A&A 663, A131 (2022)
https://doi.org/10.1051/0004-6361/202243288

Astronomy & Astrophysics

An analysis of the most distant cataloged open clusters

Re-assessing fundamental parameters with *Gaia* EDR3 and ASteCA*

G. I. Perren[1,3], M. S. Pera[1,3], H. D. Navone[2,3], and R. A. Vázquez[1,4]

[1] Instituto de Astrofísica de La Plata, IALP (CONICET-UNLP), 1900 La Plata, Argentina
e-mail: gabrielperren@gmail.com
[2] Instituto de Física de Rosario, IFIR (CONICET-UNR), 2000 Rosario, Argentina
[3] Facultad de Ciencias Exactas, Ingeniería y Agrimensura (UNR), 2000 Rosario, Argentina
[4] Facultad de Ciencias Astronómicas y Geofísicas (UNLP), 1900 La Plata, Argentina

Received 8 February 2022 / Accepted 24 March 2022

ABSTRACT

Context. Several studies have been presented in the last few years applying some kind of automatic processing of data to estimate the fundamental parameters of open clusters. These parameters are then employed in larger scale analyses, for example the structure of the Galaxy's spiral arms. The distance is one of the most straightforward parameters to estimate, yet enormous differences can still be found among published data. This is particularly true for open clusters located more than a few kiloparsecs away.
Aims. We cross-matched several published catalogs and selected the 25 most distant open clusters (>9000 pc). We then performed a detailed analysis of their fundamental parameters, with emphasis on their distances, to determine the agreement between the catalogs and our estimates.
Methods. Photometric and astrometric data from the *Gaia* EDR3 survey was employed. The data were processed with our own membership analysis code, pyUPMASK, and our package for the automatic estimation of fundamental cluster parameters, ASteCA.
Results. We find differences in the estimated distances of up to several kiloparsecs between our results and those cataloged, even for the catalogs that show the best matches with ASteCA values. Large differences are also found for the age estimates. As a by-product of the analysis we find that vd Bergh-Hagen 176 could be the open cluster with the largest heliocentric distance cataloged to date.
Conclusions. Caution is thus strongly recommended when using cataloged parameters of open clusters to infer large-scale properties of the Galaxy, particularly for those located more than a few kiloparsecs away.

Key words. methods: statistical – galaxies: star clusters: general – open clusters and associations: general – techniques: photometric – astronomical databases: miscellaneous

1. Introduction

The unprecedented amount of high precision data for parallaxes, proper motions, and photometry provided by the *Gaia* mission in successive deliveries (DR2 and EDR3, Gaia Collaboration 2016, 2021b) offers us a unique opportunity to estimate the fundamental parameters of open clusters (OCs): metal content, age, total mass, binary fraction, distance, and extinction. The arrival of new techniques for analyzing massive quantities of data, combined with the increasing data precision, promises more reliable results than those obtained with the old techniques. The latter were mostly based on the visual inspection of their color–magnitude diagrams and isochrone fittings (Phelps & Janes 1994) or on direct comparison with HR diagrams of synthetic clusters (Siess et al. 1997). Automated processes such as that applied by Kharchenko et al. (2012) have also played a very important role in determining cluster parameters. The continuous increase of high quality data means that a variety of new analyses are being considered including artificial neural networks (Cantat-Gaudin et al. 2020), combined with new strategies for determining cluster memberships (Krone-Martins & Moitinho 2014; Cantat-Gaudin et al. 2018) or dynamical evolution analysis such as that applied by (Gregorio-Hetem et al. 2015).

The intrinsic value of studying OCs has been profusely described in several studies; we give here only a brief summary of the importance of these objects. The oldest OCs allow us to investigate the height and radial extension of the Galactic disk; old OCs tell us about the chemical history (age–metallicity relation), the mixing processes (radial metallicity gradient), and the processes of cluster destruction by interaction with other populations of the Galaxy (Friel 1995; Tosi et al. 2004; Lamers et al. 2005). The youngest OCs, on the other hand, are not only used as laboratories to investigate stellar evolution (they allow the boundary conditions necessary to create new generations of stars to be studied in detail; Lada & Lada 2003), but are also routinely employed in the analysis of the structure of the Milky Way (Loktin & Matkin 1992; Moitinho et al. 2006; Vázquez et al. 2008; Moitinho 2010) and are particularly useful in the tracing of spiral arms (Carraro 2013; Molina Lera et al. 2018).

Young OCs are arranged along the Galactic disk, where the strong visual absorption and the contamination by field stars very often prevent us from observing stars in the lower part of their main sequence. The situation is not much better for the older OCs which do not have very luminous stars in the main

* Table 2 is only available at the CDS via anonymous ftp to cdsarc.u-strasbg.fr (130.79.128.5) or via http://cdsarc.u-strasbg.fr/viz-bin/cat/J/A+A/663/A131

sequence, although they do in the giant branch. Stars in the lower part of the main sequences, as well as those belonging to the giant branch, share similar photometric characteristics with field stars making it rather difficult to unravel to which population each star belongs (Hayes et al. 2015).

The situation worsens as the distance to the older OCs increases because the limiting magnitude increases, which results in only a small portion of the lower part being visible. However, it is not only the photometric data dimensions that are disturbed by distance. The proper motions of distant OCs are extremely difficult to separate from those characterizing the field population against which we see them projected, therefore introducing an additional degree of confusion in determining memberships.

Our interest in this current topic is twofold. First, we focus on reexamining the distances and properties of the most distant OCs cataloged in our Galaxy. A total of 25 clusters that satisfy this requirement were found after inspecting four different recognized catalogs and databases, as we explain below. However, these catalogs display enormous differences in the estimated distances and ages. In part, these differences for the same cataloged object may be due to the varying techniques used to perform the analysis, combined with the problem of the very large distance at which they are located. We want to contribute to the task of resolving these differences.

Second, we want to test our new membership estimation technique pyUPMASK[1] (Pera et al. 2021) in combination with the Automated Stellar Cluster Analysis (`ASteCA`) package[2] (Perren et al. 2015) on clusters with proper motions not easily distinguishable from those of surrounding stars, and that are composed of a small number of members and with non-trivial sequences in the photometric space.

This article is structured as follows. In Sect. 2 we introduce the stellar cluster catalogs, the clusters selected to be analyzed (cross-matched from those catalogs), and the photometric and astrometric data used to perform the analysis. Section 3 presents the methods employed in the study of all the clusters. The comparison of the estimated parameters with the cataloged values for each cluster is done in Sect. 4. Finally, our conclusions are highlighted in Sect. 5.

2. Catalogs, clusters, and data

We selected four catalogs to cross-match and subsequently use to identify the most distant clusters: Dias et al. (2002, the New Catalog of Optically Visible Open Clusters and Candidates, hereafter OC02); Netopil et al. (2012, hereafter WEBDA[3]); Kharchenko et al. (2012, Milky Way Star Clusters Catalog, hereafter MWSC); and Cantat-Gaudin et al. (2020, hereafter CG20).

The first two (OC02 and WEBDA) are compilations of open cluster fundamental parameters from the literature. They contain around 1700 (WEBDA) and 2100 (OC02) entries, and are heavily used in the field of open cluster research. The parameter values in the two catalogs are heterogeneous, being compiled from various sources. The MWSC catalog is the largest (~3000 entries) and, similarly to the CG20 catalog (~2000 entries), is composed of homogeneous fundamental parameter values obtained for all its entries. The method employed by the authors of the MWSC catalog is a semi-automated isochrone fit applied on clusters and candidate clusters, while the CG20 catalog was generated employing an artificial neural network only on verified clusters (trained on parameter values taken from the literature).

Since we are interested in the open clusters most distant from the Sun, we select from these cross-matched catalogs those that are located at a distance of 9000 pc or more in either of them. This is an arbitrary value that results in enough clusters to draw general conclusions, but not too many that would impede their detailed analysis. The final 25 clusters that were studied in this work are shown in Table 1.

Our full list initially consisted of 38 open clusters, 11 of which were found only in the MWSC catalog with distances greater than 9000 pc. These clusters are either listed with substantially smaller distances in the other catalogs, or were too sparse and/or dubious, and were thus removed from the cross-matched list.

Two other clusters were also removed from the initial list: Shorlin 1 (α_{2000} = 166.44, δ_{2000} = −61.23) and FSR0338 (α_{2000} = 327.93, δ_{2000} = 55.33). The latter appears in WEBDA and MWSC at a distance of 12 600 pc and 5600 pc, respectively, while the former is listed only in MWSC at a distance of 14 655 pc. Shorlin1 is studied in Carraro & Costa (2009) and Turner (2012); in both cases the authors conclude that this is not a real cluster, but a grouping of young stars. FSR0338 is analyzed in Froebrich et al. (2010) where a distance of 6000 pc is assigned, but with large uncertainties. In both cases we find no evidence of a true stellar cluster in these regions. We base our conclusion on two findings: the large proper motion dispersion of the stars that occupy the overdensity around the central coordinates assigned to either object and the lack of a clear sequence in their respective color-magnitude diagrams (CMDs). These two clusters were thus also discarded from further analysis.

Most of the 25 selected clusters are located in the third quadrant; all of them are in the latitude range of $[-12°, 8°]$, relatively close to the Galactic plane. The final list thus contains 24 clusters present in the MWSC catalog, 21 in WEBDA, 19 in OC02, and 16 in CG20.

There are two other major works where a large catalog of analyzed open clusters is presented: Liu & Pang (2019) and Dias et al. (2021). The former does not contain clusters with such large distances, and was not used. The latter lists only four clusters that are also present in our set of 25 selected clusters. None of their distances meet the criteria of our selection filter, hence this database was not included. Nevertheless, their distance values are mentioned in the discussion of the results in Sect. 4.

Data from *Gaia* EDR3 (Gaia Collaboration 2016, 2021b) were retrieved for a box of 20 arcmin in length around the central coordinates for all the clusters. We employed equatorial coordinates, parallax, proper motions, and photometry (G, $G_{BP} - G_{RP}$) from this survey. In Fig. 1 we show the 25 selected clusters for each of the four catalogs, positioned on the face-on view of the Galaxy (top), and two edge-on views (center, bottom). The spiral arms are those presented in Momany et al. (2006). The large dispersion for the distances assigned to each cluster in different catalogs is clearly visible, where ideally the position of all the clusters would overlap for the four catalogs.

