Resolved photometry of extragalactic young massive star clusters[*][**]

S. S. Larsen[1], S. E. de Mink[2][***], J. J. Eldridge[3], N. Langer[4], N. Bastian[5], A. Seth[6], L. J. Smith[2,9],
J. Brodie[7], and Yu. N. Efremov[8]

[1] Astronomical Institute, University of Utrecht, Princetonplein 5, 3584 CC Utrecht, The Netherlands
e-mail: s.s.larsen@astro-uu.nl
[2] Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA
[3] Institute of Astronomy, The Observatories, University of Cambridge, Madingley Road, Cambridge, CB3 0HA, UK
[4] Argelander-Institut für Astronomie, Bonn University, Auf dem Hügel 71, 53121 Bonn, Germany
[5] Excellence Cluster Universe, Boltzmannstr. 2, 85748 Garching, Germany
[6] Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA
[7] UCO/Lick Observatory, University of California, Santa Cruz, CA 95064, USA
[8] Sternberg Astronomical Institute, MSU, Moscow 119992, Universitetsky pr. 13, Russia
[9] European Space Agency, Research and Scientific Support Department, Baltimore, MD 21218, USA

ABSTRACT

Aims. We present colour–magnitude diagrams (CMDs) of young massive star clusters in several galaxies located well beyond the Local Group. The richness of these clusters allows us to obtain large samples of post-main sequence stars and test how well the observed CMDs are reproduced by canonical stellar isochrones.

Methods. We use imaging of seven clusters in the galaxies NGC 1313, NGC 1569, NGC 1705, NGC 5236 and NGC 7793 obtained with the Advanced Camera for Surveys on board the *Hubble* Space Telescope and carry out PSF-fitting photometry of individual stars in the clusters. The clusters have ages in the range $\sim(5-50) \times 10^6$ years and masses of $\sim 10^5\ M_\odot$–$10^6\ M_\odot$. Although crowding prevents us from obtaining photometry in the inner regions of the clusters, we are still able to measure up to 30–100 supergiant stars in each of the richest clusters. The resulting CMDs and luminosity functions are compared with photometry of artificially generated clusters, designed to reproduce the photometric errors and completeness as realistically as possible.

Results. In agreement with previous studies, our CMDs show no clear gap between the H-burning main sequence and the He-burning supergiant stars, contrary to predictions by common stellar isochrones. In general, the isochrones also fail to match the observed number ratios of red-to-blue supergiant stars, although the difficulty of separating blue supergiants from the main sequence complicates this comparison. In several cases we observe a large spread (1–2 mag) in the luminosities of the supergiant stars that cannot be accounted for by observational errors. We find that this spread can be reproduced by including an age spread of $\sim(10-30) \times 10^6$ years in the models. However, age spreads cannot fully account for the observed morphology of the CMDs and other processes, such as the evolution of interacting binary stars, may also play a role.

Conclusions. Colour–magnitude diagrams can be successfully obtained for massive star clusters out to distances of at least 4–5 Mpc. Comparing such CMDs with models based on canonical isochrones we find several areas of disagreement. One interesting possibility is that an age spread of up to $\sim$30 Myr may be present in some clusters. The data presented here may provide useful constraints on models for single and/or binary stellar evolution.

Key words. open clusters and associations: general – galaxies: star clusters: general – galaxies: spiral – Hertzsprung-Russell and C-M diagrams

1. Introduction

Stars profoundly affect the properties of the galaxies they inhabit. They are the primary sites of nucleosynthesis and drivers of chemical evolution and they contribute to the energy budget of the interstellar medium through feedback processes such as supernovae and stellar winds. In most galaxies they also deliver most of the ultraviolet, optical and near-infrared light. In order to correctly interpret observations of galaxies and constrain their past evolution (e.g., in terms of star formation histories and chemical evolution) it is therefore essential to strive towards the best possible understanding of the properties of individual stars. This is true not only for integrated light measurements, but also for modelling of colour–magnitude diagrams (CMDs) of resolved stellar populations (Gallart et al. 2005).

Although stars spend most of their lifetimes in the core-hydrogen burning phase on the main sequence (MS), post-MS stars provide a significant fraction of the bolometric luminosity of a stellar population at all except the very youngest (few Myr) ages (e.g., Renzini & Buzzoni 1986). Intermediate- and high mass post-MS stars contribute significantly to the energy output

---

of young stellar populations, and the modelling of such stars is particularly uncertain. One example is the relative numbers of blue and red core-He burning stars, which is very sensitive to assumptions about input parameters such as mass loss and rotation (Langer & Maeder 1995; Maeder & Meynet 2000). Models are not yet capable of reproducing trends in the red/blue ratio e.g., as a function of metallicity seen in star clusters (Eggenberger et al. 2002), although somewhat better agreement is found when comparing trends in this ratio versus *age* (Dohm-Palmer et al. 1997; Dohm-Palmer & Skillman 2002). Another issue concerns the extent of the "blue loops" in the core-He burning phase, which depends on the detailed structure and chemical profile of a star. These, in turn, are affected e.g., by the treatment of convective overshooting (Matraka et al. 1982; Ritossa 1996), with the net effect that more overshooting tends to lead to less extended (redder) blue loops. For standard assumptions, models tend to underpredict the extent of the blue loops. While most stellar evolution models predict a gap in the CMD between the H and He burning stars (the "Blue Hertzsprung Gap" or BHG; Chiosi 1998; Salasnich et al. 1999), the post-MS radius evolution is very uncertain for massive stars and star cluster CMDs often show significant numbers of stars in the BHG (e.g., Fischer et al. 1993). The nature of these stars is poorly understood and it is unclear whether they are even core H or He-burning (Vink et al. 2010), although in some cases their CNO abundances suggest that they have undergone a first dredge-up as red supergiants (Salasnich et al. 1999). Some models do predict more extended blue loops, but this requires that a very large fraction of the H-rich envelope is lost in the red supergiant stage (Hirschi et al. 2004; Salasnich et al. 1999). This may not be compatible with the fact that SN 1987A still had a massive H-rich envelope when it exploded as a blue supergiant (Woosley 1988). Complicating matters further a large fraction of massive stars are members of binary systems (Sana et al. 2008; Mason et al. 2009), many of which may interact or even merge during a phase of common envelope evolution (Taam & Sandquist 2000; Eldridge et al. 2008; Sansom et al. 2009).

Historically, star clusters have long served as important testbeds for stellar models. A prime reason for this has been the assumption that all stars in a cluster are very nearly coeval, so that a cluster CMD can be directly compared against a model isochrone. Furthermore, under this assumption all post-MS stars in a given cluster will have very similar (initial) masses, so that the step of interpolation in the tracks to produce isochrones can often be eliminated for comparison with observed CMDs or calculation of integrated properties (Maraston 1998). Models of low-mass stars in particular have been extensively tested by comparison with globular cluster CMDs (Renzini & Fusi Pecci 1988; Palmieri et al. 2002; Salaris et al. 2007), in some cases providing evidence of multiple stellar populations (Piotto 2009). Similarly, high precision photometry has shown that the CMDs of several intermediate-age ($\sim$1–2 Gyr) clusters in the Magellanic Clouds are not well fitted by standard isochrones, but instead show a significant spread in the main sequence turn-off magnitudes and luminosities of red clump stars. One possible explanation for these features is a significant age spread in these clusters (Mackey et al. 2008; Milone et al. 2009), although other effects (such as stellar rotation) may also play a role (Bastian & de Mink 2009; but see also Girardi et al. 2011). However, interpretation of cluster CMDs remains easier than the complex mix of stellar populations that make up the field, and clusters remain attractive targets for comparison with stellar models. In fact, such a comparison may yield important information on possible
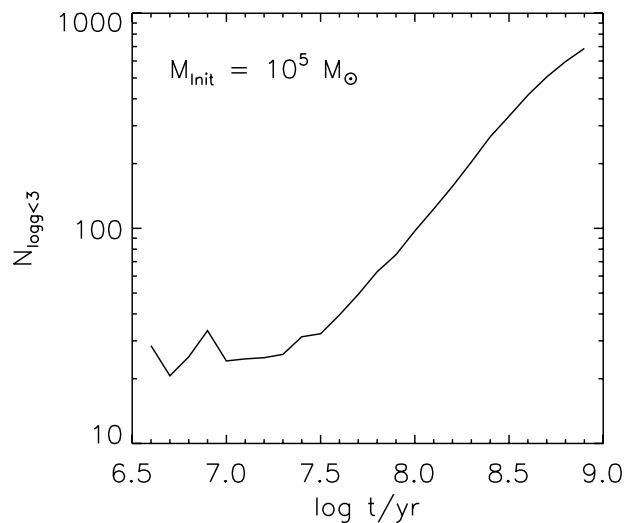
Fig. 1. Number of stars with $\log g < 3.0$ for an initial total stellar mass of $10^5 M_\odot$.

age spreads within clusters, in addition to providing constraints on the stellar models.

One difficulty is that a typical low-mass young (open) cluster contains no or only a few post-MS stars. As an illustration, Fig. 1 shows the number of stars with $\log g < 3.0$ vs. age, predicted for a single-age stellar population with an initial mass of $10^5 M_\odot$. We have assumed a Salpeter IMF over the mass range $0.15 < M/M_\odot < 100$ and used Padua isochrones (Bertelli et al. 2009) to obtain $\log g$ values versus age and (initial) stellar mass. The $\log g$ limit adopted in this figure comfortably includes all local supergiants (Lyubimkov et al. 2010). The figure shows that, for ages less than about 30 Myr, only 20–30 such stars are expected per $10^5 M_\odot$, with an increase at older ages. The figure is consistent with observations of post-MS stars in massive star clusters, e.g., of the rich LMC cluster NGC 1866 which has 100 post-MS stars, an age of about 160 Myr and a mass of about $1.3 \times 10^5 M_\odot$ (Fischer et al. 1992; Barmina et al. 2002), or the "Red Supergiant Cluster 2" (age $\sim$10 Myr, initial mass $\sim$40 000 $M_\odot$) which hosts 26 red supergiant stars (RSGs; Davies et al. 2007). It is thus clear that cluster masses of at least several times $10^4 M_\odot$ are required in order to get significant samples of massive post-MS stars. Such clusters are rare in the Milky Way and even in the Local Group, making it difficult to obtain samples that cover a range in age (or, equivalently, main sequence turn-off mass), metallicity, etc.

If the search volume is extended beyond the boundaries of the Local Group several galaxies hosting young clusters with $M \gtrsim 10^5 M_\odot$ become available. The trade-off is that photometry of individual stars becomes extremely challenging due to severe crowding, even if imaging from the *Hubble* Space Telescope (HST) is used. One of the closest examples of a galaxy hosting such massive, young star clusters is NGC 1569 at $\sim$3.4 Mpc (Grocholski et al. 2008). Using imaging obtained with the High Resolution Channel (HRC) of the Advanced Camera for Surveys (ACS) on board HST, Larsen et al. (2008) constructed a CMD for the cluster NGC 1569-B and identified about 60 RSGs. This demonstrates that photometry can indeed be obtained for individual stars in clusters well beyond the Local Group, and motivated us to search for additional candidates suitable for such analysis. Several other galaxies located within distances of a few

Table 1. Data for the clusters analysed in this paper.

