Eyes on K2-3: A system of three likely sub-Neptunes characterized with HARPS-N and HARPS[*]

M. Damasso[1], A. S. Bonomo[1], N. Astudillo-Defru[2], X. Bonfils[3], L. Malavolta[4,5], A. Sozzetti[1], E. Lopez[6], L. Zeng[7], R. D. Haywood[8], J. M. Irwin[8], A. Mortier[9], A. Vanderburg[8,10,**], J. Maldonado[11], A. F. Lanza[12], L. Affer[11], J.-M. Almenara[2], S. Benatti[5], K. Biazzo[12], A. Bignamini[13], F. Borsa[14], F. Bouchy[2], L. A. Buchhave[15], A. C. Cameron[9], I. Carleo[5], D. Charbonneau[8], R. Claudi[5], R. Cosentino[16], E. Covino[17], X. Delfosse[3], S. Desidera[5], L. Di Fabrizio[16], C. Dressing[18], M. Esposito[17], R. Fares[12], P. Figueira[19,20], A. F. M. Fiorenzano[16], T. Forveille[3], P. Giacobbe[1], E. González-Álvarez[11], R. Gratton[5], A. Harutyunyan[16], J. Asher Johnson[8], D. W. Latham[8], G. Leto[12], M. Lopez-Morales[8], C. Lovis[2], A. Maggio[11], L. Mancini[1,21,22], S. Masiero[11], M. Mayor[2], G. Micela[11], E. Molinari[16,23], F. Motalebi[2], F. Murgas[24,25], V. Nascimbeni[4,5], I. Pagano[12], F. Pepe[2], D. F. Phillips[8], G. Piotto[4,5], E. Poretti[14], M. Rainer[14], K. Rice[26], N. C. Santos[20,27], D. Sasselov[8], G. Scandariato[12], D. Ségransan[2], R. Smareglia[13], S. Udry[2], C. Watson[28], and A. Wünsche[3]

(Affiliations can be found after the references)

ABSTRACT

Context. M-dwarf stars are promising targets for identifying and characterizing potentially habitable planets. K2-3 is a nearby (45 pc), early-type M dwarf hosting three small transiting planets, the outermost of which orbits close to the inner edge of the stellar (optimistic) habitable zone. The K2-3 system is well suited for follow-up characterization studies aimed at determining accurate masses and bulk densities of the three planets.

Aims. Using a total of 329 radial velocity measurements collected over 2.5 years with the HARPS-N and HARPS spectrographs and a proper treatment of the stellar activity signal, we aim to improve measurements of the masses and bulk densities of the K2-3 planets. We use our results to investigate the physical structure of the planets.

Methods. We analysed radial velocity time series extracted with two independent pipelines using Gaussian process regression. We adopted a quasi-periodic kernel to model the stellar magnetic activity jointly with the planetary signals. We used Monte Carlo simulations to investigate the robustness of our mass measurements of K2-3 c and K2-3 d, and to explore how additional high-cadence radial velocity observations might improve these values.

Results. Even though the stellar activity component is the strongest signal present in the radial velocity time series, we are able to derive masses for both planet b ($M_{\rm b} = 6.6 \pm 1.1\ M_\oplus$) and planet c ($M_{\rm c} = 3.1^{+1.3}_{-1.2}\ M_\oplus$). The Doppler signal from K2-3 d remains undetected, likely because of its low amplitude compared to the radial velocity signal induced by the stellar activity. The closeness of the orbital period of K2-3 d to the stellar rotation period could also make the detection of the planetary signal complicated. Based on our ability to recover injected signals in simulated data, we tentatively estimate the mass of K2-3 d to be $M_{\rm d} = 2.7^{+1.2}_{-0.8}\ M_\oplus$. These mass measurements imply that the bulk densities and therefore the interior structures of the three planets may be similar. In particular, the planets may either have small H/He envelopes (<1%) or massive water layers, with a water content ≥50% of their total mass, on top of rocky cores. Placing further constraints on the bulk densities of K2-3 c and d is difficult; in particular, we would not have been able to detect the Doppler signal of K2-3 d even by adopting a semester of intense, high-cadence radial velocity observations with HARPS-N and HARPS.

Key words. stars: individual: K2-3 – planets and satellites: detection – planets and satellites: composition – techniques: radial velocities – stars: individual: EPIC 201367065 – stars: individual: 2MASS 11292037-0127173

1. Introduction

In the last decade, the search for potentially habitable exoplanets has focussed particularly on M-dwarf stars. In this context, we regard a potentially habitable exoplanet as, in the broadest sense, one that orbits within or close to the habitable zone (HZ) of the parent star. Using the two main indirect detection methods, photometric transits and radial velocity (RV), potentially habitable exoplanets can be detected more easily around M dwarfs than around earlier type stars.

However, even M dwarfs present some challenges because of their faintness, their magnetic activity, and the difficulty in measuring accurate stellar parameters, which are essential if we wish to accurately determine the planetary parameters. Several exoplanet surveys have been devised to target M dwarfs specifically, such as the current ground-based photometric experiments MEarth (Irwin et al. 2015), APACHE (Sozzetti et al. 2013), and TRAPPIST (Gillon et al. 2017); the upcoming

---

SPECULOOS[1] and ExTrA (Bonfils et al. 2015) projects; and surveys exploiting high-resolution and high-stability spectrographs (e.g. Bonfils et al. 2013; Delfosse et al. 2013b; Affer et al. 2016 and the HADES paper series; Quirrenbach et al. 2016).

Results from the *Kepler* and K2 missions have also been used to provide estimates for the occurrence of planets in the HZ of M dwarfs (e.g. Dressing & Charbonneau 2013, 2015; Howell et al. 2014; Dressing et al. 2017a, b). These analyses suggest that small, low-mass planets are abundant around M dwarfs, and a significant percentage of these planets may have properties suitable for the emergence of life (Dressing & Charbonneau 2015; Tuomi et al. 2014). Space-based missions planned for the very near future, such as TESS (Ricker et al. 2014), CHEOPS (Fortier et al. 2014), and PLATO (Rauer et al. 2014) will target bright and nearby M dwarfs to detect potentially habitable planets, while JWST (Beichman et al. 2014) and ground-based 25–40 m class telescopes will characterize the atmospheres of these planets.

Some of the most intriguing exoplanet discoveries of the last few years are temperate planets around nearby M dwarfs, which orbit close to, or within, the predicted circumstellar HZ. For example, temperate, low-mass planets have been discovered around a number of nearby M-dwarf stars including Proxima Centauri (Anglada-Escudé et al. 2016), TRAPPIST-1 (Gillon et al. 2017), LHS 1140 (Dittmann et al. 2017), GJ 273 (Astudillo-Defru et al. 2017), K2-18 (Cloutier et al. 2017), GJ 667 C (Anglada-Escudé et al. 2013; Delfosse et al. 2013a; Feroz & Hobson 2014), and Ross 128 (Bonfils et al. 2018). These new discoveries have stimulated many theoretical studies addressing the habitability of planets orbiting M dwarfs (see e.g. Shields et al. 2016 for a review on this topic). Open questions regarding the true habitability of these temperate planets include the influence of the spectral energy distribution and the activity of M dwarfs on planetary atmospheres, as well as the effect of tidal locking. Potentially habitable planets around M dwarfs are likely subject to physical conditions that are very different from those experienced on Earth, and their properties may also depend strongly on the M-dwarf spectral subtype.

Indeed, M dwarfs that host temperate rocky planets exhibit a wide range of physical properties that might influence habitability. GJ 273, Ross 128, LHS 1140, Proxima Centauri, and TRAPPIST-1 are mid- to late-M dwarfs (M3.5V, M4V, M4.5V, M5.5V, and M8V, respectively). GJ 273 and Ross 128 have weak magnetic activity, while Proxima Centauri and TRAPPIST-1 have high activity levels, which probably has a large impact on the potential habitability of their planets. K2-18 is, instead, an earlier type M2.5 dwarf with low chromospheric activity, and GJ 667 C is a quiet M1.5 dwarf. The fact that these temperate planets are located at different distances from host stars spanning different spectral subtypes makes their comparative characterization particularly interesting. Unfortunately, many of these planets either do not transit or their parent star is too faint to permit detailed follow-up, making it difficult to robustly characterize these planets and hence study their interior structures and compositions.

Within this context, the planetary system around K2-3 (EPIC 201367065), a nearby (45 pc) M0 dwarf ($V = 12$ mag; $J = 9.4$ mag), presents an interesting opportunity for follow-up studies. Observations from the K2 mission revealed that K2-3 hosts at least three transiting small planets (Crossfield

et al. 2015): K2-3 b ($R_p = 2 R_\oplus$, $P_{orb} = 10$ days), K2-3 c ($R_p = 1.7 R_\oplus$, $P_{orb} = 24.6$ days), and K2-3 d ($R_p = 1.6 R_\oplus$, $P_{orb} = 45.5$ days). According to the optimistic HZ boundaries derived by Kopparapu et al. (2013, 2014), planet K2-3 d orbits close to the inner edge of the HZ of its host star. A particularly intriguing property of the K2-3 system is that the host star is bright enough to estimate the masses of these planets using existing high-resolution stabilized spectrographs.

Measuring the masses of the K2-3 planets would be interesting for several reasons. First, by determining the mass and bulk density of the temperate planet K2-3d, we can extend the study of planets orbiting close or within the HZ to earlier type host stars than those discussed earlier. Moreover, because of their measured sizes, planets K2-3 c and K2-3 d are optimal targets to test the results of Rogers (2015), who found that the majority of the observed planets with radius $R_p \geq 1.6 R_\oplus$ have densities too low to have a bulk rocky composition. Fulton et al. (2017) cast light on this result by finding evidence for a bimodal distribution of small planet sizes orbiting stars with $T_{eff} > 4500$ K with periods less than 100 days. Planets are preferentially found with radii around $\sim 1.3 R_\oplus$ and $\sim 2.4 R_\oplus$, with a region between 1.5 and 2.0 $R_\oplus$ where planet occurrence is rare. This bi-modality likely reflects a separation between purely rocky planets and rocky cores surrounded by varying amounts of lower density volatile material. K2-3 c and K2-3 d fall within the gap between the two peaks in the planet radius distribution and orbit a star cooler than 4000 K, making them interesting test cases for understanding the composition of planets in this size regime at lower stellar irradiation than most of the planets analysed by Fulton et al. (2017).

Previously, other groups have recognized the appeal of the K2-3 system and have begun conducting RV follow-ups (Almenara et al. 2015; Dai et al. 2016). Almenara et al. (2015) found that stellar activity has a major impact on RV observations of K2-3 and, based on their single semester of monitoring, were unable to measure the masses of K2-3 c and K2-3 d robustly. Accurate determination of the masses of K2-3 c and K2-3 d evidently requires a larger dataset and denser sampling to trace out the activity signal.

We present the results of an intense RV follow-up of K2-3 conducted over three seasons with the HARPS-N and HARPS spectrographs. Despite having data from several independent teams, and having precise transit ephemeris from high-precision photometry (Beichman et al. 2016; Fukui et al. 2016), measuring the mass of K2-3 c and K2-3 d has been challenging.

The paper is organized as follows. We first introduce the RV datasets and discuss the significant signals present in the data through a frequency analysis (Sect. 2). In Sect. 3 we present updated stellar parameters and analyse the photometric light curves of K2-3 and the Hα line activity indicator time series. The analysis of the RVs with Gaussian processes (GP) is described in Sect. 4. The significance of our best-fit solution for the masses of K2-3 c and K2-3 d is investigated through Monte Carlo simulations and is discussed in Sect. 5. Finally, we present the mass-radius diagram for the K2-3 planets and discuss the implications of our findings for their bulk composition.

