ± 2.5	1.2 ± 0.5	0.2
	7	2.47 ± 0.32	7.9	3.30 ± 0.05	0.37	−168.7 ± 2.7	−169.8 ± 2.7	3.7 ± 0.7	0.2
D_2	3	3.03 ± 0.48	7.9	3.43 ± 0.07	0.37	−199.0 ± 3.4	−200.1 ± 3.4	–	0.4
	4	2.81 ± 0.50	7.9	3.40 ± 0.09	0.37	−194.1 ± 3.7	−195.3 ± 3.7	5.1 ± 1.2	0.4
	5	2.52 ± 0.49	7.9	3.30 ± 0.11	0.37	−187.6 ± 3.3	−188.7 ± 3.3	6.9 ± 2.1	0.4
	6	2.56 ± 0.50	7.9	3.29 ± 0.11	0.37	−186.5 ± 3.2	−187.6 ± 3.1	1.1 ± 0.5	0.4
	7	2.41 ± 0.51	7.9	3.29 ± 0.12	0.37	−182.9 ± 3.3	−184.1 ± 3.3	3.6 ± 1.0	0.4

Table 8. Best TRAPPIST-1 planetary system mass estimates for data sets sampled with the schemes listed in Table 4 at a white noise of 1 or 2 m s^{-1} rms. In each case, the values are averaged over ∼50 RV time-series with different white noise realization. Case TTV stands for the best mass estimates from G18.

Case	m_b ($M_\oplus$)	m_c ($M_\oplus$)	m_g ($M_\oplus$)	m_f ($M_\oplus$)	m_e ($M_\oplus$)	m_d ($M_\oplus$)	m_h ($M_\oplus$)
TTV	$1.017^{+0.154}_{-0.143}$	$1.156^{+0.142}_{-0.131}$	$1.148^{+0.098}_{-0.049}$	$0.934^{+0.08}_{-0.078}$	$0.772^{+0.079}_{-0.075}$	$0.297^{+0.039}_{-0.035}$	$0.331^{+0.056}_{-0.049}$
A_1	0.988 ± 0.083	1.173 ± 0.049	1.130 ± 0.113	0.943 ± 0.077	0.762 ± 0.071	0.325 ± 0.133	0.332 ± 0.260
A_2	0.991 ± 0.109	1.194 ± 0.093	1.116 ± 0.175	0.907 ± 0.147	0.773 ± 0.131	0.331 ± 0.155	0.402 ± 0.292
B_1	1.040 ± 0.096	1.221 ± 0.079	1.127 ± 0.147	0.965 ± 0.124	0.837 ± 0.102	0.251 ± 0.129	0.565 ± 0.261
B_2	1.059 ± 0.137	1.214 ± 0.130	1.143 ± 0.237	1.026 ± 0.207	0.816 ± 0.174	0.289 ± 0.178	0.624 ± 0.338
C_1	1.190 ± 0.129	1.120 ± 0.134	1.329 ± 0.230	0.774 ± 0.195	1.066 ± 0.165	0.459 ± 0.183	0.441 ± 0.353
C_2	1.222 ± 0.178	1.105 ± 0.188	1.339 ± 0.316	0.918 ± 0.274	1.137 ± 0.236	0.549 ± 0.234	0.553 ± 0.428
D_1	1.078 ± 0.173	1.262 ± 0.189	1.646 ± 0.271	1.173 ± 0.232	0.940 ± 0.144	0.306 ± 0.174	0.909 ± 0.336
D_2	1.050 ± 0.201	1.285 ± 0.228	1.771 ± 0.330	1.138 ± 0.295	0.908 ± 0.221	0.291 ± 0.228	1.032 ± 0.412

Downloaded from https://academic.oup.com/mnras/article/488/4/5114/5532370 by guest on 28 May 2023