Cluster	RA and Dec (2000.0)		Distance (Mpc)	V (mag)	A_B (mag)	log age (years)	Mass ($M_\odot$)	M_{TO} ($M_\odot$)	*FWHM*	η
NGC 1569-A	04 30 48.08	+64 50 57.2	3.4[a]	14.6	2.30[f]	6.7	7.6×10^5	~25	0″057/0.9 pc	1.12
NGC 1569-B	04 30 48.88	+64 50 51.2	3.4[a]	15.4	2.30[f]	7.2	1.4×10^6	~12.0	0″084/1.4 pc	1.04
NGC 1705-1	04 54 13.48	−53 21 39.4	5.1[b]	14.7	0.04[g]	7.1	1.1×10^6	~12.5	0″036/0.9 pc	1.0
NGC 1313-F3-1	03 17 47.76	−66 30 18.7	4.1[c]	17.4	0.47[g]	7.74	2.1×10^5	5.9	0″19/3.8 pc	0.83
NGC 5236-F1-1	13 37 01.37	−29 50 49.2	4.5[d]	17.2	0.28[g]	7.44	1.9×10^5	7.9	0″14/3.1 pc	1.09
NGC 5236-F1-3	13 37 04.67	−29 50 35.2	4.5[d]	18.1	0.28[g]	7.45	8.4×10^4	7.9	0″15/3.3 pc	1.13
NGC 7793-F1-1	23 57 35.88	−32 35 40.9	3.3[e]	17.4	0.08[g]	7.69	9.6×10^4	6.3	0″23/3.7 pc	1.49

Notes. [a] Grocholski et al. (2008) ; [b] Tosi et al. (2001) ; [c] Méndez et al. (2002) ; [d] Thim et al. (2003) ; [e] Carignan (1985) ; [f] Israel (1988) ; [g] Schlegel et al. (1998) .

Mpc have ACS imaging, obtained as part of our own and other programmes. We have inspected these images and looked for clusters that appeared sufficiently well resolved into individual stars that photometry would be feasible. In this paper we present the resulting CMDs and compare with theoretical isochrones. Our aim is to quantify whether or not the observations can be well approximated by single isochrones.

2. Data reduction and analysis

2.1. Selection of the sample

Even with the spatial resolution offered by HST, we cannot resolve individual stars in clusters more distant than a few Mpc. For this work, we inspected images taken as part of our Cycle 12 programme (GO-9774, P.I. S. S. Larsen). We found potentially suitable clusters in three out of five galaxies observed in this programme (NGC 1313, NGC 5236 and NGC 7793), while no suitable clusters were found in the two other galaxies (NGC 45, NGC 4395). The images obtained under this programme were all taken with the wide field channel (WFC) of the ACS, which has a pixel scale of 0″050 pixel⁻¹. At the typical distances of ~ 4 Mpc, this corresponds to about 1 pc pixel⁻¹ so it is clear that we cannot obtain photometry in the cluster cores but only in the outer regions. The filters used for the ACS/WFC observations were *F435W, F555W* and *F814W*, roughly equivalent to the Johnson-Cousins *BVI* filters. For a detailed description of the data, see Mora et al. (2009). In addition, we include images of the clusters NGC 1569-A, NGC 1569-B and NGC 1705-1, taken with the ACS/HRC in the filters *F330W, F555W* and *F814W* as part of programme GTO-9300 (P.I.: H. Ford). The *F330W* data are less suitable for stellar photometry due to lower S/N (especially for the heavily reddened NGC 1569), and we only make use of the *F555W* and *F814W* photometry in this paper. We will from now on refer to the ACS *F330W, F435W, F555W* and *F814W* bands as *U*, *B*, *V* and *I*, although we do not actually transform the photometry to the standard Johnson-Cousins system.

Table 1 lists basic data for the clusters we have examined in detail. Coordinates were measured on the ACS images; we caution that the absolute HST astrometry is only accurate to 1″–2″. Ages for the clusters in NGC 1313, NGC 5236 and NGC 7793 were estimated from integrated (ground-based) *UBV* photometry (Larsen 1999), using the empirical calibration of *UBV* colour vs. age in Girardi et al. (1995) which we note is strictly valid only for LMC-like metallicity. The clusters in NGC 1569 and NGC 1705 are significantly younger than those in the spirals, and age estimates from integrated light are more uncertain for these clusters due to the rapid, metallicity-dependent changes in integrated colours at very young ages. For NGC 1705-1 we

initially adopt an age of 12 Myr, based on STIS spectroscopy (Vazquez et al. 2004). For NGC 1569-A we assume an age of 5 Myr (Origlia et al. 2001; Westmoquette et al. 2007) and for NGC 1569-B we assume 15 Myr (Anders et al. 2004; Larsen et al. 2008). The *V* magnitudes in Table 1 were measured on the *F555W* images in an aperture with a radius of 2″. The cluster masses were determined as a byproduct of the synthetic cluster experiments described below (Sect. 2.3). Note that our mass estimate for NGC 1569-B is significantly higher than the photometric mass estimate of ~ $5.5 \times 10^5 M_\odot$ of Larsen et al. (2008) for this cluster; this is mainly due to the greater distance assumed here (3.4 Mpc vs. 2.2 Mpc). We also list the main sequence turn-off mass (M_{TO}) corresponding to the estimated age of each cluster, as indicated in the Padua isochrone tables (Bertelli et al. 2009). Since these refer to the bluest point reached during the core H-burning phase, the most massive stars still present in a cluster will be somewhat more massive.

2.2. Photometry

PSF-fitting photometry was done with the DAOPHOT/ALLFRAME package (Stetson 1994). First, the FIND task in DAOPHOT was run on 1000×1000 pixels sections of the *F555W* images, centered on the clusters, and aperture photometry was obtained in all three bands with the PHOT task. Point-spread functions (PSFs) were constructed for each band based on typically 10–20 stars, and ALLFRAME was then run on all three images simultaneously to produce PSF-fitting magnitudes. In a second iteration of this process, improved PSFs were generated by subtracting all stars except the PSF stars from the images and redetermining the PSFs, and the FIND task was applied once again on the star-subtracted *F555W* images. Finally, ALLFRAME was run with the improved PSFs and combined object lists from the first and second iterations as input. The photometry was calibrated by obtaining aperture photometry of the PSF stars on the cleaned images and applying (small) offsets to the PSF magnitudes to make them match the aperture photometry. The photometry was calibrated to the VEGAMAG system using zero-points published on the ACS web pages. We corrected for foreground extinction according to Schlegel et al. (1998), except for NGC 1569 where we assumed $A_B = 2.30$ mag (Israel 1988).

Photometry was extracted in a circular region with a radius of 2″ around each cluster, excluding the region within 5 pixels of the centre (0″25 for the WFC and 0″125 for the HRC) where crowding gets too severe to obtain useful photometry. Given that we are generally measuring stars located in the outskirts of the clusters, we expect some contamination of the cluster colour–magnitude diagrams by field stars. In order to correct for this,
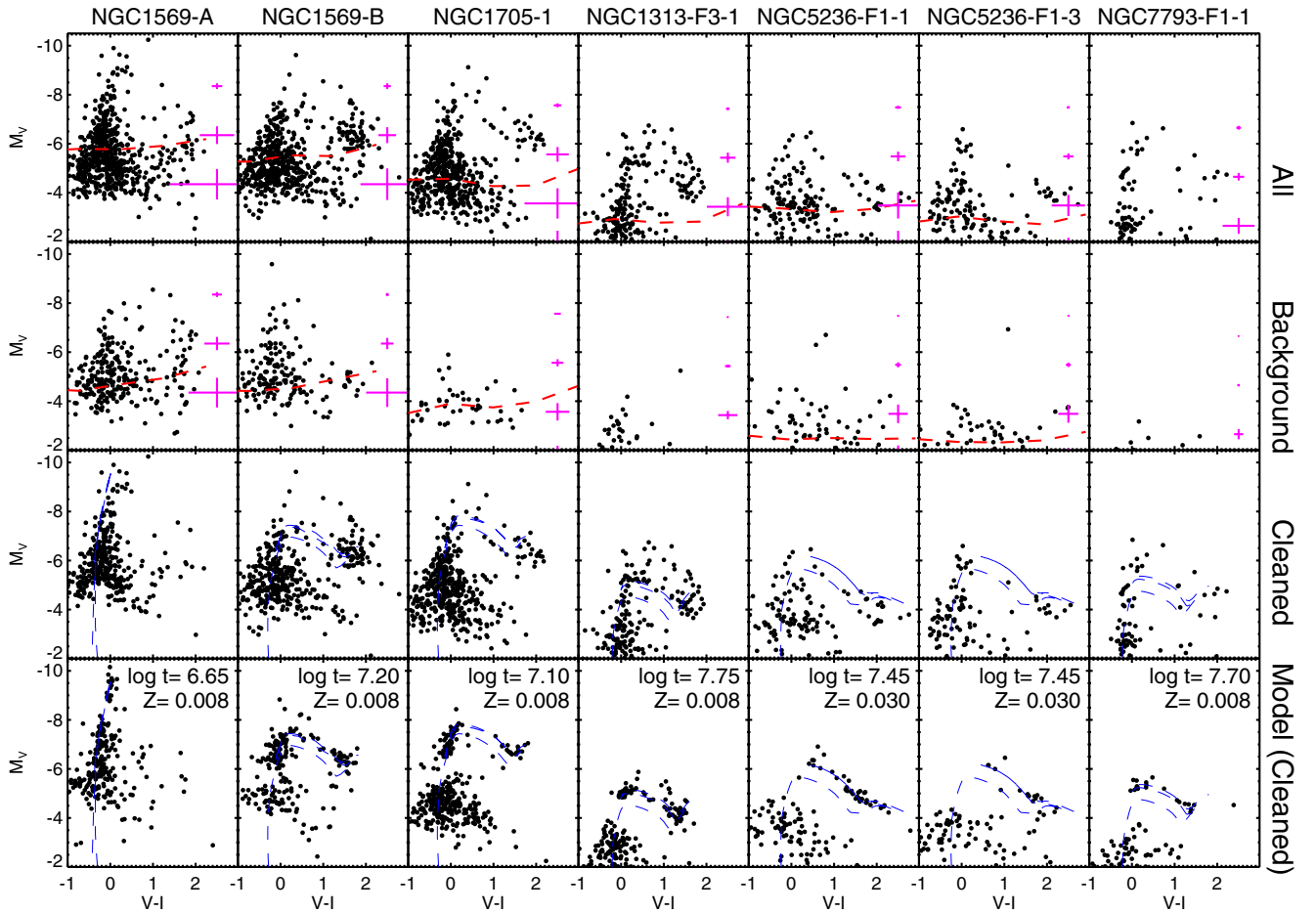
Fig. 2. *Top row*: observed $(V - I)$, M_V colour–magnitude diagrams for the seven clusters, showing all stars for which photometry was obtained. The red dashed curves indicate the approximate 50% completeness limits, while error bars show typical photometric errors. *Second row from top*: Colour–magnitude diagrams for background annuli. *Third row*: Background-subtracted CMDs (see text for details). *Bottom row*: simulated CMDs, using Bertelli et al. (2009) isochrones for the ages and metallicities listed in the legends.

we used a large comparison annulus with an inner radius of typically 7″.5 and with a similar area as the region where cluster CMDs were obtained. For each star in the comparison annulus, the star in the CMD region which had the most similar B, V and I magnitudes was removed (allowing a maximum mismatch of 0.5 mag). For clusters where no B data were available, the matching was done using V and I only.

