A&A, 682, A136 (2024)
https://doi.org/10.1051/0004-6361/202347079
© The Authors 2024

Astronomy & Astrophysics

The GAPS Programme at TNG

LI. Investigating the correlations between transiting system parameters and host chromospheric activity

R. Claudi[1,2], G. Bruno[3], L. Fossati[4], A. F. Lanza[3], A. Maggio[5], G. Micela[5], J. Maldonado[5], S. Benatti[5], K. Biazzo[6], A. Bignamini[7], L. Cabona[1], I. Carleo[8,9], C. Danielski[10,11], S. Desidera[1], L. Malavolta[1,12], L. Mancini[13,14,9], M. Montalto[3], D. Nardiello[1], M. Rainer[15], G. Scandariato[3], A. Sozzetti[9], R. Cosentino[16], E. Covino[17], L. Di Fabrizio[16], A. Ghedina[16], V. Lorenzi[16], E. Molinari[18], M. Molinaro[7], I. Pagano[3], G. Piotto[12], and E. Poretti[15]

(Affiliations can be found after the references)

Received 2 June 2023 / Accepted 14 November 2023

ABSTRACT

Context. Stellar activity is the most relevant types of astrophysical noise that affect the discovery and characterization of extrasolar planets. On the other hand, the amplitude of stellar activity could hint at an interaction between the star and a close-in giant planet. Progress has been made in recent years in understanding how to deal with stellar activity and search for observational evidence of star-planet interactions.

Aims. The aim of this work is to characterize the chromospheric activity of stars hosting short-period exoplanets by studying the correlations between the chromospheric emission (CE) in the Ca II H&K and the planetary parameters.

Methods. We measured CE in the Ca II H&K lines using more than 1900 high-resolution spectra of a sample composed of 76 targets, observed with the HARPS-N spectrograph between 2012 and 2020. We transformed the fluxes into bolometric- and photospheric-corrected chromospheric emission ratios, $R'_{\rm HK}$. Furthermore, we completed the sample of hosts digging for data in previous works. Stellar parameters $T_{\rm eff}$, $B-V$, and V were retrieved homogeneously from the *Gaia* DR3. Then, $M_\star$, $R_\star$, and ages were determined from isochrone fitting. We retrieved planetary data from the literature and catalogs. The search for correlations between the $\log(R'_{\rm HK})$ and planetary parameters have been performed through both Spearman's rank and its statistics as well as the more sophisticated Gaussian mixture model method.

Results. We found that the distribution of $\log(R'_{\rm HK})$ for the transiting planet hosts is different from the distribution of field main-sequence and sub-giant stars. The $\log(R'_{\rm HK})$ of planetary hosts is correlated with planetary parameters proportional to the planetary radius to the power of n ($R_{\rm P}^n$), indicating a common origin for the correlations. The statistical analysis has also highlighted four clusters of host stars with different behavior in terms of their stellar activity with respect to the planetary surface gravity. Some of the host stars have a value of $\log(R'_{\rm HK})$ that is lower than the basal level of activity for main sequence stars. The planets of these systems are very close to filling their Roche lobe, suggesting that they evaporate through hydrodynamic escape under the strong irradiation of the host star, creating shrouds that absorb the core of the chromospheric resonance lines.

Key words. planets and satellites: fundamental parameters – planet-star interactions – stars: activity – planetary systems

1. Introduction

The majority of the 5000 exoplanets known thus far have been discovered thanks to transit and radial velocity (RV) methods[1] (Schneider et al. 2011). The exoplanet population displays a tremendous diversity in terms of masses, radii, and system architectures. The broad diversity among the extrasolar planets is also reflected in the range of atmospheric compositions that has been emerging from several studies on transmission and emission spectroscopy data (for a review, see, e.g., Madhusudhan 2019). The atmospheric chemical compositions of these planets are expected to provide information on different scenarios of planetary formation and evolution. A severe obstacle in gathering this information is presented by the stellar noise, mainly due to the chromospheric activity (CA) of stars. The induced stellar variability due to the interaction between the turbulent stellar plasma and the existent magnetic field causes modifications to the stellar spectral features that could be confused with planetary ones during a planetary transit or the observation of a phase curve (e.g., Ballerini et al. 2012; Robertson et al. 2015; Cegla et al. 2018; Barclay et al. 2021).

On the other hand, the stellar CA is not only a nuisance with the capability of hiding or mimicking the presence of a planetary companion and its atmospheric features; in fact, the enhancement of the CA of a stellar host is also considered to be an effect brought on by the presence of a hot-Jupiter orbiting very close to its host star via star-planet interactions (SPI) or it could be modulated or even affected by the presence of evaporating material shrouds lost by the close-in planet. The SPI can be induced by tidal and magnetic interactions (Cuntz et al. 2000). In the former case, planets can induce tidal bulges on the star with an interaction, depending on the semi-major axis of the planet, which could lead to an enhancement of the CA of the star ($\propto a_{\rm P}^{-3}$). In the latter case, the SPI is driven by the planetary magnetic field that interacts with the stellar magnetic field

---

[1] http://exoplanet.eu/

and the intensity of the interaction is $\propto a_P^{-2}$. Also, included in this case is a possible enhancement of the host CA. The first studies on the induced activity on planet hosts were performed by Bastian et al. (2000) in the radio wavelength range and by Saar & Cuntz (2001) in the optical, with poor results due to the insufficient sensitivity of their measurements. Later, Shkolnik et al. (2003, 2005, 2008), Kashyap et al. (2008), Scharf (2010), Gurdemir et al. (2012), Pillitteri et al. (2014), and Maggio et al. (2015) reported the detection of a significant increase of stellar activity in stars hosting hot Jupiters (hot gas-giant planets orbiting very close to their host stars). However, these studies were disputed by several authors, such as Poppenhaeger et al. (2010), Poppenhaeger & Schmitt (2011), Scandariato et al. (2013), and Miller et al. (2015), who attributed those detections to biases or selection effects. Furthermore, Canto Martins et al. (2011) searched for possible effects of SPI on the stellar chromosphere and did not find any significant correlations between the chromospheric activity (CA) indicator $\log(R'_{\rm HK})$ and planetary parameters.

Jupiter-like planets have a hydrogen-dominated atmosphere that evaporates when subjected to the EUV radiation and X-ray ($\lambda < 91.2$ nm) produced by their stellar host. In particular, the study of hot-Jupiters gave some evidence that mass loss can be a common phenomenon (Lammer et al. 2003; Lecavelier des Étangs et al. 2004; Lecavelier des Étangs 2007; Penz et al. 2008; Sanz-Forcada et al. 2010, 2011; Locci et al. 2019). A typical example is the extreme hot-Jupiter WASP-12 ($P_{\rm orb}$ = 1.09 d, Hebb et al. 2009), which is characterized by an exosphere surrounding the planet that fills its Roche lobe (Fossati et al. 2010; Haswell et al. 2012). The resulting mass loss seems to reinforce the circumstellar gas shroud that veils the host stars in some systems.

The most widely used CA indicator for F, G, and K stars is $\log(R'_{\rm HK})$, derived from the ratio between the Ca II H&K line-core chromospheric excess flux, evaluated by the Mount Wilson index, S, and the bolometric emission of the star (Noyes et al. 1984). Main sequence stars typically show $\log(R'_{\rm HK}) > -5.1$, a basal level exhibited by the quiet Sun. The determination of the $\log(R'_{\rm HK})$ index for the Sun has been discussed by, for instance, Egeland et al. (2017), while the recent work by Milbourne et al. (2019) shows that our star had a rotational modulation amplitude $\Delta \log(R'_{\rm HK}) \sim 0.015$, while the maximum variation in its chromospheric index was ~0.05 between July 2015 and October 2017, when the Sun was observed as a star by the HARPS-N solar telescope. Such amplitudes of variation appear to be broadly similar to those observed by Baliunas et al. (1995) in low-activity main sequence stars with a colour index comparable to the Sun over three decades. Therefore, they can be regarded as representative of the intrinsic variations in the $\log(R'_{\rm HK})$ index of quiet solar-like stars.

The anomalously low value of the WASP-12's Ca II H&K line-core flux ($\log(R'_{\rm HK}) \sim -5.5$) is interpreted as due to the presence of a circumstellar shroud of gas that is able to depress the stellar chromospheric flux to the observed level (Fossati et al. 2013). In subsequent works, other hosts of close-in planets were presented with values of $\log(R'_{\rm HK}) < -5.1$ and they all are likely viewed through shrouds of circumstellar gas coming from planetary atmosphere ablation (Staab et al. 2017; Haswell et al. 2020).

Following the study of Knutson et al. (2010) and considering a sample of 39 transiting planets, Hartman (2010) reported a correlation between the stellar CA ($\log(R'_{\rm HK})$) and the planetary surface gravity of transiting hot Jupiters (HJs) with a confidence level of 99.5%. Specifically, this author found that the chromospheric emission index is lower for stars with planets with lower surface gravity. In the following years, Figueira et al. (2014) reviewed the $\log(R'_{\rm HK}) - \log(g_{\rm P})$ correlation by collecting the available data on $\log(R'_{\rm HK})$ and stellar and planet properties and re-evaluated the correlation in the light of an extended dataset (108 transiting planets). A theoretical model was elaborated by Lanza (2014) to explain this correlation. In that model, the planetary material, ablated by the action of the UV radiation of the host star spreads towards the star along the magnetic field lines of the extended stellar corona in which the planet is lodged. The material is retained by the magnetic field lines forming prominence-like structures able to absorb at the core of chromospheric resonance lines (Mg II h&k, Ca II H&K). Strong absorption of the chromospheric resonance lines reduces the measured value of the $\log(R'_{\rm HK})$ indicator and the effect is stronger when the gravity of the planet is lower, explaining the correlation.

Exploiting a set of 41 transiting hosts, Fossati et al. (2015) improved the Lanza (2014) model by treating the sample of stars as a mixture of two subsets with intrinsic CA above or below the Vaughan-Preston gap (Vaughan & Preston 1980) observed for solar-like stars. Collier Cameron & Jardine (2018) highlighted an alternative explanation for the existence of the CA versus planetary surface gravity correlation, interpreting it as a direct consequence of the resulting distribution of orbital distances for transiting planets. Collier Cameron & Jardine (2018) found that the detection probabilities for transiting planets scale inversely with the increase of the planetary tidal migration rate, creating a statistical bias toward more massive planets being more likely to be observed around young and more active stars. Because the surface gravity of giant planets is mostly directly proportional to their mass, this hypothesis explains the correlation between CA and the surface gravity found by Hartman (2010) and confirmed by Figueira et al. (2014). However, the bimodality of the correlation proposed by Fossati et al. (2015) does not seem to be interpretable in the framework of this hypothesis.

In this paper, we evaluate $\log(R'_{\rm HK})$ for an extended sample of 76 transiting planet hosts observed over a time frame of about 9 yr by the Global Architecture of Planetary Systems (GAPS; Covino et al. 2013) collaboration. To this sample, we added original $S_{\rm HK}$ data from other 60 transiting hosts drawn from the literature. Finally, we evaluated, in a uniform way, the $\log(R'_{\rm HK})$ value for both samples to review several correlations between $\log(R'_{\rm HK})$ and planetary parameters using this extended dataset.

The observations and data of this study are described in Sect. 2. In Sect. 3, we provide a description and characterization of the host stars. In Sect. 4, we explain how we obtained the chromospheric index for each star and the method used to evaluate the same for the data obtained by previous works. In Sect. 5, we discuss the time variation in the CA for the GAPS targets and we put the transiting hosts into the context of the other stars in Sect. 6. Finally, in Sects. 7 and 8, we present our analysis. In Sect. 9, we give our summary and conclusions.

2. Observations and data

The GAPS project is an Italian collaboration aimed at exploiting the HARPS-N spectrograph (Cosentino et al. 2012) at the Telescopio Nazionale *Galileo* (TNG) to search for new planets in known planetary systems and around other peculiar stars (of low metallicity and M stars) by means of high precision radial velocity time series. The project began in 2012 and completed its first phase in 2017, after 5 yr of observations. The second phase of GAPS started in 2017, triggered by

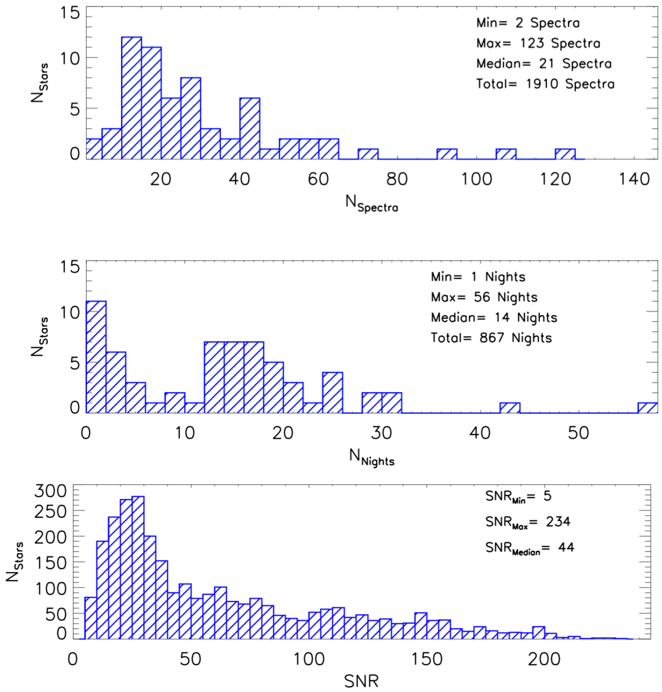

Fig. 1. GAPS observations of transiting hosts considered in this work. Top: number of spectra per star of the GAPS sample of transiting hosts. Middle: number of observing nights per star. Bottom: S/N value per star.

the implementation of the new TNG observing mode GIARPS (Claudi et al. 2016, 2017), which allows for the simultaneous use of HARPS-N and GIANO-B, covering both the optical (380–690 nm) and near-infrared (970–2945) wavelength ranges at the same time. In this second phase, GAPS has been devoted to studies of young planets (Carleo et al. 2018, 2020; Damasso et al. 2020; Benatti 2021; Nardiello et al. 2022) and planetary atmospheres (Borsa et al. 2019; Pino et al. 2020; Guilluy et al. 2020; Carleo et al. 2022), and this project is still ongoing. In both phases of GAPS, planetary systems with transiting planets were observed, in some cases collecting quite long time series of spectral data that have also proven useful in characterizing the CA level of such targets. In this study, we consider 76 transiting planet hosts observed in the past years by GAPS, to which we added 19 objects available from Knutson et al. (2010), and Staab et al. (2017), along with another 41 objects taken from Figueira et al. (2014). The complete sample contains 136 objects.

2.1. GAPS sample

The 76 stars in the original GAPS sample were regularly monitored with the spectrograph HARPS-N since 2012, collecting data for about five to 8 yr, in some cases. In the following, we discarded four objects (KELT-7, KELT-9, KELT-20, and WASP-33) for the reasons explained in Sect. 4. The journal of the observations for the remaining 72 objects is reported in Table A.1, including the date, T_0, in which the target has been observed for the first time, the total time span in days covered by GAPS data, and the number of available HARPS-N spectra. Figure 1 shows the distribution of spectra per star (top panel) and the distribution of the number of observing nights per star (central panel). For the transiting planets hosts, we have a total of 1910 spectra gathered in 867 observing nights, with a minimum of two spectra for WASP-3 and the highest number of spectra (123) for HD189733. The median is 21 spectra per star. The high number of spectra obtained for some stars is due to the frequent observing cadence during the planetary transit night we observed to detect the Rossiter-McLaughlin effect on the radial velocity curves (e.g., Esposito et al. 2017). This is the case of WASP-11 with 26 spectra taken in one night, while the maximum number of nights has been devoted to XO-2N observed for 56 nights in about 3 yr.

The main aim of the GAPS observations is to obtain precise radial velocity (RV) measurements for new exoplanet detections. To achieve this goal, in our observations, we exploited the simultaneous reference technique using both optical fibers: the target fed the first fiber, while the second fiber was illuminated by a stable calibration source, either ThAr or Fabry Perot lamp[2]. If the stars are fainter than 11 mag, we used the second fiber instead for sky subtraction to avoid lamp pollution. The typical exposure time was 900 s to average out solar-like pulsations. The median value of the V magnitude of GAPS targets is 11.2 mag, so the median value of the reached signal-to-noise ratio (S/N) is 44 (see the bottom panel of Fig. 1). For the brighter tail of the target distribution, the single observations can reach $S/N > 100$ per pixel at 550 nm, to achieve a radial velocity precision of $1\,\mathrm{m\,s^{-1}}$ or better, depending on the airmass of the observation and the weather conditions. The spectra were reduced through the HARPS-N data reduction software (DRS), including the radial velocity extraction based on the cross-correlation function (CCF) method (Pepe et al. 2002), where the observed spectrum is correlated with a binary mask matching the spectral type of each object. With a dedicated pipeline implemented at the INAF Trieste Observatory[3] through the YABI workflow interface (Hunter et al. 2012), for the GAPS component of the sample, we also evaluated the CA index from the H&K lines of Ca II (S-index and $\log(R'_{\mathrm{HK}})$), through the procedure described in Sect. 4.

2.2. Objects from previous similar works

We gathered other transiting planet hosts for which S-indices and/or $\log(R'_{\mathrm{HK}})$ values are available in the literature (see Table 1).

We did not take into consideration the set from Canto Martins et al. (2011) because the data are not homogeneous with the previous ones. Furthermore, those objects are not transiting their stars and the measured planetary masses are just minimum masses.

In summary, we selected from Knutson et al. (2010) 18 objects (with the measured S-index value) that are not included in the list of GAPS targets and another 40 objects (with no S-index) from Figueira et al. (2014) that are not included in either GAPS or Knutson's lists. Furthermore, we take into consideration the other two objects (WASP-72 and WASP-103) from Staab et al. (2017) for which the S-index values are given. Finally, the sample of 72 GAPS objects includes 43 stars already listed in previous works and 29 new entries. The complete target list considered in this study consists of 132 objects.

For homogeneity, we would like to evaluate the $\log(R'_{\mathrm{HK}})$ from the original value of the S-index, using the stellar parameters that we obtained for all systems by the *Gaia* DR3 (see Sect. 3). Generally, the discovery papers reported in a few cases the complete time series of the S_{HK}. This was the case

---

[2] Since mid-2016, the Fabry Perot lamp has been operative, replacing the ThAr lamp due to a sharp decrease in the lamp flux.

[3] http://ia2.inaf.it

Table 1. Compilations of $\log(R'_{\rm HK})$ measurements for planetary host stars.