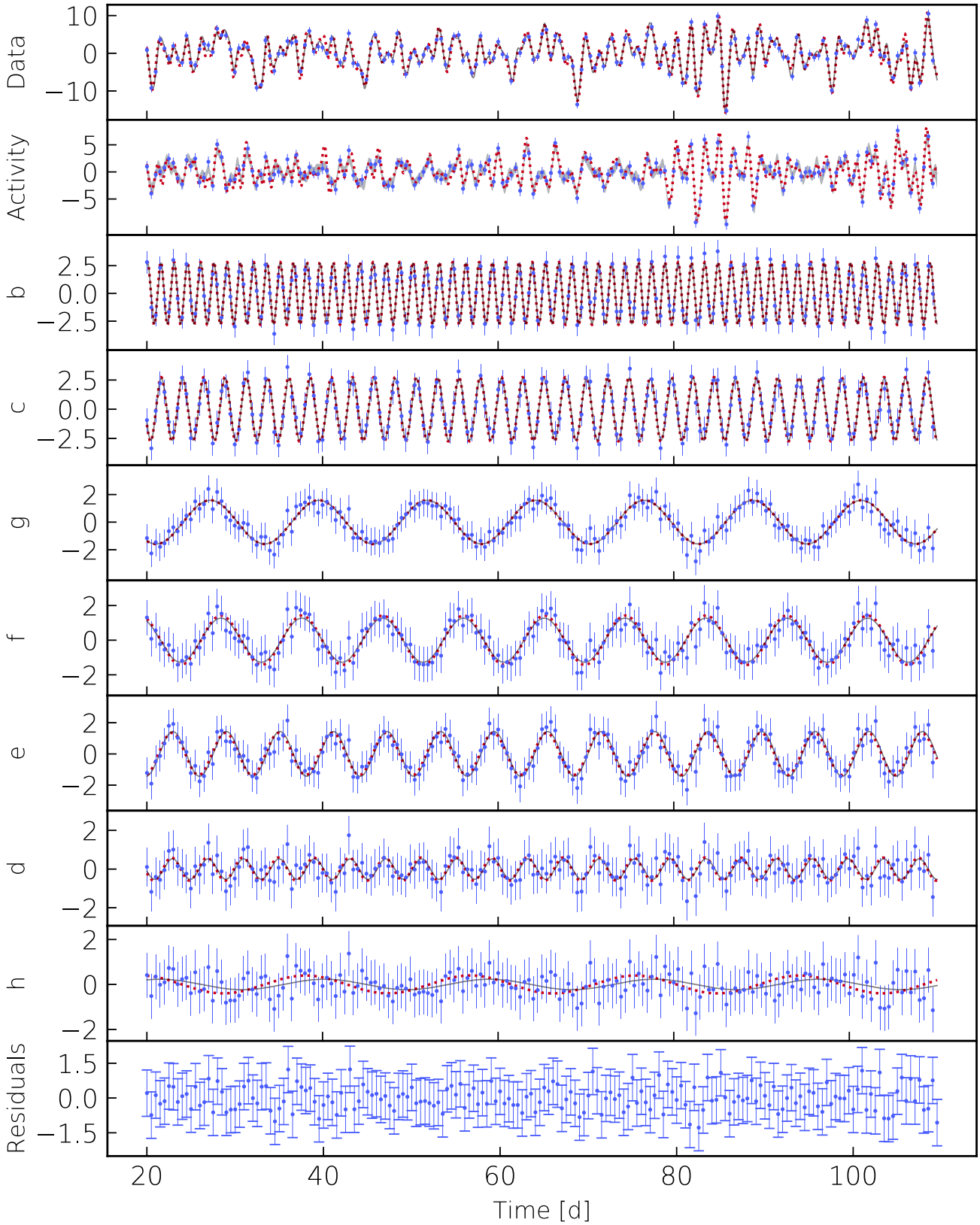


Figure 6. Fit of a data set sampled with A_1 scheme with a given realization of the 1 m s^{-1} white noise. All RVs are expressed in m s^{-1}. From top to bottom: raw data set, stellar activity RV signal, RV signatures from planets b to h in a decreasing order of semi-amplitude and residuals (rms of 0.5 m s^{-1}). In each panel, the red dotted line is the injected RV curve and the grey solid line is the best prediction from the model. For the stellar activity and planetary signals, the blue dots are obtained by subtracting all the components of the model except the one displayed in the data.