In what follows we only show the figures for a single representative cluster (Berkeley 29) to avoid clutter and to improve the readability of the article. The plots for the remaining clusters can be found in the appendix.

[1] https://github.com/msolpera/pyUPMASK
[2] http://asteca.github.io/
[3] https://webda.physics.muni.cz/

Table 1. Selected open clusters with a cataloged distance ≥9000 pc, ordered by right ascension.

Cluster	α_{2000}	δ_{2000}	OC02		CG20		WEBDA		MWSC	
			Age	Dist	Age	Dist	Age	Dist	Age	Dist
Berkeley 73 (BER73)	95.5	−6.35	9.18	9800	9.15	6158	9.36	6850	9.15	7881
Berkeley 25 (BER25)	100.25	−16.52	9.70	11 400	9.39	6780	9.60	11 300	9.70	11 400
Berkeley 75 (BER75)	102.25	−24.00	9.60	9100	9.23	8304	9.48	9800	9.30	6273
Berkeley 26 (BER26)	102.58	+5.75	9.60	12 589	–	–	9.60	4300	8.71	2724
Berkeley 29 (**BER29**)	103.27	16.93	9.03	14 871	9.49	12 604	9.03	14871	9.10	10 797
Tombaugh 2 (TOMB2)	105.77	−20.82	9.01	6080	9.21	9316	9.01	13 260	9.01	6565
Berkeley 76 (BER76)	106.67	−11.73	9.18	12 600	9.22	4746	9.18	12 600	8.87	2360
FSR 1212 (F1212)	106.94	−14.15	–	–	9.14	9682	–	–	8.65	1780
Saurer 1 (**SAU1**)	110.23	+1.81	9.70	13 200	–	–	9.85	13 200	9.60	13 719
Czernik 30 (CZER30)	112.83	−9.97	9.40	9120	9.46	6647	9.40	6200	9.20	6812
Arp-Madore 2 (**ARPM2**)	114.69	−33.84	9.34	13341	9.48	11 751	9.34	13 341	9.34	13 338
vd Bergh-Hagen 4 (**BH4**)	114.43	−36.07	–	–	–	–	8.30	19 300	–	–
FSR 1419 (F1419)	124.71	−47.79	–	–	9.21	11 165	–	–	8.38	7746
vd Bergh-Hagen 37 (BH37)	128.95	−43.62	8.84	11 220	8.24	4038	8.85	2500	7.50	5202
ESO 092 05 (E9205)	150.81	−64.75	9.30	5168	9.65	12 444	9.78	10 900	9.30	5168
ESO 092 18 (E9218)	153.74	−64.61	9.02	10 607	9.46	9910	9.02	607	9.15	9548
Saurer 3 (SAU3)	160.35	−55.31	9.30	9550	–	–	9.45	8830	9.30	7075
Kronberger 39 (KRON39)	163.56	−61.74	–	11 100	–	–	–	–	6.00	4372
ESO 093 08 (E9308)	169.92	−65.22	9.74	14 000	–	–	9.65	3700	9.80	13 797
vd Bergh-Hagen 144 (BH144)	198.78	−65.92	8.90	12 000	9.17	9649	8.90	12 000	9.00	7241
vd Bergh-Hagen 176 (BH176)	234.85	−50.05	–	–	–	–	–	13 400	9.80	18 887
Kronberger 31 (**KRON31**)	295.05	+26.26	–	11 900	–	–	–	–	8.50	12 617
Saurer 6 (SAU6)	297.76	+32.24	9.29	9330	–	–	9.29	9330	9.20	7329
Berkeley 56 (**BER56**)	319.43	+41.83	9.60	12 100	9.47	9516	9.60	12 100	9.40	13 180
Berkeley 102 (BER102)	354.66	+56.64	9.50	9638	9.59	10 519	8.78	2600	9.14	4900

Notes. The ages are expressed as the logarithm, and the distances are in parsec. The short names used for the clusters throughout the article in tables and figures are in parentheses. Clusters with no distances below the 9000 pc limit in any of the catalogs are in bold.

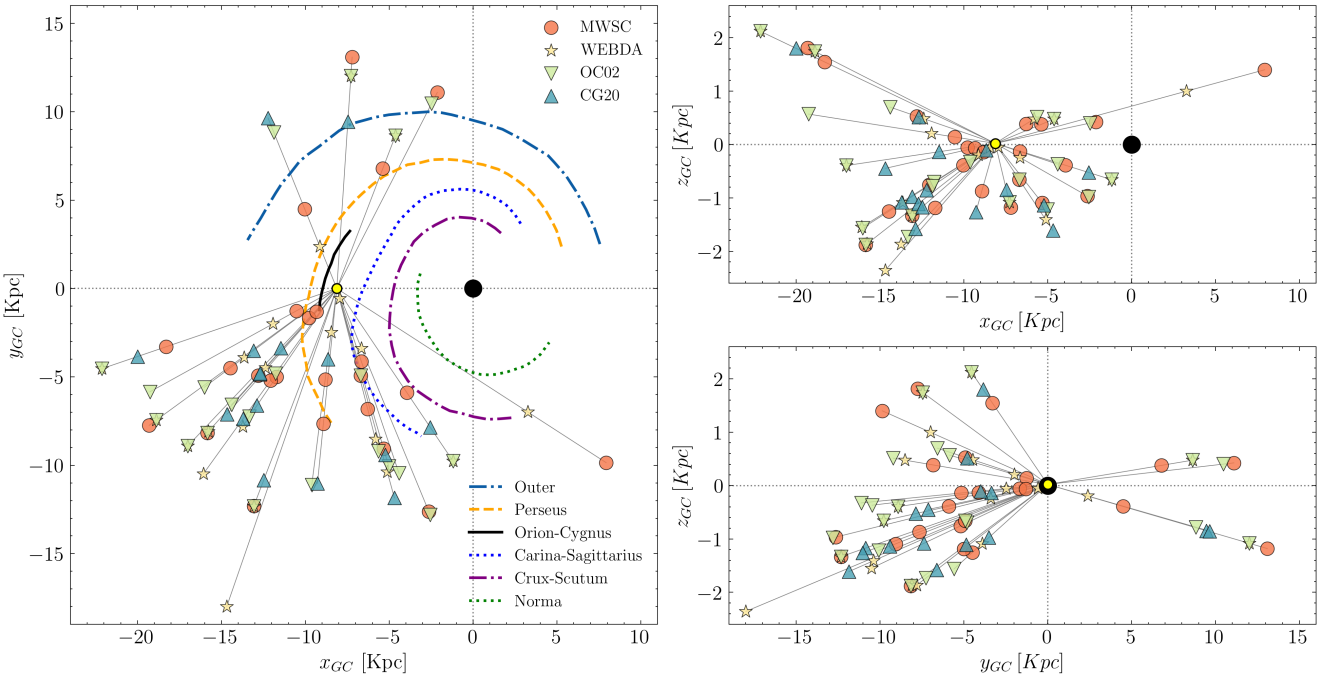

Fig. 1. Position of the 25 clusters selected from the four catalogs mentioned in the text. *Left*: face-on view of the Milky Way. The Sun and the center of the Galaxy are shown as a yellow filled circle and a black filled circle, respectively. *Right, top and bottom*: same as left, but for edge-on views. The sight lines are shown in gray for each cluster.

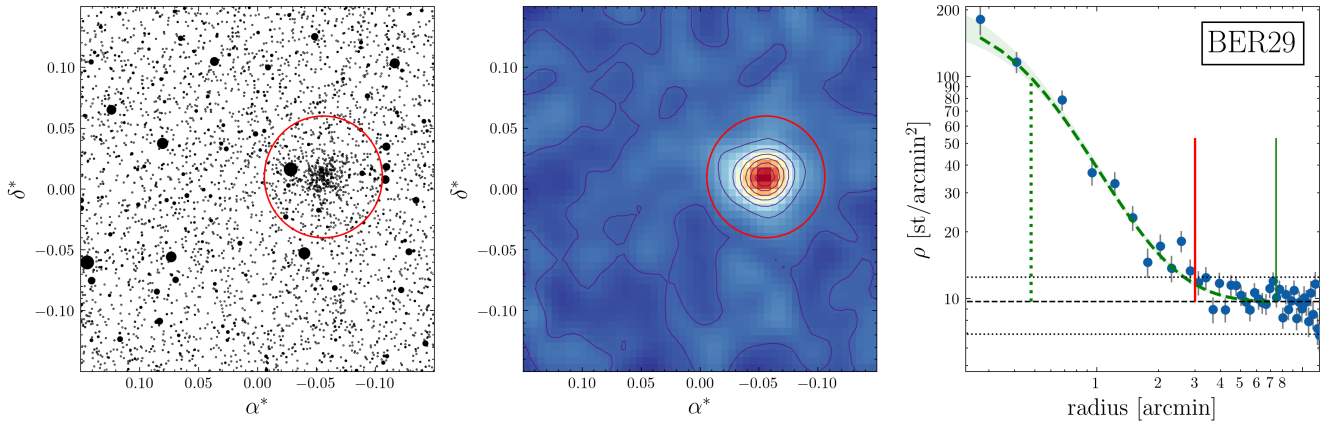

Fig. 2. Structural analysis for the cluster Berkeley 29. *Left*: analyzed 20′ × 20′ arcmin frame with the estimated cluster region enclosed in a red circle. Asterisks in the equatorial coordinates of the left plot indicate that these were shifted and transformed so that the center of the frame is located at (0, 0) and to remove projection artifacts. *Center*: same frame, but shown as a 2D density map. *Right*: radial density analysis, axis shown in logarithmic scale. The dashed green line and the shaded green area are the King profile fit and its 16th–84th uncertainty region, respectively. The green dotted vertical line, solid red vertical line, and solid green vertical line are respectively the core (r_c), adopted (r_a), and tidal (r_t) radii. The dashed and dotted horizontal black lines are the field density estimate and its $\pm 1\sigma$ region, respectively.