Compilation by	Obj.	$\log(R'_{\rm HK})_{\rm MAX}$	Host	$\log(R'_{\rm HK})_{\rm MIN}$	Host	Note
Knutson et al. (2010)	50	−4.331	CoRoT-2	−5.500	WASP-12	a
Hartman (2010)	39	−4.331	CoRoT-2	−5.500	WASP-12	b
Figueira et al. (2014)	95	−4.100	WASP-59	−5.500	WASP-12	c
Fossati et al. (2015)	41	−4.100	WASP-59	−5.319	HAT-P-13	d
Canto Martins et al. (2011)	74	−4.42	HD128311	−5.22	HD4203	e

Notes. Columns indicate the compilation origin, number of objects, maximum and the minimum values, and the corresponding objects, from left to right. The last column gives explicative notes on each compilation. a: S-index available; b: subset of Knutson et al. (2010) compilation. No error is given; c: set of Hartman (2010) plus other multi planetary systems; d: same list of Figueira et al. (2014) but trimmed for systems with a ≤0.1 au and $M_{\rm p} \geq 0.1\ M_{\rm J}$; e: no transiting systems.

of the following systems HAT-P-27, HAT-P-28, HAT-P-33, HAT-P-34, HAT-P-35, HAT-P-39, HAT-P-40, HAT-P-41, HAT-P-44, HAT-P-45, HAT-P-46, and HD97658 – for which we evaluate the mean value of $S_{\rm HK}$. Instead, for the following systems, 55 Cnc, HAT-P-25, HAT-P-38, Kepler-19, Kepler-62, Kepler-68, WASP-22, and WASP-23, the bibliography reported only a single value of $S_{\rm HK}$. For the remaining systems, when possible, we evaluated the $S_{\rm HK}$ value from the available values of $\log(R'_{\rm HK})$ and the $B-V$ used by the paper authors. This was the case of: Kepler-10, Kepler-17, Kepler-78, WASP-5, WASP-15, WASP-58, WASP-70 A, and WASP-84. Not all the papers outline the value of $B-V$ used in evaluating the $\log(R'_{\rm HK})$, but most of them report the $T_{\rm eff}$. We used the latter parameter to obtain the value of $B-V$ and then the value of $S_{\rm HK}$ from Modern Mean Dwarf Stellar Color and Effective Temperature Sequence by Mamajek (2021, Pecaut & Mamajek 2013)[4]. These systems are: Kepler-25, Kepler-48, Kepler-93, WASP-7, WASP-16, WASP-41, WASP-42, WASP-52, WASP-117, and Kepler-16. Due to the binarity of the latter system, the value of $S_{\rm HK}$ obtained is pretty uncertain. Finally, for Kepler-20, we obtain the value of $B-V$ from Zacharias et al. (2012). All the hosts are listed in Table 2 with the bibliography used.

3. Stellar parameters

We used *Gaia*/DR3 (Gaia Collaboration 2016, 2018, 2023) to obtain the values of the G magnitude and Bailer-Jones et al. (2021) for the distance of all the targets in our sample[5]. We found that only four stars of our sample (55 Cnc, HD 189733, HD 209458, and HD 97658) require the saturation correction described by Riello et al. (2021) for the G mag. Once applied, the correction for G is 0.002 mag for all four stars. The retrieved distances range from a minimum of 9.77 pc (GJ 436) to a maximum of 990.44 pc (Kepler-8) with a median of 221.12 pc (see the upper panel of Fig. 2). Thus, 85% of the sample is at a distance larger than 100 pc from the Sun.

We used the procedure exploited in developing the PLATO Input Catalog (PIC, Montalto et al. 2021) to retrieve the stellar parameters of the hosts homogeneously. We derived the $(B-V)_0$ colors from the intrinsic $G_{\rm BP} - G_{\rm RP}$ *Gaia* DR3 color by inverting the cubic equation reported in Table 2 of Evans et al. (2018). The reddening correction was obtained by interpolating the reddening map of Lallement et al. (2019, 2022), using the distances from Bailer-Jones et al. (2021). The effective temperatures of the stars were obtained from our custom-calibrated *Gaia* intrinsic color-effective temperature relationship used for the PIC. The error in the color (median uncertainty of 3.6%) accounts for the reddening uncertainties as deduced from the reddening map, while the error on the temperature is derived from a Monte Carlo simulation, randomizing the input color within its uncertainties assuming normal error distribution and repeating the temperature calculation as described above. A normal random perturbation with a standard deviation equal to 200 K is added to the Monte Carlo temperature estimates to account for the expected difference between our temperature scale and literature results. All the details about this procedure are described in Montalto et al. (2021). In this way, we have a homogeneous determination of the $T_{\rm eff}$ for most of the targets in the sample, with a trade-off of a median uncertainty of 3.5%. This is about four times the relative error for the determination of the $T_{\rm eff}$ listed in the SWEET Catalog[6] (Santos et al. 2013; Sousa et al. 2018, 2021), but as these authors warn, these data are not homogeneous. In the catalog, only those systems with the `SWFlag=` 1 can identify the stars with derived spectroscopic parameters in a homogeneous way by Sousa et al. (2021). Furthermore, we transformed the *Gaia* G magnitude to Johnson V magnitude using the relation by Evans et al. (2018). The middle and bottom panels of Fig. 2 show the distribution of the apparent V magnitude and of the $T_{\rm eff}$ of the whole sample, respectively. The stars have apparent magnitude ranging $5.9 \leq V \leq 14.3$ mag, with a median of 11.6 mag, while the temperature range is $3461 \leq T_{\rm eff} \leq 6781$ K. We evaluated the spectral type of each member of the sample by exploiting the table from Pecaut & Mamajek (2013). The sample comprises 49 F stars, 42 G stars, 39 K stars, and 2 M stars.

For a small subset of our sample (27 stars), we considered the metallicity homogeneously derived through the methodology described in Biazzo et al. (2022) and based on iron lines. For the other hosting stars, since no homogeneous metallicities were obtained in the literature, we considered the values listed in SWEET-Cat and derived through similar spectroscopic methods identified by `SWFlag=` 1 (see, e.g., Sousa et al. 2021). Figure 3 shows the distribution of [Fe/H] spanning between −0.30 and 0.48 dex with a median value of 0.010 dex and a median error of 0.04 dex.

We retrieved the values of the mass, the radius, log g, and age of the stars through the PARAM web interface[7], which makes use of a Bayesian method, described in da Silva et al. (2006) applied on the theoretical isochrone set PARSEC (Bressan et al. 2012). The input parameters to the PARAM tool were the apparent magnitude V, the parallax, the $T_{\rm eff}$, and [Fe/H].

---

[4] `https://www.pas.rochester.edu/~emamajek/`

[5] Using SIMBAD (Wenger et al. 2000) and VIZIER (Ochsenbein et al. 2000).

[6] `http://sweetcat.iastro.pt/`

[7] `http://stev.oapd.inaf.it/cgi-bin/param_1.3`

Table 2. For all systems, the table lists *Gaia* DR3 id., and $(B-V)$, and $T_{\rm eff}$ obtained by Gaia Collaboration (2018) as explained in the text.

System	*Gaia* DR3 Id.	$(B-V)$	$T_{\rm eff}$ (K)	$S_{\rm MWO}$	$\log(R'_{\rm HK})_N$	$\log(R'_{\rm HK})_M$	$R'_{\rm HK} \times 10^5$ rms
55Cnc	704967037090946688	0.841 ± 0.029	5214 ± 217	$0.165^{(1)}$	−5.0308	−4.8557	–
CoRoT-1	3105507886130792448	0.554 ± 0.021	6066 ± 205	$0.124^{(2)}$	−5.2761	−5.2975	–
CoRoT-2	4287820848378092672	0.821 ± 0.028	5265 ± 192	$0.435^{(2)}$	−4.4536	−4.2139	–
CoRoT-7	3107267177757848576	0.824 ± 0.028	5256 ± 187	$0.225^{(2)}$	−4.8207	−4.6104	–
GJ3470$^{(a)}$	654687847820642560	1.917 ± 0.014	3558 ± 204	1.755 ± 0.321	−5.0625 ± 0.0843	−4.6696 ± 0.0852	0.158
GJ436$^{(a)}$	4017860992519744384	2.072 ± 0.014	3461 ± 208	0.735 ± 0.060	−5.3263 ± 0.0347	−5.3727 ± 0.0455	0.038
HAT-P-1	1928431764627661440	0.612 ± 0.023	5869 ± 199	0.151 ± 0.014	−5.0408 ± 0.1151	−4.8964 ± 0.1404	0.220
HAT-P-2$^{(a,b)}$	1380825667768742144	0.447 ± 0.019	6459 ± 202	0.181 ± 0.006	−4.7539 ± 0.0271	−4.6819 ± 0.0388	0.111
HAT-P-3	1510191594552968832	0.903 ± 0.031	5062 ± 185	0.220 ± 0.036	−4.8967 ± 0.1287	−4.6748 ± 0.1901	0.266
HAT-P-4	1291120362349158016	0.609 ± 0.023	5880 ± 203	0.147 ± 0.016	−5.0702 ± 0.1058	−4.8645 ± 0.1062	0.242
HAT-P-5	4606030169272920320	0.622 ± 0.023	5838 ± 207	$0.148^{(2)}$	−5.0683	−4.9209	–
HAT-P-6$^{(a)}$	1925321658551399040	0.378 ± 0.017	$6570 \pm 80^{(3)}$	0.181 ± 0.032	−4.6872 ± 0.1006	−4.6464 ± 0.1525	0.648
HAT-P-7	2129256395211984000	0.469 ± 0.019	6376 ± 204	0.140 ± 0.004	−5.0318 ± 0.0390	−5.0585 ± 0.0744	0.078
HAT-P-8	1891507552826485632	0.521 ± 0.020	6181 ± 171	0.148 ± 0.004	−5.0072 ± 0.0276	−4.9462 ± 0.0395	0.062
HAT-P-9	898130030131443584	0.510 ± 0.020	6220 ± 204	$0.141^{(2)}$	−5.0567	−5.0381	–
HAT-P-11	2086512227851023872	1.075 ± 0.036	4681 ± 229	0.544 ± 0.010	−4.6808 ± 0.0081	−4.3017 ± 0.009	0.039
HAT-P-12	1499514786891168640	1.110 ± 0.037	4612 ± 213	0.362 ± 0.063	−4.9148 ± 0.0789	−4.5383 ± 0.0946	0.218
HAT-P-13	1014520826353577088	0.699 ± 0.025	5602 ± 193	$0.141^{(2)}$	−5.1400	−4.8945	–
HAT-P-14	1340497608486742400	0.452 ± 0.019	6439 ± 185	0.164 ± 0.003	−4.8479 ± 0.0184	−4.8152 ± 0.0295	0.058
HAT-P-15	179498266829041664	0.910 ± 0.031	5043 ± 204	0.182 ± 0.019	−5.0100 ± 0.0616	−4.7565 ± 0.0708	0.134
HAT-P-16	381592313648387200	0.548 ± 0.021	6085 ± 190	0.164 ± 0.060	−4.9157 ± 0.1950	−4.7965 ± 0.2226	1.036
HAT-P-17	1849786481032616960	0.824 ± 0.028	5257 ± 191	0.156 ± 0.042	−5.0629 ± 0.1392	−4.8963 ± 0.2363	0.391
HAT-P-18	1334573817793362560	1.045 ± 0.035	4744 ± 216	0.334 ± 0.079	−4.8553 ± 0.1295	−4.5282 ± 0.164	0.349
HAT-P-20	869913435026514688	1.203 ± 0.040	4436 ± 213	1.142 ± 0.107	−4.5609 ± 0.0389	−4.0656 ± 0.0409	0.259
HAT-P-21	770622651659107712	0.687 ± 0.025	5636 ± 200	0.306 ± 0.044	−4.5313 ± 0.0713	−4.3084 ± 0.0703	0.588
HAT-P-22	846946629987527168	0.812 ± 0.028	5290 ± 183	0.157 ± 0.016	−5.0590 ± 0.0702	−4.819 ± 0.0727	0.161
HAT-P-24	3167323052618369408	0.470 ± 0.019	6370 ± 197	0.146 ± 0.011	−4.9793 ± 0.1005	−4.9962 ± 0.2171	0.213
HAT-P-25	111322601672419712	0.854 ± 0.029	5182 ± 179	$0.166^{(4)}$	−5.0322	−4.8562	–
HAT-P-26	3668036348641580288	0.901 ± 0.031	5066 ± 246	0.172 ± 0.079	−5.0372 ± 0.2043	−4.8528 ± 0.2747	0.581
HAT-P-27$^{(c)}$	1159336403336463872	0.865 ± 0.030	5153 ± 202	$0.231 \pm 0.017^{(5)}$	−4.8368	−4.6176	–
HAT-P-28	363702817083391232	0.821 ± 0.028	5265 ± 224	$0.163 \pm 0.013^{(6)}$	−5.0318 ± 0.0690	−4.7908 ± 0.172	–
HAT-P-29	359058441314838528	0.595 ± 0.022	5925 ± 227	0.160 ± 0.020	−4.9705 ± 0.1778	−4.8274 ± 0.147	0.316
HAT-P-30	3096441729861716224	0.518 ± 0.020	6191 ± 195	0.142 ± 0.012	−5.0585 ± 0.2019	−5.0176 ± 0.0975	0.211
HAT-P-31	4583004815239650176	0.608 ± 0.022	5884 ± 231	0.130 ± 0.011	−5.2268 ± 0.1293	−5.0236 ± 0.1308	0.174
HAT-P-32	356348286886230272	0.469 ± 0.019	6376 ± 219	0.228 ± 0.010	−4.5879 ± 0.0327	−4.4472 ± 0.0397	0.194
HAT-P-33$^{(a)}$	893550942158776832	0.429 ± 0.018	6533 ± 226	$0.172 \pm 0.001^{(7)}$	−4.7806 ± 0.0230	−4.7494 ± 0.084	–
HAT-P-34$^{(a)}$	1810218734058821632	0.423 ± 0.018	6555 ± 210	$0.174 \pm 0.002^{(8)}$	−4.7656 ± 0.0290	−4.7349 ± 0.088	–
HAT-P-35	3094733088793018240	0.554 ± 0.021	6064 ± 182	$0.129 \pm 0.002^{(8)}$	−5.2118 ± 0.0370	−5.1918 ± 0.087	–
HAT-P-36	1541532207133249920	0.706 ± 0.025	5582 ± 189	0.263 ± 0.028	−4.6378 ± 0.0671	−4.4364 ± 0.0699	0.364
HAT-P-38	324889227693298560	0.842 ± 0.029	5211 ± 141	$0.144^{(9)}$	−5.1290	−4.9821	–
HAT-P-39	3169391577586325120	0.470 ± 0.019	6371 ± 227	$0.179 \pm 0.002^{(10)}$	−4.7795 ± 0.0250	−4.6761 ± 0.089	–
HAT-P-40	1962153854973972096	0.600 ± 0.022	5910 ± 188	$0.145 \pm 0.001^{(10)}$	−5.0794 ± 0.0200	−4.8822 ± 0.076	–
HAT-P-41	4290415081653653632	0.428 ± 0.018	6534 ± 242	$0.148 \pm 0.002^{(10)}$	−4.9237 ± 0.0360	−5.0068 ± 0.096	–
HAT-P-44	1506159170017461888	0.824 ± 0.028	5258 ± 212	$0.125 \pm 0.011^{(11)}$	−5.2453 ± 0.0730	−5.1386 ± 0.213	–
HAT-P-45	4270547520706275328	0.483 ± 0.019	6322 ± 200	$0.121 \pm 0.008^{(11)}$	−5.2540	−5.1684	–
HAT-P-46	4177560894633502720	0.651 ± 0.065	5826 ± 322	0.127 ± 0.004	−5.2756 ± 0.0490	−5.173 ± 0.182	–
HAT-P-47$^{(a)}$	131624808999798656	0.449 ± 0.019	6451 ± 215	0.177 ± 0.013	−4.7751 ± 0.1040	−4.71 ± 0.0564	0.247
HD 149026	1331356474971716992	0.566 ± 0.021	6024 ± 190	$0.152^{(2)}$	−5.0048	−4.8367	–
HD 17156	545560867790611072	0.584 ± 0.022	5961 ± 201	0.146 ± 0.022	−5.0654 ± 0.1362	−4.883 ± 0.1447	0.353
HD 189733	1827242816201846144	0.921 ± 0.031	5017 ± 186	0.521 ± 0.037	−4.4815 ± 0.0480	−4.1955 ± 0.0545	0.257
HD 209458	1779546757669063552	0.553 ± 0.021	6066 ± 180	0.156 ± 0.003	−4.9685 ± 0.0210	−4.8578 ± 0.0261	0.054
HD 80606	1019003226022657920	0.737 ± 0.026	5492 ± 203	0.149 ± 0.012	−5.0887 ± 0.1070	−4.9221 ± 0.1321	0.148
HD 97658	3997075206232885888	0.848 ± 0.029	5197 ± 231	$0.175 \pm 0.001^{(12)}$	−4.9943 ± 0.0190	−4.8099 ± 0.122	–
KELT-3$^{(a)}$	806492023789218688	0.464 ± 0.019	6393 ± 198	0.189 ± 0.009	−4.7287 ± 0.0361	−4.6289 ± 0.0471	0.170
KELT-6	1464700950221781504	0.494 ± 0.020	6280 ± 215	0.151 ± 0.003	−4.9580 ± 0.0217	−4.9219 ± 0.0323	0.055

Notes. The Ca II H&K line strength is listed for both the original GAPS sample and the sample obtained by literature. We report in the $\log(R'_{\rm HK})_{\rm N}$ column the classical value (Noyes et al. 1984) of the activity as obtained in this work or in the literature while in the $\log(R'_{\rm HK})_{\rm M}$ the values corrected with the Mittag et al. (2013) procedure as described in Sect. 4.2. In the last column, the rms of the several measured time series are reported. $^{(a)}$The R-indices obtained with the relation of Mittag et al. (2013) are derived in the color range of $0.44 < B-V < 1.6$. Nevertheless, we give values for all the objects in our samples, and values for spectral types out of this color range should be taken with a pinch of salt. $^{(b)}$Also known as HD147506. $^{(c)}$Also known as WASP-40. $^{(d)}$Also known as HAT-P-10.

References. (1) Wright et al. (2004); (2) Knutson et al. (2010); (3) Noyes et al. (2008); (4) Quinn et al. (2012); (5) Béky et al. (2011); (6) Buchhave et al. (2011); (7) Hartman et al. (2011); (8) Bakos et al. (2012); (9) Sato et al. (2012); (10) Hartman et al. (2012); (11) Hartman et al. (2014); (12) Howard et al. (2011); (13) Dumusque et al. (2014); (14) Winn et al. (2011); (15) Bonomo et al. (2012); (16) Malavolta et al. (2017); (17) Gautier et al. (2012); (18) Borucki et al. (2012); (19) Marcy et al. (2014); (20) Borucki et al. (2013); (21) Gilliland et al. (2013); (22) Pepe et al. (2013); (23) Triaud et al. (2010); (24) Albrecht et al. (2012); (25) Brown et al. (2012); (26) Anderson et al. (2011); (27) Triaud et al. (2011); (28) Maxted et al. (2011); (29) Lendl et al. (2012); (30) Hébrard et al. (2013); (31) Anderson et al. (2014); (32) Staab et al. (2017); (33) Lendl et al. (2014); (34) Tsantaki et al. (2014).

Table 2. continued.

