Froebrich, Dirk, Scholz, Aleks, Campbell-White, Justyn, Vanaverbeke, Siegfried, Herbert, Carys, Eislöffel, Jochen, Urtly, Thomas, Long, Timothy P, Walton, Ivan L, Wiersema, Klaas and others (2024) *A survey for variable young stars with small telescopes – VIII. Properties of 1687 Gaia selected members in 21 nearby clusters.* Monthly Notices of the Royal Astronomical Society, 529 (2). pp. 1283-1298. ISSN 0035-8711.

The version of record is available from

https://doi.org/10.1093/mnras/stae311

This document version

Publisher pdf

DOI for this version

Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **529,** 1283–1298 (2024) https://doi.org/10.1093/mnras/stae311
Advance Access publication 2024 February 1

A survey for variable young stars with small telescopes – VIII. Properties of 1687 *Gaia* selected members in 21 nearby clusters

Dirk Froebrich,[1]★ Aleks Scholz,[2] Justyn Campbell-White,[3] Siegfried Vanaverbeke,[4,5,6]
Carys Herbert,[1] Jochen Eislöffel,[7] Thomas Urtly,[8]† Timothy P. Long,[9]† Ivan L. Walton,[8]†
Klaas Wiersema,[10,11]† Nick J. Quinn,[8]† Tony Rodda,[8]† Juan-Luis González-Carballo,[12,13,14]† Mario
Morales Aimar,[13,15]† Rafael Castillo García,[13,14,16]† Francisco C. Soldán Alfaro,[13,14,17]† Faustino García
de la Cuesta,[18]† Domenico Licchelli,[19,20]† Alex Escartin Perez,[13,21,22]† José Luis Salto González,[13,23,24]†
Marc Deldem,[14]† Stephen R. L. Futcher,[8,14,25,26]† Tim Nelson,[25,27]† Shawn Dvorak,[14,28]†
Dawid Moździerski,[29]† Krzysztof Kotysz,[29]† Przemysław Mikołajczyk,[29,30]† George Fleming,[8]†
Mark Phillips,[8,31]† Tony Vale,[8,14,32,33,34]† Yenal Öğmen,[35]† Franky Dubois,[4,5]† Samantha M. Rolfe,[36]†
David A. Campbell,[36]† Heinz-Bernd Eggenstein,[37]† Franz-Josef Hambsch,[5,14,38,39]† Michael
A. Heald,[14]† Pablo Lewin,[40]† Adam C. Rose,[14]† Geoffrey Stone,[14,41]† Martin Valentine Crow,[8,42,43]†
Simon Francis Dawes,[8,43]† Derek OKeeffe,[44]† Adam Popowicz,[45]† Krzysztof Bernacki,[45]†
Andrzej Malcher,[45]† Slawomir Lasota,[45]† Jerzy Fiolka,[45]† Adam Dustor,[46]† Amritanshu Vajpayee,[47,48]†
Pat Devine,[31]† Matthias Kolb,[38,49]† Jean-Baptiste Marquette,[50]† Gregg L. Ruppel,[51]† Dan R. Crowson,[52]†
Cledison Marcos da Silva,[14,53,54]‡ Michel Michaud,[55]‡ Aashini L. Patel,[1]$ Matthew D. Dickers,[1]$
Lord Dover,[1]$ Ivana I. Grozdanova,[1]$ James S. Urquhart[1] and Chris J.R. Lynch[1]

Affiliations are listed at the end of the paper

Accepted 2024 January 29. Received 2024 January 22; in original form 2023 November 23

ABSTRACT

The Hunting Outbursting Young Stars (HOYS) project performs long-term, optical, multifilter, high cadence monitoring of 25 nearby young clusters and star-forming regions. Utilizing *Gaia* DR3 data, we have identified about 17 000 potential young stellar members in 45 coherent astrometric groups in these fields. Twenty one of them are clear young groups or clusters of stars within 1 kpc and they contain 9143 *Gaia* selected potential members. The cluster distances, proper motions, and membership numbers are determined. We analyse long-term ($\approx$ 7 yr) *V*-, *R*-, and *I*-band light curves from HOYS for 1687 of the potential cluster members. One quarter of the stars are variable in all three optical filters, and two-thirds of these have light curves that are symmetric around the mean. Light curves affected by obscuration from circumstellar materials are more common than those affected by accretion bursts, by a factor of 2–4. The variability fraction in the clusters ranges from 10 per cent to almost 100 per cent, and correlates positively with the fraction of stars with detectable inner discs, indicating that a lot of variability is driven by the disc. About one in six variables shows detectable periodicity, mostly caused by magnetic spots. Two-thirds of the periodic variables with disc excess emission are slow rotators, and amongst the stars without disc excess two-thirds are fast rotators – in agreement with rotation being slowed down by the presence of a disc.

Key words: stars: formation – stars: pre-main-sequence – stars: rotation – stars: variables: T Tauri, Herbig Ae/Be.

1 INTRODUCTION

Variability is a characteristic feature of young stellar objects (YSOs, Joy 1945). It is caused by strong magnetic activity (i.e. spots and flares), accretion from protoplanetary discs and its variations, or obscurations by disc material. In many objects multiple sources of variability can be identified. Because variability is prevalent in YSOs, it is a helpful indicator to identify regions of active star formation, even with only a few epochs of data. Beyond that,

---

★ E-mail: df@kent.ac.uk
† HOYS Observer.
‡ HOYS Data Processor
$ Beacon Observer.

extensive monitoring campaigns have been used to study rotation, accretion, and disc structure in YSOs (see Protostars and Planets reviews by: Herbst et al. 2007; Bouvier et al. 2007, 2014; Audard et al. 2014; Fischer et al. 2023).

Variability studies of young stars are ideally conducted quasi-simultaneously in multiple bands in the optical, to gain information on colour changes, which helps to disentangle the various sources of photometric variations. The CSI 2264 program (e.g. Cody et al. 2014) has set the tone for comprehensive multifilter campaigns, including optical and infrared photometry, albeit limited to one region and a 30 d time window. To obtain a comprehensive picture, the typical time-scales of the variability, from hours to years, need to be covered. In the past, datasets that combine these two characteristics were rarely available. In particular, the long-term coverage was often missing, or only available in one band (with exceptions, see Grankin et al. 2007). The optical time-domain coverage of the sky has much improved over the last decade, thanks to dedicated missions to find supernovae (e.g. ASAS-SN: Kochanek et al. 2017) or to hunt for exoplanets (e.g. *Kepler* and *TESS*), and thanks to the time-domain capacity of *Gaia*. None of these however is particularly well suited for YSO variability. Therefore, there is still a need for dedicated multifilter, multitime-scale studies of specific star-forming regions.

The project Hunting Outbursting Young Stars (HOYS) is one of the endeavours designed to overcome this deficit, by collecting observations from a multitude of small telescopes from professional and amateur observatories, all using a pre-defined observing strategy and a specific set of target regions (see overview paper by Froebrich et al. 2018). In the first HOYS papers we have, for example, presented a study of rotation in one star-forming region (Froebrich et al. 2021), a follow-up study of spot properties in the same region (Herbert, Froebrich & Scholz 2023), and an analysis of the variability in a specific source that is caused by eclipses from structures in a protoplanetary disc (Evitts et al. 2020), among other findings. In this paper, we take a more holistic view – we aim to select prospective young stars in all 25 HOYS target regions, using data from *Gaia*, collate information about these samples, and study general variability properties among those samples. The precise kinematic information from *Gaia* is key for this paper, as it enables a systematic and robust selection of members of young clusters based primarily on kinematic information, parallax, and proper motion, as demonstrated for example by Kuhn et al. (2020), Pavlidou, Scholz & Teixeira (2021), Kounkel, Deng & Stassun (2022), and Luhman (2023).

This paper is structured as follows. In Section 2. we detail the analysis methods used to identify cluster members, obtain their light curves and characterize them. We discuss the light curve properties in general and on a cluster-by-cluster basis in Section 3.

2 DATA ANALYSIS

In this section, we describe the selection of our samples using kinematics and photometry, and the light curve analysis. In all steps, we aim to identify reliable cluster members with good photometry in the HOYS data base, by including several conservative limits. That means in turn that our samples do not constitute a complete list of cluster members for individual regions – this is not the goal of this analysis.

2.1 Cluster selection

We are basing our selection of cluster members on the astrometry supplied by *Gaia* DR3 (Gaia Collaboration 2016, 2023). For each of the 25 HOYS fields,[1] we downloaded all *Gaia* sources within a radius of 0.6 deg from the cluster centre. All sources with a parallax of less than 0.3 mas, a parallax signal-to-noise ratio (SNR) below five and $G_{\rm mag}$ fainter than 18 mag were removed from our analysis. This ensures only stars with accurate distances and photometry are included in the analysis. Furthermore, we remove all blue sources with colours $\mathrm{BP} - G_{\rm mag} < -0.2$ mag and $G_{\rm mag} - \mathrm{RP} < 0.0$ mag. This is the colour range for white dwarfs and cataclysmic variables, which we do not expect in our target regions.

We manually identified the peaks in the histogram of the distances (assumed to be one over the parallax) for each field, using a bin width of 20 pc. All stars around each peak were selected. The distance range selected for each cluster is listed in Table 1. The proper motion distributions for the selected stars were searched for groups with coherent proper motions. All stars within a group showing coherent parallax and proper motion were selected as potential cluster members. Fig. 1 shows one example (the main cluster in the IC 348 field) for this manual selection process. In Table 1, we list for each field and cluster the ranges for distances and proper motion used in the selection. For each group, we determined the median parallax and proper motion, together with their standard deviation and the standard error of the median. These are also listed in Table 1. Note that for some clusters we only used a subpopulation within a small region centred around the cluster coordinates to estimate the median and scatter for distance and proper motion. That radius is listed in Table 1. The final selection of potential *Gaia* detected cluster members, that is, all stars that are within three sigma from the median values in distance and proper motion, was done within 0.6 deg from the cluster centre. The number of these *Gaia* selected potential cluster members is also listed in Table 1.

In many of the HOYS target regions this procedure resulted in more than one group of stars. In total 45 such groups or clusters of stars have been identified in the 25 HOYS target regions. These are all listed in Table 1. Typically there is a main cluster of objects dominating the field, which is the intended HOYS target. The secondary clusters are usually subpopulations with different proper motions at the same distance or fore/background clusters/groups in the same field of view. In Appendix A, we briefly discuss all the clusters/groups identified in each HOYS target field.

In total, about 17 000 unique Gaia sources are selected as potential members of the 45 clusters. A subsample of these could in principle be considered as members of several of the clusters due to spatial and/or proper motion range overlap of the populations. Hence, every *Gaia* source has been assigned to the cluster to which it most likely belongs, based on the deviation of its parallax and proper motion values from the median value of all potential cluster members. For more details on the samples for individual clusters, we refer again to Table 1.

2.2 HOYS light curve selection

For each identified cluster/group, we extracted all light curves from the HOYS data base,[2] on 2022 October 21. Prior to this, all pairs of *Gaia* sources with separations of less than 3 arcsec were identified and the fainter of the sources were removed. This was done to avoid false cross-matches of HOYS photometry to fainter *Gaia* targets, which are most likely not correct, given the typical seeing of three arcseconds in the HOYS data (Froebrich et al. 2018;

---

[1] HOYS Target list: https://hoys.space/target-lists/

[2] HOYS data base: http://astro.kent.ac.uk/HOYS-CAPS/

Table 1. This table contains the results from the *Gaia* DR3 data selection for all HOYS fields. The first column is the cluster name. The main cluster (a) is the most populous young cluster associated with the HOYS target field. All other subpopulations are numbered consecutively and do have either a different proper motion and/or distance from the main cluster. In some cases, the secondary clusters are known and we used their common names (e.g. NGC 7142 in the NGC 7129 field, or IC 4996 in the P Cyg field).