2. Description and first look on the HARPS(-N) radial velocity datasets

Northern and southern spectroscopic observations of K2-3 were carried out between January 22, 2015 and July 7, 2017, producing a total of 211 HARPS-N and 138 HARPS spectra. Of

---

[1] http://www.speculoos.ulg.ac.be/cms/c_3272698/en/speculoos-portail

the 349 spectra, 283 are unpublished observations while 66 HARPS observations were previously published by Almenara et al. (2015). The HARPS-N spectra come from two independent programmes: the HARPS-N Collaboration Guaranteed Time Observations (GTO)[2] and the Global Architecture of Planetary Systems programmes (GAPS; Benatti et al. 2016). The two collaborations shared observing time on this target to maximize the number of RV measurements and to optimize the observing strategy. The spectra were reduced with the version 3.7 of the HARPS-N Data Reduction Software (DRS; Cosentino et al. 2014). The extraction and wavelength calibration of the HARPS spectra were performed using the on-line pipeline (Lovis & Pepe 2007). The exposure time was fixed at 1800 s for both instruments, producing a typical signal-to-noise ratio $S/N = 30$ at a wavelength $\lambda \sim 550$ nm.

2.1. Definition of the final dataset

We excluded some observations from our analysis for two reasons. We did not consider spectra with a $S/N \leq 11$ as measured at $\lambda = 550$ nm (echelle orders 46 and 49 for HARPS-N and HARPS, respectively). We then identified RV measurements potentially contaminated by scattered moonlight. This is particularly important for faint targets such as K2-3, which lie close to the ecliptic plane. Using the procedure described in Sect. 2.1 of Malavolta et al. (2017), we identified five HARPS-N spectra that are probably contaminated by scattered moonlight, and we discarded them from the dataset[3]. We could not perform the same analysis on the full HARPS dataset because nearly half of those spectra were acquired with simultaneous Fabry-Perot and not with fibre B on sky. Nonetheless, we adopted a heuristic approach[4] to identify four potentially contaminated measurements, which we removed from the final dataset[5]. After these cuts, our final dataset includes 197 HARPS-N spectra and 132 HARPS spectra.

2.2. Radial velocity extraction

In this study we primarily use RVs extracted with the TERRA pipeline (Anglada-Escudé & Butler 2012). The TERRA pipeline is commonly used to extract RVs of M dwarfs because it typically provides measurements with better precision and lower scatter than those extracted for low-mass stars with the CCF recipe of the on-line DRS pipeline (Perger et al. 2017). The HARPS fibre link was upgraded with octagonal fibres on May 28, 2015 (Lo Curto et al. 2015), introducing an RV offset between data acquired before and after the upgrade. We accounted for this offset by producing *pre-* and *post-upgrade* spectral templates to extract the RVs, and by introducing a velocity zero point offset for each dataset as free parameter when fitting the time series.

The TERRA RV time series are listed in the on-line Tables A.1 and A.2 for HARPS-N and HARPS, respectively. The median values of the internal errors are 1.88 m s^{-1} for HARPS-N and 2.04 m s^{-1} for HARPS.

---

[2] https://plone.unige.ch/HARPS-N/science-with-harps-n
[3] They correspond to the epochs BJD$_{UTC}$ 2457407.638254, 2457412.711410, 2457412.732546, 2457413.703766, and 2457413.724936.
[4] We identified potentially contaminated observations as those where the absolute difference between the RV of the Moon (which we assumed was equal to the Earth's barycentric RV) and the target star was less than 15 km s^{-1} and more than 99% of the Moon's disk was illuminated.
[5] We discarded observations on the epochs BJD$_{UTC}$ 2457056.724361, 2457056.842204, 2457057.713776, and 2457057.849294.
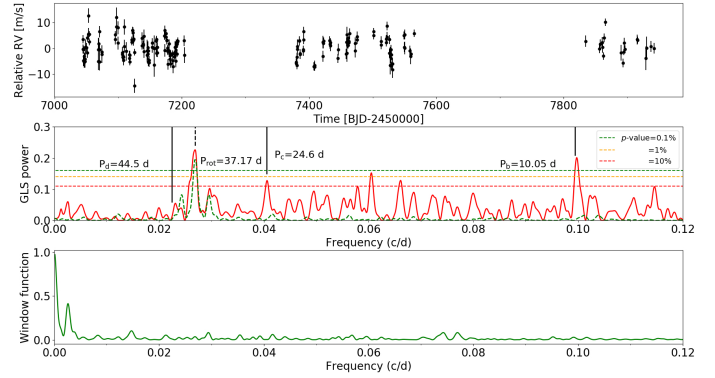

Fig. 1. *Upper plot*: time series of the K2-3 RVs extracted with the pipeline TERRA using the HARPS-N spectra. *Middle plot*: GLS periodogram of the RV time series (red line) is shown. Three levels of *p*-values are indicated by the horizontal dashed lines. The green dashed line overplotted on the RV periodogram represents the window function of the measurements (*bottom plot*) shifted in frequency to be superimposed on the strongest peak of the RV periodogram. This helps to identify the alias frequencies of the most relevant peaks. The stellar rotation and planetary orbital frequencies are indicated by vertical lines and corresponding labels.

2.3. Preliminary RV frequency analysis

We conducted a preliminary analysis of the RV datasets using the generalized Lomb-Scargle (GLS) algorithm (Zechmeister & Kürster 2009) to identify significant signals. We performed the GLS frequency analysis of the northern and southern datasets separately along with the combined RV dataset, removing RV offsets as determined by the analysis described in Sect. 4. We show the resulting periodograms in Figs. 1–3.

The HARPS-N periodogram is dominated by a signal at 37.2 ± 0.1 days[6], which we identify as the rotation period of the star (see Sect. 3). This signal has a *p*-value = 0.01%, as estimated by boot strapping the data, i.e. by randomly drawing the RV measurements (with replacement) and generating 10 000 mock datasets. Its semi-amplitude is 2.9 ± 0.3 m s^{-1}, as estimated with GLS, slightly lower than the RMS of the data (4.1 m s^{-1}). A significant peak with a slightly lower power appears at the orbital period of K2-3 b, while the peak corresponding to the orbital period of K2-3 c appears with less significance (*p*-value ~1%). The orbital period of K2-3 d is undetected in the HARPS-N periodogram.

The HARPS periodogram is dominated by the orbital period of K2-3 b (Fig. 2; *p*-value = 0.01%). The signal produced by the stellar rotation has a *p*-value around 1%, therefore it is much less significant than in the HARPS-N data. The window function is responsible for a pattern of alias frequencies around these signals. In addition to the one-year aliases, one-month aliases (synodic month frequency $f_{s.m.} = 0.03386$ c/d) are also probably present, which is expected for a star near to the ecliptic. For example, a one-month alias of the 37-day signal occurs at 0.061 c/d (=1/37+$f_{s.m.}$ c/d), while a one-month alias of P_b occurs at 0.065 c/d. Both aliases are visible in Fig. 2.

The signals that most clearly emerge in the periodogram of the combined dataset are those of the planet K2-3 b and the stellar rotation period, which are highly significant and have almost the same power (Fig. 3). The signal of planet K2-3 c is more weakly present with a *p*-value ~1%.

---

[6] In this work we adopt the uncertainties calculated with GLS as formal errors of the periods.
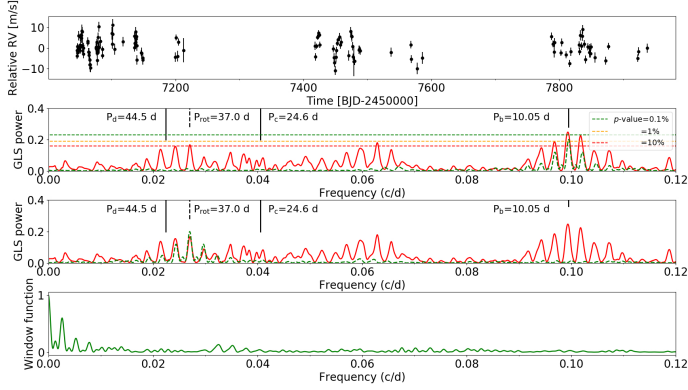
Fig. 2. As in Fig. 1 but for TERRA RVs extracted from the HARPS spectra. *Pre-* and *post-*upgrade RV offsets have been applied, as derived from our analysis described in Sect. 4.
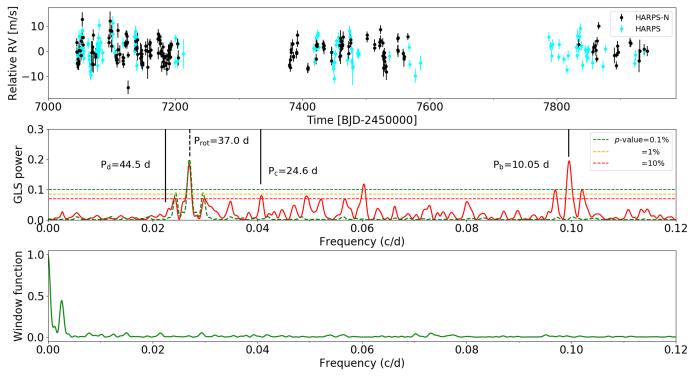

Fig. 3. As in Fig. 1 but for all TERRA HARPS-N and HARPS RVs. We applied an offset to each separate dataset, as derived from our analysis described in Sect. 4.

3. Stellar parameters and activity

The main properties of the star K2-3 are summarized in Table 1. While K2-3 has previously been characterized by Crossfield et al. (2015) and Almenara et al. (2015), we independently determined the stellar parameters from our HARPS-N spectra using the method developed by Maldonado et al. (2015)[7]. We estimated the effective temperature and iron abundance of the star using measurements of the pseudo-equivalent widths of the spectral features. We then determined stellar mass, radius, and surface gravity using empirical relationships. Our results are consistent with, but slightly more precise than, previous estimates. We adopted these new values for the analysis in the rest of the paper.

3.1. Photometry

The K2 light curve of K2-3, with all the transit signals removed, is shown in Fig. 4. We used the K2SFF[8] light curve processed as described by Vanderburg & Johnson (2014) and Vanderburg et al. (2016a). The light curve shows a quasi-periodic modulation with a flux semi-amplitude of ~0.1%. There is also clear evidence for changes from one rotation to the next, likely due to the evolution of active regions. In the bottom panel of Fig. 4 we show the GLS periodogram of the binned K2 light curve (one point

---

[7] https://github.com/jesusmaldonadoprado/mdslines
[8] https://archive.stsci.edu/prepds/k2sff/

Table 1. Stellar parameters for K2-3

Parameter	Value	Ref.
RA [deg, ICRS J2015]	172.3353579	
Dec [deg, ICRS 2015]	−1.4551256	
Mass [$M_\odot$]	0.62±0.06	(1)
	0.60±0.09	(2)
	0.61±0.09	(3)
	0.60±0.09	(4)
Radius [$R_\odot$]	0.60±0.06	(1)
	0.56±0.07	(2)
	0.55±0.04	(3)
	0.56±0.07	(4)
Effective temperature, T_{eff} [K]	3835±70	(1)
	3896±189	(2)
[Fe/H]	−0.01±0.09	(1)
	−0.32±0.13	(2)
Surface gravity, $\log g$ [$\log_{10}$(cgs)]	4.66±0.05	(1)
	4.73±0.06	(3)
	4.72±0.13	(4)
Density [$\rho_\odot$]	3.51±0.61	(3)
$\log(L/L_\odot)$	−1.15±0.09	(1)
Age [Gyr]	≥1	(2)

Notes. (1) This work: derived from HARPS-N spectra using the method described by Maldonado et al. (2015); (2) Crossfield et al. (2015); (3) Almenara et al. (2015); and (4) Sinukoff et al. (2016).

per day), which shows a strong peak at $P_{rot} = 38.3 \pm 0.7$ days[9]. Even though the K2 light curve was obtained about six months before the first spectroscopic observations and only covers about two rotation periods, the high S/N photometry is useful for constraining the stellar rotation period. It is interesting to note that, while the amplitude of the photometric rotational variability of the star is fairly low, the stellar rotation frequency is nonetheless the strongest signal in the RV time series of the star.

