

# Homogeneous analysis of 239 Magellanic clusters with ASt<sub>E</sub>CA

G. I. Perren,<sup>1,3</sup>★ A. E. Piatti,<sup>2,3</sup> and R. A. Vázquez<sup>1,3</sup>

<sup>1</sup>Facultad de Ciencias Astronómicas y Geofísicas (UNLP), IALP-CONICET, La Plata, Argentina

<sup>2</sup>Observatorio Astronómico, Universidad Nacional de Córdoba, Córdoba, Argentina

<sup>3</sup>Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Buenos Aires, Argentina

Accepted XXX. Received YYY; in original form ZZZ

## ABSTRACT

We present a catalogue of 239 resolved Magellanic Clouds clusters observed in the Washington photometric system. The entire cluster sample was processed with the recently introduced ASt<sub>E</sub>CA package, which ensures both an automatized and fully reproducible treatment, together with a statistically based obtention of their fundamental parameters and uncertainties. The main parameters determined for each cluster with this tool are: metallicity, age extinction, true distance modulus, and total mass. Our results allow us to generate a truly objective and homogeneous catalogue composed of structural properties for each cluster, along with a complete determination of their fundamental parameters. A detailed internal error analysis and a thorough comparison with values taken from twenty-six published articles was performed. This lead us to conclude that ASt<sub>E</sub>CA can be applied to the unsupervised determination of fundamental cluster parameters; a task of increasing importance as more data becomes available through upcoming surveys. We present results for the distribution of all fundamental parameters in both Clouds, along with the obtention of the age-metallicity relation from our homogeneous set of star cluster's ages and metallicities.

**Key words:** galaxies: star clusters: general – Magellanic Clouds – techniques: photometric – methods: statistical

## 1 INTRODUCTION

The study of a galaxy's structure, dynamics, star formation history, chemical enrichment history, etc., depends heavily on the analysis of its star clusters. These galactic building blocks are made up of a varying number of coeval stars which share a chemical composition, are located at the same distance, and affected by roughly the same amount of reddening. All these factors greatly facilitate the obtention of their fundamental parameters, and thus the properties of the host galaxy. New developments in astrophysical software allow the homogeneous processing of different types of star clusters' databases. The article series by Kharchenko et al. (see Kharchenko et al. 2005; Schmeja et al. 2014, and references therein) and the integrated photometry based derivation of age and mass for 920 clusters presented in Popescu et al. (2012)<sup>1</sup>, are examples of semi-automated and automated packages applied on a large number of clusters.

Such tools notwithstanding, there is no guarantee that

employing a homogeneous method will result in similar – much less unique – parameter values across different studies. This is particularly true when the methods require user intervention, which inevitably makes the results subjective to some degree. In Netopil et al. (2015) the parameters age, extinction, and distance are contrasted throughout seven published databases, for open clusters matched across them (metallicity is not inspected as this quantity is usually assumed rather than adjusted). The authors find that all articles show non-negligible offsets among each other, in the reported parameter values for the studied star clusters. This result underscores an important issue. Most of the times CMD isochrone fits are done by-eye, adjusting correlated parameters independently and often omitting any kind of proper error treatment (see von Hippel et al. 2014, for a more detailed description of this problem). When statistical methods are employed, the code used is seldom publicly shared to allow scrutiny by the community. There is then no objective way to asses the underlying reliability of each set of results. Lacking this basic audit, the decision of which database values to use becomes a matter of preference.

As demonstrated by Hills et al. (2015), assigning precise fundamental parameters for an observed cluster (OC) is

★ E-mail: gabrielperren@gmail.com

<sup>1</sup> Based on their MASSCLEAN package: <http://www.massclean.org/>

not a straightforward task. Using multiple combinations of up to eight filters ( $UBVRIJHK_s$ ) and three stellar evolution models, they process an open cluster with their Bayesian isochrone fitting technique<sup>2</sup>, and arrive at statistically different results for its final parameter values. The above study was performed on NGC 188, a  $\sim 4$  Gyr Milky Way OC with a well defined sequence that is rather unaffected by field star contamination, and with proper motions and radial velocities data available. One can therefore expect that analysing clusters with a complex or non-standard morphology, affected by a significant amount of field star contamination, observed with fewer filters, and without information about its dynamics (i.e.: a rather standard situation), will be significantly more complicated. A mismatch between several theoretical evolutionary models among themselves, along with an inability to reproduce OCs in the unevolved main sequence range, had already been reported in [Grocholski & Sarajedini \(2003\)](#).

The aforementioned difficulties in the analysis of an OC will only increase if the study is done by-eye, since: a) the number of possible solutions manually fitted is several orders of magnitude smaller than what a code can handle, b) correlations between parameters are almost entirely disregarded, c) uncertainties can not be assigned through valid mathematical means – and are often not assigned at all –, and d) the final values are necessarily highly subjective. The need is clear for an automated general method with a fully open and extensible code base, that takes as much information into account as possible, and is capable of generating reproducible results.

In [Perren et al. \(2015\)](#) (hereafter Paper I) we presented the Automated Stellar Cluster Analysis (**AStECA**<sup>3</sup>) package, aimed towards allowing a more accurate and comprehensive study of star clusters. Through an unassisted process the code analyses an OC's positional and photometric data to derive its fundamental parameters, along with their uncertainties. **AStECA** was applied on a group of 20 observed Milky Way open clusters in Paper I. As shown in that article, the code is able to assign precise parameter values for OCs with low to medium field star contamination, and gives reasonable estimations for heavily contaminated OCs. Every part of this astrophysical package is open and publicly available, and its development is ongoing.

In the present work we apply **AStECA** on 239 OCs from the Small and Large Magellanic Clouds (S/LMC), distributed up to  $\sim 5^\circ$  and  $\sim 8^\circ$  in angular distance from their centres, respectively. The Magellanic Clouds (MCs) are located close enough to us to allow the study of their resolved star clusters. The large number of catalogued OCs –  $\sim 4000$  are listed in the [Bica et al. \(2008\)](#) catalogue – makes them an invaluable resource for investigating the properties of the two most massive galaxies that orbit the Milky Way. The reddening that affects the MCs is relatively small, except for a few regions like 30 Doradus in the LMC, where  $E_{(B-V)}$  extinction can reach values above 0.4 mag ([Piatti et al. 2015b](#)). The

overall low levels of extinction further simplifies the research of these two galaxies, through the obtention of their clusters' parameters. We make use of the  $CT_1$  filters of the Washington photometric system ([Canterna 1976; Geisler 1996](#)), which are known to produce a colour that is highly sensitive to metal abundance ([Geisler & Sarajedini 1999](#)). The results obtained here regarding the metal content of the clusters in our sample, are thus of clear relevance for the analysis of the MCs chemical enrichment history.

This is the first study where such a large sample of resolved star clusters is homogeneously analysed in an automatic way, with all of their fundamental parameters statistically estimated rather than fitted by eye or fixed a priori. Having metal content assigned for 100% of our sample is particularly important, especially considering the state of other catalogues. The latest version of the well known DAML02 database<sup>4</sup> ([Dias et al. 2002](#)) for example, reports abundances for only 13% of the 2167 clusters it contains.

In Sect. 2 we present the OC sample used in this work along with the twenty-six studies used to compare and validate our results. Sections 3 and 4 describe the obtention of the fundamental parameters found with **AStECA**, and analyse their uncertainties respectively. A detailed comparison of our results with published values from the aforementioned articles is performed in 5. Sect. 6 shows how the fundamental parameters obtained with **AStECA** for both Clouds are distributed, along with the galaxies' age-metallicity relations. We give in Sect. 7 a summary of our results and concluding remarks.

## 2 OBSERVED CLUSTERS SAMPLE

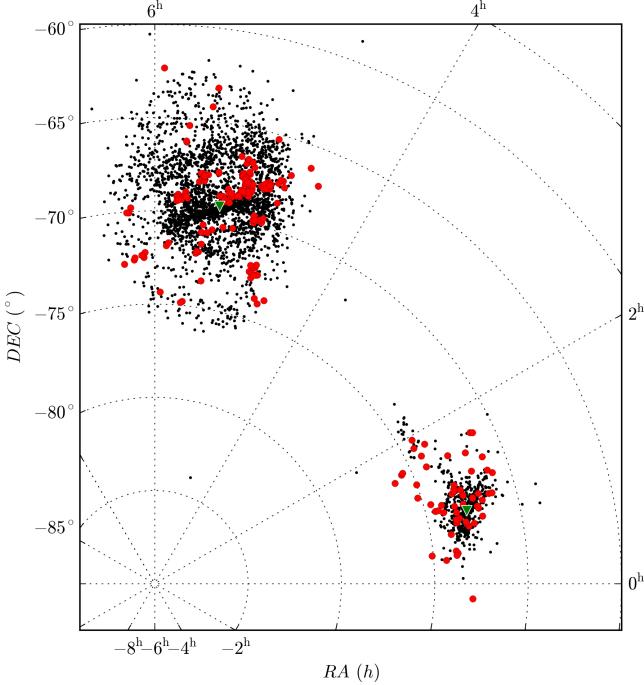
The data used in this work consists of  $CT_1$  Washington photometric system magnitudes and their errors, for the 239 OCs that comprise our sample. The majority of the OCs, 150 out of the 239, belong to the LMC while the remaining 89 belong to the SMC. In Fig. 1 we show the distribution of all OCs in both MCs, along with the entire set of star clusters from the [Bica et al. \(2008\)](#) database for reference.

All OCs have been previously analysed and the results published in the articles shown in Table 1. These nineteen articles can be merged into a single group, whose results arise from the analysis of the same  $CT_1$  photometry used by **AStECA** in the current study. From here on we will refer to this group as the “literature” and the values presented in each of the articles in Table 1 as “literature values”. Details on the data handling and reduction process can be found in the literature, and will not be repeated here. Metallicity, age, reddening, and true distance modulus ( $\mu_o$ ) values are assigned for all OCs in the literature, except for the 36 clusters presented in [Piatti \(2011b\)](#) which had only their ages estimated via the  $\delta T_1$  index ([Phelps et al. 1994; Geisler et al. 1997](#)). In most works metallicity is fixed to  $z = 0.004$  and  $z = 0.008$  ( $[Fe/H] = -0.7$ ,  $[Fe/H] = -0.4$ ), while the distance modulus values are always fixed to  $\mu_o = 18.5$  mag and  $\mu_o = 18.9$  mag, for the S/LMC respectively. Ages and extinctions reported in the literature were in all cases obtained by

<sup>2</sup> BASE-9: <http://webfac.db.erau.edu/~vonhippt/base9/>

<sup>3</sup> **AStECA** is released under a general public license (GPL v3; <https://www.gnu.org/copyleft/gpl.html>) and can be downloaded from its official site: <http://asteca.github.io>

<sup>4</sup> Latest version: v3.5, 2016 Jan 28; <http://www.wilton.unifei.edu.br/ocdb/>



**Figure 1.** Distribution of our set of analysed OCs, shown in red over the Bica et al. (2008) database of 3740 star clusters (black dots) for both MCs. The assumed centres of the clouds are marked with a green triangle.

eye, either through the standard isochrone technique or applying the  $\delta T_1$  index method. Maia et al. (2013) is the only article that presents total mass estimations for their OCs sample.

Our set of 239 OCs was also cross-matched with databases from seven published articles, as shown in Table 2. These can be combined into a single group of works, were different photometry was used to obtain the parameter values of each OC. This group will be referred from here on as the ‘‘databases’’ (DBs), as a way to distinguish them from the ‘‘literature’’ articles mentioned above. In Sect. 5 we show how the parameter values obtained by *ASteCA* for our sample of OCs, compares to those given in both the literature and the databases.

### 3 OBTENTION OF OBSERVED CLUSTERS PARAMETERS

All the fundamental – metallicity, age, true distance modulus, extinction, mass – and structural parameters – centre coordinates, radius, contamination index, approximate number of members, membership probabilities, true cluster probability – associated with the OCs in our database, were obtained either automatically or semi-automatically by the *ASteCA* package. A detailed description of the functions built within this tool can be found in Paper I, and in the code’s online documentation<sup>5</sup>.

The final catalogue, composed of the fundamental and

<sup>5</sup> *ASteCA* documentation: <http://asteca.rtfd.org>

**Table 1.** Set of articles where the same photometric data was used as that employed by *ASteCA* in this work.  $N$  is the number of clusters in this combined ‘‘literature’’ database that are present in the corresponding article. The CTIO 0.9m, 1.5m, and Blanco 4m telescopes are located at Cerro Tololo (Chile); the Danish 1.54m telescope is located at La Silla Observatory (Chile).

Article	$N$	Galaxy	Telescope
Geisler et al. (2003)	8	LMC	CTIO 0.9m
Piatti et al. (2003a)	5	LMC	CTIO 0.9m
Piatti et al. (2003b)	6	LMC	CTIO 0.9m
Piatti et al. (2005)	8	SMC	CTIO 0.9m
Piatti et al. (2007a)	4	SMC	CTIO 0.9m
Piatti et al. (2007b)	2	SMC	Danish 1.54m
Piatti et al. (2007c)	2	SMC	Danish 1.54m
Piatti et al. (2008)	6	SMC	Danish 1.54m
Piatti et al. (2009)	5	LMC	CTIO 0.9m / Danish 1.54m
Piatti et al. (2011a)	3	LMC	CTIO 0.9m
Piatti et al. (2011b)	14	SMC	CTIO 1.5m
Piatti (2011a)	9	SMC	Blanco 4m
Piatti (2011b)	36	LMC	Blanco 4m
Piatti (2011c)	11	SMC	Blanco 4m
Piatti (2012)	26	LMC	Blanco 4m
Piatti & Bica (2012)	4	SMC	Blanco 4m
Palma et al. (2013)	23	LMC	Blanco 4m
Maia et al. (2013)	29	SMC	Blanco 4m
Choudhury et al. (2015)	38	LMC	Blanco 4m

**Table 2.** Set of articles where different photometric data was used to analyse the OCs sample.  $N$  is the number of clusters in each database that could be cross-matched with our set of 239 clusters. The photometry used in each article is shown in the ‘‘Phot’’ column.

Article	$N$	Galaxy	Phot
Pietrzynski & Udalski (1999), P99	7	SMC	BVI
Pietrzynski & Udalski (2000), P00	25	LMC	BVI
Hunter et al. (2003), H03	62	S/LMC	UBVR
Rafelski & Zaritsky (2005), R05	24	SMC	UBVI
Chiosi et al. (2006), C06	16	SMC	VI
Glatt et al. (2010), G10	61	S/LMC	UBVI
Popescu et al. (2012), P12	48	LMC	UBVR

structural parameters for the 239 resolved Magellanic star clusters in our sample, can be accessed via Vizier<sup>6</sup>. We’ve made available online the entire Python codebase developed to analyse the data obtained with *ASteCA* and generate the figures in the article.<sup>7</sup> In addition, the full output image generated by *ASteCA* for each processed cluster together with its corresponding membership probability file, can be accessed separately through a public code repository.<sup>8</sup>

#### 3.1 Ranges for fitted fundamental parameters

Before *ASteCA* is able to process an OC, the user is required to provide a suitable range for each free (fitted)

<sup>6</sup> <http://vizier.XXXX>

<sup>7</sup> <https://github.com/Gabriel-p/mc-catalog>

<sup>8</sup> <https://github.com/Gabriel-p/mc-catalog-figs>

fundamental parameter. This is accomplished by setting a minimum, a maximum, and a step value in the appropriate input data file. As explained in Sect. 3.4, each single value combination for each of the five fundamental parameters fitted represents a unique synthetic cluster or model. Hence an interval too broad between the minimum and maximum values for a given parameter, or a very small step – or a combination of both – will result in a large number of possible values. The larger the number of values a parameter can access, the larger the number of models the code will have to process to find the one with the best match to the OC. For this reason, the allowed ranges and steps were selected to provide a balance between a reasonably large interval, and a computationally manageable number of total models. Especial care must be taken to avoid defining a range that could bias the results towards a particular region of any fitted parameter. All ranges can be seen in Table 3, with the rationale behind each one explained below.

Unlike most previous works where the metallicity is a pre-defined fixed value (usually  $z = 0.008$  for the LMC and  $z=0.004$  for the SMC), we do not make any assumptions on the cluster’s metal content. The minimum and maximum [Fe/H] values selected are [ $\sim -2.2$ , 0]; with a step of  $\sim 0.1$  dex. This interval covers completely the usual metallicities reported for OCs in the MCs. The age range is set to 1 Myr–12.6 Gyr with a step of  $\log(\text{age}/\text{yr}) = 0.05$ , encompassing almost the entire allowed range of the CMD service<sup>9</sup> where the theoretical isochrones were obtained from (see Sect. 3.4).

The maximum allowed value for the extinction of each cluster was determined through the MCEV<sup>10</sup> reddening maps (Haschke et al. 2011), while the minimum value is always zero. These maps contain a large number of observed areas with assigned  $E_{V-I}$  extinction values. The TOPCAT<sup>11</sup> tool was used to query a region as small as possible around the positions of each our OCs for MCEV reddening values. For approximately 85% of our sample we found several regions with an associated reddening value, within a box of 0.5 deg centred on the OC’s position. For the remaining OCs, larger boxes had to be used to find regions with assigned  $E_{V-I}$  values. The two most extreme cases are NGC1997 ( $\alpha=5^h30^m34^s$ ,  $\delta=-63^\circ12'12''$ , [J2000.0]) and OHSC28 ( $\alpha=5^h55^m35^s$ ,  $\delta=-62^\circ20'43''$ , [J2000.0]) in the outliers of the LMC, where boxes of 4 deg and 6 deg respectively where needed to find a region with an assigned reddening value. Once all close-by MCEV regions for an OC are found, the maximum value  $\text{MCEV}_{\max}$  among all regions is selected as the extinction range’s upper limit. Three steps are used to ensure that the reddening range is partitioned similarly for all  $\text{MCEV}_{\max}$  values: 0.01 for  $\text{MCEV}_{\max} > 0.1$ , 0.02 if  $0.05 \leq \text{MCEV}_{\max} \leq 0.1$ , and 0.005 for  $\text{MCEV}_{\max} < 0.05$ . The  $E_{V-I}$  extinction is converted to  $E_{B-V}$  following Tammann et al. (2003):  $E_{V-I}=1.38E_{B-V}$ , with an extinction law of  $R_V=3.1$  applied throughout the analysis.

The true distance modulus ranges for the S/LMC Clouds were assumed to be  $18.96 \pm 0.1$  mag and  $18.46 \pm$

<sup>9</sup> CMD: <http://stev.oapd.inaf.it/cgi-bin/cmd>

<sup>10</sup> Magellanic Clouds Extinction Values (MCEV): <http://dc.zah.uni-heidelberg.de/mceextinct/q/cone/form>

<sup>11</sup> TOPCAT: <http://www.star.bris.ac.uk/~mbt/topcat/>

**Table 3.** Fundamental parameters’ ranges used by ASteca on our set of 239 OCs. The approximate the number of values used for each parameter is N, which gives total of  $\sim 2.3 \times 10^7$  possible models.

Parameter	Min	Max	Step	N
[Fe/H]	$\sim -2.2$	0	$\sim 0.1$	23
$\log(\text{age}/\text{yr})$	6.	10.1	0.05	82
$E_{B-V}$	0.0	$\text{MCEV}_{\max}$	$[0.5, 1, 2] \times 10^{-2}$	$\sim 12$
$\mu_{SMC}$	18.86	19.06	0.02	10
$\mu_{LMC}$	18.4	18.6	0.02	10
Mass ( $M_\odot$ )	10	$[1, 3] \times 10^3$	200	50/150

0.1 mag respectively; where the mean values are taken from de Grijs & Bono (2015) and de Grijs et al. (2014). The 0.1 mag deviations allowed give depths of  $\sim 5.7$  kpc and  $\sim 4.6$  kpc for the S/LMC. This covers entirely the line of sight depths found in Subramanian & Subramaniam (2009) for both MCs.<sup>12</sup>

For the total cluster mass, the range was set to a minimum of  $10 M_\odot$  and a maximum of  $10000 M_\odot$ . This is true for almost all the OCs in our sample, except for 15 visibly massive clusters for which the maximum mass limit was increased to  $30000 M_\odot$ . The step used to divide the mass range was  $200 M_\odot$ , which also sets the minimum possible uncertainty for a cluster’s mass estimate.

Finally, the binary fraction was fixed to a value of 0.5 – considered a reasonable estimate for OCs (von Hippel 2005; Sollima et al. 2010) – to avoid introducing an extra degree of complexity into the fitting process. Secondary masses are randomly drawn from a uniform mass ratio distribution of the form  $0.7 \leq q \leq 1$ , where  $q=M_2/M_1$ , and  $M_1$ ,  $M_2$  are the primary and secondary masses. This range for the secondary masses was found for example for the LMC cluster NGC 1818 in Elson et al. (1998), and represents a value commonly used in analysis regarding the MCs (see Rubele et al. 2011, and references therein).

### 3.2 Centre and radius assignment

In fully automatic mode, ASteca uses a Gaussian kernel density estimator (KDE) function to determine the centre of the OC as the point of maximum star density. A radial density profile (RDP) is used in this same mode to estimate the OC’s radius, as the point where the RDP reaches the star density of the surrounding field.

The first function requires that the density of cluster members make the OC stand out over the combination of foreground and background field stars, and that no other over-density is present in the observed frame.<sup>13</sup> The second function will give good results when the RDP generated is reasonably smooth, and only if the OC does not occupy the entire observed frame; i.e.: some portion of the surrounding field must be visible. When these conditions are not met

<sup>12</sup> Line of sight depth, SMC:  $4.9 \pm 1.2$  kpc (bar),  $4.23 \pm 1.48$  kpc (disk); LMC:  $4.0 \pm 1.4$  kpc (bar),  $3.44 \pm 1.16$  kpc (disk).

<sup>13</sup> The single over-density limitation is planned to be lifted in upcoming versions of ASteca.

for either function, the user needs to run the code on semi-automatic mode. In this mode the centre coordinates can be either manually fixed, or found by the code based on an initial set of approximate values. Similarly, in semi-automatic mode the radius must be fixed by the user.

For over  $\sim 66\%$  of our OCs sample the centre coordinates were obtained via a KDE analysis, based on initial approximate values passed to the code. Radii values were calculated by *AStéCA* for  $\sim 60\%$  of the OCs, through an RDP analysis of the observed surrounding field. The remaining OCs are those that are highly contaminated, contain very few members, and/or occupy most of the observed frame. This means that an automatic obtention of their centres and/or radii is not possible, and the semi-automatic mode must be used.

The contamination index ( $CI$ ) calculated by the code, is a parameter related to the number of foreground/background stars present in the defined cluster region. A value of  $CI=0$  means that the OC is not contaminated by field stars,  $CI=0.5$  means the overdensity is indistinguishable from the surrounding field (i.e.: the average number of field stars expected within the cluster region equals the total number of stars in it), and  $CI>0.5$  means more field stars are expected within the cluster region than true cluster members.<sup>14</sup> For reference, the average  $CI$  of the set of OCs with manual radii assignment is  $CI \approx 0.9 \pm 0.2$ , while the same average is  $CI \approx 0.6 \pm 0.2$  for those OCs where the radius was automatically estimated. Naturally, more contaminated OCs are the ones that require more often that the radius be set manually, to prevent the code from employing a wrong value in its analysis.

### 3.3 Field-star decontamination

A decontamination algorithm (DA) was employed on the  $CT_1$  colour-magnitude diagram (CMD) of each processed OC. The DA allows cleaning the CMD of field-star contamination, before the isochrone fitting function (see 3.4) is used. The DA must therefore process all stars within the defined cluster region, i.e: the circular region of centre and radius either automatically determined by the code or manually fixed by the user.

The Bayesian DA presented in Paper I was improved for this article, the new DA works in two steps. First, the original Bayesian membership probability (MP) assignation is applied to the CMD of all stars within the cluster region (see Paper I for more details). After that, a cleaning algorithm is used to remove stars of low MPs from the CMD, as shown in Fig. 2. To do this, the full CMD is divided into smaller cells, according to a given binning method. By default *AStéCA* uses the Bayesian blocks method<sup>15</sup> introduced in Scargle et al. (2013), via the implementation of the astroML package (Vanderplas et al. 2012).<sup>16</sup> This second step

<sup>14</sup> See Paper I, Sect. 2.3.2 for a complete mathematical definition of the index.

<sup>15</sup> [http://www.astroml.org/examples/algorithms/plot\\_bayesian\\_blocks.html](http://www.astroml.org/examples/algorithms/plot_bayesian_blocks.html)

<sup>16</sup> While Bayesian blocks binning is the default setting in *AStéCA*, several others techniques for CMD star removal are available, as well as five more binning methods.

is similar to the field-star density based cell-by-cell removal algorithm applied in Bonatto & Bica (2007, B07), which uses a simpler rectangular grid. The main difference, aside from the binning method employed, is that *AStéCA* can remove stars cell-by-cell based on their previously assigned MPs, not randomly as done in B07.

For over  $\sim 70\%$  of our sample the default settings in *AStéCA* were used in the decontamination process, this is: Bayesian DA followed by the removal of low MP stars, based on a Bayesian block binning of the cluster region's CMD. The remaining OCs were processed with modified settings to allow a proper field-star decontamination, mainly for those OCs with a low number of members or a heavily contaminated evolutionary sequence. Changes in the settings can go from selecting a different binning method (often a rectangular grid using Scott's rule, Scott 1979), to skipping the Bayesian MP assignation and only performing a density based cell-by-cell statistical field star removal (in this last case, the DA works very similarly to the B07 algorithm).

A proper field-star decontamination is of the utmost importance, as the resulting – hopefully – clean cluster sequence will determine the OCs' fundamental parameters via the isochrone fitting process that follows.

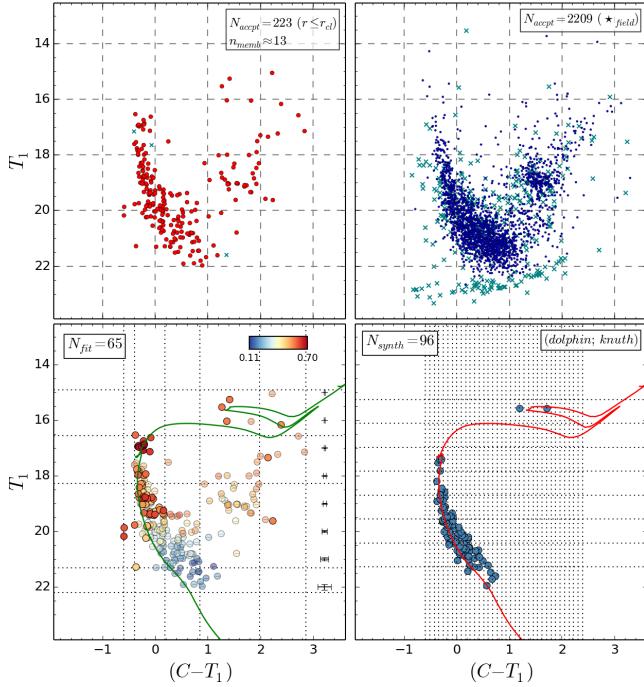
### 3.4 Isochrone fitting

The last part of the automatic analysis performed by *AStéCA* is the isochrone fitting of the OC, to estimate its fundamental parameters. In rigour, the process does not employ isochrones but rather synthetic clusters (SC) generated from theoretical isochrones.<sup>17</sup> In this work we use a PARSEC v1.1 set of isochrones (Bressan et al. 2012, B12). An isochrone of given metallicity and age is combined with a total mass value and with the log-normal form of the initial mass function (IMF) defined by Chabrier (2001), to generate a SC with a correct mass distribution. The IMF is stochastically sampled until the desired total cluster mass is reached. The position of this SC in the CMD is then affected by reddening and moved a certain distance modulus value, to obtain the final SC – or model – that will be compared to the OC.