Downloaded from https://academic.oup.com/mnras/article/488/4/5114/5532370 by guest on 28 May 2023

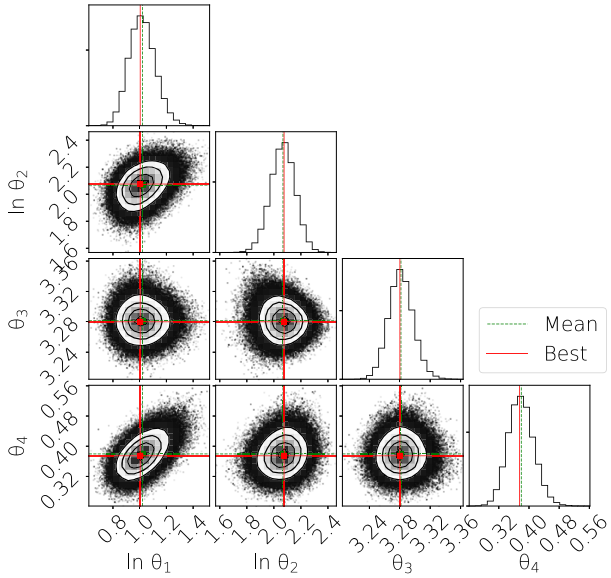


Figure 7. Posterior distribution of the four GP hyperparameters returned by the MCMC sampling for a given data set sampled with scheme A_1. The red solid lines indicate the best hyperparameter values, i.e. the values that maximize the likelihood of the MCMC. The green dashed lines indicate the median hyperparameter values over the posterior density. This plot was made using the CORNER PYTHON module (Foreman-Mackey 2016).

a factor of ∼2, implying that the main limitation comes from the white noise itself. In contrast, the precision on the mass estimates of planets b, d, and h is dominated by the stellar activity signal itself, as evidenced by the low increase in the error bars when switching from case A_1 to A_2. With a white noise of 2 m s^{-1}, only the mass estimates of planets b and c remain more precise that those derived from TTVs.

The fit to the RV time-series is noticeably degraded for sampling scheme D. In this case, the MCMC sampling of the parameter space does not converge anymore. The GP prefers to either minimize the smoothing factor θ_4 or shrink the time-scale θ_2 so that the correlation between consecutive data points becomes ridiculously small leading to a non-physical modelling of the stellar activity RV signal. This implies that the RV curve is no longer sampled at a high enough rate to allow for a reliable fit of the data with the four-hyperparameter GP we used.

When no stellar activity component is added to the data, data sets sampled on the D_1 scheme provide unbiased mass estimates with a precision of $>5\sigma$ for all TRAPPIST-1 planets, demonstrating that sampling D performs a satisfactory coverage of the planetary phase curves. This is also notable in Fig. 8 showing examples of orbit-folded RV curves of planet signals for sampling schemes A_1, C_1, and D_2, with activity taken into account and filtered out from the data as outlines in Section 3.3. Fig. 9 shows the histograms of the number of observations per stellar rotation cycle for each considered sampling scheme. Scheme D performs worst with only ∼3 data points per stellar rotation period. This is an insufficient coverage of the stellar rotation cycle before the spot pattern evolves significantly, leading to the non-convergence of the MCMC process. To address this issue, we have to freeze two of the 4 GP parameters, namely the time-scale and smoothing parameters, θ_2 and θ_4, to their best estimates when fitting the stellar activity signal only (see Table 6).

Fig. 10 shows the mass estimates as a function of the injected masses for all the considered cases. With scheme D, the masses of

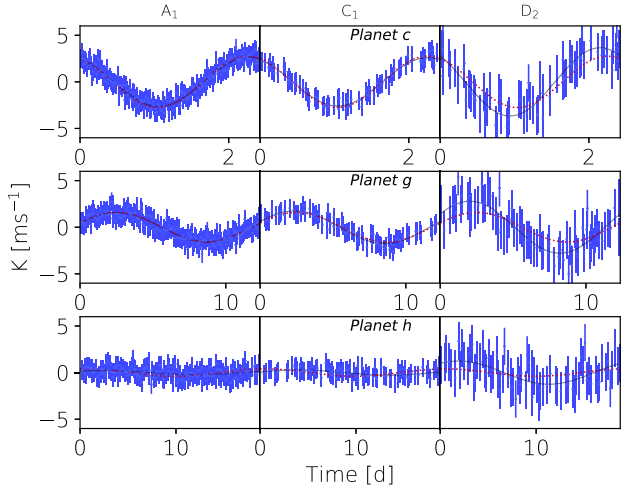


Figure 8. Orbit-folded recovered RV curves for TRAPPIST-1 planets c (top row), g (middle row), and h (bottom row), and sampling schemes A_1 (first column), C_1 (middle column), and D_2 (right column) when stellar activity is taken into account, each for a given realization of the white noise. For each planet, the X-axis covers one orbital period on which all data points are folded.