System	*Gaia* DR3 Id.	$(B-V)$	$T_{\rm eff}$ (K)	$S_{\rm MWO}$	$\log(R'_{\rm HK})_N$	$\log(R'_{\rm HK})_M$	$R'_{\rm HK} \times 10^5$ rms
Kepler-4	2132152916856093952	0.637 ± 0.023	5792 ± 204	0.168(2)	-4.9437	-4.7219	–
Kepler-5	2079018300195390464	0.605 ± 0.022	5892 ± 240	0.148(2)	-5.0584	-4.856	–
Kepler-6	2086636884980514304	0.715 ± 0.025	5556 ± 215	0.160(2)	-5.0160	-4.7755	–
Kepler-7	2102117871259036672	0.561 ± 0.021	6042 ± 202	0.155(2)	-4.9801	-4.8148	–
Kepler-8	2116730994965905280	0.532 ± 0.021	6141 ± 225	0.153(2)	-4.9764	-4.8897	–
Kepler-10	2132155017099178624	0.667 ± 0.024	5697 ± 189	0.164(13)	-4.9782	-4.8045	–
Kepler-16	2133476355197071616	1.373 ± 0.015	4111 ± 186	0.913(14)	-4.9293	-4.7535	–
Kepler-17	2086449761846310784	0.703 ± 0.025	5591 ± 201	0.420(15)	-4.3611	-4.1517	–
Kepler-19	2051106987063242880	0.724 ± 0.026	5530 ± 216	0.164 ± 0.011(16)	-4.9951 ± 0.0680	-4.8165 ± 0.169	–
Kepler-20	2102548708017562112	0.750 ± 0.026	5456 ± 215	0.183 ± 0.005(17)	-4.9173 ± 0.0340	-4.7287 ± 0.131	–
Kepler-22	2127941757262806656	0.687 ± 0.025	5637 ± 188	0.149 ± 0.004(18)	-5.0792 ± 0.0320	-4.913 ± 0.11	–
Kepler-25	2100451630105041152	0.502 ± 0.020	6251 ± 213	0.121(19)	-5.2757	-5.1947	–
Kepler-48	2075112109039378688	0.894 ± 0.030	5082 ± 223	0.226(19)	-4.8733	-4.6502	–
Kepler-62	2107001283431633408	0.953 ± 0.032	4945 ± 154	0.212 ± 0.002(20)	-4.9622 ± 0.0340	-4.731 ± 0.104	–
Kepler-68	2129550445852902656	0.634 ± 0.023	5799 ± 196	0.139(21)	-5.1500	-5.0069	–
Kepler-78	2078373642776670080	0.918 ± 0.031	5025 ± 178	0.467(22)	-4.5297	-4.2495	–
Kepler-93	2052747119115620352	0.689 ± 0.025	5630 ± 230	0.165(19)	-4.9801	-4.8031	–
Qatar-1	2244830490514284928	1.031 ± 0.035	4774 ± 211	0.597 ± 0.079	-4.5713 ± 0.0594	-4.2177 ± 0.065	0.368
Qatar-2	3620030644476623616	1.194 ± 0.040	4451 ± 186	1.122 ± 0.696	-4.5543 ± 0.2054	-4.0655 ± 0.2157	1.736
TrES-1	2098964849867337856	0.831 ± 0.029	5239 ± 235	0.229 ± 0.042	-4.8148 ± 0.1275	-4.6021 ± 0.1516	0.388
TrES-2	2131314401800665344	0.634 ± 0.023	5800 ± 198	0.156 ± 0.015	-5.0150 ± 0.1294	-4.8531 ± 0.1552	0.222
TrES-3	4609131509318715136	0.732 ± 0.026	5507 ± 214	0.305(2)	-4.5663	-4.3571	–
TrES-4	4609062308806929152	0.507 ± 0.020	6232 ± 211	0.137 ± 0.008	-5.0888 ± 0.0757	-5.0365 ± 0.1156	0.134
WASP-1	2862548428079638912	0.570 ± 0.022	6011 ± 213	0.136 ± 0.071	-5.1429 ± 0.3168	-4.9851 ± 0.3438	1.186
WASP-2	1748596020745038208	0.892 ± 0.030	5087 ± 220	0.159(2)	-5.0789	-4.9109	–
WASP-3	2095108312831835648	0.484 ± 0.019	6317 ± 218	0.147 ± 0.006	-4.9875 ± 0.0513	-4.9795 ± 0.0889	0.112
WASP-4	6535499658122055552	0.753 ± 0.026	5448 ± 210	0.194(2)	-4.8750	-4.6819	–
WASP-5	6533307957785865856	0.679 ± 0.024	5661 ± 196	0.224(23)	-4.7297	-4.5368	–
WASP-7	6681720724498802176	0.449 ± 0.019	6453 ± 196	0.447(24)	-4.1655	-3.9722	–
WASP-10	1909762228985058944	1.083 ± 0.036	4666 ± 209	0.145 ± 0.074	-5.2953 ± 0.1494	-5.0891 ± 0.2144	0.285
WASP-11 (d)	123376685084303360	0.997 ± 0.033	4846 ± 208	0.295 ± 0.053	-4.8489 ± 0.0655	-4.56 ± 0.076	0.281
WASP-12	3435282862461427072	0.570 ± 0.022	6009 ± 218	0.112 ± 0.059	-5.5035 ± 0.4301	-5.3623 ± 0.5349	0.989
WASP-13	702295464353437952	0.591 ± 0.022	5940 ± 172	0.140 ± 0.008	-5.1172 ± 0.0905	-4.9306 ± 0.0966	0.127
WASP-14	1242084170974175232	0.447 ± 0.019	6458 ± 200	0.152 ± 0.004	-4.9130 ± 0.0293	-4.9334 ± 0.0555	0.077
WASP-15	6171009049549197824	0.491 ± 0.020	6289 ± 198	0.165 ± 0.015(23)	-4.8704 ± 0.1040	-4.7925 ± 0.197	–
WASP-16	6283723285046532864	0.704 ± 0.025	5587 ± 178	0.145(25)	-5.1102	-4.9466	–
WASP-17 (a)	6042793005779654656	0.345 ± 0.016	6561 ± 287	0.121(2)	-5.0183	-4.8822	–
WASP-18	4955371367334610048	0.479 ± 0.019	6337 ± 238	0.116(2)	-5.3271	-5.2571	–
WASP-19	5411736896952029568	0.775 ± 0.027	5386 ± 188	0.252(2)	-4.7152	-4.5051	–
WASP-21	2831084391023184128	0.637 ± 0.023	5789 ± 214	0.160 ± 0.007	-4.9904 ± 0.0432	-4.8248 ± 0.048	0.099
WASP-22	5086537022856406272	0.581 ± 0.022	5972 ± 165	0.170 ± 0.010(26)	-4.8985	-4.7533	–
WASP-23	5557345496687437696	0.895 ± 0.030	5080 ± 212	0.320 ± 0.040(27)	-4.6895 ± 0.0960	-4.4376 ± 0.191	–
WASP-24	1153682508388170112	0.542 ± 0.021	6105 ± 179	0.159 ± 0.015	-4.9446 ± 0.1074	-4.8376 ± 0.143	0.260
WASP-26	2416782701664155008	0.604 ± 0.022	5897 ± 192	0.147 ± 0.009	-5.0687 ± 0.0777	-4.9361 ± 0.095	0.139
WASP-31	3545873945303316224	0.518 ± 0.020	6192 ± 238	0.150 ± 0.023	-4.9873 ± 0.1462	-4.9222 ± 0.2079	0.411
WASP-32	2546413408888429696	0.563 ± 0.021	6034 ± 164	0.175 ± 0.023	-4.8599 ± 0.1110	-4.7194 ± 0.1305	0.385
WASP-35	3211188618762023424	0.608 ± 0.022	5884 ± 180	0.157 ± 0.036	-4.9917 ± 0.1411	-4.8433 ± 0.1671	0.571
WASP-38	4453211899986180352	0.527 ± 0.020	6161 ± 200	0.152 ± 0.003	-4.9801 ± 0.0232	-4.901 ± 0.0325	0.052
WASP-39	3643098875168270592	0.781 ± 0.027	5371 ± 240	0.172 ± 0.035	-4.9778 ± 0.1726	-4.795 ± 0.2549	0.372
WASP-41	6160877943251930368	0.764 ± 0.027	5416 ± 191	0.261(28)	-4.6853	-4.4754	–
WASP-42	6139698733664726784	0.955 ± 0.032	4940 ± 185	0.202(29)	-4.9899	-4.7652	–
WASP-43	3767805209112436736	1.360 ± 0.015	4162 ± 204	1.618 ± 0.125	-4.6610 ± 0.0341	-4.0522 ± 0.0349	0.849
WASP-48	2141754578242371584	0.520 ± 0.020	6184 ± 267	0.169 ± 0.011	-4.8648 ± 0.0601	-4.724 ± 0.0713	0.192
WASP-49	2991284369063612928	0.728 ± 0.026	5517 ± 188	0.158 ± 0.012	-5.0330 ± 0.0633	-4.8587 ± 0.071	0.142
WASP-50	5160557726183065984	0.771 ± 0.027	5399 ± 180	0.161 ± 0.002	-5.0260 ± 0.0108	-4.8508 ± 0.0125	0.842
WASP-52	2666015878575546496	0.914 ± 0.031	5034 ± 191	0.585(30)	-4.4180	-4.1318	–
WASP-54	3661983850663908608	0.541 ± 0.021	6111 ± 192	0.141 ± 0.010	-5.0825 ± 0.1189	-4.9593 ± 0.1665	0.173
WASP-57	6340377202215639936	0.668 ± 0.024	5695 ± 216	0.213 ± 0.046	-4.7592 ± 0.1804	-4.5697 ± 0.1939	0.639
WASP-58	2115245554756763392	0.593 ± 0.022	5932 ± 207	0.347(30)	-4.3887	-4.1979	–
WASP-59	2841043011433969152	1.280 ± 0.014	4270 ± 193	0.707 ± 0.103	-4.8952 ± 0.0660	-4.3494 ± 0.0719	0.185
WASP-60	2868528637464028160	0.641 ± 0.023	5779 ± 209	0.147 ± 0.009	-5.0877 ± 0.0704	-4.8597 ± 0.0675	0.134
WASP-69	6910753016653587840	1.033 ± 0.035	4769 ± 228	0.671 ± 0.013	-4.5221 ± 0.0082	-4.1625 ± 0.0089	0.057
WASP-70A	6888469489370835840	0.650 ± 0.024	5750 ± 220	0.132(31)	-5.2200	-5.0829	–
WASP-72	5065640460769428224	0.545 ± 0.021	6097 ± 164	0.124 ± 0.009(32)	-5.2701 ± 0.1430	-5.1909 ± 0.205	–
WASP-76	2512326349403275520	0.532 ± 0.021	6141 ± 211	0.140 ± 0.009	-5.0802 ± 0.0696	-4.9708 ± 0.0855	0.158
WASP-80	4223507222112425344	1.505 ± 0.016	3971 ± 199	1.773 ± 0.087	-4.8242 ± 0.0210	-4.1707 ± 0.0213	0.073
WASP-84	3078836109158636928	0.843 ± 0.029	5210 ± 175	0.445(31)	-4.4652	-4.2175	–

Table 2. continued.

System	*Gaia* DR3 Id.	$(B-V)$	$T_{\rm eff}$ (K)	$S_{\rm MWO}$	$\log(R'_{\rm HK})_N$	$\log(R'_{\rm HK})_M$	$R'_{\rm HK}\times10^5$ rms
WASP-103	4439085988769170432	0.601 ± 0.022	5908 ± 191	$0.247 \pm 0.030^{(32)}$	-4.6106 ± 0.0990	$-4.4066\pm$	–
WASP-106	3788394461991295488	0.547 ± 0.021	6089 ± 214	0.182 ± 0.006	-4.8130 ± 0.0318	-4.6738 ± 0.0378	0.107
WASP-107	3578638842054261248	1.272 ± 0.016	4262 ± 194	1.191 ± 0.089	-4.6556 ± 0.0324	-4.0914 ± 0.0339	0.166
WASP-117	4746157737910069888	0.566 ± 0.021	6023 ± 182	$0.160^{(33)}$	-4.9497	-4.8242	–
WASP-127	3778075717162985600	0.638 ± 0.023	5788 ± 211	0.156 ± 0.004	-5.0157 ± 0.0280	-4.7914 ± 0.0271	0.064
XO-1	1316708918505350528	0.690 ± 0.025	5630 ± 236	0.182 ± 0.011	-4.8929 ± 0.0506	-4.7086 ± 0.0547	0.143
XO-2N	934346809278715776	0.827 ± 0.028	5248 ± 222	0.175 ± 0.031	-4.9833 ± 0.1272	-4.7398 ± 0.1346	0.291
XO-3 [a]	470650560779277952	0.325 ± 0.016	$6781 \pm 44^{(34)}$	0.221 ± 0.036	-4.4877 ± 0.0802	-4.4124 ± 0.1164	0.789
XO-4	990291507088739072	0.488 ± 0.020	6300 ± 176	0.146 ± 0.055	-5.0001 ± 0.1775	-4.991 ± 0.2271	1.009

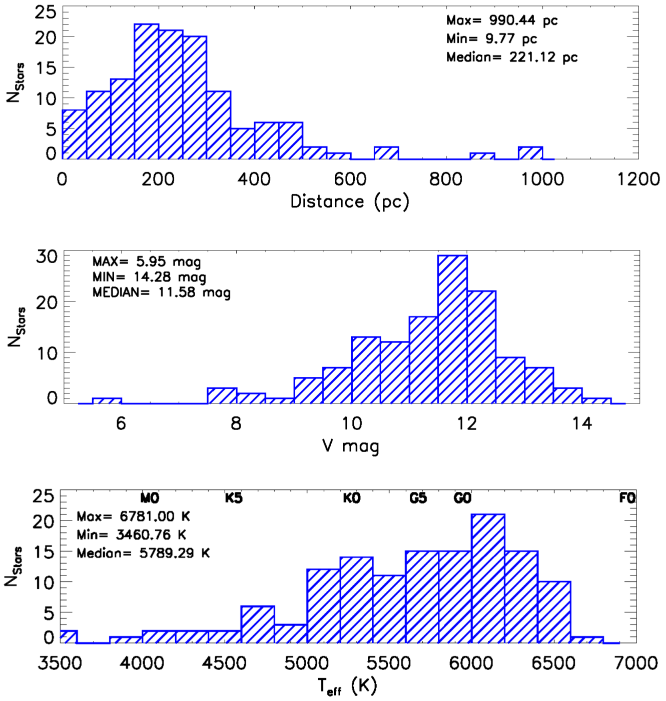

Fig. 2. Distributions of the stellar parameters of the GAPS sample. Upper panel: distribution of the distances of the star of the sample. Middle panel: distribution of the apparent V mag. Bottom panel: $T_{\rm eff}$ distribution. The corresponding spectral types are indicated at the very top.

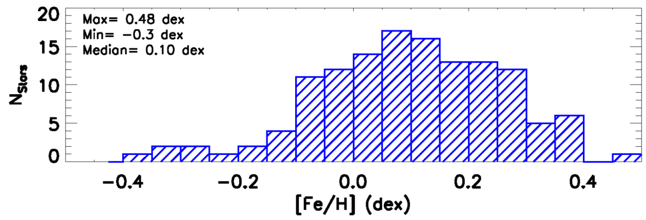

Fig. 3. Distribution of the spectroscopic stellar parameter: metallicities of the hosts.

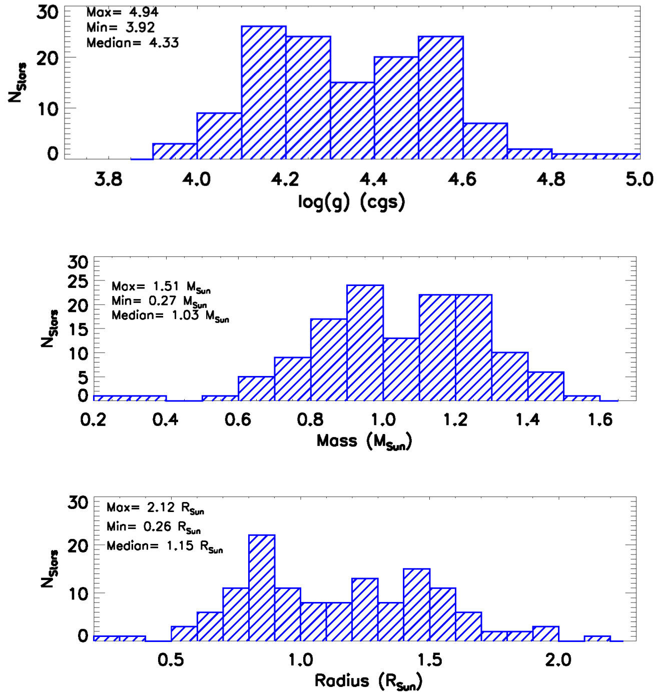

Fig. 4. Distribution of the stellar parameters: surface gravity (top panel), masses (middle panel), and radii (bottom panel).

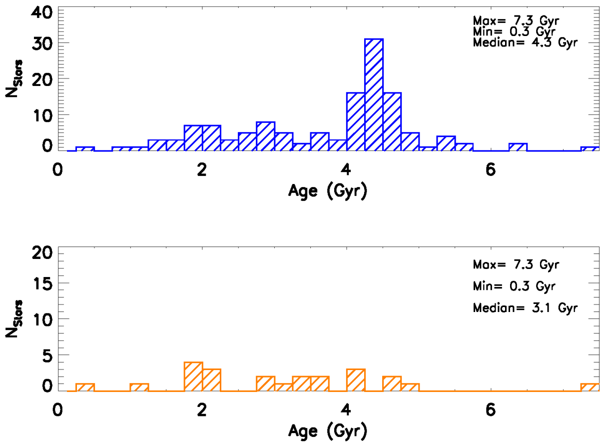

Fig. 5. Distribution of the stellar ages. Top panel: all hosts. Bottom panel: age distribution, considering only those with Δage/age ≤ 50%.

In Fig. 4 (from top to bottom), the distribution of surface gravity, masses, and radii are reported. They range between $3.92 \leq \log g \leq 4.94$ with a median uncertainty of 1%, $0.27 \leq M/M_{\odot} \leq 1.51$, with a median relative error of 4%, and $0.26 \leq R/R_{\odot} \leq 2.12$ with an uncertainty of 5%, respectively.