Name	r	d_1	d_2	PM$_\alpha$	PM$_\delta$	PM$_r$	d	d^e	d^s	μ_α	μ_α^e	μ_α^s	μ_δ	μ_δ^e	μ_δ^s	*N Gai a*		N	Incl?
	[deg]	[pc]	[pc]		[mas yr^{-1}]		[pc]	[pc]	[pc]		[mas yr^{-1}]			[mas yr^{-1}]		man.	3 σ	LC	Y/N
Clusters/subgroups used in analysis																			
IC 348 a	0.3	250	360	4.5	−6.4	1.4	318.1	1.19	23.6	4.478	0.039	0.524	−6.345	0.039	0.518	178	272	70	YYY
IC 348 b	0.6	250	360	6.8	−9.4	1.0	297.0	4.0	17.9	6.723	0.097	0.433	−9.738	0.057	0.253	20	33	5	YYN
λ-Ori a	0.4	330	450	0.7	−2.0	0.4	398.9	1.6	14.8	0.780	0.018	0.168	−2.005	0.017	0.154	85	191	25	YYY
λ-Ori b	0.6	330	450	1.8	−2.1	0.4	403.7	1.85	13.5	1.652	0.024	0.172	−2.138	0.021	0.151	53	91	7	YYN
M 42 a	0.3	300	500	1.3	−0.1	3.0	400.4	0.89	23.2	1.307	0.029	0.76	0.215	0.04	1.036	684	2157	250	YYY
L 1641 N a	0.3	320	440	1.0	−0.3	1.4	388.4	1.29	16.8	1.034	0.038	0.49	0.327	0.045	0.579	168	978	97	YYY
σ-Ori a	0.3	350	450	1.7	−0.4	0.7	404.5	1.36	15.0	1.618	0.031	0.347	−0.491	0.026	0.29	122	210	61	YYY
σ-Ori c	0.6	330	400	1.9	−1.6	0.5	368.5	1.68	12.9	1.803	0.022	0.165	−1.475	0.023	0.175	59	132	16	YYN
NGC 2264 a	0.6	650	850	−1.5	−3.7	0.5	732.0	2.1	38.6	−1.597	0.01	0.181	−3.635	0.011	0.201	339	423	163	YYY
NGC 2264 b	0.6	650	850	−2.5	−3.7	0.5	737.3	2.74	41.5	−2.425	0.015	0.231	−3.682	0.012	0.187	229	302	128	YYY
V 898 Ori a	0.6	300	500	0.5	−0.3	1.4	402.8	3.77	35.0	0.222	0.065	0.601	−0.31	0.064	0.592	86	194	22	YYY
YY Ori a	0.3	330	430	1.3	0.5	0.5	386.9	0.94	13.4	1.29	0.013	0.179	0.553	0.014	0.202	204	712	66	YYY
V 555 Ori a	0.3	340	460	1.2	0.3	1.9	393.4	0.88	16.8	1.229	0.028	0.543	−0.275	0.043	0.833	367	1569	118	YYY
IC 5070 a	0.6	750	900	−1.4	−3.1	0.75	832.4	2.7	33.6	−1.328	0.028	0.347	−3.076	0.026	0.320	155	252	84	YYY
IC 5070 b	0.6	750	900	−0.9	−4.3	0.7	824.7	4.43	36.0	−0.977	0.036	0.294	−4.148	0.041	0.335	66	114	47	YYY
IC 1396 A a	0.6	850	1050	−2.4	−4.8	0.8	942.0	2.39	45.4	−2.418	0.016	0.311	−4.724	0.015	0.293	361	549	221	YYY
IC 1396 N a	0.6	850	1050	−1.9	−4.1	1.2	955.0	3.1	43.9	−2.021	0.03	0.428	−3.964	0.04	0.568	200	642	168	YYY
IC 1396 N b	0.6	580	650	−6.2	−5.9	0.4	619.8	2.69	10.8	−6.139	0.031	0.124	−5.826	0.031	0.125	16	30	13	YYN
NGC 7129 a	0.25	850	950	−1.7	−3.4	0.7	910.0	4.11	21.3	−1.727	0.037	0.191	−3.38	0.054	0.282	27	55	28	YYY
IC 5146 a	0.3	600	900	−2.8	−2.6	1.0	784.8	5.07	51.2	−2.845	0.034	0.342	−2.704	0.039	0.397	102	178	75	YYY
Clusters/subgroups identified but not analysed in detail																			
σ-Ori b	0.6	350	450	−2.2	1.1	0.9	421.9	3.66	19.0	−2.135	0.063	0.328	1.27	0.066	0.341	27	49	15	NYN
NGC 2068 a	0.6	300	480	−1.2	−1.0	0.9	420.0	2.34	18.1	−0.982	0.05	0.385	−1.015	0.042	0.324	60	73	0	NYN
NGC 2068 b	0.6	300	480	0.1	−0.6	0.8	417.9	2.44	17.8	0.066	0.051	0.369	−0.555	0.044	0.323	53	69	0	NYN
NGC 2068 c	0.6	300	480	−2.9	1.2	0.7	379.3	6.86	24.7	−2.882	0.048	0.173	1.194	0.074	0.268	13	32	0	NYN
NGC 2244 a	0.3	1420	1600	−1.7	0.2	0.5	1512.7	3.94	51.1	−1.751	0.013	0.166	0.221	0.015	0.198	168	502	136	YNY
NGC 2244 b	0.6	590	710	−1.7	−4.7	0.6	658.3	4.52	26.0	−1.707	0.033	0.188	−4.762	0.027	0.154	33	69	18	NYN
ASASSN-13DB a	0.6	290	440	1.3	−1.0	0.5	387.3	3.16	18.4	1.306	0.029	0.171	−0.964	0.042	0.244	34	57	0	NYN
ASASSN-13DB b	0.6	290	440	0.5	−1.4	0.6	346.2	5.28	23.6	0.367	0.043	0.193	−1.24	0.057	0.254	20	37	0	NYN
V 555 Ori b	0.6	340	460	−1.8	1.4	0.7	419.4	3.75	14.5	−1.737	0.047	0.182	1.401	0.056	0.215	15	25	0	NYN
Gaia 17 bpi a	0.3	1200	1350	−1.1	−5.4	1.1	1300.3	6.26	42.9	−0.997	0.049	0.338	−5.654	0.058	0.396	47	289	39	YNY
Gaia 19 fct a	0.3	1000	1300	0.2	−0.1	0.35	1127.1	4.43	58.0	0.251	0.009	0.114	−0.158	0.009	0.112	172	247	100	YNY
Gaia 19 fct b	0.6	1000	1400	−3.5	0.9	0.9	1230.9	7.61	88.1	−3.565	0.037	0.424	1.017	0.028	0.319	134	411	57	YNY
Gaia 19 eyy a	0.3	1300	1500	−1.8	1.1	0.3	1405.5	5.48	45.2	−1.771	0.011	0.092	1.107	0.011	0.089	68	131	50	YNY
Gaia 19 eyy b	0.4	1200	1400	−5.3	5.0	0.4	1274.2	5.69	39.0	−5.407	0.016	0.109	4.96	0.023	0.155	47	101	21	YNY
Gaia 19 eyy c	0.6	1900	3000	−4.9	4.5	0.3	2487.9	25.09	257.1	−4.872	0.01	0.1	4.476	0.01	0.103	105	274	22	YNY
MWSC 3274 a	0.6	830	970	−1.1	−1.6	0.3	884.5	3.45	29.9	−1.145	0.012	0.101	−1.589	0.009	0.081	75	93	30	NYY
IC4996	0.25	1900	2400	−2.6	−5.3	0.3	2158.7	8.84	117.3	−2.604	0.008	0.108	−5.346	0.008	0.108	176	962	294	YNY
Berkeley 86 a	0.3	1600	2000	−3.5	−5.5	0.35	1839.0	7.82	92.2	−3.451	0.012	0.137	−5.479	0.013	0.149	139	990	217	YNY
Berkeley 86 b	0.4	1600	2000	−3.2	−6.0	0.2	1852.5	7.89	75.6	−3.246	0.008	0.077	−6.008	0.008	0.078	92	323	93	YNY
Berkeley 86 c	0.4	1440	1840	−0.7	−3.1	0.2	1636.0	12.66	86.8	−0.695	0.01	0.069	−3.081	0.009	0.061	47	113	16	YNN
NGC 7142	0.30	2300	2800	−2.7	−1.4	0.25	2563.8	8.45	129.8	−2.675	0.005	0.076	−1.355	0.005	0.076	236	329	286	NNY
IC 5146 b	0.6	1100	1900	1.3	−2.4	0.3	1531.0	15.61	120.9	1.276	0.017	0.135	−2.376	0.014	0.106	60	126	40	NNY
FSR 408 a	0.6	750	950	−0.8	−2.3	1.0	860.2	1.78	41.9	−0.876	0.017	0.405	−2.29	0.014	0.325	556	2418	0	NYN
FSR 408 b	0.6	750	950	−2.3	−2.5	0.4	875.6	3.78	34.9	−2.334	0.017	0.159	−2.484	0.019	0.178	85	344	0	NYN

Notes. The other columns are: r: the radius around the central coordinates from which *Gaia* DR3 sources are analysed when determining the cluster distance and proper motion; d_1, d_2: distance range in which cluster members are selected from; PM$_\alpha$, PM$_\delta$, PM$_r$: proper motions and range in which cluster members are selected from for distance and proper motion determination; d, d^e, d^s: median distance, standard error of the median, rms scatter of the cluster members; $\mu_{\alpha/\delta}$, $\mu^e_{\alpha/\delta}$, $\mu^s_{\alpha/\delta}$: median proper motion in RA/Dec., standard error of the median, rms scatter of the cluster members; *N Gaia* man.: number of *Gaia* stars selected manually for the cluster property determination; *N Gaia* 3σ: number of *Gaia* stars selected as potential cluster members; and *N* LC: number of stars with HOYS light curves; incl? Y/N: inclusion flag; Does the CMD look like a young cluster Y/N? / Is the cluster closer than 1 kpc Y/N?/Are there more than 20 stars with HOYS light curves Y/N?.

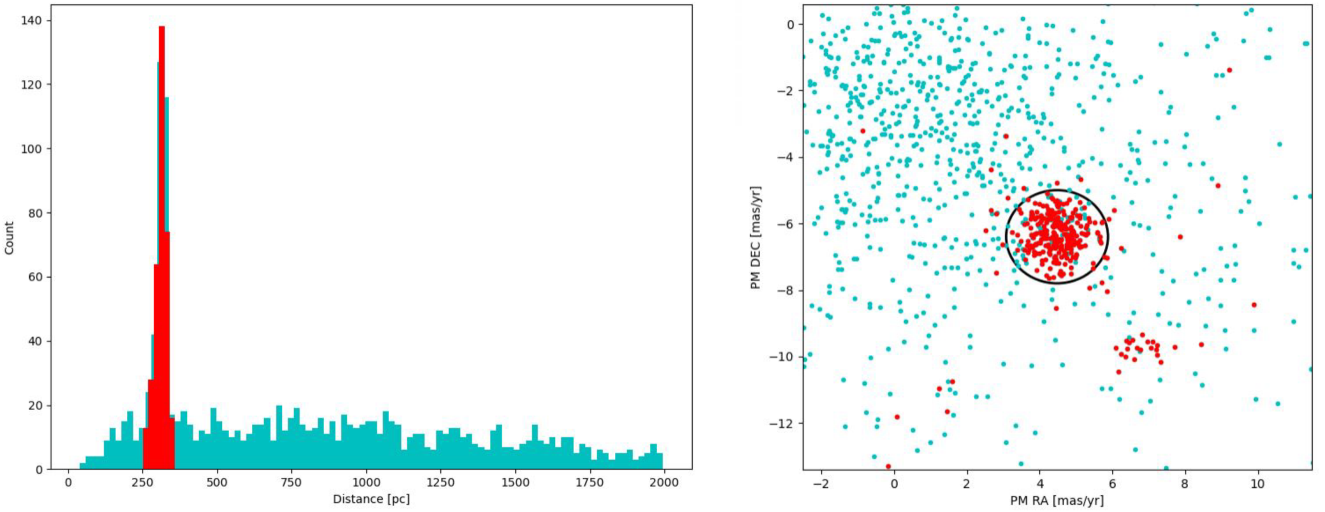

Figure 1. Example of the cluster member selection in *Gaia* DR3 data for the main cluster in the IC 348 field. The left panel shows the distance histogram with the stars selected based on their parallax highlighted in red. In the right panel, the proper motion values for all stars in the field are shown, with the sources selected in the left panel highlighted in red. The black circle encompasses the proper motion range used in the initial manual selection. The red-highlighted group of stars to the bottom right from the main group represents a secondary population of stars in this field – see Table 1 for details on its properties.

Evitts et al. 2020). Only HOYS data points that had a photometric uncertainty below 0.2 mag and seeing of better than 5 arcsec were used. Furthermore, only stars that had at least 100 photometry data points in each of the *V*, *R*, and *I* filters were considered for analysis. Furthermore, we find that light curves of stars that are within 5 arcmin from very bright sources (G_{mag} <6 mag) can have unreliable photometry. These light curves have hence been removed from the analysis. All data were calibrated following the procedure described in Evitts et al. (2020). This leaves just over 3000 *V*-, *R*-, and *I*-band light curves of potential cluster members for investigation.

In Table 1, we list for each cluster how many HOYS light curves are available for analysis. The *Gaia* colour–magnitude diagrams (CMDs) with overlaid isochrones (Husser et al. 2013) were visually inspected for each cluster. Note that we do not apply an extinction correction for any of the CMDs. In Fig. 2, we show an example CMD for the main cluster in the IC 348 field. Note that the extinction vector in these plots is almost parallel to the direction of the young (few Myr) isochrones near BP − RP = 2 mag. We selected all clusters that clearly showed a population of young stars, that is, the members aligned with a young isochrone. Furthermore, all clusters with a median distance of the potential *Gaia* members of less than 1 kpc were selected to ensure the samples of stars in all clusters are comparable in mass. All stars in these clusters were analysed in our general YSO sample. Finally, we selected all clusters with at least 20 HOYS light curves for a more detailed analysis on a cluster-by-cluster basis. All these selections are summarized in the final column in Table 1. There are spatial overlaps between some of the HOYS target fields. For the analysis on a cluster-by-cluster basis, we have hence merged these. In particular, the cluster M 42 contains all stars from the HOYS fields M 42, L 1641 N, V 898 Ori, YY Ori, and V 555 Ori; and IC 1396 contains all objects from the IC 1396 A and IC 1396 N fields.

2.3 HOYS light curve analysis

In this section, we briefly describe the parameters determined for each HOYS light curve for every selected potential cluster member.