3. Cluster analysis

3.1. Structural analysis

The first step in the cluster analysis is the estimation of their structural properties (i.e., center coordinates and limiting radius). Although the centers and diameters are present in some of the catalogs, not all of these values are correct. We use our `ASteCA` package throughout this work to perform the structural and fundamental parameters analysis. We applied this tool to the study of hundreds of clusters in previous articles, with excellent results (Perren et al. 2017, 2020).

The center values are obtained applying a two-dimensional kernel density analysis (KDE) on each of the cluster's coordinates. This method assigns the center of the cluster to the point with the highest density in the frame. As shown in previous articles (Perren et al. 2015, 2017, 2020), this approach is robust even when applied on frames with star densities that are highly non-uniform (see, e.g., the case of van den Bergh-Hagen 37 in Fig. A.4).

A King's profile fit (King 1962) is performed on the radial density profile (RDP) of each cluster to estimate their core and tidal radii (r_c, r_t). The adopted radius r_a is the limiting distance from the center used to define the studied cluster region for each cluster. These radii are estimated applying a process that compares the ratio of the approximated number of true members for increasing radii values with the number of stars in a concentric ring centered on each radius. The approximated number of members is obtained as the total number of stars within the radius, minus the expected number (field density times circle area). This method produces an overdensity around the value where the radial density approaches the field density, maximizing the contrast between members included within the radius and contaminating field stars. The method is also useful for heavily contaminated clusters and/or clusters with very few true members. All radius values are shown in Table A.1.

The adopted radius r_a is on average 50% smaller than the tidal radius (see Table A.1). This allows us to reduce the field star contamination, while ensuring that only a small number of true members (cluster stars located as far from the center as the tidal radius) are lost. The fraction of lost members can be estimated integrating King's profile. This fraction depends on the concentration of the cluster (r_t/r_c) and the value of the adopted radius as a fraction of the tidal radius (r_a/r_t). In our case, less than 20% of the members could be lost in a worst-case scenario. Since these are clusters that are strongly contaminated (particularly in the parallax and proper motion spaces), the trade-off between losing a small portion of members and improving the contrast of the true members over the field noise is positive. Because the r_a values used in our analysis are smaller than the tidal radius, the total estimated mass for each cluster shown in Sect. 4 must be thought of as a lower limit.

In Fig. 2 we show the structural analysis and the center and radius estimation processes for the cluster Berkeley 29. The plots for the remaining clusters can be seen in Appendix A.

3.2. Membership and fundamental parameters

Before we can estimate the fundamental parameters with `ASteCA`, we needed to select the set of most probable members for each cluster. For this task we employed our recently developed pyUPMASK algorithm which performed very well for very contaminated clusters, even outperforming UPMASK (Krone-Martins & Moitinho 2014), as shown in Pera et al. (2021). Internal tests showed that pyUPMASK also outperforms `ASteCA`'s own membership algorithm, hence the reason why we selected the former over the latter.

pyUPMASk requires an input data set composed of (α, δ) coordinates and at least two dimensions of data of any type to estimate the membership probabilities. We chose to make use of the proper motion data dimensions only, thus excluding photometric and parallax data. We made the decision to leave out these extra data dimensions because, although they can sometimes be useful in the process of singling out the most probable members, for this type of very distant clusters they tend to add more noise than information. This is particularly true for the parallax data which rapidly tends to zero for stars beyond ∼2 kpc, where the parallax values for the cluster members become almost indistinguishable from the contaminating field stars. The selected clustering method in pyUPMASK is a Gaussian mixture model, which was demonstrated to have the best performance in Pera et al. (2021, see Sect. 4).

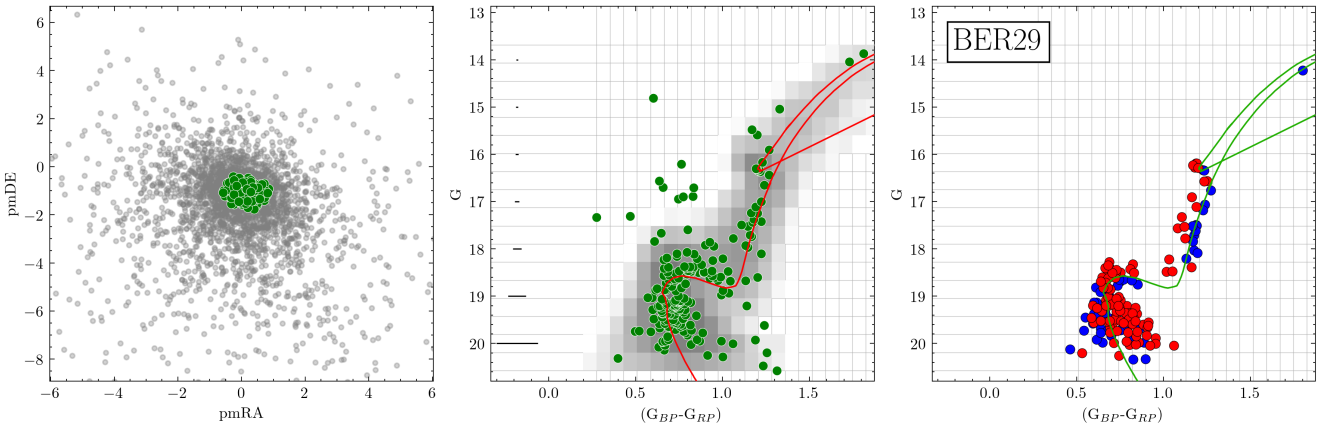

Fig. 3. Estimation of the membership probabilities and the fundamental parameters for the cluster Berkeley 29. *Left*: VPD for stars in the analyzed Berkeley 29 frame. The green and gray circles show the selected true members and the field stars, respectively. *Center*: CMD for the cluster members with the isochrone associated with the best synthetic cluster fit drawn in red to guide the eye. *Right*: Best synthetic cluster fit found by `ASteCA` with the same isochrone, now show in green. The blue and red circles are single and binary systems, respectively.

Once pyUPMASK has assigned membership probabilities to all the stars in the frame, we must select the stars that most likely belong to the cluster (i.e., true members). This selection is performed within the cluster region, defined as $r \leq r_a$, where r is the distance to the cluster center. This step is usually handled by selecting an arbitrary cutoff probability value; in Cantat-Gaudin et al. (2020), for example, the authors fix this value to $P = 70\%$. Instead of setting an ad hoc value, we performed an analysis that combines the membership probabilities with the stellar density inside and outside the cluster region. This allowed us to estimate the number of cluster members expected within the cluster region. Combining this number with the membership probabilities given by pyUPMASK we selected those stars with the highest probabilities within the cluster region, such that the resulting total number of members was as close as possible to the expected number (i.e., that obtained through the stellar density analysis).

Using a physically reasonable number of members not only reduces the probability of excluding true members (by only selecting those with the highest membership probabilities), it also ensures that the estimation of the total mass parameter is properly performed by `ASteCA`.

After selecting the set of true members for all the clusters as described above, we fed this data directly to the final section of our `ASteCA` package bypassing its internal membership algorithm. The goal of this section is to estimate the fundamental parameters: metallicity, age, total mass, fraction of binary systems, distance, and extinction. The code uses the ptemcee parallel tempering Bayesian inference algorithm (Vousden et al. 2016) to sample the distributions of the fundamental parameters. The likelihood function employed to assess the fit between the observed cluster and the synthetic clusters is the Bayesian Poisson ratio defined in Tremmel et al. (2013). The theoretical isochrones used to generate the synthetic clusters used to match the observed clusters are the PARSEC tracks (Bressan et al. 2012). Priors are uniform for all the parameters using the following limiting ranges:

- metallicity ([Fe/H]): [−0.60, 0.30];
- logarithmic age: [8, 10.1];
- total mass: [1e2, 2e5] $M_\odot$;
- binarity fraction: [0, 1];
- distance modulus: [10, 20] mag;
- E_{BV} extinction: [0, E_{BV}^{max}].

The maximum value for the extinction priors, E_{BV}^{max}, was set on a per cluster basis selecting the values given by the Schlegel et al. (1998) extinction maps with the re-calibration by Schlafly & Finkbeiner (2011). The logarithmic abundances [Fe/H] were obtained using the approximation given in the CMD service for [M/H][4] given that the PARSEC isochrones are generated using Z.

In Fig. 3 we show the result of the membership probabilities estimation done with pyUPMASK, plus the fundamental parameter estimation performed by `ASteCA`. We only show here the plots for the cluster Berkeley 29; the remaining clusters can be seen in Appendix B. The plot on the left shows the vector point diagram (VPD) with the proper motion distributions for both the selected clusters members, and the field stars. The members are clearly very much embedded within the field star distribution, which is expected for distant clusters. The center plot shows the CMD traced by the selected members, and the right plot a sampling of the best fit synthetic cluster. The grid in the center and right plots is the two-dimensional binning used to estimate the likelihood, obtained using Knuth's rule (Knuth 2006). The gray region represents the uncertainty in the fit. The isochrone drawn in the center and right plots is associated with the synthetic cluster, but it is there merely to guide the eye; the fit is performed for the CMD of the observed cluster versus the CMD of synthetic clusters, not versus theoretical isochrones (this is further explained in Perren et al. 2015, 2017, 2020).

4. Results and discussion

We present the general results for the fundamental parameters contrasted with values taken from the above-mentioned databases, with particular emphasis on the distances.

In Appendix C we discuss each cluster individually, commenting on the most relevant studies published in the literature and how these compare to the results obtained in this article.

Henceforth we employ the default values for the galactocentric coordinate frame given by the astropy package[5]:

[4] Assuming [M/H] ∼ [Fe/H], [Fe/H] = log(Z/X) − log(Z/X)$_o$, where (Z/X)$_o$ = 0.0207; Y = 0.2485 + 1.78Z. CMD service: `http://stev.oapd.inaf.it/cgi-bin/cmd`

[5] `https://docs.astropy.org/en/stable/coordinates/galactocentric.html`

Table 2. Fundamental parameters estimated with ASteCA for the 25 analyzed clusters.