The age of the sample is between $0.3 \leq$ age ≤ 7.3 Gyr, with 4.3 Gyr as the median age (see Fig. 5). The median relative uncertainty on age determination is 83%. Such a high uncertainty is generally given by two orders of problems tied both to the isochrones method. The age determination with the

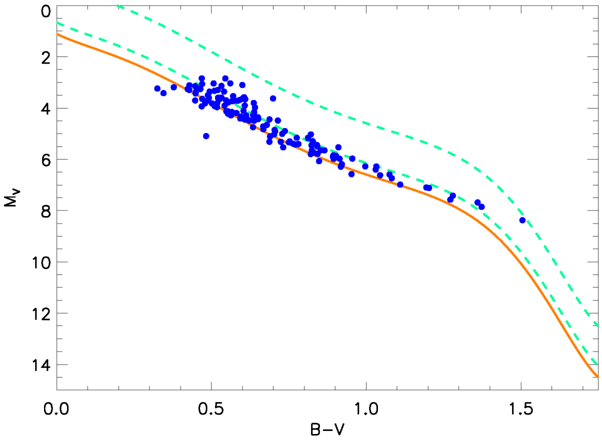

Fig. 6. HIPPARCOS average MS by Wright (2004) and the transiting planet hosts. Light-green dashed lines indicate loci 0.45 and 2.0 mag above the MS.

isochrones method introduces large uncertainties due to the poor knowledge of star luminosities caused by the errors in the values of the parallaxes; furthermore, a second effect is due to the long permanence of the star on the main sequence (MS). In our case, the parallax comes from the *Gaia* DR3 with a small uncertainty of 0.3%. Nevertheless, $T_{\rm eff}$ is also an input parameter useful for evaluating the star's age. In this case (as already described), for most of the objects of our sample, we used the values of the $T_{\rm eff}$ from the PIC, with a median error of 203 K. Such a high uncertainty on $T_{\rm eff}$ has an impact on the choice of the right isochrones. Taking this into account, we selected all the ages with an uncertainty ≤50% (see bottom panel of Fig. 5). The peak of the distribution at 4.2 Gyr disappears, leaving a quite uniform distribution of ages.

To discern between evolved and unevolved stars in our sample, we evaluated the absolute visual magnitude for each host star taking into account their parallax. The parallax errors allow us to distinguish between M-S and slightly evolved objects. In Fig. 6, we compare the objects with the HIPPARCOS average M-S fitted by Wright (2004) in the (B−V) vs. M_V plane. The two dashed curves indicate the portion of the plane above the M-S by 0.45 and 2.0 mag, where moderately evolved stars fall. Some objects (33) are in the sub-giant region. Objects with a height greater than 2.0 mag are considered evolved objects, but none of the stars fall in this region. It is worth noting that this is a selection effect that limits the number of evolved stars with respect to MS stellar hosts. In fact, the planet detection around a star and the determination of the planetary mass is made hard by the RV jitter induced by magnetic activity (e.g., Wright 2005) and stellar oscillations (e.g., Hatzes 1996; Beck et al. 2015). The latter ones are typically enhanced in amplitude and periods in evolved stars together with the increasing strength of surface convective motions. Both cause an increase in the RV jitter that may conceal the Keplerian signal of a planetary companion.

4. Chromospheric activity of the hosts

4.1. Extracting S-index and log R'_{HK}

To evaluate the S-index for each spectrum, we used the procedure described by Lovis et al. (2011). After the standard data reduction obtained by the DRS on-line at the telescope, this procedure performs background subtraction to correct for the diffuse light due to the ThAr or Fabry Perot lamp when we observe in simultaneous reference mode. The correction for the background is essential for the spectrophotometric accuracy and precision in evaluating the S-index. The S-index is then evaluated in the same way and with the same spectral passbands as defined for Mt Wilson Observatory spectrophotometer (Duncan et al. 1991). In particular, the Mt Wilson index is defined as:

$$S = \alpha \frac{H + K}{R + V}, \tag{1}$$

where H and K are the total flux in the CaII H&K line cores evaluated over triangular bands centered at 393.3664 nm (K) and 396.8470 nm (H), with an FWHM of 0.109 nm. Instead, V and R are the continuum bandpasses at 390.1070 nm and 400.1070 nm, respectively, with a width of 2.0 nm. The α value is a historical conversion factor discussed in depth by Hall et al. (2007), and fixed either at 2.40 or 2.30 (Duncan et al. 1991), or adjusted following the calibration by Baliunas et al. (1995). Lovis et al. (2011) found that this S-index is equal (or very close) to the ratio of the mean fluxes per wavelength interval ($\overline{X} = X/\Delta\lambda_X$, with X representing any bandpass, and $\Delta\lambda_X$ its width in wavelength). This reduces edge effects from finite sampling at the boundaries of the bands and it leads to a value of $\alpha \sim 1$ (Lovis et al. 2011). Together with the S value, its uncertainty is also computed considering the photon shot noise through error propagation. The value of the S-index obtained with HARPS (or HARPS-N) is indeed very close to the $S_{\rm MWO}$ index, as can be seen by Eq. (3) in Lovis et al. (2011), with an observed difference between the two indices of 0.03.

The value of $R'_{\rm HK}$ is the chromospheric excess flux of the CaII H&K lines, $\mathcal{F}'_{\rm HK}$, normalized by the bolometric flux of the star (Noyes et al. 1984):

$$R'_{\rm HK} = \left(\frac{\mathcal{F}'_{\rm HK}}{\sigma T^4_{\rm eff}}\right) = R_{\rm HK} - R_{\rm HK,phot}, \tag{2}$$

where $R_{\rm HK}$ is the stellar surface flux $\mathcal{F}_{\rm HK}$ in the CaII H&K lines normalized by the bolometric stellar flux and $R_{\rm HK,phot}$ is the photospheric flux contribution, $\mathcal{F}_{\rm HK,phot}$, in the line center normalized by the stellar bolometric flux. The former is tied to the measured $S_{\rm MWO}$ index via the following relation:

$$\mathcal{F}_{\rm HK} = \frac{\mathcal{F}_{\rm RV}}{\alpha} S_{\rm MWO}, \tag{3}$$

where $\mathcal{F}_{RV}$ is the surface flux in both the continua, while α is the same factor of Eq. (1).

The procedure on YABI measures in an automatic way the value of the S-indices of the time series obtained by GAPS for each object. In the period between 2012 Sep. 28 (JD=2456199) and 2012 Oct. 25 (JD=2456228), the HARPS-N CCD had a failure in the red chip. We did not consider the spectra taken in that period. We exploited the S-index values obtained with the procedure on YABI and derived the $\log(R'_{\rm HK})_{\rm N}$ values following Noyes et al. (1984). Some stars of the sample show high dispersion in the time series of $\log(R'_{\rm HK})_{\rm N}$, due to the intrinsic activity cycle or flares. These variations can be occasionally observed during intense flares, but in most of the transient events, the ranges of the $\log(R'_{\rm HK})$ changes are comparable or smaller than those associated with the rotational modulation or the activity cycles. Nevertheless, the sparseness of the time series in a significant fraction of our targets prevents us from identifying the individual sources of variations as seen in the $\log(R'_{\rm HK})$ datasets of the individual stars. To characterize the mean activity level of each of our targets, we decided to adopt the median of the observed $\log(R'_{\rm HK})$ timeseries of each star as the best indicator for that level. We computed also the arithmetic mean and compared it with the median to provide information on the deviation of the distribution of the $\log(R'_{\rm HK})$ measurements from a symmetric distribution that may suggest a preponderance of transient events. In general, these tend to increase the mean level of the $\log(R'_{\rm HK})$ with respect to its median. We found only three hosts (HAT-P-50, WASP-43, and WASP-50) with the average $\log(R'_{\rm HK})$ values differing more than 0.1 dex from the corresponding median. Moreover, the S/N of HAT-P-50 spectra is so low that we decided to exclude it from our analysis. The S-index and the corresponding $\log(R'_{\rm HK})_{\rm N}$ median values are reported for each object of the GAPS sample in Table 2. We discarded KELT-7, KELT-9, KELT-20, and WASP-33 from our sample because of their high temperatures and very blue colors that are very far from the bluer limit of the calibration range of $\log(R'_{\rm HK})$. Moreover, we evaluate the $\log(R'_{\rm HK})_{\rm N}$ also for all the hosts for which the $S_{\rm HK}$ values have been extracted by the literature (see Sect. 2.2, and Table 2).

4.2. The correction for the photospheric line flux

The $\log(R'_{\rm HK})_{\rm N}$ index obtained with the Lovis' procedure underestimates the value of the activity of the star, because the photospheric flux line contribution is not fully taken into consideration in the correction by Noyes et al. (1984). This choice results in an under-correction of the photospheric flux contribution by a factor of 2 (Mittag et al. 2013), thereby leading to an underestimation of the true flux excess in the Ca II H&K lines.

To convert the S-index into absolute CaII H&K chromospheric flux, Mittag et al. (2013) derived new color-dependent photospheric flux relation for the MS and evolved (subgiant and giant) stars in the color range between $0.44 \leq (B-V) \leq 1.6$. This will allow us to compare different compilations of the CA indicator from the literature.

Mittag et al. (2013) derived the following relation between the continuum flux and color index $B{-}V$ for the MS stars (with $0.44 \leq B-V \leq 1.60$) and subgiant stars (with $0.44 \leq B-V \leq 1.10$):

$$\log\left(\frac{\mathcal{F}_{\rm RV}}{\alpha}\right) = 8.25 - 1.67(B-V). \tag{4}$$

With respect to the $R_{\rm HK,phot}$ in Mittag et al. (2013), there are also the following relations with the color index for stars of different luminosity classes:

$$\begin{cases} \text{MS stars} & \\ \log\mathcal{F}_{\rm HK,phot} = 7.49 - 2.06(B-V) & 0.44 \leq B-V < 1.28 \\ \log\mathcal{F}_{\rm HK,phot} = 6.19 - 1.04(B-V) & 1.28 \leq B-V < 1.60 \\ & \\ \text{Subgiant stars} & \\ \log\mathcal{F}_{\rm HK,phot} = 7.57 - 2.25(B-V) & 0.44 \leq B-V < 1.10 \\ & \\ \text{Giant stars} & \\ \log\mathcal{F}_{\rm HK,phot} = 7.61 - 2.37(B-V) & 0.76 \leq B-V < 1.18. \end{cases} \tag{5}$$

This method allows for a simple and reliable conversion of the S-index and it will allow us to compare historical and new S-indices (Mittag et al. 2013).

Using the relations from Eqs. (2) to (5), we estimated the value of $\log(R'_{\rm HK})_{\rm M}$ for all the hosts (see Table 2). We use the suffix M to distinguish the value of $\log(R'_{\rm HK})$ obtained with the procedure described by Mittag et al. (2013) from that ($\log(R'_{\rm HK})_{\rm N}$) obtained with the procedure described by Noyes et al. (1984).

We compared the values obtained with both methods to check if they are different in a significant way. We test the difference between the two mean $\log(R'_{\rm HK})$ values against the null hypothesis using the t-student statistic technique. It results that most of the means are significantly different with a confidence level of <0.001. The objects: HAT-P-6, HAT-P-24, HAT-P-30, WASP-3, WASP-12, WASP-31, and XO-4, have averages that cannot be considered different due to large errors that hamper the evaluation of the two $\log(R'_{\rm HK})$ values.

The obtained values of the CA allow us to find an empirical relation between $\log(R'_{\rm HK})_{\rm N}$ and $\log(R'_{\rm HK})_{\rm M}$ for the different evolutionary states of the hosts. The empirical relation will be useful to correct the classic index for the CA ($\log(R'_{\rm HK})_{\rm N}$) for the photospheric line flux obtaining $\log(R'_{\rm HK})_{\rm M}$ when the S-indices are missing in the literature. We separated the sample into MS stars and slightly evolved (subgiant) stars as they come out from the comparison with the HIPPARCOS average MS (see Sect. 3 and Fig. 6) and for both the subsamples we search for a linear correlation between them. We take into consideration all the significant different values for all the objects in the MS with color ranging in the $0.44 \leq B{-}V \leq 1.60$ interval. For the subgiant stars instead, we considered those objects with color $0.44 \leq B{-}V \leq 1.10$. Figure 7 displays the comparison between the $\log(R'_{\rm HK})_{\rm N}$ and $\log(R'_{\rm HK})_{\rm M}$ for the MS objects (top panel) and sub-giant stars (bottom panel) and the correlation is described by the following:

$$\begin{cases} \text{Main sequence stars} \\ \log(R'_{\rm HK})_M = (1.255 \pm 0.229) + (1.223 \pm 0.047) \times \log(R'_{\rm HK})_N \\ \\ \text{Subgiant stars} \\ \log(R'_{\rm HK})_M = (1.205 \pm 0.485) + (1.202 \pm 0.096) \times \log(R'_{\rm HK})_N. \end{cases} \tag{6}$$

In the following of this work, we will use the values of $\log(R'_{\rm HK})_{\rm M}$, and consequently, we drop the subscript M. The distribution of the $\log(R'_{\rm HK})$ is shown in Fig. 8. Most of the transiting planet hosts are in the inactive and active region of the plot as they are defined by Wright (2004) and Henry et al. (1996), fewer are in the very active region, while some objects fall in the very inactive region with $\log(R'_{\rm HK}) \leq -5.1$ (basal activity limit for MS stars). The upper and lower limits of our sample are respectively individuated by WASP-7 ($\log(R'_{\rm HK}) = -3.9722$) and GJ436 ($\log(R'_{\rm HK}) = -5.3727$). The median value of the sample is $\log(R'_{\rm HK}) = -4.8099$ held by HD 97658. This distribution follows the typical bias against very active stars of RV and (to a lesser degree) transit surveys.

5. $R'_{\rm HK}$ raw dispersion and chromospheric emission variability

For the targets observed by GAPS, we have long time series of the CA indicator, but their inhomogeneity does not allow a global analysis. Nevertheless, we used the recorded variation of CA to better characterize this sample. Information on the activity variability can be obtained via the standard deviation of the stellar activity indicator. The left panel of Fig. 9 shows the distribution

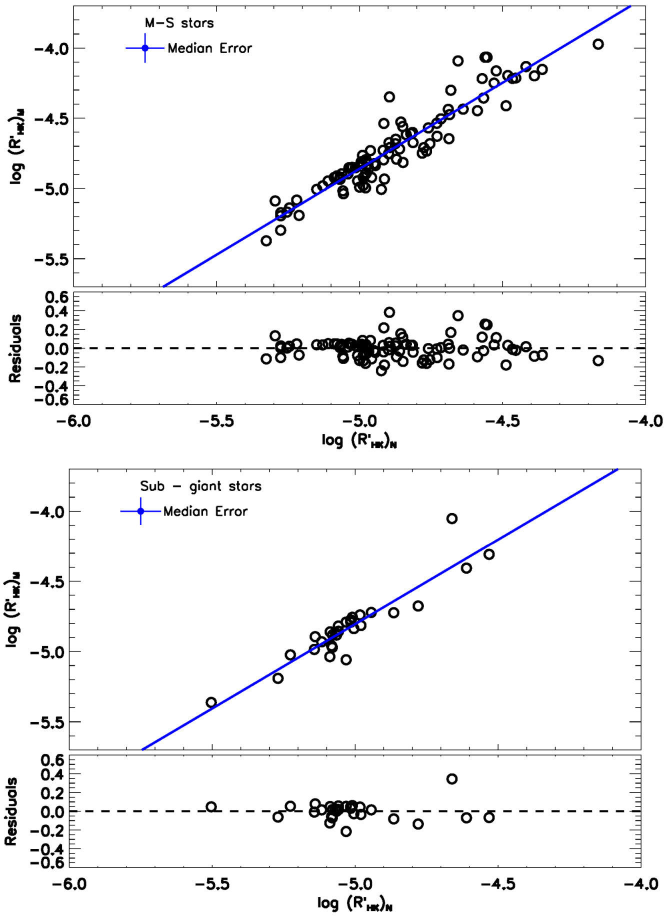

Fig. 7. Comparison between the values of $\log(R'_{\mathrm{HK,Mittag}})$ and $\log(R'_{\mathrm{HK,Noyes}})$ for the objects of the sample. The continuous blue line represents the best linear fit and the residuals of the fit are plotted beneath. The median error bars are indicated at the top of the plot. Top panel: comparison for the MS objects. Bottom panel: same as the top panel but for the sub-giant stars.

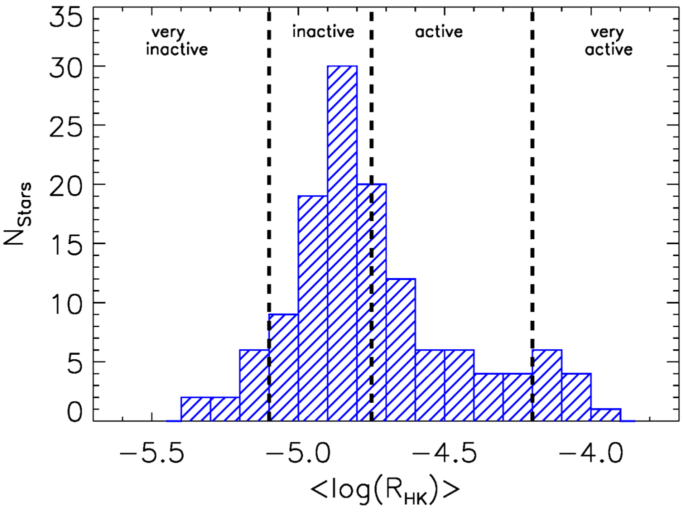

Fig. 8. Distribution of the $\log(R'_{\mathrm{HK}})$ values of the considered sample of transiting objects. Plot also indicates the activity regions following Wright (2004) and Henry et al. (1996).

of the standard deviation of $10^5 \times R'_{\mathrm{HK}}$ (hereafter, σR_5) as evaluated by the GAPS time series. The median of the distribution is 0.2 dex, and the distribution shows a tail of stars with high dispersion, possibly due to short-term variability (e.g., flares) or long-term magnetic cycles. The M2.5 dwarf GJ436 records the minimum value (0.038 dex) of the chromospheric emission (CE) variability. We observed this star on two distinct nights about 2 yr apart, so it could not be considered proof of the lowest dispersion level of our sample. Following other authors (e.g., Lovis et al. 2011; Gomes da Silva et al. 2021), we selected hosts observed on more than 20 nights on a time basis longer than 1000 days. The host with the lowest relative dispersion is HAT-P-7 (F6V) with an absolute $\sigma R_5 = 0.079$ dex and a relative dispersion for S_{MW} of 2.0%. Comparing HAT-P-7 with the values for the quiet star τ Ceti: $S_{\mathrm{MW}}/\sigma S_{\mathrm{MW}} = 0.35\%$, obtained by Lovis et al. (2011) on a time bases of 7 yr of observations (157 nights), we can see that also HAT-P-7 could not be considered a long-term quiet star. So we have just a mere indication of the precision of our R'_{HK} measurement.

Stellar activity compilations and catalogs measured for a high number of stars show that the higher envelope of activity dispersion decreases with decreasing activity level and evolutionary stage. The majority of stars have a variation in activity that ranges $0.01 \leq \sigma R_5 \leq 0.32$ dex (Gomes da Silva et al. 2021). In our sample, there are systems with a dispersion larger than 0.5 dex (see Table 3). Among them, Qatar-2 shows the highest amplitude of variation (1.736 dex) but because of the very short time span (less than 1 day), it could be due to a flare. All the others have longer lapstime, instead.