We also obtained time-series photometry of K2-3 from the MEarth survey (Irwin et al. 2015). We monitored K2-3 with one of the 40 cm telescopes of the Southern MEarth array from January 21, 2015 to April 4, 2016. With a total of 8669 data points, the light curve overlaps with the first two seasons of the spectroscopic observations. The fact that the MEarth data are contemporaneous with many of our RV observations and their high photometric precision (due in part to the excellent observing conditions at Cerro Tololo), make the MEarth light curve potentially useful for characterizing the stellar activity. We analysed the MEarth data (in nightly bins) by calculating a GLS periodogram. The highest peak in the periodogram is at 31.8 ± 0.2 days with a semi-amplitude of ~1 mmag (Fig. 5). This period does not correspond to the more robust stellar rotation period derived from the uninterrupted observations of K2 or using the H_α spectroscopic indicator (Sect. 3.2); nonetheless the result is worth mentioning because the data have been collected with uneven sampling by a ground-based small-aperture telescope and they show very low amplitude modulation compared to the typical photometric error. The periodogram also shows an additional peak at a period of about ~230 days, without any counterpart in the window function. While the signal could be due to systematics present in the MEarth data, its nature

---

[9] Since the K2 data barely cover two rotation cycles of K2-3, this estimate should not be considered particularly accurate.
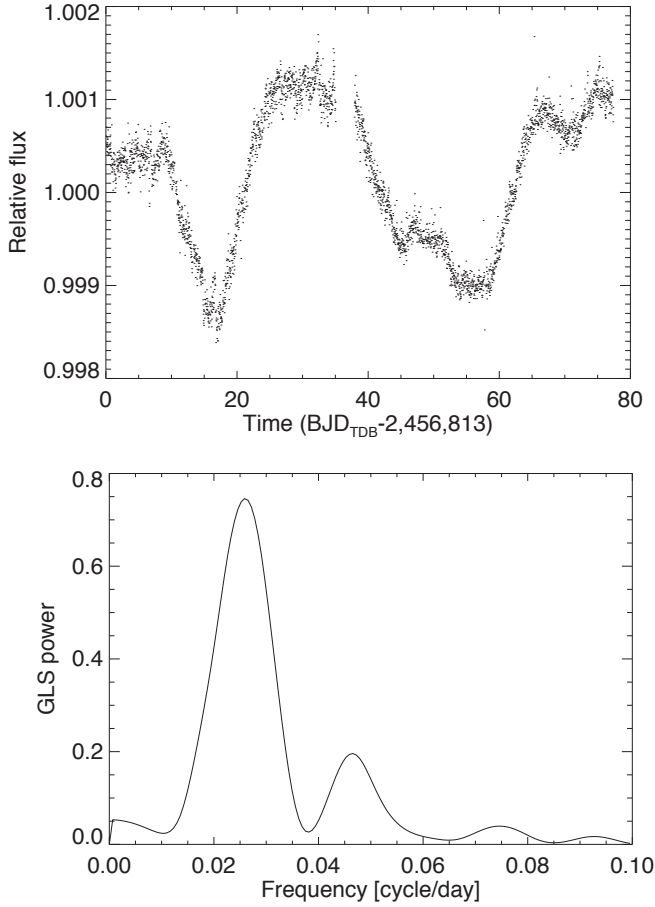
Fig. 4. *Top plot*: K2 light curve of K2-3 with the planetary transits removed. *Bottom plot*: GLS periodogram of the binned light curve (one point per day), showing a peak at $P_{rot} = 38.3$ days.

is unclear and an astrophysical origin cannot be ruled out (see Sect. 3.2).

3.2. Spectroscopic activity indexes

We studied the activity level of K2-3 during the time period of our observations using the Hα line as a spectroscopic activity indicator, which we extracted from our spectra using the method described by Gomes da Silva et al. (2011)[10]. The time series of the Hα activity indicator is shown in Fig. 6 and the data are listed in Table A.5. We analysed the time series by calculating a GLS periodogram to identify periodicities related to a possible activity cycle and stellar rotation. The GLS periodogram is shown in the second and third panels of Fig. 6.

The highest peak in the periodogram occurs at $P = 211 \pm 3$ days, and has a power comparable to that of its one-year alias frequency at $\sim$450 days. The existence of a long-term modulation is especially clear after looking at the data of the first two seasons. A bootstrap (with replacement) Monte Carlo analysis based on 10 000 mock datasets reveals that these peaks are statistically significant and have false alarm probabilities lower than 0.1%. We are, however, unable to ascertain whether the 211-day period or its alias at 450 days is the true underlying period. The origin of this long-period signal is unclear. If the signal is astrophysical, one possible

---

[10] We did not analyse the activity indicator based on the CaII H&K lines because of the very low S/N in this spectral region.
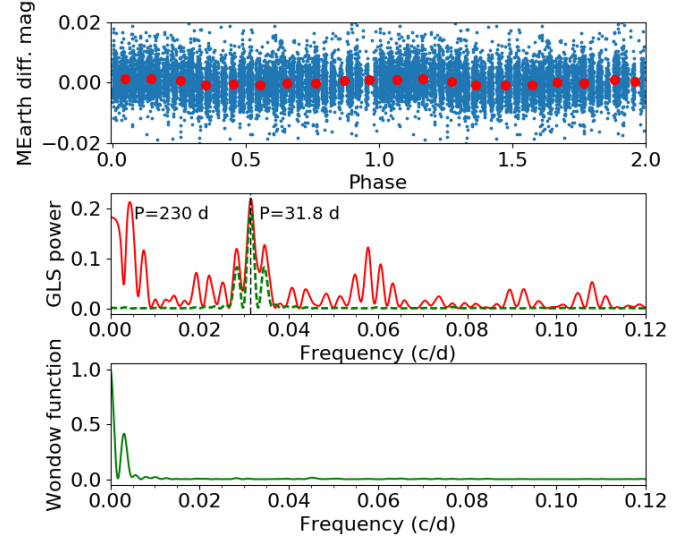

Fig. 5. *Top plot*: light curve of K2-3 from the MEarth-south survey, folded at the best period $P = 31.8$ days found using GLS. *Middle plot*: GLS periodogram of the MEarth light curve. The green line corresponds to the window function of the measurements (*bottom plot*), which is shifted in frequency so that the peak is superimposed on the major peak of the RV periodogram to directly identify alias frequencies.

explanation is an intermediate-duration activity cycle. Some tentative evidence exists for such cycles in low-mass stars (Savanov 2012; Robertson et al. 2013) and could represent sub-cycles superimposed on longer duration activity cycles, as observed for the Sun (so-called Rieger cycles). We note that the 211-day period is close to the 230-day signal observed in the MEarth photometry.

The second highest peak in the periodogram of the Hα indicator is close to the expected stellar rotation period and is highly significant ($P = 40.3 \pm 0.1$ days, p-value $< 0.1\%$). This signal is particularly strong in the last season of observations; during this time, a clear modulation related to P_{rot} is visible in the Hα time series, which covers nearly four stellar rotations. A GLS analysis of only the last season of observations identifies a periodicity of 43.5 ± 1.0 days. Folding the data at this period reveals that the modulation does not have a simple sinusoidal shape (Fig. 7).

4. Gaussian process regression analysis of radial velocities

We used the stellar activity information derived in Sect. 3 to perform a detailed analysis of the combined HARPS-N and HARPS RV datasets within a Bayesian framework based on GP regression. It has now become standard in RV analysis to use GPs to model the stellar contribution to the RV variations jointly with a number of Keplerian functions describing the planetary orbital motion (see Haywood et al. 2014 for the first application of this technique). This approach has proven to be a powerful to retrieve planetary masses from high-precision RV measurements when a signal closely related to the stellar rotation period, or its harmonics, is present in the RV time series (Dumusque et al. 2017).

For a general description of the GP method, and its performance when applied to RV time series, we refer, amongst others, to the recent works by López-Morales et al. (2016), Cloutier et al. (2017), Damasso & Del Sordo (2017), and Dittmann et al. (2017). The GP regression is based on the choice
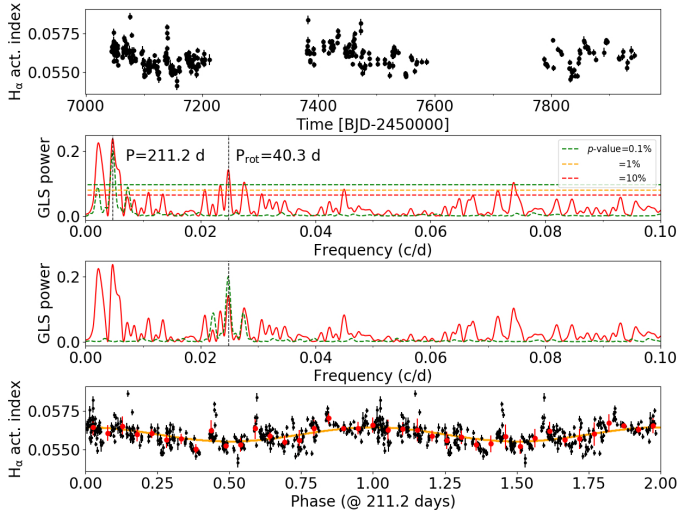
Fig. 6. *Top plot*: time series of the activity indicator based on the Hα line extracted from the HARPS and HARPS-N spectra. *Second and third plots*: GLS periodogram of the dataset is shown. The dashed horizontal lines indicate the p-value levels as derived from a bootstrap analysis. The green curve corresponds to the window function of the measurements, which is shifted in frequency so that the peak is superimposed on the major peak of the RV periodogram (*second plot*), to directly identify alias frequencies. In the *third plot* the window function is shifted to be superimposed on the P ∼ 40 days rotational signal. *Bottom*: time series phase-folded at the period $P = 211$ days. Red points represent the average of the data within 20 bins in the phase range [0,1].

of a specific kernel; i.e. a covariance matrix describing the correlation between measurements taken at two different epochs. For the case of K2-3, the so-called *quasi-periodic* (q-p) kernel is particularly useful because a signal very likely related to the stellar rotation period P_{rot} dominates the RV time series (Sect. 2), and there is evidence for an evolutionary timescale of the active regions close to P_{rot} (Sects. 3.1 and 3.2). Each element of the covariance matrix has the form

$$K(t, t') = h^2 \cdot \exp\left[-\frac{(t-t')^2}{2\lambda^2} - \frac{sin^2(\frac{\pi(t-t')}{\theta})}{2w^2} \right]$$
$$+ \left[\sigma_{RV,instr}^2(t) + \sigma_{jit,instr}^2 \right] \cdot \delta_{t,t'}, \qquad (1)$$

where t and t' represent two different epochs. The first term represents the quasi-periodic kernel, which is composed of a periodic term and an exponential decay term. This functional form is suitable for modelling a recurrent signal linked to stellar rotation and takes into account the finite lifetime of the active regions. Eq. (1) contains four covariance matrix hyperparameters: h represents the amplitude of the correlations; θ represents the rotation period of the star; w is the length scale of the periodic component, linked to the size evolution of the active regions; and λ is the correlation decay timescale, which can be physically related to the active regions lifetime. The remaining parameters in Eq. (1) are $\sigma_{RV,instr}(t)$, which is the RV internal error at time t for each spectrograph (or independent dataset); $\sigma_{jit,instr}$, which are additional uncorrelated jitter terms, one for each instrument (or independent dataset), which we add in quadrature to the internal errors to account for additional instrumental effects and noise sources neither included in $\sigma_{RV,instr}(t)$ nor modelled by the q-p kernel; and $\delta_{t,t'}$, which is the Kronecker delta function.