The top row in Fig. 2 shows the $(V - I, V)$ CMDs for all stars within the central apertures, while CMDs for the comparison annuli are shown in the second row from the top and the decontaminated cluster CMDs are in the third row. Also shown in the third row are theoretical Padua isochrones for the ages given in Table 1 and metallicities indicated in each panel (see references in Sect. 2.4). Equivalent plots are shown in Fig. 3 for the $(B - V, V)$ photometry. The cleaned CMDs all show a clear excess above the background and one can readily identify the RSGs at $V - I \approx 2$ in several clusters (NGC 1569-B, NGC 1705-1, NGC 1313-F3-1, NGC 5236-F1-1, NGC 5236-F1-3, NGC 7793-F1-1). The CMD for NGC 1569-A also shows a clear excess above the background, although this cluster appears too young to host a significant number of red supergiants. In the remainder of this paper we concentrate on the $(V - I, V)$ photometry, which constitutes a homogeneous dataset for the whole sample.

A first rough indication of the 50% completeness limits (red dashed lines in Fig. 2) and typical photometric errors was obtained by adding artificial stars to the cluster images, randomly distributed within the central aperture, and redoing the photometry. This was done for V magnitudes between 21 and 27 and $V–I$ colours between −1 and +3, in steps of 1 mag and for 100 stars at each step. We required a minimum separation of 5 pixels between two artificial stars. The completeness fraction was then determined as the fraction of added stars for which a star was recovered within a separation of 1 pixel. Since these artificial stars are added on top of existing images, there is some probability of recovering a star even where none was added. This number was determined by looking for counterparts to the artificial stars in the original input images and was typically on order ~10%. An alternative approach would be to simply compare the number of artificial stars added to an image with the additional number of detections. The problem with this approach is that a pre-existing star might be "masked" by an artificial star, leading to an underestimate of the number of recovered stars. It is clear that the classical artificial-star method is less than ideal for our situation and below we will adopt a more realistic approach to deal with the severe crowding in our fields. The classical artificial star approach was also applied to the background annuli (see Fig. 2). These are less affected by incompleteness than the
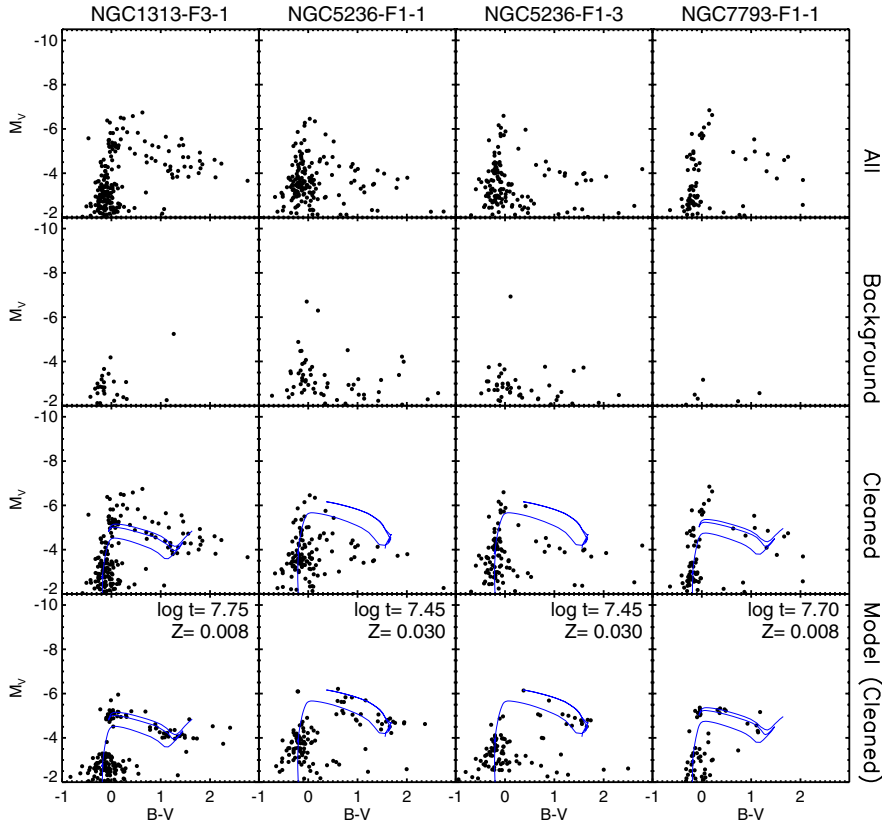
Fig. 3. Similar to Fig. 2, but showing $(B-V)$, M_V for those clusters where *F435W* band data are available.

central apertures. Given the approximative nature of these completeness determinations, we have not attempted to take the different degrees of incompleteness into account when correcting for the background contribution, so we are likely over-subtracting it. However, from Figs. 2 and 3 it is clear that this mainly affects the CMDs near the completeness limits where errors are large anyway. The post-main sequence stars are, in most cases, hardly affected by the background correction, and well above the 50% completeness limits.

2.3. Synthetic cluster experiments

A simple comparison with isochrones is of limited use for interpreting data such as those in Fig. 2, since this fails to take into account the strongly varying density of stars along a particular isochrone. More information can be gained by plotting the isochrones as Hess diagrams, as shown in Fig. 4 (for the Padua isochrones) and Fig. 5 (for the Geneva isochrones; Lejeune & Schaerer 2001). It should be noted that the Geneva isochrones are shown for the HST *Wide Field / Planetary Camera 2 F555W* and *F814W* filters, which are not completely identical to the filters with the same names on the ACS. However, for the purpose of illustration the differences are marginal.

The main characteristics of the Padua and Geneva isochrones are similar, although there are differences in the details. The predicted Blue Hertzsprung Gap is evident at all ages relevant to the data analysed here. It is already clear from Fig. 2 that this feature is not nearly as evident in the data. In what follows, we will restrict the comparison to the Padua isochrones, for which photometry in the ACS filters is directly available.

Even though the outskirts of the clusters appear resolved into individual stars in the ACS images, crowding still represents a major challenge and we cannot directly compare our observed CMDs with the Hess diagrams in Figs. 4 and 5. A realistic assessment of photometric errors and detection limits is crucial in order to carry out a meaningful comparison between the observed CMDs and model predictions. As we have discussed above, the classical approach of adding stars at random positions within an image is insufficient to fully capture the complicated effects of crowding in our data. In addition to the effects mentioned in Sect. 2.2, the background and crowding vary strongly with distance from the cluster centre, and a different approach is needed. Rather than relying on "typical" completeness limits and photometric errors, we therefore opted to generate artificial counterparts of the real cluster images, and carry out photometry on those. In this way, we could directly compare the CMDs of our *artificial* clusters with those of the actual clusters.

To generate the artificial clusters, we first derived structural parameters for each real cluster using the ISHAPE task in the BAOLAB package (Larsen 1999). ISHAPE assumes an analytic model of each cluster, where for this work we use profiles of the following type, found by Elson et al. (1987) to give a good fit to clusters in the Large Magellanic Cloud:

$$I(r) = I_0 \left[1 + \left(\frac{r}{r_c} \right)^2 \right]^{-\eta}. \tag{1}$$

Here, r_c is a core radius and η determines how steeply the envelope falls off at large radii. ISHAPE convolves the analytic cluster model with the PSF and iteratively adjusts the structural parameters of the cluster model until the best match to the data is
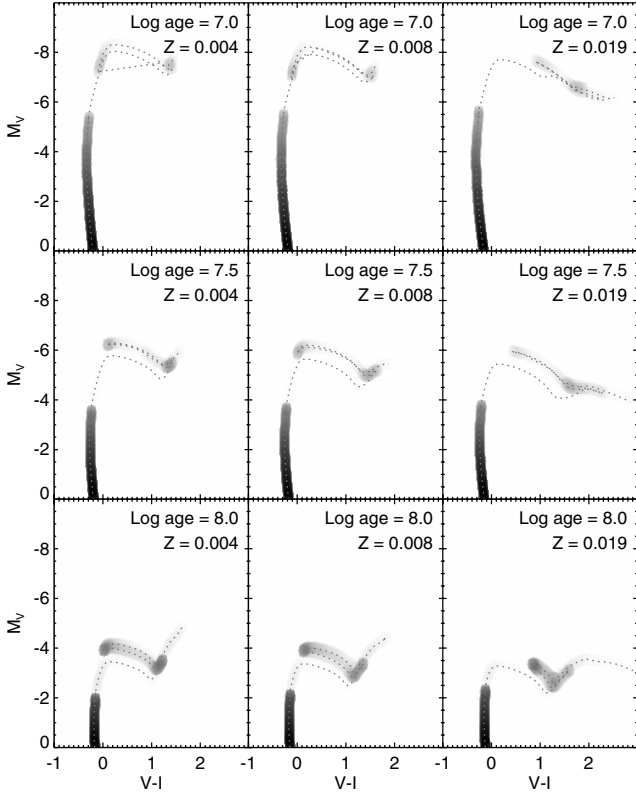
Fig. 4. Hess diagrams for Padua isochrones (Bertelli et al. 2009) of various ages and metallicites. The dotted line in each panel connects each point on the isochrones, while the shaded areas indicate the density of stars (assuming a Kroupa IMF).
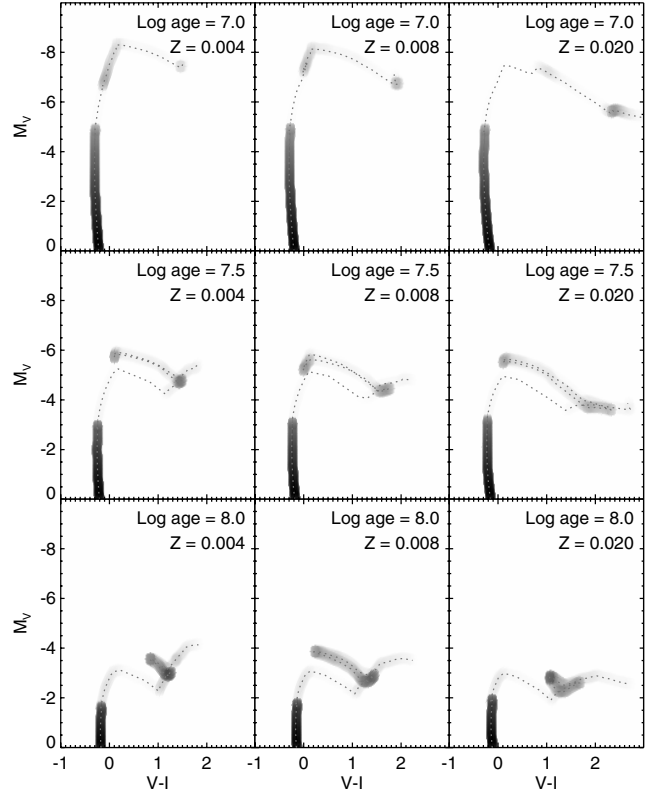
Fig. 5. Same as Fig. 4, but for Geneva isochrones (Lejeune & Schaerer 2001). Note that the Geneva isochrones are shown for *WFPC2* filter definitions.

achieved. In our case, the signal-to-noise was sufficiently high to allow ISHAPE to solve for both r_c and η. The last two columns in Table 1 list the best-fitting η and *FWHM* (Full Width at Half Maximum) for each cluster, the latter related to r_c and η as

$$FWHM = 2\,r_c\,\sqrt{2^{1/\eta - 1}} \qquad (2)$$

for the Elson et al. (1987) profiles. For the clusters located in spiral galaxies (NGC 1313, NGC 5236, NGC 7793) the *FWHM* values are fairly typical compared to other clusters of similar masses in spirals (Larsen 2004), while the clusters in NGC 1569 and NGC 1705 have extraordinarily compact cores. This is particularly remarkable given that the latter are also the most *massive* clusters in our sample, implying very high stellar densities in the centre of these clusters ($\gtrsim 10^5\ M_\odot\ \mathrm{pc}^{-3}$). We do not tabulate half-light radii, which are very uncertain because the η values are close to unity for most of the clusters (the half-light radii are, in fact, undefined for $\eta \leq 1$, unless the profiles are truncated at some finite outer radius).