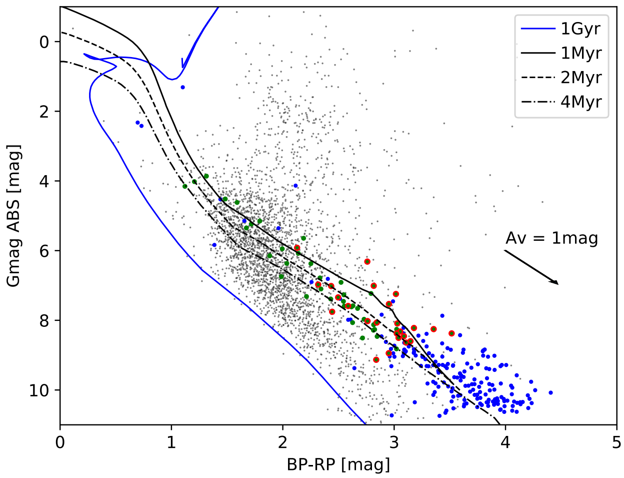

Figure 2. Example of a *Gaia* CMD for the main cluster in the IC 348 field. We converted the apparent G_{mag} values into absolute magnitudes using each star's parallax value. Note that no extinction correction has been applied. All *Gaia* DR3 sources in the field are shown as small grey dots, potential cluster members as large blue dots, all stars with HOYS light curves are shown as large green dots and all stars identified as variable are shown as large red dots. The overplotted isochrones are based on the PHOENIX models (Husser et al. 2013). The arrow represents an extinction of 1 mag in A_V and is based on the average extinction coefficients for the *Gaia* filters from Casagrande & VandenBerg (2018).

A visual inspection of the data showed that despite the strict quality selection applied to the photometry (see above), a few light curves contained a small number of erroneous, outlying data points. We define outliers as single brightness measurements that are either much brighter or much fainter than any of the other magnitudes of the star. In the absence of independent data taken in the same nights, we have to assume they are caused by photometric errors. Most of these occur in objects that are situated on a spatially highly variable background emission (as in e.g. M 42). We remove all such points in all light curves if they were more than 3σ away from the median magnitude in the light curve. Note that these could be correct measurements of short-lived flares or very narrow dips. Furthermore, there is a small subset of light curves which do show highly unusual variability patterns which cannot be attributed to any specific issues in the photometry. These objects do remain in the sample, and might hence cause small, but statistically insignificant changes in the results due to potential misclassification of their properties.

For each light curve and filter we determined a general Lomb–Scargle periodogram (Scargle 1982; Zechmeister & Kürster 2009) using the implementation in ASTROPY. The periodograms use test frequencies between one over 1.3 d and one over half the light curve length. They are distributed homogeneously in frequency space with a frequency spacing of one divided by 20 times the light curve length. Despite the many participants and their wide-spread geographic distribution (Froebrich et al. 2018), the observing cadence of the ground-based observatories involved in the HOYS project is usually close to 1 d during periods of continuous clear weather. Therefore periodograms are often dominated by 1 d periods and their harmonics/aliases. We therefore compute the window function of the light curves together with the periodograms (following equation A10 in Froebrich et al. 2021). We remove all periods with peaks of the window function which are higher than three standard deviations of the window function and their aliases from consideration in the periodograms. After this step, the strongest period is logged for each object and photometric filter. The corresponding false-alarm probability (FAP) is computed using the bootstrapping implementation in ASTROPY. This is the same procedure as described in detail in Froebrich et al. (2021). In a follow-up paper (Herbert et al., in preparation), we will analyse all periodic light curves in the entire sample of cluster members, including the identification of periods with a number of different period finding and evaluation methods, as detailed in Froebrich et al. (2021).

In order to classify the light curves in the HOYS sample according to their general properties, we compute the variability metrics Q and M following Cody et al. (2014), which quantify the quasi-periodicity and flux asymmetry, respectively.

$$Q = \frac{\sigma_{\rm red}^2 - \sigma_{\rm phot}^2}{\sigma_m^2 - \sigma_{\rm phot}^2}. \qquad (1)$$

The quasi-periodicity index Q is defined in equation (1). There, $\sigma_{\rm red}^2$ is the variance of the data after a smoothed periodic component has been subtracted, $\sigma_{\rm phot}$ is the measurement uncertainty (set to the average photometry uncertainty of all light curve data points), and σ_m^2 is the variance of the original light curve. The smoothed periodic component is determined using a boxcar filter with a width equal to one-fourth of the period in a phase-folded light curve. The Q values close to zero indicate light curves which are close to strictly periodic, whereas Q values close to 1 are found for signals that are completely stochastic. Values of the Q index in between those extremes correspond to various levels of quasi-periodic behaviour.

$$M = \frac{\langle m_{10\ \rm per\ cent} \rangle - m_{\rm med}}{\sigma_m} \qquad (2)$$

The flux asymmetry index M is defined in equation (2), where $\langle m_{10\ \rm per\ cent} \rangle$ denotes the mean of the 10 per cent brightest and faintest data points in a given light curve, $m_{\rm med}$ is the median magnitude, and σ_m is the standard deviation of the light curve. Stars with negative values of M are predominantly objects undergoing brightening, while stars with positive values of M are predominantly stars going through dimming. Light curves with values of M close to zero are variables with symmetric light curves.

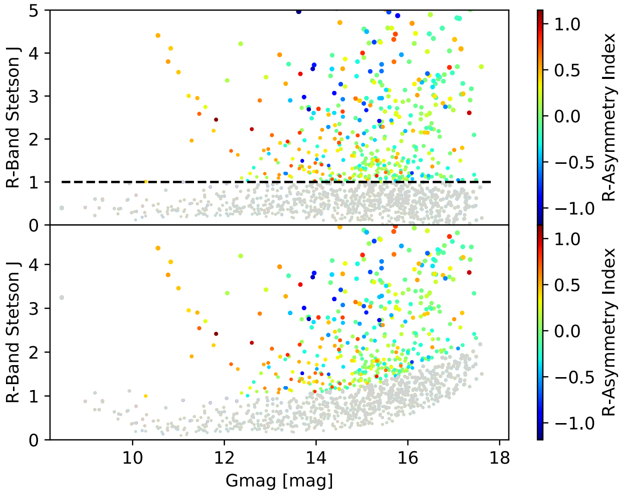

Figure 3. Stetson J index distribution for the light curves of the investigated clusters, using the R-band data. In the bottom panel, the original values are displayed, which show a magnitude dependence. In the top panel, the values after our applied correction (see the text for details) are shown and the dashed line marks Stetson $J = 1$, above which we consider stars variable. The colour code for all points represents the asymmetry index M, except for the non-variable objects which are shown in grey.

For each light curve, we also compute the Stetson J variability index (Stetson 1996) as a main indicator of the overall level of variability of a light curve. This index basically determines how much more variable compared to the photometric uncertainty a star is. As has been found in Evitts et al. (2020), the HOYS calibration procedure slightly underestimates the photometric uncertainties for fainter stars. Thus, the Stetson J index for our stars shows a slight magnitude dependence. We have thus fit a magnitude dependent running median to the Stetson J index values and subtracted this function from the values. We then added the median of all Stetson J indices in the magnitude range 12–15 mag to all values to ensure the correct values for this range (which does not show any magnitude dependence) are restored. In Fig. 3, we show the Stetson J values for the R-band data before and after the above-described correction. Stars with values of this corrected Stetson J index above one are considered variables in our sample.

2.4 Auxiliary data and definitions

In order to investigate the evolutionary stages of the selected cluster members, we have utilized the cross-matches to 2-micron All Sky Survey (2MASS) (Skrutskie et al. 2006) and the *Wide-field Infrared Survey Explorer (WISE)* All-sky catalogue (Cutri et al. 2013) provided in *Gaia* DR3, and downloaded the J, H, and K as well as the $W1$, $W2$, $W3$, and $W4$ magnitudes when available for all sources. Only data with uncertainties of less than 0.1 mag in *WISE* were included in the analyses.

In this work, we analyse the light curve properties of cluster members with and without a $K - W2$ excess. We consider sources with excess infrared emission, or detectable inner disc, as objects that have a $K - W2$ colour in excess of 0.5 mag (Teixeira et al. 2012). Objects with a bluer $K - W2$ colour will be referred to as non-excess sources or photospheres.

Furthermore, we consider stars as variable if they have a Stetson J index (corrected as detailed above) larger than one. All other objects are considered as non-variable. Light curves of variable stars with an M index above 0.5 are considered as dippers. Objects with an M index below −0.5 are classified as bursters. All other sources with an M index between −0.5 and 0.5 are considered symmetric. Non-variable sources cannot be classified in this way. Note that in principle flare stars could be classified as bursters, if our data catch the objects during a number of such flare events. Furthermore, we use the above values to distinguish dippers, bursters, and symmetric light curves for consistency in following Cody et al. (2014). However, the distribution of M values (as discussed in Section 3.1.1) shows that there is a continuum of values with no clear 'underpopulation' at ±0.5. Thus, this is an arbitrary cut, and other works, such as the simulations by Robinson, Espaillat & Owen (2021) use different values (±0.25) for the cut and slightly different definitions for the M index. Hence, absolute values for the fraction of dipper and burster light curves should be interpreted with caution.

Finally, variable objects with a Q index between zero and 0.45 are classified as periodic objects. If the Q index is between 0.45 and 0.85 the source is considered semi-periodic. Irregular or stochastic light curves have a Q value between 0.85 and one. As for the M index, non-variable sources cannot be classified based on the Q index. The periods determined as part of the evaluation for the Q index are considered significant if the associated FAP is below 1×10^{-6}. Following Froebrich et al. (2021), we use a period of 5.5 d to distinguish fast and slow rotators. The final note in the last paragraph, about the specific values distinguishing classes of sources also applies to the Q-values.

3 RESULTS

Our selection of potential clusters and groupings in the 25 HOYS target fields has identified 45 different clusters and/or groups of comoving stars. In most fields, these are dominated by a main cluster – the original target of the HOYS project. The median distances and proper motions for each of these groups are listed in Table 1. For the analysis in this paper, we only select clusters/groups which clearly appear young in the *Gaia* CMD (see Fig. 2 for an example) and have a distance of less than 1 kpc. There are 21 such clusters and groups. All of these are listed in the top part of Table 1. All other groups are listed in the bottom part of the table for completeness. We will not analyse their potential members in this work.

In total, the 21 clusters/groups contain 9143 unique *Gaia* DR3 sources as potential members. The HOYS data base contains light curves with at least 100 photometry data points in each of the V-, R-, and I-band filters for 1687 (18 per cent) of these. This is what we consider the *full sample* of young stars. We are analysing the overall properties of this full sample and their light curves in Section 3.1. In Section 3.2, we investigate the properties of the stars in the clusters that have at least 20 members. These clusters/groups are indicated with 'YYY' in the last column of Table 1.

3.1 General light curve results

3.1.1 Light curve variability and symmetry

We show the distribution of Stetson J indices for the full sample of stars in the top panel of Fig. 3 for the R-band data. The plots in the other filters are quantitatively similar. As one can see, the majority of sources is not variable. However, a fraction of 40/38/29 per cent of the stars are considered variable with a Stetson J index above one in V, R, and I, respectively. We consider these sources the *variable sample* of young stars in each filter. The drop in the I-band variability fraction is due to the fact that generally variability amplitudes for the majority of physical mechanisms (surface spots, line-of-sight extinction changes, and mass accretion rate variations) are lower at longer wavelengths. Note that 25 per cent of all stars are considered variable in all three filters simultaneously, and 44 per cent are variable in at least one of the filters. In Fig. 3, we only show values for Stetson J up to five, and a fraction of 3.7/5.0/3.0 per cent of the stars is more variable than this, up to a maximum Stetson J index of 25/34/56 in V, R, and I, respectively. The symbol colour in the plot indicates the asymmetry index M for each star. Note that this is only meaningful for the variable sample, that is, the sources with a Stetson J index above one (see below). In Fig. B1 in Appendix B, we show a few examples of symmetric, dipper, and burster light curves.

The number of objects in our full and variable sample per magnitude bin starts dropping for $G_{\rm mag} = 16$ mag and fainter. Thus, the samples become incomplete for sources fainter than that limit. This is caused by the variations in limiting magnitude and field of view in our data. Similarly, objects brighter than $G_{\rm mag} = 10$ mag are at or near saturation in some of the data. Independent of this, or the filter used, the dominant type of light curve in the variable sample is symmetric ($-0.5 < M < +0.5$). Typically about 50 per cent–80 per cent of the variable stars are categorized as symmetric (depending on $G_{\rm mag}$ and filter). Within the $G_{\rm mag}$ = 10–14 mag range (where we are certainly complete), the fraction of symmetric light curves is 0.67/0.62/0.56 for V, R, and I, respectively. The remaining objects are bursters and dippers. The ratio of bursters to dippers for the variable sources in the $G_{\rm mag}$ = 10–14 mag range is 0.26/0.62/0.00, for V, R, and I, respectively. Thus, in all filters, the brighter sources are dominated by dipper light curves, which outnumber the burster light curves by a factor of 2–4 or more. At fainter magnitudes (above about 14/15/15 in V, R, and I) the burster light curves start to dominate over dippers.

All of these statistics are caused by the typical bias of a magnitude limited sample, like ours. Most of the young stars in our sample are red, that is, they are brighter in the I band than in the V filter. Furthermore, there are many more fainter (less-massive) stars than bright ones. Thus, many dipper stars will not be detected in their faint state in the V band, and they hence get misclassified as symmetric light curves. This is less of an issue in the R and I bands, since the objects are brighter in these. Similarly, some outbursting stars will be saturated in the I-band data but not in the V-filter, which can lead to an erroneous classification as symmetric light curve instead of burster in that filter. This issue is further influenced by the above discussed bias that changes in brightness due to accretion rate changes and/or line-of-sight extinction variations are smaller in the I band. The R-band data hence suffer the least amount of bias. If we consider this, and only the magnitude range for which we most likely have a complete sample, then we find that dipper light curves amongst our variable sample outnumber the bursters by a factor of 2−4. However, as indicated above, the symmetric light curves dominate the variable sample with 55 per cent–65 per cent being classified as symmetric.