Our GP analysis was based on a Markov chain Monte Carlo (MCMC) algorithm. The model, algorithms, and statistical
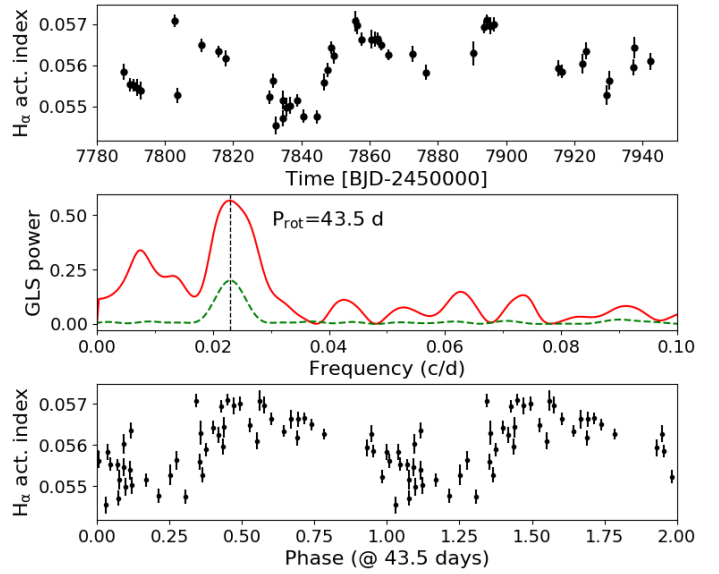
Fig. 7. *Top plot*: time series of the activity indicator based on the Hα line extracted from the HARPS and HARPS-N spectra. Here, only the dataset of the third season is shown. *Middle plot*: GLS periodogram of the dataset is shown. The green line corresponds to the window function of the measurements, which is shifted in frequency so that the peak superimposes to the major peak of the RV periodogram, to directly identify possible alias frequencies. *Bottom plot*: time series phase-folded at the period $P = 43.5$ days.

framework used in this work are the same as described in Damasso & Del Sordo (2017), and we refer to that work for a detailed description. In this work, we assumed circular orbits for all the planets. The best-fit values and uncertainties for each jump parameter were calculated as the median of the marginal posterior distributions and the 16% and 84% quantiles.

4.1. Choice of the priors

The priors adopted in our analysis are listed in Table 2. In this subsection, we describe and justify some of our choices for the priors.

In our fits, we allowed the semi-amplitude h of the correlated stellar signal to vary up to a value of 5 m s^{-1}, which represents nearly twice the value estimated by GLS at $P = 37$ days. The prior range used for the stellar rotation hyperparameter θ is defined based on our rotation period estimates from the K2 photometry and Hα activity indicator. This also takes into account the results of a trial MCMC analysis that was run after the conclusion of the second observing season, which adopted a larger prior range and indicated that the posterior distribution was well constrained within the range [35, 43] days.

For the active regions evolutionary timescale λ we adopted a uniform prior between 20 and 60 days, corresponding nearly to $P_{rot}/2$ and $1.5 \cdot P_{rot}$. This choice was first motivated by the marginal posterior obtained with the trial MCMC analysis of the dataset for the first and second seasons, which is symmetric around $\lambda = 36$ days ($\sigma \sim 8$ days). An evolutionary timescale of the order of the stellar rotation period could also be guessed directly by looking at the K2 light curve, which shows changes in its pattern from one rotation to the next[11]. We also used the results of Giles et al. (2017; Eq. (8)) to get a loose estimate

---

[11] Due to the short time baseline, we could not constrain λ through the analysis of the autocorrelation function, as done by López-Morales et al. (2016) using the longer baseline of the *Kepler* photometry.

of the timescale based on the RMS of the K2 data and the stellar effective temperature. The equation for determining the timescale was derived by Giles et al. (2017) using a calibration sample not biased by spectral type and, although faster rotators than K2-3 were analysed in the Giles et al. work, this result could be tentatively used for a star with $P_{rot} = 40$ days. With an $RMS = 7.6 \times 10^{-4}$ mag and $T_{eff} = 3835$ K, we get a timescale of 38^{+19}_{-13} days.

For the planetary orbital parameters, we fixed the upper limit of the Keplerian semi-amplitude to 5 m s^{-1} for K2-3 b and K2-3 c, and to 3 m s^{-1} for planet K2-3 d. For K2-3 b, $K = 5$ m s^{-1} is more than twice the value estimated by GLS after removing the stellar rotation signal, and represents a conservative upper limit also for the planet K2-3 c. For the outermost planet, K2-3 d, the absence of a signal in the GLS periodogram suggests that K_d should be significantly lower. The priors on the orbital period and time of transit were taken from Beichman et al. (2016), who determined the ephemerides by combining K2 observations with additional transits observed with the *Spitzer* space telescope.

4.2. Analysis of the combined radial velocity dataset

We analysed the full RV dataset without binning the data when more than one observation is available during the same night. To guarantee a wide exploration of the parameter space, we adopted 150 independent chains properly initialized to start from well separated locations. We discarded the first 3 000 steps of each chain by resetting the sampler. The MCMC chains reached the convergence according to the Gelman-Rubin statistics after 18 000 steps, and a further burn-in (0.75% of the total steps) was applied to calculate the best-fit values of the parameters (see Eastman et al. 2013 and references therein).

We show the best-fit results for all the free parameters and derived quantities in Table 3. The mass of the super-Earth K2-3 b is robustly determined with a S/N of $\sim 7\sigma$, while the mass of K2-3 c is detected with a lower significance of 2.6σ. The mass of K2-3 d remains undetermined, with a 1σ upper limit of about $\sim 2 M_{\oplus}$. This suggests that the mass of K2-3 d is too low to be detected, but it is also plausible that the low amplitude of the Doppler signal, compared to that of the stellar activity component, and its proximity to the stellar rotation period allow the signal to be absorbed into our GP activity model (Vanderburg et al. 2016b). We investigate the robustness of this upper limit in Sects. 5.1 and 5.2.

We show the RV curves in Fig. 8, with the stellar activity contribution subtracted, and folded at the best-fit orbital periods. The stellar contribution to the RV time series resulting from our fit is shown in Fig. 9. The stellar activity component of the model is reliably described by our best-fit results. The stellar rotation period and the evolutionary time scale of the active regions appear to be well characterized for both datasets. As we have found from the analysis of the RV datasets of other targets, the GP quasi-periodic regression tends to suppress low frequencies in the residuals. This means that if an additional long-term modulation is actually present in the original data, and it is not explicitly included in the fitted model, it would be absorbed by the stellar activity component and disappear in the residuals, after removing the planetary solutions. However, we note that the data in Fig. 9 do not show any long-term trend residual suggestive of an activity cycle or a longer period companion to K2-3 not included in our global model. We also note that the stellar rotation period, as retrieved by the GP regression, differs from that derived with GLS by ~ 3 days. This difference

Table 2. Prior probability distributions for the three-planet circular model parameters.

Jump parameter	Prior
h [m s^{-1}]	$\mathcal{U}$ (0.5, 5)
λ [days]	$\mathcal{U}$ (20, 60)
w	$\mathcal{U}$ (0.001, 1)
θ [days]	$\mathcal{U}$ (35, 43)
K_b [m s^{-1}]	$\mathcal{U}$ (0.05, 5)
P_b [days]	$\mathcal{N}$ (10.054544, 0.000029^2) [a]
$T_{0,b}$ [BJD-2 400 000]	$\mathcal{N}$ (56813.42024, 0.00094^2) [a]
K_c [m s^{-1}]	$\mathcal{U}$ (0.05, 5)
P_c [days]	$\mathcal{N}$ (24.64638, 0.00018^2) [a]
$T_{0,c}$ [JD-2 400 000]	$\mathcal{N}$ (56812.2777, 0.0026^2) [a]
K_d [m s^{-1}]	$\mathcal{U}$ (0.05, 3)
P_d [days]	$\mathcal{N}$ (44.55765, 0.00043^2) [a]
$T_{0,d}$ [JD-2 400 000]	$\mathcal{N}$ (56826.2248, 0.0038^2) [a]
$\gamma_{HARPS-N}$ [m s^{-1}]	$\mathcal{U}$ (−10, +10)
$\gamma_{HARPS-pre}$ [m s^{-1}]	$\mathcal{U}$ (−10, +10)
$\gamma_{HARPS-post}$ [m s^{-1}]	$\mathcal{U}$ (−10, +10)
$\sigma_{jit,HARPS-N}$ [m s^{-1}]	$\mathcal{U}$ (0.05, 5)
$\sigma_{jit,HARPS-pre}$ [m s^{-1}]	$\mathcal{U}$ (0.05, 5)
$\sigma_{jit,HARPS-post}$ [m s^{-1}]	$\mathcal{U}$ (0.05, 5)

Notes. [a]Ephemeris from Beichman et al. (2016), derived from transit observations with K2 and *Spitzer*. T_0 is the time of inferior conjunction.

is because the functional form of the quasi-periodic kernel used to model the stellar contribution is different from the simple sinusoid used by the GLS periodogram. We calculated the GLS periodogram of the GP quasi-periodic stellar activity timeseries and found the strongest peak at $P = 37$ days, reproducing the period found in our GLS analysis of the RV dataset. The RMS of the residuals is 2.6 m s^{-1}, only slightly higher than the median of the internal errors.

4.3. Cross-check with the alternative RV extraction method

In order to test the robustness of the results using RVs obtained with the TERRA dataset, we also extracted the RVs using an independent pipeline described by Astudillo-Defru et al. (2015, 2017). In brief, this pipeline works by aligning all the spectra to a common reference frame by removing the Earth's barycentric RV and by using the RVs of K2-3 measured by the DRS as an initial guess. Then a median template is computed from the aligned spectra and regions of the spectra contaminated by tellurics are rejected. The template is used to calculate a chi-square profile as a function of RVs for each individual spectrum, whose minimum corresponds to the stellar RV.

The RVs from this alternative pipeline are listed in Tables A.3, A.4. We performed the same analysis as described in Sect. 4.2 with the data from this alternative pipeline, and found that the results of this analysis (shown in Table 3) are in very good agreement with those obtained with TERRA. In particular, the planetary parameters are all in agreement within 1σ, strengthening our findings.

5. Assessing the reliability of the derived planetary masses

The results of the analysis presented in Table 3 show that while the Doppler signal of the innermost planet is retrieved with a

Table 3. Best-fit solutions for the quasi-periodic GP model applied to the combined HARPS/HARPS-N RV time series extracted with TERRA and an alternative pipeline.