Having determined the intrinsic intensity profile of each cluster, we then proceeded to sample stars at random from the model profiles. The profiles in Eq. (1) extend to infinity, but we imposed a cut-off at a radius of 50 pixels (2″5 for the WFC images and 1″25 for the HRC images). Each star was assigned a mass drawn at random from a Kroupa (2002) IMF, and corresponding magnitudes in each band were looked up in the Padua isochrones. Isochrones were selected according to the cluster ages given in Table 1 and the total number of stars in each artificial cluster was tuned such that the integrated *V*-band magnitudes matched those of the real clusters. The stars belonging

to these artificial clusters were then added, one by one, to the actual images using the MKSYNTH task in BAOLAB. To the extent possible, we added the artificial clusters at locations that resembled those of the real clusters in terms of the background level and stellar density. The photometric analysis described in Sect. 2.2 was then repeated on the images with the artificial clusters added.

Real and simulated *F555W* images of each cluster are shown in Fig. 6. Although our simulated clusters closely resemble the real clusters, a couple of caveats should be noted: first, while we have taken great care to model the photometric (in)completeness and errors as realistically as possible using simulated images, there are always simplifications involved in the generation of artificial datasets. For example, the PSFs used in the artificial cluster images are approximations to the actual PSFs of the parent images, so that stars in the artificial clusters will have slightly different PSFs than the real stars present in the images. This may introduce subtle differences in the goodness of the ALLFRAME fits, possibly affecting the photometric errors. This would be a concern especially if we used the same PSFs to generate artificial clusters *and* carry out PSF-fitting photometry for them. To alleviate this concern and mimic the process by which PSFs were derived for real clusters as closely as possible, we re-derived the PSFs for the synthetic images using a set of synthetic reference stars. These were placed in similar locations within the images as the real PSF stars in terms of crowding and background level. Another potential concern is our implicit neglect of mass segregation. This effect causes massive stars to move towards the centre of a cluster and low-mass stars to move to the outskirts. However, since the light profiles of both the artificial
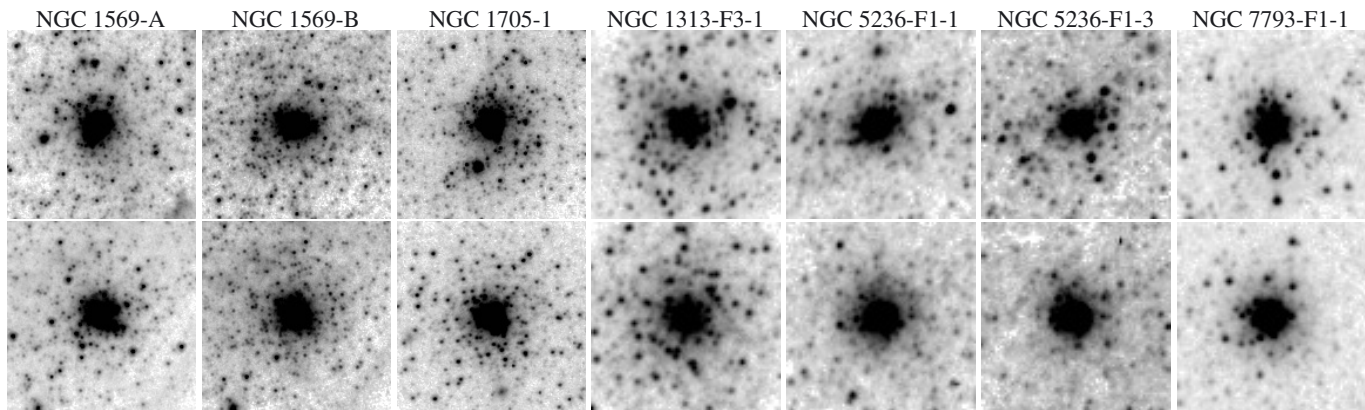
| NGC 1569-A | NGC 1569-B | NGC 1705-1 | NGC 1313-F3-1 | NGC 5236-F1-1 | NGC 5236-F1-3 | NGC 7793-F1-1 |

Fig. 6. *Top panels*: *F555W* images of the star clusters. *Bottom panels*: simulated clusters with the same apparent magnitudes, ages and sizes. *Each panel* measures $4'' \times 4''$.

and real clusters are dominated by the massive (luminous) stars, we do not expect this simplification to affect our analysis very strongly. Potentially more problematic is our neglect of binary stars. While random blends are automatically accounted for, physical binaries are not. This is a complex problem: a full account of binaries needs to consider not just the blending of light from two stars, but also possible effects of stellar evolution in binary systems (mass transfer, etc.) for realistic distributions of mass ratios and orbital periods. We will comment further on some possible effects of binary evolution on the CMDs below (Sect. 3.5).

2.4. Simulated versus observed CMDs

CMDs for the simulated clusters are shown in the bottom rows of Figs. 2 and 3. Background contamination has been statistically subtracted in the same way as for the real clusters. Also shown are the isochrones used to generate the lists of input magnitudes. We have somewhat arbitrarily, but guided by the CMDs in Fig. 2, made a division between "red" and "blue" stars at $V - I = 0.8$. In Fig. 7 we compare the observed and model luminosity functions for blue and red stars. The "turn-overs" in both the simulated and observed luminosity functions generally agree fairly well with the approximate 50% completeness limits indicated in Fig. 2.

For the artificial clusters, both the LFs and the synthetic CMDs are averages of 100 individual realizations of each cluster. For each realization, a different random number seed was used for the sampling of the IMF and the cluster profile, and the centre of the cluster was shifted around randomly within a box of ±50 pixels. This approach allows us to put one-sigma error bars on the simulated LFs that include not only stochastic fluctuations in the clusters themselves, but also background variations. It should once again be emphasized that the luminosity functions have not been corrected for completeness effects, so we do not attempt to derive quantitative estimates of the shapes of these distributions. Instead, we keep the comparison with the simulated data strictly differential. Similarly, the simulated CMDs shown in the figures are "average" CMDs generated by first combining the 100 individual CMDs for each artificial cluster, and then randomly sampling 1% of the stars.

Even though we employ PSF-fitting photometry, there is still the possibility that chance superpositions of two or several stars (as opposed to physical binaries) overlap so closely that they will be fitted as a single star. This is likely responsible for some of the data points in the synthetic CMDs that fall significantly

off the input isochrones. The effect will be present both in the simulated and real data, and should not affect the relative comparison. Using our synthetic clusters, we can obtain some rough estimates of the amount of blending. The percentage of stars affected by blending will clearly depend on the specific criterion for a blend to occur. If we (somewhat arbitrarily) define a "blend" as a star with a companion that is closer than 1 pixel and at most one magnitude fainter than the primary star, we find that roughly 20% of the brightest stars in our CMDs are affected by blending. However, this fraction depends very strongly on position and magnitude.

In terms of the magnitude limits and scatter in the colour of main sequence stars, the simulated CMDs resemble the observed CMDs quite well, giving us some confidence that the modelling procedure captures the most important observational effects. However, a detailed comparison shows several significant differences between the modelled and observed CMDs. We will now comment further on these. We first discuss similarities and differences between the observed and model CMDs for each cluster, and then consider how various modifications to our initial model assumptions might affect the model CMDs.

NGC 1569-A: comparison with the other CMDs clearly shows this cluster to be the youngest in our sample, in agreement with analyses based on integrated light. It is actually a binary cluster, as already hinted at in pre-refurbishment mission HST *Wide Field / Planetary Camera* data (O'Connell et al. 1994). In our model CMDs we have assumed a metallicity of $Z = 0.008$ for this cluster and NGC 1569-B (see discussion in Larsen et al. 2008). Spectroscopic studies have revealed the presence of both Wolf-Rayet and RSGs in NGC 1569-A (Origlia et al. 2001). However, while the raw CMD in the top panel of Fig. 2 does show a number of RSGs, the background annulus shows a similar population so we find no statistically significant population of RSGs in NGC 1569-A itself. We do however note that a number of the brightest stars (at $M_V \sim -9$) appear somewhat redder than any stars appearing in the model CMD. These stars cannot be accounted for by any choice of isochrone: the isochrone used for the modelling (log(age) = 6.65) is the oldest one for which no RSGs are present. For log(age)=6.70, the RSGs appear suddenly at $V - I \sim 1.0$. The observed and modelled LFs do not match particularly well either; on the blue side the simulated LF is significantly shallower than the observed one. The LF for red stars is most likely due to residual background contamination.
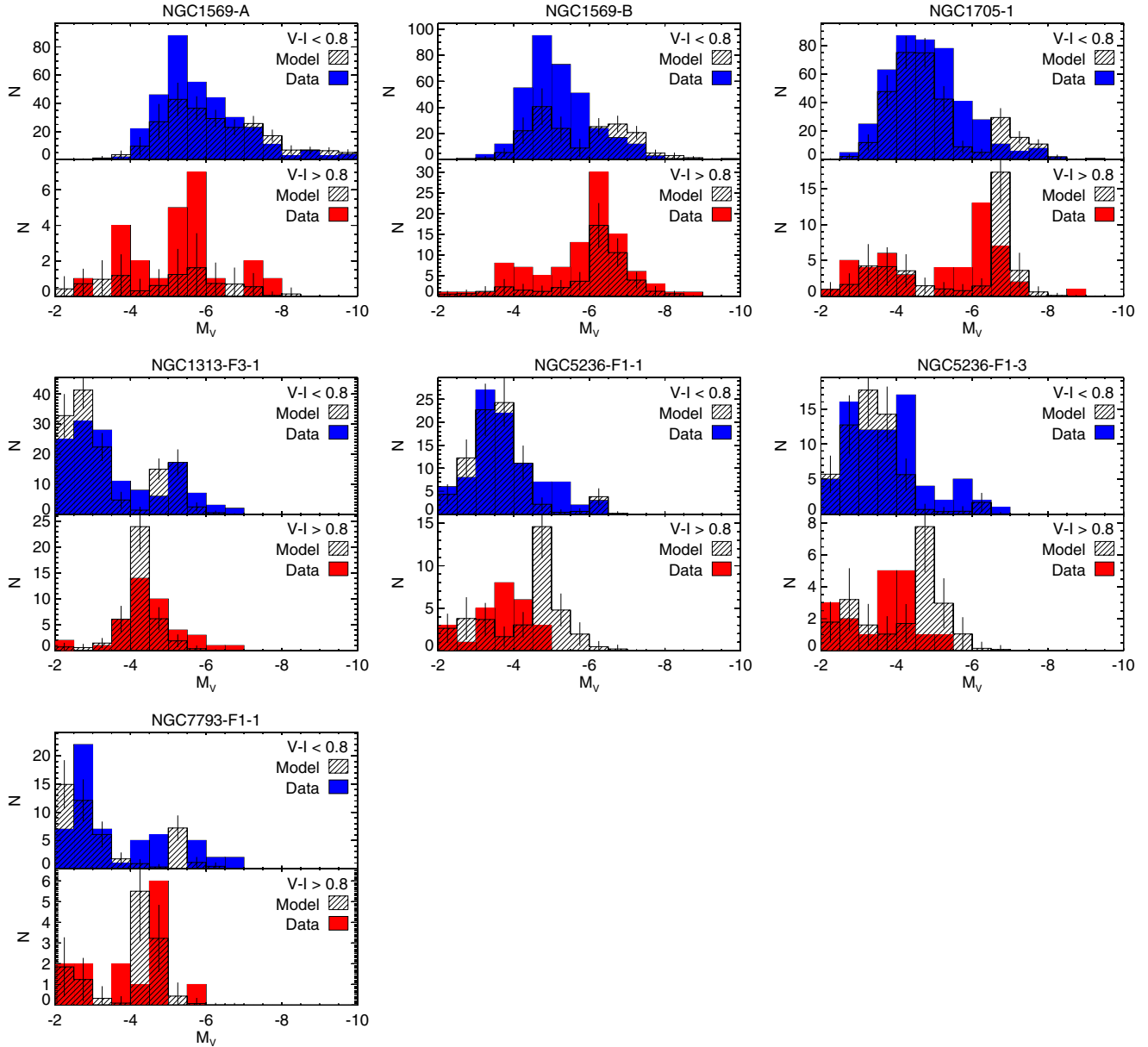
Fig. 7. Observed and simulated luminosity functions for red and blue stars in the clusters.