The Poisson likelihood rate (PLR) developed in Dolphin (2002) is employed to asses the goodness of match between the OC and a SC. As the PLR is a binned statistic, it requires binning the CMDs of the OC and all the SCs that can be generated according to the fundamental parameter ranges defined in Sect. 3.1. This part of the analysis uses Knuth's rule (Knuth 2006, also implemented via the astroML package) as the default binning method, see bottom right plot in Fig. 2. The aforementioned Bayesian blocks method results in noticeably larger cells, which means that if applied here some defining features of the evolutionary sequence could be lost. The inverted logarithmic form of the PLR can be written as

$$LPLR = -\ln PLR = \sum_i m_i - n_i + n_i \ln \frac{n_i}{m_i}, \quad (1)$$

<sup>17</sup> For a detailed description on the *AStéCA* technique for the generation of synthetic clusters from theoretical isochrones see: Paper I, Sect. 2.9.1



**Figure 2.** CMDs for the SMC-L62 cluster. *Top:* CMD of cluster region to the left, where  $n_{memb}$  is the approximate number of true members based on the star density of the OC compared to the star density of the field. The CMD of the surrounding field region is shown to the right. In both cases  $N_{accept}$  is the number of stars that were not rejected due to their large photometric errors. Rejected stars with large errors are shown as green crosses. *Bottom:* cluster region after the DA is applied is shown at the left; MPs vary according to the colorbar at the top right. Dotted horizontal and vertical lines show the binning used to reject low MP stars cell-by-cell, as obtained via the Bayesian blocks method. Stars that are drawn semi-transparent in each cell are those that were rejected, i.e.: not used in the isochrone fitting process that follows.  $N_{fit}$  is the number of stars left after the cell-by-cell rejection. The best fit isochrone is shown in green and red in this CMD, and in the CMD of the best match synthetic cluster to the right (generated from that isochrone), respectively.  $N_{synth}$  is the number of stars in the synthetic cluster, and the dotted lines represent the binning obtained using Knuth’s rule.

where  $m_i$  and  $n_i$  are the number of stars in the  $i$ th CMD cell of the SC and the OC, respectively.<sup>18</sup>

The best match for an OC is given by the SC that minimizes the LPLR. Since we have five free fundamental parameters, a 5-dimensional surface of solutions is determined by all the possible SCs that can be matched to each OC. **ASteCA** applies a genetic algorithm (GA) on this surface to find the SC that best matches the OC. This SC thus determines the fundamental parameter values that will be assigned to the OC under analysis.

The GA uses a set of reasonable default options which indicate how it performs the search for the best OC-SC match. As with the DA, these options had to be modified – mainly extending the depth of the search for the best

<sup>18</sup> If for any given cell we have  $n_i \neq 0$  and  $m_i = 0$ , a very small number is used instead ( $m_i = 1 \times 10^{-10}$ ) to avoid a mathematical inconsistency with the factor  $\ln m_i$ .

match – for  $\sim 30\%$  of the OCs in the sample. This sub-sample is composed of clusters with particularly complicated morphologies, i.e.: very low number of members, high field star contamination, little to no MS present above the maximum magnitude value observed, etc.

Once the GA returns the optimal fundamental parameter values found for an OC, uncertainties are estimated via a standard bootstrap technique. This process takes a significant amount of time to complete, since it involves running the GA several more times on a randomly generated OC with replacement.<sup>19</sup> Ideally, the bootstrap process would be run hundreds and even thousands of times. This is not possible in practice, as it would be prohibitively costly timewise, so we must settle with running it ten times for each OC in our set.

#### 4 ERRORS IN FITTED PARAMETERS

It is well known that a parameter given with no error estimation is meaningless from a physical standpoint (Dolphin 2002; Andrae 2010). This fact notwithstanding, a detailed error treatment is usually ignored in articles that deal with star clusters analysis (Paunzen & Netopil 2006). As explained in Sect. 3.4, **ASteCA** employs a bootstrap method to assign standard deviations for each fitted parameter in our OCs sample.

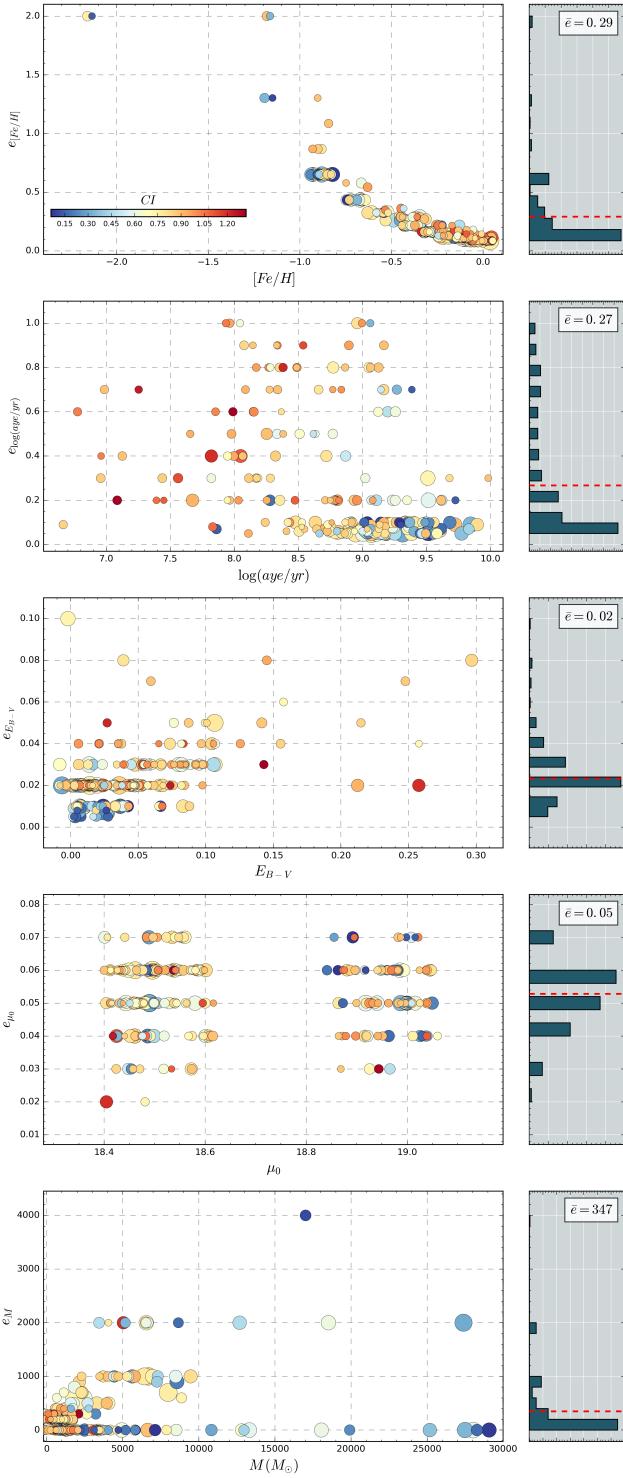
Since the code must simultaneously fit a large number of free parameters – five in this case – within a wide range of allowed values, and using only a 2-dimensional space of observed data (i.e.: the  $CT_1$  magnitudes)<sup>20</sup>, the uncertainties involved will expectedly be somewhat large. It is worth noting that unlike manually set errors, these are statistically valid uncertainty estimates obtained via a bootstrap process. This is an important point to make given that the usual by-eye isochrone fitting method not only disregards known correlations among all OCs parameters, it is also fundamentally incapable of producing a valid error analysis (Naylor & Jeffries 2006). Any uncertainty estimate produced by-eye serves only as a mere approximation, which will often be biased towards smaller figures. The average logarithmic age error for the OCs in the literature is 0.16 dex, in contrast with the almost twice as large average value estimated by **ASteCA** (see below).

In Fig. 3 we show the distribution of the standard deviations versus the five fundamental parameters fitted by the code, for the entire sample (SMC and LMC OCs). Colours follow the  $CI$  obtained for each cluster.

The apparent dependence of the metallicity error with decreasing  $[Fe/H]$  values arises from the fact that **ASteCA** uses  $z$  values – the linear metallicity measure – to find the best OC-SC match. The error in  $z$  is found to be on average

<sup>19</sup> Generating a new OC “with replacement”, means randomly picking stars one by one from the original OC, where any star can be selected more than once. The process stops when the same number of stars as those present in the original OC have been picked.

<sup>20</sup> We plan to upgrade the code to eventually allow more than just two observed magnitudes, therefore extending the 2-dimensional CMD analysis to an  $N$ -dimensional one.



**Figure 3.** *Left:* Distribution of errors versus the five parameters fitted by ASteCA. Colours are associated to the *CI* (see bar in top plot), sizes are proportional to the actual cluster sizes. A small random scatter in the x axis is added for clarity. *Right:* Error histogram. The mean error value for each parameter is shown in the top right corner, and drawn in the plot with a dashed red line.

$e_z \approx 0.003$  for all analysed clusters<sup>21</sup>. This means that, when converting to  $e_{[Fe/H]}$  using the relation

$$e_{[Fe/H]} = e_z / [z * \ln(10)], \quad (2)$$

the  $z$  in the denominator makes  $e_{[Fe/H]}$  grow as it decreases, while  $e_z$  remains more or less constant. For very small  $z$  values (e.g.:  $z = 0.0001$ ), the logarithmic errors can easily surpass  $e_{[Fe/H]} = 2 \text{ dex}$ . In those cases, as seen in Fig. 3, the error is trimmed to 2 dex which is enough to cover the entire metallicity range.

Other than the mathematical dependence of the logarithmic metallicity explained above, there are no visible trends in the arrangement of errors for any of the fitted parameters. This is a desirable feature for any statistical method. If the uncertainties of a parameter varied (increase/decrease) with it, it would indicate that ASteCA was introducing biases in the solutions.

Histograms plotted to the right in Fig. 3 show the distribution of errors and their arithmetic means as a dashed red line. The average metallicity and age errors for the full sample are 0.3 dex and 0.27 dex respectively. For the SMC  $e_{[Fe/H]} < 0.2$  for  $\sim 34\%$  of the calculated uncertainties. The same relation is true for  $\sim 69\%$  of the LMC OCs, which means that OCs in the SMC have a larger dispersion in their assigned metallicity errors. This is due to the lower overall metal content of the SMC OCs, combined with the effect described above by which the logarithmic metallicity error increases for smaller  $[Fe/H]$  values. Approximately 53% of the combined S/LMC sample show  $e_{\log(\text{age}/\text{yr})} < 0.1$ , a very reasonable uncertainty estimate. Error estimates for the remaining parameters are all within acceptable ranges.

Numerically, errors could be lowered applying several different techniques: increase the number of bootstrap runs, increase the number of models evaluated in the GA (“generations”), or reduce the value of the steps in the parameters grid. All of these methods will necessarily extend the time needed to process each cluster, in particular increasing the number of bootstrap runs. Limited computational time available requires a balance between the maximum processing power allocated to the calculations, and the precision one is aiming at. The error values presented here should then be considered a conservative upper limit of the accuracy with which the fundamental parameters of our sample can be obtained by ASteCA.

## 5 COMPARISON WITH PUBLISHED DATA

We compare in Sect. 5.1 the resulting parameter values obtained after running ASteCA on our set of Magellanic OCs, with those taken from the original reference articles (those that used the same Washington photometry used in this work, see Table 1) referred as the “literature”. The parameters age, extinction, and mass are also analysed in Sect 5.2 for a subset of 142 OCs that could be cross-matched with seven articles (where *UBVRI* photometry was used, see Table 2), referred as the “databases”.

<sup>21</sup> Approximately 70% and over 80% of the OCs in the S/LMC have assigned  $e_z = 0.003$  errors

## 5.1 Literature values

The comparison of **AStECA** versus literature values for the parameters metallicity, age, extinction and true distance modulus, is presented in Fig. 4, one parameter per row. The first two plots show the 1:1 identity line for the LMC and the SMC respectively. The rightmost diagram shows a Bland-Altman (BA) plot for our combined set of OCs in the MCs, with the variation in the x axis proposed by Krouwer (Bland & Altman 1986; Krouwer 2008).<sup>22</sup> The BA plot shows differences in the y axis in the sense  $\Delta = (\text{AStECA} - \text{literature})$ , versus values found by the code in the x axis. Errors in  $\Delta$  values are calculated combining the errors for both estimates. The mean of the differences,  $\bar{\Delta}$ , is shown as a dashed line and its standard deviation as a grey band.

An offset is noticeable for the abundance estimates (first row of Fig. 4), where **AStECA** tends to assign values  $\sim 0.22$  dex larger than those found in the literature. On average, this offset is of  $\sim 0.27$  dex for the SMC and  $\sim 0.18$  dex for the LMC. This could be explained by two different processes, in light of our knowledge that the **AStECA**'s OC-SC best fit matching introduces no biases into the solutions. First, the MC's star clusters are generally considered to have low metallicities. For example, the metal content values assumed in the 19 literature articles, are exactly  $[\text{Fe}/\text{H}] = -0.7$  and  $-0.4$  for  $\sim 60\%$  and  $\sim 75\%$  of the S/LMC clusters. The by-eye fit is thus very likely biased towards the assignment of lower abundances. This type of "confirmation bias" in published values has been studied recently by de Grijs et al. (2014), in relation to distance measurements reported for the LMC. Researchers also tend to fit isochrones adjusting it to the lower envelope on an OC's sequence; which also contributes to the selection of isochrones of smaller metallicity.<sup>23</sup> Second, the code writes an obtained parameter value to file, using as many significant figures as those given by the rounded standard deviation. The standard deviation in turn, is rounded following the convention of keeping only its first significant figure. This means for example that a value of  $z = 0.0005 \pm 0.00312$  will be rounded to  $z = 0.001 \pm 0.003$ . Although this effect is an issue only for cases where the uncertainty is larger – or close to – the parameter value, it will nonetheless affect smaller metallicities shifting them towards larger estimates.<sup>24</sup> These two effects can therefore explain the offset found for the abundances assigned to OCs in the literature versus those estimated by **AStECA**. It is important to notice that neither effect is intrinsic to the best likelihood matching method used by the code.

<sup>22</sup> The BA is also called a "difference" plot or more commonly a "Tukey Mean-Difference" plot. In the original BA plot the default x axis displays the mean values between the two methods being compared. The Krouwer variation changes the means for the values of one of those methods, called the "reference". In our case, the reference method is **AStECA** so we use its reported values in the x axis.

<sup>23</sup> The reason is that increasing an isochrone's metal content moves it towards redder (greater colour) values in a CMD, see for example Bressan et al. (2012), Fig 15.

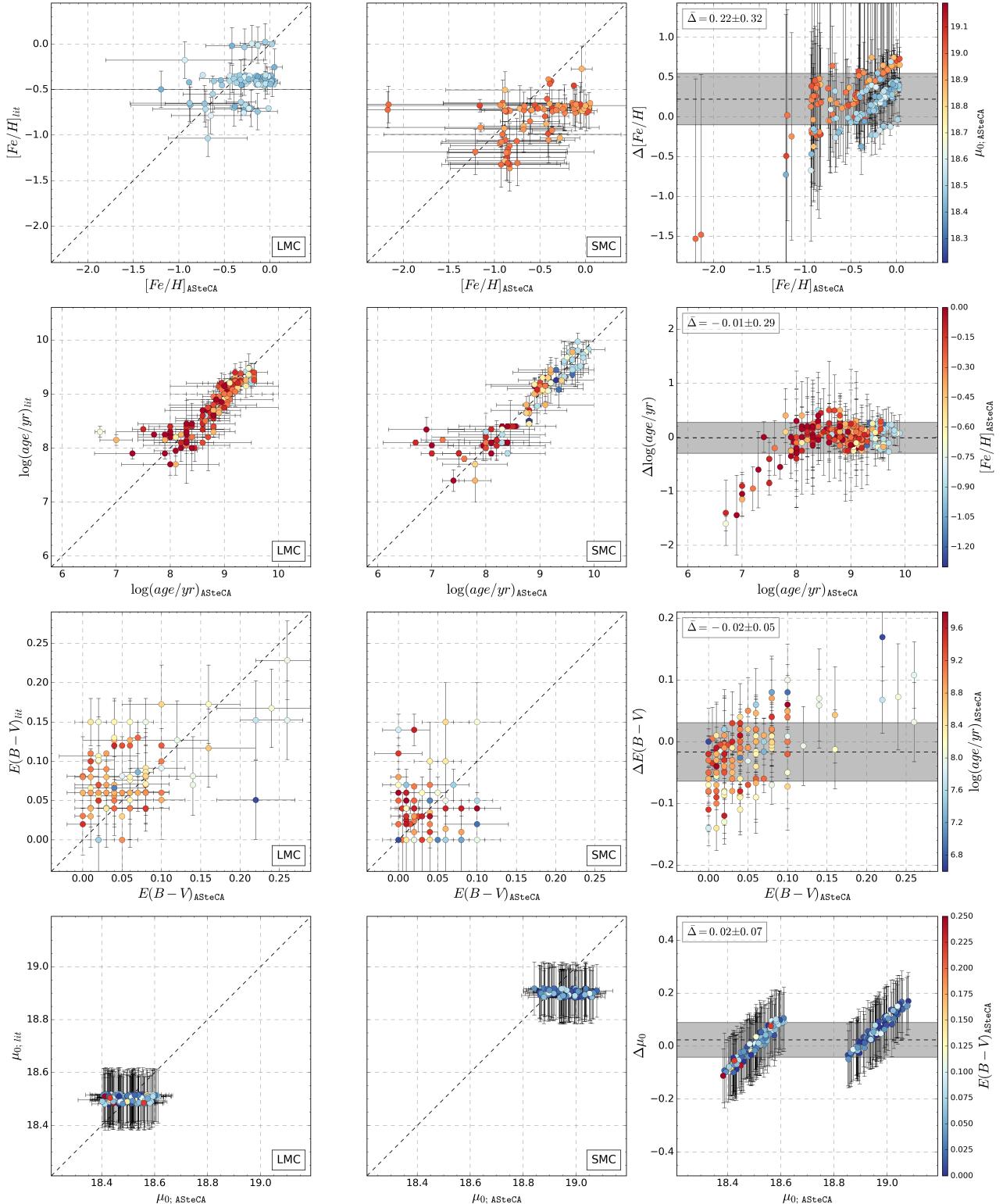
<sup>24</sup> This behaviour will be changed in the next release of **AStECA**. The parameter estimates will be written to file keeping more significant figures, to avoid this issue.

There are two SMC OCs which present an extremely low metallicity: NGC 294 ( $\alpha=0^h53^m06^s$ ,  $\delta= -73^\circ22'49''$ , [J2000.0]) and HW85 ( $\alpha=1^h42^m28^s$ ,  $\delta= -71^\circ16'45''$ , [J2000.0]). These OCs are positioned close to the SMC's centre and towards its periphery, respectively. Both are found to have metallicities of  $[\text{Fe}/\text{H}] \approx -2.2$  dex – with large errors given by **AStECA** – versus their smaller literature values of  $-0.7$  dex. The code assigns  $\log(\text{age}/\text{yr})$  values of  $8.8 \pm 0.06$  for NGC294, and  $9.3 \pm 0.7$  for HW85; while the respective values in the literature are  $8.51 \pm 0.47$  and  $9.26 \pm 0.28$ . NGC 294 has been assigned abundances as low as  $-1.2$  dex (see the integrated spectroscopy study by Dias et al. 2010), although the value obtained by **AStECA** is substantially lower. HW85 is a little studied cluster with very few members – approximately 20 in our photometric data set – and thus prone to present variations in its estimated parameters. No other published estimation of this cluster's metallicity could be found. The sequence of an OC is broadened towards redder colours by the presence of unresolved binaries. We thus investigate the possibility that our selected fraction of unresolved binaries ( $b_f = 0.5$ , see Sect. 3.1) could be influencing the selection of these low metallicities. Both OCs are processed five more times setting the  $b_f$  factor to  $[0., 0.25, 0.5, 0.75, 1.]$ , leaving the metallicity to vary as a free variable, and fixing the remaining parameters to the values found by **AStECA**'s original run. If we average these results we find that, although in both cases the mean  $[\text{Fe}/\text{H}]$  is larger than the original value  $-1.36 \pm 0.08$  dex and  $-1.5 \pm 0.2$  dex for the NGC 294 and HW85, respectively –, these are still very low metallicity estimates. We can thus rule out the binary fraction used as the responsible factor of these low abundances, and conclude that these are indeed OCs with low metallicities.

The general dispersion between literature and **AStECA** values is quantised by the standard deviation of the  $\Delta$  differences in the BA plot. This value is  $\sim 0.32$  dex (top left of BA plot), in close agreement with the mean internal uncertainty found in Sect. 4 for this parameter. Mean metallicity estimates for the MCs using **AStECA** values are  $[\text{Fe}/\text{H}]_{SMC} \approx -0.52 \pm 0.44$  dex, and  $[\text{Fe}/\text{H}]_{LMC} \approx -0.26 \pm 0.24$  dex. These are considerably more metal rich and disperse averages than the ones obtained using literature values:  $[\text{Fe}/\text{H}]_{SMC} \approx -0.78 \pm 0.23$  dex, and  $[\text{Fe}/\text{H}]_{LMC} \approx -0.42 \pm 0.16$  dex

The second row in Fig. 4 shows the age distribution for both Clouds. The overall agreement is very good, with larger deviations from the identity line appearing for younger ages. There are ten OCs for which **AStECA** assigned  $\log(\text{age}/\text{yr})$  values that differ more than 0.5 dex from their literature values. These are referred to as "outliers", and are treated separately in Sect. 5.1.1.

In Palma et al. (2015) an offset is found for star clusters in the LMC fitted with the Girardi et al. (2002, G02) and Bressan et al. (2012, B12) set of isochrones, where OCs fitted with the latter presented consistently larger values. Here we confirm this trend, for ages assigned to our set of OCs for both the SMC and the LMC. While **AStECA** uses the B12 set of theoretical isochrones, eleven out of the nineteen articles the literature set (see Table 1) used G02 isochrones. The Lejeune & Schaerer (2001, L01) and Marigo et al. (2008) isochrone sets were used by four and two literature articles,



**Figure 4.** Left column: parameters comparison for the LMC. Centre column: idem for the SMC. Right column: BA plot with differences in the sense ASteCA minus literature, for the combined S/LMC sample. For clarity, a small random scatter is added to both axes for the metallicity and distance modulus plots. Mean and standard deviation are shown as a dashed line and a grey band, respectively; its values are displayed in the top left of the plot. Colours following the coding shown in the bar at the right of the figure, for each row. Piatti (2011b) OCs which contain only age information are plotted with  $E_{B-V}=0$  colour coding.

respectively; with the remaining two articles applying the  $\delta T_1$  index to derive ages. In Piatti et al. (2003b,a, 2007a) no significant differences were found between the L01 and the G02 sets, so we assume the same offset will be present in the former set. Excluding the outlier OCs described above, the  $\Delta \log(\text{age}/\text{yr})$  offset for the S/LMC are  $\sim 0.02$  dex and  $\sim 0.05$  dex, respectively.

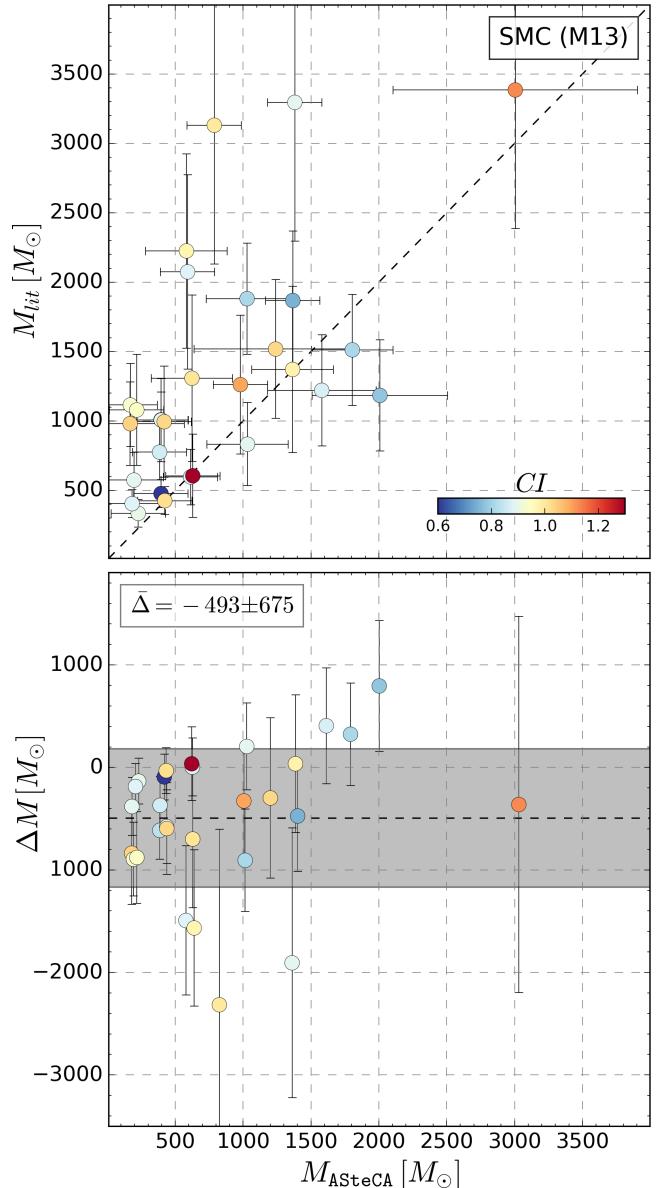
The mean value for the  $\Delta \log(\text{age}/\text{yr})$  parameter in the BA plot, including all S/LMC OCs, is  $\sim 0.01$  dex. This points to an excellent agreement among literature and AStECA values for the  $\log(\text{age}/\text{yr})$ . If we exclude the outliers, the mean of the differences increases by a small amount to  $\sim 0.04$  dex. Similarly to what was found for the metallicity, the dispersion between literature and AStECA values is almost exactly the internal uncertainty found for errors assigned by the code, i.e.:  $\sim 0.3$  dex.

The reddening distribution is relatively small for OCs in both MCs. Maximum  $E_{B-V}$  values are  $\sim 0.15$  mag and  $\sim 0.3$  mag for the S/LMC respectively, as shown in the AStECA versus literature identity plots. The  $\Delta$  differences are well balanced with a mean of -0.02 and a standard deviation of 0.05 mag, slightly larger than the 0.02 mag average uncertainty found for the errors assigned by the code. We obtain average  $E_{B-V}$  extinctions for the S/LMC of  $0.03 \pm 0.03$  and  $0.05 \pm 0.05$ , noticeably lower estimates – approximately a third – than those used for example in the Hunter et al. (2003) study of the MCs.

The true distance moduli ( $\mu_o$ ) found by AStECA in both MCs show a clear displacement from literature values. This is expected, as the distance to OCs in the latter is always assumed to be a fixed constant equivalent to the distance to the centre of the corresponding galaxy. The distribution of  $\mu_o$  values found by the code covers the entire range allowed in Sect. 3.1. Distances obtained by AStECA thus vary up to  $\sim \pm 0.1$  mag from the default fixed  $\mu_o$  values used in the literature. It is worth noting that this variation appears to have no substantial effect on any of the remaining parameters, an effect that could in principle be expected due to the known correlations between all fundamental parameters (see Paper I, Sect. 3.1.4). This reinforces the idea that using a fixed value for the distance modulus, as done in the literature, is a valid way of reducing the number of free variables at no extra cost.