Jump parameter	Best-fit value	
	TERRA	alternative pipeline
Stellar activity GP model		
h [m s^{-1}]	$2.9^{+0.4}_{-0.3}$	$3.1^{+0.5}_{-0.3}$
λ [days]	$40.0^{+10.4}_{-9.0}$	$43.5^{+9.8}_{-8.9}$
w	$0.18^{+0.11}_{-0.04}$	0.26 ± 0.07
θ [days]	$40.4^{+1.1}_{-1.9}$	$39.8^{+1.1}_{-0.9}$
Uncorrelated jitter		
$\sigma_{\mathrm{jit,HARPS-N}}$ [m s^{-1}]	1.4 ± 0.3	0.6 ± 0.4
$\sigma_{\mathrm{jit,HARPS-pre}}$ [m s^{-1}]	2.2 ± 0.5	2.5 ± 0.5
$\sigma_{\mathrm{jit,HARPS-post}}$ [m s^{-1}]	2.0 ± 0.6	1.6 ± 0.6
RV offset		
$\gamma_{\mathrm{HARPS-N}}$ [m s^{-1}]	$-0.06^{+0.54}_{-0.57}$	$30149.2^{+0.6}_{-0.7}$
$\gamma_{\mathrm{HARPS-pre}}$ [m s^{-1}]	$0.20^{+0.75}_{-0.79}$	$30480.2^{+0.8}_{-0.9}$
$\gamma_{\mathrm{HARPS-post}}$ [m s^{-1}]	$0.002^{+0.670}_{-0.688}$	30479.2 ± 0.8
Planetary orbital parameters		
K_{b} [m s^{-1}]	2.7 ± 0.4	2.9 ± 0.4
P_{b} [days]	10.05454 ± 0.00003	10.05454 ± 0.00003
$T_{0,\mathrm{b}}$ [BJD-24 00 000]	56813.42022 ± 0.00095	56813.42025 ± 0.00095
K_{c} [m s^{-1}]	0.95 ± 0.37	0.98 ± 0.34
P_{c} [days]	24.64638 ± 0.00017	24.64638 ± 0.00018
$T_{0,\mathrm{c}}$ [BJD-24 00 000]	56812.2777 ± 0.0026	56812.2778 ± 0.0026
K_{d} [m s^{-1}]	$0.29^{+0.34}_{-0.18}$ [<0.43 (68.3th perc.)]	$0.31^{+0.35}_{-0.20}$ [<0.47 (68.3th perc.)]
P_{d} [days]	44.55764 ± 0.00042	44.55766 ± 0.00043
$T_{0,\mathrm{d}}$ [BJD-24 00 000]	56826.2248 ± 0.0037	56826.2247 ± 0.0038
Planetary radii [a]		
$R_{\mathrm{p,b}}$ ($R_{\oplus}$)	2.29 ± 0.23	
$R_{\mathrm{p,c}}$ ($R_{\oplus}$)	1.77 ± 0.18	
$R_{\mathrm{p,d}}$ ($R_{\oplus}$)	1.65 ± 0.17	
Quantities derived from RVs [b]		
$M_{\mathrm{p,b}}$ ($M_{\oplus}$)	6.6 ± 1.1	7.0 ± 1.0
$M_{\mathrm{p,c}}$ ($M_{\oplus}$)	$3.1^{+1.3}_{-1.2}$	$3.2^{+1.2}_{-1.1}$
$M_{\mathrm{p,d}}$ ($M_{\oplus}$)	$1.2^{+1.4}_{-0.7}$ [<1.8 (68.3th perc.)]	$1.3^{+1.5}_{-0.8}$ [<1.9 (68.3th perc.)]
	$2.7^{+1.2}_{-0.8}$ (see Sect. 5.1)	
$\rho_{\mathrm{p,b}}$ [g cm^{-3}]	$3.0^{+1.3}_{-0.9}$	$3.2^{+1.3}_{-0.9}$
$\rho_{\mathrm{p,c}}$ [g cm^{-3}]	$3.1^{+1.9}_{-1.3}$	$3.2^{+1.8}_{-1.3}$
$\rho_{\mathrm{p,d}}$ [g cm^{-3}]	$1.6^{+2.1}_{-1.0}$ [<2.4 (68.3th perc.)]	$1.6^{+2.1}_{-1.0}$
	$3.1^{+2.0}_{-1.2}$ (see Sect. 5.1)	
$a_{\mathrm{p,b}}$ [AU]	$0.0777^{+0.0024}_{-0.0026}$	
$a_{\mathrm{p,c}}$ [AU]	$0.1413^{+0.0044}_{-0.0047}$	
$a_{\mathrm{p,d}}$ [AU]	$0.2097^{+0.0065}_{-0.0070}$	

Notes. Our global model includes three orbital equations (circular case). [a]Radii are derived using our estimate for the stellar radius, R = 0.60 ± 0.06 $R_{\odot}$, and the ratios $R_{\mathrm{planet}}/R_{\mathrm{star}}$ derived by Beichman et al. (2016) from K2 and *Spitzer* data. [b]Derived quantities from the posterior distributions. We used the following equations (assuming $M_{\mathrm{s}} + m_{\mathrm{p}} \cong M_{\mathrm{s}}$): $m_{\mathrm{p}}\sin i \cong (K_{\mathrm{p}} \cdot M_{\mathrm{s}}^{\frac{2}{3}} \cdot \sqrt{1 - e^2} \cdot P_{\mathrm{p}}^{\frac{1}{3}})/(2\pi G)^{\frac{1}{3}}$; $a \cong [(M_{\mathrm{s}} \cdot G)^{\frac{1}{3}} \cdot P_{\mathrm{p}}^{\frac{2}{3}}]/(2\pi)^{\frac{2}{3}}$, where G is the gravitational constant.
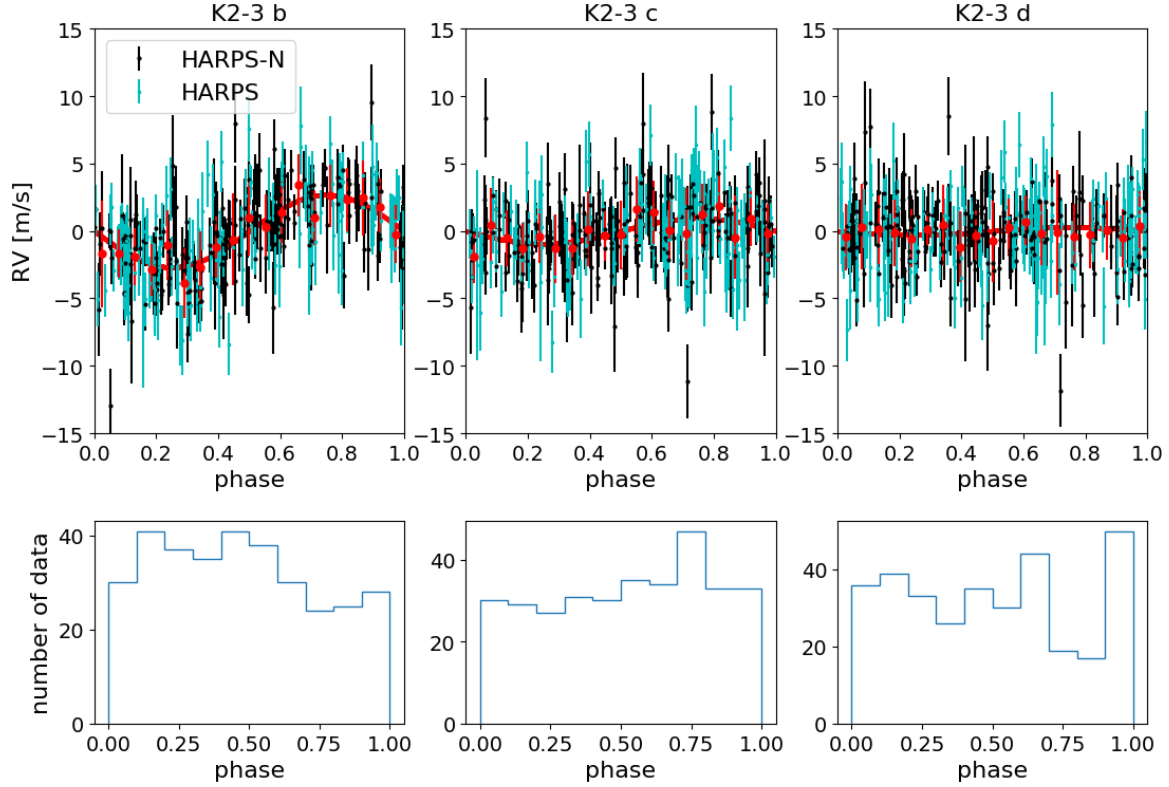
Fig. 8. *First row:* TERRA RV residuals, after removing our best-fit stellar component, phase-folded to the three planetary solutions found for the quasi-periodic GP model (represented by a red curve). *Second row:* histograms of the number of data divided in bins of phase are represented, showing a fairly uniform coverage for each planet.
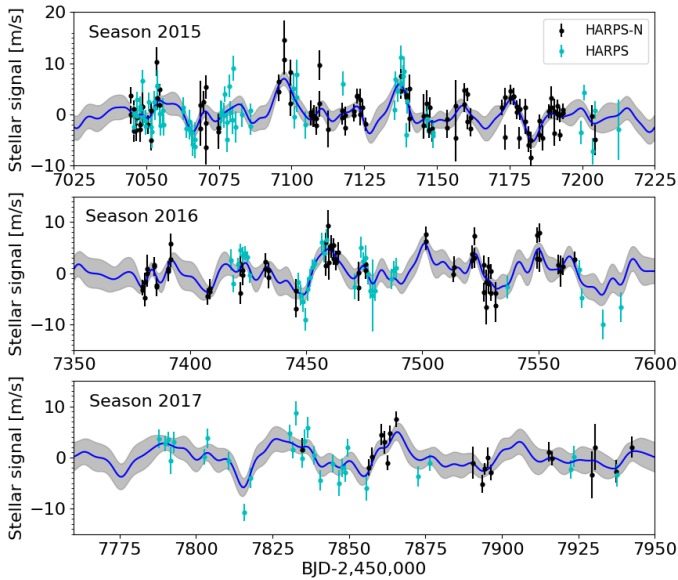

Fig. 9. Stellar signal contribution to the TERRA RV times series, for each observing season, as fitted with our GP quasi-periodic model (Table 3). The blue line indicates the best-fit curve, while the shaded grey area represents the ±1σ confidence interval.

Fig. 10. TERRA RV residuals, after removing the planetary and stellar activity signals, as fitted within a GP quasi-periodic framework (Table 3).

significance of ∼7σ, the Doppler signal of K2-3 c is retrieved with a significance lower than 3σ, and K2-3 d is undetected. With the present dataset and the adopted model, we cannot determine a bulk density of K2-3 c accurately enough to put reliable constraints on its possible composition, and we can only place an upper limit on the bulk density of K2-3 d. This impacts our understanding of the formation and evolutionary scenarios,

which led to the observed system architecture, and makes inferences about the habitability of K2-3 d especially challenging.

Why is the characterization of this system so challenging, despite the large dataset we have collected with two world-class spectrographs? Can we rule out a rocky composition for K2-3 d? To provide answers to these questions we performed simulations to investigate the impact of the observing sampling, stellar activity, and internal RV uncertainties on our ability to retrieve masses for the K2-3 planets.

5.1. Simulations using real epochs

We ran a first set of simulations based on the real epochs of our observations with the objective to improve the estimate of K2-3 d's mass, for which we could provide only an upper limit, and to assess the robustness of our result for K2-3 c's mass, despite its low significance. First, we adopted our best-fit values of the GP hyperparameters (TERRA dataset) to generate the stellar activity RV signal. We then injected planetary signals (circular orbits) with semi-amplitudes $K_b = 2.7$ m s^{-1}, $K_c = 0.95$ m s^{-1}, and $K_d = 1$ m s^{-1}, and the known periods and times of transit (Table 3) into the simulated activity signal. We refer to the dataset built in this way as the exact solution. While the simulated Doppler semi-amplitudes for K2-3 b and K2-3 c are those we derived from real data, for K2-3 d we assumed a value corresponding approximately to a purely rocky composition, given the measured radius[12].

We then created $N = 156$ mock RV time series[13]. Each dataset was obtained by randomly drawing RV values normally distributed around the exact solution, where $\sigma^2_{RV}(t) + \sigma^2_{jit,instr}$ was used as the variance of the normal distributions at each epoch. We analysed each synthetic dataset following the same procedure as for the real data (Sect. 4.2) and recorded the best-fit values of the free parameters once the MCMC chains reached convergence.

Results from these simulations depend on the actual properties of the stellar activity observed during our campaign. More complex simulations to explore in detail the effects of the stellar activity could also be carried out, where each hyperparameter is randomly drawn within the uncertainties while keeping the others fixed. However, such simulations, which would require a large number of mock datasets (i.e. thousands) and a correspondingly huge amount of computational time, are beyond the scope of this paper. Thus, our simulations do not explore the possibility that the non-detection is due to the proximity of the stellar rotation period θ to the orbital period of K2-3 d, which could be a limiting factor (Vanderburg et al. 2016b). We note that the amplitude h of the stellar activity term is precisely known ($\sim 10\sigma$, Table 3), so we expect that drawing values from the posterior distribution of this hyperparameter would not significantly change the results of our simulations.