Cluster	[Fe/H]	log age	E_{BV}	$dm_{\odot}$	Dist (kpc)	M ($M_{\odot}$)	$b_{\rm fr}$	N
BER73	$-0.41^{-0.15}_{-0.50}$	$9.60^{10.02}_{9.27}$	$0.16^{0.35}_{0.06}$	$13.70^{14.30}_{13.08}$	$5.49^{7.25}_{4.14}$	$3.3E+03^{5.9E+03}_{2.1E+03}$	$0.53^{0.81}_{0.22}$	103
BER25	$-0.20^{0.00}_{-0.49}$	$9.72^{9.79}_{9.62}$	$0.39^{0.46}_{0.35}$	$14.34^{14.45}_{14.25}$	$7.37^{7.76}_{7.08}$	$1.5E+04^{2.2E+04}_{9.1E+03}$	$0.82^{0.94}_{0.60}$	213
BER75	$-0.39^{-0.03}_{-0.55}$	$9.74^{9.84}_{9.68}$	$0.11^{0.16}_{0.05}$	$14.52^{14.72}_{14.25}$	$8.03^{8.80}_{7.08}$	$6.8E+03^{1.3E+04}_{3.0E+03}$	$0.77^{0.96}_{0.19}$	95
BER26	$0.07^{0.24}_{-0.20}$	$9.94^{10.06}_{9.79}$	$0.55^{0.59}_{0.50}$	$13.30^{13.67}_{12.96}$	$4.57^{5.43}_{3.91}$	$4.6E+03^{8.5E+03}_{2.3E+03}$	$0.78^{0.95}_{0.45}$	76
BER29	$-0.21^{-0.09}_{-0.32}$	$9.57^{9.61}_{9.52}$	$0.07^{0.10}_{0.04}$	$15.79^{15.88}_{15.67}$	$14.41^{15.03}_{13.64}$	$1.1E+04^{1.8E+04}_{7.4E+03}$	$0.56^{0.82}_{0.34}$	202
TOMB2	$-0.48^{-0.47}_{-0.49}$	$9.33^{9.34}_{9.31}$	$0.39^{0.40}_{0.39}$	$14.70^{14.74}_{14.67}$	$8.73^{8.88}_{8.59}$	$2.1E+04^{2.2E+04}_{1.8E+04}$	$0.45^{0.49}_{0.40}$	845
BER76	$-0.11^{0.02}_{-0.34}$	$9.26^{9.32}_{9.20}$	$0.60^{0.66}_{0.55}$	$13.66^{13.81}_{13.53}$	$5.40^{5.77}_{5.08}$	$4.3E+03^{6.8E+03}_{3.0E+03}$	$0.61^{0.81}_{0.42}$	156
F1212	$-0.12^{0.11}_{-0.34}$	$9.10^{9.16}_{9.05}$	$0.66^{0.72}_{0.60}$	$15.01^{15.20}_{14.85}$	$10.05^{10.97}_{9.34}$	$5.0E+03^{8.4E+03}_{3.6E+03}$	$0.51^{0.76}_{0.33}$	99
SAU1	$-0.08^{0.16}_{-0.34}$	$9.82^{9.92}_{9.71}$	$0.13^{0.17}_{0.07}$	$15.46^{15.66}_{15.22}$	$12.37^{13.54}_{11.07}$	$1.0E+04^{1.7E+04}_{4.8E+03}$	$0.81^{0.96}_{0.46}$	84
CZER30	$-0.32^{-0.04}_{-0.45}$	$9.56^{9.65}_{9.48}$	$0.29^{0.33}_{0.22}$	$14.08^{14.21}_{13.90}$	$6.54^{6.96}_{6.03}$	$7.2E+03^{1.1E+04}_{4.3E+03}$	$0.79^{0.93}_{0.58}$	119
ARPM2	$-0.33^{-0.07}_{-0.50}$	$9.61^{9.69}_{9.56}$	$0.63^{0.66}_{0.58}$	$15.19^{15.30}_{15.10}$	$10.91^{11.46}_{10.46}$	$9.8E+03^{1.4E+04}_{7.1E+03}$	$0.34^{0.57}_{0.17}$	195
BH4	$-0.29^{0.15}_{-0.44}$	$9.10^{9.28}_{8.88}$	$0.34^{0.43}_{0.13}$	$14.55^{15.04}_{14.22}$	$8.12^{10.20}_{6.97}$	$1.8E+03^{3.3E+03}_{1.1E+03}$	$0.64^{0.88}_{0.32}$	66
F1419	$0.02^{0.15}_{-0.38}$	$9.62^{9.85}_{9.49}$	$0.57^{0.65}_{0.50}$	$14.82^{14.99}_{14.43}$	$9.21^{9.95}_{7.71}$	$1.1E+04^{2.0E+04}_{6.8E+03}$	$0.63^{0.88}_{0.40}$	142
BH37	$-0.06^{0.20}_{-0.41}$	$8.87^{9.64}_{8.63}$	$1.22^{1.34}_{0.80}$	$12.28^{12.79}_{11.00}$	$2.85^{3.61}_{1.58}$	$2.5E+03^{4.3E+03}_{1.4E+03}$	$0.52^{0.80}_{0.31}$	90
E9205	$-0.12^{0.13}_{-0.44}$	$9.78^{9.80}_{9.72}$	$0.11^{0.17}_{0.06}$	$15.52^{15.58}_{15.43}$	$12.70^{13.04}_{12.22}$	$3.3E+04^{4.3E+04}_{2.7E+04}$	$0.74^{0.86}_{0.65}$	378
E9218	$-0.30^{-0.30}_{-0.33}$	$9.68^{9.71}_{9.68}$	$0.24^{0.25}_{0.23}$	$15.25^{15.31}_{15.23}$	$11.24^{11.53}_{11.12}$	$5.2E+04^{5.9E+04}_{4.5E+04}$	$0.60^{0.69}_{0.52}$	721
SAU3	$0.02^{0.19}_{-0.34}$	$9.81^{9.91}_{9.51}$	$0.71^{0.78}_{0.64}$	$13.94^{14.30}_{13.77}$	$6.12^{7.26}_{5.68}$	$1.4E+04^{2.0E+04}_{8.3E+03}$	$0.86^{0.96}_{0.67}$	146
KRON39	$-0.11^{0.15}_{-0.43}$	$9.45^{9.67}_{9.34}$	$0.78^{0.86}_{0.69}$	$15.57^{15.80}_{15.25}$	$13.03^{14.44}_{11.22}$	$1.3E+04^{2.6E+04}_{6.4E+03}$	$0.75^{0.94}_{0.41}$	55
E9308	$-0.32^{0.02}_{-0.52}$	$9.91^{10.05}_{9.54}$	$0.67^{0.72}_{0.60}$	$15.61^{15.85}_{15.38}$	$13.25^{14.80}_{11.91}$	$3.4E+04^{6.9E+04}_{1.4E+04}$	$0.68^{0.89}_{0.32}$	60
BH144	$-0.53^{-0.48}_{-0.55}$	$9.02^{9.03}_{8.99}$	$0.83^{0.84}_{0.80}$	$15.01^{15.09}_{14.95}$	$10.06^{10.41}_{9.79}$	$6.9E+03^{7.6E+03}_{6.1E+03}$	$0.32^{0.41}_{0.25}$	307
BH176	$0.15^{0.26}_{-0.03}$	$9.70^{9.71}_{9.62}$	$0.52^{0.58}_{0.49}$	$16.31^{16.39}_{16.24}$	$18.27^{18.96}_{17.72}$	$1.7E+05^{1.9E+05}_{1.3E+05}$	$0.48^{0.61}_{0.35}$	277
KRON31	$-0.37^{0.13}_{-0.49}$	$9.00^{9.06}_{8.92}$	$1.32^{1.35}_{1.22}$	$14.40^{14.74}_{14.23}$	$7.57^{8.87}_{7.02}$	$8.3E+03^{1.2E+04}_{5.7E+03}$	$0.80^{0.93}_{0.65}$	133
SAU6	$-0.11^{0.15}_{-0.40}$	$9.14^{9.44}_{9.05}$	$0.94^{1.04}_{0.85}$	$14.82^{15.10}_{14.26}$	$9.19^{10.48}_{7.10}$	$5.2E+03^{8.2E+03}_{3.5E+03}$	$0.52^{0.74}_{0.33}$	129
BER56	$-0.34^{-0.34}_{-0.35}$	$9.72^{9.72}_{9.71}$	$0.51^{0.51}_{0.50}$	$15.23^{15.25}_{15.18}$	$11.12^{11.23}_{10.87}$	$6.1E+04^{6.7E+04}_{5.3E+04}$	$0.70^{0.75}_{0.60}$	843
BER102	$-0.17^{0.07}_{-0.45}$	$9.69^{9.84}_{9.63}$	$0.45^{0.51}_{0.39}$	$14.33^{14.58}_{14.14}$	$7.35^{8.22}_{6.72}$	$5.8E+03^{9.2E+03}_{4.2E+03}$	$0.55^{0.75}_{0.36}$	156

Notes. Sub and supra indexes indicate the 16th and 84th percentiles, respectively. The last column indicates the number of true members used in the analysis.

- ICRS coordinates of the Galactic center: (266.4051°, −28.936175°);
- Distance from the sun to the Galactic center: 8.122 pc;
- Distance from the sun to the Galactic midplane: 20.8 pc;
- Velocity of the sun in the galactocentric frame as Cartesian velocity components: (12.9, 245.6, 7.78) km s^{-1}.

Table 2 shows the fundamental parameters along with their uncertainties estimated by ASteCA. The Bayesian inference process was allowed to run for enough steps to achieve convergence.

In Fig. 4 we show how the map of the Galaxy shown in Fig. 1 looks, but with the distance parameter values found in this work. The arrows represent the velocity vectors for all the clusters with available radial velocity. The sizes correspond to the estimated masses, and the colors follow the distribution of ages, metallicities, and binary fraction as shown in the color bars to the right of each plot. The values used to construct this figure are presented in Table D.1. It is worth noting that only about half of the clusters are truly beyond the 9 kpc (∼14.8 mag) limit originally used to perform the selection from the published databases.