In general, the activity variability is proportional to the level of stellar activity, and therefore, regarding age, young stars have greater activity variation than old ones. This behavior has been also highlighted by Scandariato et al. (2017) for a sample of low-activity early-type M dwarfs. The right panel of Fig. 9 reports the σR_5 as a function of the $\log(R'_{\mathrm{HK}})$. The different colors represent the time span of the measurements. Considering only the red squares, corresponding to values with a time baseline longer than 100 days, it is possible to see that active stars have a greater spread of the CE variation. Nevertheless, some stars in the inactive zone ($-5.1 \lesssim \log(R'_{\mathrm{HK}}) \lesssim -4.75$) exhibit a broad variation. They are WASP-1 (F9.5V), HAT-P-16 (F9V), XO-4 (F7V), HAT-P-26 (K2V), and WASP-35 (G1V). This situation seems to be corroborated by Gomes da Silva et al. (2021), who observed a high activity variation for stars of spectral types F and K with a low level of activity, as well as a larger spread of variation for the former spectral type than for the latter.

For the time being, we do not have the capacity to dig deeply into the activity variation behavior, and we leave it for future works; this is also because it is not the focal topic of this paper. In any case, the data indicate that the more active the stars, the more the activity varies. Furthermore, non-active stars could have high activity variation, in particular F stars.

6. Activity of planetary hosts in a general context

Figure 10 shows the distribution of the mean $\log(R'_{\mathrm{HK}})$ obtained for each host as a function of the T_{eff} of the stars. In this plot, it is possible to note that F stars (blue circles) have a scattered distribution with a concentration around $\log(R'_{\mathrm{HK}}) = -5.0$. Differently, the G stars (green circles) range between $-5.1 \leq \log(R'_{\mathrm{HK}}) \leq -4.1$, with a slightly high concentration around $\log(R'_{\mathrm{HK}}) \sim -4.9$. Then, K stars (orange circles) show a scattered

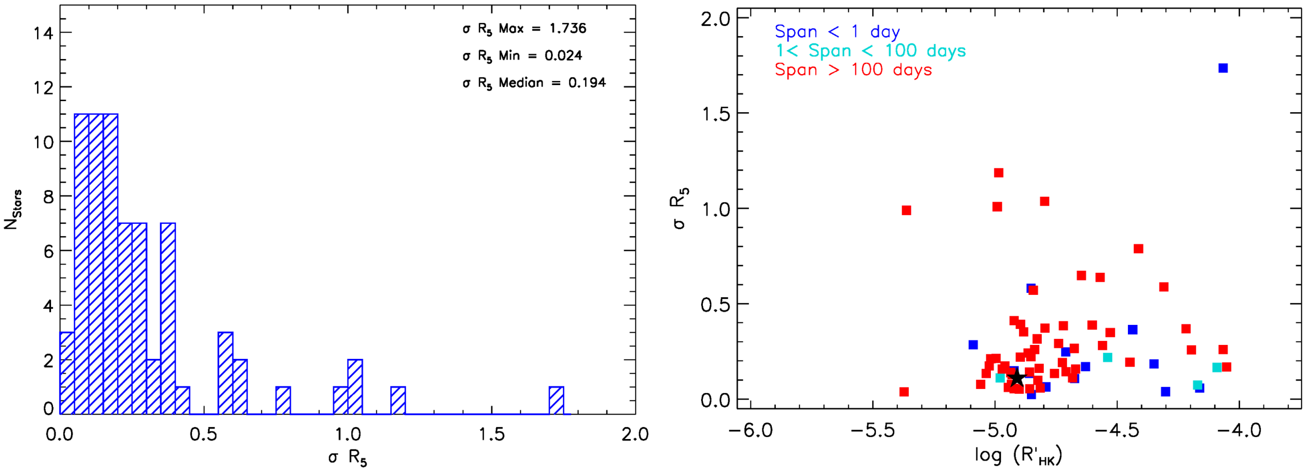

Fig. 9. Standard deviation of σR_5 for the GAPS transiting planet hosts:. Left panel: distribution of σR_5. Right panel: standard deviation of R_5 versus the $\log(R'_{\rm HK})$ for the GAPS stars. The different colors indicate the time baselines of the measurements as reported in the legend. The position of the Sun is marked with a black star.

Table 3. GAPS transiting systems with $\sigma R_5 > 0.5$.

System	Time span (d)	σR_5
Qatar-2	0.1713	1.736
WASP-1	1433.944	1.186
HAT-P-16	1849.973	1.036
XO-4	1164.830	1.009
WASP-12	1955.973	0.989
XO-3	1481.935	0.789
HAT-P-6	1859.903	0.648
WASP-57	426.063	0.639
HAT-P-21	1494.9679	0.588
HAT-P-26	0.1882	0.581
WASP-35	2661.838	0.571

distribution in the activity range of $-5.0 \leq \log(R'_{\rm HK}) \leq -4.0$. There are only two M stars (red circles), too few to say anything about stars of this spectral type. We can just highlight their outlying position in the diagram. The outer atmosphere of these stars, due to their faintness in the visible wavelength range, still remains poorly understood. Despite this, there are several analyses of the activity of M stars (e.g., Maldonado et al. 2017, and reference therein). The FGK host stars' behavior seems to recall what Gomes da Silva et al. (2021) found by analyzing the distribution of the activity level of about 1650 stars for the different spectral types. In particular, for K stars, they found a distribution with three peaks: an inactive peak at −4.9 dex, an intermediate peak near −4.75 dex and an active peak near −4.5 dex.

We have too few K stars to recognize the peaks reported by Gomes da Silva et al. (2021), but it is possible to observe that also for our sample, the K stars show a large scatter.

6.1. Comparison with MS and evolved stars

To put the activity of host stars into context, we exploited the sample composed by Pace (2013) with a compilation of S-index measurements for about 2000 field stars from the literature. To remain homogeneous with the planetary host data, we search each field star in *Gaia* DR3 and eventually we evaluate the $(B-V)$ color and the temperature for each star in the same way as described in Sect. 3. Since MS and evolved stars have very different activity distributions (Wright 2004; Mittag et al. 2013; Staab et al. 2017), to separate them in the field star sample, we used the absolute magnitude obtained from the XHIP catalog (Anderson & Francis 2012) and built an HR diagram (Fig. 11) using the average MS defined by Wright (2004) as in Sect. 3.

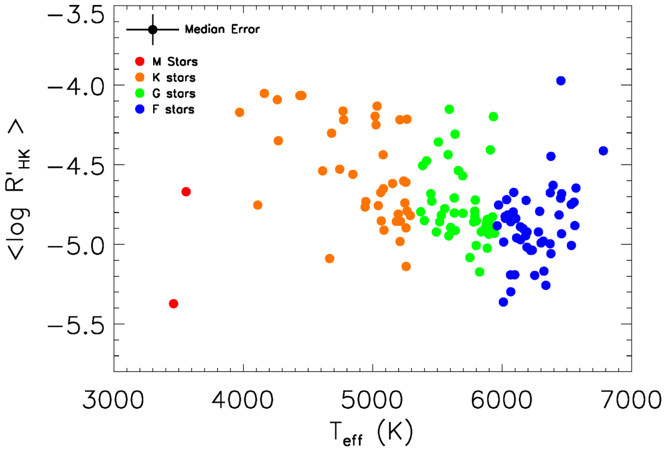

Fig. 10. Values of $\log(R'_{\rm HK})$ as a function of the $T_{\rm eff}$ of the star of sample. The color code is explained in the legend and the median error of both $T_{\rm eff}$ and $\log(R'_{\rm HK})$ are also indicated (in black).

We computed the $\log(R'_{\rm HK})$ values for both the MS and evolved stars identified in the Pace (2013) sample, following the procedure described in Sect. 4.2. The two left panels of Fig. 12 show the $\log(R'_{\rm HK})$ measurements for the planetary hosts compared with MS field stars (top panels) or evolved stars (bottom panels). In the top left panel, the basal limit for MS stars is also indicated with a horizontal black line. The basal level of activity is the stellar analog of the quiet Sun reached by stars with no chromospherically active regions (e.g., Schrijver 1987; Schröder et al. 2012). The basal activity limit of $\log(R'_{\rm HK}) > -5.1$ is applied only for dwarfs with solar metallicity. It is a little bit different (−5.15) for MS stars with super-solar metallicity

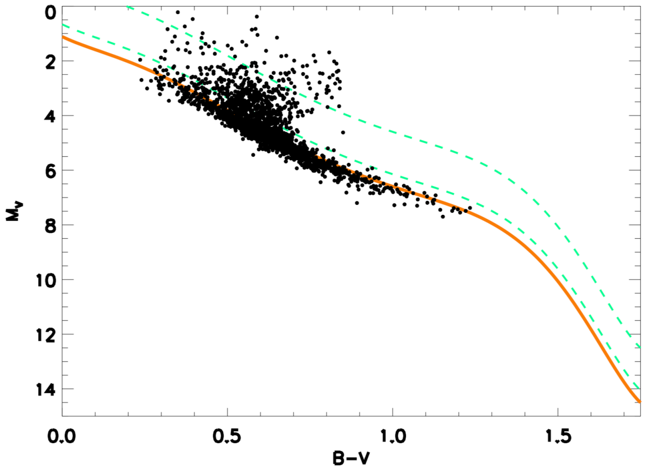

Fig. 11. HR Diagram for the field stars in the Pace (2013) sample. The HIPPARCOS average MS by Wright (2004) and the loci 0.45 and 2 mag above it are indicated in orange and green, as in Fig. 6.

(Henry et al. 1996; Wright 2004; Saar 2011). Furthermore, the basal level is also a slightly complicate function of the $(B-V)$ of the star (Mittag et al. 2013, and references therein). Considering the polynomial model of the basal level described in Table 3 of Mittag et al. (2013), we superimposed it (the blue dashed curve) on the top-left panel of Fig. 12. The largest difference between the two descriptions of the basal level is visible in the red portion of the plot where the model has a quadratic dependence on the $B-V$ color. The presence of a relative maximum in the range $1.1 \leq B-V \leq 1.45$ has been justified by Mittag et al. (2013) with the absence of inactive stars in that color range. Instead, in the bluer range of the diagram, the difference between the two is inside the median measurement error. In any case, the polynomial description does not represent the basal limit in a better way than a constant value of −5.1 (e.g. Henry et al. 1996; Wright 2004; Staab et al. 2017; Haswell et al. 2020).

The right panels of Fig. 12 compare the histogram and cumulative distribution function of $\log(R'_{\rm HK})$ for both the host and MS stars (top – right), and for host and slightly evolved stars (bottom – right). A comparison of the hosts with MS and evolved star populations shows that they actually come from different parent distributions in both cases (see Appendix B). This difference can be interpreted as due to observational selection biases that make planet detection more likely for less active hosts.

Among the MS stars in Pace (2013), there are very few stars (about 0.3%) below the basal activity level of −5.1, and about 1% in the complete sample. In our sample of transiting-planet hosts, ten stars (CoRoT-1, GJ436, HAT-P-35, HAT-P-44, HAT-P-45, HAT-P-46, Kepler-25, WASP-12, WASP-18, WASP-72), corresponding to 8% of our sample, show activity levels below the basal limit. Two of these, WASP-12, and WASP-72, are reported to be sub-giants having the $\Delta M_{\rm V} > 0.45$ (0.74, and 1.26, respectively, see Sect. 3)[8]. Evolved stars can be inactive stars well below the basal limit for MS stars (e.g., Wright 2004). The remaining host stars are all MS stars and in those cases, a mechanism that suppresses the star activity, making it lower than the basal limit should be at work. The activity suppression seems more common in our sample of planet-transiting hosts than in the field stars. This mechanism could also veil the CE of more active stars, but in this case, it will be very difficult to identify it. Hence, the activity suppression due to evaporated shrouds could be a more common phenomenon than it appears.

The presence of some planetary hosts below the MS activity basal limit has been interpreted as an indication that the presence of short-period massive planets mimics an activity suppression, but it is actually due to circumstellar absorption linked to planetary mass loss, which is sensitive to the surface gravity of the planet (Lanza 2014; Fossati et al. 2015). The same behavior of some MS stars has been considered, with success, which is indicative of undiscovered mass-losing planets (Haswell et al. 2020). A model that reproduces the observed correlation was proposed by Lanza (2014) and Fossati et al. (2015). On the other hand, since the transiting planet hosts are also distant stars (see the upper panel of Fig. 2 for the sample under study), a contribution to this suppression of the observed chromospheric index may be also given by the interstellar medium (Staab et al. 2017).

6.2. Considering whether the ISM absorption could bias the chromospheric activity level

The value of $\log(R'_{\rm HK})$ of 8% of our sample of stars is lower than the activity basal limit for the MS. One of the causes that can contribute to justifying this low value of the $\log(R'_{\rm HK})$ is the possible absorption of the HK flux by the interstellar medium (ISM).

Figure 13 shows the distribution of the transit hosts into the Galaxy. The red stars indicate the line of view of the hosts with an activity level lower than the basal one. In the plot, there is no evidence of a preferred direction. As we saw in Sect. 3, most of the stars we are considering have distances greater than 100 pc. For these systems, the ISM absorption has a significant influence on the measurement of the activity. As a matter of fact, beyond 100 pc the column density of the ISM Ca II is about 10^{12} cm^{-2} (Welsh et al. 2010; Wyman & Redfield 2013; Fossati et al. 2017).

Following the work of Fossati et al. (2017), for each target we evaluate the column density of ISM Ca II. We utilized the ISM maps by Lallement et al. (2019, 2022) to evaluate $E(B-V)$ and successively the relation by Diplas & Savage (1994) to derive the column density of H I. The relation between the $N_{\rm H\,I}$ and the ionic gas-phase abundance of ISM by Wakker & Mathis (2000) gave us the column density of Ca II. Finally, to obtain the correction of the $\log(R'_{\rm HK})$ for the ISM absorption, we used the IDL procedure `CORRECTION.PRO`, described by Fossati et al. (2017, more details are given in Table A.2).

We note that with the only exception of HAT-P-44 and HAT-P-45, all the other eight MS systems with $\log(R'_{\rm HK}) \leq -5.1$ remain below the basal level after the application of the correction (see the systems highlighted in Table A.2). We can conclude that for these systems, the ISM absorption only partially explains the suppression of the CA index below the basal level, so that other mechanisms, such as absorption from evaporated shrouds, are at work.

---

[8] We note in passing that the classification of WASP-12 as a subgiant, that is supported by the modeling of our HARPS-N spectra, may account for the decay rate of the orbit of its transiting planet due to the dissipation of tidal inertial gravity waves excited by the planet inside the star. According to the present theory of dynamic tide, those waves have a sufficient dissipation rate only if the core of WASP-12 is radiative, as generally predicted by stellar structure models in the case of a subgiant, but not in the case of a MS star because the latter would have a convective core leading to a much smaller dissipation (e.g., Bailey & Goodman 2019; Barker 2020).

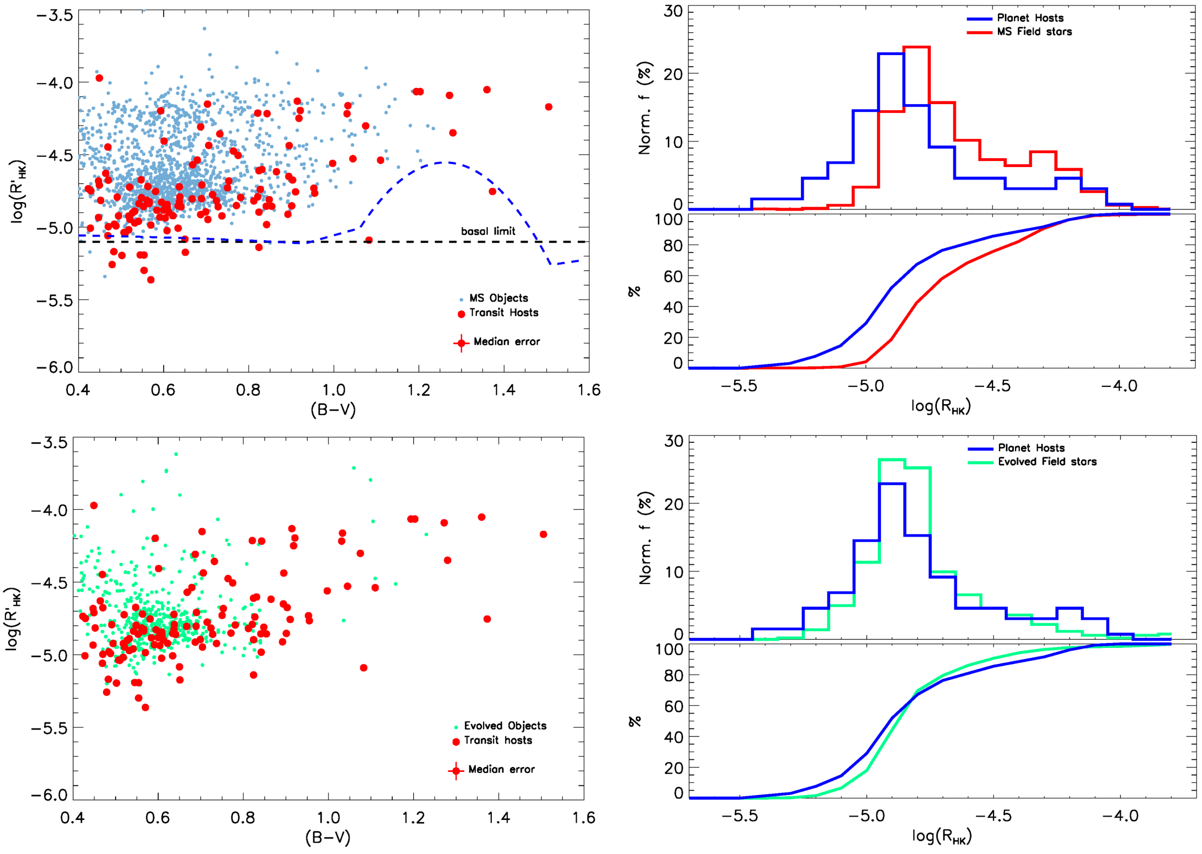

Fig. 12. Comparison of the CA of host stars with the CA of MS and evolved stars. Top panel: distribution of the MS-field stars with $\Delta M_V \leq 0.45$ above the average HIPPARCOS MS (cyano small filled circles) and the transiting planet hosts (left). In the plot, we also indicated two representations of the activity basal limit for MS stars. The black dashed line is the generally considered constant limit. The blue dashed line is the polynomial model of the basal level (Mittag et al. 2013). The comparison of the normalized histograms and the cumulative distribution functions of the two distributions in the left panel (right). Bottom panel: evolved field stars with $0.45 \leq \Delta M_V \leq 2.00$ above the average Hypparcos MS (small filled green circles) and the transiting planet hosts (of both samples) are on the left. A comparison of the normalized histograms and the cumulative distribution functions of the two distributions in the left panel is on the right.

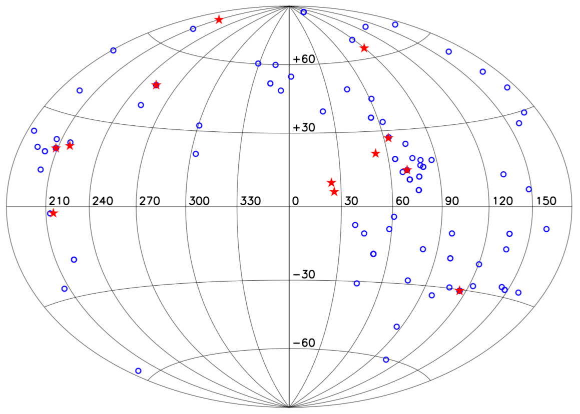

Fig. 13. Aitoff projection of the distribution of transit hosts in our sample in Galactic coordinates l and b of sight lines. The Galactic center is located at the center of the plot and the longitude increases counterclockwise. In the plot, the hosts with $\log(R'_{\mathrm{HK}}) \leq -5.1$ are indicated as red stars.