Analysis framework. For each ith simulated dataset we derive a posterior distribution for the Doppler semi-amplitudes of planets c and d that we call $K_{p,i}$, where $p = (c, d)$. Each semi-amplitude is characterized by a median value $K^{med}_{p,i}$ and upper and lower uncertainties $\sigma^+_{p,i}$ and $\sigma^-_{p,i}$ (as derived from the 16th and 84th percentiles). We then calculate the median recovered semi-amplitude $K^{med}_{p,N}$ of all the $K^{med}_{p,i}$. We compare the median recovered semi-amplitude $K^{med}_{p,N}$ with the injected value $K_{p,inj}$ to draw conclusions about the results obtained for the real RV dataset.

For K2-3 d, we define the ratio $r_{d,i} = (K_{d,inj} - K^{med}_{d,i})/\sigma^+_{d,i}$ to measure the discrepancy between the best-fit estimate and the injected value in units of $\sigma^+_{d,i}$. The term $(K_{d,inj} - K^{med}_{d,i})$ is weighted by $\sigma^+_{d,i}$ to take the skewness of each posterior distribution into account. By averaging $r_{d,i}$ over the number, N, of simulated

datasets we get the metric $\bar{r}_d$, that we propose as a way to correct the measured semi-amplitude $K_{d,meas}$ using the equation

$$K_{d,real} = K_{d,meas} + \bar{r}_d \cdot \sigma^+_{d,meas}, \qquad (2)$$

where $K_{d,meas}$ and $\sigma^+_{d,meas}$ come from our best-fit solution (Table 3).

As an alternative approach, for each marginal distribution $K_{d,i}$ we calculate the percentile corresponding to the position of $K_{d,inj}$. We use the median over N of these percentiles to derive an estimate for $K_{d,real}$ from the posterior distribution obtained for the real dataset.

Results for K2-3 c. The distribution of the median values $K^{med}_{c,i}$ is shown in Fig. 11. The median of this distribution is $K^{med}_{c,N} = 0.96^{+0.27}_{-0.22}$ m s^{-1}, where the uncertainties represent the 16th and 84th percentiles. Looking at the values for $\sigma^+_{c,i}$ and $\sigma^-_{c,i}$ over all the marginal posteriors $K_{c,i}$, we note that $\langle(\sigma^+_{c,i} - \sigma^-_{c,i})\rangle = 0.01$ m s^{-1}, indicating that the distributions are normal-shaped and their average is $\langle\sigma^+_{c,i}\rangle = \langle\sigma^-_{c,i}\rangle = 0.33$ m s^{-1}. These results show that the injected signal $K_{c,inj} = 0.95$ m s^{-1} is well recovered, and indicate that our estimate of K_c from the real dataset is reliable, despite our detection having a significance $< 3\sigma$.

Results for K2-3 d. A sample of the N posterior distributions $K_{d,i}$ is shown in Fig. 12. The distribution of $K^{med}_{d,i}$ is shown in Fig. 13. The median of this distribution is $K^{med}_{d,N} = 0.54^{+0.18}_{-0.14}$ m s^{-1}, while we get $\langle\sigma^+_{d,i}\rangle = 0.47$ m s^{-1} and $\langle\sigma^-_{d,i}\rangle = 0.35$ m s^{-1} for the 68.3% confidence interval of each distribution. Therefore, the semi-amplitude K_d is generally underestimated with respect to $K_{d,inj}$. This result is suggestive when compared to what we get for K_c. As shown in Fig. 8, the phase coverage of our data is uniform for planet K2-3 c, and fairly uniform for planet K2-3 d. Then, we would expect two signals with an equal semi-amplitude (here $K_{c,inj} \simeq K_{d,inj}$) to be recovered with similar significance in absence of other hampering factors. For the metric $\bar{r}_d$ we find $\bar{r}_d = 0.99 \pm 0.04$, where the error is calculated as RMS[$r_{d,i}$]/$\sqrt{N}$. We use this result and our best-fit value for $K_{d,meas} = 0.29^{+0.34}_{-0.18}$ to draw $N = 10\,000$ random values for $\bar{r}_d$ and $K_{d,meas}$, obtaining a distribution of N samples for $K_{d,real}$ from Eq. (2). By taking the median of this distribution and the 68.3% confidence interval we get $K_{d,real} = 0.63^{+0.32}_{-0.18}$ m s^{-1}. This corresponds to $M_{d,real} = 2.7^{+1.2}_{-0.8}$ $M_\oplus$ and $\rho_{d,real} = 3.1^{+2.0}_{-1.2}$ g cm^{-3} for the planet mass and density. Using the second approach mentioned above, we find that $K_{d,real}$ corresponds on average at the 84th percentile of the K_d posterior distribution for the real dataset, that is $K_{d,real} = 0.62$ m s^{-1}, which is a result equal to that obtained with the first method.

Our statistical analysis shows that $K_{d,real}$ is < 1 m s^{-1} with 1σ confidence. This in turn suggests that the planets in the K2-3 system may have very similar bulk densities and thus share a similar composition, although this outcome should be taken with caution because of the large uncertainties in the masses and densities for K2-3 c and K2-3 d.

5.2. Role of the observing sampling

We devised new simulations to investigate how the detection significance of the signals induced by the planets K2-3 c and K2-3 d would change by increasing the number of the RV data collected with HARPS-N and HARPS, still assuming $K_d = 1$ m s^{-1}. We simulated an intensive observing strategy conducted during the third season, which has the lowest amount of real data, in a similar way as carried out for the high-cadence campaign devised

---

[12] $K = 1$ m s^{-1} corresponds to $M_p = 4$ $M_\oplus$.

[13] The number of simulations is as large as possible given the computational expense of a GP analysis on each simulated dataset (up to ~ 10 h each).
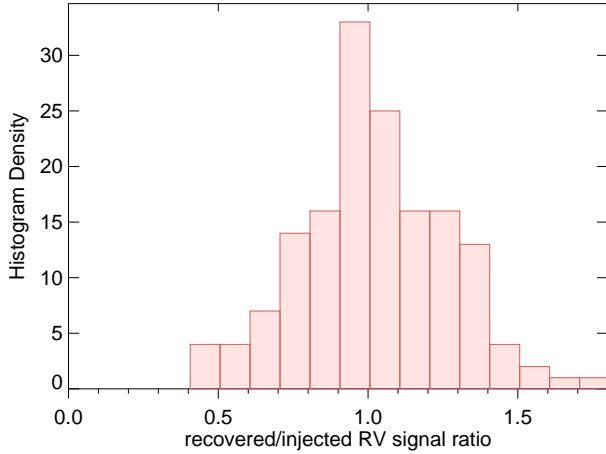
Fig. 11. Distribution of $N = 156$ median values for the Doppler semi-amplitude K_c of planet K2-3 c, normalized to the injected value $K_{c,inj}$, as derived from posterior distributions of the N mock RV datasets. This result refers to the case of real observing epochs.
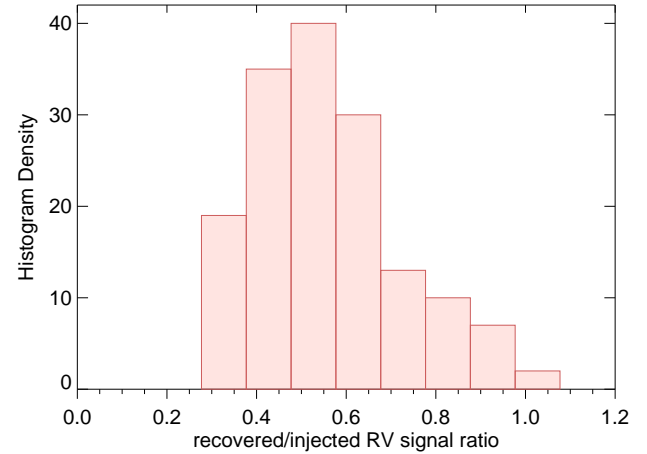

Fig. 13. Distributions of the 50th percentiles for all the posterior distributions of the semi-amplitude K_d obtained from 156 mock datasets, normalized to the injected value $K_{d,inj}$. This result refers to the case of real observing epochs.
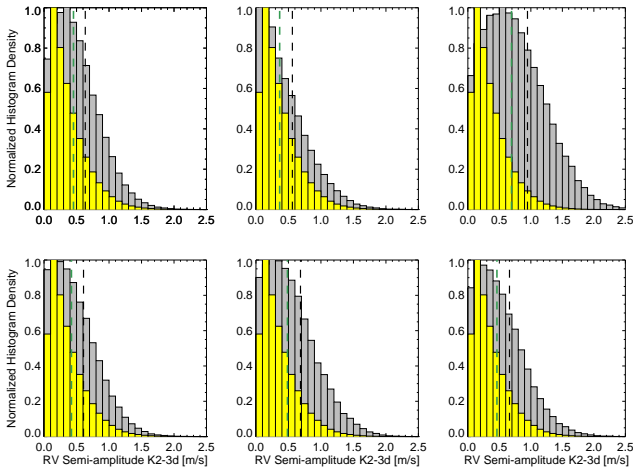
Fig. 12. Sample of posterior distributions (grey histograms) for the Doppler semi-amplitude K_d of planet K2-3 d obtained from the GP regression analysis of the mock RV datasets described in Sect. 5.1. Each plot shows the marginal distribution for a single mock dataset, and this is compared to the posterior distribution obtained for the real TERRA RV dataset, represented by the yellow histogram. The vertical dashed lines indicate the 50th (green) and 68.3th (black) percentiles for the posterior distributions of the mock datasets.

to detect Proxima b with HARPS (Anglada-Escudé et al. 2016). In order to keep our simulation realistic, we did not include mock epochs later than the 2017 observing season because our representation of the RV stellar signal cannot be considered predictive in the far future.

We created the mock datasets as described in Sect. 5.1. We generated all the epochs suitable for observations during the 2017 observing season in addition to the real epochs. We simulated only one measurement per night avoiding superposition with the epochs corresponding to the real observations. Every random epoch was selected by placing constraints on the Moon phase illumination and distance from the target (they have to be <90% and >45°, respectively), and on the altitude of the star above the horizon (airmass <1.7). Using these criteria we got 112 and 100 additional new epochs for HARPS-N and HARPS, respectively. We randomly removed 10% of these epochs at each simulation run to account for bad weather, thus simulating an

optimistic scenario for a feasible follow-up[14]. The uncertainties $\sigma_{RV}(t)$ of the mock RV data were randomly drawn from normal distributions with mean and σ equal to the average and RMS values of the HARPS-N and HARPS *post*-upgrade internal errors derived with TERRA. The final mock dataset is obtained by randomly shifting each data point of the exact solution within the error bars by a quantity $\Delta RV(t)$ drawn from a normal distribution with mean zero and σ equal to $\sqrt{\sigma_{RV}^2(t) + \sigma_{jit}^2}$. An example mock dataset is show in Fig. 14.

Our final sample is composed of $N = 97$ mock datasets. Also in this case, we analysed each simulated dataset within the same GP quasi-periodic framework applied to the real dataset, except for the σ_{jit} terms that were not included as free parameters, and we analysed the outcomes as described in Sect. 5.1.

Results for K2-3 c. The median and 68.3% confidence interval of the distribution for the $K_{c,i}$ semi-amplitudes are $K_c = 0.96^{+0.27}_{-0.26}$ m s^{-1}. For the upper and lower uncertainties we get $\langle (\sigma_{c,i}^+ - \sigma_{c,i}^-) \rangle = 0.006$ m s^{-1} over all the posterior distributions, indicating that they are generally normal-shaped. In addition, $\langle \sigma_{c,i}^+ \rangle = \langle \sigma_{c,i}^- \rangle = 0.26$ m s^{-1}. This result not only confirms that the estimate obtained from real data is robust, but also represents an improvement in the significance of the detection that is now increased to 3.7σ.