The cluster vd Bergh-Hagen 176, located in the fourth quadrant in Fig. 4, turns out to be the most distant open cluster cataloged to date with a heliocentric distance greater than 18 kpc. Its status as a bona fide open cluster is nonetheless still questioned; a more detailed discussion is presented in Appendix C.

In a recent study (Anders et al. 2022) per star parameters such as distance, extinction, metallicity, and age were estimated. Comparing the results from this analysis with those from Cantat-Gaudin et al. (2020), the authors find differences in the distance values larger than 3 kpc for clusters located at 6 kpc or more from the Sun. We find even larger discrepancies between our analysis and those taken from the four databases. As shown in Fig. 5, all but the CG20 database show differences of up to 10 kpc for clusters spanning the full distance range. The CG20 database, the one with the better overall match to our values, only shows differences larger than 2 kpc for clusters located beyond ∼10 kpc from the Sun. Taking the uncertainties of both estimates into account, these differences are expected; particularly for such distant clusters.

We see no evident trend that correlates the differences in the distance with the ages (used to color the symbols in the right plots of Fig. 5).

To further investigate the various ways to estimate the distance, we performed two more analyses. First, we re-ran ASteCA for all the clusters using four different combinations of settings for the metallicity and binary fraction parameters. We chose these two parameters because in isochrone-fit analyses they are usually either fixed (e.g., the metallicity is set to solar) or neglected altogether (e.g., the binary fraction).

Second, we compared the distances estimated in this work with those obtained via parallax analysis using three different methods: ASteCA's own Bayesian inference estimation (described in Perren et al. 2020), the distance inferred by

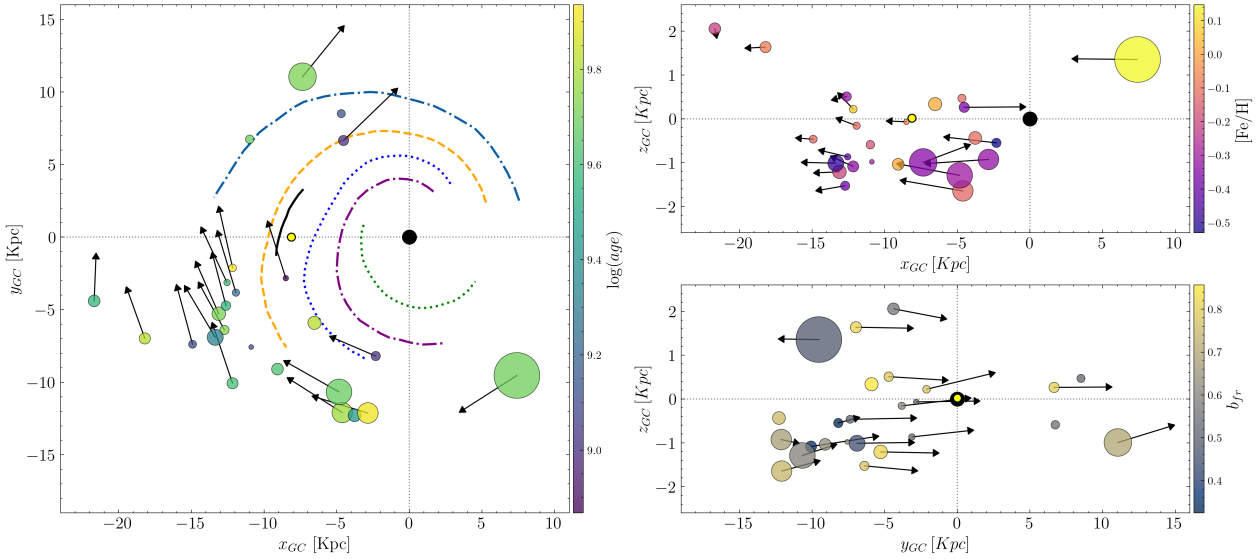

Fig. 4. Same as Fig. 1, but showing the positions given by our analysis with ASteCA. The velocity vectors are drawn for those clusters with available radial velocities. The length of the vectors are proportional to the velocity modules in each 2D projection. Sizes follow masses and colors follow ages, metallicities, and binary fractions for the *left, top right, and bottom right* plots, respectively.

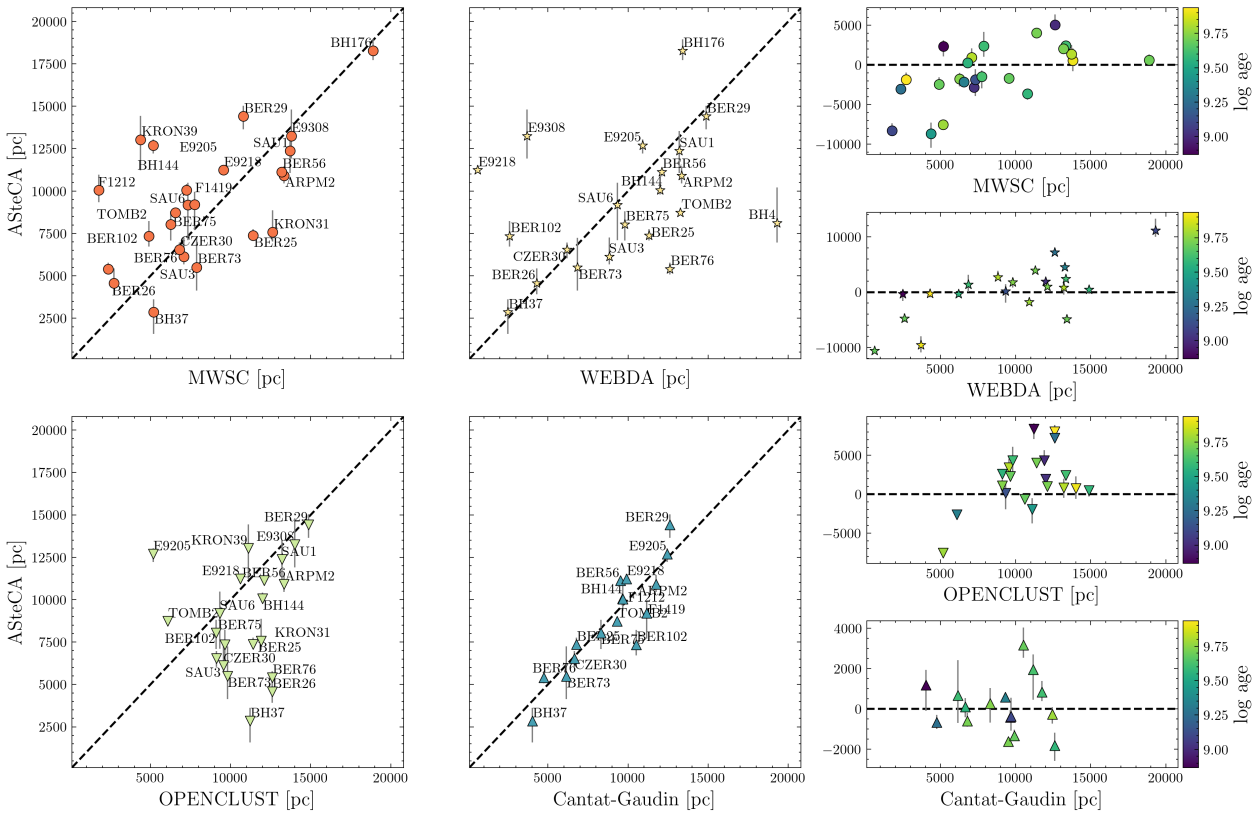

Fig. 5. ASteCA vs. database distances. The plots to the right stacked vertically are the ASteCA distances vs. the differences in the sense (ASteCA – database). Clusters are colored according to the log age assigned by ASteCA.

the Kalkayotl package (Olivares et al. 2020), and the median of a simple inversion of the parallax values of the selected members. The parallax values were previously corrected using the method described in Lindegren et al. (2021)[6].

The results of these two extra analyses can be seen in Fig. 6 compared to the main ASteCA run (i.e., the one whose estimated parameter values are shown in Table 2). In the top plot we show the variation in the distance estimates between our main ASteCA run and four different runs: metallicity fixed to solar and binary

[6] Analytical functions to compute the expected parallax zero-point as a function of ecliptic latitude, magnitude, and color for any *Gaia* (e)DR3 source: https://gitlab.com/icc-ub/public/gaiadr3_zeropoint

fraction fixed to 0 ($Z = Z_{\odot}, b_{\rm fr} = 0.0$; blue left facing triangles); metallicity fixed to solar and binary fraction as a free parameter ($Z = Z_{\odot}$; green right facing triangles); metallicity as a free parameter and binary fraction fixed to 0 ($b_{\rm fr} = 0.0$; orange squares); and metallicity fixed to solar and binary fraction fixed to 0.5 ($Z = Z_{\odot}, b_{\rm fr} = 0.5$; red diamonds), where 50% is chosen to be a typical estimate for binary fraction in open clusters (von Hippel 2005). The median difference with the main ASteCA run is largest when the binary fraction is fixed to 0.0 (∼1100 pc), smaller when we fix this parameter to 0.5 (∼100 pc), and smallest when it is allowed to vary (∼50 pc). This is another indicator that a proper binary fraction fit is of utmost importance for a correct estimation of the cluster's fundamental parameters, particularly for the distance. Even when the binary fraction is free, fixing just the metallicity to solar values can have a non-negligible impact on the estimated distances, as shown in Fig. 6 (green right facing triangles).

The bottom plot in Fig. 6 shows the parallax value analysis. Here the distance estimates obtained by ASteCA processing the *Gaia* EDR3 photometry are compared to three methods to estimate distances using parallaxes: ASteCA's own Bayesian inference, the Kalkayotl package estimate, and the inversion of the median of the selected member's parallaxes. It is clear that a trend arises where the most distant clusters have their distances enormously underestimated by any of the parallax-based methods. This is expected as the parallax values of the most distant clusters are associated with very large uncertainties and are also heavily affected by noise from non-removed field stars that mostly contaminate the lower mass region. It is surprising to see that the naive approach of inverting the median of the member's parallaxes is the method that most closely approximates the photometric distances estimated by ASteCA: the mean difference is only ∼600 pc, where the other two methods show median differences more that twice as large (∼1200 pc).