7. Host activity and planetary parameters analysis

We searched for correlations between the $\log(R'_{\mathrm{HK}})$ and planetary parameters, such as the mass (M_{P}), the radius (R_{P}), the density (ρ_{P}), and the orbital distance (a_{P}) of the planets. We checked also for correlation of the activity level with the planet's insolation ($S_\star/S_0$; S_0 is the Earth's insolation at the top of the atmosphere), the equilibrium temperature ($T_{\mathrm{eq,P}}$) of the planet assuming zero bond albedo. Considering also the star-planet interaction (SPI), we take into account the possibility that tidal interaction between the planet and its star can generate an enhancement of local turbulent velocity, resulting in an increment of the heating and a rise in stellar activity level (Cuntz et al. 2000). Hence, we searched for a correlation with the stellar gravitational perturbation by the planet (Cuntz et al. 2000):

$$\frac{\Delta g_\star}{g_\star} = \frac{M_p}{M_\star}\frac{2R_\star^3}{a_P^3}$$

where M_{P}, $M_\star$, and $R_\star$ are the mass of the planet, the mass of the star, and the radius of the star, respectively; $g_\star$ is the surface gravity of the host star, and a_{p} is the orbital axis of the planet.

While the former set of variables consists of planetary quantities, the latter set contains quantities evaluated with both star and planet parameters. We used planetary data determined in a previous work (Bonomo et al. 2017), while for planets in our sample not considered in that work, we used the Exoplanet Encyclopaedia[9] (Schneider et al. 2011), searching the literature, to

---

[9] Accessible from `http://exoplanet.eu/`

Table 4. Statistical analysis of the correlation between the CA indicator $< \log(R'_{\rm HK}) >$ and different planetary parameters of the transiting planetary systems.

Parameters	Condition	$\nu = N - 2$	$\rho^{(a)}$	FAP
M_P	None	140	0.029	73%
	$a_{\rm p} \leq 0.1$ au and $M_{\rm p} \geq 0.1\ M_{\rm J}$	109	0.081	40%
$R_{\rm P}$	None	140	−0.216	1.0%
	$a_{\rm p} \leq 0.1$ au and $M_{\rm p} \geq 0.1\ M_{\rm J}$	109	−0.315	0.09%
$\rho_{\rm P}$	None	140	0.186	3.0%
	$a_{\rm p} \leq 0.1$ au and $M_{\rm p} \geq 0.1\ M_{\rm J}$	109	0.256	0.7%
a_P	None	137	−0.179	4.0%
	$a_{\rm p} \leq 0.1$ au and $M_{\rm p} \geq 0.1\ M_{\rm J}$	109	−0.149	12%
$\log(g_P)$	None	140	0.204	1.5%
	$a_{\rm p} \leq 0.1$ au and $M_{\rm p} \geq 0.1\ M_{\rm J}$	109	0.219	2.3%
$S_\star/S_0$	None	137	−0.213	1.0%
	$a_{\rm p} \leq 0.1$ au and $M_{\rm p} \geq 0.1\ M_{\rm J}$	109	−0.322	0.06%
$T_{\rm eq,P}$	None	137	−0.213	1.0%
	$a_{\rm p} \leq 0.1$ au and $M_{\rm p} \geq 0.1\ M_{\rm J}$	109	−0.334	0.04%
$\frac{\Delta g_\star}{g_\star}$	None	140	0.007	94%
	$a_{\rm p} \leq 0.1$ au and $M_{\rm p} \geq 0.1\ M_{\rm J}$	108	−0.047	63%

Notes. In the second column, the condition of the sample is given; N is the total number of planets in the sample that fulfils the condition. In the last two columns we reported the Spearman's coefficient, ρ, and the value of the probability of the null hypothesis. [(a)] Rank correlation of two populations; the false alarm probability (FAP, in the last column) denotes the two-sided significance of its deviation from 0 by random chance, i.e., small values indicate a significant correlation.

obtain the mass, and radius and all the other useful parameters ($a_{\rm P}$, $\log(g_P)$, etc.). We are aware that the derivation of the planetary parameters obtained in this way has a degree of inhomogeneity, which can produce biased results, and we stress that the investigation that follows should be seen with this caveat in mind.

The analysis of all the possible correlations has been performed using the Spearman's rank coefficient (Press et al. 1992), as was already done on smaller samples of transiting planets by Hartman (2010) and Figueira et al. (2014). For each correlation coefficient, we obtained the false alarm probability (FAP) or the probability of having a larger or equal correlation coefficient under the hypothesis that the data pairs are uncorrelated (our null hypothesis). The tests were performed considering the whole sample of host-planet pairs with no restrictions and a second sub-set responding to the condition: $a_{\rm p} \leq 0.1$, and $M_{\rm p} \geq 0.1\ M_{\rm J}$. In Table 4, we summarize the results.

Among the planetary parameters, the correlation between $\log(R'_{\rm HK})$ and M_P has a value of Spearman's correlation coefficient that is not different from zero with a FAP of 73% considering the whole sample. This seems to contradict the statistical explanation of Collier Cameron & Jardine (2018) of the correlation reported by Hartman (2010) and Figueira et al. (2014) between the CE and the surface gravity of close-orbiting planets based on the statistical bias towards younger systems produced by the faster inward migration of more massive planets because of their stronger tidal interaction with their host stars. The correlation with the orbital semi-major axis a_P has a FAP = 4%. This is the only SPI correlation that has been considered so far by further dedicated studies (e.g. Figueira et al. 2014), to the best knowledge of the authors. Instead, the correlation with the $R_{\rm P}$ (FAP = 1.0% and FAP = 0.09%) have a confidence level greater than 99%, while $\log g_{\rm P}$ (FAP = 1.5% whole sample and 2.3% the limited one), and the density of the planet $\rho_{\rm P}$ (FAP = 3% and FAP = 0.7%), have all a confidence in the range between 97% and 99%.

Concerning the parameter containing stellar and planetary quantities, the correlation with insolation $S_\star/S_0$, and the planetary $T_{\rm eq,P}$ show also a confidence greater than 99%. This is not the same result, however, for the quantity $\Delta g_\star/g_\star$ (FAP = 94% and 63%). In Fig. 14, we reported the value of $\log(R'_{\rm HK})$ of the object as a function of the ratio, η, between the semi-major axis, $a_{\rm p}$, of the planet and the value of the Roche limit of the system (Faber et al. 2005; Ford & Rasio 2006):

$$a_{\rm Roche} = 2.16 R_{\rm P} \left(\frac{M_\star}{M_{\rm P}}\right)^{1/3}, \tag{7}$$

where $a_{\rm Roche}$ is the critical separation at which the planet fills its Roche's lobe and it begins to lose mass. This limit is a function of the planetary radius $R_{\rm p}$, and of the stellar ($M_\star$) and planetary ($M_{\rm P}$) masses.

The analysis of the correlation between η and $\log(R'_{\rm HK})$ does not give a significant correlation (Fig. 14). The Spearman's rank coefficient results $\rho = 0.02$ with a FAP $\sim$ 83%, but it is worth noting that all the hosts with $\log(R'_{\rm HK})$ lower than the basal limit have planets with a mean density ($1.7{\times}10^3$ kg m^{-3}) that is lower than the average density ($2.2{\times}10^3$ kg m^{-3}) of the planets with the ratio η in the interval [1, 10]. Actually, in the same range of η values, there is a high spread in the activity level of the hosts. The spread could be due to the dispersion of the host age and/or to the spread of the planetary companions' densities. In fact, in both cases, they are the stellar and planetary parameters with higher standard deviations by their mean values. There is not a clear indication of which parameter can account for it. Another reason that accounts for the spread in this region of the plot could

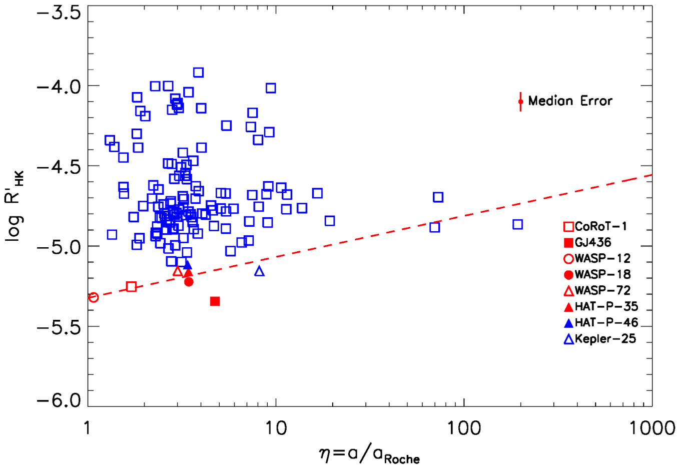

Fig. 14. Activity level as a function of the ratio η between the semi-major axis and the Roche lobe radius. The hosts with $\log(R'_{\rm HK}) < -5.1$ are clearly identified.

be the presence of matter evaporated by the planet that, depending on its quantity, could veil (with different levels of efficiency) the CE of the stars.

In the region of the plot for $\eta > 10$, the distribution of the activity levels has a reduced spread, and it is possible to define a border-line ($\log(R'_{\rm HK}) = 0.27\log(\eta) - 5.41$) that separates the space in the diagram between a region that is populated by the hosts of transiting planets and one that is empty. Because we are considering only transiting planets, the population of the $\eta > 10$ zone in the η–$\log(R'_{\rm HK})$ plane could be the result of a selection effect. Planets observed with the RV method (e.g., the sample discussed by Canto Martins et al. 2011) could help to populate this zone and unveil whether there is a selection effect present. For those planets, the value of the radius is unknown and the mass obtained is the minimum mass, making it impossible to evaluate the η value for those systems using measured quantities.

To conclude this section, we can say that there is no evidence for a strong single correlation between the activity level of the planet host star and one of the planetary and stellar parameters. Instead, a slight correlation seems to arise when we considered planetary parameters depending on the radius of the planet (surface gravity, density, insolation). All the parameters that have a correlation with $\log(R'_{\rm HK})$ with a confidence value greater than 98% are $\propto R_{\rm P}^n$ suggesting a common origin. The same is true for the insolation and the $T_{\rm eq,P}$ that show a negative correlation with the activity. The analysis of these correlations confirms that close-in transiting planets swell under the irradiation from their host stars corroborating the hypothesis already discussed in preceding works (e.g., Lanza 2014; Fossati et al. 2013, 2015) that the planetary atmosphere evaporates under the action of the stellar irradiation and the matter lost by the planet forms shrouds that absorb at the core of the chromospheric resonance lines (e.g., Mg II h&k, Ca II H&K), making the observed chromospheric emission correspondingly lower. Due to the role of the planetary radius in these correlations, the mechanism at work could be the hydrodynamic escape to which some of these planets (those with the $\log(R'_{\rm HK})$ lower than the basal value) should be subject. In this direction, we get also an indication that hosts with $\log(R'_{\rm HK})$ lower than the basal limit for MS stars are very close to filling their Roche lobe even if other unknown phenomena could have their effects.

On the other hand, it is worth noting that the analyzed correlations with the possible star-planet interaction (a_P and $\Delta g_\star/g_\star$) are not significant ones. This seems to indicate that the higher activity level of some of the considered transit hosts in this study is not due to the SPI, but it appears natural.

8. Machine learning classification

In order to test the Lanza (2014) model of planetary material absorbing at the core of chromospheric resonance lines, Fossati et al. (2015) applied a cluster-weighted model (Ingrassia et al. 2012) to distinguish different sub-populations in their data. This model reproduces the bimodal distribution observed in the CA level of solar-like stars in the field, the so-called *Vaughan-Preston gap* (e.g., Wright 2004). Specifically, the existence of two linear regressions can be considered to be the natural outcome of the underlying bimodal distribution of the chromospheric index. These authors explored the likelihood of one to three components in their data set and found two components to be the most likely representation.

Following these results, we also classified our data based on the $\log(R'_{\rm HK})$ and the inverse of the planetary surface gravity, $g_{\rm p}^{-1}$. In this analysis, we only considered systems with main sequence stars and excluded those with an eccentric orbit and planetary masses in the Neptunian range. For the latter, in particular, there would not be enough planetary material available to create the absorbing torus predicted by Lanza (2014). Moreover, as the effect of planetary material on stellar chromospheric lines is only expected for close-in planets, we excluded systems with $a_p > 0.1$ AU.

Our data set is larger than that of Fossati et al. (2015)'s by almost a factor of 3, and we visually identified four possible components: the same high- and low-activity components they highlighted, a third component with average activity and high inverse of the surface gravity $g_{\rm p}^{-1}$, namely, low planetary-surface gravity and a high-activity component with low correlation with $g_{\rm p}^{-1}$. We hypothesize the third population might represent a "saturated" regime for stellar irradiation-driven planet evaporation, where planetary mass loss $\dot{M}_{\rm p}$ becomes proportional to the square root of the stellar UV flux (Murray-Clay et al. 2009), rather than being proportional to the UV flux as assumed in the energy-limited evaporation model by Lanza (2014). Therefore, for this subset, a lower evaporation rate for a given $g_{\rm p}^{-1}$ than assumed by Lanza (2014) would hold and this would imply a lower absorption by the evaporated matter than predicted by his model. The fourth population might instead represent a set of stars with an intrinsically high level of activity and a strong enough stellar wind to blow away all evaporated matter, making the measured stellar chromospheric emission virtually unaffected by planet evaporation.

In light of the already available information, as well as our aforementioned hypotheses, we implemented a supervised machine-learning classification scheme to determine a statistically-motivated separation between the subsets. The details of its training and implementation can be found in Appendix C. This resulted in the grouping presented in Fig. 15, where each data point is represented by a marker and colour corresponding to its predicted cluster membership, and where white-filled data points are those with an assignment probability <70%: a similar threshold was used by Fossati et al. (2015) to identify points that might likely belong to a different cluster. The high- and low-activity, low $g_{\rm p}^{-1}$ populations (recalling those of Fossati et al. 2015) were then used to carry out linear fits in the $\log(R'_{\rm HK}) - g_{\rm p}^{-1}$ space, following the Lanza (2014) model. This required including the data point uncertainties both in the classification and fits: to do this, we first conservatively expanded

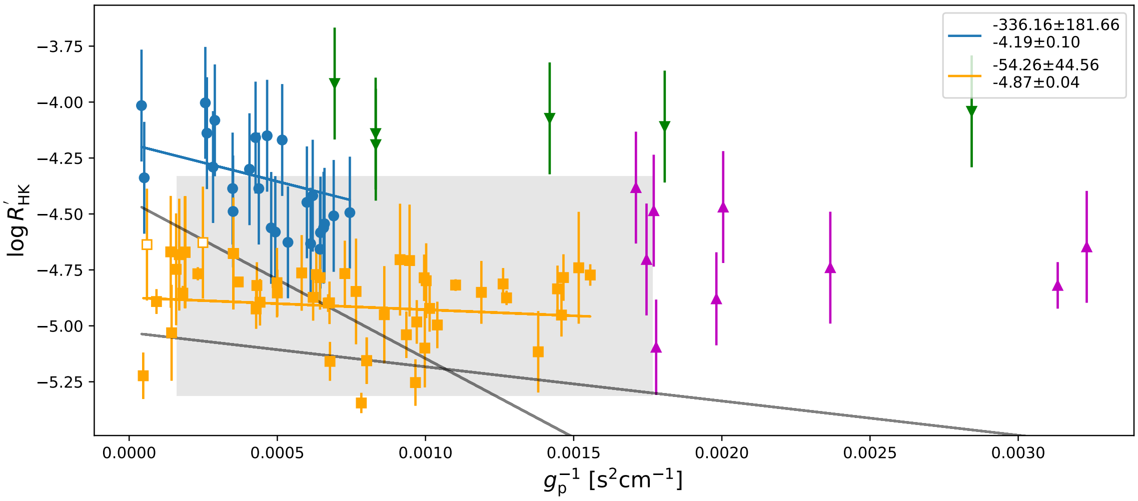

Fig. 15. Clustering in the $g_{\mathrm{p}}^{-1}-\log(R'_{\mathrm{HK}})$ space provided by our decision tree implementation. Each cluster is represented by a different color and marker. Linear fits of $\log(R'_{\mathrm{HK}})$ as a function of g_{p}^{-1}, as discussed in the text, are also shown. The portion of parameter space and the linear relationships found by Fossati et al. (2015) are shown with the gray region and black lines, respectively.

the error bars of the high-activity systems ($\log(R'_{\mathrm{HK}}) > -4.75$) to 0.25 dex, based on the values shown in Fig. 3 of Wright (2004). This allowed us to take into account realistic variations due to the activity cycle. We then bootstrapped our data set 1000 times, on the basis of each data point's Gaussian uncertainty in $\log(R'_{\mathrm{HK}})$. At every iteration, the previously-trained classification was carried out on the new data set, and linear fits were performed on the synthetic data set. For every realization, we only considered points with a >70% probability of being assigned to a given cluster, to then fit the populations with the linear model $\log R'_{\mathrm{HK}} = \log R'^{(0)}_{\mathrm{HK}} + \gamma g_{\mathrm{p}}^{-1}$. We obtained:

- $\gamma = -336.16 \pm 181.66$, $\log R'^{(0)}_{\mathrm{HK}} = -4.19 \pm 0.10$ for the high-activity cluster;
- $\gamma = -54.26 \pm 44.56$, $\log R'^{(0)}_{\mathrm{HK}} = -4.87 \pm 0.04$ for the low-activity cluster.

Both populations are in 3σ agreement with Fossati et al. (2015): the parameter space investigated by these authors and the trends they identified are shown in Fig. 15. The γ coefficients of both populations are, however, less significant than those found by Fossati et al., while the one on the high-activity population is more significant than the one found by Danielski et al. (2022). Such differences might relate to different factors, such as the kind of classification algorithm, the details of the conversion between the S and the $\log(R'_{\mathrm{HK}})$ activity indexes, and, most importantly, the size of the sample. All are aspects that play a key role in the classification process and identified correlations. Another relevant aspect is the probability of assignment to the high- or low-activity cluster or even to the high-activity and higher g_{p}^{-1} group, of single data points. For these reasons, we highlight the need to perform further investigations which use a larger homogeneous sample, to shed light on how planetary systems populate this parameter space. This might reduce the reliance on training information and allow for the use of unsupervised algorithms to increase the robustness of the clustering and, therefore, the estimates of the correlation coefficients.

9. Summary and conclusions

In this work, we have presented the characterization of a part of the hosts of transiting planets of the GAPS sample, mainly analyzing their stellar chromospheric activity (CA). In the process, we added to the observed data others found in the literature obtained for other hosts of transiting planets, enlarging the data set on which the cited previous works were based.