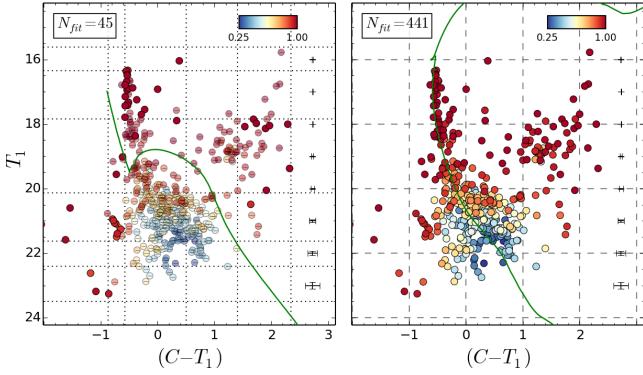
Since masses are only assigned in Maia et al. (2013, M13) for its sample of 29 SMC OCs, the comparison with AStECA is presented separately in Fig. 5. The identity plot (top) shows a trend for AStECA masses to be smaller than those from M13. This is particularly true for two of the OCs with the largest mass estimates – and the largest assigned errors – in M13: H86-97 ( $3300 \pm 1300 M_\odot$ ), and H86-87 ( $3100 \pm 1700 M_\odot$ ). The BA plot (bottom) shows that, on average, M13 masses are  $\sim 500 \pm 700 M_\odot$  larger. To explain these differences, we need to compare the way total cluster masses are obtained by M13 and by AStECA in this work.

In M13 masses were determined using two methods, both based on estimating a mass function via the  $T_1$  luminosity function (LF) of an OC. In both cases a field star cleaning process was applied. The first one employs a CMD decontamination procedure (described in Maia et al. 2010),



**Figure 5.** Top: Mass comparison for AStECA versus literature values. Bottom: BA plot with mean and standard deviation of the differences shown as a dashed horizontal line, and a grey band respectively. Values are displayed in the top left corner.

and the second one cleans the cluster region's LF by subtracting it a field star LF. The results obtained with these two methods are averaged to generate the final mass values. Although any reasonable field star cleaning algorithm should remove many or most – depending on the complexity of the OC's CMD – of the foreground/background stars in a cluster region, some field stars are bound to remain. For heavily contaminated OCs this effect will be determinant in shaping the “cleaned” CMD sequence, as true cluster members will be very difficult to disentangle from contaminating field stars. The set of 29 OCs in the M13 sample are indeed heavily affected by field star contamination. This can be seen in Fig. 5, where the colour assigned to each OC corresponds to its  $CI$ . The minimum value is  $CI=0.55$ , which



**Figure 6.** Left: Best fit isochrone for B48 found by *ASteCA* when a cell-by-cell removal is applied on its sequence, following the Bayesian MP assignation (removed stars are drawn semi-transparent). The estimated age and total mass are  $\log(\text{age}/\text{yr})=6.2 \pm 0.6$ , and  $M=400 \pm 200 M_{\odot}$ . Right: Best fit isochrone found when no removal of stars is performed, and the full cluster region is used in the search for the best synthetic cluster match. The estimated age and total mass are  $\log(\text{age}/\text{yr})=7.5 \pm 0.3$ , and  $M=3000 \pm 900 M_{\odot}$ .

means all OCs in the set are expected to contain on average more field stars within the analysed cluster region, than true members. The presence of a large number of contaminating field stars not only makes the job much harder for the DA, it also necessarily implies that the overall LF will be overestimated, therefore leading – in the case of M13 – to an overestimation of the total mass. In contrast, *ASteCA* assigns OC masses taking their values directly from the best match synthetic clusters (SCs). Field star contamination will thus have a much lower influence on the code’s mass estimate, limited just to how effective the DA is in cleaning the cluster region.

The case of B48 is worth mentioning, as it is the OC with the largest total mass given in M13 ( $3400 \pm 1600 M_{\odot}$ ). After being cleaned of possible field stars by the DA – see Sect. 3.3 – low mass stars are entirely removed and B48 is left only with its upper sequence ( $T_1 < 18.4$  mag). This happens both in the literature and in the analysis done by *ASteCA*, see left CMD in Fig. 6, and Fig. C8 in M13. The likelihood defined in Eq. 1 sees then no statistical benefit in matching the OC with a SC of a similar age and mass, which will contain a large number of low mass stars. This leads the GA to select as best match SCs of considerably younger ages ( $\log(\text{age}/\text{yr}) < 7.0$ ) than that assigned in M13 ( $7.9 \pm 0.05$ ), and with much lower mass estimates (see caption of Fig. 6). A good match both in age and in mass could be found by the code, only if the DA was applied with no cell-by-cell removal of low MP stars as shown in the right CMD of Fig. 6. This means that all stars within the cluster region – including field stars – are used in the SC matching process, which inevitably questions the reliability of the total mass estimate. Dealing with this statistical effect is not straightforward and will probably require an extra layer of modelling added to the SC generation algorithm. As discussed in Sect. 5.1.1 this effect also plays an important role in the significant age differences found between *ASteCA* and the literature, for a handful of OCs.

**Table 4.** OCs with large differences in their assigned literature ages versus the values found by the code (“outliers”). Equatorial coordinates are expressed in degrees for the J2000.0 epoch. Ages are given as  $\log(\text{age}/\text{yr})$  for literature (L) and *ASteCA* (A). The difference between both estimates (L-A) is given in the last column as  $\Delta$ .

Cluster	$\alpha(^{\circ})$	$\delta(^{\circ})$	L	A	$\Delta$
L-KMHK975	82.49583	-67.87889	8.30	6.70	1.60
L-SL579	83.55417	-67.85639	8.15	7.00	1.15
L-BSDL631	76.64167	-68.42722	8.35	7.50	0.85
L-KMHK979	82.41250	-70.98389	7.90	7.30	0.60
L-H88-316	85.41250	-69.22944	8.25	7.70	0.55
S-L35	12.00417	-73.48611	8.34	6.90	1.44
S-H86-188	15.05833	-72.45833	8.10	6.70	1.40
S-L39	12.32500	-73.37167	8.05	7.00	1.05
S-B134	17.25417	-73.20667	8.15	7.20	0.95
S-K47	15.79583	-72.27361	7.90	7.00	0.90

### 5.1.1 Outliers

Ten of the analysed OCs in this work – approximately 4% of the total 239 clusters present in the set – show a difference in age with literature values of  $\Delta \log(\text{age}/\text{yr}) > 0.5$ . Such a large age difference translates into two very dissimilar isochrones fitted to the same observed coeval star sequence, which makes this sub-sample of OCs stand out from the rest. For these “outliers” no configuration of the DA plus the employed binning methods could be found, that resulted in SC matches with age values more similar to those found in the literature. In Appendix A, literature and *ASteCA* isochrone matches are shown for these ten outliers.

Five of these OCs belong to the LMC and the remaining five to the SMC, as shown in Table 4. All the OCs in the outliers sample had smaller ages assigned by the code, compared to their literature values. These differences go from 0.55 dex up to 1.6 dex in the most extreme case of the LMC cluster KMHK975. Fig. A1 (see CMD diagrams a) shows the best matches for this OC, where the reason for the different isochrones selected can be clearly seen. While the by-eye isochrone fit done in the literature aligned the brighter part of the OC’s sequence with the turn off point of a  $\sim 200$  Myr isochrone, *ASteCA* decided instead that this was the top portion of a much younger cluster ( $\sim 5$  Myr) with no discernible turn off. For almost all of the outliers, the same process can be identified as the main cause responsible for the observed age differences. The statistical mismatch due to the removal of low mass stars by the DA – discussed in Sect. 5.1 – can also be seen to affect some of the fits here. In particular the SMC OCs SL579 and H86-188 show signs of this effect in the best match SCs selected by *ASteCA* (see isochrone fits in Fig. A1, CMD diagrams b and h).

These age estimates could be improved – in the sense that they could be brought closer to literature values – if a more restrictive age range was used (e.g.: a minimum value of  $\log(\text{age}/\text{yr}) = 7.5$  instead of 6 dex as used in this work, see Table 3). Lacking external evidence to substantiate this a priori restriction, we choose to keep the values obtained by *ASteCA*, with this section acting as a cautionary note.

## 5.2 Databases values

In addition to the analysis performed in Sect. 5.1, we compare our results with those taken from seven articles – the “databases” or DBs – where a different photometric system was used; see Table 2. We can further separate these seven DBs into two groups: those where the standard by-eye isochrone fitting method was applied – P99, P00, C06, and G10 – and those where integrated photometry was employed to derive the OCs fundamental parameters – H03, R05, and P12. A total of 142 individual OCs from our sample could be cross-matched. Where names were not available – P00, P99, and R05 – we employed a 20 arcsec radius to find matches, based on the equatorial coordinates of the OCs in each DB.

The comparison of **ASteCA** ages with those from the four isochrone fit DBs is shown in Fig. 7, left and centre plot. P99 and P00 analyse SMC and LMC clusters respectively, using Bertelli et al. (1994) isochrones and fixed metallicities for the S/LMC of  $z=0.004, 0.008$ . While P99 derives individual reddening estimates based on red clump stars, P00 use extinction values determined for 84 lines-of-sight in the Udalski et al. (1999) LMC Cepheids study. The mean S/LMC extinction in P99 and P00 are  $E_{B-V} \approx 0.08 \pm 0.02$  and  $E_{B-V} \approx 0.14 \pm 0.02$ , both larger estimates than those found in this work ( $E_{B-V} \approx 0.03 \pm 0.03$  and  $E_{B-V} \approx 0.05 \pm 0.05$  for the S/LMC, see Sect. 5.1). Both studies attempt to eliminate field star contamination following the statistical procedure presented in Mateo & Hodge (1986). The distance moduli for the S/LMC of  $\mu_0 = 18.65, 18.24$  mag employed in these DBs, are approximately  $\sim 0.25$  mag smaller than the canonical distances assumed for each Cloud, which has a direct impact on their obtained ages. In de Grijs & Anders (2006) the authors estimate that had P00 used a value of  $\mu_0 = 18.5$  mag instead for the LMC, their ages would be  $\sim 0.2\text{--}0.4$  dex younger; a similar conclusion is reached by Baumgardt et al. (2013).<sup>25</sup> The same reasoning can be applied to the P99 age estimates. P99 and P00 ages are displaced on average from **ASteCA** values (in the sense **ASteCA** minus DB) by  $-0.13 \pm 0.6$  dex and  $0.37 \pm 0.5$  dex respectively, as seen in Fig. 7 left plot. In the case of P99, the distance modulus correction mentioned above would bring the age values to an overall agreement with those obtained by **ASteCA**, although with a large scatter around the identity line. P00 age values on the other hand, would end up  $\sim 0.7$  dex below the code’s age estimates after such a correction. Such a large deviation is most likely due to the overestimated extinction values used by P00, as will be shown below.

C06 studied 311 SMC clusters via isochrone fitting applying two methods: visual inspection and a Monte Carlo based  $\chi^2$  minimization. The authors also employ a decontamination algorithm to remove contaminating field stars, making this the article that more closely resembles this present work. Distance modulus is assumed to be  $\mu_0 = 18.9$  mag. Reddening and metallicity values of  $E_{B-V} = 0.08$  mag and  $z = 0.008$  dex are used, adjusted when necessary to improve the fit. It is worth noting that the  $[\text{Fe}/\text{H}] = -0.4$

<sup>25</sup> Notice that in Baumgardt et al. (2013) the authors correct the age bias that arises in P00 due to the small distance modulus used, increasing P00 age estimates by 0.2 dex. This is incorrect, ages should have been decreased by that amount.

dex abundance employed in C06 is closer to the average  $[\text{Fe}/\text{H}] = -0.52 \pm 0.44$  dex value found for the SMC by **ASteCA**, than the canonical value of  $[\text{Fe}/\text{H}] = -0.7$  used by default in most works. This is the DB – out of the four isochrone-fit and the three integrated photometry DBs – that best matches **ASteCA** age values, with a mean deviation from the identity line of  $0.02 \pm 0.58$  dex.

G10 analysed over 1500 OCs with ages  $< 1\text{Gyr}$  in both Clouds via by-eye isochrone fitting. They assumed distance moduli of (18.9, 18.5) mag, and metallicities of (0.004, 0.008), for the S/LMC respectively. Extinction was adjusted also by-eye on a case-by-case basis. The G10 database presents a systematic bias where smaller logarithmic ages are assigned compared to our values, with an approximate deviation of  $\Delta \log(\text{age}/\text{yr}) \approx 0.23 \pm 0.46$ . This is consistent with the results found in Choudhury et al. (2015) (see Fig. 5), and later confirmed in Piatti et al. (2015a,b). G10 does not apply any decontamination method, by which probable field stars are removed from the cluster’s CMD prior to its analysis. Instead, they plot over the cluster region a sample of surrounding field stars 0.1 arcmin away from the cluster’s radius, taken from a 0.5 arcmin concentric annulus. The lack of a proper statistical removal of contaminating foreground/background stars can cause the isochrone fit to be skewed by their presence.

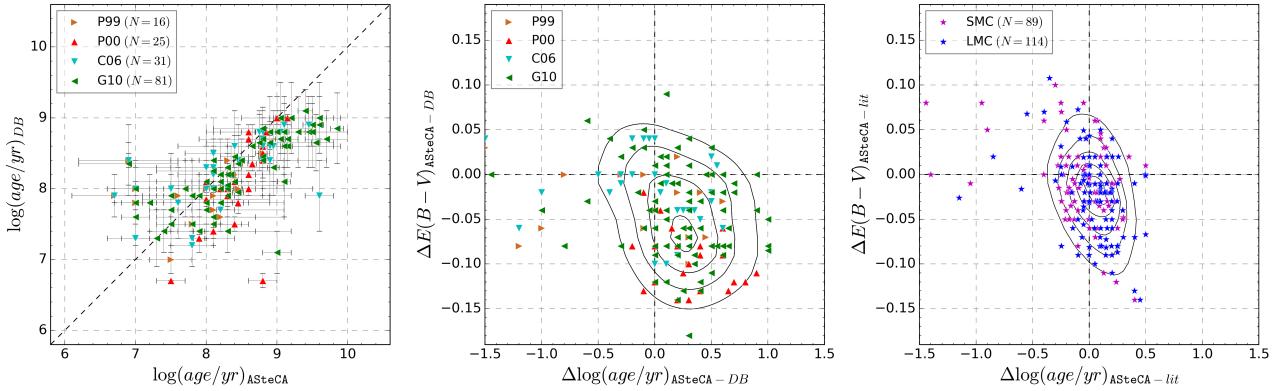
As seen in the centre plot of Fig. 7, these four DBs taken as a single group present a clear age-extinction bias, when compared to **ASteCA** values. This known degeneracy was found in Paper I to be the one with the largest correlation value (see Paper I, Table 3), meaning it is the process most likely to affect isochrone fit studies. The maximum density in this “delta” plot is located around  $\Delta E_{B-V} \approx -0.07$  mag and  $\Delta \log(\text{age}/\text{yr}) \approx 0.3$  dex. This trend is most obvious for P00 where a rather large average extinction value was employed, compared to the mean value found by **ASteCA** as stated above. For comparison purposes we show in the right plot of Fig. 7 the same delta plot, this time generated subtracting literature values from **ASteCA** age estimations. It can be clearly seen that afore mentioned bias is basically non-existent here, pointing to a consistent overall assignation of extinction and ages by the code.

In Appendix B we show the CMDs of each cross-matched OC for this four DBs (153 in total). The isochrone plotted is the best fit proposed by the corresponding DB, and is compared to the best match found by **ASteCA**. Those OCs that present the largest age discrepancies between **ASteCA** and DBs values, are those where the same effect mentioned in Sect. 5.1.1 takes place. A good example of this is SMC-L39, as seen in Figs. B9 and B18 for C06 and G10 respectively.

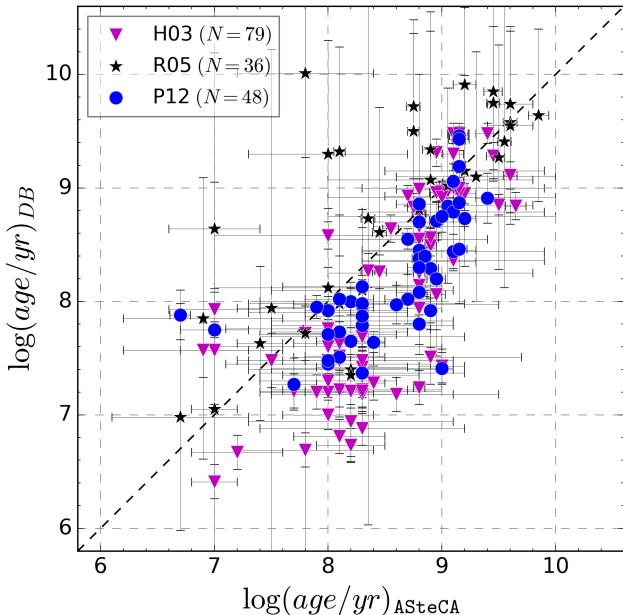
Our age and mass estimates are also compared with three DBs – H03, R05, and P12 – which used integrated photometry to obtain these parameters (see Table 2). Only two of these, H03 and P12, obtained total mass values for the OCs in their sample.

H03 studied approximately 1000 OCs in both MCs – 748 in the LMC and 196 in the SMC<sup>26</sup> – via *UBVR* inte-

<sup>26</sup> To these numbers, 140 and 76 “questionable” (according to H03) S/LMC clusters respectively, can be added to their sample.



**Figure 7.** Left: age comparison for DBs that used the isochrone fit method, versus ASteca, where  $N$  is the number of OCs cross-matched in each DB. Centre: “delta” plot, showing the differences between extinction and age, in the sense ASteca minus DB. Curves represent regions of iso-densities after fitting a 2-dimensional Gaussian Kernel. Right: same as previous plot, now showing ASteca minus literature values for both Clouds.



**Figure 8.** Age comparison for DBs that used the integrated photometry method, versus ASteca, where  $N$  is the number of OCs cross-matched in each DB.

grated photometry. Ages were assigned based on the Starburst99 model (Leitherer et al. 1999), assuming metallicities, distance moduli, and average  $E_{B-V}$  extinction values of (0.004, 0.008), (18.94, 18.48) mag, and (0.09, 0.13) mag, for the S/LMC, respectively. The masses for each OC were derived through their absolute magnitudes  $M_V$  and the mass-luminosity relation, assuming -14.55 mag to be the  $M_V$  of a 10 Myr old  $10^6 M_\odot$  cluster with  $z = 0.008$ . This article represents, as far as we are aware, the largest published database of MCs cluster masses to date.

R05 used two models – GALEV (Anders & Fritze-v. Alvensleben 2003) and Starburst99 – combined with three metallicities values – (0.004, 0.008) and (0.001, 0.004, 0.008), used in each model respectively – to obtain ages for 195 SMC clusters. This results in five age estimates for each cluster.

Individual reddening values are obtained in the same manner as done in the Harris & Zaritsky (2004) study, assigning extinctions according to fixed age ranges. We averaged all extinction-corrected age values for each matched cluster, and assigned an error equal to the midpoint between the lowest and highest error bound among all reported ages in the article.

P12 uses the same dataset from H03 to analyse 920 LMC clusters through their MASSCLEAN<sub>colors</sub> and MASSCLEAN<sub>age</sub> packages (Popescu & Hanson 2010a,b). Ages and masses from duplicated entries in the P12 sample are averaged.

As seen in the identity plot in Fig. 8, H03 visibly underestimates ages for younger clusters. In de Grijs & Anders (2006, see Fig. 1) this effect was also registered, which the authors assigned to the photometry conversion done in H03. The average dispersion between H03 and ASteca age values is  $0.44 \pm 0.56$  dex. The same happens for P12 ages, albeit with a smaller mean dispersion of the logarithmic age of approximately  $0.35 \pm 0.44$  dex. In P12 their own age estimates are compared with those taken from H03. They find a clear systematic difference with H03 (see P12, Fig. 8), where MASSCLEAN ages are larger than H03 estimates, for OCs with  $\log(\text{age}/\text{yr}) < 8$ . In our case, most of the OCs cross-matched with P12 are older than 8 dex. Nevertheless the same trend is confirmed, with P12 age values located below the identity line in Fig 8 – i.e.: younger ages compared to ASteca – but still closer than those from H03. This bias towards smaller age estimates by P12 is consistent to what was found in Choudhury et al. (2015). Contrary to the what was found for H03 and P12, the R05 study slightly underestimates ages compared to our results, with a mean dispersion of  $-0.25 \pm 0.63$  dex around the identity line. The standard deviation is the largest for the three integrated photometry DBs. In R05 the authors mention the lack of precision in their age measurements, due to the use of integrated colours, and the lack of constraints for the metallicity.

Expectedly, the four isochrone fit studies analysed previously show a more balanced distribution of ages around the 1:1 relation, in contrast with the DBs that employ integrated photometry. Ages taken from integrated photometry studies are known to be less accurate, and

should be taken as a rather coarse approximation to the true values. As can be seen in P12, integrated colours present large scatters for all age values, leading inevitably to degeneracies in the final solutions. The added noise by contaminating field stars is also a key issue, as it is very difficult to remove properly from integrated photometry data. A single overly bright field star can also substantially modify the observed cluster's luminosity, leading to incorrect estimates of its parameters (Baumgardt et al. 2013; Piatti 2014). A detailed analysis of some of the issues encountered by integrated photometry studies, and the accuracy of their results, is presented in Anders et al. (2013).

Masses are obtained in H03 and P12 via integrated photometry analysis. Baumgardt et al. (2013) also derives masses, but as stated in that article, their results are in good agreement with those from P12, so we do not add this database to our work.

There are 127 OCs in the combined H03 and P12 cross-matched samples. In Fig. 9 we show DBs masses for all cross-matched OCs, versus their relative differences<sup>27</sup> defined as:

$$\overline{\Delta M_r} = (M_{\text{ASteCA}} - M_{\text{DB}})/(M_{\text{ASteCA}} + M_{\text{DB}}) \quad (3)$$

$$= \Delta M/(M_{\text{ASteCA}} + M_{\text{DB}}),$$

where  $\text{DB}$  represents either H03 or P12, and cross-matched OCs are divided into three regions according to the masses given in either database. Sizes are scaled with each OC's radius in parsecs, and colours follow the difference in assigned ages  $\Delta \log(\text{age}/\text{yr})$ , in the sense ASteCA minus DB values, as given in the colorbar in the right plot. The minimum CI value for OCs in this set is  $\sim 0.6$ , meaning that they comprise a highly contaminated sample of clusters. The grey band in each plot is the mean of the  $\overline{\Delta M_r}$  relative differences in the assigned masses for these OCs – combining both DBs –, extended one standard deviation upwards and downwards.

The mean relative differences in the  $M_{\text{DBs}} \leq 5000 M_\odot$  low-mass region – left plot in Fig. 9 – is very close to zero, with a standard deviation of almost  $\sim 0.5$ , equivalent to a multiplicative factor of 3 between ASteCA and DBs masses. The mean value of  $\Delta M$ , the standard mass difference, is  $\sim 40 M_\odot$  which points to a very reasonable scatter around the identity line. The standard deviation of  $\Delta M$  is  $\sim 1700 M_\odot$ , a somewhat large number considering the maximum  $5000 M_\odot$  limit for the OCs analysed. This is nevertheless expected for a set composed of clusters with such small masses. As stated for example in P12 and Baumgardt et al. (2013), low mass OCs – i.e., those with  $M \lesssim 5000$  or  $10000 M_\odot$  – will tend to have their estimated integrated photometry masses largely dominated by stochastic processes.

A surprising systematic trend arises when we take a closer look at the mass values assigned by the DBs. It can be evidently noticed that the larger the mass estimated by either DB, the larger the relative difference with ASteCA's derived value. The left plot in Fig. 9 already shows this trend,

<sup>27</sup> For clarity we employ relative differences here instead of differences as in previous BA plots, since the masses span a large range of values. A  $5000 M_\odot$  discrepancy will not carry the same weight if it happens for an OC with estimated masses below  $10000 M_\odot$ , than if it happens for an OC with estimated masses above  $1 \times 10^5 M_\odot$ .

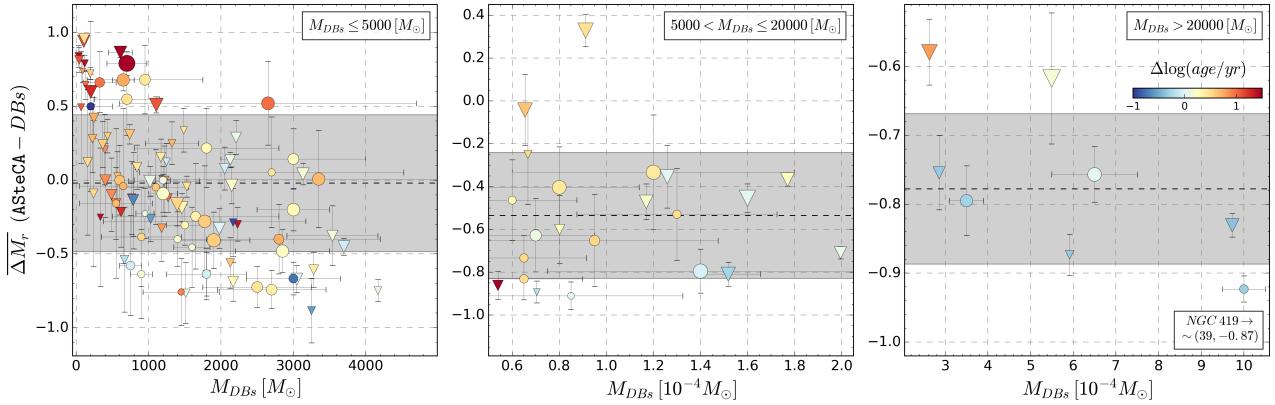
albeit disguised by some OCs with positive  $\overline{\Delta M_r}$  differences – i.e., with larger masses given by the code – beyond  $2000 M_\odot$ . If we look at the centre and right plots in Fig. 9, where DBs mass estimates  $> 5000 M_\odot$  are shown, the trend becomes unmistakable. In these plots we see OCs with DB mass estimates up to  $1 \times 10^5 M_\odot$  with  $\Delta M$  differences of almost  $9 \times 10^4 M_\odot$ . The most discrepant case of SMC cluster NGC419 had to be left out of the right plot in Fig. 9 for clarity, as it is given a mass of  $3.9 \times 10^5 M_\odot$  by H03 and only  $2.8 \times 10^4 M_\odot$  by the code. This OC is one of the nine clusters – five in the LMC and four in the SMC – identified by H03 as “extreme” clusters due to their low absolute magnitude values (see end of Sect. 4 in H03).

The mean and standard deviation for  $\overline{\Delta M_r}$  in the medium and large-mass regions shown in the centre and right pots of Fig. 9, is  $\sim -0.5 \pm 0.3$ , and  $\sim -0.8 \pm 0.1$  respectively. A mass value estimated by the DBs is thus on average between 4 and 9 times larger than the one calculated by ASteCA. The smaller standard deviations for  $\overline{\Delta M_r}$  in both these regions, would seem imply that this is not just a stochastic effect but rather a systematic one.

Considering that integrated photometry studies assign more credibility to higher mass estimates – as they will tend to be less influenced by stochastic fluctuations – this constitutes indeed an unexpected result. In Table 5 we show the five OCs with the largest ASteCA-DBs mass discrepancies,  $\Delta M > 2 \times 10^4 M_\odot$ , ordered locating the ones with the larger DB masses on top. CMD plots for each one of these OCs are presented in Appendix C, along with the best match synthetic clusters generated by the code. For the most extreme cases, both DBs assign total masses that are over an order of magnitude greater than the value found by ASteCA. Ages are largely in good agreement across the two DBs and this work, for these five OCs. As mentioned previously for the SMC cluster NGC419, H03 assigns a mass  $3.6 \times 10^5 M_\odot$  – equivalently 14 times – larger than the mass obtained by the code. Similarly, P12 gives LMC cluster NGC1917 a mass 25 times larger –  $\Delta M = 9.6 \times 10^4 M_\odot$  – than the one derived by ASteCA.