All the analyzed clusters are rather old; the youngest one (vd Bergh-Hagen 37) has an assigned age of ∼0.7 Gyr, although we note the very large uncertainty associated with it.

The comparisons between our age estimation and those from the four databases are shown in Fig. 7. The top plot shows that ASteCA systematically assigns ages that are older on average than those from the databases. A logarithmic difference of ∼0.23 dex (the average value for all the catalogs) translates to a difference of ∼1.5 Gyr for an age of 3 Gyr, which is a reasonable uncertainty given the complexities associated with the clusters under investigation. The catalog with the smallest logarithmic difference is CG20 with a median of 0.11 dex, again displaying the best match to the values given by ASteCA.

The largest age difference arises for Kronberger 39, for which ASteCA finds an age of log(age) = 9.45, but has an age of log(age) = 6 assigned in the MWSC database (the youngest age by far in the four databases).

As can be seen in the bottom plot of Fig. 7, there appears to be a slight correlation between the difference in age estimates and the binarity fractions. The trend shows that the higher the percentage of binary systems present in the cluster, the larger on average the difference between the age value obtained by ASteCA and those in the databases. This effect can be explained by noticing that the TO in the CMD is pushed downward by the presence of binaries, which are located above the brightest point of the main sequence of single stars. This in turn forces the fit to adjust toward older ages, hence producing the systematic trend seen in the analysis. This result points to the importance of taking binary systems into account when performing stellar clusters' parameter estimations.

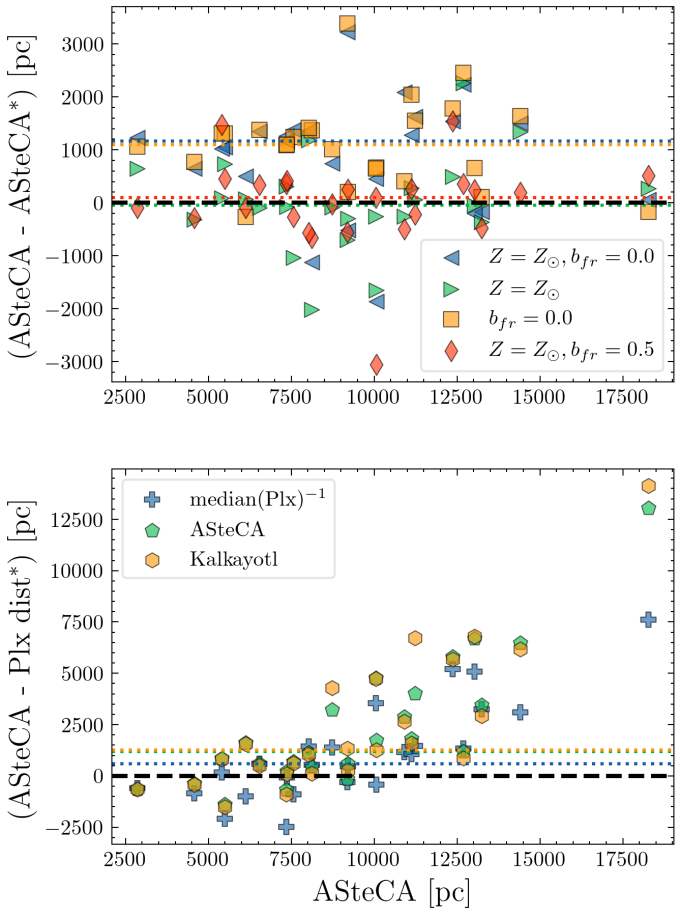

Fig. 6. Distances obtained by the main run of ASteCA compared to other configurations and methods. In both plots the abscissa is the main run ASteCA distance values and the ordinate shows the difference in the sense (ASteCA – ASteCA*), where the asterisk represents either of the four runs from the top plot or either method from the bottom plot. In both plots the black dashed line indicates zero difference and the colored dotted lines the median differences for each run. *Top*: main run distances vs. four different runs where the metallicity and binary fraction were either fixed or allowed to be fitted. *Bottom*: main run distances vs. the difference with the distances estimated using parallax values and three different methods.

The metal content of a cluster is the hardest parameter to estimate photometrically, which is why it is usually fixed to solar value in this type of analysis. We found six clusters from our sample that are also investigated spectroscopically in very recent works: Berkeley 25, Berkeley 29, Berkeley 73, Czernik 30, Saurer 1, and Tombaugh 2 studied in Donor et al. (2020), Netopil et al. (2022), and Spina et al. (2021). The abundances are shown in Table 3 along with the ASteCA estimates. Uncertainties are around 0.02, 0.06, and 0.04 dex for Donor et al. (2020), Netopil et al. (2022), and Spina et al. (2021), respectively. The uncertainties associated with the ASteCA values are substantially larger, averaging 0.2 dex (see Table 2).

In Fig. 8 we show the metallicity vs. galactocentric distance ($R_{\rm GC}$) distribution for the clusters in our sample, plus ten verified clusters from Perren et al. (2020). This distribution (also called radial metallicity distribution or metallicity gradient) is a key tracer of the Galaxy's chemical evolution. Open clusters have been used as a tool to investigate this relation for several decades (Janes 1979). The gradient is usually taken to be around -0.05 dex kpc^{-1} for the inner clusters, with a break

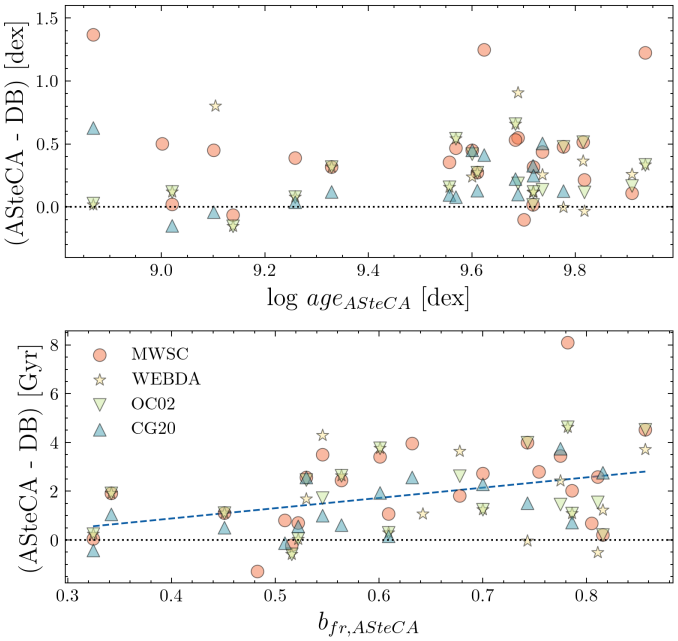

Fig. 7. ASteCA ages vs. ages from the four databases. *Top*: differences in logarithmic ages between ASteCA and the databases in the sense (ASteCA – database). The cluster Kronberger 39 was left out of the plot to improve visibility. *Bottom*: linear age differences vs. binarity fraction assigned by ASteCA. The blue dashed line is the regression trend.

Table 3. Six cluster from our sample whose [Fe/H] metal content is also analyzed in recent works.

Cluster	ASteCA	Donor	Netopil	Spina
BER25	−0.20	–	−0.20 (6)	–
BER29	−0.21	−0.49 (3)	–	−0.48 (1)
BER73	−0.41	–	−0.23 (2)	−0.319 (1)
CZER30	−0.32	−0.40 (2)	–	−0.396 (2)
SAU1	−0.08	−0.42 (1)	–	–
TOMB2	−0.48	–	−0.30 (17)	–

Notes. In parentheses is the number of stars used to estimate each value.

beyond $R_{GC} \approx 10$ kpc into a shallower slope (Donor et al. 2020). In Donor et al. (2018) it is shown that the metallicity gradient is also (as expected) highly dependent on the database used fix the distances, varying by as much as 40% depending on the database used. This result was confirmed in Donor et al. (2020), where a database-dependent variation of 15% was found.

We see in Fig. 8 that ASteCA assigns on average a slightly higher metallicity (∼0.06 dex) than that expected for clusters located below $R_{GC} \approx 14$ kpc.

The case of vd Bergh-Hagen 144 is interesting because its estimated metallicity of [Fe/H] = $-0.53^{-0.55}_{-0.48}$ is well below the expected solar value at that distance ($R_{GC} \sim 8.5$ kpc). There are two other articles where a similar markedly sub-solar metallicity was found for this cluster: Frinchaboy et al. (2006b) and Fragkou et al. (2019). In these studies the reported metallicity values are [Fe/H] = −0.51 ± 0.3 (spectroscopic metallicity from two stars) and [Fe/H] ≈ −0.40 (photometric estimate)[7] for Frinchaboy et al. (2004) and Fragkou et al. (2019), respectively.

[7] Transformed from the fitted $Z = 0.006$ assuming $Z_\odot = 0.0152$.

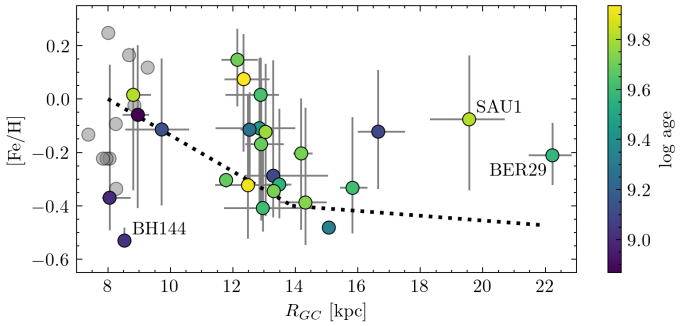

Fig. 8. Metallicity gradient for the set of 25 analyzed clusters. Points are colored according to the log(age). The gray vertical lines are the 16th and 84th percentiles. The dotted line is the broken relation from Donor et al. (2020, Fig. 7). The gray dots are the ten verified clusters from Perren et al. (2020).

Fragkou et al. (2019) assigned a distance of 12 ± 0.5 kpc, ∼4 kpc larger than that found by ASteCA, while in Frinchaboy et al. (2004) the estimated distance is 9.35 kpc, which is a much closer value to ours.