In Fig. 4, we show the quasi-periodicity index Q versus the asymmetry index M of the variable sample based on the R data. The colour coding indicates the M index and the symbol size the Stetson J variability index, in line with the other figures. As discussed above, the most common variability is symmetric, around $M = 0$. However, it can be seen that amongst the symmetric light curves, there are more objects with positive values for M than with negative ones. There is a more or less homogeneous distribution of Q values in the range from 0.0 to 0.6. A much larger fraction of sources has higher Q-values. Typically, sources with larger Stetson J indices can have more asymmetric light curves. About 1.5 per cent–3.0 per cent of the variables have Q-values outside the nominal range

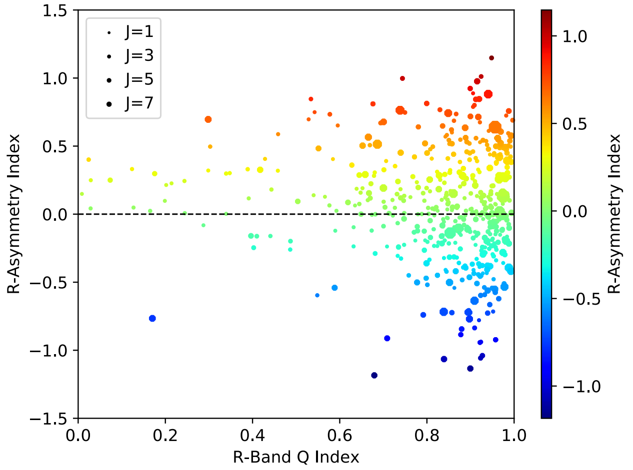

Figure 4. *R*-band *Q* index versus *M* index for all variable sources. The colour code also represents the *R*-band asymmetry index *M* and the symbol size is proportional to the *R*-band Stetson *J* index for each source. The legend gives an indication of the respective symbol sizes.

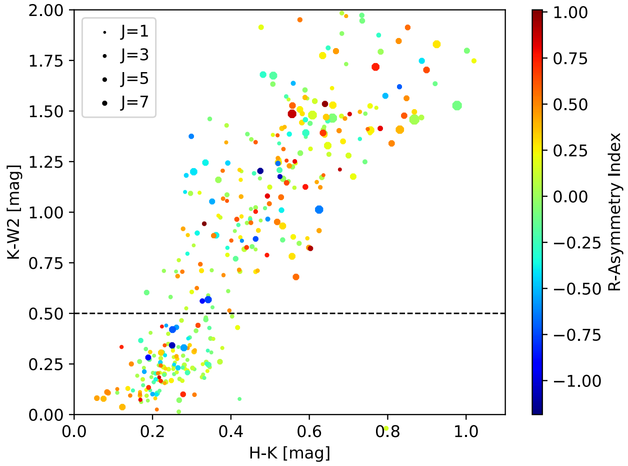

Figure 5. Near- and mid-infrared colour–colour plot using 2MASS and *WISE* data for all investigated variable cluster members. The dashed line marks the separation between objects with a detectable inner disc (above) and without (below). The colour code represents the *R*-band asymmetry index *M* and the symbol size is proportional to the *R*-band Stetson *J* index for each source. The legend gives an indication of the respective symbol sizes.

between 0 and 1, and are hence not shown. The reason for these is the underestimation of the photometric uncertainties towards fainter magnitudes, as discussed in Section 2.3.

3.1.2 *Near- and mid-infrared colours*

In Fig. 5, we show the distribution of near- and mid-infrared colours of the full sample of variable stars investigated, which have a detection at those wavelengths with a minimum SNR of 10. The colour code of the symbols indicates the asymmetry index *M* and the symbol size represents the Stetson *J* variability index. There is a clear separation of the objects into two groups. One group shows a clear $K-W2$ excess emission, that is, represents sources with detectable inner discs. The other group has near zero mid-infrared colours, hence represents sources with no detectable inner disc, that is, photospheres. The dashed line in Fig. 5 indicates the dividing line between these two groups at $K-W2 = 0.5$ mag (Teixeira et al. 2012). As discussed in the last section, both groups of sources are dominated by symmetric light curves. They also contain a mix of dippers and bursters, but there is an over abundance of dippers compared to bursters in general.

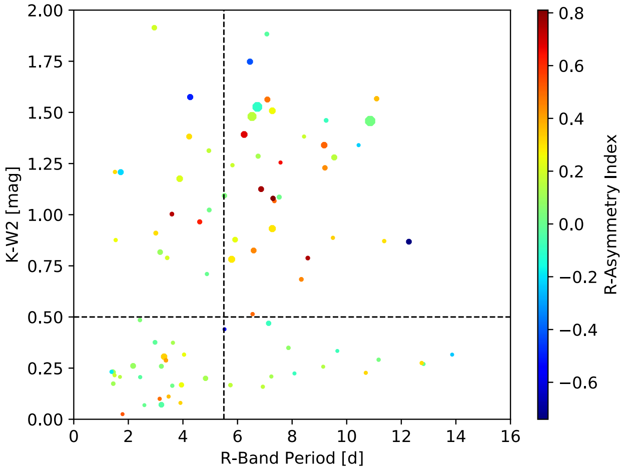

Figure 6. Significant periods in variable sources versus $K-W2$ colour. Sources above the horizontal dashed line are considered to have an excess in the $K-W2$ colour. The vertical dashed line separates fast from slow rotators. The symbol colour represents the *M* index and the symbol size the Stetson *J* variability index.

On average our full sample contains 32 per cent of objects with a $K-W2$ excess and 68 per cent without. However, if only the variable sample of stars is considered (the ones shown in Fig. 5), the fraction of sources with detectable inner discs increases to 60 per cent–65 per cent, depending on the filter for which the Stetson *J* index is determined. If we consider only objects that are variable in all three filters, then the disc fraction is 71 per cent. Amongst the variables, the symmetric light curves dominate the population as discussed above, with 70 per cent–80 per cent. However, the dippers outnumber the bursters in the $K-W2$ excess variables by 3:1 (in *V* and *R*) and 10:1 in *I*. For the variables without $K-W2$ excess emission, the bursters outnumber the dippers by 3:2 in *V*, and the populations are equal the *R* and *I* bands. Some of the burster light curves could be caused by flares. We note that the HOYS survey is not designed to detect these reliably and refer back to our comments on the removal of photometry outliers and unusual light curves in Section 2.3.

3.1.3 *Light curve periodicity*

In Fig. 6, we show all variable sources for which a significant period has been identified. This period (determined from the *R* data) is plotted against the $K-W2$ colour. The colour coding and symbol size is the same as in the other figures. Note that the periods and associated FAPs have been determined for the entire, up to 7 yr long, light curves. Thus, spotted stars which show detectable periodic variations only during a part of their light curve, due to variable spot properties, are likely not included. Only sources with periodic variability, which is stable during the entire light curve are identified. We show the data (light curve and periodogram) for one such source in Fig. C1 in Appendix C.

We find that amongst the periodic variables the fraction of $K-W2$ excess sources is 59/56/72 per cent (for *V*, *R*, and *I*, respectively). This is roughly the same fraction that we found for all variable sources.

Thus, the identification of periodic variability is not dependent on the presence of a detectable inner disc. Of all the variable stars, between 15 per cent and 20 per cent of the stars do have detectable periodic variability. Considering the above, these sources are hence most likely stars with stable, large, cold spots on their surface, or AA-Tau like objects with warped inner discs. The properties of the entire sample of periodic variables and their spots will be investigated in detail in Herbert et al. (in preparation). This will include a search for periodic signals in shorter (6 months) parts of each light curve.

There is a difference in how the fast and slow rotators are distributed amongst the variables with and without $K - W2$ excess. For the inner disc sources, between 65 per cent and 70 per cent (depending on the filter) are slow rotators, while 30 per cent–35 per cent are fast rotators. Hence, the slow rotators outnumber the fast rotators by about 2:1 if the sources have a detectable inner disc. The situation is reversed for the objects without $K - W2$ excess. Here, between 57 per cent and 81 per cent are fast rotators, while between 19 per cent and 43 per cent have longer rotation periods. Thus, the fast rotators outnumber the slow rotators by 2:1. Compared to the investigation of IC 5070 by Froebrich et al. (2021), these distributions match within the statistical uncertainties, but are based on a much larger sample. We note however, that the period identification for our current paper has been done using only one methodology and uses the entire 7 yr long light curves. The Herbert et al. (in preparation) paper will include a full discussion of this aspect. This change of the ratio of fast to slow rotators for objects with and without inner discs is consistent with the idea of disc breaking (see e.g. the review by Herbst et al. 2007).

The periodic variables have preferentially (86 per cent, same in all filters) symmetric light curves. This is expected if most periodic variations are due to spotted objects. Amongst the remaining sources, the dippers outnumber the bursters by about 3:1. Thus, only a fraction of 2 per cent–6 per cent of objects are periodic bursters. The much larger fraction of periodic dippers (8 per cent–13 per cent) is most likely dominated by AA-Tau type sources. This is supported by the fact that the periodic dippers are almost exclusively found amongst the $K - W2$ excess sources.

3.2 Cluster population analysis

Here, we discuss the light curve analysis results by cluster. Only the 17 clusters/groups with more than 20 stars with HOYS light curves are considered. Furthermore, we merged the stars in clusters/groups that have a spatial and proper motion overlap. This leaves a total of 12 clusters/groups of stars for our analysis. These are: IC 5070 a and b, NGC 1333, σ-Ori, λ-Ori, NGC 7129, M 42 (we merged M 42, L 1641 N, V 898 Ori, YY Ori, and V 555 Ori), IC 348, IC 5146, NGC 2264 a and b, and IC 1396 (we merged IC 1396 A and IC 1396 N). Unless indicated otherwise in the name, only the stars from the main, most populous group in each target region are included in this analysis.

We aim to compare the light curve properties of the stars in each cluster to investigate how they potentially depend on the cluster properties. For this purpose we estimate the disc fraction of stars in each cluster. This has been determined as the fraction of all potential Gaia detected cluster members which have $K - W2$ excess emission due to circumstellar dust in the inner disc. We consider only stars with a detection in the 2MASS *K* band and the *WISE W*2 filter, with uncertainties below 0.1 mag. This disc fraction can be considered an indicator for the typical age of the cluster. As the $K - W2$ excess emission only traces the inner discs, our disc fractions have to be considered a lower limit. All our clusters are closer than 1 kpc. Thus, we approximately detect stars down to the same mass in all clusters. Thus, any systematic biases caused by a potential mass dependence of the disc fraction, or by incompleteness for low-mass stars, are very similar in all clusters.

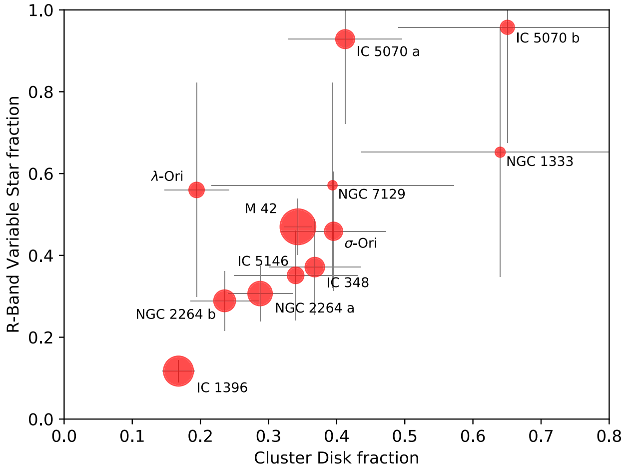

Figure 7. Disc fraction versus fraction of variable stars determined from the HOYS *R*-band light curves. The thin grey lines represent the uncertainties of the plotted ratios. The clusters are labelled as described in Section 3.2.

3.2.1 Cluster variability fraction

We determine the fraction of stars that are variable in each cluster. This was done for each of the three filters (*V*, *R*, and *I*) separately. We show how this variability fraction compares to the disc fraction in each cluster in Fig. 7 for the *R* band. The plots in the other filters do qualitatively look similar and some quantitative differences are discussed below. We have labelled the clusters with their common names and the circle size of the symbols is proportional to the log of the number of stars in the cluster used in our analysis. We over plot uncertainties for the determined fractions, which are based on the estimation that the uncertainty in the number (N) of sources is equal to $\sqrt{N}$.

In general (with some exceptions discussed below), the variability fraction correlates with the fraction of detectable inner discs in the cluster. Clusters with a higher inner disc fraction do show on average a larger fraction of variable stars. This indicates that younger clusters contain a larger fraction of variable stars. The correlation is stronger in the *V* and *R* filters, compared to the *I* band. Furthermore, on average the variability fraction decreases with increasing wavelength of the observation. The latter is most likely caused by the fact that most causes for the photometric variability are stronger at shorter wavelengths. Both, accretion rate and line-of-sight extinction changes cause larger magnitude changes in *V* and *R*, compared to the *I* band. Thus, stars are more likely to be detected as variable above the photometric uncertainty in those filters. This is in part counter-balanced by the fact that the young stars are fainter in the *V* and *R* filters, which in total leads to only a small dependence of the variability fraction on the filter used, especially for the most populous clusters. This correlation of disc and variability fraction is in line with the finding from Section 3.1.2 that the fraction of disc excess sources doubles amongst the variable stars compared to all cluster members in our sample.