A possible explanation for this large divergence in the calculated masses, is the presence of contaminating field stars. To test this hypothesis we re-processed with ASteCA all OCs in Table 5, this time with no previous decontamination process applied. This means that all stars within the cluster region, including field stars in the same line of sight, will be employed in the obtention of the best synthetic cluster match. The upper limit for the total mass is set to  $5 \times 10^5 M_\odot$ , to avoid biasing the results by setting a low total mass value. For LMC cluster SL244 the total mass value obtained this way is  $3 \times 10^4 M_\odot$ , meaning its average DB mass is recovered when no field stars cleaning is performed. The best synthetic cluster match for SMC cluster L27 results in an even higher total mass of  $1 \times 10^5 M_\odot$ . These two cases clearly highlight the importance of a proper field star decontamination, before the method to derive an OC's mass is applied.

For the remaining three OCs – S-NGC419, L-NGC1917, and L-NGC1751; the most massive according to H03 and P12 – the masses derived by ASteCA using the contaminated cluster region fall short from the values assigned by the DBs. The closest match is found for LMC clusters NGC1751 and NGC1917, for which a mass of  $2 \times 10^4 M_\odot$  is estimated by the code while their average DB mass is  $\sim 8 \times 10^4 M_\odot$ . In the case



**Figure 9.** *Left:* BA mass plot, showing the differences between estimated masses in the H03 and P12 DBs and the code, in the sense AStECA minus DB; symbols as in Fig. 8. Only DB masses  $\leq 5000 M_{\odot}$  are shown here. Colours are assigned according to the contamination index (CI) of each OC (colorbar is shown in the right plot), sizes are proportional to the actual sizes in parsecs. The grey band is the mean  $\pm 1\sigma$  for the  $\Delta M$  values (notice the axis is scaled by  $10^{-4} M_{\odot}$ ). *Centre:* same as previous plot, now showing DB mass values in the range  $5000 - 20000 M_{\odot}$ . *Right:* same as previous plot, for DB mass values  $> 20000 M_{\odot}$ .

**Table 5.** OCs with large differences ( $\Delta M > 20000 [M_{\odot}]$ ) in their assigned AStECA masses, versus the values found in the DBs. Equatorial coordinates are expressed in degrees for the J2000.0 epoch. Ages are given as  $\log(\text{age}/\text{yr})$ .

Cluster	$\alpha(^{\circ})$	$\delta(^{\circ})$	H03 $_{age}$	P12 $_{age}$	AStECA $_{age}$	H03 $M [M_{\odot}]$	P12 $M [M_{\odot}]$	AStECA $M [M_{\odot}]$
S-NGC419	17.07917	-72.88417	$9.31 \pm 0.12$	—	$8.95 \pm 0.05$	$\sim 3.9 \times 10^5$	—	$2.8 \pm 0.3 \times 10^4$
L-NGC1917	79.7583	-69.001	$9.48 \pm 0.09$	$9.46 \pm 0.08$	$9.15 \pm 0.08$	$\sim 5.9 \times 10^4$	$1 \pm 0.05 \times 10^5$	$4 \pm 1 \times 10^3$
L-NGC1751	73.550	-69.80694	$9.48 \pm 0.09$	$9.06 \pm 0.01$	$9.1 \pm 0.05$	$\sim 9.7 \times 10^4$	$6.5 \pm 1 \times 10^4$	$9 \pm 1 \times 10^3$
S-L27	10.35	-72.89083	$9.28 \pm 0.12$	—	$9.45 \pm 0.06$	$\sim 5.5 \times 10^4$	—	$1.3 \pm 0.4 \times 10^4$
L-SL244	76.90417	-68.54194	$9.48 \pm 0.09$	$9.43 \pm 0.01$	$9.15 \pm 0.09$	$\sim 2.9 \times 10^4$	$3.5 \pm 0.4 \times 10^4$	$4 \pm 1 \times 10^3$

of MC cluster NGC419, the recovered AStECA mass using the entirety of stars in its observed field is  $5.5 \times 10^4 M_{\odot}$ , still seven times smaller than the value given by H03. The radius used for NGC 419 in our case is larger than the one employed in H03 by more than 20% ( $\sim 85''$  versus  $\sim 70''$ ), so we can be sure that this effect is not related to a lack of stars included within the cluster region.

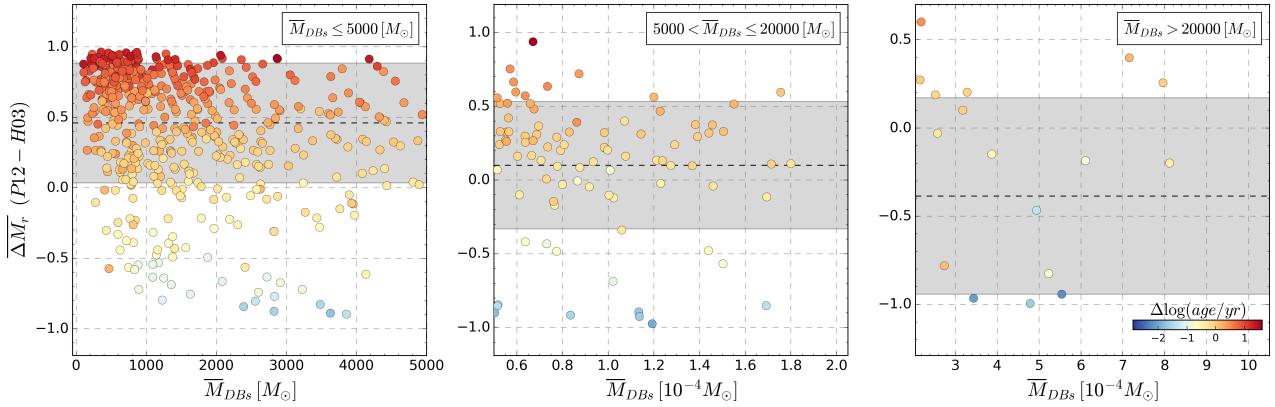
These three OCs share a common extra attribute beyond being the ones with the largest assigned masses in both DBs, and showing the largest discrepancies with AStECA mass values: they are all identified as clusters presenting the controversial “dual red clump” (dRC) structure (Girardi et al. 2009). This feature was predicted in Girardi et al. (1998) as a grouping of stars with enough mass to ignite helium, while avoiding e<sup>-</sup>-degeneracy. The quantitative effect of the dRC structure, on the integrated magnitude of a cluster, was tested on synthetic MASSCLEAN clusters of varying masses. We found that adding a secondary RC composed of about  $\sim 15\%$  of the stars present in the main RC (the approximate value found for NGC 419 in Girardi et al. 2009), has a very small effect on the synthetic cluster’s integrated V magnitude, as well as its (B-V) colour.<sup>28</sup> For a 1 Gyr synthetic cluster of  $10000 M_{\odot}$ , adding stars to form the dRC – located  $\sim 0.3$  mag below the RC, and  $\sim 0.04$  mag towards the bluer region of the (B-V) vs. V CMD – results in the integrated V band (and the integrated (B-V)

colour) increasing only a few hundredths of a magnitude. At the same time, just doubling the mass of the synthetic cluster, i.e.  $20000 M_{\odot}$ , increases the integrated V band value almost a full magnitude. This difference is large enough to assume that the H03 and P12 models will not mistakenly assign large masses, based on such a small variance in integrated photometry as that produced by a dRC. The excess brightness generated by stars in the dRC region of these three OCs, would thus appear to not be enough to explain the overestimated total mass values given by H03 and P12 (in particular to NGC 419 by H03).

In addition to the presence of a dRC, both NGC 419 and NGC 1751 show extended or multiple main-sequence turnoffs (MSTO; see: Glatt et al. 2008; Milone et al. 2009; Rubele et al. 2010, 2011; Girardi et al. 2011) while NGC 1917 is known to possess a broadened MSTO (Milone et al. 2009). The origin of this structure is still under debate, as seen in Piatti & Bastian (2016), Milone et al. (2016), and Li et al. (2016). Its influence on the derived masses from integrated photometry studies is not straightforward to assess, nor can its impact be easily discarded.

If we examine the H03 and P12 databases, we find that the aforementioned bias – by which integrated photometric studies increasingly overestimate masses for larger mass OCs – exists even when comparing these studies among themselves. After removing duplicated entries in both DBs, and cross-matching them with a maximum search radius of 20 arcsec, we are left with 670 unique LMC OCs across H03 and

<sup>28</sup> Where both V and B filters correspond to the Johnson photometric system.



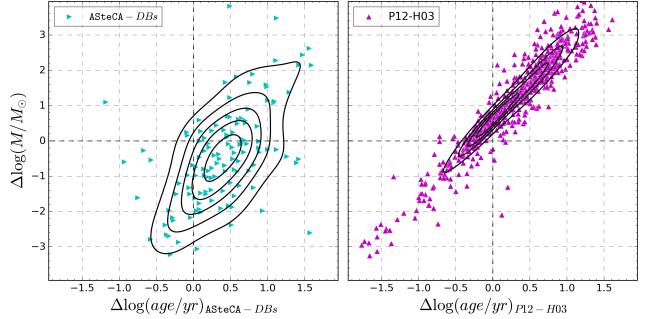
**Figure 10.** *Left:* BA plot for the relative difference between P12 minus H03 masses, for average mass values below  $5000 M_\odot$ . OCs are coloured according to the difference in their assigned logarithmic ages by each DB – i.e.:  $\Delta \log(\text{age}/\text{yr})$  – in the sense P12 minus H03; see colorbar in the rightmost plot. The mean and standard deviation for  $\overline{\Delta M}_r$  is shown as a dashed black line and a grey region, respectively. *Centre:* idem, for average DB masses  $5000 < \overline{M}_{DB} < 20000 (M_\odot)$ . *Right:* idem, for average DB masses  $\overline{M}_{DB} > 20000 (M_\odot)$ .

P12. Fig. 10 shows BA diagrams for these cross-matched OCs. We plot here the average P12-H03 mass  $\overline{M}_{DBs}$  versus their relative difference  $\overline{\Delta M}_r$ , as defined in Eq. 3, in the sense P12 minus H03. As done previously, masses are separated into three regimes for clarity. There are only five OCs that show average masses larger than  $100000 M_\odot$ , and they are all massive globular clusters which are incorrectly assigned a low age and mass by P12.<sup>29</sup> For example, the LMC globular cluster NGC1835 is correctly identified by H03 as an old  $\sim 5$  Gyr system, with a total mass of  $\sim 1.4 \times 10^6 M_\odot$  (a reasonable value, although a bit overestimated, according to Dubath et al. 1990). P12 on the other hand classifies this as an extremely young  $\sim 6.3$  Myr OC with a very low total mass estimate of  $1700 M_\odot$ .

In Fig. 10 we see that, as the average OC mass given by these DBs increases, so do their relative differences. The mean values of  $\overline{\Delta M}_r$  decreases from  $\sim 0.5 \pm 0.4$  in the low mass region, to  $\sim 0.1 \pm 0.4$  in the medium OC mass region, to  $\sim -0.4 \pm 0.6$  in the large mass region. Masses go from being overestimated a factor of 3 by P12 – in relation to H03 – in the low mass region, to being underestimated by a factor of more than  $\sim 2$  in the large mass region. These differences in total mass grow with larger average masses, in a way that is closely related to the difference in the  $\log(\text{age}/\text{yr})$  values estimated by each DB. Where P12 estimates larger masses than H03 – i.e., the low mass region – it also assigns larger ages by more than 1.5 dex. Conversely, in the large average mass region, P12 ages can reach differences of up to 3 dex lower than H03. This age-mass positive correlation, by which an older large cluster can be incorrectly identified as a much younger and less massive one or vice-versa, is also noticeable in Fig. 9 albeit to a lesser extent. In Fig. 11 we show how the age-mass correlation affects DBs estimates compared to those taken from ASteca (left plot), and how this effect is much stronger in P12 and H03 estimates (right plot).

About  $\sim 17\%$  of the OCs in H03 and P12, are assigned masses above  $5000 M_\odot$  in their respective DBs. A

<sup>29</sup> These five LMC globular clusters are: NGC 1916, NGC 1835, NGC 1786, NGC 1754, and NGC 1898. P12 assigns masses below  $2000 M_\odot$  in all cases.



**Figure 11.** *Top:* Differences plot for ASteca minus the combined P12 and H03 DBs. Horizontal and vertical axis show differences in  $\log(\text{age}/\text{yr})$ , and  $\log(M/M_\odot)$ , respectively. A 2-dimensional Gaussian kernel density estimate is shown as iso-density black curves. *Bottom:* idem, for ages and masses of P12 and H03 cross-matched OCs, in the sense P12 minus H03.

smaller percentage, less than  $\sim 8\%$  and  $\sim 5\%$ , are assigned by those works mass values above  $10000 M_\odot$  and  $20000 M_\odot$  respectively. Similar proportions are found when inspecting ASteca's obtained masses. Clusters with relatively large masses represent – as demonstrated by the aforementioned percentages – a small portion of OCs in the DBs. Nonetheless, care should be taken when applying their integrated photometry estimated masses to the study of properties such as the initial cluster mass function (ICMF). Large discrepancies in the mass value assigned for the most massive OCs, could have a non-negligible impact on the slope of the IMCF.

<– AGREGAR ALGO MAS? –>

## 6 PARAMETERS DISTRIBUTION

We present here a summary of the distribution of fundamental parameter values obtained with ASteca, for our set of 239 Magellanic clouds clusters.

In astrophysics analysis, histograms are widely used to derive a large number of properties when a substantial amount of data is available. A galaxy's star formation his-

tory (SFH) is a good example of such a property, almost always quantified via a one-dimensional histogram. Their widespread use notwithstanding, the generation of a histogram is affected by well known statistical issues; see Silverman (1986), Simonoff & Udina (1997). Different selected bin widths and anchor positions (point of origin) can make histograms built from the exact same data look utterly dissimilar. In the worst cases, completely spurious sub-structures may appear, leading the analysis towards erroneous conclusions. We bypass these issues by constructing an adaptive (or variable) Gaussian kernel density estimate (KDE) in one and two dimensions, using the standard deviations associated to a given parameter as the bandwidth estimates. The formulas for both KDEs are:

$$KDE_{1D}(x) = \frac{1}{N\sqrt{2\pi}} \sum_{i=1}^N \frac{1}{\sigma_i} e^{-\frac{(x-x_i)^2}{2\sigma_i^2}}, \quad (4)$$

$$KDE_{2D}(x, y) = \frac{1}{2\pi N} \sum_{i=1}^N \frac{1}{\sigma_{xi}\sigma_{yi}} e^{-\frac{1}{2}\left(\frac{(x-x_i)^2}{\sigma_{xi}^2} + \frac{(y-y_i)^2}{\sigma_{yi}^2}\right)}, \quad (5)$$

where  $N$  is the number of observed values,  $x_i$  is the  $i$ th observed value of parameter  $x$ , and  $\sigma_{xi}$  its assigned standard deviation (same for  $y_i$  and  $\sigma_{yi}$ ). The 1D version of these KDEs is similar to the “smoothed histogram” used in the Rafelski & Zaritsky (2005) study of SMC clusters. The use of standard deviations as bandwidth estimates means that the contribution to the density map (in 1D or 2D) of parameters derived with large errors, will be smoothed (or “spread out”) over a large portion of the parameter’s domain. Precise parameter values on the other hand – i.e. those with small assigned errors – will contribute to a much more narrow region, as one would expect.

Replacing one and two-dimensional histogram analysis with these KDEs has two immediate benefits: a- it frees us from having to select an arbitrary value for the bandwidth (the most important component of a KDE, equivalent to the bin width of a regular histogram), and b- it naturally incorporates the errors obtained for each studied parameter into its derived probability density function. Figs. 12 and 13 show 1D and 2D density maps constructed via Eqs. 4 and 5 for two paired parameters, for each of the MCs. Being probability density functions means that the area under the curve integrates to 1. This makes the distributions for a given parameter comparable for both galaxies, even if the number of observed points – clusters in our case – is not the same (equivalent to a normalized histogram).

Based on the analysis of a database of over 1500 OCs, G10 reports two periods of enhanced cluster formation in the MCs: around 160 Myr and 630 Myr for the SMC, and around 125 Myr and 800 Myr for the LMC. A third period is reported at approximately 8 Myr for both MCs. To obtain these results, the authors construct histograms of their data, combined with data from P00 and C06 (see Figs. 5 and 6 in G10). They analyse the peaks found in both histograms and conclude that the formation episodes are correlated, as they happened around the same period of time. This is a good example of the issues mentioned at the beginning of this section, regarding histogram construction. If we look at the 1D  $\log(\text{age}/\text{yr})$  KDEs in Fig. 12 (top), we see that most

of the enhanced formation periods reported in G10 are not present. A distinct period of cluster formation is visible in the LMC starting around the  $\sim 5$  Gyr mark, which lasted up to  $\sim 1.3$  Gyr ago. A similar, but much less pronounced peak is seen for the SMC, with a clear drop in cluster formation around  $\sim 2$  Gyr. The height difference between the SMC and LMC KDEs is related to the decline in cluster formation. While the LMC sharply drops to almost zero from  $\sim 1$  Gyr to present times, the SMC shows a much softer descent with smaller peaks around  $\sim 250$  Myr and  $\sim 130$  Myr. The well known “age gap” in the LMC between 3–10 Gyrs (Babbinot et al. 2010) is present, visible as a marked drop in the  $KDE_{\log(\text{age}/\text{yr})}$  curve at approximately  $\sim 9.5$  dex. No clusters older than this age are found in our work.

The 2D KDE age-metallicity map shows how spread out these values are for OCs in the SMC, compared to those in the LMC which are much more heavily clustered together. Although in this map the abundance of the SMC can be seen to reach substantially lower values than the LMC, the 1D KDE to the right reveals that the [Fe/H] parameter peaks between 0 dex and -0.2 dex, for OCs in both clouds.

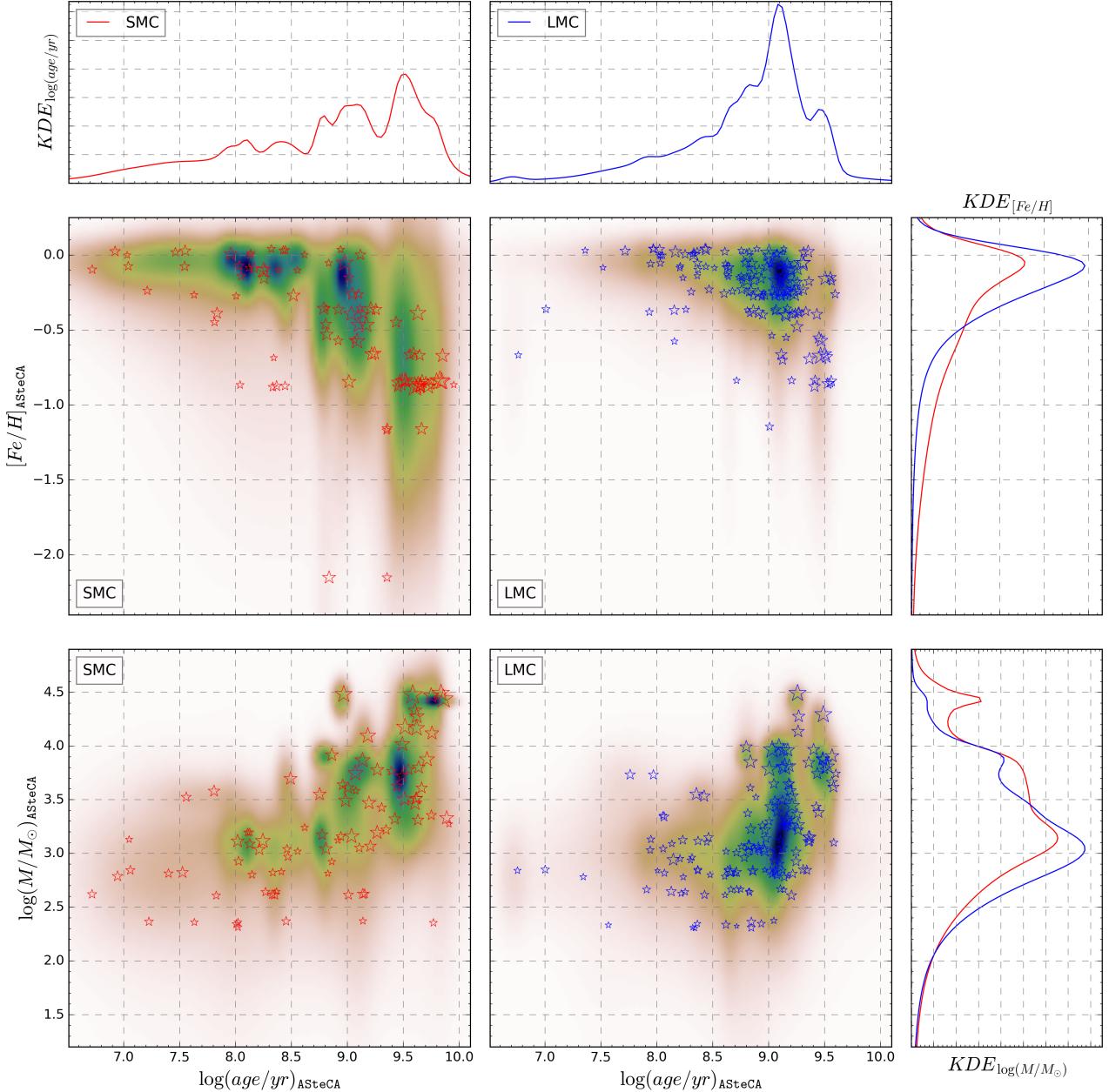
The age-mass 2D map shows a clustering around younger ages and smaller masses for the LMC, relative to the SMC. The OC seen in the bottom right corner of the age-mass SMC map is HW42 ( $\alpha=1^h01^m08^s$ ,  $\delta=-74^\circ04'25''$ , [J2000.0]), a small cluster ( $r_{cl}<20$  pc) located close to the SMC’s centre. Though its position in the map is somewhat anomalous, the  $1\sigma$  error in its age and mass estimates could move it to  $\log(\text{age}/\text{yr})\approx 9.4$  and  $\log(M/M_\odot)\approx 2.6$ . This OC is classified as a possible emissionless association by Bica & Schmitt (1995). In both clouds there’s a tendency for the mass, and the size of the OC, to grow with the estimated age, as expected. In the 1D mass KDE, we see that the LMC accumulates most OCs in the low mass regime ( $\sim 3000 M_\odot$ ). This is also true for the SMC which has a larger proportion of large mass OCs, with a distinctive peak around  $\sim 30000 M_\odot$ .

As seen in Fig. 13 (top), the 1D KDEs of the true distance moduli are well behaved and clearly normal in their distribution. A Gaussian fit to these curves results in best fit values of  $18.96\pm 0.08$  mag for the SMC, and  $18.49\pm 0.08$  mag for the LMC. The initial ranges given in Table 3, as well as the literature mean distances, are thus properly recovered. Extinction values are much more concentrated in the SMC around  $E_{B-V}\approx 0.015$  mag. The OCs in the LMC on the other hand, show that most values are dispersed below  $E_{B-V}\approx 0.1$  mag, with a shallower peak located at  $\sim 0.03$  mag.

## 6.1 Age-metallicity relation

A stellar system’s age-metallicity relation (AMR) is an essential tool to learn about its chemical enrichment evolution. This relation is usually presented either as scattered single points in the age-metallicity space, or as a function created by grouping and averaging metallicity estimates in arbitrary age bins. In Piatti (2010) a method was devised to generate an AMR able to take into account the errors in age values, to produce bins of different sizes. This method has been applied to the obtention of AMRs in Piatti & Geisler (2013), and also adapted to derive star cluster frequency distributions (e.g., Piatti 2013).

We can take advantage of the KDE technique described in Sect. 6 – used to produce two-parameters density maps –

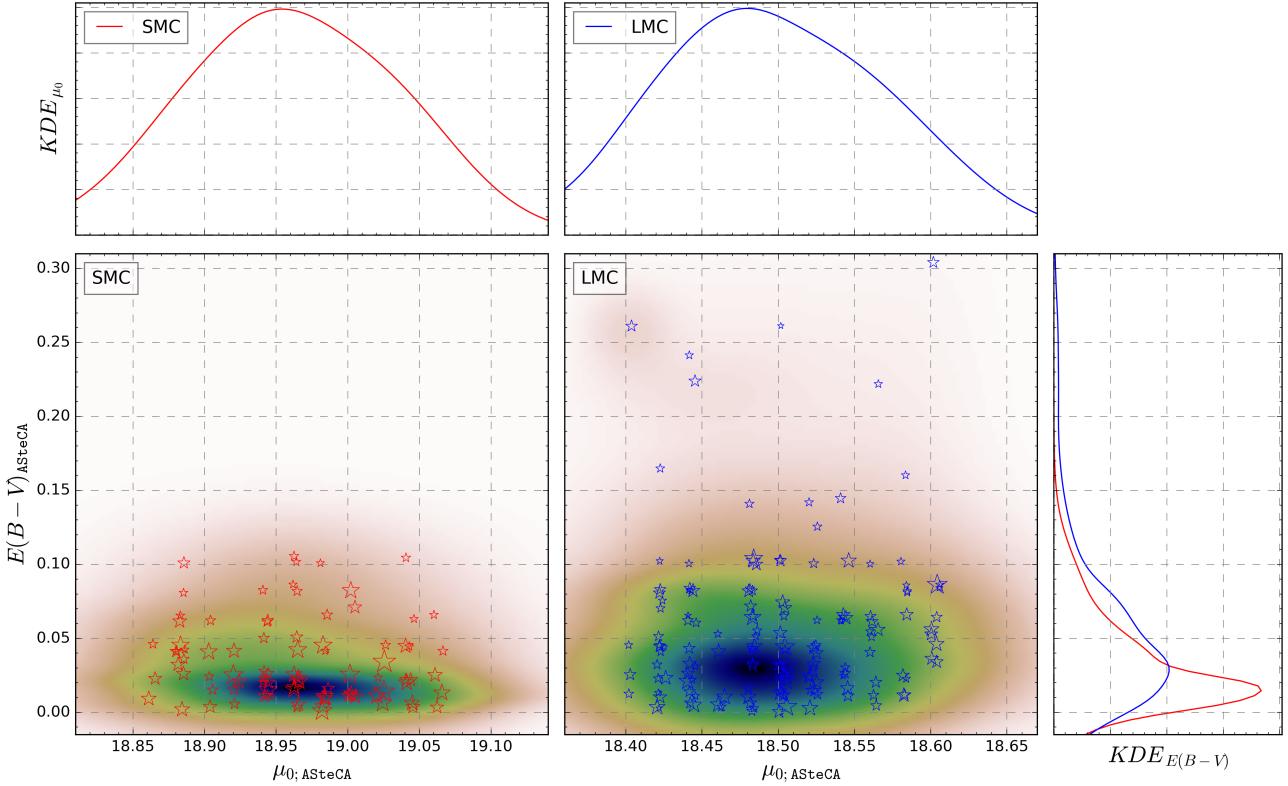


**Figure 12.** One and two-dimensional Gaussian adaptive KDEs for the age, metallicity and mass parameters. Top and right plots are 1D KDEs while the centre plots are 2D KDEs. Observed clusters are plotted as red and blue stars for the S/LMC, respectively in the 2D KDEs. Sizes are scaled according to each OC's radius in parsecs. A small scatter is introduced for clarity.

to generate an AMR that is truly representative of the observed data, with some important improvements over previous methods. First, a Gaussian density map has no dependence on the number, size (fixed or variable) or location of bins, as a regular histogram or the aforementioned method would. Second, the errors in the two parameters used to obtain the density map (age and metallicity), are organically included in the function that generates it (as explained in Sect. 6). This means that no ad-hoc procedure is required to incorporate the important information carried by these values, into the final AMR.