For the seven clusters beyond this galactocentric distance the difference with ASteCA is larger, where our code assigns abundances on average 0.20 dex above the Donor et al. gradient. Saurer 1 is the cluster with the largest value in this group, with an abundance assigned by ASteCA of [Fe/H] = $-0.08^{0.16}_{-0.34}$, which conflicts with the value ∼−0.42 dex expected for its galactocentric distance.

There are several articles where this cluster was assigned a markedly sub-solar [Fe/H] value: −0.27 (Carraro & Baume 2003), −0.38 ± 0.14 (Carraro et al. 2004), −0.50 (Frinchaboy et al. 2004), −0.38 (Frinchaboy et al. 2006a), −0.42 ± 0.01 (Donor et al. 2020). The distances given to Saurer 1 in these studies are located in the range [12, 13.2] kpc, a reasonable match for the distance estimated by ASteCA of $12.4^{13.5}_{11.1}$ kpc. It is thus clear that ASteCA has overestimated the metal content for this cluster. Saurer 1 is the third oldest cluster in our sample (∼6.6 Gyr) and one of the most distant from the Sun, which results in less than a full magnitude visible below the TO with a total of only 84 members present in the CMD. This is very likely the reason for the large difference in the abundance estimated by ASteCA versus the value predicted by the radial metallicity trend and the spectroscopic analyses.

The clusters analyzed in Perren et al. (2020) on the other hand (shown as gray circles in the plot) display a much more balanced distribution around the [Fe/H] ≈ 0.0 dex expected value for their galactocentric distance of ∼9 kpc.

The abundances presented here should therefore be taken with caution. It is always preferable to refer to spectroscopic estimates whenever available, particularly when dealing with very distant, old, scarcely populated, and/or heavily contaminated clusters.

The assigned binary fractions range from 32% to ∼86%, with a mean value of 63% for the entire sample. Although this value is not that far off from the typical value expected for open clusters (50%, as stated previously), it is a bit high. On the other hand, the uncertainties are also rather large and the binary fractions assigned to most of the clusters drop below 50% within their 1σ range. In Sollima et al. (2010) the binary fraction within the core radius of five clusters is estimated. The authors find values in the range 35%–70% with a combined mean value of 56%, somewhat similar to ours. We did not estimate core binary fractions, but we

did employ radii about 50% smaller than the cluster's tidal radii, so the large binarity found could be related to this. In any case, as these are rather distant and old clusters, most of which have a very small portion of their sequence observed, these values should also be used with caution.

The total estimated masses are found in the range [2000, 60 000] $M_\odot$ with the exception of vd Bergh-Hagen 176, for which a much higher mass of ∼170 000 $M_\odot$ is given. This is a high mass value for an open cluster, which would suggest that this object is closer to being classified as a young globular cluster. It is worth noting that these are lower limit mass estimates since ASteCA does not take into account the experienced dynamical mass loss, which can be significant for old stellar clusters (Martinez-Medina et al. 2017).

5. Conclusions

Taking advantage of the precise photometry and proper motions from the most recent *Gaia* data release (Gaia Collaboration 2021b), the fundamental parameters of the 25 most distant cataloged clusters (>9 kpc) have been reassessed using pyUPMASK and ASteCA. The results for the fundamental parameters of metallicity, age, distance, extinction, total mass, and binary fraction are shown in Table 2. In this table we can see that these are rather old clusters: with the exception of just two (vd Bergh-Hagen 37 and Kronberger 31) the remaining clusters are all older than 1 Gyr. Only 13 clusters out of 25 turn out to be at a distance greater than 9 kpc from the Sun, thus reducing the number of clusters that fit the minimum distance criterion by almost half.

Regarding the distribution in the Galactic plane and the galactocentric distance, we see that 14 clusters are placed in the third Galactic quadrant, 10 of which with negative latitudes and thus located below the formal Galactic equator. The remaining four clusters that are above the Galactic plane are Berkeley 26 ($Z \approx 0.2$ kpc), Saurer 1 ($Z \approx 1.6$ kpc), Czernik 30 ($Z \approx 0.5$ kpc), and Berkeley 29 ($Z \approx 2$ kpc). The cluster with the largest galactocentric distance, Berkeley 29, is also the cluster with the largest vertical distance. This cluster is on its course to cross the Galactic disk, as shown by its velocity vector seen in Fig. 4. Maximum vertical distances of ∼1.8 kpc for Berkeley 29 and ∼1.6 kpc for Saurer 1 are estimated in Gaia Collaboration (2021a). These values are in reasonable agreement with the vertical distances obtained here, meaning that both clusters are currently at their maximum height above the Galactic plane.

Despite some bias effect (e.g., lower dust absorption, particularly along the Fitzgerald window; Fitzgerald 1968), it appears that a large number of clusters in the third quadrant of the Galaxy follow the warp defined by the diffuse blue population (Carraro et al. 2005c; Moitinho et al. 2006), whose maximum height above the Galactic equator is located at about [lat = −8°, lon = 240°].

The four databases list 14 clusters with galactocentric distances larger than 15 kpc (in either one of them), the assumed limit for the Galactic disk radius (see Carraro et al. 2010, and references therein). One of the most relevant results emerging from our new distance estimation is that five clusters were confirmed to be located beyond this value: Tombaugh 2 ($R_{GC} \approx$ 15.1 kpc), Arp-Madore 2 ($R_{GC} \approx$ 15.8 kpc), FSR 1212 ($R_{GC} \approx$ 16.7 kpc), Saurer 1 ($R_{GC} \approx$ 19.6 kpc), and Berkeley 29 ($R_{GC} \approx$ 22.2 kpc). These values are in line with recent findings where evidence of populations more than 15 kpc away from the Galactic center was presented (Liu et al. 2017; López-Corredoira et al. 2018, and references therein).

In this work we reported distant open clusters (older tan 1.2 Gyr) instead of single stars, as shown in most of the papers referred to above. A recent review of the spatial distribution of star clusters and the impact of the subsequent releases of *Gaia* data on the topic can be found in Cantat-Gaudin (2022).

When comparing the results given by our analysis with ASteCA with those present in the MWSC, WEBDA, and OC02 databases, we clearly see substantial disagreements in age and distance (the fundamental parameters available in these catalogs). The best overall agreement in distances, the main objective of this article, is found with the database presented in CG20. For the 16 clusters in common with this work the differences range from −2 to +3 kpc, without any apparent dependency on the ages. Within the limits of the uncertainties associated with the distance parameter, we can say that the agreement with CG20 is good. The differences with other databases are substantially larger, spanning a range from −10 to +10 kpc, and are present in the case of MWSC and WEBDA for clusters whose cataloged distance is even below 5 kpc. The age parameter also suffers from important inconsistencies, with the best match found again with the CG20 catalog. Caution is hence advised when making use of these databases for large-scale analyses. We thus recommend choosing the CG20 database over the rest whenever possible.

Acknowledgements. We are grateful for the suggestions and comments given by the referee, which greatly improved the final version of this paper. G.I.P., M.S.P., and R.A.V. acknowledge the financial support from CONICET (PIP317) and the UNLP (PID-G148 project). 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. This research has made use of the WEBDA database, operated at the Department of Theoretical Physics and Astrophysics of the Masaryk University. This research has made use of the VizieR catalog access tool, operated at CDS, Strasbourg, France (Ochsenbein et al. 2000). This research has made use of "Aladin sky atlas" developed at CDS, Strasbourg Observatory, France (Bonnarel et al. 2000; Boch & Fernique 2014). This research has made use of NASA's Astrophysics Data System. This research made use of the Python language v3.7.3 (van Rossum 1995) and the following packages: NumPy (http://www.numpy.org/) (Van Der Walt et al. 2011); SciPy (http://www.scipy.org/) (Jones et al. 2001); Astropy (http://www.astropy.org/), a community-developed core Python package for Astronomy (Astropy Collaboration 2013, 2018); matplotlib (http://matplotlib.org/) (Hunter 2007); scikit-learn (https://scikit-learn.org/) (Pedregosa et al. 2011); ASteCA (https://github.com/asteca); ptemcee (https://github.com/willvousden/ptemcee); Kalkayotl (https://github.com/olivares-j/Kalkayotl); sfdmap (https://github.com/kbarbary/sfdmap).

References

Anders, F., Khalatyan, A., Queiroz, A. B. A., et al. 2022, A&A, 658, A91
Astropy Collaboration (Robitaille, T. P., et al.) 2013, A&A, 558, A33
Astropy Collaboration (Price-Whelan, A. M., et al.) 2018, AJ, 156, 123
Bica, E., Ortolani, S., & Barbuy, B. 1999, A&AS, 136, 363
Boch, T., & Fernique, P. 2014, in Astronomical Data Analysis Software and Systems XXIII, eds. N. Manset, & P. Forshay, ASP Conf. Ser., 485, 277
Bonnarel, F., Fernique, P., Bienaymé, O., et al. 2000, A&AS, 143, 33
Bressan, A., Marigo, P., Girardi, L., et al. 2012, MNRAS, 427, 127
Cantat-Gaudin, T. 2022, Universe, 8, 111
Cantat-Gaudin, T., Jordi, C., Vallenari, A., et al. 2018, A&A, 618, A93
Cantat-Gaudin, T., Anders, F., Castro-Ginard, A., et al. 2020, A&A, 640, A1
Carraro, G. 2013, Proc. Int. Astron. Union, 9, 7
Carraro, G., & Baume, G. 2003, MNRAS, 346, 18
Carraro, G., & Costa, E. 2007, A&A, 464, 573
Carraro, G., & Costa, E. 2009, A&A, 493, 71
Carraro, G., Vallenari, A., & Ortolani, S. 1995, A&A, 300, 128
Carraro, G., Bresolin, F., Villanova, S., et al. 2004, AJ, 128, 1676