There are some exceptions to the general trends seen in Fig. 7. In NGC 1333, the variability fraction is slightly below the general trend. This could in part be a statistical effect since only 23 cluster members have a HOYS light curve with sufficient data points to analyse, and the uncertainties do support this. Both, the λ-Ori cluster and IC 5070 a are outliers above the general trend. They have a far higher variability fraction in V and R, with respect to the overall trend. Again, this could be a statistical effect, especially for λ-Ori as there are only 25 cluster members with a HOYS light curve in it, but not all of the deviation from the trend can be fully explained that way. One other possibility is that the correlation between variability fraction and disc fraction breaks down for very young samples with a disc fraction above 40 per cent (which includes two of the three outliers just mentioned).

3.2.2 *Other light curve properties*

We have investigated a number of other light curve properties for each sample of stars in the 12 clusters and groups. In particular, these are the fractions of burster or dipper light curves amongst the variables sample, as well as the fraction of periodic or stochastic light curves. Typically there are some systematic trends of all these fractions depending on the filter used in obtaining the light curve. All of these are in line with the systematic biases in our data discussed above and the results obtained for the overall sample of stars analysed, including the occasional outliers from any trends.

For all the investigated light curve properties, there are however no correlations of any of the properties with the cluster disc fraction or any other cluster property. In other words, the dipper or burster fraction varies from cluster to cluster by a factor of a few, around the same average as found for the entire sample of variables, the values obtained for the same cluster in different filters correlate, but there are no correlations of those with any of the cluster properties. The same applies to the fraction of periodic or stochastic light curves. This might in part be due to the small number statistics in some clusters, but generally there seems to be indeed no significant trends even considering just the more populated clusters.

4 DISCUSSION AND CONCLUSIONS

The HOYS project is providing long-term, optical, multifilter, high-cadence monitoring of 25 target fields centred on presumed, nearby young star clusters and star-forming regions. In this work, we have utilized the *Gaia* DR3 astrometry to identify potential cluster members and to analyse their optical light curves. A total of 45 groups/clusters with ≈17 000 *Gaia* selected members (≈3000 with well-sampled HOYS light curves) have been identified. We selected 21 of these groups for detailed analysis, as they are within 1 kpc and clearly represent a population of young stars in the *Gaia* CMD. Thus, our detailed analysis includes 9143 potential cluster members, with well-sampled optical light curves in V, R, and I for 1687 (18 per cent) of them.

In many of the HOYS target fields, multiple populations of young stars were identified, and their distances and proper motions are summarized in Table 1. For many of these this is the first time they are determined from *Gaia* DR3 data. However, for several of the clusters our values are in agreement with previous works using earlier versions of *Gaia* data, such as DR2 and EDR3. These are for example: NGC 1333 and IC 348 (Pavlidou et al. 2021); M 42 and other regions in the Orion complex (Chen et al. 2020); σ-Ori (Froebrich et al. 2022); IC 5070 (Kuhn et al. 2020); and IC 1396 (Pelayo-Baldárrago et al. 2023). Several of our regions are also included in Mendigutía et al. (2022).

Our work represents a large, homogeneous sample of nearby (within 1 kpc) young stars with long-term (≈7 yr) light curves in at least three optical filters (V, R, and I). Similarly to the astrometric properties, in many cases this represents the first time such data have been made available and/or analysed. Previous works presenting the analysis of optical light curves (but mostly either using single filter data, or shorter survey duration) for a large number of members of some of our clusters include for example: NGC 1333 and IC 348 (e.g. Wang et al. 2023); IC 348 (e.g. Herbst, Maley & Williams 2000; Cohen, Herbst & Williams 2004; Kızıloğlu, Kızıloğlu & Baykal 2005; Nordhagen et al. 2006; Fritzewski et al. 2016); NGC 2264 (e.g. Cody et al. 2014; Stauffer et al. 2016; Lakeland & Naylor 2022); and σ-Ori (e.g. Scholz 2004; Scholz & Eislöffel 2004; Cody & Hillenbrand 2010). Note that for many of the regions time-series analysis has also been conducted in the infrared, such as for NGC 1333 (Rebull et al. 2015), IC 1396 (Meng et al. 2019), and many others.

We identify light curves as variable if their Stetson J index is above one. Depending on the filter used, the variability fraction varies. About one in four objects is considered variable in all three filters (V, R, and I), while almost half the objects are variable in at least one of the filters. The variability fraction in the sample increases with decreasing wavelength of the filter used. The asymmetry index M of the light curves shows that roughly two-thirds of all variable light curves can be classified as symmetric. The remainder are made up of dipper and burster light curves. For a subsample of bright sources (which is complete), the dippers outnumber the bursters by a factor of 2–4. Furthermore, the majority of light curves are classified as stochastic, based on the quasi-periodicity index Q. Strictly periodic light curves are rare and quasi-periodic light curves are uncommon.

The potential cluster members with reliable K and W2 photometry show a clear split into objects with and without detectable inner discs. They can be separated at $K - W2 = 0.5$ mag. Our full sample contains 32 per cent with disc excess emission and 68 per cent without. Amongst the variable stars, however, the fraction of disc excess sources increases to 60 per cent–65 per cent. In the disc excess sample, the dippers outnumber the bursters by just over three to one, while in the non-disc excess sample the bursters and dippers occur in roughly the same amount. Note that the lack of disc excess emission only means that they are dust depleted inner discs. They can still be gas rich and have colder dust further out (see e.g. Manara et al. 2014).

About one in six variables shows a detectable periodic variability. Sixty per cent of these are disc excess objects. Thus, the detection rate of periodic variability is independent of the presence of a disc. The period distributions of objects with and without discs clearly support the idea of disc breaking (Herbst et al. 2007). About two in three objects with detectable inner disc are slow rotators, while two thirds of the objects without disc are fast rotators.

As we found that the fraction of variable stars increases amongst the disc excess population, we investigated this on a cluster-by-cluster basis. We find that there is a clear correlation of the disc fraction in a cluster and the variability fraction. There are, however, a few outliers to this trend, which are explainable by small number statistics.

ACKNOWLEDGEMENTS

We would like to thank all contributors of observational data for their efforts towards the success of the HOYS project. This research has been partially supported by the supercomputing infrastructure of the NLHPC (ECM-02) in Chile. SV acknowledges the help of the IT support team of the NLHPC while working on the calculations described in this paper. CH is supported by the Science and Technology Facilities Council under grant number STFC Kent 2020 DTP ST/V50676X/1. JCW is funded by the European Union (ERC, WANDA, 101039452). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them. 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.

DATA AVAILABILITY STATEMENT

The data underlying this article are available in the HOYS data base at http://astro.kent.ac.uk/HOYS-CAPS/.

REFERENCES

Audard M. et al., 2014, in Beuther H., Klessen R. S., Dullemond C. P., Henning Th., eds, Protostars and Planets VI. University of Arizona Press, Tucson, AZ, p. 387
Bouvier J., Alencar S. H. P., Harries T. J., Johns-Krull C. M., Romanova M. M., 2007, in Reipurth B., Jewitt D., Keil K., eds, Protostars and Planets V. University of Arizona Press, Tucson, AZ, p. 479
Bouvier J., Matt S. P., Mohanty S., Scholz A., Stassun K. G., Zanni C., 2014, in Beuther H., Klessen R. S., Dullemond C. P., Henning Th., eds, Protostars and Planets VI. University of Arizona Press, Tucson, AZ, p. 433
Casagrande L., VandenBerg D. A., 2018, MNRAS, 479, L102
Chen B., D'Onghia E., Alves J., Adamo A., 2020, A&A, 643, A114
Cody A. M., Hillenbrand L. A., 2010, ApJS, 191, 389
Cody A. M., et al., 2014, AJ, 147, 82
Cohen R. E., Herbst W., Williams E. C., 2004, AJ, 127, 1602
Cutri R. M. et al., 2013, VizieR Online Data Catalog, 2328
Evitts J. J. et al., 2020, MNRAS, 493, 184
Fischer W. J., Padgett D. L., Stapelfeldt K. L., Sewiło M., 2016, ApJ, 827, 96
Fischer W. J., Hillenbrand L. A., Herczeg G. J., Johnstone D., Kospal A., Dunham M. M., 2023, in Inutsuka S., Aikawa Y., Muto T., Tomida K., Tamura M., eds, ASP Conf. Ser., Vol. 534, Protostars and Planets VII. Astron. Soc. Pac., San Francisco, p. 355
Fritzewski D. J. et al., 2016, MNRAS, 462, 2396
Froebrich D. et al., 2018, MNRAS, 478, 5091
Froebrich D. et al., 2021, MNRAS, 506, 5989
Froebrich D., Eislöffel J., Stecklum B., Herbert C., Hambsch F.-J., 2022, MNRAS, 510, 2883
Froebrich D. et al., 2023, MNRAS, 520, 5413
Gaia Collaboration, 2016, A&A, 595, A1
Gaia Collaboration, 2023, A&A, 674, A1
Grankin K. N., Melnikov S. Y., Bouvier J., Herbst W., Shevchenko V. S., 2007, A&A, 461, 183
Herbert C., Froebrich D., Scholz A., 2023, MNRAS, 520, 5433
Herbst W., Maley J. A., Williams E. C., 2000, AJ, 120, 349
Herbst W., Eislöffel J., Mundt R., Scholz A., 2007, in Reipurth B., Jewitt D., Keil K., eds, Protostars and Planets V. University of Arizona Press, Tucson, AZ, p. 297
Hillenbrand L. A. et al., 2018, ApJ, 869, 146
Husser T. O., Wende-von Berg S., Dreizler S., Homeier D., Reiners A., Barman T., Hauschildt P. H., 2013, A&A, 553, A6
Joy A. H., 1945, ApJ, 102, 168
Kharchenko N. V., Piskunov A. E., Schilbach E., Röser S., Scholz R. D., 2013, A&A, 558, A53
Kızıloğlu Ü., Kızıloğlu N., Baykal A., 2005, AJ, 130, 2766
Kochanek C. S. et al., 2017, PASP, 129, 104502
Kounkel M., Deng T., Stassun K. G., 2022, AJ, 164, 57
Kuhn M. A., Hillenbrand L. A., Carpenter J. M., Avelar Menendez A. R., 2020, ApJ, 899, 128
Lakeland B. S., Naylor T., 2022, MNRAS, 514, 2736
Luhman K. L., 2023, AJ, 165, 269
Manara C. F., Testi L., Natta A., Rosotti G., Benisty M., Ercolano B., Ricci L., 2014, A&A, 568, A18
Mendigutía I., Solano E., Vioque M., Balaguer-Nuñez L., Ribas A., Huélamo N., Rodrigo C., 2022, A&A, 664, A66
Meng H. Y. A., Rieke G. H., Kim J. S., Sicilia-Aguilar A., Cross N. J. G., Esplin T., Rebull L. M., Hodapp K. W., 2019, ApJ, 878, 7
Nordhagen S., Herbst W., Rhode K. L., Williams E. C., 2006, AJ, 132, 1555
Pavlidou T., Scholz A., Teixeira P. S., 2021, MNRAS, 503, 3232
Pelayo-Baldárrago M. E., Sicilia-Aguilar A., Fang M., Roccatagliata V., Kim J. S., García-Álvarez D., 2023, A&A, 669, A22
Rebull L. M. et al., 2015, AJ, 150, 175
Robinson C. E., Espaillat C. C., Owen J. E., 2021, ApJ, 908, 16
Scargle J. D., 1982, ApJ, 263, 835
Scholz A., 2004, PhD thesis. Karl Schwarzschild Observatory, Tautenburg, Germany
Scholz A., Eislöffel J., 2004, A&A, 419, 249
Sewiło M. et al., 2019, ApJS, 240, 26
Skrutskie M. F., et al., 2006, AJ, 131, 1163
Stauffer J. et al., 2016, AJ, 151, 60
Stetson P. B., 1996, PASP, 108, 851
Teixeira P. S., Lada C. J., Marengo M., Lada E. A., 2012, A&A, 540, A83
Wang X.-L., Fang M., Herczeg G. J., Gao Y., Tian H.-J., Zhou X.-Y., Zhang H.-X., Chen X.-P., 2023, Res. Astron. Astrophys., 23, 075015
Zechmeister M., Kürster M., 2009, A&A, 496, 577

APPENDIX A: NOTES ON INDIVIDUAL CLUSTERS IN THE MANUAL *GAIA* SELECTION

In this section, we briefly discuss the clusters and groups of astrometrically coherent populations of stars identified in the HOYS target fields. Each section is identified by the name of the main target cluster. If uncertainties for distances or proper motions are quoted, then these are the standard errors for the mean as in Table 1 and not the RMS scatter of the cluster members from the median. All comparisons to isochrones are based on the PHOENIX models by Husser et al. (2013).

NGC 1333

There is only one main peak in the histogram of *Gaia* distances and only one coherent cluster in proper motion space. Thus, this is a simple field. This main cluster has about the same proper motion ($\mu_\alpha = +7.09 \pm 0.12$ mas yr^{-1} and $\mu_\delta = -9.92 \pm 0.07$ mas yr^{-1}) and distance ($d = 298.5 \pm 1.7$ pc) as the spatially distributed foreground group in the IC 348 field (see below). Hence, the same distributed population might exist in this field, but might not be identifiable due to this overlap and the small number of members. The scatter and uncertainty of μ_α is about 70 per cent larger than μ_δ. This might hint at a mix of populations in the cluster with slightly different proper motions. The same is the case for the foreground population in the IC 348 field (see below).