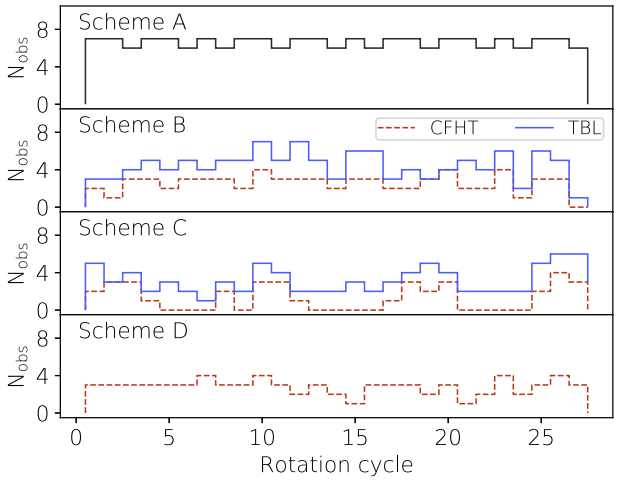
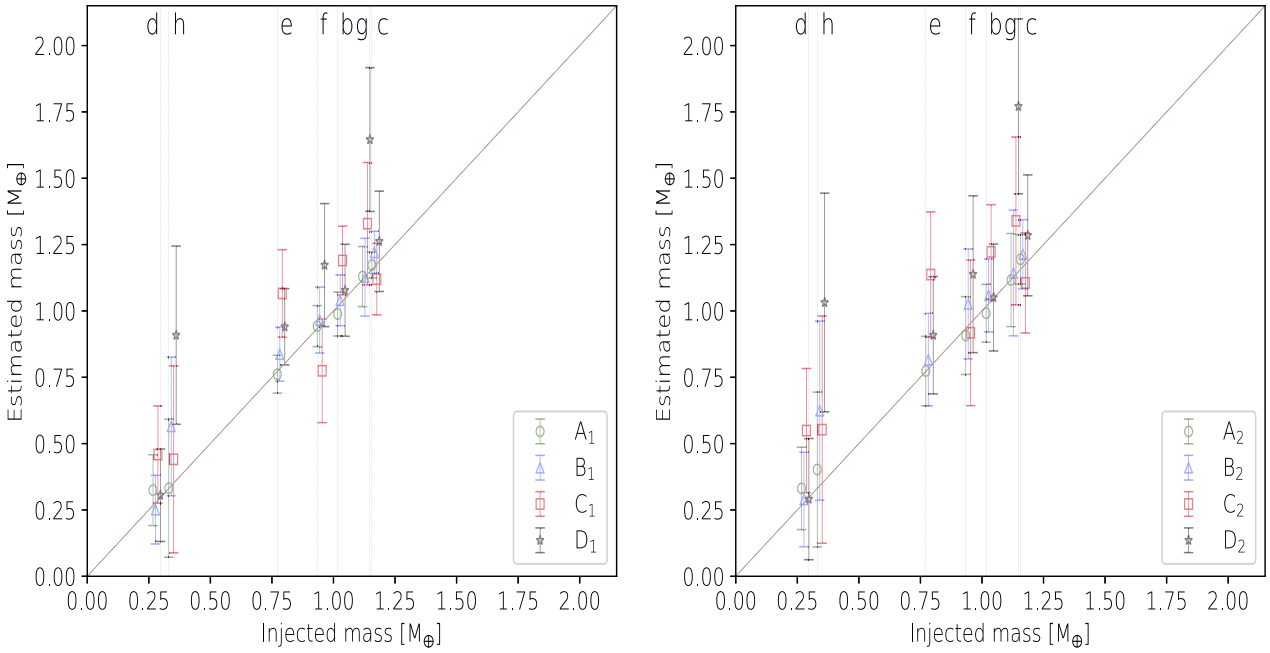


Figure 9. Histograms of the number of observations (N_{obs}) per stellar rotation cycle for each sampling scheme. When relevant, the observations are shown in dashed red and solid blue lines when from CFHT and TBL, respectively.