The process of creating an AMR function<sup>30</sup> requires that we associate a unique  $[Fe/H]$  to a single age value, for the available age range. After generating the age-metallicity 2D density map, a dense 2D grid is created dividing it into N steps of 0.01 dex width, covering the ranges of both parameters. Every point in this grid is evaluated in the KDE map and its value ( $w_i$ ) is stored, along with its age-metallicity coordinates ( $age_i$ ,  $[Fe/H]_i$ ). Each of the N ages defined in the grid is then associated to a single representative  $[Fe/H]$

<sup>30</sup> By “AMR function” we mean a curve that spans the entire observed age range, mapping each age value to a single metallicity value.



**Figure 13.** Same as Fig. 12 for the extinction and distance modulus parameters.

value. This representative metallicity for a given age is obtained as the mean metallicity value, weighted by the KDE function at that particular age. The formal equation for obtaining it, can be written as

$$\overline{[Fe/H]}_{age_i} = \frac{\sum w_i [Fe/H]_i}{\sum w_i} \quad (6)$$

where the summations are performed over each  $i$  step for the  $N$  defined steps in the metallicity range,  $[Fe/H]_i$  is the metallicity value at step  $i$ , and  $w_i$  is the value of the 2D KDE map for that fixed age and metallicity coordinates. The  $age_i$  subindex in Eq. 6 indicates that this mean metallicity was calculated for a fixed age value, and thus represents a unique point in the AMR. A similar version of this weighted metallicity was employed in Noël et al. (2009, see Eq. 3) to derive AMR estimates for three observed fields. We apply the above formula to all the  $N$  ages in the grid defined at the beginning of the process. The standard deviation for each  $\overline{[Fe/H]}_{age_i}$  value is calculated through the equation

$$\sigma_{age_i}^2 = \frac{\sum w_i \sum [w_i ([Fe/H]_i - \overline{[Fe/H]}_{age_i})]^2}{(\sum w_i)^2 - \sum w_i^2} \quad (7)$$

where again all summations are applied over  $N$ , and the descriptions given for the parameters in Eq. 6 apply. At this point, this method already gives us an AMR function estimate, since every age step is mapped to a unique metallicity. The downsides are that the AMR is noisy due to the very small step of 0.01 dex used, and that the associated errors are quite large. This latter effect arises because the weighted

standard deviation, Eq. 7, will be affected not only by the errors in both measured parameters – through the constructed 2D KDE map – but also by the intrinsic dispersion in the metallicity values found for any given age. To solve this, we calculate the average  $[Fe/H]$  for a given age interval, rather than assigning a metallicity value to each age step in the grid. Dividing the age range into intervals requires a decision about the step width, much like when constructing a histogram, bringing back the issue of binning. We have two advantages here: a- we use Knuth’s algorithm (see Sect. 3.4) to obtain the optimal binning for our data, and b- the final AMR function is very robust to changes in the binning method selected, so even the previous choice is not crucial in determining the shape of our AMR. Finally, the  $\overline{[Fe/H]}_{age_i}$  values obtained for every  $age_i$  within a defined age interval, are averaged. Errors are propagated through the standard formula, disregarding covariant terms (Bevington & Robinson 2003, Eq. 3.14).

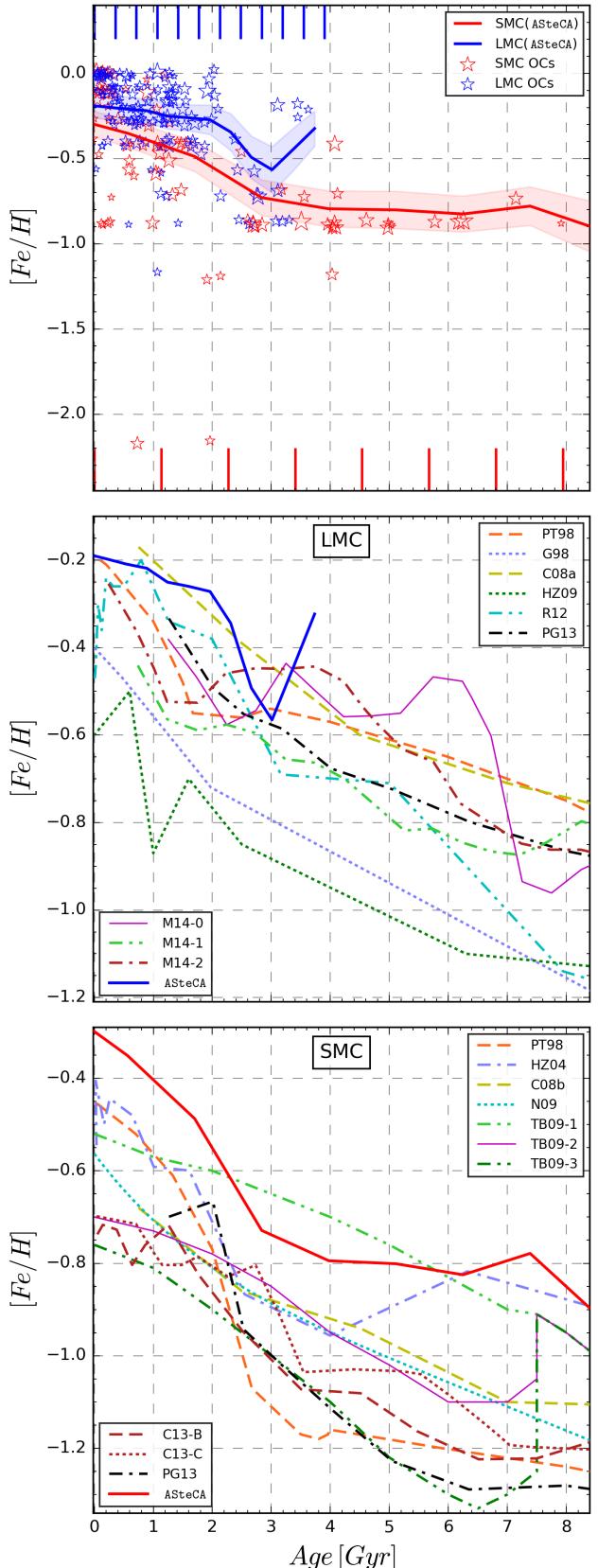
The final *ASteCA* AMRs for the S/LMC can be seen in Fig. 14 as red and blue continuous lines, respectively. Stars show the position of all OCs in our sample for each galaxy, with sizes scaled according to their radii. The shaded regions represent the  $1\sigma$  standard deviations of the AMR functions. These regions span a  $[Fe/H]$  width of approximately 0.2 dex for both Clouds, for the entire age range. The blue (top) and red (bottom) vertical segments in the top plot are the bin edges determined for each age interval by Knuth’s algorithm, for the LMC and the SMC respectively. As stated previously, the final AMR functions are mostly unaffected by the chosen binning method. Using Knuth’s algorithm results in approximately 10 age intervals of 1 Gyr width, and

the AMRs seen in Fig. 14. If instead we use 100 intervals of  $\sim 0.1$  dex width, the only substantial change is that the SMC curve is raised by  $\sim 0.1$  dex, for ages below 500 Myr. The two SMC clusters with extremely low metallicities –  $[\text{Fe}/\text{H}] < -2$  dex – are NGC294 and HW85; their abundances were already analysed in Sect. 5.1. Having such small metallicity values means that their associated uncertainty will be quite large, as discussed in Sect. 4. This has the effect of spreading their positions in the density map, preventing these low values from affecting the AMR substantially. If these two OCs are excluded from our data, the resulting AMR for the SMC moves upwards in the  $[\text{Fe}/\text{H}]$  axis by less than 0.02 dex.

Several chemical evolution models and empirically estimated AMRs can be found in the literature for both Magellanic Clouds. Many of the studies performed on the age-metallicity relation of the S/LMC present their results as scattered points in the age vs metallicity plot, rather than fitting a unique AMR function to describe the distribution of those observed values. To allow a straightforward comparison with our AMR functions, we show in Fig. 14 – centre and bottom plots – the functions presented in twelve other works. These studies constitute a representative sample of the different methods and data used in the literature over the past twenty years, for these two galaxies: [Pagel & Tautvaisiene \(1998, PT98; bursting models\)](#), [Geha et al. \(1998, G98; closed-box model with Holtzman SFH\)](#), [Harris & Zaritsky \(2004, HZ04\)](#), [Carrera et al. \(2008a, C08a; average of four disk frames\)](#), [Carrera et al. \(2008b, C08b; average of thirteen frames\)](#), [Harris & Zaritsky \(2009, HZ09\)](#), [Noël et al. \(2009, N09; 5th degree polynomial fit to the AMRs of their three observed regions\)](#), [Tsujimoto & Bekki \(2009, TB09; 1: no merger model, 2: equal mass merger, 3: one to four merger\)](#), [Rubele et al. \(2012, R12; four tiles average\)](#), [Cignoni et al. \(2013, C13; B: Bologna, C: Cole\)](#), [Piatti & Geisler \(2013, PG13\)](#), and [Meschin et al. \(2014, M14; 0: field LMC0, 1: field LMC1, 2: field LMC2\)](#). Details on how these AMRs were constructed will not be given here, as they can be consulted in each reference. All of the above mentioned articles used field stars for the obtention of their AMRs. This is, as far as we are aware, the first work were the AMR function for both galaxies is derived entirely from observed star clusters.

The overall trend of both AMRs coincides with what has already been found in the literature, namely that the metallicity increases rather steadily with younger ages. On average, our AMR estimates are displaced slightly towards more metal rich values, particularly in the case of the SMC. This tendency was already mentioned in Sect. 5.1, where possible causes for at least some part of the effect were also given.

For the LMC galaxy, Fig. 14 centre plot, we see a marked drop in metallicity from  $\sim -0.3$  dex beginning around 3.8 Gyr, and ending 3 Gyrs ago at  $\sim -0.55$  dex. The former high  $[\text{Fe}/\text{H}]$  value is not found in any of the remaining functions for the LMC. The M14-0 curve seems to reproduce this behaviour, but shifted  $\sim 0.8$  Gyr towards younger ages, and with a much less pronounced peak that reaches its maximum of  $\sim -0.5$  dex at  $\sim 3.2$  Gyr. The high metallicity value in the LMC AMR at 3.8 Gyr, is caused by the four OCs located around  $\sim 3.3$  Gyr with  $[\text{Fe}/\text{H}] \approx -0.25$  dex; when they are excluded, the peak disappears. These four OCs are: SL33, H3, SL5, and KMHK586, their CMDs



**Figure 14.** ASTeCA’s age-metallicity relation for the SMC (red solid line) and the LMC (blue solid line). Shaded areas are the  $1\sigma$  regions for each AMR. Blue and red stars are the LMC and SMC clusters in our set. See Sect. 6.1 for more details.

are shown in Appendix D, Fig D1. The last three of these OCs are analysed in Piatti (2011b) and have no metallicities or extinction values assigned, as only ages were estimated in this article via the  $\delta T_1$  index. Ages obtained by *AStéCA* for this group of OCs are between 0.1–0.3 dex larger than the ones given in Piatti (2011b). According to this article, the oldest OCs in the group have ages of  $\sim 2.5$  Gyr. This means that if literature's ages were used instead of the ones produced by the code, the peak would not show. In fact, the entire AMRs functions derived using literature values – i.e., those taken from the articles in Table 1 – are markedly different from the ones shown in Fig. 14. This is expected as most OCs in these works are assigned fixed metallicities of -0.7 dex and -0.4 dex for the S/LMC, respectively; particularly for estimated ages below 1 Gyr. Literature AMRs for both MCs can be seen in Appendix D, Fig D2.

After the aforementioned drop in the LMC's AMR, there is a steep climb from 3 Gyr to 2 Gyr reaching almost  $[Fe/H] \sim -0.25$  dex, and then a sustained but much more shallow increase up to the present day's metal content of  $\sim -0.2$  dex. The AMR functions for the LMC that differ the most from the one obtained with *AStéCA* values, are those taken from HZ09 and G98. These two curves are visibly separated, not only from our AMR, but also from the rest of the group. Our average metallicity value for present day OCs, coincides reasonably well with those from PT98, C08a, and M14-2. The PT98 bursting model and field 2 from M14, show nonetheless a very different rate of increase from 2 Gyr to present times, compared to *AStéCA*'s AMR. The average AMR from C08a, while lacking finer details, provides a better match for this age range.

The AMR function for the SMC obtained using *AStéCA*'s age and metallicity values, is shown along ten AMRs taken from the published literature in Fig. 14, bottom. Our AMR shows an increased rate of  $[Fe/H]$  that reaches a peak around  $\sim 7.5$  Gyr ago, followed by a marked dip between 6–7 Gyr. This feature was reported by TB09 in its two merger models. The TB09-2 model (equal mass merger) appears to best mimic our AMR for the SMC, although ours is shifted upwards towards higher abundances. The TB09-3 model (1:4 mass merger) shows the same peak, but followed by a much more pronounced dip that goes well below most AMRs, including our own. The TB09-2 model and our AMR follow a very similar path up to 3 Gyr, where they start to deviate. For present day ages, where the AMRs differ the most, our curve is more metal rich by about 0.4 dex compared to the TB09-2 model.

Following the dip beyond  $\sim 6$  Gyr, abundances in our AMR plateau around a value of  $[Fe/H] \approx -0.8$  dex until approximately 3 Gyrs ago, where the rate of growth for the metallicity increases considerably. From that point up to the present day, the average metallicity for OCs in the SMC grows by about 0.4 dex, according to the *AStéCA* AMR. The increased rate of growth behaviour for ages younger than 3 Gyrs, is only reproduced by the PT8 model, and the HZ04 function. In both cases, the maximum  $[Fe/H]$  value attained by these AMRs for the youngest ages is lower by  $\sim 0.12$  dex, compared to our own estimate for this galaxy.

The canonical metallicity values usually accepted –  $[Fe/H] = -0.7$  dex,  $[Fe/H] = -0.4$  dex, for the S/LMC – are shown to be off by  $\sim 0.3$  dex on average, for clusters younger

than 2.5 Gyr in both galaxies. Overall our AMRs can not be explained by any single model or empirical AMR function, and are best reproduced by a combination of several of these. A similar result was found in Piatti & Geisler (2013), although their AMRs – derived from field star population – are significantly different from ours, mainly for the SMC case.

It is important to remember that the AMRs estimated using the ages and metallicities derived via *AStéCA*, are averaged over the structure of both galaxies. In Fig. 1 we showed that our set of OCs covers a large portion of the surface of these galaxies. If more OCs were available so that the AMRs could be instead estimated by sectors in the S/LMC, it is entirely possible that different results would arise. When clusters in our set are divided by sectors, statistically low numbers are assigned to each – particularly for the SMC – which makes this more detailed study not feasible.

<– AGREGAR/MODIFICAR ALGO? –>

## 7 SUMMARY AND CONCLUSIONS

We presented an homogeneous catalogue for a set of 239 star clusters in the Large and Small Magellanic Clouds, observed with the Washington photometric system. The OCs span a wide range in metallicity, age, and mass, and are spatially distributed throughout both galaxies. The fundamental parameters metallicity, age, reddening, distance modulus, and total mass were determined using the *AStéCA* package. This tool allows the automated processing of an OC's positional and photometric data, resulting in estimates of both its structural and intrinsic/extrinsic properties. As already shown in Paper I, the advantages of using this package include reproducible and objective results, along with a proper handling of the uncertainties involved in the synthetic cluster matching process. This permits the generation of a truly homogeneous catalogue of observed clusters, with their most important parameters fully recovered. Our resulting catalogue is complete for all the analysed parameters, including metallicity and mass, two properties often assumed or not obtained at all.

The analysis of our results, demonstrate that the assigned values for the OCs are in good agreement with published literature which used the same Washington photometry. The metallicity parameter showed to be the most discrepant one, with *AStéCA* values being on average  $\sim 0.22$  dex larger than those present in the same photometry literature. The most likely explanation for this is the confirmation bias effect, by which most OC studies will assume the canonical  $[Fe/H]$  values rather than derive them through statistically valid means. We also compared our results with those taken from articles that used different photometric systems. While the differences in age in this case are somewhat larger, they can be mostly explained by effects outside the code. Internal errors show no biases present in our determination of fundamental parameters.

We performed a detailed comparative study of masses obtained through integrated photometry studies, with our own estimates from CMD analysis. The total mass values from the latter studies are shown to be systematically over-

estimated. This result is opposite to the expectation that larger clusters can have their masses recovered with moderate accuracy, via integrated photometry.

A method for deriving the distribution of any fundamental parameter – or a combination of two of them – is presented. This method takes into account the information contained by the uncertainties, often excluded from the analysis. By relying on Gaussian kernels, it is robust and independent of ad-hoc binning choices. An age-metallicity relation is derived using the above mentioned method and the parameter values obtained by **AStECA**, for both galaxies. The AMRs generated can not be fully matched by any model or empirical determination found in the recent literature.

We demonstrated that the **AStECA** package is able to produce proper estimations of observed star clusters, with their fundamental parameters covering a wide range of values. A necessary statistically valid error analysis can be performed, thanks to its built-in bootstrap error assignment method. The tool is also proven to be capable of operating almost entirely unassisted, on large databases of clusters. This is an increasingly essential feature of any astrophysical analysis tool, given the growing importance of big data and the necessity to conduct research on large astronomical data sets.

## ACKNOWLEDGMENTS

GIP would like to thank the help and assistance provided throughout the redaction of several portions of this work by: D. Hunter, A. E. Dolphin, M. Rafelski, D. Zaritsky, T. Palma, F. F. S. Maia, B. Popescu, and H. Baumgardt. This research has made use of the VizieR<sup>31</sup> catalogue access tool, operated at CDS, Strasbourg, France (Ochsenbein et al. 2000). This research has made use of “Aladin sky atlas”<sup>32</sup> developed at CDS, Strasbourg Observatory, France (Bonnarel et al. 2000; Boch & Fernique 2014). This research has made use of NASA’s Astrophysics Data System<sup>33</sup>. This research made use of the Python language v2.7<sup>34</sup> (van Rossum 1995), and the following packages: NumPy<sup>35</sup> (Van Der Walt et al. 2011); SciPy<sup>36</sup> (Jones et al. 2001); Astropy<sup>37</sup>, a community-developed core Python package for Astronomy (Astropy Collaboration et al. 2013); scikit-learn<sup>38</sup> (Pedregosa et al. 2011); matplotlib<sup>39</sup> (Hunter et al. 2007). This research made use of the Tool for OPerations on Catalogues And Tables (TOPCAT)<sup>40</sup> (Taylor 2005).

## REFERENCES

- Anders P., Fritze-v. Alvensleben U., 2003, *A&A*, **401**, 1063
- Anders P., Kotulla R., De Grijs R., Wicker J., 2013, *The Astrophysical Journal*, **778**, 138
- Andrae R., 2010, preprint, ([arXiv:1009.2755](https://arxiv.org/abs/1009.2755))
- Astropy Collaboration et al., 2013, *A&A*, **558**, A33
- Balbinot E., Santiago B. X., Kerber L. O., Barbuy B., Dias B. M. S., 2010, *MNRAS*, **404**, 1625
- Baumgardt H., Parmentier G., Anders P., Grebel E. K., 2013, *MNRAS*, **430**, 676
- Bertelli G., Bressan A., Chiosi C., Fagotto F., Nasi E., 1994, *A&AS*, **106**
- Bevington P. R., Robinson D. K., 2003, McGraw-Hill
- Bica E. L. D., Schmitt H. R., 1995, *ApJS*, **101**, 41
- Bica E., Bonatto C., Dutra C. M., Santos J. F. C., 2008, *Monthly Notices of the Royal Astronomical Society*, **389**, 678
- Bland J. M., Altman D. G., 1986, *Lancet*, **1**, 307
- Boch T., Fernique P., 2014, in Manset N., Forshay P., eds, *Astronomical Society of the Pacific Conference Series Vol. 485, Astronomical Data Analysis Software and Systems XXIII*. p. 277
- Bonatto C., Bica E., 2007, *MNRAS*, **377**, 1301
- Bonnarel F., et al., 2000, *AAPS*, **143**, 33
- Bressan A., Marigo P., Girardi L., Salasnich B., Dal Cero C., Rubele S., Nanni A., 2012, *MNRAS*, **427**, 127
- Canterna R., 1976, *The Astronomical Journal*, **81**, 228
- Carrera R., Gallart C., Hardy E., Aparicio A., Zinn R., 2008a, *The Astronomical Journal*, **135**, 836
- Carrera R., Gallart C., Aparicio A., Costa E., Méndez R. A., Noël N. E. D., 2008b, *The Astronomical Journal*, **136**, 1039
- Chabrier G., 2001, *ApJ*, **554**, 1274
- Chiosi E., Vallenari A., Held E. V., Rizzi L., Moretti A., 2006, *Astronomy and Astrophysics*, **452**, 179
- Choudhury S., Subramaniam A., Piatti A. E., 2015, *The Astronomical Journal*, **149**, 52
- Cignoni M., Cole A. A., Tosi M., Gallagher J. S., Sabbi E., Anderson J., Grebel E. K., Nota A., 2013, *ApJ*, **775**, 83
- Dias W. S., Alessi B. S., Moitinho A., Lépine J. R. D., 2002, *A&A*, **389**, 871
- Dias B., Coelho P., Barbuy B., Kerber L., Idiart T., 2010, *A&A*, **520**, A85
- Dolphin A. E., 2002, *Monthly Notices of the Royal Astronomical Society*, **332**, 91
- Dubath P., Meylan G., Mayor M., Magain P., 1990, *A&A*, **239**, 142
- Elson R. A. W., Sigurdsson S., Davies M., Hurley J., Gilmore G., 1998, *MNRAS*, **300**, 857
- Geha M. C., et al., 1998, *The Astronomical Journal*, **115**, 1045
- Geisler D., 1996, *AJ*, **111**, 480
- Geisler D., Sarajedini A., 1999, *The Astronomical Journal*, **117**, 308
- Geisler D., Bica E., Dottori H., Claria J. J., Piatti A. E., Santos Jr. J. F. C., 1997, *AJ*, **114**, 1920
- Geisler D., Piatti A. E., Bica E., Claria J. J., 2003, *Monthly Notices of the Royal Astronomical Society*, **341**, 771
- Girardi L., Groenewegen M. A. T., Weiss A., Salaris M., 1998, *MNRAS*, **301**, 149
- Girardi L., Bertelli G., Bressan A., Chiosi C., Groenewegen M. A. T., Marigo P., Salasnich B., Weiss A., 2002, *A&A*, **391**, 195
- Girardi L., Rubele S., Kerber L., 2009, *MNRAS*, **394**, L74
- Girardi L., Eggenberger P., Miglio A., 2011, *MNRAS*, **412**, L103
- Glatt K., et al., 2008, *The Astronomical Journal*, **136**, 1703
- Glatt K., Grebel E. K., Koch A., 2010, *Astronomy and Astrophysics*, **517**, A50
- Grocholski A. J., Sarajedini A., 2003, *MNRAS*, **345**, 1015
- Harris J., Zaritsky D., 2004, *AJ*, **127**, 1531
- Harris J., Zaritsky D., 2009, *The Astronomical Journal*, **138**, 1243
- Haschke R., Grebel E. K., Duffau S., 2011, *The Astronomical Journal*, **141**, 158