IC 348

Similar to NGC 1333, there is only one single identifiable peak in the histogram of the *Gaia* distances. However, in the proper motion distribution there are two, well-separated coherent groups. The more populated one, representing the main IC 348 cluster, has proper motions of $\mu_\alpha = +4.48 \pm 0.04$ mas yr^{-1} and

$\mu_\delta = -6.35 \pm 0.04$ mas yr^{-1}. Its distance is $d = 318.1 \pm 1.2$ pc. This places is about 20 pc further away than NGC 1333. The second coherent group is much less populated and clearly separated from the main cluster in proper motion space with $\mu_\alpha = +6.72 \pm 0.10$ mas yr^{-1} and $\mu_\delta = -9.74 \pm 0.06$ mas yr^{-1}. The distance of this group is $d = 297.0 \pm 4.0$ pc. This places these stars at almost the same distance and proper motion as the NGC 1333 cluster. Furthermore, the members of the second group are homogeneously distributed across the field and thus represent a larger scale foreground population. Indeed this group is corresponding to the population identified and named Alcaeus in Pavlidou et al. (2021), using *Gaia* DR2 data.

λ-Ori

The distance histogram of this field only shows a single peak. However, in proper motion space there are two clear coherent populations. Their total proper motions differ by about 0.9 mas yr^{-1}. Most of that difference is in right ascension. The more populated, and hence main group of stars has $\mu_\alpha = +0.78 \pm 0.02$ mas yr^{-1} and $\mu_\delta = -2.01 \pm 0.02$ mas yr^{-1}, while the second group has $\mu_\alpha = +1.65 \pm 0.02$ mas yr^{-1} and $\mu_\delta = -2.14 \pm 0.02$ mas yr^{-1}. The distances of both groups are almost identical with $d = 398.9 \pm 1.6$ pc for the main group and $d = 403.7 \pm 1.9$ pc for the secondary group. The main difference between the groups is that the main one is clustered around λ-Ori itself, while the other objects are far more distributed, mostly to the South of the main group.

M 42

The cluster properties for M 42 (Orion Nebula) have been determined only from *Gaia* sources within 0.3 deg from the cluster centre. There is only a single peak in the distance histogram, and only one coherent group in proper motion space. While the mean proper motion ($\mu_\alpha = +1.31 \pm 0.03$ mas yr^{-1} and $\mu_\delta = +0.22 \pm 0.04$ mas yr^{-1}) can be evaluated quite accurately due to the large number of member stars, the scatter around that mean is quite large. The population of objects in proper motion space is also quite inhomogeneous with a lot of substructure, hinting at a number of overlapping subgroups in the cluster. There is also a large number of objects with proper motions far away from the mean, potentially representing stars in the process of being ejected from the cluster. The median distance of the stars is $d = 400.4 \pm 0.9$ pc. However, if one includes stars from the entire 0.6 deg field, then this distance drops by about 5 pc. This suggests that the more distributed populations just North (see V 555 Ori) and South (see L 1641 N), which overlap the M 42 field, are closer to Earth than the main cluster. We note that all the groups identified in all the HOYS fields covering the Orion star-forming complex (λ-Ori, M 42, L 1641 N, σ-Ori, NGC 2068, V 898 Ori, YY Ori, and V 555 Ori), do coincide with stellar groups identified by Chen et al. (2020) based on the *Gaia* DR2 data set.

L 1641 N

To avoid including any stars from the neighbouring YY Ori field (see below), we only selected stars within 0.3 deg from the field centre for the determination of the distance and proper motions for this field. The distance of the stars in this field is $d = 388.4 \pm 1.3$ pc, that is, about 12 pc closer than M 42 (see above). In the distribution of proper motions, we find a main compact group and two other less populated groups which are up to 1.5 mas yr^{-1} offset. However, due to the small number of members in those, they do not show a clear sequence in a *Gaia* colour–magnitude plot. We hence determine the median proper motion for all stars in those groups together ($\mu_\alpha = +1.03 \pm 0.04$ mas yr^{-1} and $\mu_\delta = +0.33 \pm 0.05$ mas yr^{-1}), but note that this region has potential substructure and should be investigated in more detail. The average proper motions in this field overlap the range found for the M 42 field, just to the north, and the YY Ori field to the south.

σ-Ori

This field has already been analysed in detail in Froebrich et al. (2022) using *Gaia* EDR3 data. Hence, our results are very similar to these. The histogram of distances shows two clear peaks, a main one containing the majority of stars and a secondary foreground population at a distance of $d = 368.5 \pm 1.7$ pc. This foreground population shows a very compact distribution (RMS scatter of only 0.17 mas yr^{-1}) in proper motion space at $\mu_\alpha = +1.80 \pm 0.02$ mas yr^{-1} and $\mu_\delta = -1.48 \pm 0.02$ mas yr^{-1}. The sequence in the CMD indicates a slightly older population of stars. These objects are also widely distributed across the field, and hence represent a foreground moving group.

The main peak in the distance histogram splits into two coherent groups in proper motion space. The main group is spatially clustered around σ-Ori at a distance of $d = 404.5 \pm 1.4$ pc. The proper motion of this group is $\mu_\alpha = +1.62 \pm 0.03$ mas yr^{-1} and $\mu_\delta = -0.49 \pm 0.03$ mas yr^{-1}. There are hints of further substructure in the proper motion distribution of this group which should be looked at in more detail. The second group of stars in the main distance histogram peak has a proper motion of $\mu_\alpha = -2.14 \pm 0.06$ mas yr^{-1} and $\mu_\delta = +1.27 \pm 0.07$ mas yr^{-1}. Thus, it is clearly different from the main group and also spatially distributed. The median distance of the stars is $d = 421.9 \pm 3.7$ pc. Thus, this might represent another group of young stars, about 20 pc in the background of the main σ-Ori cluster.

NGC 2068

The field of NGC 2068 (or M 78) shows two peaks in the distance histogram, one at ∼250 pc and other at ∼410 pc. The stars in the 250 pc peak show no coherent structure in proper motion space, hence this is not a foreground group/cluster. The stars in the 410 pc peak split into three coherent groups in proper motion space. The two more populated groups (a and b) are close to each other with proper motions of $\mu_\alpha = -0.98 \pm 0.05$ mas yr^{-1} and $\mu_\delta = -1.02 \pm 0.04$ mas yr^{-1}, and $\mu_\alpha = +0.07 \pm 0.05$ mas yr^{-1} and $\mu_\delta = -0.56 \pm 0.04$ mas yr^{-1}, respectively. Similarly, the distances are $d = 420.0 \pm 2.3$ pc for group a and $d = 417.9 \pm 2.4$ pc for group b. Spatially, group a corresponds to the northern subcluster, while stars from group b are mostly located in the southern subcluster. The third group (c) is offset in proper motion space with $\mu_\alpha = -2.88 \pm 0.05$ mas yr^{-1} and $\mu_\delta = +1.19 \pm 0.07$ mas yr^{-1}. These stars are spatially distributed and seem to align with a roughly 10 Myr isochrone in a CMD. At a median distance of $d = 379.3 \pm 6.9$ pc they seem to form an older foreground moving group.

NGC 2244

In this field (Rosette Nebula), there are two peaks in the distance histogram, one at ∼640 pc and other near 1500 pc. The latter is representing the main cluster in the field. The members have a median distance of $d = 1512.7 \pm 3.9$ pc and a proper motion of $\mu_\alpha = -1.75 \pm 0.01$ mas yr^{-1} and $\mu_\delta = +0.22 \pm 0.02$ mas yr^{-1}, and are concentrated towards the centre of the cluster. The foreground population has a median distance of $d = 658.3 \pm 4.5$ pc, with a proper motion of $\mu_\alpha = -1.71 \pm 0.03$ mas yr^{-1} and $\mu_\delta = -4.76 \pm 0.03$ mas yr^{-1}. It aligns with a roughly 20 Myr old isochrone in the CMD and is spatially distributed towards the south of the field. It thus represents an older foreground moving group.

NGC 2264

In the NGC 2264 (Christmas Tree Cluster) field, there is only one peak in the distance histogram. In proper motion space this population in the peaks splits into two coherent groups which are about 1.2 mas yr^{-1} apart (mostly μ_α). Both groups are well populated. The slightly more populated one (a) represents the main cluster around NGC 2264, while the second group (b) is spatially clustered in the south of the field near the Cone Nebula. Both groups have similar distances with $d = 732.0 \pm 2.1$ pc for group a and $d = 737.3 \pm 2.7$ pc for group b, respectively. Their median proper motions are $\mu_\alpha = -1.60 \pm 0.01$ mas yr^{-1} and $\mu_\delta = -3.64 \pm 0.01$ mas yr^{-1} and $\mu_\alpha = -2.43 \pm 0.02$ mas yr^{-1} and $\mu_\delta = -3.68 \pm 0.01$ mas yr^{-1}, respectively.

V898 Ori

This is the most Southern of our target fields in Orion. It has a single peak in the distance histogram. The proper motion distribution shows three potential subgroups in close vicinity to each other, up to 1.6 mas yr^{-1} separation. But all them individually have too few members to analyse them separately. The distance of the group of $d = 402.8 \pm 3.9$ pc places it at about the same distance as M 42. However, the proper motions of $\mu_\alpha = +0.22 \pm 0.07$ mas yr^{-1} and $\mu_\delta = -0.31 \pm 0.06$ mas yr^{-1} are different. There is no spatial clustering of members and a lack of objects in the north-western part of the field.

YY Ori

This target field is situated between the M 42 and L 1641 N fields. Thus, to avoid overlaps, the group properties are determined for stars within 0.3 deg from the field centre only. There is a single peak in the distance histogram and a single compact group is evident in the proper motion distribution. Though, there are a number of stars scattered significantly away from the main proper motion group. These are most likely members of the M 42 population, which shows a similar proper motion distribution. The main compact group has a proper motion of $\mu_\alpha = +1.29 \pm 0.01$ mas yr^{-1} and $\mu_\delta = +0.55 \pm 0.01$ mas yr^{-1}, with a median distance of $d = 386.9 \pm 0.9$ pc, which places it foreground to M 42 in this astrometrically complex field.

ASASSN-13DB

There are two peaks in the distance histogram, one at ∼360 pc and other at ∼760 pc. The stars in the more distant peak show no coherent structure in the proper motion distribution. Thus they are most likely field stars. The stars in the nearby group are distributed as two coherent proper motion groups, separated by about 1.1 mas yr^{-1}. The main group (a) has a proper motion of $\mu_\alpha = +1.31 \pm 0.03$ mas yr^{-1} and $\mu_\delta = -0.96 \pm 0.04$ mas yr^{-1}, while the secondary group (b) has a proper motion of $\mu_\alpha = +0.37 \pm 0.04$ mas yr^{-1} and $\mu_\delta = -1.24 \pm 0.06$ mas yr^{-1}. The main group (a) has a distance of $d = 387.3 \pm 3.2$ pc, while the second group (b) is closer at $d = 346.2 \pm 5.3$ pc. Both groups do not show any spatial clustering. However, stars in group a are mostly found in the North-East of the field, while stars in group b are placed more central.

V555 Ori

This field is just north of M 42. Thus, to avoid overlap, only stars within 0.3 deg from the field centre are included in the determination of the cluster properties. There is a single peak in the distance histogram. The proper motion distribution of the stars has a very similar shape to the M 42 field. It is very extended with potential substructure. The median proper motions of the potential members are $\mu_\alpha = +1.23 \pm 0.03$ mas yr^{-1} and $\mu_\delta = -0.28 \pm 0.04$ mas yr^{-1}, very similar to M 42. However, their median distance is $d = 393.4 \pm 0.9$ pc, placing them slightly closer than M 42.

If one analyses the entire 0.6 deg field around the central coordinates, there is a not well populated (∼15 members), but coherent subgroup visible in proper motion space. These stars align well with a 10 Myr isochrone in a CMD. Their proper motions are $\mu_\alpha = -1.74 \pm 0.05$ mas yr^{-1} and $\mu_\delta = +1.40 \pm 0.06$ mas yr^{-1}, and their distance is $d = 419.4 \pm 3.8$ pc. They are spatially distributed and hence seem to correspond to a slightly older, background moving group.

Gaia 17 bpi

In this field there is no clear peak in the distance histogram. We know from Hillenbrand et al. (2018) that the approximate distance of the source is 1.27 kpc. When using a distance range around that value for the selection of *Gaia* sources, a coherent group of stars is identifiable in proper motion space. Though the group aligns with a young isochrone in a CMD, there is a large amount of scatter. This is most likely caused by strong and variable extinction, given that the majority of objects are situated in the south-western part of the field, which coincides with a dark cloud. This group of stars has a proper motion of $\mu_\alpha = -1.00 \pm 0.05$ mas yr^{-1} and $\mu_\delta = -5.65 \pm 0.06$ mas yr^{-1}, and a distance of $d = 1300 \pm 6.3$ pc.