planets g, f, and h are strongly overestimated, regardless of the white noise level. In contrast, the amplitude of the GP is lower than when fitting the stellar activity signal alone, implying an underestimation of the latter (see Table 7). Moreover, the planet mass estimates are considerably less precise than those obtained from TTV except for planet b (see Table 8). In an attempt to improve the planet mass estimates while still using a single telescope, we doubled the number of observations per CFHT night. However, this configuration only marginally improves the precision on the mass estimates and the masses of planets g and h remain largely overestimated. This demonstrates that doubling the number of observations per night in this configuration still leads to an incomplete coverage of the stellar rotation cycle. The same goes for sampling strategies aimed at phases where the planetary signal is expected to be maximum.

In contrast with scheme D, the fit of data sets sampled with scheme B_1 provides remarkably good mass estimates with unbiased

Downloaded from https://academic.oup.com/mnras/article/488/4/5114/5532370 by guest on 28 May 2023



(a) 1 m s^{-1} rms white noise (b) 2 m s^{-1} rms white noise

Figure 10. Estimated planetary masses as a function of the injected masses for the four sampling cases listed in Table 4 at 1 m s^{-1} (left-hand panel) and 2 m s^{-1} (right-hand panel) rms white noise. For clarity purposes, we slightly shifted the mass estimates for the different sampling schemes to the left (for planets g and d) or to the right (remaining planets).

estimates at a precision $>8\sigma$ for planets b, c, g, f, and e. However, the mass estimates of planets d and h are measured at a precision $<3\sigma$ with an overestimation of the mass of planet h. Case B_2 leads to unbiased mass estimates at a precision of $\gtrsim 5\sigma$ for planets b, c, g, f, and e. Compared to sampling B, scheme C provides mass estimates degraded in precision due to a significant decrease in the number of observations. However, the mass estimates for planets g and f remain much less overestimated than in case D (see Fig. 8). For these planets in particular, the mass estimates are less degraded when moving from C_1 to C_2 than from C_2 to D_2, which implies that the observational sampling mainly explains the erroneous mass estimates recovered in case D_2.

We finally generated ∼40 data sets sampled in case C_1 with different realizations of white noise and modelled each of them using the procedure described in Section 3.3, but this time freezing the planet orbital phases to their best estimates from D18. For this sampling scheme, it turns out that freezing the planet orbital phases when modelling the data set has virtually no impact on the detectability of the planets and no more than marginally increases the precision on the planet masses (by about 2 per cent).

4.2 Planet detection

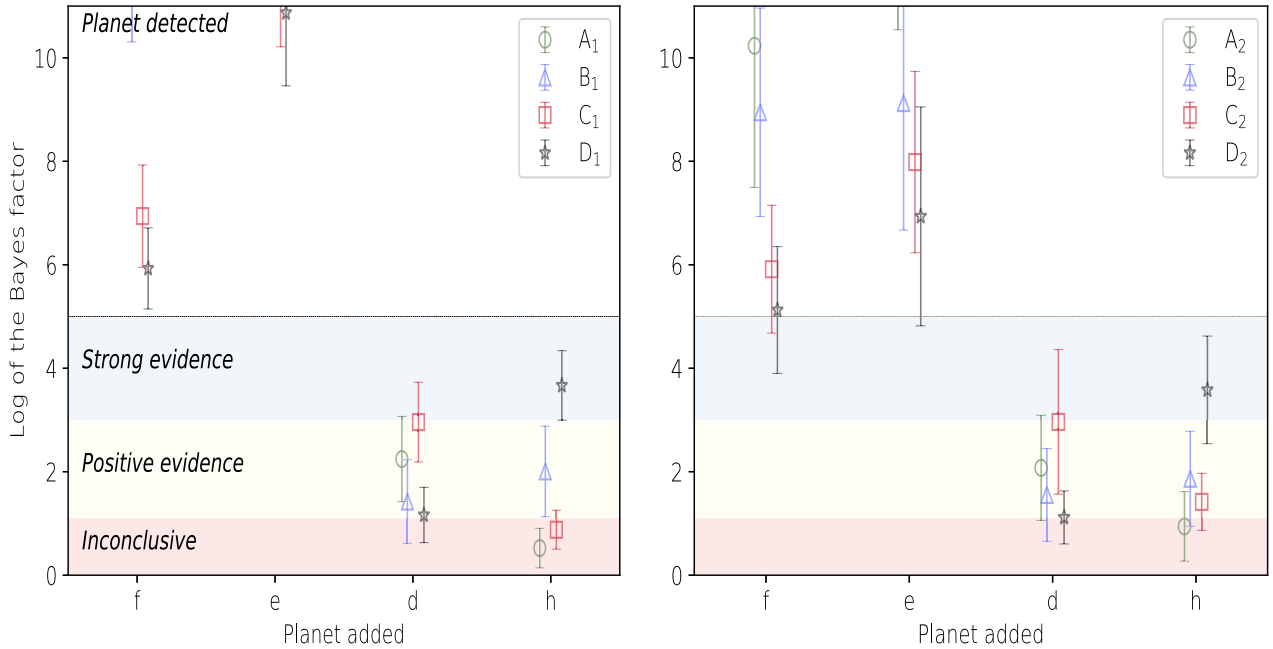
We computed the Bayes factors relative to each of the seven TRAPPIST-1 planets added in a decreasing order of RV semi-amplitude for data sets sampled with the schemes listed in Table 4 with white noise levels of 1 and 2 m s^{-1} rms. We ran our estimation algorithm on 50 data sets with different realization of white noise and computed the average and standard deviation of the distribution using a 3σ clipping process to prevent the average and standard deviation from being affected by strong outliers.