NGC 1569-B: the observed CMD reveals a large number of RSGs (about 60) which have been discussed in detail in Larsen et al. (2008). Figure 7 shows that the shape of the luminosity function of the RSGs is fairly well matched by the simulation for the assumed age and metallicity, although the detailed morphologies of the simulated and observed CMDs do differ somewhat. In particular, the observed luminosities of the RSGs show more scatter at fixed colour than the simulation. Overall, the simulated CMD appears sparser than the observed one. Although, by construction, the *total* luminosities of the simulated clusters match those of the observed ones, the integrated (*V*-band) luminosity of the *measured* stars is about 20% lower in the simulation for NGC 1569-B. Thus, the DAOPHOT photometry appears to be slightly less complete for the simulated cluster. This may be due to slight inaccuracies in the modelling of the cluster profiles or differences in the CMDs.

Similarly to several other clusters, the simulated CMD shows a clear BHG which is not present in the real data. This difference

is also clear when comparing the observed and simulated LFs. A small number of very luminous stars ($M_V < -8$) are observed; however, some such stars are also present in the simulation indicating that they may well be blends or other artifacts due to the extreme crowding in this very dense cluster.

NGC 1705-1: the image of this cluster shows considerable substructure, with two "filaments" and another fainter cluster located about 1″ from the centre of the main core. The metallicity of NGC 1705 is intermediate between those of the SMC and LMC (Lee & Skillman 2004) and for the simulation we have assumed $Z = 0.008$. The observed CMD of this cluster is very similar to that of NGC 1569-B, indicating similar ages for the two clusters, and also for NGC 1705-1 the simulated CMD is a rather poor match to the observed one (although in this case the integrated *V*-band luminosities of the measured stars agree within ~5% for the observed and simulated clusters).

Again, a BHG is clearly expected at $M_V \sim -6$, but not observed. Figure 7 shows that the observed and simulated numbers of red supergiants match quite nicely, although the observations show the RSGs to be somewhat cooler and fainter than predicted by the models for this age and metallicity.

NGC 1313-F3-1: this is a remarkable object. With a core radius of about 4 pc, this is a relatively extended cluster and the HST images show it to be resolved into numerous individual stars. The metallicity of NGC 1313 is 12+log O/H = 8.4±0.1 (Walsh & Roy 1997), which translates to [O/H] = −0.43 ± 0.1 (Grevesse & Sauval 1998) or $Z = 0.007$ assuming Solar abundance ratios. We thus use the $Z = 0.008$ Padua models for the simulated CMD, which to first order results in quite a good match to the data. This is one of the few cases where the observed CMD and LF show a hint of the BHG that is expected at $M_V \sim -4$. However, the observed CMD shows a spread of nearly 2 mag in the M_V magnitudes of the supergiant stars, a significantly larger spread than in the simulated CMDs.

NGC 5236-F1-1 and NGC 5236-F1-3: the metallicity of most of the NGC 5236 disc is well above Solar (Bresolin & Kennicutt 2002) and we have used the $Z = 0.03$ isochrones for the simulations. However, the simulated CMDs would not look much different for the Solar metallicity isochrones ($Z = 0.019$). The integrated colours suggest identical ages for these two clusters, and their observed CMDs are indeed quite similar and in neither case well reproduced by the simulations. NGC 5236-F1-1 is the brighter of the two by a magnitude, and the CMD shows about twice as many RSGs compared to NGC 5236-F1-3. The observed M_V magnitudes of the RSGs are about a magnitude fainter than the simulated ones, possibly an indication that the clusters are older than we have assumed. The mismatch between models and data is even worse on the blue side of the CMD. The models predict a clear BHG both in M_V magnitude and colour, whereas the data show no such separation. This problem cannot be solved by adopting an older age.

NGC 7793-F1-1: the observed CMD is similar to that for NGC 1313-F3-1 but more sparsely populated, consistent with the lower estimated total mass. We have assumed $Z = 0.008$ (Edmunds & Pagel 1984). Many of the same remarks given for NGC 1313-F3-1 apply to this cluster, but at lower statistical significance. Also here, there appears to be a significant 1–2 mag spread in the M_V magnitudes of the supergiant stars.

3. Discussion

3.1. General comments on CMD comparisons

Comparing our observed and simulated CMDs, we first note that they do agree in some respects. The colour of the main sequence is, in general, quite well reproduced, an indication that our adopted extinction values are reasonable. In addition, the CMDs show the red supergiants to be fairly well separated from the blue loop and main sequence stars, in agreement with model predictions for the relevant mass/age range (Figs. 4 and 5). However, there are also several areas of disagreement: 1) Even after accounting for observational errors, we confirm the presence of stars in the BHG where none are expected. This is a long-standing problem (Chiosi 1998; Salasnich et al. 1999) and has previously been noted e.g., for the ~50 Myr old

cluster NGC 1850 in the LMC (Fischer et al. 1993). Likewise, inspection of the CMD for the 25-Myr old cluster vdB0 in M31 (Perina et al. 2009) shows no clear gap, although statistics are poorer here. 2) In several cases, the observations show a larger spread in the luminosities of the supergiant stars than expected from our artificial-cluster experiments. 3) The observed colours (i.e., effective temperatures) of the RSGs generally tend to be somewhat redder than the predictions. This is seen both at super-solar metallicities (i.e., in NGC 5236), as well as in the more metal-poor (but younger) clusters NGC 1569-B and NGC 1705-1 (see also Larsen et al. 2008). The agreement between models and data is best for the clusters NGC 1313-F3-1 and NGC 7793-F1-1, which are the oldest in our sample and probably have slightly sub-solar metallicities.

In the following we discuss several modifications to our initial, simple model assumptions that might affect the CMDs.

3.2. Age spreads?

In recent years, the traditional view of star clusters as simple stellar populations has been challenged by various pieces of observational evidence. In addition to the long-standing puzzle of anomalous abundance ratios of light elements in globular clusters (e.g., Kraft 1979), this includes multiple main sequence turn-offs and subgiant branches in old GCs (Milone et al. 2010), as well as bi-modal or broad distributions of main sequence turn-off points in intermediate-age clusters in the Magellanic Clouds (Mackey et al. 2008; Milone et al. 2009) and the presence of dual red clumps in such clusters (Girardi et al. 2009). For ancient GCs there are few observational constraints on any actual age spread, and the observed CMD features also appear to require strongly enhanced He abundances (Norris 2004; Piotto et al. 2005; D'Antona & Ventura 2007). The LMC and SMC clusters where age spreads have been claimed all have ages of >1 Gyr, and the implied age spreads ($>10^8$ years) would have dramatic effects on the CMDs of younger clusters. Theoretically, it is a challenge to understand how extended star formation could occur within star clusters on long time scales. After a few Myr, the winds and radiation pressure from OB stars (or eventually, supernovae) are likely to evacuate any remaining gas, and the material subsequently returned by stellar evolution is insufficient to form the observed numbers of second-generation stars unless the first generation has a very top-heavy IMF. This problem may be partly alleviated if a large fraction of the first generation of stars is preferentially lost via dynamical effects, and/or mass lost from stars is supplemented by additional accreted material (e.g., Bekki 2011). It has been proposed that gas may indeed be accreted onto massive clusters as they plough through the interstellar medium of their parent galaxies (Pflamm-Altenburg & Kroupa 2009; Conroy & Spergel 2011).

So far, evidence of age spreads greater than a few Myr is scant in younger clusters. NGC 1850, one of the most luminous young clusters in the LMC, hosts a small number of stars with ages of only a few Myr that are located within a small subcluster (Fischer et al. 1993). Multiple stellar generations (10–16 Myr and 30–100 Myr) have also been reported for the cluster Sandage 96 in the nearby Sc-type spiral NGC 2403 (Vinkó et al. 2009) although the older population in this case may be associated with the surrounding field. The giant HII region NGC 604 in M33 appears to host distinct populations with ages of ~3 and ~12 Myr (Eldridge & Relaño 2011), and star formation also seems to have occurred in the star forming region NGC 346 in the Small Magellanic Cloud over a period of 5–10 Myr (Hennekemper et al. 2008; Cignoni et al. 2011) and possibly even up to ~20 Myr

(De Marchi et al. 2011). However, the latter two objects are probably too extended to evolve into single bound clusters.

An age spread would result in different main sequence turn-off points within a given cluster, thus providing a way to blur the CMDs, thereby potentially helping to fill the BHG and would also give a spread in the luminosities of the supergiant stars. We therefore now investigate to what extent this would help improve the agreement between our observations and the synthetic CMDs.

As a starting point, we use the cluster NGC 1313-F3-1. As can be seen in Fig. 2, the observed colours of the supergiant stars are fairly well approximated by the models for this cluster. It therefore appears plausible that a better overall match to the CMD could be obtained by including a range of ages. We first use the CMD-fitting code FITSFH (Silva-Villa & Larsen 2010) to search for the combination of isochrones that provides the best match to the observed CMD. This code first constructs a Hess diagram for the observed data, and then attempts to match this by linearly combining Hess diagrams for individual isochrones (such as those in Fig. 4) after accounting for data incompleteness and observational errors. We used the errors estimated by ALLFRAME as input to FITSFH. Because of the large uncertainties in the completeness estimates, we have not actually included a detailed modelling of detection incompleteness across the CMDs in the SFH fits presented here, but simply restricted the fits to regions of the CMDs above the 50% completeness limits in Fig. 2.