<sup>31</sup> <http://vizier.u-strasbg.fr/viz-bin/VizieR>

<sup>32</sup> <http://aladin.u-strasbg.fr/>

<sup>33</sup> <http://www.adsabs.harvard.edu/>

<sup>34</sup> <http://www.python.org/>

<sup>35</sup> <http://www.numpy.org/>

<sup>36</sup> <http://www.scipy.org/>

<sup>37</sup> <http://www.astropy.org/>

<sup>38</sup> <http://scikit-learn.org/>

<sup>39</sup> <http://matplotlib.org/>

<sup>40</sup> <http://www.starlink.ac.uk/topcat/>

- Hills S., von Hippel T., Courteau S., Geller A. M., 2015, *The Astronomical Journal*, 149, 94
- Hunter D. A., Elmegreen B. G., Dupuy T. J., Mortonson M., 2003, *Astron J*, 126, 1836
- Hunter J. D., et al., 2007, Computing in science and engineering, 9, 90
- Jones E., Oliphant T., Peterson P., et al., 2001, SciPy: Open source scientific tools for Python, <http://www.scipy.org/>
- Kharchenko N. V., Piskunov A. E., Röser S., Schilbach E., Scholz R.-D., 2005, *A&A*, 440, 403
- Knuth K. H., 2006, ArXiv Physics e-prints,
- Krouwer J. S., 2008, Stat Med, 27, 778
- Leitherer C., et al., 1999, *ApJS*, 123, 3
- Lejeune T., Schaerer D., 2001, *A&A*, 366, 538
- Li C., Grijjs Richard d., Bastian N., Deng L., Niederhofer F., Zhang C., 2016, preprint, ([arXiv:1606.05394](https://arxiv.org/abs/1606.05394))
- Maia F. F. S., Corradi W. J. B., Santos Jr. J. F. C., 2010, *MNRAS*, 407, 1875
- Maia F. F. S., Piatti A. E., Santos J. F. C., 2013, *Monthly Notices of the Royal Astronomical Society*, 437, 2005
- Marigo P., Girardi L., Bressan A., Groenewegen M. A. T., Silva L., Granato G. L., 2008, *A&A*, 482, 883
- Mateo M., Hodge P., 1986, *ApJS*, 60, 893
- Meschin I., Gallart C., Aparicio A., Hidalgo S. L., Monelli M., Stetson P. B., Carrera R., 2014, *MNRAS*, 438, 1067
- Milone A. P., Bedin L. R., Piotto G., Anderson J., 2009, *A&A*, 497, 755
- Milone A. P., Marino A. F., D'Antona F., Bedin L. R., Da Costa G. S., Jerjen H., Mackey A. D., 2016, *MNRAS*, 458, 4368
- Naylor T., Jeffries R. D., 2006, *MNRAS*, 373, 1251
- Netopil M., Paunzen E., Carraro G., 2015, *Astronomy & Astrophysics*
- Noël N. E. D., Aparicio A., Gallart C., Hidalgo S. L., Costa E., Méndez R. A., 2009, *ApJ*, 705, 1260
- Ochsenbein F., Bauer P., Marcout J., 2000, *A&AS*, 143, 23
- Pagel B. E. J., Tautvaisiene G., 1998, *MNRAS*, 299, 535
- Palma T., Clariá J. J., Geisler D., Piatti A. E., Ahumada A. V., 2013, *Astronomy & Astrophysics*, 555, A131
- Palma T., Gramajo L. V., Clariá J. J., Lares M., Geisler D., Ahumada A. V., 2015, preprint, ([arXiv:1511.05451](https://arxiv.org/abs/1511.05451))
- Paunzen E., Netopil M., 2006, *MNRAS*, 371, 1641
- Pedregosa F., et al., 2011, *Journal of Machine Learning Research*, 12, 2825
- Perren G. I., Vázquez R. A., Piatti A. E., 2015, *Astronomy & Astrophysics*, 576, A6
- Phelps R. L., Janes K. A., Montgomery K. A., 1994, *AJ*, 107, 1079
- Piatti A. E., 2010, *Astronomy and Astrophysics*, 513, L13
- Piatti A. E., 2011a, *Monthly Notices of the Royal Astronomical Society: Letters*, 416, L89
- Piatti A. E., 2011b, *Monthly Notices of the Royal Astronomical Society: Letters*, 418, L40
- Piatti A. E., 2011c, *Monthly Notices of the Royal Astronomical Society: Letters*, 418, L69
- Piatti A. E., 2012, *Astronomy & Astrophysics*, 540, A58
- Piatti A. E., 2013, *Monthly Notices of the Royal Astronomical Society*, 437, 1646
- Piatti A. E., 2014, *MNRAS*, 445, 2302
- Piatti A. E., Bastian N., 2016, preprint, ([arXiv:1603.06891](https://arxiv.org/abs/1603.06891))
- Piatti A. E., Bica E., 2012, *Monthly Notices of the Royal Astronomical Society*, 425, 3085
- Piatti A. E., Geisler D., 2013, *The Astronomical Journal*, 145, 17
- Piatti A. E., Geisler D., Bica E., Claria J. J., 2003a, *Monthly Notices of the Royal Astronomical Society*, 343, 851
- Piatti A. E., Bica E., Geisler D., Claria J. J., 2003b, *Monthly Notices RAS*, 344, 965
- Piatti A. E., Sarajedini A., Geisler D., Seguel J., Clark D., 2005, *Monthly Notices of the Royal Astronomical Society*, 358, 1215
- Piatti A. E., Sarajedini A., Geisler D., Clark D., Seguel J., 2007a, *Monthly Notices of the Royal Astronomical Society*, 377, 300
- Piatti A. E., Sarajedini A., Geisler D., Gallart C., Wischnjewsky M., 2007b, *Monthly Notices of the Royal Astronomical Society: Letters*, 381, L84
- Piatti A. E., Sarajedini A., Geisler D., Gallart C., Wischnjewsky M., 2007c, *Monthly Notices of the Royal Astronomical Society*, 382, 1203
- Piatti A. E., Geisler D., Sarajedini A., Gallart C., Wischnjewsky M., 2008, *Monthly Notices of the Royal Astronomical Society*, 389, 429
- Piatti A. E., Geisler D., Sarajedini A., Gallart C., 2009, *Astronomy and Astrophysics*, 501, 585
- Piatti A. E., Clariá J. J., Parisi M. C., Ahumada A. V., 2011a, *Publications of the Astronomical Society of the Pacific*, 123, 519
- Piatti A. E., Clariá J. J., Bica E., Geisler D., Ahumada A. V., Girardi L., 2011b, *Monthly Notices of the Royal Astronomical Society*, 417, 1559
- Piatti A. E., de Grijjs R., Rubele S., Cioni M.-R. L., Ripepi V., Kerber L., 2015a, *MNRAS*, 450, 552
- Piatti A. E., et al., 2015b, *MNRAS*, 454, 839
- Pietrzynski G., Udalski A., 1999, *AcA*, 49, 157
- Pietrzynski G., Udalski A., 2000, *AcA*, 50, 337
- Popescu B., Hanson M. M., 2010a, *ApJ*, 713, L21
- Popescu B., Hanson M. M., 2010b, *ApJ*, 724, 296
- Popescu B., Hanson M. M., Elmegreen B. G., 2012, *ApJ*, 751, 122
- Rafelski M., Zaritsky D., 2005, *Astron J*, 129, 2701
- Rubele S., Kerber L., Girardi L., 2010, *MNRAS*, 403, 1156
- Rubele S., Girardi L., Kozhurina-Platais V., Goudfrooij P., Kerber L., 2011, *MNRAS*, 414, 2204
- Rubele S., et al., 2012, *A&A*, 537, A106
- Scargle J. D., Norris J. P., Jackson B., Chiang J., 2013, *ApJ*, 764, 167
- Schmeja S., Kharchenko N. V., Piskunov A. E., Röser S., Schilbach E., Froebrich D., Scholz R.-D., 2014, *A&A*, 568, A51
- Scott D. W., 1979, *Biometrika*, 66, 605
- Silverman B., 1986, Density Estimation for Statistics and Data Analysis. Chapman & Hall/CRC Monographs on Statistics & Applied Probability, Taylor & Francis, <https://books.google.com.ar/books?id=e-xsrjsL7WkC>
- Simonoff J. S., Udina F., 1997, Computational Statistics & Data Analysis, 23, 335
- Sollima A., Carballo-Bello J. A., Beccari G., Ferraro F. R., Pecci F. F., Lanzoni B., 2010, *MNRAS*, 401, 577
- Subramanian S., Subramanian A., 2009, *A&A*, 496, 399
- Tammann G. A., Sandage A., Reindl B., 2003, *A&A*, 404, 423
- Taylor M. B., 2005, in Shopbell P., Britton M., Ebert R., eds, *Astronomical Society of the Pacific Conference Series Vol. 347, Astronomical Data Analysis Software and Systems XIV*. p. 29
- Tsujimoto T., Bekki K., 2009, *The Astrophysical Journal Letters*, 700, L69
- Udalski A., Soszynski I., Szymanski M., Kubiak M., Pietrzynski G., Woźniak P., Zebrun K., 1999, *Acta Astron.*, 49, 223
- Van Der Walt S., Colbert S. C., Varoquaux G., 2011, *Computing in Science & Engineering*, 13, 22
- Vanderplas J., Connolly A., Ivezić Ž., Gray A., 2012, in Conference on Intelligent Data Understanding (CIDU). pp 47–54, doi:10.1109/CIDU.2012.6382200
- de Grijjs R., Anders P., 2006, *MNRAS*, 366, 295
- de Grijjs R., Bono G., 2015, *The Astronomical Journal*, 149, 179
- de Grijjs R., Wicker J. E., Bono G., 2014, *The Astronomical Journal*, 147, 122
- van Rossum G., 1995, Report CS-R9526, Python tutorial. pub-CWI, pub-CWI:adr
- von Hippel T., 2005, *ApJ*, 622, 565
- von Hippel T., van Dyk D. A., Stenning D. C., Robinson E.,

## APPENDIX A: COLOR-MAGNITUDE DIAGRAMS FOR OUTLIERS

We present CMDs for the ten OCs referred to as “outliers”, given their  $\log(\text{age}/\text{yr}) > 0.5$  differences between **AStECA** and literature values. In Fig. A1 the CMDs for these OCs are plotted in pairs, two OCs per row.

The left CMD in each pair shows the cluster region with the literature isochrone fit shown in red. For nine of the OCs we fit Marigo et al. (2008) isochrones, since this set of theoretical isochrones were used in all but one of the articles where these OCs were analysed. The exception is SMC-L35 which was fit using Girardi et al. (2002) isochrones, so the same set is used for this OC.

The right CMD shows the same cluster region with the isochrone that generated the best match SC found by **AStECA**, shown in green. The colours of the stars in this CMD correspond to the MPs assigned by the DA, when applied. Semi-transparent stars are those removed by the cell-by-cell density based cleaning algorithm (see Sect. 3.3), also when applied.

Values for the fundamental parameters in both fits are shown to the bottom left of each CMD.

## APPENDIX B: COLOR-MAGNITUDE DIAGRAMS FOR P99, P00, C06, AND G10 DATABASES

We present in this appendix the CMDs of all OCs cross-matched with our own sample, in the databases P99, P00, C06, and G10, i.e.: those that used the isochrone fit method in their analysis.

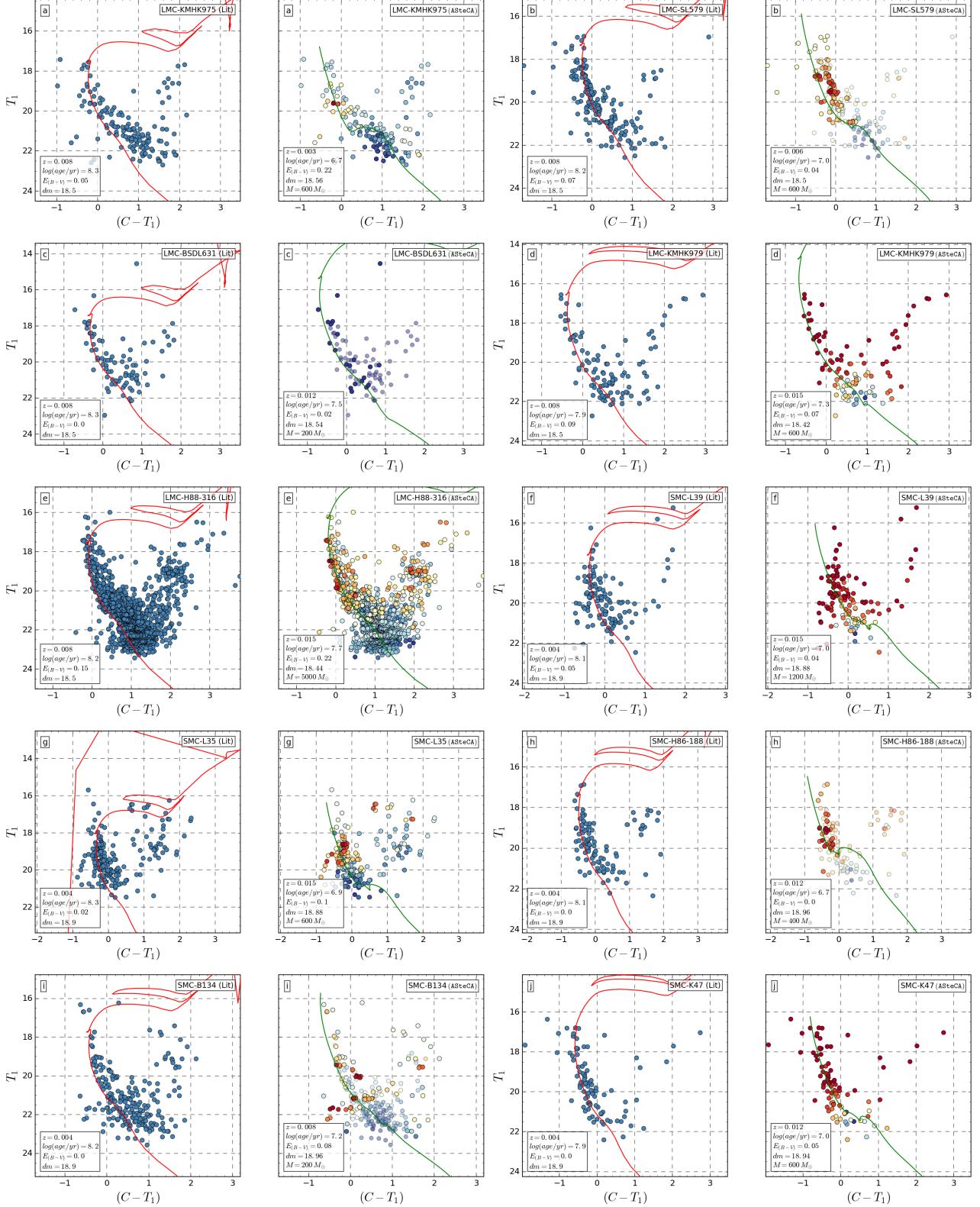
## APPENDIX C: COLOR-MAGNITUDE DIAGRAMS FOR LARGE MASS OCs

We present here CMDs for the five OCs with the largest mass estimates by either the H03 or the P12 database. These are also the OCs with the largest differences in assigned masses when compared with those derived by **AStECA**. See Table 5 for more information about the fundamental parameters assigned to these OCs.

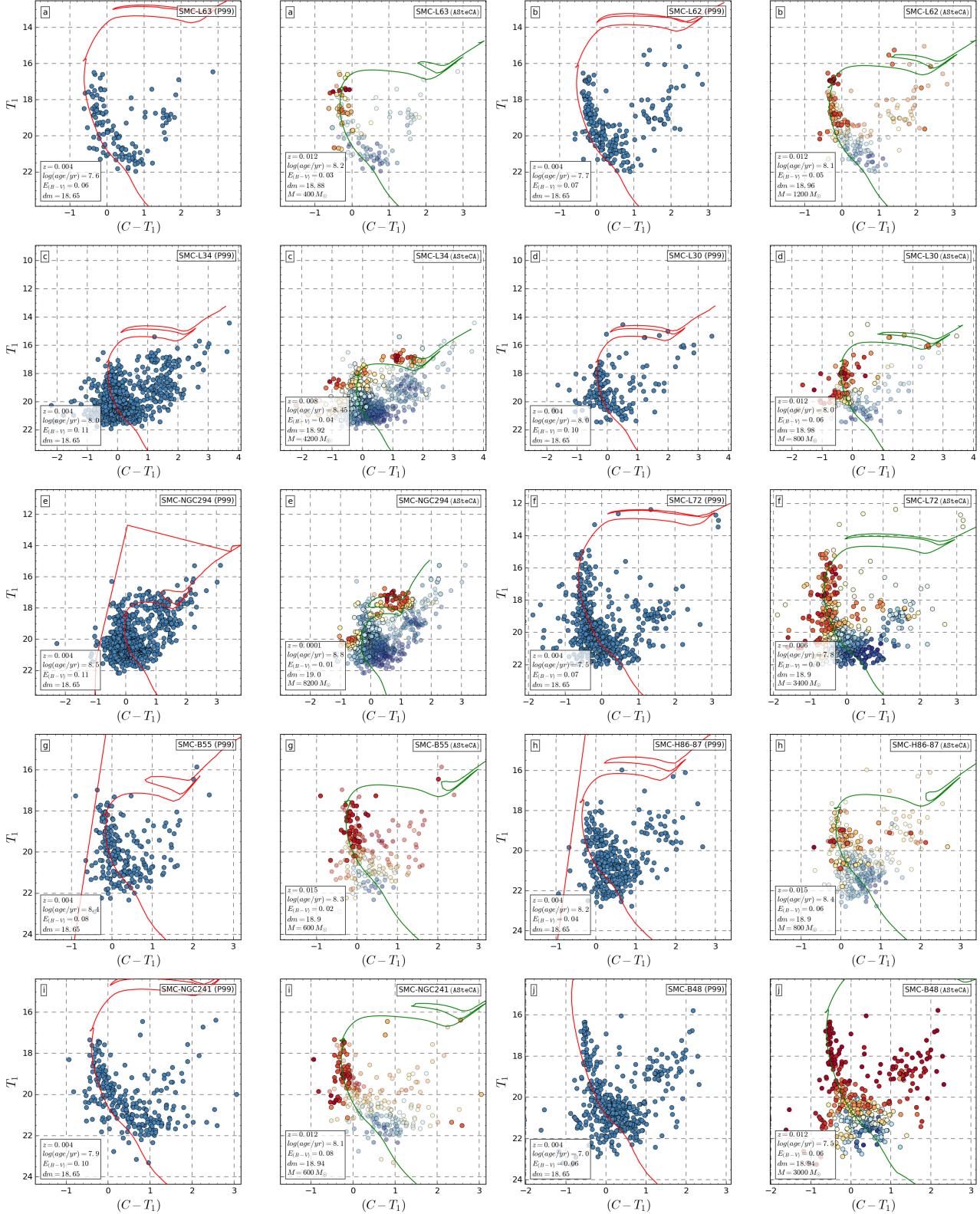
## APPENDIX D: AGE-METALLICITY RELATIONS FOR LITERATURE VALUES

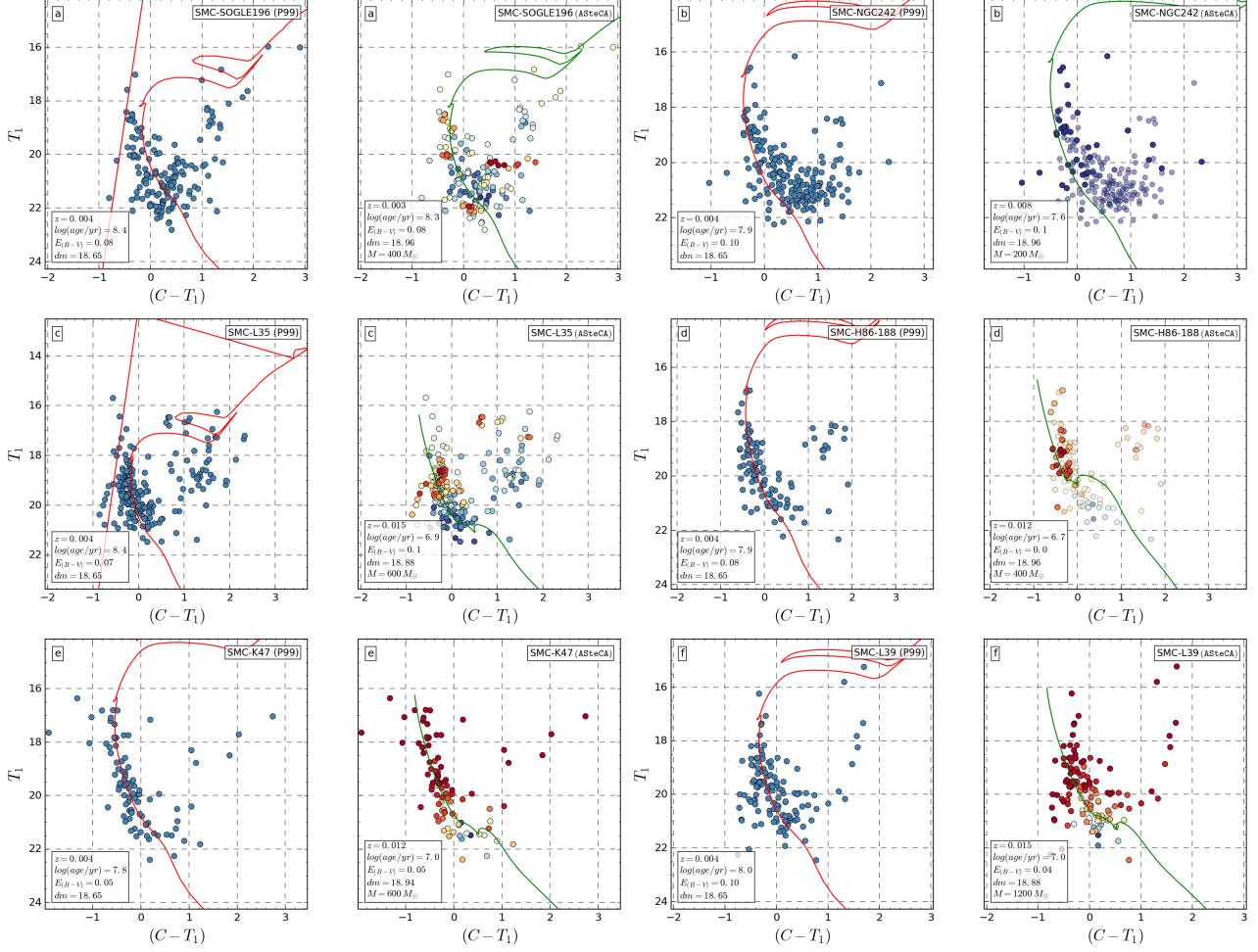
In Fig. D1 the CMDs for the four LMC OCs with metallicities and ages in the vicinity of  $\sim 3.3$  Gyr and  $[\text{Fe}/\text{H}] \approx -0.25$  dex, are shown. These are the OCs responsible for the high metallicity found for the AMR of this galaxy, for ages close to 4 Gyr.

Fig. D2 shows the AMRs for both Clouds, generated using the metallicity and age values taken from those articles presented in Table 1.