Results for K2-3 d. The median of the $K_{d,i}$ best-fit values is now $K_d = 0.58^{+0.18}_{-0.15}$ m s^{-1}, while $\langle \sigma_{d,i}^+ \rangle = 0.40$ m s^{-1} and $\langle \sigma_{d,i}^- \rangle = -0.34$ m s^{-1}. This result is not very different from that presented in Sect. 5.1, and shows that, despite the 190 additional data to the real dataset, the Doppler signal of K2-3 d is still underestimated and not significantly detected.

6. Discussion and conclusions

This work was focussed on deriving the masses and bulk densities of the three planets transiting the nearby M dwarf K2-3, using 329 RV measurements collected with HARPS and HARPS-N over a period of 2.5 years. We found that stellar activity makes a significant contribution to the RV variations over the entire time period. We have also shown that, for the case

---

[14] It must be considered that 10% refers to the epochs when K2-3 could actually be observed, not to all the nights of the third season.
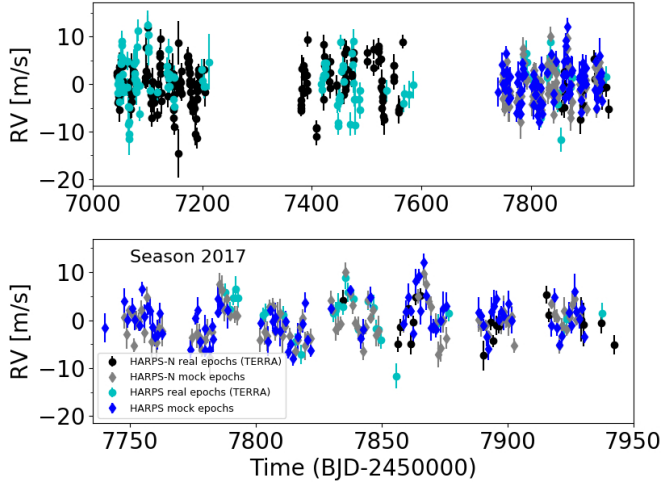
Fig. 14. Example of a simulated RV dataset used to explore the effects on the characterization of the K2-3 planets of additional measurements taken over the 2017 season. The *upper plot* shows the complete mock dataset, while only the third season is shown in the *second plot*, to better appreciate the intensive simulated sampling.

of K2-3, this can be effectively mitigated using a GP regression with a quasi-periodic kernel. The results of our global model describe the stellar activity component in a plausible way, and this allowed us to derive a precise and accurate mass estimate for K2-3 b. We also derive a mass for K2-3 c with a significance of less than 3σ. However, using simulations, we demonstrate that our estimate is accurate. Conversely, we do not detect, in our data, the Doppler signal induced by the temperate planet K2-3 d.

Figure 15 shows a planetary mass-radius diagram that includes planets for which the mass and radius were both measured with a relative error better than 30%. Theoretical mass-radius curves for various chemical compositions (Zeng & Sasselov 2013; Zeng et al. 2016) are shown with solid lines. The precision of the radii of the K2-3 planets is mainly limited by that of the stellar radius (all the relative uncertainties are ~10%; see Table 1). One firm outcome of our analysis is that, for K2-3 b, an Earth-like composition (~33% of iron and ~67% of silicates) is rejected with high confidence; we note that in Fig. 15 the masses are represented on a logarithmic scale. Concerning K2-3 c, our mass determination excludes an Earth-like composition with a confidence level of ~4σ (assuming $R_p = 1.77\ R_\oplus$). The non-detection of K2-3 d was explored in detail through simulations showing that the real Doppler semi-amplitude K_d is likely less than $1\ \mathrm{m\,s^{-1}}$ and its corresponding mass is $M_d = 2.7^{+1.2}_{-0.8}\ M_\oplus$ (Sect. 5). Looking at its position on the mass-radius diagram as derived from our simulations, the interior composition of K2-3 d would differ from that of the Earth with a confidence level greater than 2σ in mass and ~2σ in radius. We note that planets K2-3 c and K2-3 d occupy a region of the mass-radius diagram in which planet occurrence is rare, when only planets with mass and radius measured with a precision better than 30% are considered.

The corresponding bulk densities of all planets ($\rho_p \sim 3\ \mathrm{g\,cm^{-3}}$) show that they may have a very similar composition. If further measurements were to confirm our density estimates, excluding rocky compositions for K2-3 c and K2-3 d with higher significance, there are two scenarios that may explain the bulk properties of the K2-3 planets: water-poor planets with H/He envelopes or water worlds.

Several recent studies have investigated the water-poor hypothesis (Lopez 2017; Owen & Wu 2017; Jin & Mordasini

2018; Van Eylen et al. 2017). Fulton et al. (2017) analysed a sample of short-period planetary candidates detected by *Kepler* ($P < 100$ days) and demonstrated that the distribution of planetary radii is bi-modal: the planet candidates have radii that are predominantly either ~1.3 $R_\oplus$ or ~2.4 $R_\oplus$, and a gap is observed between 1.5 and 2 $R_\oplus$. The evolutionary model of Owen & Wu (2017) reproduced the observed bi-modal radius distribution in terms of two populations of planets: those consisting of a bare core resulting from photoevaporation, and those with twice the core radius, where the size is doubled by a H/He envelopes. The "gap" detected by Fulton et al. (2017) is actually observed for planets orbiting FGK stars, while the M-dwarf regime was not explored. However, K2-3 is more similar to the FGK sample than to a late-M dwarf, and the results from *Kepler* could still be applicable[15]. Therefore, K2-3 b could have a significant volatile envelope, large enough to measurably change its radius with respect to that of a rocky core. The same may be the case for the other two planets in the K2-3 system, but neither of their radius estimates allow us to unambiguously associate them with one of the groups.

If all planets share the same composition, their densities can be explained by modest primordial hydrogen and helium envelopes atop Earth-like iron and silicate cores. Using the stellar and planet properties derived in this work and assuming a rocky Earth-like core and a solar composition H/He envelope, we find that K2-3 b, c, d are best fit with H/He envelopes comprising 0.7%, 0.3%, and 0.4% of their masses, respectively (Lopez & Fortney 2014). Moreover, using the planetary evolution models of Lopez (2017) we find that none of the planets in this system are vulnerable to losing significant mass through photo-evaporative atmospheric escape.

Even though the water-poor scenario has received a great deal of attention, alternatives should also be considered. Water worlds are planets having massive water envelopes comprising $\geq$50% of the planet total mass. Recently, bi-modal radius distributions have been derived for the complete *Kepler* sample of Q1–Q17 small exoplanet candidates with radii $R_p <$ 4 $R_\oplus$ (Zeng et al. 2017a,b). These distributions show that the limits and extent of the radius gap depends on the spectral type of the host star. One proposed explanation for the observed bi-modal distributions is the existence of two populations of planets: rocky worlds, with the lowest radii, and water worlds. The two populations likely share the same underlying rocky component by mass, but differ in the presence of a H_2O-dominated mantle, which is similar to, or slightly more massive than, the rocky component. According to Zeng et al. (2017b), the radius gap for an M0 dwarf such as K2-3 is located at 1.6–1.7 $R_\oplus$. Therefore K2-3 b has an observed mass and radius consistent with that expected for a water world. This could also be the case for K2-3 c and K2-3 d. However, the accuracy of their radii places these planets close to the transition limit that defines the gap, making their water-world membership not highly significant.

We therefore conclude that all three planets in the K2-3 system are likely sub-Neptunes, defined as small, non-rocky planets that have enough volatiles to change their bulk composition measurably. Both the H/He gas envelope and water-world scenarios are possible, particularly for K2-3 b. Based on our data,

---

[15] An interesting counterexample is represented by planets orbiting the HZ of the late-M-dwarf TRAPPIST-1, which are thought to harbour significant amounts of water (Bourrier et al. 2017), in particular TRAPPIST-1 f (Quarles et al. 2017), showing the diversity of the possible scenarios within the M-dwarf class.
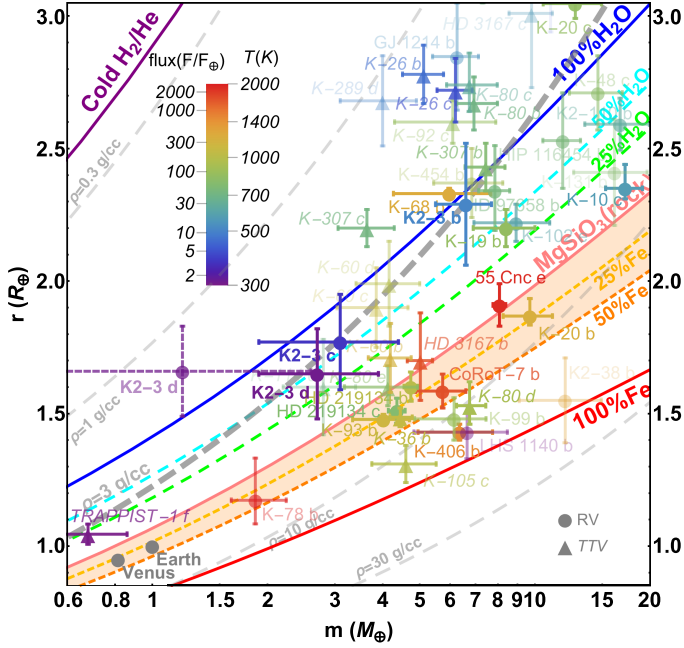
Fig. 15. Mass-radius diagram for exoplanets for which the mass and radius have been both measured with a relative error better than 30%. The location of the K2-3 planets is emphasized. For K2-3 d we also plot the mass derived from the GP analysis (shaded point in violet), and the value corrected using the result of the simulations described in Sect. 5.1. The curve for bulk density $\rho = 3$ g cm^{-3} is shown in grey passing through the three positions occupied by the K2-3 planets. The planetary data are taken from the NASA exoplanet archive and updated to August 30, 2017.

however, we cannot rule out that K2-3 c and K2-3 d have bare cores of purely rocky, Earth-like composition.

Within the H/He envelope scenario, planets likely formed within the first ~10 Myr before the dispersion of the gaseous proto-planetary disk. These water-poor planets could have formed in situ, and their similar masses would suggest similar formation histories (Lee & Chiang 2015, 2016). On the other hand, following Ginzburg et al. (2016) one may expect that the planets, during the cooling phase that follows their formation beyond the snow line (e.g. Selsis et al. 2007), were characterized by an intrinsic luminosity that could blow off, over a billion year timescale, any H/He envelope less than about ~5% by mass. Since a H_2O-layer has a much higher heat capacity than a H/He envelope, this evolutionary pathway could result in water worlds.

If K2-3 d is surrounded by a gaseous envelope, with the properties estimated here, this would likely result in surface pressure and temperatures that are too high to support a habitable planet scenario. A better characterization of K2-3 c and K2-3 d is left to the next generation of high-precision, high-stability spectrographs and to new photometric transit observations. We have shown that an intensive observing sampling over one season with two of the best spectrographs now available would still not have detected a signal with $K = 1$ m s^{-1} and a period $P = 44.5$ days. Assuming that our GP result is a good representation of the stellar activity contribution, this means that the true mass of K2-3 d is expected to remain unmeasurable even with a dataset of more than 500 RVs, which is currently not possible for a single target and without a collaboration among various teams. Thus, detecting the real signal induced by K2-3 d is currently very challenging. K2-3 is, however, an ideal target for characterization studies with the VLT/ESPRESSO spectrograph (Pepe et al. 2014), or with near-infrared (NIR) spectrographs such as CARMENES

(Quirrenbach et al. 2016), SPIRou (Artigau et al. 2014), and HPF (Mahadevan et al. 2014), provided that they reach their design RV precision and assuming, as expected, that RVs extracted from NIR spectra are less affected by stellar activity.