Gaia 19 fct

There is a wide peak around ∼1200 pc in the distance histogram. In proper motion space, these stars split into three coherent groups. One of them is compact (group a), while the other two are more distributed and close to each other. We thus merge them together into group b. The main group (a) is centred in the south-east of the field and has a proper motion of $\mu_\alpha = +0.25 \pm 0.01$ mas yr^{-1} and $\mu_\delta = -0.16 \pm 0.01$ mas yr^{-1}. The distance of this group is $d = 1127 \pm 4.4$ pc. The secondary group has a proper motion of $\mu_\alpha = -3.57 \pm 0.04$ mas yr^{-1} and $\mu_\delta = +1.02 \pm 0.03$ mas yr^{-1}, with a distance of $d = 1231 \pm 7.6$ pc. However, it does not show a coherent alignment with an isochrone in a CMD, possibly due to variable line-of-sight extinction in this field. The distances are in agreement with the roughly 1 kpc quoted for young stars in this field in, for example, Fischer et al. (2016) and Sewiło et al. (2019).

Gaia 19 eyy

There is no reason for any clusters to be found in this field, as it has been centred on the periodically erupting Be-star *Gaia* 19 eyy (Froebrich et al. 2023). Indeed there is no peak in the distance histogram for this field. However, a more systematic search identified three coherent populations of stars in proper motion space. The main group (a) seems to be a cluster in the north-east of the field near the coordinates RA = 127.45 deg and Dec. = −41.9 deg. The stars have a proper motion of $\mu_\alpha = -1.77 \pm 0.01$ mas yr^{-1} and $\mu_\delta = +1.11 \pm 0.01$ mas yr^{-1}, and a distance of $d = 1406 \pm 5.5$ pc. In the CMD-, the cluster members align along the main sequence. The second group (b) forms a more spatially loose grouping in the south of the field near the coordinates RA = 127.70 deg and Dec. = −41.9 deg. Their proper motion is $\mu_\alpha = -5.41 \pm 0.02$ mas yr^{-1} and $\mu_\delta = +4.96 \pm 0.02$ mas yr^{-1}, with a distance of $d = 1274 \pm 5.7$ pc. Finally, the third group (c) seems to represent a more distributed population of stars that aligns with a reddened main sequence in the CMD. It has a proper motion of $\mu_\alpha = -4.87 \pm 0.01$ mas yr^{-1} and $\mu_\delta = +4.48 \pm 0.01$ mas yr^{-1}, with a distance of $d = 2488 \pm 25$ pc.

MWSC 3274

There is a single peak in the distance histogram for this region. The stars in the peak also form a single coherent group in proper motion space with $\mu_\alpha = -1.15 \pm 0.01$ mas yr^{-1} and $\mu_\delta = -1.59 \pm 0.01$ mas yr^{-1}. The median distance of the group is $d = 884.5 \pm 3.5$ pc. In the CMD the stars follow a sequence that can be interpreted as a reddened older cluster of stars. In Kharchenko et al. (2013), the object is listed as a 4.5 Myr old cluster at a distance of 1580 pc, which is about twice our distance.

P Cyg

In this field, there are two known compact clusters, Dolidze 41 and IC 4996. Dolidze 41 is situated at RA = 304.65 deg and Dec. = 37.75 deg. It has an apparent radius of 0.15 deg and a distance of $d = 4.6$ kpc. Thus, with the limits in our SNR requirement for the parallax values in our analysis, the stars of this cluster are too far way. The cluster IC 4996 is situated at RA = 304.135 deg and Dec. = 37.65 deg. It has a distance of 1.9 kpc and an age of about 9 Myr. This cluster is easily identified and we determine its proper motion as $\mu_\alpha = -2.60 \pm 0.01$ mas yr^{-1} and $\mu_\delta = -5.35 \pm 0.01$ mas yr^{-1}. The median distance of the cluster stars is $d = 2159 \pm 8.8$ pc.

The main cluster associated with P Cyg (MWSC 3301), should be at a distance of 1457 pc and have an age of 5 Myr (Kharchenko et al. 2013). There is no indication in the *Gaia* DR3 data of this cluster. In other words, selecting stars with distances within about 200 pc from the supposed distance, there are no coherent groups of stars in proper motion space. There is a peak in the distance histogram of the field at about 540 pc, but again there is no coherent group in proper motion space for these sources either.

Berkeley 86

The proper motion distribution of the region shows three coherent groups of sources. Two of then (a and b) are very close together with proper motions of $\mu_\alpha = -3.45 \pm 0.01$ mas yr^{-1} and $\mu_\delta = -5.48 \pm 0.01$ mas yr^{-1} for group a and $\mu_\alpha = -3.25 \pm 0.01$ mas yr^{-1} and $\mu_\delta = -6.01 \pm 0.01$ mas yr^{-1} for group b. The median respective distance for these groups are $d = 1839 \pm 7.8$ pc for group a and $d = 1853 \pm 7.9$ pc for group b. Hence, these are two spatially close clusters with very similar proper motions. Group a stars are distributed in the centre of the field, while the second group is mostly located in the south-west of the area. The third group (c) seems to be a cluster in the north-west of the field. Its stars have a distance of $d = 1636 \pm 13$ pc and proper motions of $\mu_\alpha = -0.70 \pm 0.01$ mas yr^{-1} and $\mu_\delta = -3.08 \pm 0.01$ mas yr^{-1}. In the CMD, the sequence shows a lot of scatter, indicating potentially variable line-of-sight extinction towards the group members.

IC 5070

The distance histogram shows a large drop in numbers for distances greater than 850 pc. This indicates a high extinction layer at the distance of the IC 5070 (or Pelican Nebula) region. The proper motion distribution shows two coherent groups with similar proper motions. The more populous group a is at $\mu_\alpha = -1.33 \pm 0.03$ mas yr^{-1} and $\mu_\delta = -3.08 \pm 0.03$ mas yr^{-1}, while the second group b has a proper motion of $\mu_\alpha = -0.98 \pm 0.04$ mas yr^{-1} and $\mu_\delta = -4.15 \pm 0.04$ mas yr^{-1}. The distances of the two groups are $d = 832.4 \pm 2.7$ pc and $d = 824.7 \pm 4.4$ pc, for a and b, respectively. The more populous group of stars is mostly situated in the west of the field, while the secondary group is in the centre of the field. Our groups a and b correspond to the groups D and C (respectively) identified by Kuhn et al. (2020) based on *Gaia* DR2 data.

IC 1396A

There is a single large peak in the distance histogram of IC 1396 A (Elephant Trunk Nebula, or Tr 37). Similarly there is one well populated, coherent group of sources in the proper motion distribution. There are hints of further subclusters in the area, but all of them do have a limited number of members in the area investigated, and all are spatially distributed. They might represent further subpopulations in this large H II region. The main group has a proper motion of $\mu_\alpha = -2.42 \pm 0.02$ mas yr^{-1} and $\mu_\delta = -4.72 \pm 0.02$ mas yr^{-1}, and a distance of $d = 942.0 \pm 2.4$ pc. The members of the group are clustered to the east of the centre of the field, around the ionising star of the H II region, HD 206267. Our group of stars corresponds to the main group A of objects identified in Pelayo-Baldárrago et al. (2023) based on *Gaia* EDR3 data. This work also presents a more detailed analysis of the substructure in this field (covering also IC 1396 N, see below). The slight differences in the determined distance might be caused by us only analysing the central one degree field of this much larger H II region.

IC 1396N

This field spatially overlaps with the IC 1396 A region (see above). Hence, a full analysis of the entire H II region is needed for a complete picture about the astrometric properties of the young stars in this region (see e.g. Pelayo-Baldárrago et al. 2023). There are two subgroups in the field. The main one has a proper motion of $\mu_\alpha = -2.02 \pm 0.03$ mas yr^{-1} and $\mu_\delta = -3.96 \pm 0.04$ mas yr^{-1}, and a distance of $d = 955.0 \pm 3.1$ pc. These are slightly different to the values for IC 1396 A, indicating some spatial variations within this large (∼25 pc diameter) H II region.

We identify a second coherent group (b) of objects in this field. These stars have a proper motion of $\mu_\alpha = -6.14 \pm 0.03$ mas yr^{-1} and $\mu_\delta = -5.83 \pm 0.03$ mas yr^{-1}, and a distance of $d = 619.8 \pm 2.7$ pc. Thus, this is a population in the foreground. The members are spatially distributed and their properties coincide with one of the potential subgroups discussed in IC 1396 A (see above). The proper motion distribution of this field is very structured and complex and a detailed analysis is beyond the scope of this work.

NGC 7129

This area contains the know old open cluster NGC 7142 in the south-west of the field, at the coordinates RA = 326.00 deg and Dec. = 65.77 deg. This can be easily identified and we find a proper motion of $\mu_\alpha = -2.68 \pm 0.01$ mas yr^{-1} and $\mu_\delta = -1.36 \pm 0.01$ mas yr^{-1} for the cluster members, as well as a distance of $d = 2564 \pm 8.5$ pc. The main cluster of young stars has a proper motion of $\mu_\alpha = -1.73 \pm 0.03$ mas yr^{-1} and $\mu_\delta = -3.38 \pm 0.05$ mas yr^{-1} and a distance of $d = 910.0 \pm 4.1$ pc. There are no other coherent groups of stars in the field.

IC 5146

The main cluster in the field is clearly identifiable as a coherent group in proper motion space at $\mu_\alpha = -2.85 \pm 0.03$ mas yr^{-1} and $\mu_\delta = -2.70 \pm 0.04$ mas yr^{-1}. The distance of the members is $d = 784.8 \pm 5.1$ pc. There are hints of substructure in the proper motion distribution, but no clear indication of, for example, several subgroups. A second group of stars can be identified in the proper motion distribution. This group is spatially distributed in the south of the field with proper motion values of $\mu_\alpha = +1.28 \pm 0.02$ mas yr^{-1} and $\mu_\delta = -2.38 \pm 0.01$ mas yr^{-1}. The median distance of the stars is $d = 1531 \pm 16$ pc. The distribution of these objects in the colour–magnitude plot is in agreement with a slightly reddened main sequence. Thus, this could potentially be a background moving group.

FSR 408

There is one main peak in the distance histogram, which splits into two distinct groups (a and b) in proper motion space. The more populated group a has a proper motion of $\mu_\alpha = -0.88 \pm 0.02$ mas yr^{-1} and $\mu_\delta = -2.29 \pm 0.01$ mas yr^{-1}, and a distance $d = 860.2 \pm 1.8$ pc. It represents the main cluster in the centre of the field and also a smaller cluster to the East of the main group of stars. The proper motion distribution is not symmetric, hinting at more than one subpopulation. The second group (b) has a proper motion of $\mu_\alpha = -2.33 \pm 0.02$ mas yr^{-1} and $\mu_\delta = -2.48 \pm 0.02$ mas yr^{-1}, and a distance $d = 875.6 \pm 3.8$ pc. Thus, these stars are slightly in the background of the main cluster and they are spatially distributed in the entire field. They are most likely part of the general population of sources in Cep OB3.

APPENDIX B: EXAMPLE LIGHT CURVES

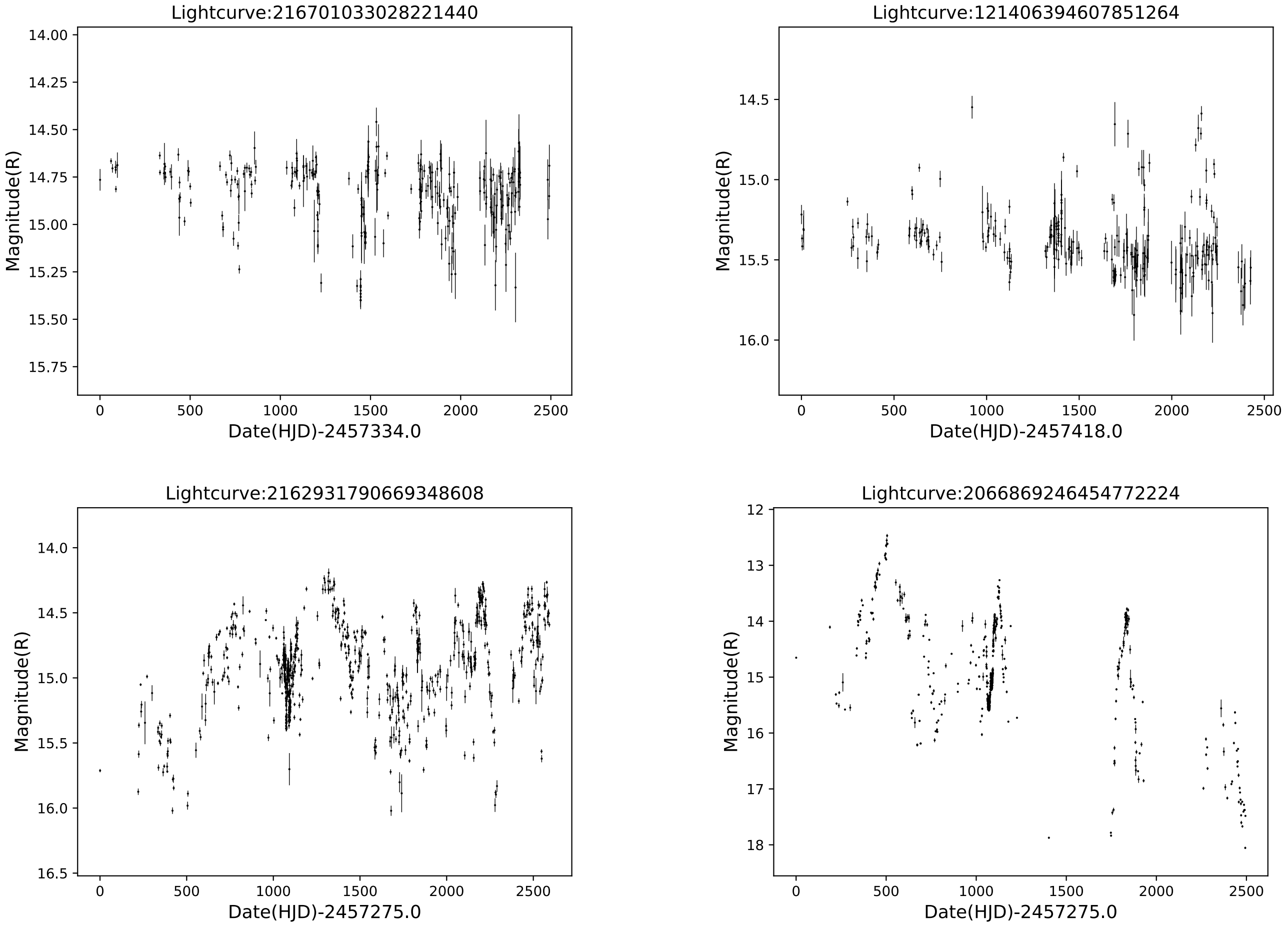

Figure B1. Example HOYS *R*-band light curves with photometric uncertainties for four non-periodic sources. Top left: dipper, *Gaia* DR3 ID = 216701033028221440, $J = 2.127$, $M = 0.736$, and $Q = 0.867$; top right: burster, *Gaia* DR3 ID = 121406394607851264, $J = 2.672$, $M = -0.836$, and $Q = 0.909$; bottom left: symmetric, *Gaia* DR3 ID = 2162931790669348608, $J = 11.569$, $M = 0.159$, and $Q = 0.974$; and bottom right: symmetric, *Gaia* DR3 ID = 2066869246454772224, $J = 33.537$, $M = 0.132$, and $Q = 0.975$.