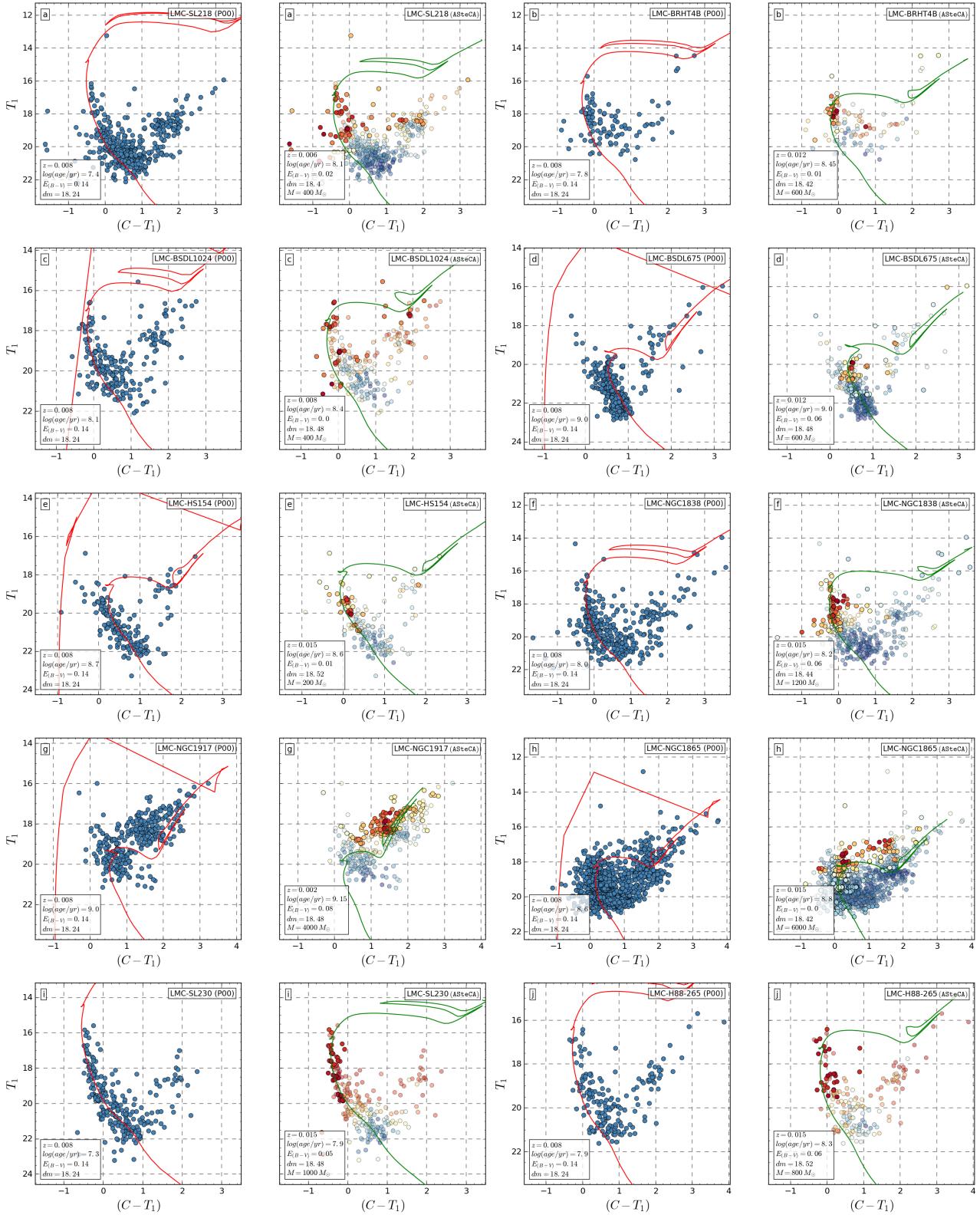


**Figure A1.** CMDs for the outliers set. See description of the plots in Appendix A.

**Figure B1.** CMDs for the XXX database. See description of the plots in Appendix B.



**Figure B2.** CMDs for the XXX database. See description of the plots in Appendix B.



**Figure B3.** CMDs for the XXX database. See description of the plots in Appendix B.

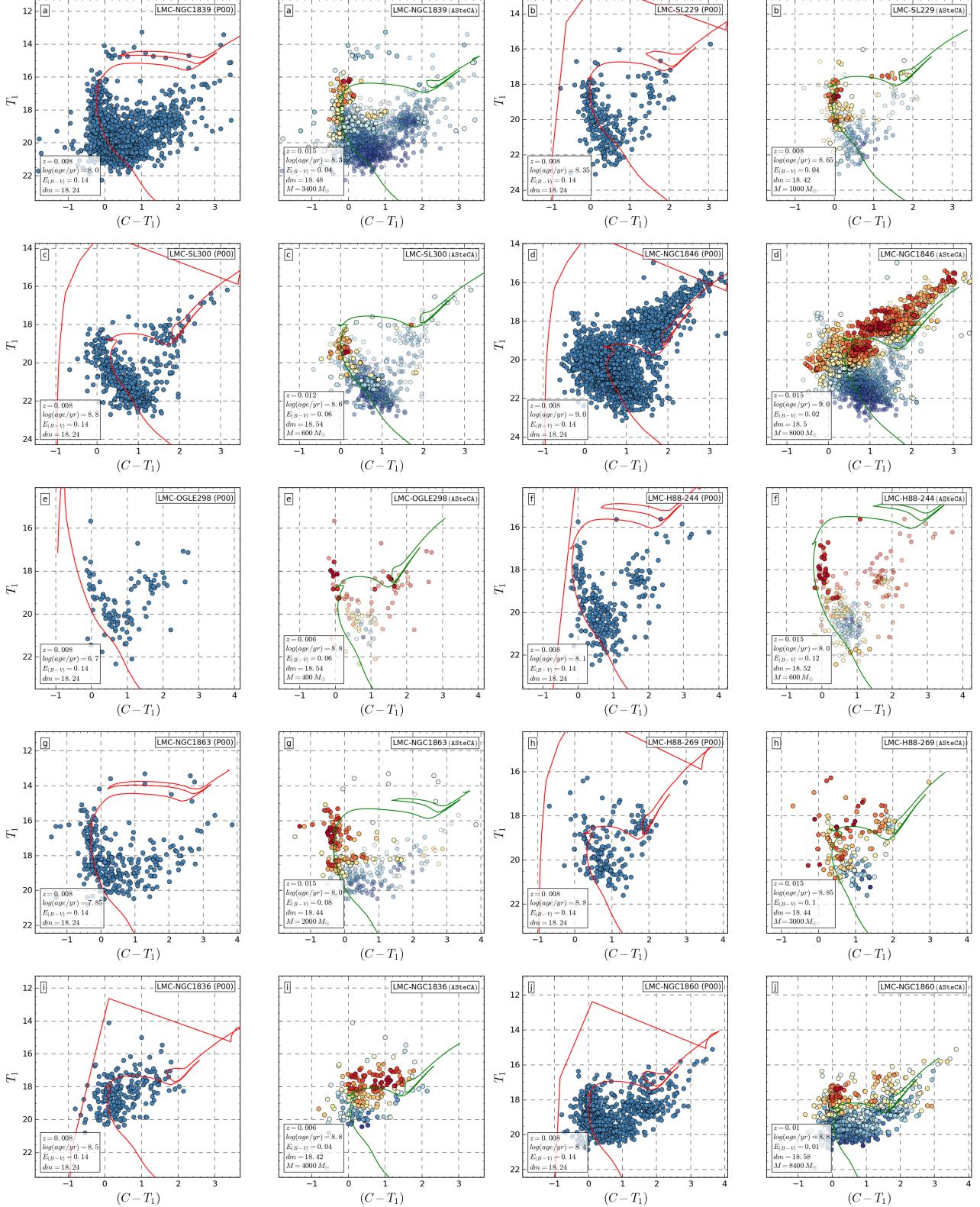
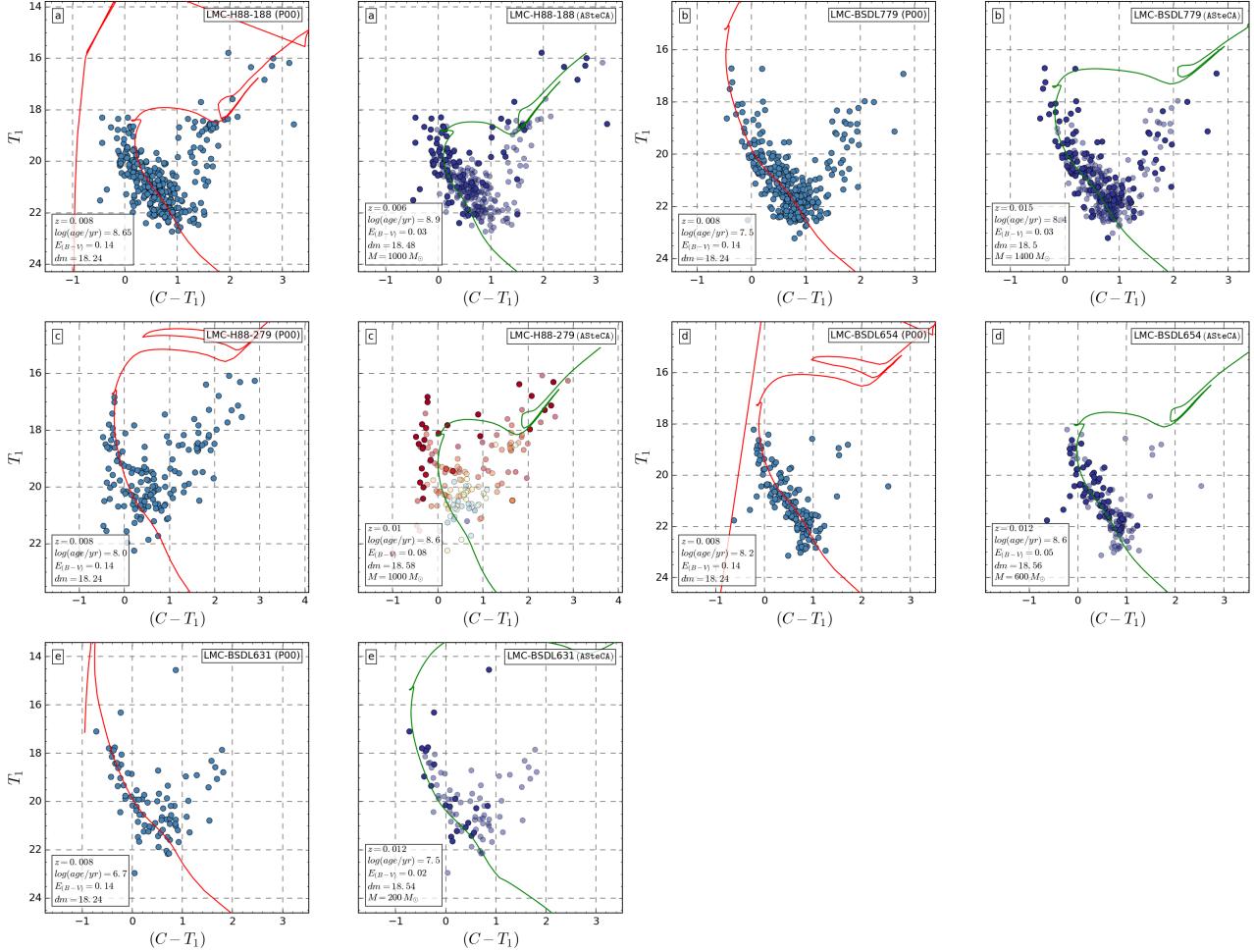
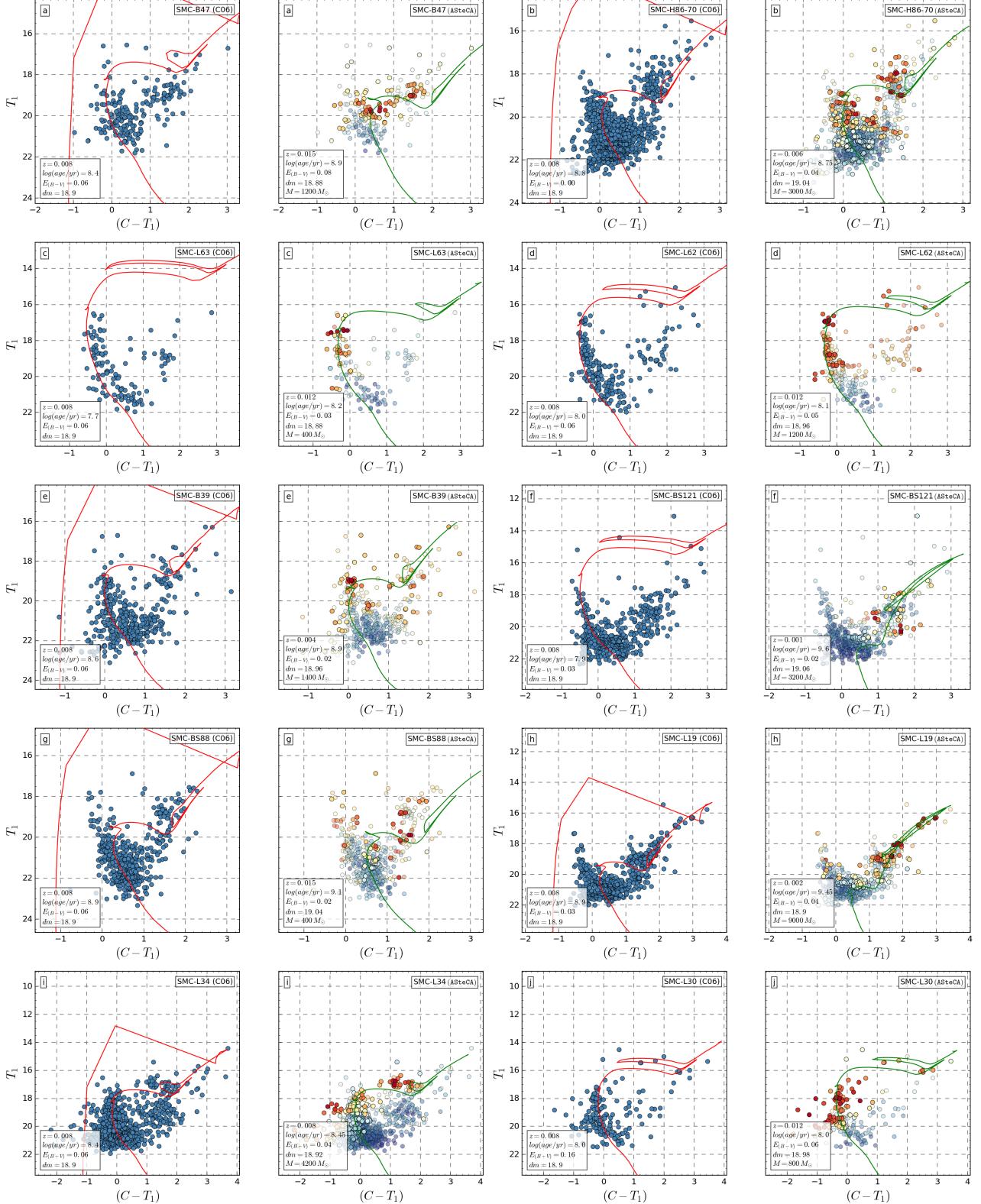


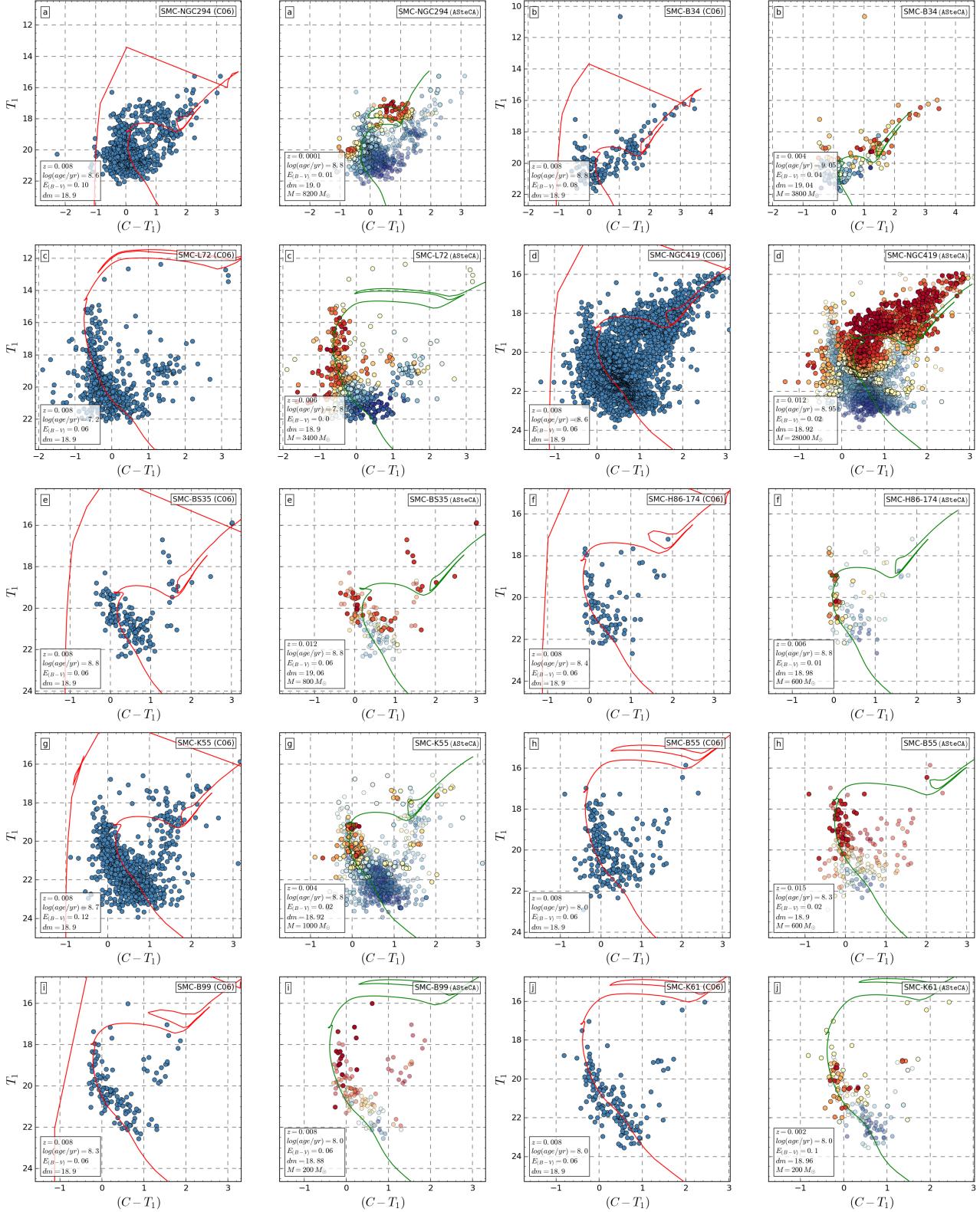
Figure B4. CMDs for the XXX database. See description of the plots in Appendix B.



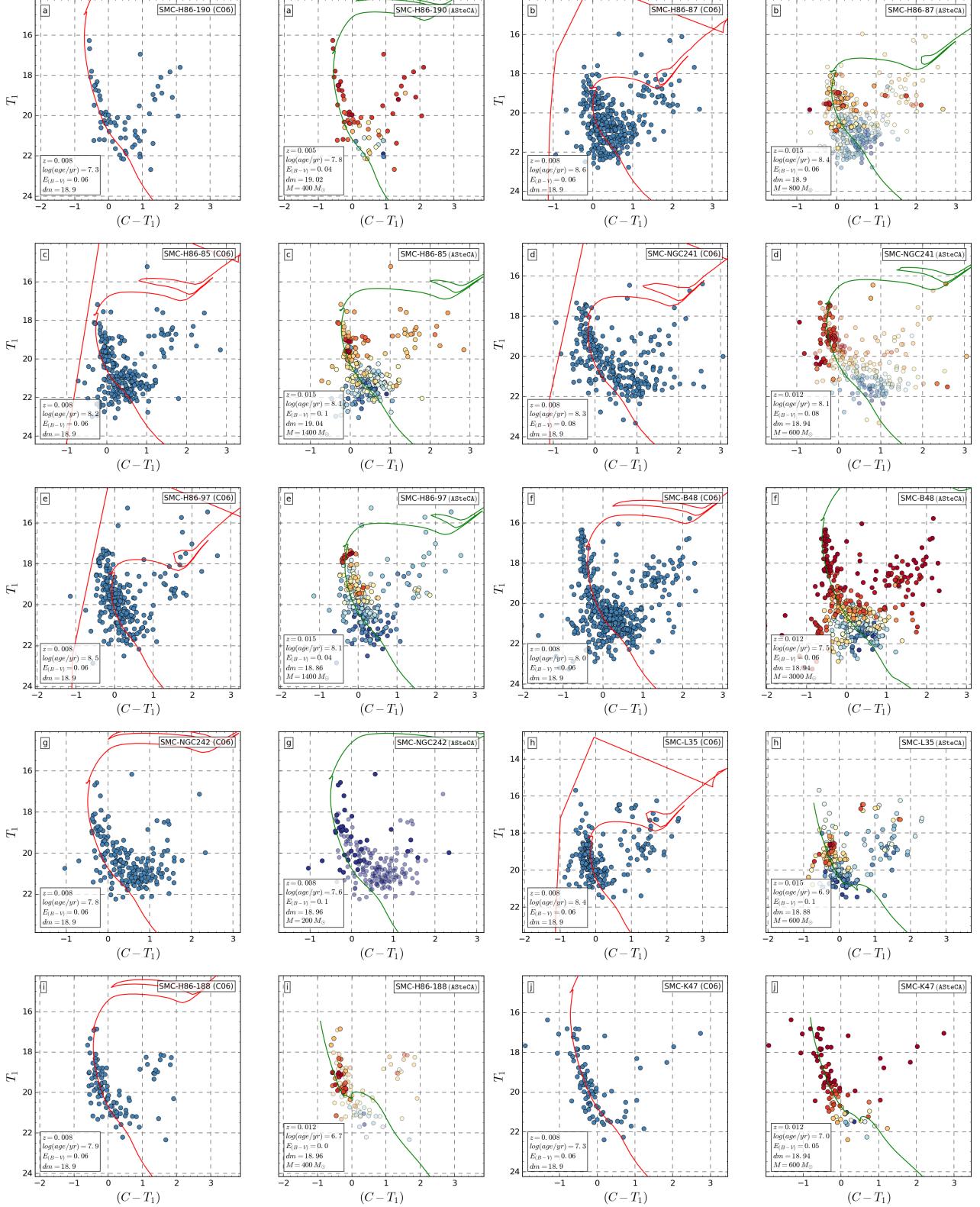
**Figure B5.** CMDs for the XXX database. See description of the plots in Appendix B.



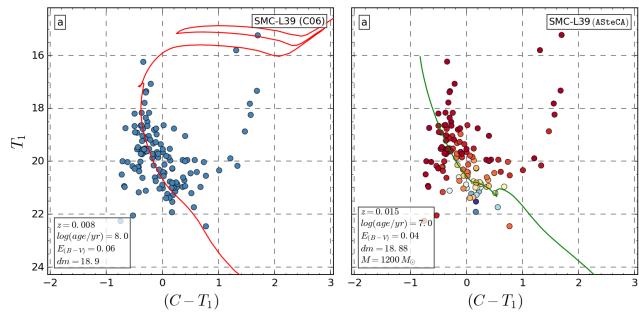
**Figure B6.** CMDs for the XXX database. See description of the plots in Appendix B.



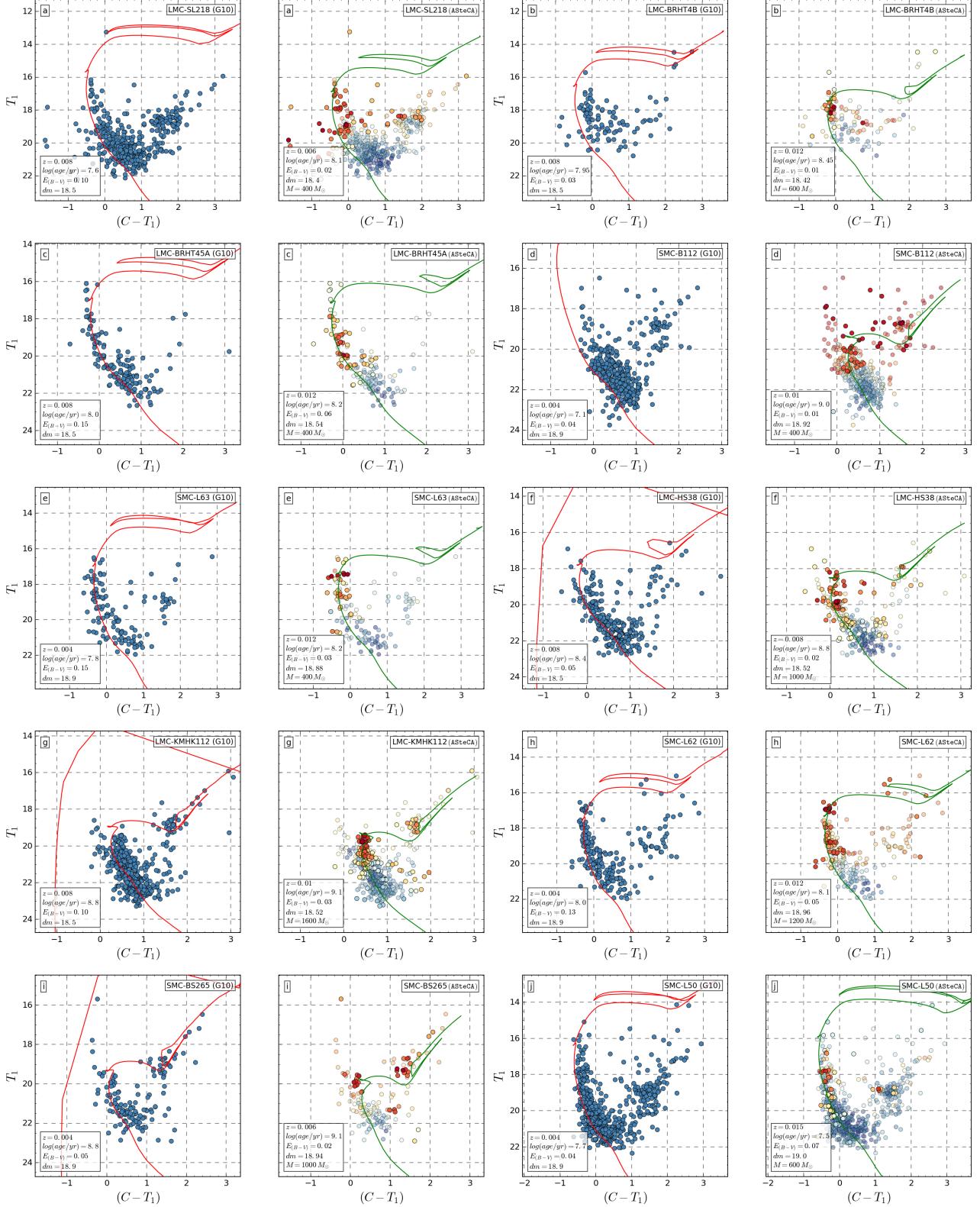
**Figure B7.** CMDs for the XXX database. See description of the plots in Appendix B.



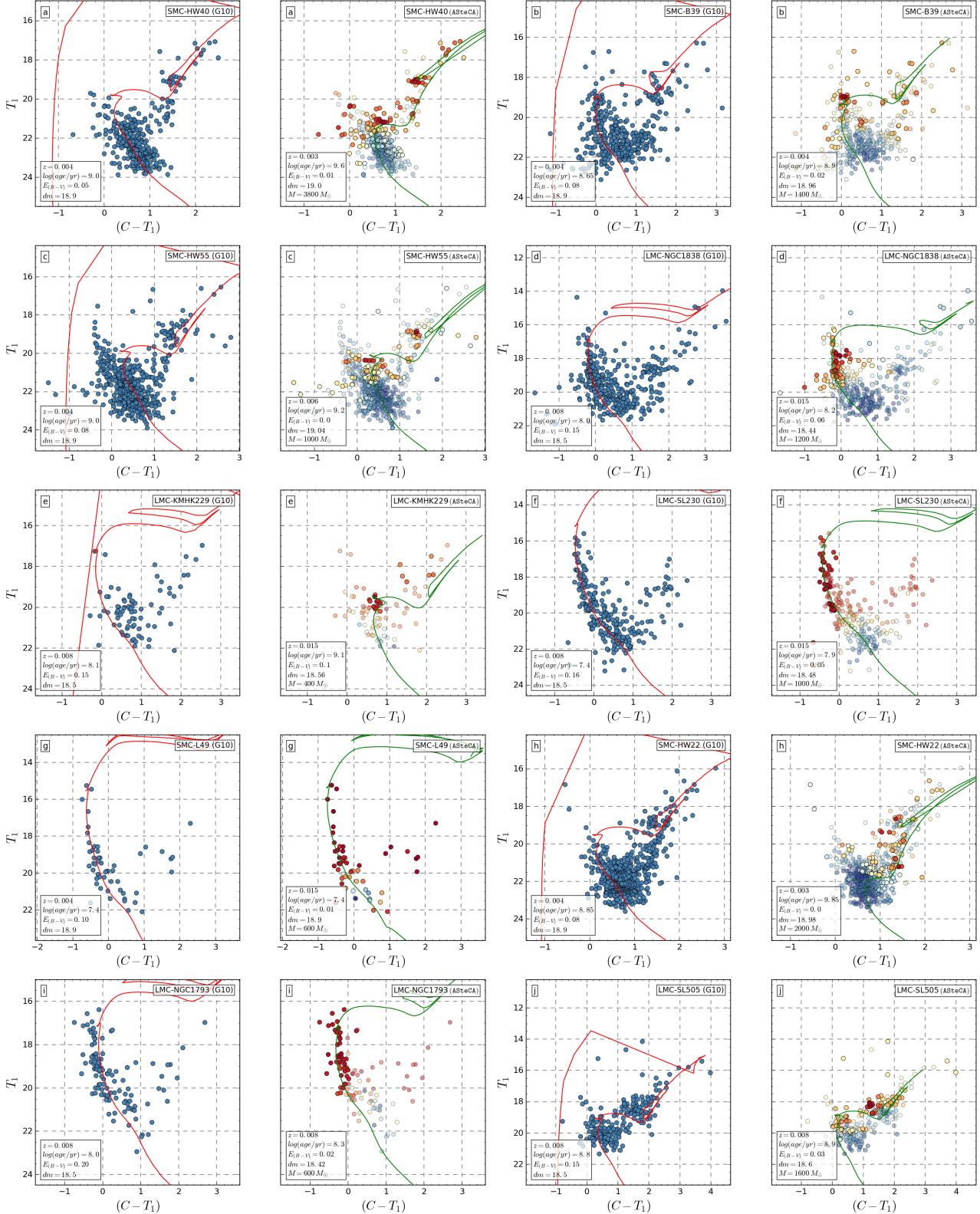
**Figure B8.** CMDs for the XXX database. See description of the plots in Appendix B.



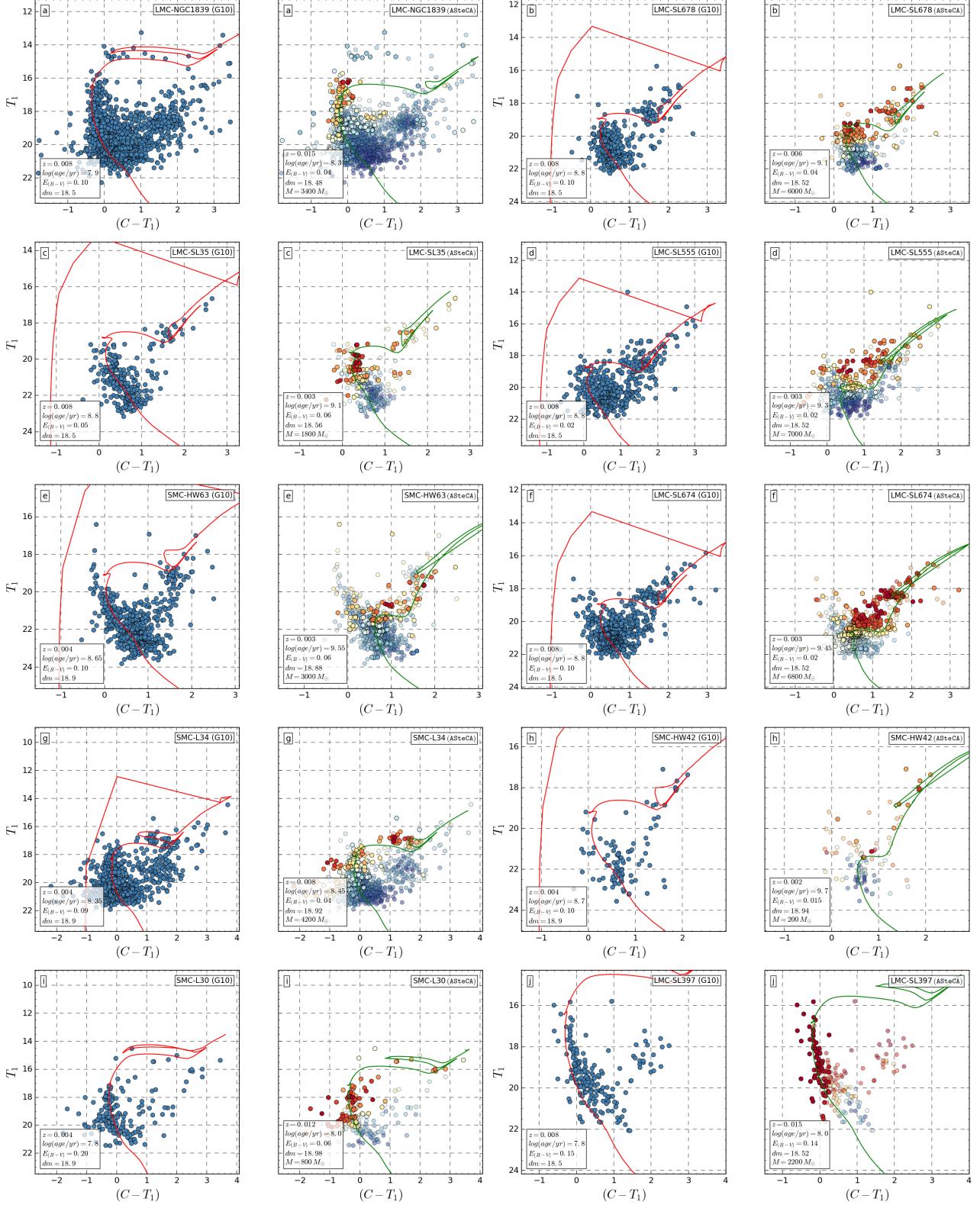
**Figure B9.** CMDs for the XXX database. See description of the plots in Appendix B.



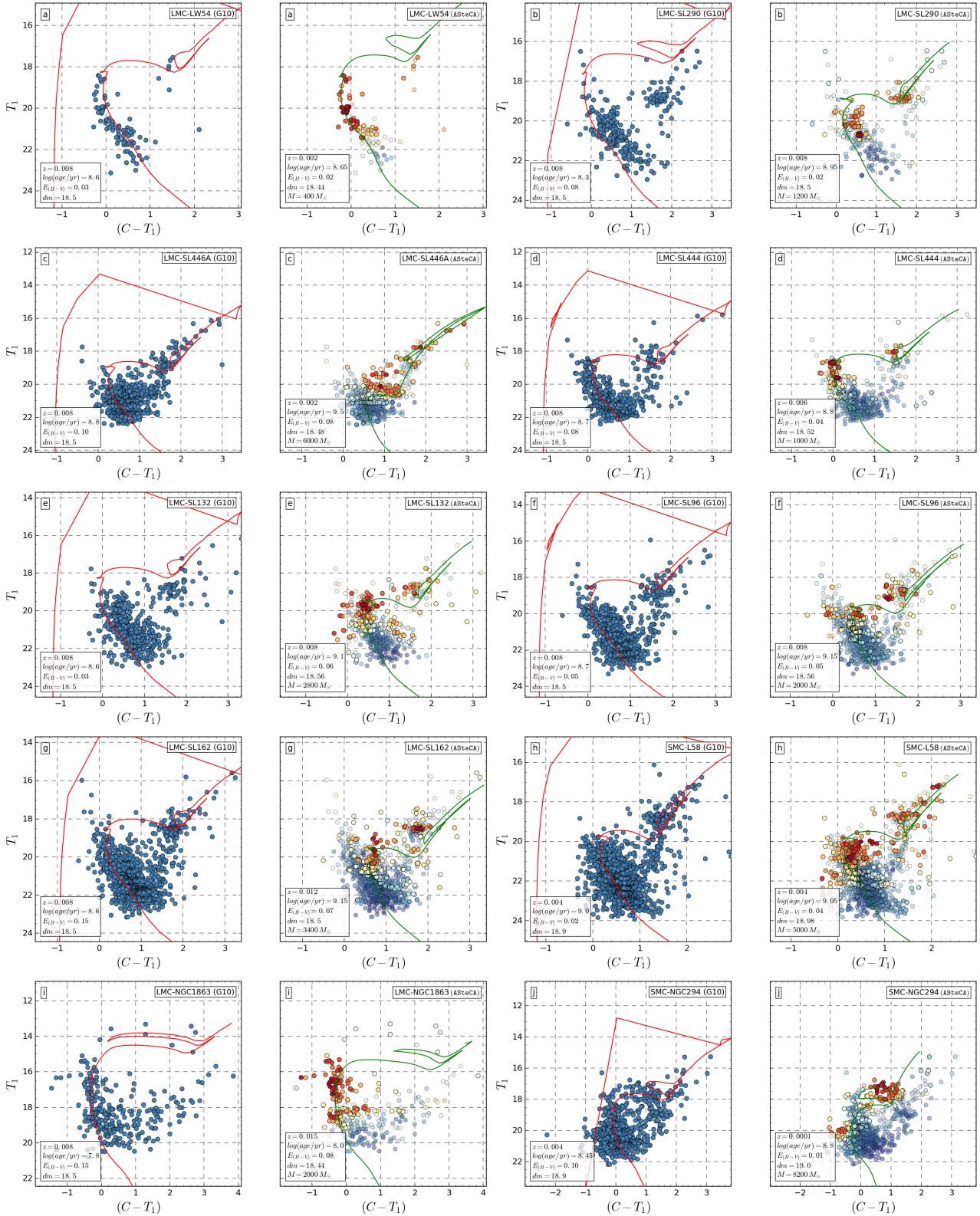
**Figure B10.** CMDs for the XXX database. See description of the plots in Appendix B.