References

Affer, L., Micela, G., Damasso, M., et al. 2016, A&A, 593, A117
Almenara, J. M., Astudillo-Defru, N., Bonfils, X., et al. 2015, A&A, 581, L7
Anglada-Escudé, G., & Butler, R. P. 2012, ApJS, 200, 15
Anglada-Escudé, G., Tuomi, M., Gerlach, E., et al. 2013, A&A, 556, A126
Anglada-Escudé, G., Amado, P. J., Barnes, J., et al. 2016, Nature, 536, 437
Artigau, É., Kouach, D., Donati, J.-F., et al. 2014, in Ground-based and Airborne Instrumentation for Astronomy V, Proc. SPIE, 9147, 914715
Astudillo-Defru, N., Bonfils, X., Delfosse, X., et al. 2015, A&A, 575, A119
Astudillo-Defru, N., Forveille, T., Bonfils, X., et al. 2017, A&A, 602, A88
Beichman, C., Benneke, B., Knutson, H., et al. 2014, PASP, 126, 1134
Beichman, C., Livingston, J., Werner, M., et al. 2016, ApJ, 822, 39
Benatti, S., Claudi, R., Desidera, S., et al. 2016, in Frontier Research in Astrophysics II, held 23-28 May, 2016, Online at https://pos.sissa.it/cgi-bin/reader/conf.cgi?confid=269, 69
Bonfils, X., Delfosse, X., Udry, S., et al. 2013, A&A, 549, A109
Bonfils, X., Almenara, J. M., Jocou, L., et al. 2015, in Techniques and Instrumentation for Detection of Exoplanets VII, Proc. SPIE, 9605, 96051L
Bonfils, X., Astudillo-Defru, N., Díaz, R., et al. 2018, A&A, 613, A25
Bourrier, V., de Wit, J., Bolmont, E., et al. 2017, AJ, 154, 121
Cloutier, R., Astudillo-Defru, N., Doyon, R., et al. 2017, A&A, 608, A35
Cosentino, R., Lovis, C., Pepe, F., et al. 2014, Ground-based and Airborne Instrumentation for Astronomy V, Proc. SPIE, 9147, 8C
Crossfield, I. J. M., Petigura, E., Schlieder, J. E., et al. 2015, ApJ, 804, 10
Dai, F., Winn, J. N., Albrecht, S., et al. 2016, ApJ, 823, 115
Damasso, M., & Del Sordo F. 2017, A&A, 599, A126
Delfosse, X., Bonfils, X., Forveille, T., et al. 2013a, A&A, 553, A8
Delfosse, X., Donati, J.-F., Kouach, D., et al. 2013b, in SF2A-2013: Proc. of the Annual meeting of the French Society of Astronomy and Astrophysics, eds. L. Cambresy, F. Martins, E. Nuss, & A. Palacios, 497
Dittmann, J. A., Irwin, J. M., Charbonneau, D., et al. 2017, Nature, 544, 333
Dressing, C. D., & Charbonneau, D. 2013, ApJ, 767, 95
Dressing, C. D., & Charbonneau, D. 2015, ApJ, 807, 45
Dressing, C. D., Newton, E. R., Schlieder, J. E., et al. 2017a, ApJ, 836, 167
Dressing, C. D., Vanderburg, A., Schlieder, J. E., et al. 2017b, AJ, 154, 207
Dumusque, X., Borsa, F., Damasso, M., et al. 2017, A&A, 598, A133
Eastman, J., Gaudi, B. S., & Agol, E. 2013, PASP, 125, 83
Feroz, F., & Hobson, M. P. 2014, MNRAS, 437, 3540
Fortier, A., Beck, T., Benz, W., et al. 2014, Space Telescopes and Instrumentation 2014: Optical, Infrared, and Millimeter Wave, Proc. SPIE, 9143, 91432J
Fukui, A., Livingston, J., Narita, N., et al. 2016, AJ, 152, 171
Fulton, B. J., Petigura, E. A., Howard, A. W., et al. 2017, AJ, 154, 109
Giles, H. A. C., Collier Cameron, A., & Haywood, R. D. 2017, MNRAS, 472, 1618
Gillon, M., Triaud, A. H. M. J., Demory, B.-O., et al. 2017, Nature, 542, 456

Ginzburg, S., Schlichting, H. E., & Sari, R. 2016, ApJ, 825, 29
Gomes da Silva, J., Santos, N. C., Bonfils, X., et al. 2011, A&A, 534, A30
Haywood, R. D., Collier Cameron, A., Queloz, D., et al. 2014, MNRAS, 443, 2517
Howell, S. B., Sobeck, C., Haas, M., et al. 2014, PASP, 126, 398
Irwin, J. M., Berta-Thompson, Z. K., Charbonneau, D., et al. 2015, in Cambridge Workshop on Cool Stars, Stellar Systems, and the Sun, eds. G. T. van Belle & H. C. Harris, 18, 767
Jin, S., & Mordasini, C. 2018, ApJ, 853, 163
Kopparapu, R. K., Ramirez, R., Kasting, J. F., et al. 2013, ApJ, 765, 131
Kopparapu, R. K., Ramirez, R. M., SchottelKotte, J., et al. 2014, ApJ, 787, L29
Lee, E. J., & Chiang, E. 2015, ApJ, 811, 41
Lee, E. J., & Chiang, E. 2016, ApJ, 817, 90
Lo Curto, G., Pepe, F., Avila, G., et al. 2015, The Messenger, 162, 9
Lopez, E. D. 2017, MNRAS, 472, 245
Lopez, E. D., & Fortney, J. J. 2014, ApJ, 792, 1
López-Morales, M., Haywood, R. D., Coughlin, J. L., et al. 2016, AJ, 152, 204
Lovis, C., & Pepe, F. 2007, A&A, 468, 1115
Mahadevan, S., Ramsey, L. W., Terrien, R., et al. 2014, in Ground-based and Airborne Instrumentation for Astronomy V, Proc. SPIE, 9147, 91471G
Malavolta, L., Borsato, L., Granata, V., et al. 2017, AJ, 153, 224
Maldonado, J., Affer, L., Micela, G., et al. 2015, A&A, 577, A132
Owen, J. E., & Wu, Y. 2017, ApJ, 847, 29
Pepe, F., Molaro, P., Cristiani, S., et al. 2014, Astron. Nachr., 335, 8
Perger, M., García-Piquer, A., Ribas, I., et al. 2017, A&A, 598, A26
Quarles, B., Quintana, E. V., Lopez, E., Schlieder, J. E., & Barclay, T. 2017, ApJ, 842, L5
Quirrenbach, A., Amado, P. J., Caballero, J. A., et al. 2016, Ground-based and Airborne Instrumentation for Astronomy VI, Proc. SPIE, 9908, 990812
Rauer, H., Catala, C., Aerts, C., et al. 2014, Exp. Astron., 38, 249
Ricker, G. R., Winn, J. N., Vanderspek, R., et al. 2014, Space Telescopes and Instrumentation 2014: Optical, Infrared, and Millimeter Wave, Proc. SPIE, 9143, 20
Robertson, P., Endl, M., Cochran, W. D., & Dodson-Robinson, S. E. 2013, ApJ, 764, 3
Rogers, L. A. 2015, ApJ, 801, 41
Savanov, I. S. 2012, Astron. Rep., 56, 716
Selsis, F., Chazelas, B., Bordé, P., et al. 2007, Icarus, 191, 453
Shields, A. L., Ballard, S., & Johnson, J. A. 2016, Phys. Rep., 663, 1
Sinukoff, E., Howard, A. W., Petigura, E. A., et al. 2016, ApJ, 827, 78
Sozzetti, A., Bernagozzi, A., Bertolini, E., et al. 2013, in EPJ Web Conf., 47, 03006
Tuomi, M., Jones, H. R. A., Barnes, J. R., Anglada-Escudé, G., & Jenkins, J. S. 2014, MNRAS, 441, 1545
Van Eylen, V., Agentoft, C., Lundkvist, M. S., et al. 2017, MNRAS, submitted [arXiv:1710.05398]
Vanderburg, A., & Johnson, J. A. 2014, PASP, 126, 948
Vanderburg, A., Latham, D. W., Buchhave, L. A., et al. 2016a, ApJS, 222, 14
Vanderburg, A., Plavchan, P., Johnson, J. A., et al. 2016b, MNRAS, 459, 3565
Zechmeister, M., & Kürster, M. 2009, A&A, 496, 577
Zeng, L., & Sasselov, D. 2013, PASP, 125, 227
Zeng, L., Sasselov, D. D., & Jacobsen, S. B. 2016, ApJ, 819, 127
Zeng, L., Jacobsen, S. B., Hyung, E., et al. 2017a, in Lunar and Planetary Inst. Technical Report, Lunar and Planetary Science Conference, 48, 1576
Zeng, L., Jacobsen, S. B., & Sasselov, D. D. 2017b, Research Notes of the AAS, 1, 32

[1] INAF – Osservatorio Astrofisico di Torino, Via Osservatorio 20, 10025 Pino Torinese, Italy
e-mail: damasso@oato.inaf.it

[2] Observatoire de Genève, Université de Genève, 51 ch. des Maillettes, 1290 Sauverny, Switzerland

[3] Université Grenoble Alpes, CNRS, IPAG, 38000 Grenoble, France

[4] Dipartimento di Fisica e Astronomia "Galileo Galilei", Università di Padova, Vicolo dell'Osservatorio 3, 35122 Padova, Italy

[5] INAF – Osservatorio Astronomico di Padova, Vicolo dell'Osservatorio 5, 35122 Padova, Italy

[6] NASA Goddard Space Flight Center, 8800 Greenbelt Rd, Greenbelt, MD 20771, USA

[7] Department of Earth and Planetary Sciences, Harvard University, Cambridge, MA 02138, USA

[8] Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA

[9] Centre for Exoplanet Science, SUPA, School of Physics and Astronomy, University of St Andrews, St Andrews KY16 9SS, UK

[10] Department of Astronomy, The University of Texas at Austin, 2515 Speedway, Stop C1400, Austin, TX 78712, USA

[11] INAF – Osservatorio Astronomico di Palermo, Piazza del Parlamento 1, 90134 Palermo, Italy

[12] INAF – Osservatorio Astrofisico di Catania, Via S. Sofia 78, 95123 Catania, Italy

[13] INAF – Osservatorio Astronomico di Trieste, Via Tiepolo 11, 34143 Trieste, Italy

[14] INAF – Osservatorio Astronomico di Brera, Via E. Bianchi 46, 23807 Merate, Italy

[15] Centre for Star and Planet Formation, Niels Bohr Institute & Natural History Museum, University of Copenhagen, 1350 Copenhagen, Denmark

[16] INAF – Fundación Galileo Galilei, Rambla José Ana Fernandez Pérez 7, 38712 Breña Baja, Spain

[17] INAF – Osservatorio Astronomico di Capodimonte, Salita Moiariello 16, 80131 Napoli, Italy

[18] Department of Astronomy, University of California, Berkeley, CA 94720, USA

[19] European Southern Observatory, Alonso de Cordova 3107, Vitacura, Santiago, Chile

[20] Instituto de Astrofísica e Ciências do Espaço, Universidade do Porto, CAUP, Rua das Estrelas, 4150-762 Porto, Portugal

[21] Dipartimento di Fisica, Università di Roma Tor Vergata, Via della Ricerca Scientifica 1, 00133 Roma, Italy

[22] Max Planck Institute for Astronomy, Königstuhl 17, 69117 Heidelberg, Germany

[23] INAF – Osservatorio di Cagliari, via della Scienza 5, 09047 Selargius, CA, Italy

[24] Instituto de Astrofísica de Canarias (IAC), 38200 La Laguna, Tenerife, Spain

[25] Departamento de Astrofísica, Universidad de La Laguna (ULL), 38206 La Laguna, Tenerife, Spain

[26] SUPA, Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH93HJ, UK

[27] Departamento de Física e Astronomia, Faculdade de Ciências, Universidade do Porto, Rua do Campo Alegre, 4169-007 Porto, Portugal

[28] Astrophysics Research Centre, School of Mathematics and Physics, Queen's University Belfast, Belfast BT7 1NN, UK