Carraro, G., Geisler, D., Moitinho, A., Baume, G., & Vázquez, R. A. 2005a, A&A, 442, 917
Carraro, G., Janes, K. A., & Eastman, J. D. 2005b, MNRAS, 364, 179
Carraro, G., Vázquez, R. A., Moitinho, A., & Baume, G. 2005c, ApJ, 630, L153
Carraro, G., Subramaniam, A., & Janes, K. A. 2006, MNRAS, 371, 1301
Carraro, G., Geisler, D., Villanova, S., Frinchaboy, P. M., & Majewski, S. R. 2007, A&A, 476, 217
Carraro, G., Vázquez, R. A., Costa, E., Perren, G., & Moitinho, A. 2010, ApJ, 718, 683
Carraro, G., Beletsky, Y., & Marconi, G. 2013, MNRAS, 428, 502
Davoust, E., Sharina, M. E., & Donzelli, C. J. 2011, A&A, 528, A70
Dias, W. S., Alessi, B. S., Moitinho, A., & Lépine, J. R. D. 2002, A&A, 389, 871
Dias, W. S., Alessi, B. S., Moitinho, A., & Lepine, J. R. D. 2007, VizieR Online Data Catalog: B/ocl
Dias, W. S., Monteiro, H., Moitinho, A., et al. 2021, MNRAS, 504, 356
Donor, J., Frinchaboy, P. M., Cunha, K., et al. 2018, AJ, 156, 142
Donor, J., Frinchaboy, P. M., Cunha, K., et al. 2020, AJ, 159, 199
Fitzgerald, M. P. 1968, AJ, 73, 983
Fragkou, V., Parker, Q. A., Zijlstra, A., Shaw, R., & Lykou, F. 2019, MNRAS, 484, 3078
Friel, E. D. 1995, ARA&A, 33, 381
Frinchaboy, P. M., & Phelps, R. L. 2002, AJ, 123, 2552
Frinchaboy, P. M., Majewski, S. R., Crane, J. D., et al. 2004, ApJ, 602, L21
Frinchaboy, P. M., Muñoz, R. R., Phelps, R. L., Majewski, S. R., & Kunkel, W. E. 2006a, AJ, 131, 922
Frinchaboy, P. M., Muñoz, R. R., Majewski, S. R., et al. 2006b, Chemical Abundances and Mixing in Stars in the Milky Way and its Satellites, 130
Frinchaboy, P. M., Marino, A. F., Villanova, S., et al. 2008, MNRAS, 391, 39
Froebrich, D., Scholz, A., & Raftery, C. L. 2007, MNRAS, 374, 399
Froebrich, D., Schmeja, S., Samuel, D., & Lucas, P. W. 2010, MNRAS, 409, 1281
Gaia Collaboration (Prusti, T., et al.) 2016, A&A, 595, A1
Gaia Collaboration (Antoja, T., et al.) 2021a, A&A, 649, A8
Gaia Collaboration (Brown, A. G. A., et al.) 2021b, A&A, 649, A1
Gregorio-Hetem, J., Hetem, A., Santos-Silva, T., & Fernandes, B. 2015, MNRAS, 448, 2504
Harris, W. E. 1996, AJ, 112, 1487
Harris, W. E. 2010, ArXiv e-prints [arXiv:1012.3224]
Hasegawa, T., Sakamoto, T., & Malasan, H. L. 2008, PASJ, 60, 1267
Hayes, C. R., Friel, E. D., Slack, T. J., & Boberg, O. M. 2015, AJ, 150, 200
Hunter, J. D., et al. 2007, Comput. Sci. Eng., 9, 90
Janes, K. A. 1979, ApJS, 39, 135
Janes, K. A., & Hoq, S. 2011, AJ, 141, 92
Janes, K. A., & Phelps, R. L. 1994, AJ, 108, 1773
Jones, E., Oliphant, T., Peterson, P., et al. 2001, SciPy: Open Source Scientific Tools for Python [Online; accessed 2016–06-21]
Kharchenko, N. V., Piskunov, A. E., Schilbach, E., Röser, S., & Scholz, R. D. 2012, A&A, 543, A156
King, I. 1962, AJ, 67, 471
King, I. R. 1964, Royal Greenwich Obs. Bullet., 82, 106
Knuth, K. H. 2006, ArXiv e-prints [arXiv:physics/0605197]
Kronberger, M., Teutsch, P., Alessi, B., et al. 2006, A&A, 447, 921
Krone-Martins, A., & Moitinho, A. 2014, A&A, 561, A57
Kubiak, M. 1991, Acta Astron., 41, 231
Kubiak, M., Kaluzny, J., Krzeminski, W., & Mateo, M. 1992, Acta Astron., 42, 155
Lada, C. J., & Lada, E. A. 2003, ARA&A, 41, 57
Lamers, H. J. G. L. M., Gieles, M., Bastian, N., et al. 2005, A&A, 441, 117
Lee, M. G. 1997, AJ, 113, 729
Lindegren, L., Bastian, U., Biermann, M., et al. 2021, A&A, 649, A4
Liu, C., Xu, Y., Wan, J.-C., et al. 2017, Res. Astron. Astrophys., 17, 096
Liu, L., & Pang, X. 2019, ApJS, 245, 32
Loktin, A. V., & Matkin, N. V. 1992, Astron. Astrophys. Trans., 3, 169
López-Corredoira, M., Allende Prieto, C., Garzón, F., et al. 2018, A&A, 612, L8
Maciejewski, G., & Niedzielski, A. 2008, Astron. Nach., 329, 602
Majaess, D., Carraro, G., Moni Bidin, C., et al. 2014, A&A, 567, A1
Martinez-Medina, L. A., Pichardo, B., Peimbert, A., & Moreno, E. 2017, ApJ, 834, 58
Moitinho, A., Vázquez, R. A., Carraro, G., et al. 2006, MNRAS, 368, L77
Moitinho, A. 2010, in Star Clusters: Basic Galactic Building Blocks Throughout Time and Space, eds. R. de Grijs, & J. R. D. Lépine, IAU Symp., 266, 106
Molina Lera, J. A., Baume, G., & Gamen, R. 2018, MNRAS, 480, 2386
Momany, Y., Zaggia, S., Gilmore, G., et al. 2006, A&A, 451, 515
Monteiro, H., Dias, W. S., Moitinho, A., et al. 2020, MNRAS, 499, 1874
Netopil, M., Paunzen, E., & Stütz, C. 2012, in Star Clusters in the Era of Large Surveys, eds. A. Moitinho, & J. Alves (Berlin, Heidelberg: Springer), 53
Netopil, M., Oralhan, İ. A., Çakmak, H., Michel, R., & Karataş, Y. 2022, MNRAS, 509, 421
Ochsenbein, F., Bauer, P., & Marcout, J. 2000, A&AS, 143, 23
Olivares, J., Sarro, L. M., Bouy, H., et al. 2020, A&A, 644, A7
Ortolani, S., Bica, E., & Barbuy, B. 1995, A&A, 300, 726
Ortolani, S., Bica, E., Barbuy, B., & Zoccali, M. 2005, A&A, 439, 1135
Ortolani, S., Bica, E., & Barbuy, B. 2008, MNRAS, 388, 723
Parker, Q. A., Frew, D. J., Miszalski, B., et al. 2011, MNRAS, 413, 1835
Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2011, J. Mach. Learn. Res., 12, 2825
Pera, M. S., Perren, G. I., Moitinho, A., Navone, H. D., & Vazquez, R. A. 2021, A&A, 650, A109
Perren, G. I., Vázquez, R. A., & Piatti, A. E. 2015, A&A, 576, A6
Perren, G. I., Piatti, A. E., & Vázquez, R. A. 2017, A&A, 602, A89
Perren, G. I., Giorgi, E. E., Moitinho, A., et al. 2020, A&A, 637, A95
Phelps, R. L., & Janes, K. A. 1994, ApJS, 90, 31
Phelps, R. L., & Schick, M. 2003, AJ, 126, 265
Phelps, R. L., Janes, K. A., & Montgomery, K. A. 1994, AJ, 107, 1079
Piatti, A. E., Clariá, J. J., & Ahumada, A. V. 2010, MNRAS, 402, 2720
Salaris, M., Weiss, A., & Percival, S. M. 2004, A&A, 414, 163
Saurer, W., Seeberger, R., Weinberger, R., & Ziener, R. 1994, AJ, 107, 2101
Schlafly, E. F., & Finkbeiner, D. P. 2011, ApJ, 737, 103
Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, ApJ, 500, 525
Sharina, M. E., Donzelli, C. J., Davoust, E., Shimansky, V. V., & Charbonnel, C. 2014, A&A, 570, A48
Siess, L., Forestini, M., & Dougados, C. 1997, A&A, 324, 556
Sollima, A., Carballo-Bello, J. A., Beccari, G., et al. 2010, MNRAS, 401, 577
Soubiran, C., Cantat-Gaudin, T., Romero-Gómez, M., et al. 2018, A&A, 619, A155
Spina, L., Ting, Y. S., De Silva, G. M., et al. 2021, MNRAS, 503, 3279
Tadross, A. L. 2008, MNRAS, 389, 285
Tarricq, Y., Soubiran, C., Casamiquela, L., et al. 2021, A&A, 647, A19
Tosi, M., Di Fabrizio, L., Bragaglia, A., Carusillo, P. A., & Marconi, G. 2004, MNRAS, 354, 225
Tremmel, M., Fragos, T., Lehmer, B. D., et al. 2013, ApJ, 766, 19
Turner, D. G. 2012, Ap&SS, 337, 303
van den Bergh, S. 2011, PASP, 123, 1044
van den Bergh, S., & Hagen, G. L. 1975, AJ, 80, 11
Van Der Walt, S., Colbert, S. C., & Varoquaux, G. 2011, Comput. Sci. Eng., 13, 22
van Rossum, G. 1995, Python Tutorial, Report CS-R9526
Vasiliev, E., & Baumgardt, H. 2021, MNRAS, 505, 5978
Vázquez, R. A., May, J., Carraro, G., et al. 2008, ApJ, 672, 930
Villanova, S., Randich, S., Geisler, D., Carraro, G., & Costa, E. 2010, A&A, 509, A102
von Hippel, T. 2005, ApJ, 622, 565
Vousden, W. D., Farr, W. M., & Mandel, I. 2016, MNRAS, 455, 1919