As a result, planets b, c, g, f, and e are well detected at a white noise level of 1 m s^{-1}. When the white noise level is increased to 2 m s^{-1}, we start to loose planet f with sampling scheme D and, to a lesser extent, with sampling scheme C. In contrast, the evidence for planets d and h lie systematically below the detection threshold. Fig. 11 shows the Bayes factors in favour of planets f, e, d, and h depending on the sampling scheme and white noise level. For each sampling scheme, we note similar trends for the evolution of the Bayes factor as a function of the planet searched at white noise levels 1 and 2 m s^{-1}. For sampling scheme A (i.e. evenly sampled datasets), all planets are firmly detected except planets d (positive evidence) and h (inconclusive). The evidence for planet d decreases with the number of data points except for sampling C where the mass of planet d is overestimated. Finally, the Bayes factor of planet h is significantly higher is scheme D than in other schemes. This trend is likely correlated to the overestimation of the mass of planet h, whose accuracy strongly decreases from schemes A/B/C to D. This observation needs however to be nuanced as planet h remains undetected, in average, for all the sampling schemes.

5 DISCUSSIONS AND CONCLUSION

In this paper, we simulated RV follow-ups of TRAPPIST-1 using various sampling strategies spread on the same 90-d period. We first generated an RV curve for TRAPPIST-1a including a realistic stellar activity signal generated from a non-magnetic stellar activity model statistically compatible with the *K2* light curve (Luger et al. 2017) and the planet RV signal using the parameters measured from photometric analyses (D18; G18). This resulted in planet and stellar activity curves with similar amplitudes (∼5 m s^{-1}). We then built time-series using various sampling schemes, namely evenly sampled observations, bi-site observations from complementary

Downloaded from https://academic.oup.com/mnras/article/488/4/5114/5532370 by guest on 28 May 2023



(a) 1 m s^{-1} rms white noise (b) 2 m s^{-1} rms white noise

Figure 11. Logarithm of the Bayes factors relative to planets f, e, d, and h for the four sampling schemes listed in Table 4 at 1 and 2 m s^{-1} rms white noise level (left-hand and right-hand panels, respectively). The blue, yellow, and red bands correspond to strong, positive, and non-conclusive evidence according to Jeffreys (1961) criterion. A Bayes factor that lies beyond the black horizontal line is interpreted as a fair detection. The values that do not appear in the figure lie beyond the *Y*-axis upper limit.

longitudes and observations from a single location, to which we added a random noise accounting for additional noise sources (photon and instrumental non-correlated noise).

The stellar activity curve is modelled using GPR with a quasi-periodic kernel. Assuming that the planet orbital periods are known from transit photometry, we obtain that masses can be estimated at a precision $\gtrsim 5\sigma$ for planets b, c, g, and e at a white noise level of 1 m s^{-1}, regardless of the considered sampling scheme. Planets d and h are systematically undetected. The increase in the white noise level impacts the errors on the planet masses in different ways among the error bars on the planet mass estimates, some being dominated by the white noise and only poorly sensible to the stellar activity signal (planets c, f, and e), while others are dominated by the stellar activity (mostly planets b, d, and h).