Figure 8 shows the resulting star formation history (SFH) for NGC 1313-F3-1. The fit suggests a somewhat extended SFH, starting about 50 Myr ago and extending until 15–20 Myr ago. Integrating the SFH in Fig. 8, we find that about 2/3 of the stellar mass was formed at $7.6 < \log t < 7.8$, although most of this may have formed in a single burst at $\log t \approx 7.7$. The remaining 1/3 of the mass is in the tail extending to $\log t \approx 7.2$. The first burst coincides quite well with the age derived from the integrated light. A similar analysis was done for three other clusters in our sample: NGC 1569-B, NGC 1705-1 and NGC 5236-F1-1. For NGC 1569-A we do not have much freedom to experiment with different SFHs because the lack of RSGs essentially only allows the youngest isochrones, and we further excluded NGC 7793-F1-1 and NGC 5236-F1-3 which have CMDs very similar to those of NGC 1313-F3-1 and NGC 5236-F1-1, respectively, but with poorer statistics.

While FITSFH will always output some best-fit star formation history, the quality of the "best fit" will be limited by the adequacy of the input assumptions, such as the stellar isochrones used, correct assessment of photometric errors, treatment of binaries, etc. For young stellar populations, it is generally difficult to obtain a very good fit to observed CMDs (e.g., Silva-Villa & Larsen 2011), and here we are faced with the additional difficulty of complicated incompleteness effects and relatively small numbers of stars. We cannot, therefore, claim to put very tight constraints on the specific star formation histories of the clusters, but merely aim to explore whether an age spread might account at least in a qualitative way for some of the observed CMD features, and roughly how large it would have to be. Our final step is therefore to construct model CMDs based on the FITSFH output and compare these with the observed CMDs. As an approximation to the extended SFHs returned by FITSFH, we model each cluster as a superposition of three individual bursts with the ages and relative amounts of mass in each burst listed in Table 2. The relative weights of each burst were determined by integrating over the SFHs from FITSFH. We see that, in general, most of the mass is in the oldest burst. The implied age
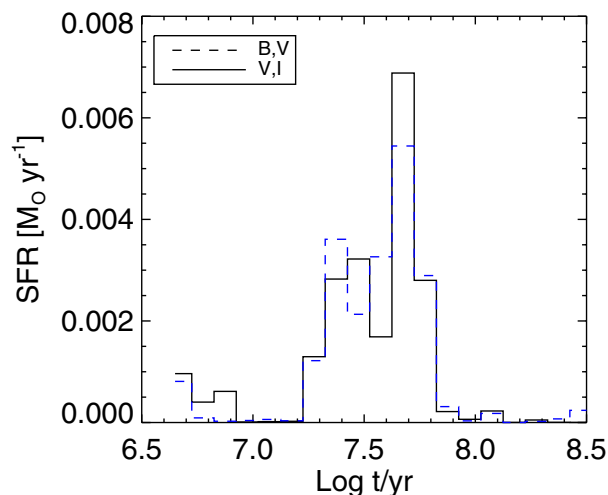
Fig. 8. Star formation history for NGC 1313-F3-1 derived via CMD synthesis. Solid curve: fit from V, I colours. Dashed curve: fit from B, V.

Table 2. Input for multiple-burst models.

Cluster	$\log t_1$	$\log t_2$	$\log t_3$	f_1	f_2	f_3
NGC 1569-B	7.20	6.95	6.60	0.70	0.15	0.15
NGC 1705-1	7.25	7.05	6.60	0.45	0.25	0.30
NGC 1313-F3-1	7.70	7.50	7.30	0.65	0.20	0.15
NGC 5236-F1-1	7.60	7.20	6.60	0.85	0.05	0.10

Notes. $\log t_1$–$\log t_3$ indicate the logarithmic ages (in years) of the three bursts and f_1–f_3 the relative amounts of stars formed in each burst.

spreads are ∼10–30 Myr. Due to the presence of younger populations in each cluster, which decreases the overall mass-to-light ratio of the cluster, the total masses of the multiple-burst models are 25–40% less than those of the single-burst models listed in Table 1. The modelling as three distinct bursts was done mostly for computational convenience, as it allowed us to easily modify the set-up used for the single-burst models by simply adding three individual "clusters" on top of each other. We could, of course, have increased the number of individual "bursts", but found that three individual sampling points in age provided an adequately smooth representation in the final CMDs. We emphasize that we do not intend to imply any preference for the clusters actually having experienced distinct bursts over smooth star formation histories.

In Fig. 9 we compare the multiple-burst models with the observations. The top row again shows the observed CMDs, while the middle row shows the CMDs for our multiple-burst models. Compared to the single-burst models, the fits are improved in the sense that we now get a larger range in luminosity for the supergiants, and the BHG is indeed blurred. However, the fits remain less than perfect. For NGC 1569-B, NGC 1705-1 and NGC 5236-F1-1, filling the BHG requires the inclusion of a younger isochrone ($\log t_3 = 6.60$) that has no RSGs, so that the number of RSGs is decreased with respect to the single-burst models. At least for NGC 5236-F1-1, this solution does not appear physically plausible, and results in a number of spuriously bright blue stars in the model CMD.

If the clusters indeed have an age spread, it would be of interest to study the spatial distributions of different sub-populations of stars within them. In Fig. 10 we attempt to make such a
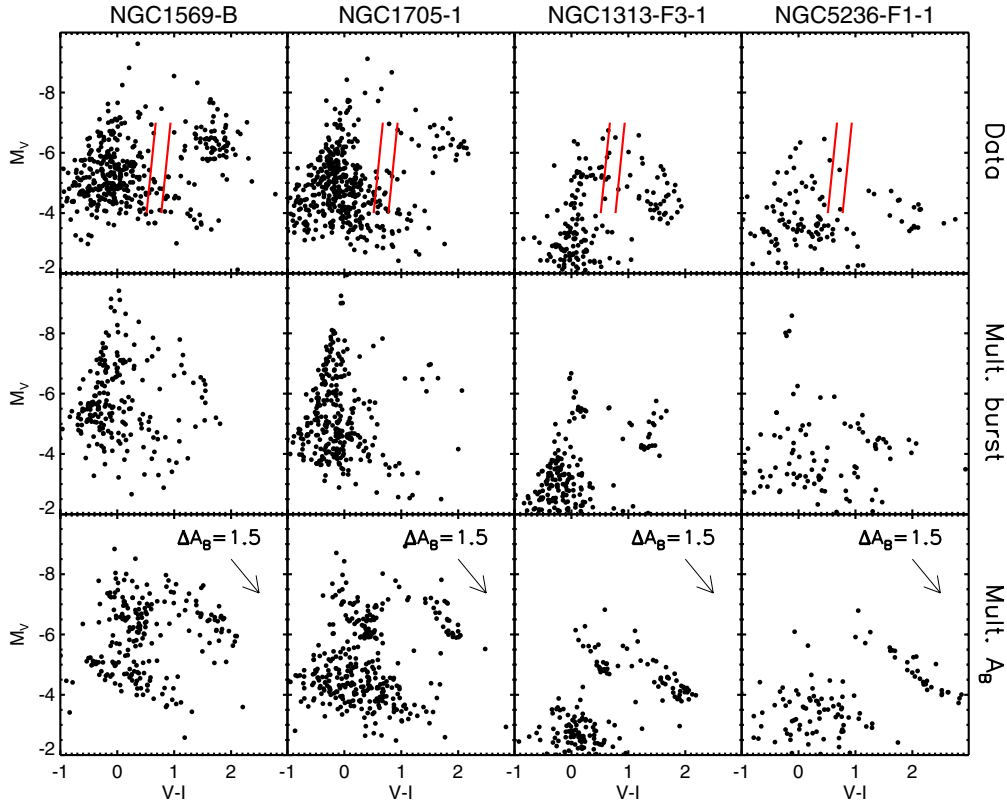
Fig. 9. *Top row*: observed CMDs. The red slanted lines indicate the approximate boundaries of the Cepheid instability strip (Sandage et al. 2004). *Middle row*: CMDs for multiple-burst models (see Table 2). *Bottom row*: CMDs for models with a range of A_B values.

comparison, by dividing the supergiant stars into a brighter and fainter subsample in each cluster. The left-hand panels illustrate our adopted division for each of the four clusters, and in the right-hand panels we plot the spatial distributions of the stars. The circles indicate the areas in which photometry was carried out. In NGC 1569-B there is a hint that the brighter stars appear somewhat more centrally concentrated, but in the other clusters there are no statistically significant differences between the radial distributions of bright and faint stars. This is quantified by the P-values from a K-S test, given in each panel. We do note that there is a curious asymmetry in the distribution of the brighter stars in NGC 1313-F3-1. The crossing time is only a few Myr, much less than the age of the cluster, so such substructure should be quickly erased by the internal motions of stars in the cluster. If the brighter stars are as young as ~20 Myr, as suggested by our CMD modelling, and formed as a coherent group, this might help explain how some substructure could remain visible. Alternatively, the apparent asymmetry may simply be the result of a chance alignment.

3.3. Extinction variations

In the discussion so far, we have implicitly assumed that there is no significant spread in the extinction towards individual cluster stars. In the bottom row of Fig. 9 we show simulated CMDs that include a range of $\Delta A_B = 1.5$ mag in the B-band extinction. As in the multiple-burst models, this was done by simulating three distinct A_B values ($A_B = 0, 0.75$ and 1.5 mag). In these models, the ages of the model clusters are significantly younger than for the models in Fig. 2 and 3. By trial-and-error, we found that the best fits to the observed CMDs were achieved for input ages of

$\log t = 7.0, 7.0, 7.50$ and 7.25 for NGC 1569-B, NGC 1705-1, NGC 1313-F3-1 and NGC 5236-F1-1, respectively, and for 70% of the stars belonging to the most extincted ($A_B = 1.5$ mag) bin.

Although a spread in extinction will clearly contribute to blurring of the CMDs, it is evident from the figure that there are several difficulties with this scenario. In all clusters, most of the supergiants would have to be relatively strongly extincted, with only a smaller fraction of them scattering upwards in the CMDs towards lower extinction values. This, however, is in contrast to the main sequence whose colours are well matched by isochrones with little or no additional extinction needed (see also Figs. 2 and 3). Furthermore, the small width of the main sequence in some clusters (especially NGC 1313-F3-1, and to some extent NGC 1569-B), also precludes a large scatter in the extinction values, unless the supergiants are systematically more subject to extinction variations than the main sequence. In conclusion, we find it unlikely that extinction variations have a major impact on the observed CMDs.