For each of the 132 stars in the sample, we obtained the stellar parameters (T_{eff}, $B-V$, V in a homogeneous way, exploiting the *Gaia* DR3 data release (Gaia Collaboration 2016, 2018, 2023). Subsequently, the complete set was compared with the activity levels of the field stars. To do this, we used the Mittag et al. (2013) procedure for calculating $\log(R'_{\mathrm{HK}})$ values. We report an activity index between −5.373 and −3.972, with a median value of −4.810. Somewhat more than half of the sample (58%) are inactive stars and this was expected, as most of our sample was selected from radial velocity follow-up campaigns, which privilege inactive stellar hosts. This study allowed us to determine a relationship between the $\log(R'_{\mathrm{HK}})$ values calculated with the classical procedure (Noyes et al. 1984) and those obtainable with Mittag's procedure, in the event that the S-index value is not available for the given object. This relationship has been determined both for MS and slightly evolved objects.

The analysis of the time series of $\log(R'_{\mathrm{HK}})$ for the GAPS objects observed for more than 20 nights showed that some of the targets have a dispersion value $\sigma R_5 > 0.5$ dex. In one case, Qatar-2, the high dispersion value (1.736 dex) and the low time-lapse of the time series seem to indicate a flare event. In general, as expected, the activity dispersion pattern is proportional to the activity level of the star, but in the cases of the F stars WASP-1, HAT-P-16, XO-4, and in that of the K star HAT-P-26, although these stars have a $\log(R'_{\mathrm{HK}})$ value included in the zone of inactive stars ($-5.1 \leq \log(R'_{\mathrm{HK}}) \leq -4.75$), they show a large variation in activity. This behavior confirms the high activity variation for K and F stars with low activity levels observed by Gomes da Silva et al. (2021).

We contrasted also the transiting planet hosts (both GAPS and those taken from the literature) with the field objects of both MS and slightly evolved by Pace (2013) after that we obtained also for field stars the stellar parameters by *Gaia*/DR3 in order to compare homogeneous samples. Kolmogorov-Smirnov tests show that the population of transiting objects is different from both populations (stars of MS and sub-giants) of the field objects. In addition, the population of transiting objects shows a higher percentage of objects (8% of the sample corresponding to 11 stars) with an activity level lower than the basal activity level for main sequence stars than in the case of field objects (1%). While three of these objects are subgiants, the other eight are stars of MS, for which there is a process of suppression of chromospheric line core fluxes that we have ascertained not to be due to the interstellar medium.

We analyzed the presence or the absence of possible correlations between the level of CA of the host stars and the stellar and planetary parameters of the systems considered. In particular, we tested the correlations of the activity level with the following planetary parameters: M_P, R_P, $\log g_\mathrm{P}$, ρ_P and a_P, the insolation at the top of the planet's atmosphere, the $T_{eq,p}$, the equilibrium temperature of the planet and the stellar equilibrium tide. The analysis has been performed for the whole sample and for a limited sub-set to $M_\mathrm{P} \geq 0.1 M_J$ and $a_p \leq 0.1$ au. We did not find significant correlations with a_P, the only planetary parameter linked to the existence of star-planet interaction. Furthermore, the lack of a correlation with M_P seems to contradict the Collier Cameron & Jardine (2018)'s explanation of the correlation firstly reported by Hartman (2010). Conversely, all correlations with FAP ≤ 1% are with planetary parameters proportional to R_p^n, indicating a common origin. We conclude that some planets evaporate by means of hydrodynamic escape under the strong irradiation of the host star and the evaporated matter forms a system of shrouds that absorbs the core of chromospheric resonance lines, as described by Lanza (2014). This is confirmed by the fact that all the systems with activity lower than the basal level host planets that are very close to filling their Roche lobe.

An analysis made with a machine learning technique supported the presence of four different clusters of hosts with different behavior concerning the activity of stars as a function of the planetary surface gravity. Two of these clusters were already individuated by Fossati et al. (2015). The third cluster, with average activity and low planetary surface gravity, has been interpreted as due to a ”saturated” regime for stellar irradiation-driven evaporation. The fourth cluster likely represents stars with a high level of activity and a strong stellar wind able to remove the shrouds of evaporated matter. So far, these interpretations are qualitative ones and they are the subject of ongoing and future works.

Acknowledgements. The authors are greateful to an anonymous referee for a careful reading of the manuscript and a very useful report that helped them to remarkably improve their work and its presentation. The GAPS project in Italy acknowledges the support by INAF through the Progetti Premiali funding scheme of the Italian Ministry of Education, University, and Research. Based on observations made with the Italian Telescopio Nazionale Galileo (TNG) operated on the island of La Palma by the Fundacion Galileo Galilei of the INAF (Istituto Nazionale di Astrofisica) at the Spanish Observatorio del Roque de los Muchachos of the Instituto de Astrofisica de Canarias This research has made use of the VizieR catalog access tool, CDS, Strasbourg, France (DOI: `10.26093/cds/vizier`) and of SIMBAD database, operated at CDS, Strasbourg, France. This work has made use of data from the European Space Agency (ESA) mission *Gaia* (`https://www.cosmos.esa.int/gaia`), processed by the *Gaia* Data Processing and Analysis Consortium (DPAC, `https://www.cosmos.esa.int/web/gaia/dpac/consortium`). Funding for the DPAC has been provided by national institutions, in particular, the institutions participating in the *Gaia* Multilateral Agreement.

References

Albrecht, S., Winn, J. N., Butler, R. P., et al. 2012, ApJ, 744, 189
Alspach, D., & Sorenson, H. 1972, IEEE Trans. Automatic Control, 17, 439
Anderson, E., & Francis, C. 2012, Astron. Lett., 38, 331
Anderson, D. R., Collier Cameron, A., Gillon, M., et al. 2011, A&A, 534, A16
Anderson, D. R., Collier Cameron, A., Delrez, L., et al. 2014, MNRAS, 445, 1114
Bailer-Jones, C. A. L., Rybizki, J., Fouesneau, M., Demleitner, M., & Andrae, R. 2021, AJ, 161, 147
Bailey, A., & Goodman, J. 2019, MNRAS, 482, 1872
Bakos, G. Á., Hartman, J. D., Torres, G., et al. 2012, AJ, 144, 19
Baliunas, S. L., Donahue, R. A., Soon, W. H., et al. 1995, ApJ, 438, 269
Ballerini, P., Micela, G., Lanza, A. F., & Pagano, I. 2012, A&A, 539, A140
Barclay, T., Kostov, V. B., Colón, K. D., et al. 2021, AJ, 162, 300
Barker, A. J. 2020, MNRAS, 498, 2270
Bastian, T. S., Dulk, G. A., & Leblanc, Y. 2000, ApJ, 545, 1058
Beck, P. G., Kambe, E., Hillen, M., et al. 2015, A&A, 573, A138
Béky, B., Bakos, G. Á., Hartman, J., et al. 2011, ApJ, 734, 109
Benatti, S. 2021, in Posters from the TESS Science Conference II (TSC2), 21
Biazzo, K., D'Orazi, V., Desidera, S., et al. 2022, A&A, 664, A161
Bonomo, A. S., Hébrard, G., Santerne, A., et al. 2012, A&A, 538, A96
Bonomo, A. S., Desidera, S., Benatti, S., et al. 2017, A&A, 602, A107
Borsa, F., Rainer, M., Bonomo, A. S., et al. 2019, A&A, 631, A34
Borucki, W. J., Koch, D. G., Batalha, N., et al. 2012, ApJ, 745, 120
Borucki, W. J., Agol, E., Fressin, F., et al. 2013, Science, 340, 587
Bressan, A., Marigo, P., Girardi, L., et al. 2012, MNRAS, 427, 127
Brown, D. J. A., Cameron, A. C., Anderson, D. R., et al. 2012, MNRAS, 423, 1503
Buchhave, L. A., Bakos, G. Á., Hartman, J. D., et al. 2011, ApJ, 733, 116
Canto Martins, B. L., Das Chagas, M. L., Alves, S., et al. 2011, A&A, 530, A73
Carleo, I., Benatti, S., Lanza, A. F., et al. 2018, A&A, 613, A50
Carleo, I., Malavolta, L., Lanza, A. F., et al. 2020, A&A, 638, A5
Carleo, I., Giacobbe, P., Guilluy, G., et al. 2022, AJ, 164, 101
Cegla, H. M., Watson, C. A., Shelyag, S., et al. 2018, ApJ, 866, 55
Claudi, R., Benatti, S., Carleo, I., et al. 2016, SPIE Conf. Ser., 9908, 99081A
Claudi, R., Benatti, S., Carleo, I., et al. 2017, Eur. Phys. J. Plus, 132, 364
Collier Cameron, A., & Jardine, M. 2018, MNRAS, 476, 2542
Cosentino, R., Lovis, C., Pepe, F., et al. 2012, SPIE Conf. Ser., 8446, 84461V
Covino, E., Esposito, M., Barbieri, M., et al. 2013, A&A, 554, A28
Cuntz, M., Saar, S. H., & Musielak, Z. E. 2000, ApJ, 533, L151
da Silva, L., Girardi, L., Pasquini, L., et al. 2006, A&A, 458, 609
Damasso, M., Lanza, A. F., Benatti, S., et al. 2020, A&A, 642, A133
Danielski, C., Brucalassi, A., Benatti, S., et al. 2022, Exp. Astron., 53, 473
Diplas, A., & Savage, B. D. 1994, ApJ, 427, 274
Dumusque, X., Bonomo, A. S., Haywood, R. D., et al. 2014, ApJ, 789, 154
Duncan, D. K., Vaughan, A. H., Wilson, O. C., et al. 1991, ApJS, 76, 383
Egeland, R., Soon, W., Baliunas, S., et al. 2017, ApJ, 835, 25
Esposito, M., Covino, E., Desidera, S., et al. 2017, A&A, 601, A53
Evans, D. W., Riello, M., De Angeli, F., et al. 2018, A&A, 616, A4
Faber, J. A., Rasio, F. A., & Willems, B. 2005, Icarus, 175, 248
Figueira, P., Oshagh, M., Adibekyan, V. Z., & Santos, N. C. 2014, A&A, 572, A51
Ford, E. B., & Rasio, F. A. 2006, ApJ, 638, L45
Fossati, L., Haswell, C. A., Froning, C. S., et al. 2010, ApJ, 714, L222
Fossati, L., Ayres, T. R., Haswell, C. A., et al. 2013, ApJ, 766, L20
Fossati, L., Ingrassia, S., & Lanza, A. F. 2015, ApJ, 812, L35
Fossati, L., Marcelja, S. E., Staab, D., et al. 2017, A&A, 601, A104
Gaia Collaboration (Prusti, T., et al.) 2016, A&A, 595, A1
Gaia Collaboration (Brown, A. G. A., et al.) 2018, A&A, 616, A1
Gaia Collaboration (Vallenari, A., et al.) 2023, A&A, 674, A1
Gautier, Thomas N., I., Charbonneau, D., Rowe, J. F., et al. 2012, ApJ, 749, 15
Geron, A. 2017, Hands-on Machine Learning with Scikit-Learn and TensorFlow (Sebastopol, CA: O'Reilly Media, Inc.)
Gilliland, R. L., Marcy, G. W., Rowe, J. F., et al. 2013, ApJ, 766, 40
Gomes da Silva, J., Santos, N. C., Adibekyan, V., et al. 2021, A&A, 646, A77
Gray, R. O., Corbally, C. J., Garrison, R. F., et al. 2006, AJ, 132, 161
Guilluy, G., Andretta, V., Borsa, F., et al. 2020, A&A, 639, A49
Gurdemir, L., Redfield, S., & Cuntz, M. 2012, PASA, 29, 141
Hall, J. C., Lockwood, G. W., & Skiff, B. A. 2007, AJ, 133, 862
Hartman, J. D. 2010, ApJ, 717, L138
Hartman, J. D., Bakos, G. Á., Torres, G., et al. 2011, ApJ, 742, 59
Hartman, J. D., Bakos, G. Á., Béky, B., et al. 2012, AJ, 144, 139
Hartman, J. D., Bakos, G. Á., Torres, G., et al. 2014, AJ, 147, 128
Haswell, C. A., Fossati, L., Ayres, T., et al. 2012, ApJ, 760, 79
Haswell, C. A., Staab, D., Barnes, J. R., et al. 2020, Nat. Astron., 4, 408
Hatzes, A. P. 1996, PASP, 108, 839
Hebb, L., Collier-Cameron, A., Loeillet, B., et al. 2009, ApJ, 693, 1920
Hébrard, G., Collier Cameron, A., Brown, D. J. A., et al. 2013, A&A, 549, A134
Henry, T. J., Soderblom, D. R., Donahue, R. A., & Baliunas, S. L. 1996, AJ, 111, 439
Howard, A. W., Johnson, J. A., Marcy, G. W., et al. 2011, ApJ, 730, 10

Hunter, A. A., Macgregor, A. B., Szabo, T. O., Wellington, C. A., & Bellgard, M. I. 2012, Source Code Biol. Med., 7, 1
Ingrassia, S., Minotti, S. C., & Vittadini, G. 2012, J. Classif., 29, 363
Kashyap, V. L., Drake, J. J., & Saar, S. H. 2008, ApJ, 687, 1339
Knutson, H. A., Howard, A. W., & Isaacson, H. 2010, ApJ, 720, 1569
Lallement, R., Babusiaux, C., Vergely, J. L., et al. 2019, A&A, 625, A135
Lallement, R., Vergely, J. L., Babusiaux, C., & Cox, N. L. J. 2022, A&A, 661, A147
Lammer, H., Selsis, F., Ribas, I., et al. 2003, ApJ, 598, L121
Lanza, A. F. 2014, A&A, 572, L6
Lecavelier des Étangs, A. 2007, A&A, 461, 1185
Lecavelier des Étangs, A., Vidal-Madjar, A., McConnell, J. C., & Hébrard, G. 2004, A&A, 418, L1
Lendl, M., Anderson, D. R., Collier-Cameron, A., et al. 2012, A&A, 544, A72
Lendl, M., Triaud, A. H. M. J., Anderson, D. R., et al. 2014, A&A, 568, A81
Locci, D., Cecchi-Pestellini, C., & Micela, G. 2019, A&A, 624, A101
Lovis, C., Dumusque, X., Santos, N. C., et al. 2011, arXiv e-prints [arXiv:`1107.5325`]
Madhusudhan, N. 2019, ARA&A, 57, 617
Maggio, A., Pillitteri, I., Scandariato, G., et al. 2015, ApJ, 811, L2
Malavolta, L., Borsato, L., Granata, V., et al. 2017, AJ, 153, 224
Maldonado, J., Scandariato, G., Stelzer, B., et al. 2017, A&A, 598, A27
Marcy, G. W., Isaacson, H., Howard, A. W., et al. 2014, ApJS, 210, 20
Maxted, P. F. L., Anderson, D. R., Collier Cameron, A., et al. 2011, PASP, 123, 547
Milbourne, T. W., Haywood, R. D., Phillips, D. F., et al. 2019, ApJ, 874, 107
Miller, B. P., Gallo, E., Wright, J. T., & Pearson, E. G. 2015, ApJ, 799, 163
Mittag, M., Schmitt, J. H. M. M., & Schröder, K. P. 2013, A&A, 549, A117
Montalto, M., Piotto, G., Marrese, P. M., et al. 2021, A&A, 653, A98
Murray-Clay, R. A., Chiang, E. I., & Murray, N. 2009, ApJ, 693, 23
Nardiello, D., Malavolta, L., Desidera, S., et al. 2022, A&A, 664, A163
Noyes, R. W., Bakos, G. Á., Torres, G., et al. 2008, ApJ, 673, L79
Noyes, R. W., Hartmann, L. W., Baliunas, S. L., Duncan, D. K., & Vaughan, A. H. 1984, ApJ, 279, 763
Ochsenbein, F., Bauer, P., & Marcout, J. 2000, A&AS, 143, 23
Pace, G. 2013, A&A, 551, L8
Pecaut, M. J., & Mamajek, E. E. 2013, ApJS, 208, 9
Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2011, J. Mach. Learn. Res., 12, 2825
Penz, T., Micela, G., & Lammer, H. 2008, A&A, 477, 309
Pepe, F., Mayor, M., Galland, F., et al. 2002, A&A, 388, 632
Pepe, F., Cameron, A. C., Latham, D. W., et al. 2013, Nature, 503, 377
Pillitteri, I., Wolk, S. J., Lopez-Santiago, J., et al. 2014, ApJ, 785, 145
Pino, L., Désert, J.-M., Brogi, M., et al. 2020, ApJ, 894, L27
Poppenhaeger, K., Robrade, J., & Schmitt, J. H. M. M. 2010, A&A, 515, A98
Poppenhaeger, K., & Schmitt, J. H. M. M. 2011, ApJ, 735, 59
Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 1992, Numerical recipes in C. The art of scientific computing (Cambridge: University Press)
Quinn, S. N., Bakos, G. Á., Hartman, J., et al. 2012, ApJ, 745, 80
Riello, M., De Angeli, F., Evans, D. W., et al. 2021, A&A, 649, A3
Robertson, P., Roy, A., & Mahadevan, S. 2015, ApJ, 805, L22
Saar, S. H. 2011, ASP Conf. Ser., 448, 1239
Saar, S. H., & Cuntz, M. 2001, MNRAS, 325, 55
Santos, N. C., Sousa, S. G., Mortier, A., et al. 2013, A&A, 556, A150
Sanz-Forcada, J., Ribas, I., Micela, G., et al. 2010, A&A, 511, L8
Sanz-Forcada, J., Micela, G., Ribas, I., et al. 2011, A&A, 532, A6
Sato, B., Hartman, J. D., Bakos, G. Á., et al. 2012, PASJ, 64, 97
Scandariato, G., Maggio, A., Lanza, A. F., et al. 2013, A&A, 552, A7
Scandariato, G., Maldonado, J., Affer, L., et al. 2017, A&A, 598, A28
Scharf, C. A. 2010, ApJ, 722, 1547
Schneider, J., Dedieu, C., Le Sidaner, P., Savalle, R., & Zolotukhin, I. 2011, A&A, 532, A79
Schrijver, C. J. 1987, A&A, 172, 111
Schröder, K. P., Mittag, M., Pérez Martínez, M. I., Cuntz, M., & Schmitt, J. H. M. M. 2012, A&A, 540, A130
Shkolnik, E., Walker, G. A. H., & Bohlender, D. A. 2003, ApJ, 597, 1092
Shkolnik, E., Walker, G. A. H., Rucinski, S. M., Bohlender, D. A., & Davidge, T. J. 2005, AJ, 130, 799
Shkolnik, E., Bohlender, D. A., Walker, G. A. H., & Collier Cameron, A. 2008, ApJ, 676, 628
Sorenson, H., & Alspach, D. 1971, Automatica, 7, 465
Sousa, S. G., Adibekyan, V., Delgado-Mena, E., et al. 2018, A&A, 620, A58
Sousa, S. G., Adibekyan, V., Delgado-Mena, E., et al. 2021, A&A, 656, A53
Staab, D., Haswell, C. A., Smith, G. D., et al. 2017, MNRAS, 466, 738
Triaud, A. H. M. J., Collier Cameron, A., Queloz, D., et al. 2010, A&A, 524, A25
Triaud, A. H. M. J., Queloz, D., Hellier, C., et al. 2011, A&A, 531, A24
Tsantaki, M., Sousa, S. G., Santos, N. C., et al. 2014, A&A, 570, A80
Vaughan, A. H., & Preston, G. W. 1980, PASP, 92, 385
Wakker, B. P., & Mathis, J. S. 2000, ApJ, 544, L107
Welsh, B. Y., Lallement, R., Vergely, J. L., & Raimond, S. 2010, A&A, 510, A54
Wenger, M., Ochsenbein, F., Egret, D., et al. 2000, A&AS, 143, 9
Winn, J. N., Albrecht, S., Johnson, J. A., et al. 2011, ApJ, 741, L1
Wright, J. T. 2004, AJ, 128, 1273
Wright, J. T. 2005, PASP, 117, 657
Wright, J. T., Marcy, G. W., Butler, R. P., & Vogt, S. S. 2004, ApJS, 152, 261
Wyman, K., & Redfield, S. 2013, ApJ, 773, 96
Zacharias, N., Finch, C. T., Girard, T. M., et al. 2012, VizieR Online Data Catalog: I/322A