Ancillary stellar activity indicators could be included to the GP modelling of the stellar activity signal in a similar approach as Rajpaul et al. (2015) and Jones et al. (2017). However, the sensitivity of each indicator to stellar activity is shown to vary from one indicator to another, as demonstrated by the obvious differences observed between the most prominent peaks in the periodograms of RV and ancillary indicators time-series (Hébrard et al. 2016). As a consequence, we ignored them when modelling the stellar activity signal.

This paper demonstrates that observational sampling is at least as critical as RV precision for velocimetric follow-ups of transiting planetary systems. For multiplanet TRAPPIST-1 analogues, we show that a dense-enough coverage of both planet orbital and stellar rotation cycles is essential to accurately recover the masses of all planets whose semi-amplitude RV wobble exceeds 1 m s^{-1}. Failing to achieve so can cause some of the system planets to remain undetected, or lead to tentative planet detections with erroneous mass estimates.

In the specific case of TRAPPIST-1, we find that an uninterrupted monitoring from CFHT can yield reliable mass estimates for planets with strongest RV semi-amplitudes and/or for which mass estimates are limited by the white noise (planets b and c, and to a lesser extent g, f, and e). However, planets with weaker RV signatures (d and h in particular, and potentially g and f as well) are likely to remain undetected, or marginally detected with erroneous mass estimates. Limiting observations to bright-time periods only is expected to further degrade the situation, whereas adding data from a second observing site with similar instrumental capabilities and located at a complementary longitude can provide most of the missing material for reliably recovering some of the less-massive planets. In this respect, SPIRou at CFHT and SPIP, its upcoming twin at TBL, are expected to form a powerful duo capable of achieving dense-enough coverage for the velocimetric follow-up of transiting planet candidates to be carried out within the SLS.

Besides that, we consider to adapt this study to other SLS-TF targets, especially young very active stars, in order to investigate how our results will be affected by significantly different activity and planet RV curves and prepare future observations with SPIRou.

ACKNOWLEDGEMENTS

This project received funding from the European Research Council (ERC) under the H2020 research & innovation programme (grant agreements #740651 NewWorlds). All plots in this paper were created using the MATPLOTLIB PYTHON module (Hunter 2007) and the ephemerides were computed using the PYEPHEM module. We warmly thank the referee, Prof. Suzanne Aigrain, for valuable comments and suggestions that helped us reaching a clearer and more robust version of the manuscript. The authors finally thank

Downloaded from https://academic.oup.com/mnras/article/488/4/5114/5532370 by guest on 28 May 2023

Hervé Carfantan and Mégane Boudineau for their help on the MCMC process.