3.4. Cepheids

Another possible reason for the spread in the luminosities of the supergiant stars is that some of these stars are variable. In Fig. 9 we have indicated the location of the Cepheid instability strip (Sandage et al. 2004). While NGC 1569-B and NGC 1705-1 are probably too young to host Cepheids, it is possible that some stars in NGC 1313-F3-1 and NGC 5236-F1-1 (as well as NGC 7793-F1-1 and NGC 5236-F1-3) may be Cepheids. The pulsation amplitudes can amount to more than 1 mag in the V-band (Sandage et al. 2004). Although there is a hint that the spread in M_V magnitude is largest near the instability strip, a
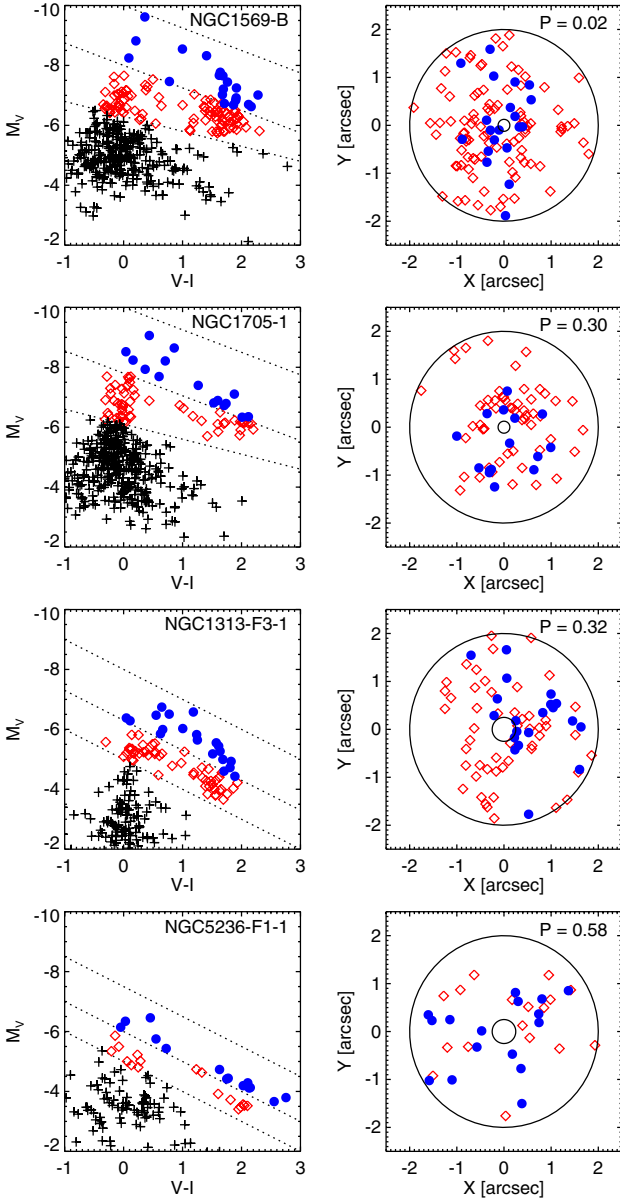
Fig. 10. *Left panels*: the division into brighter and fainter supergiant (SG) stars. *Right panels*: spatial distribution of the fainter and brighter SGs. K-S tests are generally consistent with no differences in the radial distributions of the two populations, as indicated by the *P*-values in *each panel*.

considerable scatter remains outside the boundaries of the strip, so it is questionable whether this can be the whole explanation. Observations of Cepheids might, however, shed light on possible age spreads within the clusters via the period-age relation (Efremov 2003; Bono et al. 2005).

3.5. Binaries

Binaries are common among massive stars (e.g., Mason et al. 2009) and can affect the CMDs in various ways, see for example the excellent discussion by Pols & Marinus (1994). Although it is beyond the scope of this paper to model these effects in detail we will briefly discuss them.

Unresolved multiples: the presence of an unresolved non-interacting companion can increase the brightness by up to a

factor two, i.e. change the magnitude by up to −0.75 mag. Although we must seriously consider the possibility that all stars have an unresolved companion, in most cases the companion will be considerably less luminous. Recent data suggest that the distribution of mass ratios (q) is roughly flat, implying that a typical companion is only half the mass of the primary (Pinsonneault & Stanek 2006; Kobulnicky & Fryer 2007; Kiminki et al. 2009). Given the steep dependence of the luminosity on the mass, the effect on the brightness of such a companion would be less than −0.1 mag for main sequence stars, and even less if the primary star has already evolved off the main sequence. If unresolved non-interacting binaries were to contribute significantly to any spread in the luminosities of the supergiants, this would require that a significant fraction of these binaries have nearly identical masses so that both stars would be in the supergiant phase at the same time.

For these clusters, where the stellar density is high, unresolved higher order multiple systems may be rather common. These may either be physically bound triple or multiple systems or multiple stars that are aligned by chance, although the latter case is in principle taken into account by our artificial cluster experiments. However, even if the light of a star is polluted with the light of two additional equally bright stars the effect on the magnitude is still only −1.2 mag. This is comparable but still smaller than the size of the BHG as predicted by stellar models, which is about ∼2 mag. A shift of 2 mag is reached only for an unresolved multiple system out of which at least 6 stars are equally luminous. We conclude that unresolved multiple systems may contribute to differences between our observed and modelled CMDs but can certainly not be the lone culprit.

Interacting binaries: Data from nearby open clusters suggest about 4 out of 10 O stars have companions that are nearby enough to interact (Sana & Evans 2010). Even though the fraction of stars whose properties have been altered by binary interaction may not be that large, the consequences of such interactions for the stellar luminosity are potentially severe due to the steep dependence of the luminosity on mass. In the most extreme cases, interaction can lead to a star that is twice as massive as the turn-off mass, which would imply an increase in the bolometric luminosity of a factor ∼15 or ∼3 mag. In practice, such cases should be rare and a more typical shift would be on the order of ∼5×, for a star that increases its mass by about 50% as a result of mass accretion or a merger of the original stars[1]. The effect on the CMD is still uncertain as it depends on various factors: the distribution of initial binary parameters, the amount of mass loss during interaction or mergers and the life time after interaction. In order to convert the changes in bolometric luminosity to magnitudes and colours, the effective temperatures of the merger products must also be known.

Binary interaction also affects the post main sequence evolution and therefore the luminosities of red and blue supergiants, as well as their relative numbers. The presence of a nearby companion can strip the envelope from a star before it reaches giant dimensions, effectively reducing the number of red supergiants by up to 40% (assuming the distribution functions of Sana & Evans 2010). In addition, the accreting star or merger resulting from binary interaction may under certain circumstances produce blue supergiants as suggested for the progenitor of SN1987A and the companion of SN1993J (Hillebrandt et al. 1987; Walborn et al. 1989; Maund et al. 2004). This second effect depends, among

---

[1] Assuming that the luminosity of the star after interaction with the mass as $L \sim M^{3.9}$.
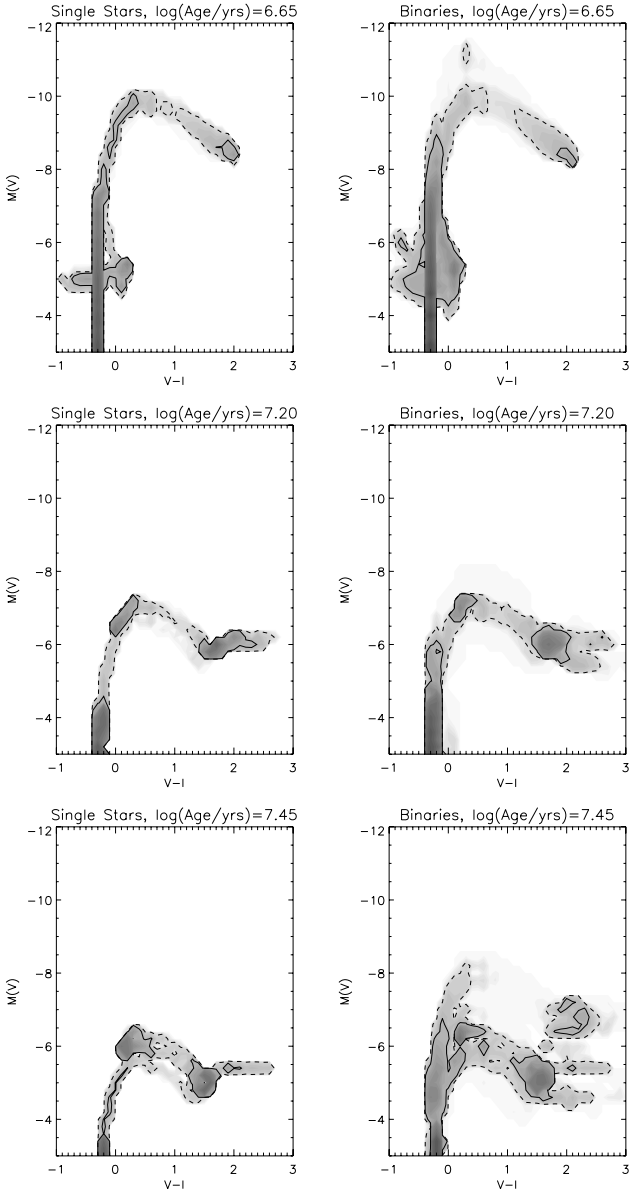
Fig. 11. Example CMDs for single stars (*left panels*) and binary (*right panels*) populations calculated with BPASS at three ages typical of clusters in our sample. Due to the nature of the binary models the CMDs are created assuming a small age spread for the population of $+/-0.05$ dex. The grid cells are 0.1 by 0.2 in $V-I$ and M_V respectively. The grey scale is logarithmic in number of stars with the cell size. The dashed and solid contours represent when the density of stars in each cell is 0.1 and 1 respectively.

general the effect of binaries, as discussed above, is to increase the luminosity of stars, including a population of "blue stragglers" that populate the region above the MS turn-off, as well as a number of brighter blue and red supergiants. We note that the biggest increase is for the older populations shown here, where it produces an effect very similar to that of an age spread (Fig. 9). These *extra* supergiants arise from binary mergers and secondaries in binaries that become more massive by accreting material from their primaries. In both cases this leads to more massive stars in the stellar population sometime after stars of such a mass would be expected to have exploded in supernovae for a simple single star population. This suggests that the creation of *apparently* younger stellar populations from binary evolutionary paths might plausibly account for some of the differences between our observed CMDs and the models based on canonical stellar isochrones (Sect. 3.2). While the code necessarily involves many simplifications in the modelling of binary evolution, and uncertainties in predicting the post-MS evolution of products of binary evolution are at least as large as for single stars, we note that the RSG to BSG ratios for these binary populations in fact do provide values similar to those found in Sect. 3.7. However, making more detailed predictions of the effects on the CMDs of these clusters is difficult due to the highly non-linear nature of binary stellar evolution, and beyond the scope of this paper.

3.6. Rotational mixing

Massive stars are observed with a wide range of rotation rates (Penny & Gies 2009). Rotation can induce mixing in the stellar interior during core hydrogen burning. This mixing brings fresh fuel towards the core of the star, which leads to an increased life time and to a larger helium core and thus higher luminosity (Maeder & Meynet 2000; Brott et al. 2011a). A distribution of initial rotational velocities will therefore lead to a dispersion of lifetimes and luminosities of stars of the same initial mass.

However, the effect is expected to be rather limited. Even for very rapid rotators, the life time is not increased by more than $\sim 30\%$ (Brott et al. 2011b), and less so for average rotators, which is smaller than the required age spread as discussed in Sect. 3.2. The spread in post-main sequence luminosity is predicted to be similar, i.e., 0.3 mag (Maeder & Meynet 2000), which is again too small to account for the observed luminosity spread in our clusters. Furthermore, while rotational mixing can shift the red edge of the main sequence band slightly redwards, the same is true for the blue edge of the blue loops (Maeder & Meynet 2001), such that the extent of the predicted blue Hertzsprung gap remains unaffected.

3.7. The red-to-blue supergiant ratio

One long-standing problem in the modelling of massive stars is the number ratio of red to blue supergiants (Langer & Maeder 1995). Determining this ratio observationally is far from straight-forward. Galactic open clusters generally contain only a small number of such stars, and in the field it is difficult to isolate clean samples of red and blue supergiants because the latter are easily confused with massive main sequence stars. Due to the richly populated post main-sequence phases, the clusters in our sample offer a welcome opportunity to address this issue.