---

[1] INAF – Osservatorio Astronomico di Padova, vicolo Osservatorio, 5, 35122 Padova, Italy
e-mail: `riccardo.claudi@oapd.inaf.it`
[2] Dipartimento di Matematica e Fisica, Università Roma Tre, Via della Vasca Navale 84, 00146 Roma, Italy
[3] INAF-Osservatorio Astrofisico di Catania, via S. Sofia, 78, 95123 Catania, Italy
[4] Space Research Institute, Austrian Academy of Sciences, Schmiedlstrasse 6, 8042 Graz, Austria
[5] INAF – Osservatorio Astronomico di Palermo, Piazza del Parlamento 1, 90134 Palermo, Italy
[6] INAF – Osservatorio Astronomico di Roma, Via Frascati, 33, 00078 Monte Porzio (RM), Italy
[7] INAF – Osservatorio Astronomico di Trieste, via Tiepolo 11, 34143 Trieste, Italy
[8] Instituto de Astrofísica de Canarias (IAC), 38205 La Laguna, Tenerife, Spain
[9] INAF – Osservatorio Astrofisico di Torino, via Osservatorio 20, 10025 Pino Torinese, Italy
[10] Instituto de Astrofísica de Andalucìa, CSIC, Glorieta de la Astronomía, 18080 Granada, Spain
[11] Sorbonne Universités, UPMC Université Paris 6 et CNRS, UMR 7095, Institut d'Astrophysique de Paris, 98 bis bd Arago, 75014 Paris, France
[12] Dipartimento di Fisica e Astronomia Galileo Galilei, Università di Padova,Vicolo dell'Osservatorio 2, 35122 Padova (Pd), Italy
[13] Dipartimento di Fisica, Università di Roma 'Tor Vergata', Via della Ricerca Scientifica 1, 00133 Rome, Italy
[14] Max Planck Institute for Astronomy, Königstuhl 17, 69117 Heidelberg, Germany
[15] INAF – Osservatorio Astronomico di Brera, via E. Bianchi 46, 23807 Merate, Italy
[16] Fundaciòn Galileo Galilei – INAF, Rambla José Ana Fernandez Pérez 7, 38712 Breña Baja, TF, Spain
[17] INAF – Osservatorio Astronomico di Capodimonte, Salita Moiariello 16, 80131, Napoli, Italy
[18] INAF – Osservatorio Astronomico di Cagliari, Via della Scienza 5, 09047, Selargius (CA), Italy

Appendix A: Additional tables

Table A.1. Journal of GAPS Observations.

System	T_0	Time span (d)	Spectra n.
GJ436	2018-04-01	698.1226	47
GJ3470	2018-01-13	744.2591	82
HAT-P-1	2012-08-08	2587.8794	55
HAT-P-2	2013-03-12	2526.9941	29
HAT-P-3	2013-01-30	791.6893	38
HAT-P-4	2013-02-01	2565.9444	30
HAT-P-6	2012-08-27	1859.9033	18
HAT-P-7	2013-04-24	2310.8561	21
HAT-P-8	2012-09-04	1423.1446	14
HAT-P-11	2019-07-07	0.1814	26
HAT-P-12	2015-03-14	41.9912	37
HAT-P-14	2013-03-25	1654.6531	14
HAT-P-15	2014-10-10	403.0163	34
HAT-P-16	2012-08-22	1849.9729	21
HAT-P-17	2013-10-13	2159.2268	53
HAT-P-18	2013-06-11	689.2777	28
HAT-P-20	2012-11-21	1136.8736	42
HAT-P-21	2013-01-03	1494.9679	59
HAT-P-22	2013-01-03	2628.7719	68
HAT-P-24	2012-11-21	1260.6462	17
HAT-P-26	2015-03-27	0.1882	27
HAT-P-29	2012-08-27	2684.7157	56
HAT-P-30	2013-01-03	1386.0594	15
HAT-P-31	2013-04-28	1142.8728	15
HAT-P-32	2017-11-17	731.139	32
HAT-P-36	2013-02-22	0.0866	9
HAT-P-47	2019-11-21	0.2101	31
HAT-P-50	2016-12-15	0.2822	40
HD17156	2012-08-22	1521.8131	25
HD80606	2020-02-02	0.4319	9
HD189733	2017-05-31	505.9913	123
HD209458	2018-07-08	423.1412	105
KELT-3	2018-01-19	0.1836	17
KELT-6	2014-02-10	854.8237	73
QATAR-1	2012-09-03	1857.105	27
QATAR-2	2014-04-27	0.1713	14
TRES-1	2013-04-23	1534.9043	20
TRES-2	2013-04-24	1533.9317	15
TRES-4	2013-03-11	1190.853	19
WASP-1	2012-08-21	1433.9444	13
WASP-3	2012-08-07	25.949	2
WASP-10	2012-11-12	0.1557	13
WASP-11	2012-10-02	2301.7128	50
WASP-12	2012-09-07	1955.9731	71
WASP-13	2012-12-05	2577.9926	41
WASP-14	2013-01-30	1229.7274	15
WASP-21	2013-07-11	1540.8895	10
WASP-24	2013-01-30	1281.6053	14
WASP-26	2012-09-05	1492.8679	12
WASP-31	2013-01-13	1166.7565	14
WASP-32	2012-09-06	1492.019	12
WASP-35	2012-09-06	2661.8378	40
WASP-38	2013-03-24	1041.062	18
WASP-39	2015-05-04	1099.1804	61
WASP-43	2012-12-27	1156.8908	54
WASP-48	2013-04-24	1533.9315	12
WASP-49	2013-12-30	1024.2611	9
WASP-50	2012-09-30	1190.6492	43
WASP-54	2013-01-13	1217.7109	16
WASP-57	2015-04-02	426.0631	20
WASP-59	2016-09-12	0.1925	19
WASP-60	2013-10-20	0.2334	21
WASP-69	2019-07-24	0.1742	25
WASP-76	2017-10-26	762.1462	110
WASP-80	2019-08-09	43.0984	32
WASP-106	2016-02-01	0.0239	3
WASP-107	2019-02-08	86.0099	38
WASP-127	2019-02-20	0.2575	25
XO-1	2013-03-25	1243.8207	13
XO-2N	2012-11-21	2263.9112	73
XO-3	2012-09-28	1481.9351	21
XO-4	2012-10-26	1164.83	17

Table A.2. Correction ($\Delta \log R'_{\rm HK}$) and corrected value of $\log R'_{\rm HK}$ for the ISM absorption for the systems of the sample. The highlighted systems have still $\log(R'_{\rm HK}) \leq -5.1$.

System	$\Delta \log R'_{\rm HK}$	$\log R'_{\rm HK}$	System	$\Delta \log R'_{\rm HK}$	$\log R'_{\rm HK}$	System	$\Delta \log R'_{\rm HK}$	$\log R'_{\rm HK}$
55Cnc	0.036	−4.8197	HAT-P-45	0.07	−5.0994	WASP-14	0.042	−4.8914
CoRoT-1	0.044	−5.2535	**HAT-P-46**	0.06	−5.1160	WASP-15	0.052	−4.7405
CoRoT-2	0.075	−4.1389	HAT-P-47	0.06	−4.6470	WASP-16	0.052	−4.8946
CoRoT-7	0.048	−4.5624	HD149026	0.04	−4.7937	WASP-17	0.063	−4.8192
GJ3470	0.040	−4.6296	HD17156	0.04	−4.8430	**WASP-18**	0.034	−5.2231
GJ436	0.028	−5.3447	HD189733	0.05	−4.1505	WASP-19	0.057	−4.4481
HAT-P-1	0.046	−4.8504	HD209458	0.04	−4.8178	WASP-21	0.052	−4.7728
HAT-P-2	0.045	−4.6369	HD80606	0.04	−4.8841	WASP-22	0.045	−4.7083
HAT-P-3	0.048	−4.6268	HD97658	0.04	−4.7709	WASP-23	0.051	−4.3866
HAT-P-4	0.043	−4.8215	KELT-3	0.05	−4.5809	WASP-24	0.053	−4.7846
HAT-P-5	0.048	−4.8729	KELT-6	0.05	−4.8749	WASP-26	0.039	−4.8971
HAT-P-6	0.063	−4.5834	Kepler-4	0.05	−4.6719	WASP-31	0.043	−4.8792
HAT-P-7	0.043	−5.0155	Kepler-5	0.06	−4.7970	WASP-32	0.050	−4.6694
HAT-P-8	0.048	−4.8982	Kepler-6	0.06	−4.7205	WASP-35	0.045	−4.7983
HAT-P-9	0.042	−4.9961	Kepler-7	0.06	−4.7558	WASP-38	0.047	−4.8540
HAT-P-11	0.053	−4.2487	Kepler-8	0.06	−4.8347	WASP-39	0.055	−4.7400
HAT-P-12	0.053	−4.4853	Kepler-10	0.05	−4.7575	WASP-41	0.057	−4.4184
HAT-P-13	0.042	−4.8525	Kepler-16	0.06	−4.6955	WASP-42	0.061	−4.7042
HAT-P-14	0.048	−4.7672	Kepler-17	0.07	−4.0817	WASP-43	0.050	−4.0022
HAT-P-15	0.077	−4.6795	Kepler-19	0.05	−4.7635	WASP-48	0.051	−4.6730
HAT-P-16	0.049	−4.7475	Kepler-20	0.05	−4.6767	WASP-49	0.046	−4.8127
HAT-P-17	0.050	−4.8463	Kepler-22	0.05	−4.8650	WASP-50	0.047	−4.8038
HAT-P-18	0.059	−4.4692	**Kepler-25**	0.04	−5.1557	WASP-52	0.059	−4.0728
HAT-P-20	0.050	−4.0156	Kepler-48	0.06	−4.5892	WASP-54	0.042	−4.9173
HAT-P-21	0.051	−4.2574	Kepler-62	0.06	−4.6710	WASP-57	0.061	−4.5087
HAT-P-22	0.042	−4.7770	Kepler-68	0.04	−4.9659	WASP-58	0.057	−4.1409
HAT-P-24	0.046	−4.9502	Kepler-78	0.06	−4.1905	WASP-59	0.059	−4.2904
HAT-P-25	0.071	−4.7852	Kepler-93	0.05	−4.7541	WASP-60	0.055	−4.8047
HAT-P-26	0.051	−4.8018	QATAR-1	0.06	−4.1587	WASP-69	0.053	−4.1095
HAT-P-27	0.056	−4.5616	QATAR-2	0.06	−4.0035	WASP-70A	0.043	−5.0399
HAT-P-28	0.057	−4.7338	TRES-1	0.06	−4.5441	**WASP-72**	0.036	−5.1549
HAT-P-29	0.061	−4.7664	TRES-2	0.05	−4.8071	WASP-76	0.042	−4.9288
HAT-P-30	0.035	−4.9826	Tres-3	0.06	−4.3001	WASP-80	0.051	−4.1197
HAT-P-31	0.046	−4.9776	TRES-4	0.04	−4.9925	WASP-84	0.048	−4.1695
HAT-P-32	0.065	−4.3822	WASP-1	0.05	−4.9381	WASP-103	0.066	−4.3406
HAT-P-33	0.046	−4.7034	WASP-2	0.06	−4.8529	WASP-106	0.046	−4.6278
HAT-P-34	0.053	−4.6819	WASP-3	0.06	−4.9245	WASP-107	0.050	−4.0414
HAT-P-35	0.033	−5.1588	WASP-4	0.05	−4.6329	WASP-117	0.040	−4.7842
HAT-P-36	0.050	−4.3864	WASP-5	0.05	−4.4888	WASP-127	0.041	−4.7504
HAT-P-38	0.060	−4.9221	WASP-7	0.06	−3.9172	XO-1	0.051	−4.6576
HAT-P-39	0.054	−4.6221	WASP-10	0.06	−5.0311	XO-2N	0.050	−4.6898
HAT-P-40	0.057	−4.8252	WASP-11	0.07	−4.4940	XO-3	0.074	−4.3384
HAT-P-41	0.055	−4.9518	**WASP-12**	0.04	−5.3203	XO-4	0.053	−4.9380
HAT-P-44	0.043	−5.0956	WASP-13	0.04	−4.8916			

Appendix B: Comparison between CA index distributions of field and transit host stars

The comparison of the four statistical moments of each distribution (see Table B.1) reveals that all are leptokurtic ($\mathrm{kurt}(X) > 0$), and asymmetric distributions towards the higher values of $\log(R'_{\rm HK})$ ($\mathrm{skew}(X) > 0$). In this contest, the host stars are more dispersed with respect to field stars, including both M-S and evolved field stars ($\sigma_{\rm host} > \sigma_{\rm all} > \sigma_{\rm M\text{-}S} > \sigma_{\rm sub\text{-}giant}$). On the contrary, the host stars distribution is more symmetric with respect to the others ($\mathrm{kurt}_{\rm host} < \mathrm{kurt}_{\rm M\text{-}S} < \mathrm{kurt}_{\rm all} < \mathrm{kurt}_{\rm sub\text{-}giant}$).

Table B.1. Comparison of the values of the distribution moments of the different populations of objects.

	Hosts	Field M-S stars	Field Sub-Giants	Field stars (all)
Mean (μ)	−4.7429	−4.5862	−4.7307	−4.6231
Variance (σ)	0.088	0.063	0.055	0.065
Skewness ($\mathrm{skew}(X)$)	0.015	0.605	3.739	0.790
Kurtosis ($\mathrm{kurt}(X)$)	0.626	0.542	1.577	0.724

A Kolmogorov-Smirnov test yields very low probabilities for the null hypothesis: the probability that the CA distributions for M-S field stars and planet hosts are drawn from the same parent population (Figure 12 top-right panel) is *Prob* = 1.29×10^{-13} (D=0.35, N_e = 120), while for evolved stars (Figure 12 bottom-right panel) we found *Prob* = 4.4×10^{-2} (D=0.1, N_e = 103). Considering instead the whole distributions for field stars and planet hosts, we obtained *Prob* = 4.14×10^{-9} (D=0.3, N_e = 122) that they are drawn from the same parent population. It is worth noting that the field stars' bimodality indicated by the different average values of $\log(R'_{\rm HK})$ for M-S (−4.5862) and sub-giant (−4.7307), is expected (Wright 2004; Gray et al. 2006). In conclusion, the M-S and evolved stars' CA distribution comes from a different parent distribution of the transit host stars.

Appendix C: Details of machine learning analysis

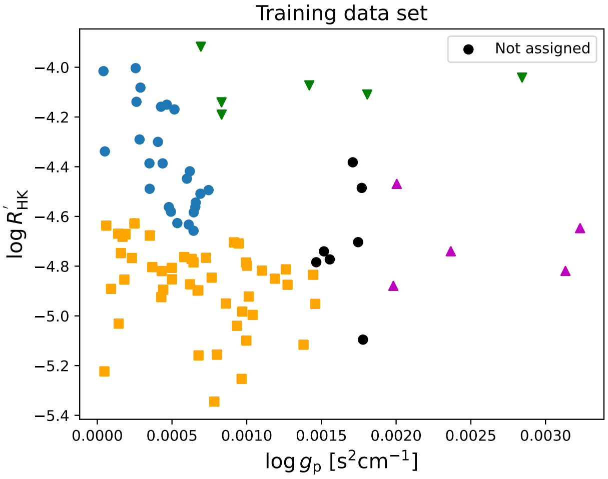

Fig. C.1. Training clustering provided to the algorithm in the $\log(R'_{\rm HK}) - g_{\rm p}^{-1}$ space. Black points were not assigned to any cluster and were used as test data.

In machine learning investigations, data points are represented by vectors containing different "features", and classification methods divide *n*-dimensional samples into clusters according to different possible metrics. Another fundamental distinction is between unsupervised and supervised methods: the former do not rely on prior, or "training", information, contrarily to the latter. Following Fossati et al. (2015), we first attempted a classification without prior information and assumed our data set could be represented by a finite number of Gaussian distributions. As shown by Danielski et al. (2022), this can be done with a Gaussian Mixture Model (GMM, Sorenson & Alspach 1971; Alspach & Sorenson 1972).

Before running the classification, we prepared our data set by normalizing it, in order to assign equal weight to each feature. This is usually done by removing the data mean and by scaling it to unit variance:

$$Z = \frac{X - \mu}{\sigma}, \tag{C.1}$$

where Z is the transformed data set, X is the raw data set, and μ and σ are its mean and standard deviation across every feature or parameter.

Then we adopted the GMM implementation in SCIKIT-LEARN (Pedregosa et al. 2011) and tested it by imposing an increasing number of sub-clusters. Even if the GMM was able to reproduce Fossati et al. (2015)'s result, we noticed its tendency to prefer three mixtures on our larger data set, where the entire sample of Fossati et al. (2015) is included in a single cluster. This can probably be explained by the different way clusters are formed by a GMM and by the cluster-weighted model adopted in their study. While in the former case, the data points are assumed to populate the $\log(R'_{\rm HK}) - g_{\rm p}^{-1}$ parameter space according to a number of multivariate Gaussian distributions, each with a center and a covariance matrix, the latter associates them to linear relationships following Lanza (2014)'s model.

We then preferred to use a supervised learning algorithm. After testing several algorithms, we chose a Decision Tree approach, which learns information from data using simple decision rules (e.g. Geron 2017). Once again, we relied on SCIKIT-LEARN's implementation of this algorithm and chose a maximum depth of 4, that is, the one that allowed us to reproduce a close clustering to the one of Fossati et al. (2015)'s on their data.

We trained the algorithm after visual inspection on how four clusters might be identified in the data set (see Sect. 8 in the main text). We assigned 94% of the data point to different clusters, as shown by the colours used in Figure C.1, leaving out a few points at the edge of the clusters. A cross-validation test (e.g. Geron 2017) yielded an accuracy of (91 ± 6)% for our combination of algorithms and training data. Then, we ran the Decision Tree on the whole data set and obtained the classification shown in Figure 15 in the main text. We remark that we could not use a classical "score test" to measure the accuracy of the prediction, as we worked in a framework of multi-label, and not binary, classification (see SCIKIT-LEARN's documentation of the Decision Tree algorithm).