REFERENCES

Aigrain S., Pont F., Zucker S., 2012, MNRAS, 419, 3147
Alibert Y., Benz W., 2017, A&A, 598, L5
Angus R., Morton T., Aigrain S., Foreman-Mackey D., Rajpaul V., 2018, MNRAS, 474, 2094
Artigau É., Malo L., Doyon R., Figueira P., Delfosse X., Astudillo-Defru N., 2018, AJ, 155, 198
Barnes J. R., Collier Cameron A., 2001, MNRAS, 326, 950
Barnes J. R., Collier Cameron A., James D. J., Donati J.-F., 2001, MNRAS, 324, 231
Basri G. et al., 2010, ApJ, 713, L155
Berdyugina S. V., 2005, LRSP, 2, 8
Berta-Thompson Z. K. et al., 2015, Nature, 527, 204
Boisse I., Bonfils X., Santos N. C., 2012, A&A, 545, A109
Bonfils X. et al., 2013, A&A, 549, A109
Bonfils X. et al., 2015, Proc. SPIE, 9605, 96051L
Borgniet S., Meunier N., Lagrange A.-M., 2015, A&A, 581, A133
Burt J., Holden B., Wolfgang A., Bouma L. G., 2018, AJ, 156, 255
Chib S., Jeliazkov I., 2001, J. Amer. Statist. Assoc., 96, 270 (CJ01)
Delrez L. et al., 2018, MNRAS, 475, 3577 (D18)
Desort M., Lagrange A.-M., Galland F., Udry S., Mayor M., 2007, A&A, 473, 983
Dittmann J. A. et al., 2017, Nature, 544, 333
Donati J.-F. et al., 2018, Handbook of Exoplanets. Springer International Publishing AG, Basel, Switzerland, p. 107
Dorn C., Mosegaard K., Grimm S. L., Alibert Y., 2018, ApJ, 865, 20
Dressing C. D., Charbonneau D., 2015, ApJ, 807, 45
Ducrot E. et al., 2018, AJ, 156, 218
Dumusque X., Udry S., Lovis C., Santos N. C., Monteiro M. J. P. F. G., 2011, A&A, 525, A140
Ford E. B., 2008, AJ, 135, 1008
Foreman-Mackey D., 2016, J. Open Source Softw., 1, 24
Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306
Gaidos E., Mann A. W., Kraus A. L., Ireland M., 2016, MNRAS, 457, 2877
Gillon M. et al., 2016, Nature, 533, 221
Gillon M. et al., 2017, Nature, 542, 456
Gregory P. C., 2007, MNRAS, 381, 1607
Grimm S. L. et al., 2018, A&A, 613, A68 (G18)
Haywood R. D. et al., 2014, MNRAS, 443, 2517 (H14)
Haywood R. D. et al., 2016, MNRAS, 457, 3637
Hébrard É. M., Donati J.-F., Delfosse X., Morin J., Moutou C., Boisse I., 2016, MNRAS, 461, 1465
Henry T. J., Jao W.-C., Subasavage J. P., Beaulieu T. D., Ianna P. A., Costa E., Méndez R. A., 2006, AJ, 132, 2360
Howell S. B. et al., 2014, PASP, 126, 398
Hunter J. D., 2007, Comput. Sci. Eng., 9, 90
Jeffreys H., 1961, Theory of Probability, 3rd edn. Clarendon Press, Oxford
Jones D. E., Stenning D. C., Ford E. B., Wolpert R. L., Loredo T. J., Dumusque X., 2017, preprint (arXiv:1711.01318)
Kipping D., 2018, RNAAS, 2, 136
Luger R. et al., 2017, NatAs, 1, 0129
Meunier N., Lagrange A.-M., Mbemba Kabuiku L., Alex M., Mignon L., Borgniet S., 2017, A&A, 597, A52
Mills S. M., Mazeh T., 2017, ApJ, 839, L8
Morley C. V., Kreidberg L., Rustamkulov Z., Robinson T., Fortney J. J., 2017, ApJ, 850, 121
Morris B. M., Agol E., Davenport J. R. A., Hawley S. L., 2018a, ApJ, 857, 39
Morris B. M. et al., 2018b, ApJ, 863, L32
Nutzman P., Charbonneau D., 2008, PASP, 120, 317
Queloz D. et al., 2001, A&A, 379, 279
Rackham B. V., Apai D., Giampapa M. S., 2018, ApJ, 853, 122
Rajpaul V., Aigrain S., Osborne M. A., Reece S., Roberts S., 2015, MNRAS, 452, 2269
Rasmussen C. E., Williams C. K. I., 2006, Gaussian Processes for Machine Learning. MIT Press, Cambridge, Massachusetts, USA
Ricker G. R. et al., 2015, JATIS, 1, 014003
Saar S. H., Donahue R. A., 1997, ApJ, 485, 319
Santerne A. et al., 2018, NatAs, 2, 393
Unno W., 1956, PASJ, 8, 108
Van Grootel V. et al., 2018, ApJ, 853, 30
Vida K., Kővári Z., Pál A., Oláh K., Kriskovics L., 2017, ApJ, 841, 124
Weiss L. M., Marcy G. W., 2014, ApJ, 783, L6
West A. A. et al., 2011, AJ, 141, 97
Wheatley P. J. et al., 2018, MNRAS, 475, 4476
Zeng L., Sasselov D. D., Jacobsen S. B., 2016, ApJ, 819, 127

SUPPORTING INFORMATION

Supplementary data are available at *MNRAS* online.

APPENDIX: Fit of the data

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

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

Downloaded from https://academic.oup.com/mnras/article/488/4/5114/5532370 by guest on 28 May 2023