APPENDIX C: PERIODOGRAM AND PHASE-FOLDED LIGHT CURVES

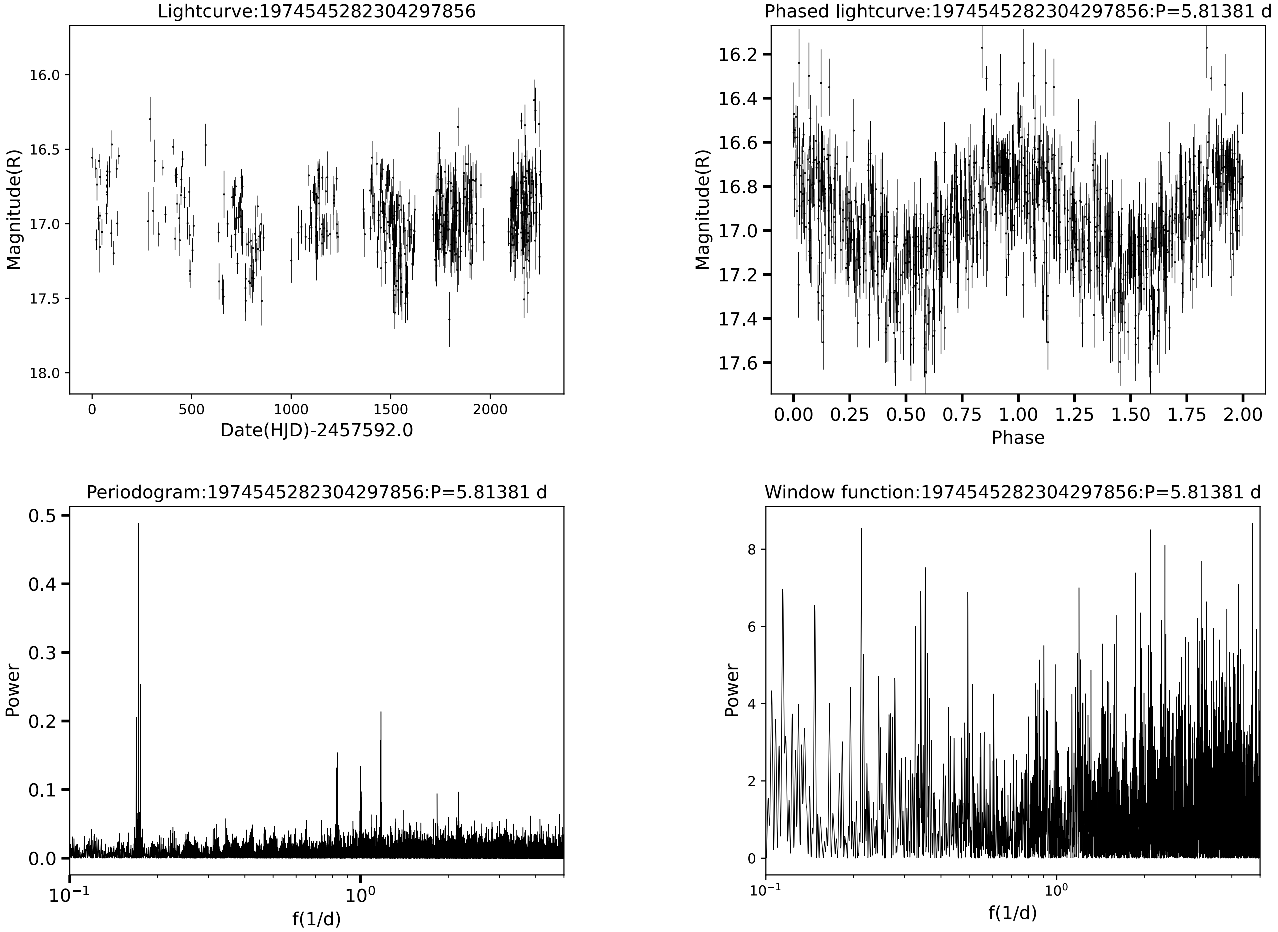

Figure C1. Example HOYS *R*-band data for a periodic source (*Gaia* DR3 ID = 1974545282304297856, period = 5.81381 d). The top left panel shows the complete light curve. The top right panel shows the phase-folded light curve. The bottom left panel shows the periodogram. The bottom right panel shows the window function for this light curve.

[1]*Centre for Astrophysics and Planetary Science, School of Physics and Astronomy, University of Kent, Canterbury CT2 7NH, UK*
[2]*SUPA, School of Physics & Astronomy, University of St Andrews, St Andrews KY16 9SS, UK*
[3]*European Southern Observatory, Karl-Schwarzschild-Strasse 2, D-85748 Garching bei München, Germany*
[4]*Public Observatory ASTROLAB IRIS, Provinciaal Domein 'De Palingbeek', Verbrandemolenstraat 5, B-8902 Zillebeke, Ieper, Belgium*
[5]*Vereniging Voor Sterrenkunde (VVS), Oostmeers 122 C, 'B-8000 Brugge, Belgium*
[6]*Centre for Mathematical Plasma-Astrophysics, Department of Mathematics, KU Leuven, Celestijnenlaan 200B, B-3001 Heverlee, Belgium*
[7]*Thüringer Landessternwarte, Sternwarte 5, D-07778 Tautenburg, Germany*
[8]*British Astronomical Association, Variable Star Section, PO Box 702, Tonbridge TN9 9TX, UK*
[9]*Timtek Systems Limited, 16 Laxton Way, Canterbury, Kent CT1 1FT, UK*
[10]*Centre for Astrophysics Research, University of Hertfordshire, Hatfield, AL10 9AB, UK*
[11]*School of Physics and Astronomy, University of Leicester, University Road, Leicester LE1 7RH, UK*
[12]*Cerro del Viento Observatory (MPC I84), Fernandez Pirfano Square 3, E-06010 Badajoz, Spain*
[13]*Observadores de Supernovas (ObSN, https://www.obsn.es/), Spain*
[14]*American Association of Variable Star Observers (AAVSO), 185 Alewife Brook Parkway, Suite 410, Cambridge, MA 02138, USA*
[15]*Observatorio de Sencelles, Sonfred Road 1, E-07140 Sencelles, Mallorca, Spain*
[16]*Asociacion Astronomica Cruz del Norte, Calle Caceres 18, E-28100 Alcobendas, Madrid, Spain*
[17]*Science Department, Seville University, Av. de la Ciudad Jardín, 20-22, E-41005 Sevilla, Spain*
[18]*La Vara, Valdes Observatory (MPC J38), Barrio La Bara, sin número Muñas de Arriba, Código Postal E-33784 Asturias, Spain*
[19]*R.P. Feynman Observatory, Piazzetta del Gesú 3, I-73034 Gagliano del Capo, Italy*
[20]*Center for Backyard Astrophysics (CBA), Piazzetta del Gesú 3, I-73034 Gagliano del Capo, Italy*
[21]*Belako, Aritz Bidea No 8, 4B Mungia Bizkaia, Basque Country, Spain*
[22]*Hosting Trevinca, Agrupación Astronómica Vizcaina, Bizkaiko Astronomia Elkartea AAVBAE.net, c/ Iparragirre 46, 4° - dpto 2, Bilbo, Basque Country, Spain*
[23]*Sociedad Malagueña de Astronomía (SMA), Centro Cultural José María Gutiérrez Romero, Cl República Argentina, no 9, Urb. El Limonar, E-29016 Málaga, Spain*
[24]*Cal Maciarol Observatory, Camí de l'Observatori S/N, E-25691 Àger, Spain*
[25]*Hampshire Astronomical Group, Hinton Manor Lane, Clanfield PO8 0QR, UK*
[26]*Royal Astronomical Society, Burlington House, Piccadilly, London W1J 0BQ, UK*
[27]*Horndean Observatory, 6 Falcon Road, Horndean, Waterlooville, Hampshire PO89BY, UK*
[28]*Rolling Hills Observatory, Clermont, FL 34711, USA*
[29]*Astronomical Institute, University of Wrocław, ul. M. Kopernika 11, PL-51-622 Wrocław, Poland*
[30]*Astronomical Observatory, University of Warsaw, Al. Ujazdowskie 4, PL-00-478 Warsaw, Poland*
[31]*Astronomical Society of Edinburgh, Edinburgh, UK*
[32]*Wiltshire Astronomical Society, The Knoll, Lowden Hill, Chippenham SN15 2BT, UK*
[33]*Bath Astronomers, 19 New King Street, Bath BA1 2BL, UK*
[34]*The Herschel Society, The Herschel Museum of Astronomy, 19 New King Street, Bath BA1 2BL, UK*
[35]*Green Island Observatory, Karaoğlaonoğlu Street 63A, Geçitkale Mağusa, North Cyprus*
[36]*Bayfordbury Observatory, School of Physics, Engineering and Computer Science, University of Hertfordshire, Lower Hatfield Road, Bayfordbury, Hertfordshire SG13 8LD, UK*
[37]*Volkssternwarte Paderborn e.V., Im Schlosspark 13, D-33104 Paderborn, Germany*
[38]*Bundesdeutsche Arbeitsgemeinschaft für Veränderliche Sterne (BAV), Munsterdamm 90, D-12169 Berlin, Germany*
[39]*Groupe Européen d'Observations Stellaires (GEOS), 23 Parc de Levesville, F-28300 Bailleau l'Evêque, France*
[40]*The Maury Lewin Astronomical Observatory, 420 N. Grand Ave., Glendora CA 91741, USA*
[41]*First Light Observatory Systems, 19807 NE 391st St, Amboy, WA 98601, USA*
[42]*Burnham Observatory, 19 Alexandra Road, Burnham on Crouch, Essex CM0 8BW, UK*
[43]*Crayford Manor House Astronomical Society, Parsonage Lane Pavilion, Parsonage Lane, Sutton-at-Hone, Dartford DA4 9HD, Kent, UK*
[44]*Cork Astronomy Club, 15 Ashdale Park, S Douglas rd. Cork, Ireland*
[45]*Department of Electronics, Electrical Engineering and Microelectronics, Silesian University of Technology, Akademicka 16, PL-44-100 Gliwice, Poland*
[46]*Department of Telecommunications and Teleinformatics, Silesian University of Technology, Akademicka 16, PL-44-100 Gliwice, Poland*
[47]*Ignited Minds VIPNET Club (VP-UP0103), 1/125-G, New Civil Lines, Fatehgarh, Farrukhabad, Uttar Pradesh 209 601, India*
[48]*Akash Ganga: Centre for Astronomy, A/2, East & West Villa, Nowroji Vakil Street, Grant Road, Mumbai, Maharashtra 400 007, India*
[49]*Sternwarte Neanderhoehe Hochdahl, Sedentaler Str. 105, D-40699 Erkrath, Germany*
[50]*Soci/'et/'e Astronomique de France, 3 rue Beethoven, F-75016 Paris, France*
[51]*Dark Skies New Mexico Observatory, 1984 W. Golden Rose Pl, Oro Valley, AZ 85737, USA*
[52]*Astronomical Society of Eastern Missouri (ASEM), 5 Nightfall Court, Dardenne Prairie, MO 63368, USA*
[53]*União Brasileira de Astronomia (UBA), Rua 227, No 25, Setor Leste Universitário, Goiânia, Goiás, CEP 74605-080, Brasil*
[54]*Variable Stars South (VSS), Royal Astronomical Society of New Zealand, PO Box 3181, Wellington, New Zealand*
[55]*MCD Observatory, 23 Langlois, Saint-Anaclet, Quebec G0K1H0, Canada*

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