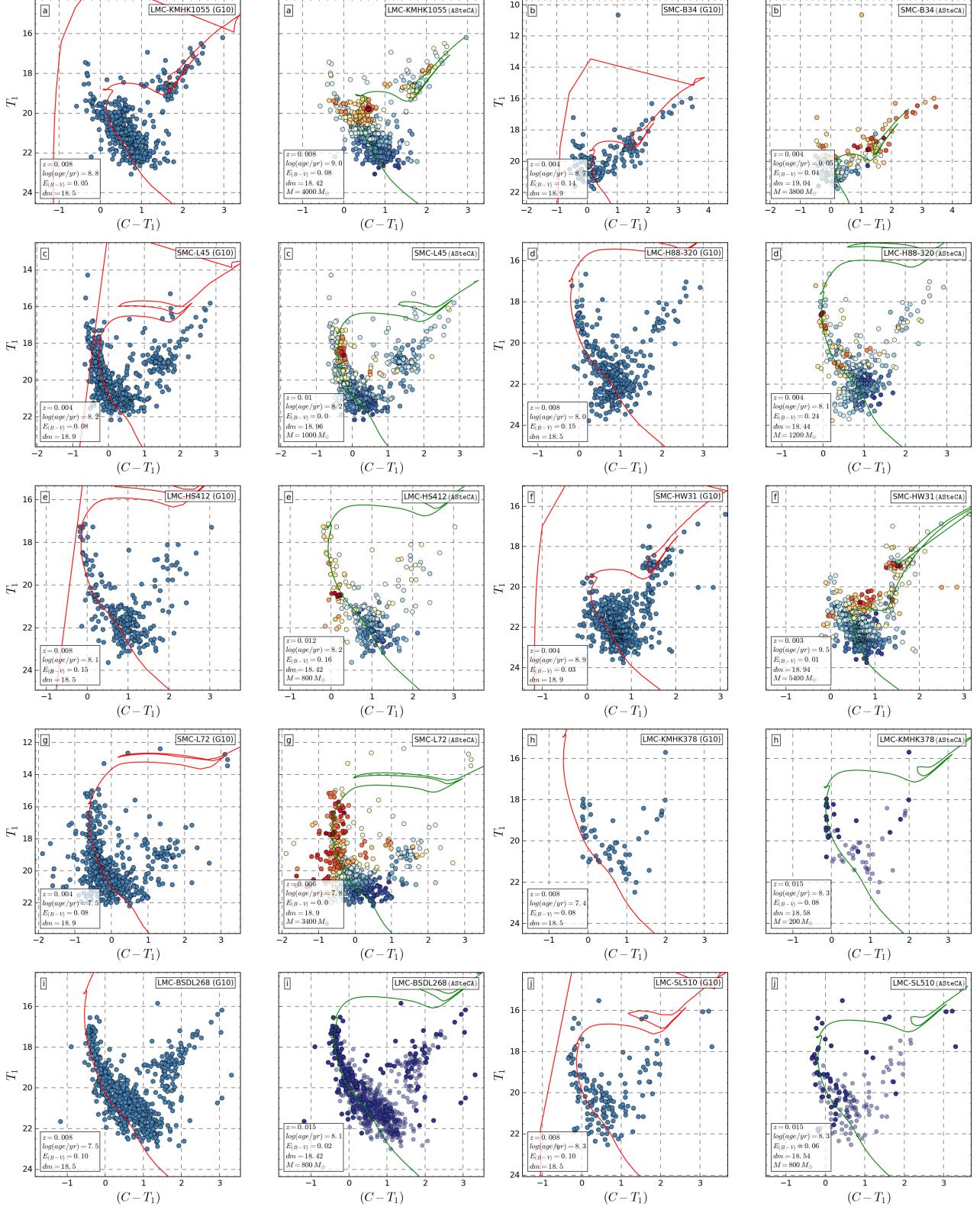
**Figure B11.** CMDs for the XXX database. See description of the plots in Appendix B.



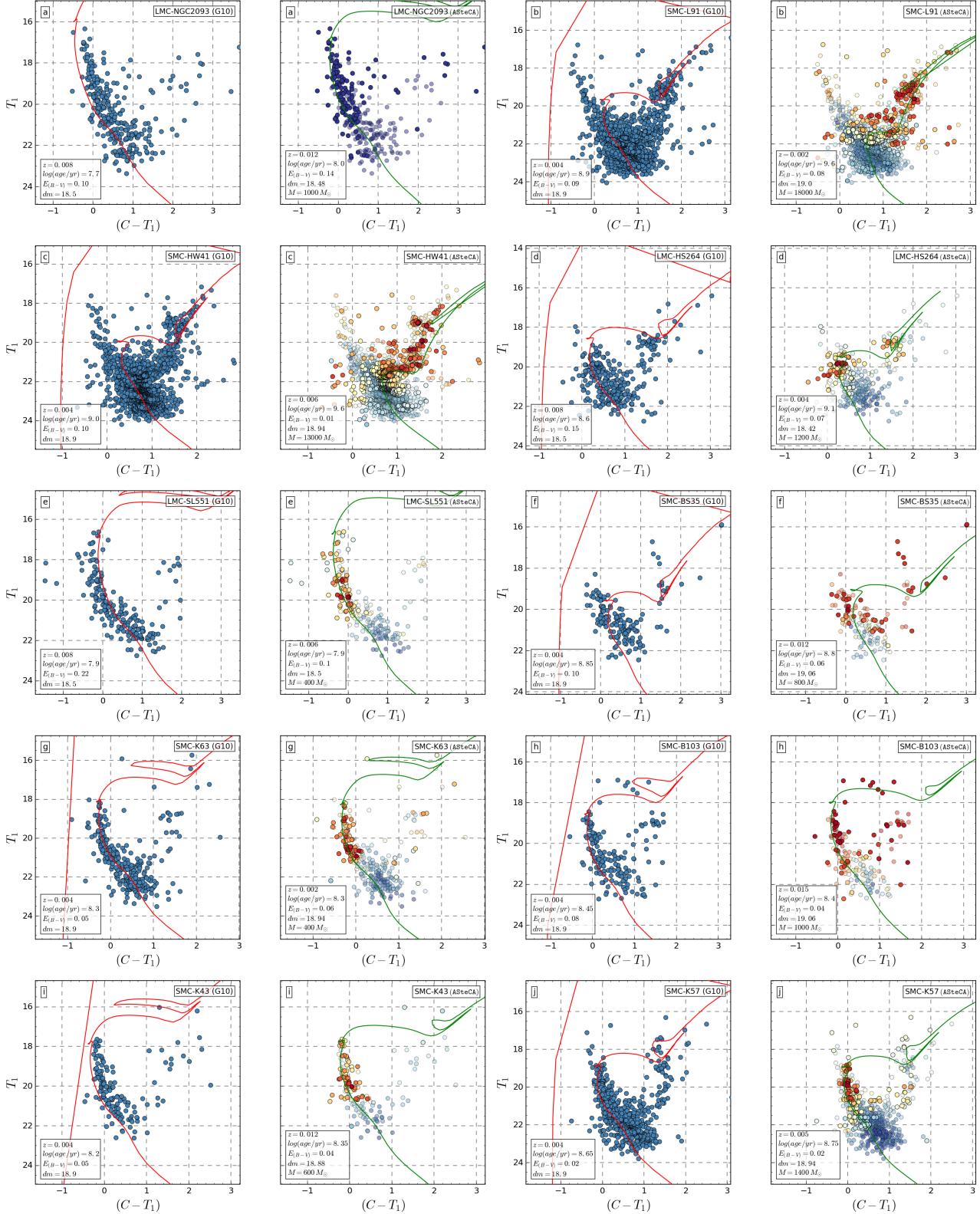
**Figure B12.** CMDs for the XXX database. See description of the plots in Appendix B.



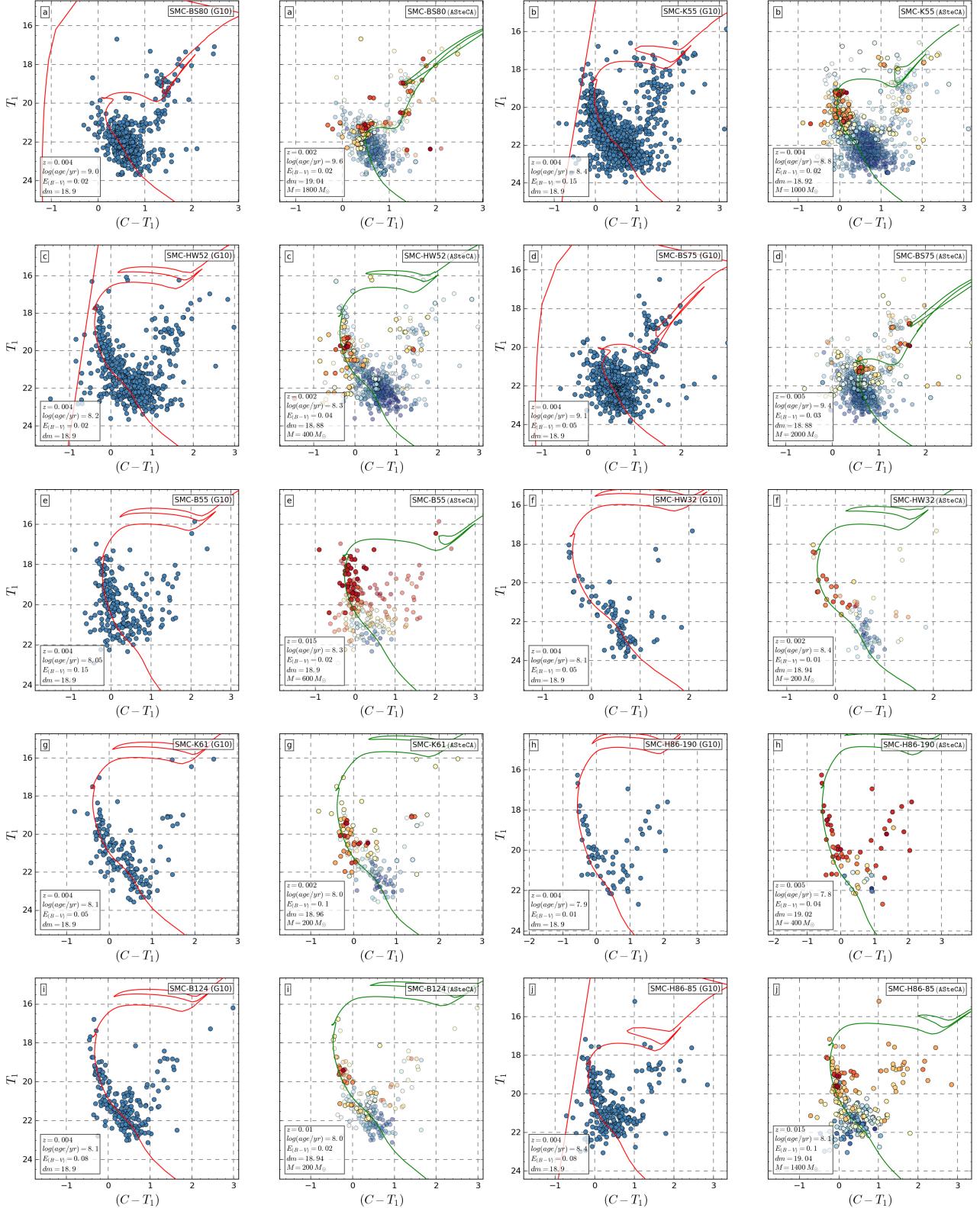
**Figure B13.** CMDs for the XXX database. See description of the plots in Appendix B.



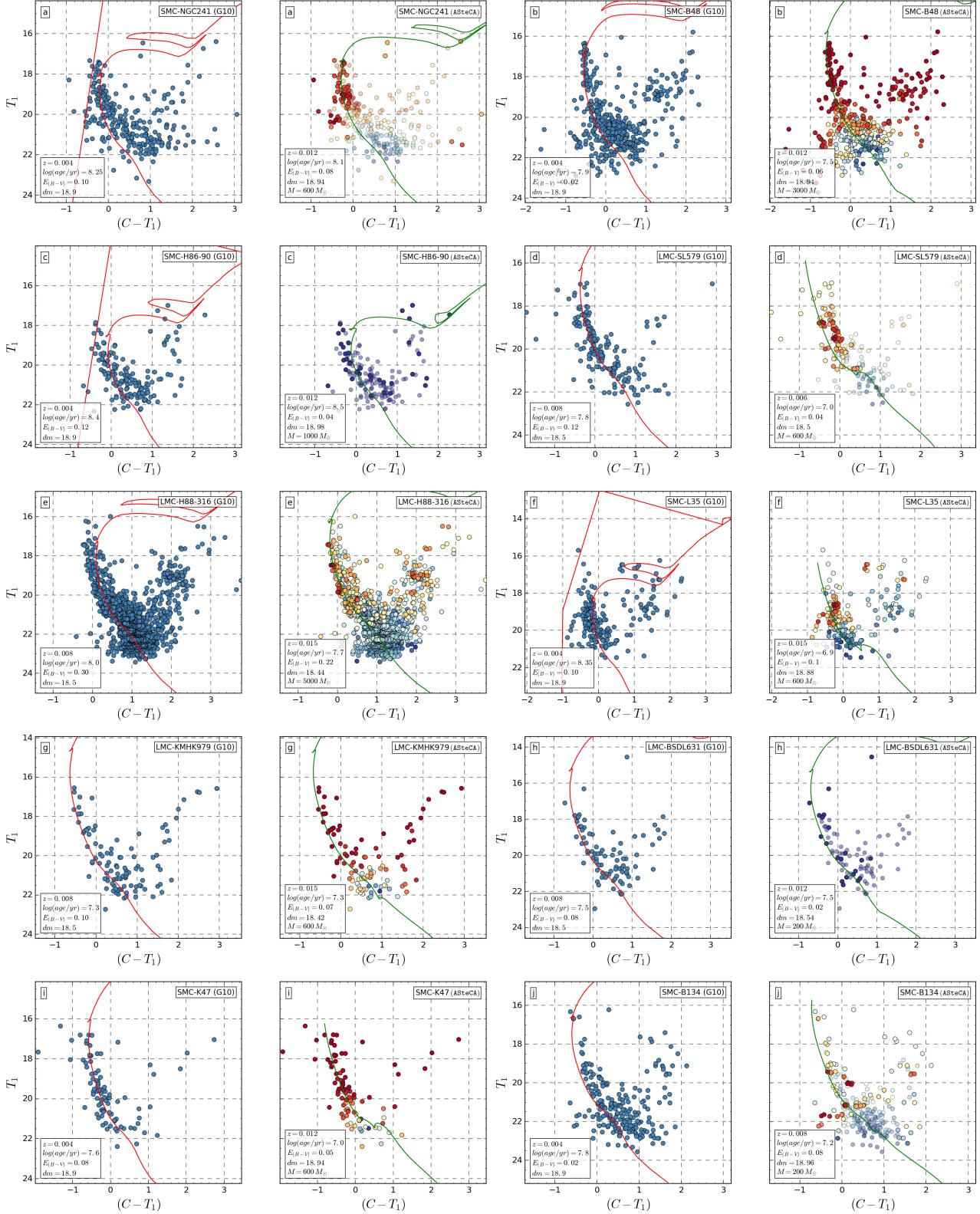
**Figure B14.** CMDs for the XXX database. See description of the plots in Appendix B.



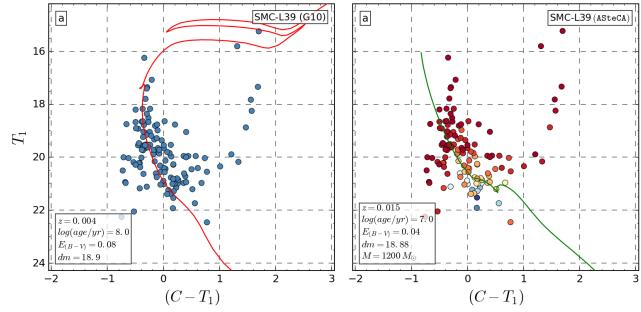
**Figure B15.** CMDs for the XXX database. See description of the plots in Appendix B.



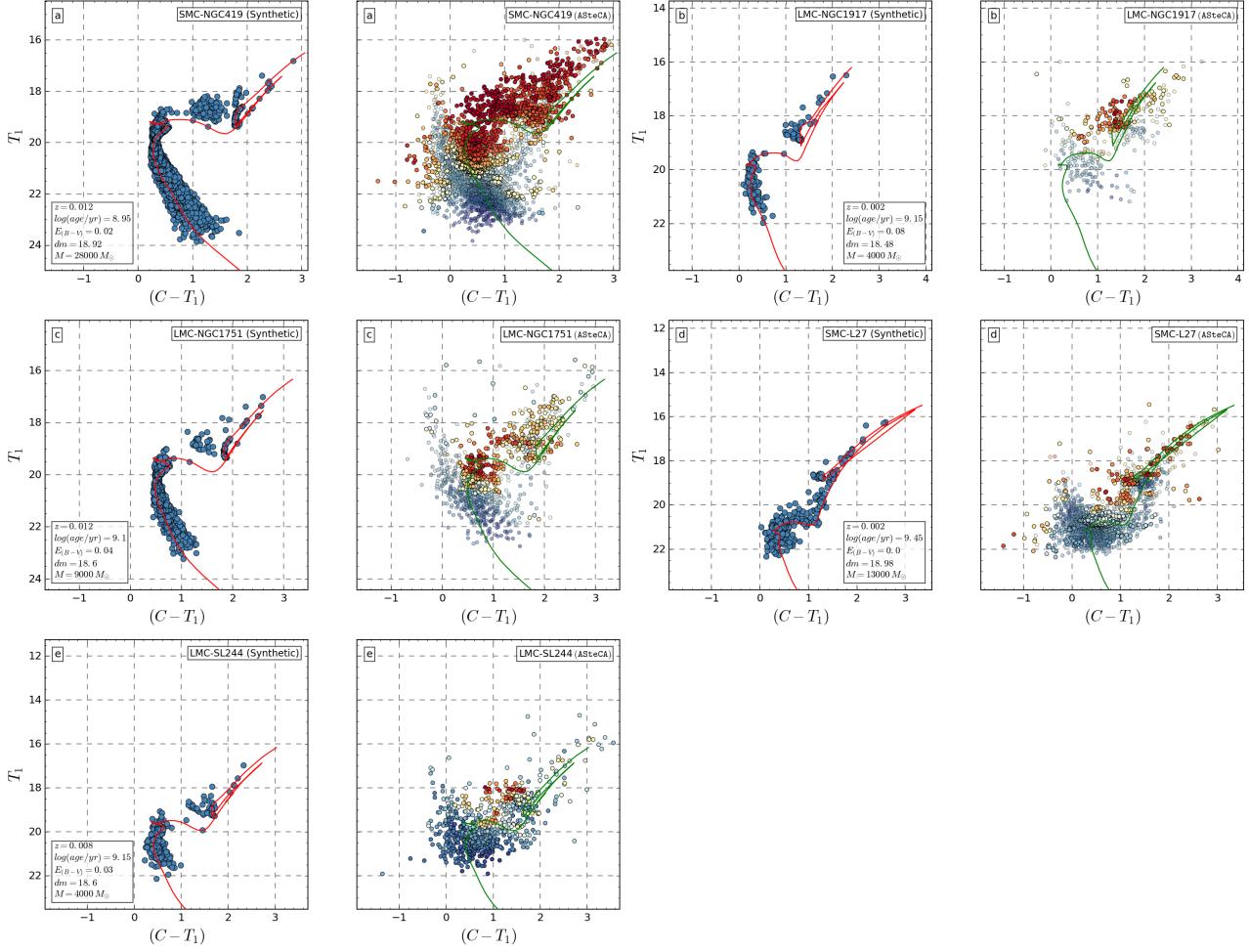
**Figure B16.** CMDs for the XXX database. See description of the plots in Appendix B.



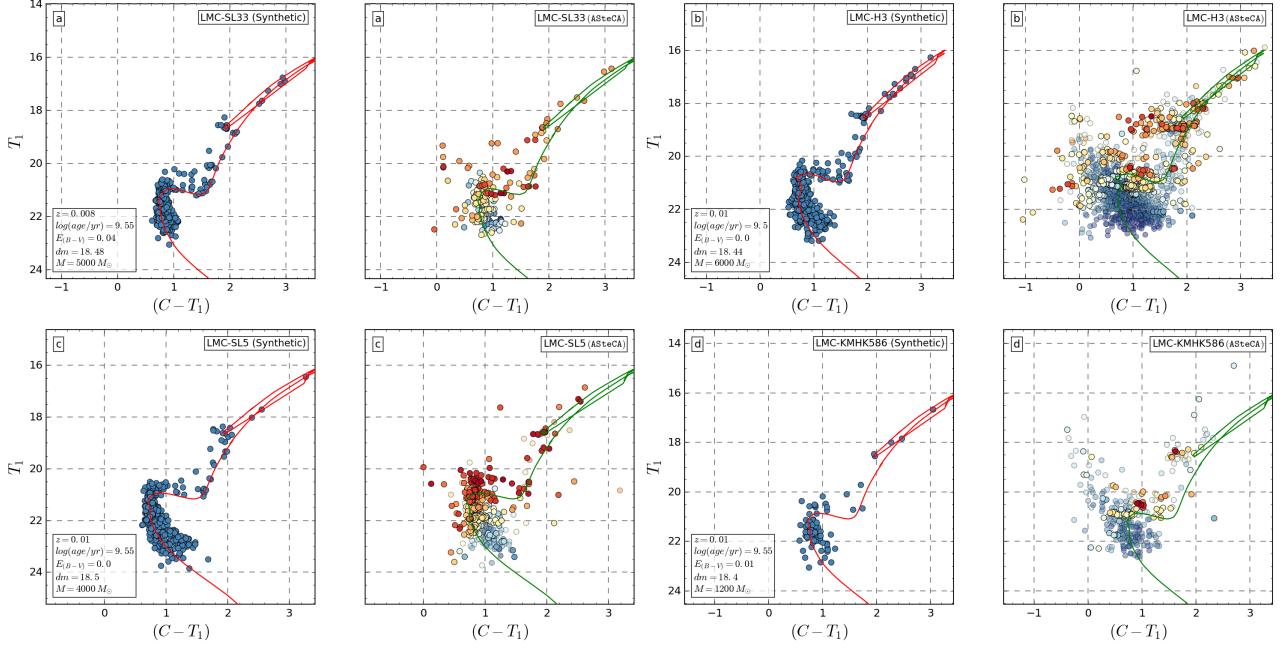
**Figure B17.** CMDs for the XXX database. See description of the plots in Appendix B.



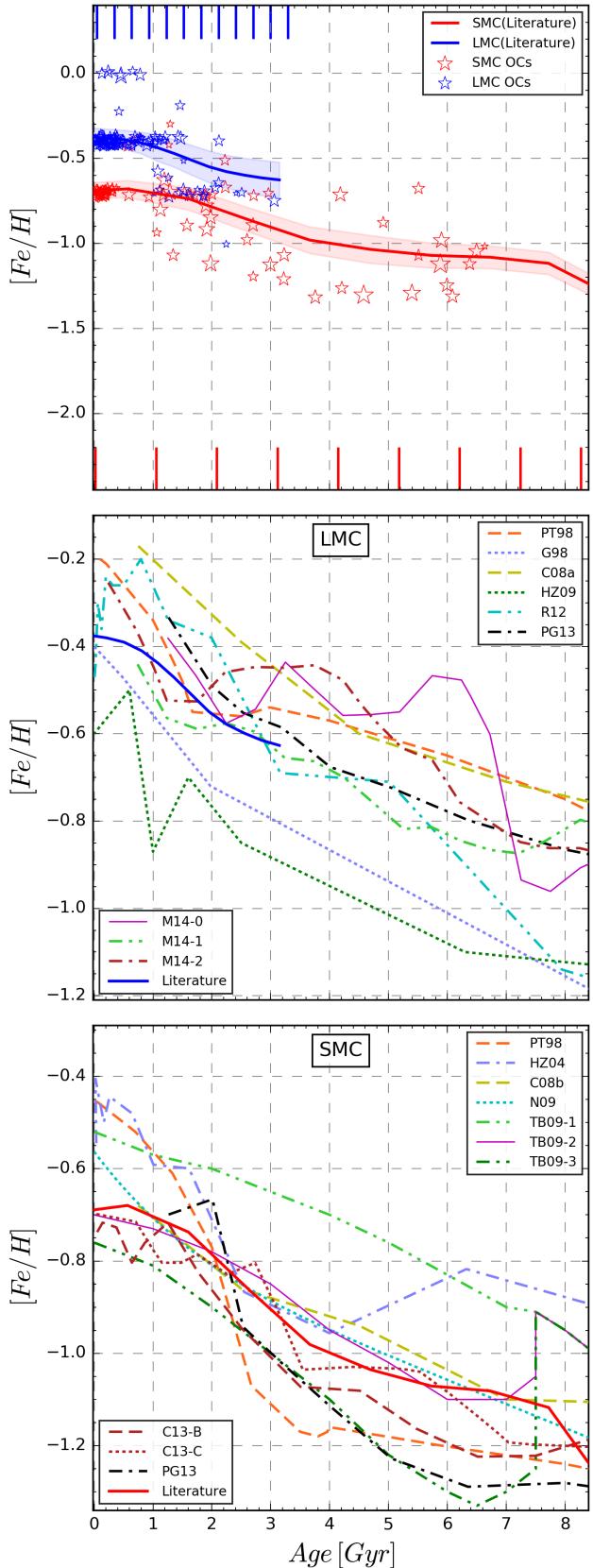
**Figure B18.** CMDs for the XXX database. See description of the plots in Appendix B.



**Figure C1.** CMDs for the OCs with the largest masses assigned by H03 and/or P12. The best match synthetic cluster is plotted to the right, and the observed cluster region CMD to the left, for each OC.



**Figure D1.** CMDs for the OCs that cause a peak in the AMR for the LMC, for ages close to 4 Gyr. The best match synthetic cluster is plotted to the right, and the observed cluster region CMD to the left, for each OC.

**Figure D2.** XXX