In Table 3 we list the numbers of red and blue supergiants ($N_{\rm RSG}$ and $N_{\rm BSG}$) found in the four richest clusters, using the colour division at $(V - I) = 0.8$ (as in Fig. 7) and the magnitude cuts indicated by the lower dotted lines in Fig. 10 (in
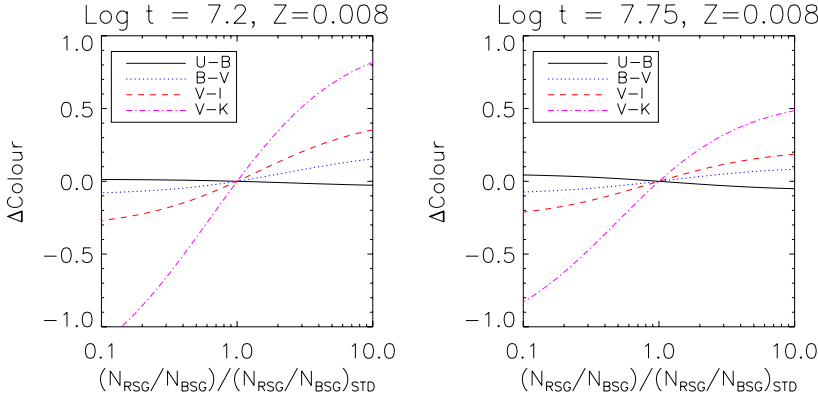
other things, on the efficiency of internal mixing processes and the number and lifetime of blue supergiants produced is therefore hard to estimate (Claeys et al. 2011). However, in combination the two effects may increase the BSG/RSG ratio by up to a factor of 4 (Eldridge et al. 2008).

In Fig. 11 we show how the CMD might be influenced if approximately two thirds of the stars are in interacting binaries. We use the binary population and spectral synthesis code BPASS (Eldridge & Stanway 2009) to create model stellar populations based on single stars (left) and binaries (right) at three different illustrative ages. The models have initial separations between $\log(a/R_\odot) = 1$ and 4 in steps of 0.25 dex, while the mass ratio takes values of $q = 0.1, 0.3, 0.5, 0.7$ and 0.9. We see that in

Fig. 12. Changes in integrated colours as a function of change in N_{RSG}/N_{BSG} ratio, relative to the ratio predicted by standard isochrones $(N_{RSG}/N_{BSG})_{STD}$.

Table 3. Statistics of blue and RSGs.

Cluster	N_{RSG}	N_{BSG}	N_{RSG}/N_{BSG}		
			Obs.	SSP	Mult.
NGC 1569-B	69	45	1.53 ± 0.09	0.53	0.27
NGC 1705-1	28	43	0.65 ± 0.03	0.39	0.17
NGC 1313-F3-1	39	32	1.22 ± 0.08	1.53	1.17
NGC 5236-F1-1	20	13	1.54 ± 0.30	5.04	2.84

Notes. The ratio N_{RSG}/N_{BSG} is given for the observations, for the single-burst models (SSP) and for the multiple-burst models (Mult.).

this comparison we do not distinguish between the "bright" and "faint" subsamples). Specifically, the cuts are:

$$\text{NGC 1569-B}: \quad M_V < 0.5(V - I) - 6.3$$
$$\text{NGC 1705-1}: \quad M_V < 0.5(V - I) - 6.1$$
$$\text{NGC 1313-F3-1}: \quad M_V < (V - I) - 5.0$$
$$\text{NGC 5236-F1-1}: \quad M_V < (V - I) - 5.0.$$

Note that the samples of supergiant stars in Table 3 are larger than those found in any individual Galactic young clusters. We also list the number ratio N_{RSG}/N_{BSG} and its associated error, assuming Poissonian statistics, along with the N_{RSG}/N_{BSG} ratio for our single- and multiple-burst simulations. The random errors on the simulated ratios are very small due to our averaging over 100 realizations of each cluster. The true uncertainty on the observed N_{RSG}/N_{BSG} is generally larger than indicated by the Poissonian errors, and typically dominated by the difficulty of isolating the BSGs due to the lack of a clear BHG. For the simulated CMDs in Fig. 2, the adopted cuts make a fairly clean selection of post-MS stars, but for the observed CMDs the number of BSGs depends on where exactly the magnitude cuts are made.

For NGC 1313-F3-1 the agreement between the observed and simulated N_{RSG}/N_{BSG} ratio is excellent, but for the other clusters there is significant disagreement. For NGC 1569-B and NGC 1705-B, the observed ratios are significantly higher than the simulated ones, a discrepancy that is exacerbated in the multiple-age models. For NGC 5236-F1-1 it is very difficult to select a sample of BSGs in a meaningful way.

3.8. Consequences for integrated colours

Because the red and blue supergiant stars contribute a significant fraction of the light from a simple stellar population, it is of interest to consider how changes in this ratio will affect integrated properties such as broad-band colours. To assess this,

we modified the N_{RSG}/N_{BSG} ratios in the Padua models and re-computed various broad-band colours. We did this by adopting a division between red and blue supergiants at $V - I = 0.8$ and changing the relative weights of stars on either side of this limit when computing integrated magnitudes. The (wavelength-dependent) integrated luminosity is then

$$L_\lambda = \int_{M_{min}}^{M_{max}} w(M) \, L_\lambda(M) \, \xi(M) \, \mathrm{d}M \qquad (3)$$

which is the usual equation for a continuously sampled SSP apart from the factor $w(M)$ which has different values depending on whether a star is on the main sequence or is a red or blue supergiant. Here, ξ is the stellar IMF and $L_\lambda(M)$ is the luminosity of a star with mass M. We kept the total number of supergiants constant, i.e.,

$$w_{RSG} \int_{M \in M_{RSG}} \xi(M) \, \mathrm{d}M + w_{BSG} \int_{M \in M_{BSG}} \xi(M) \, \mathrm{d}M$$
$$= \int_{M \in M_{RSG}} \xi(M) \, \mathrm{d}M + \int_{M \in M_{BSG}} \xi(M) \, \mathrm{d}M.$$

Figure 12 shows the result of this exercise for two different ages ($\log t = 7.2$ and $\log t = 7.75$) and four different broad-band colours ($U - B$, $B - V$, $V - I$ and $V - K$). This calculation was done for actual Johnson-Cousins colours. We plot the change in each colour with respect to the colour computed for the unmodified isochrones, as a function of the change in the N_{RSG}/N_{BSG} ratio (i.e., w_{RSG}/w_{BSG}). For $B - V$, $V - I$ and $V - K$ we see that the integrated colours indeed do become redder when the N_{RSG}/N_{BSG} ratio is increased. However, for $U - B$ the opposite happens. This somewhat counterintuitive result may be understood as follows: Taking the standard $\log t = 7.75$ model as an example, the contribution of the RSGs to the integrated light is (0.8%, 4.0%, 13%, 33%, 70%) in the (U, B, V, I, K) passbands, while the BSGs contribute (17%, 25%, 26%, 23%, 13%) in the same bands. Because both the blue and red supergiants are cooler than the main sequence stars, the net effect of both is to produce redder integrated colours than would be the case for a cluster without any supergiant stars. At long wavelengths the red supergiants dominate the light, and increasing the N_{RSG}/N_{BSG} therefore has the expected effect of making the integrated colours redder. However, at short wavelengths the contribution from the RSGs to the integrated light becomes very small, and only BSGs are able to produce a significant change in the integrated colours. Hence, increasing the N_{RSG}/N_{BSG} ratio in this case has the net effect of decreasing the overall contribution of supergiant stars to the integrated light at shorter wavelengths, thus making the integrated $U - B$ colours bluer rather than redder.

Fortunately, the changes in integrated photometry appear relatively modest even for rather extreme (factor of 10) changes in the $N_{\mathrm{RSG}}/N_{\mathrm{BSG}}$ ratio, at least for optical colours. $U - B$ is the most age-sensitive of the colours shown in Fig. 12 (at least at young ages), and a change of $\Delta(U - B)$ of 0.1 mag corresponds to a difference of about 0.1 dex in the age of a cluster. It therefore appears that age estimates from integrated colours may not be too strongly affected by uncertainties in the RSG/BSG ratio. However, it would be worth investigating in more detail how other properties estimated from integrated colours (e.g., extinction, metallicities, etc.) would be affected. In particular, the large effect on the near-IR colours might lead to substantial uncertainties in fits involving these colours that attempt to simultaneously constrain age, metallicity and extinction. Combined with the fact that age determinations involving near-IR colours are also more sensitive to stochastic sampling of the stellar IMF (Fouesneau & Lançon 2010), this implies that the inclusion of more passbands (in particular extending to the near-IR) does not necessarily lead to a better determination of physical cluster properties from integrated colours.

4. Summary

We have presented resolved photometry for seven young massive star clusters in five nearby galaxies, located at distances of 3–5 Mpc. While it is challenging to carry out photometry for individual stars in clusters at such distances, the richness of these clusters allows us to obtain CMDs for large samples of supergiant stars. In order to assess the effects of crowding, we have carried out extensive artificial-cluster tests where we generated synthetic clusters with properties as similar as possible to those of our real clusters. By comparing the CMDs of our real and simulated clusters (the latter based on Padua isochrones), we reach the following conclusions:

–   We confirm previously known problems with canonical stellar isochrones:
    1. The standard models predict a clear separation in the CMD between the H-burning main sequence and the He-burning "blue loop" stars (i.e., the blue Hertzsprung gap) while no such gap is seen in the data.
    2. The observed red-to-blue supergiant ratio is in general not well reproduced by the models. However, the lack of a clear BHG makes it difficult to identify clean samples of blue supergiants for this comparison
–   We have investigated the effect of changing the red-to-blue supergiant ratio on integrated colours and find that even large changes in this ratio have relatively modest effects on optical colours that would affect age estimates from integrated photometry by less than 0.1–0.15 dex. However, colours involving infrared passbands (such as $V - K$) are much more strongly affected, strengthening the conclusion that inclusion of such colours may actually increase, rather than decrease, the systematic errors when deriving cluster properties from integrated photometry (Fouesneau & Lançon 2010).
–   We observe a significant scatter (~ 2 mag) in the luminosities of the supergiant stars in some clusters. Comparison with our simulated clusters indicates that this scatter cannot be explained by observational errors alone. It may be partly due to variability, since some of the clusters are in the age range where Cepheid variables are expected. However, the scatter is seen also for stars outside the instability strip.
–   By including an age spread of 10–30 Myr in our model clusters we can achieve a better match to the observations. This

is, however, much smaller than the spread that has been reported for older clusters in the Magellanic Clouds, and an age spread cannot fully explain the morphologies of all the CMDs. While uncertainties in stellar models may account for the differences between predictions and observations of the BHG, this is a less likely explanation for a spread in the supergiant luminosities. We discuss several other mechanisms that might contribute to differences between observed and model CMDs, and conclude that extinction variations and unresolved binaries are unlikely to be the main culprits, while interacting binaries might also have a significant effect on the CMDs.
–   In general, we find no significant differences in the spatial distributions of the brighter and fainter supergiant stars. In NGC 1569-B there is a hint that the brighter stars are somewhat more centrally concentrated, and in NGC 1313-F3-1 there is a curious asymmetry in the distribution of the brighter stars.

Each of the clusters studied in this paper has its own unique characteristics and the current sample of seven objects is clearly still too small to draw general conclusions. However, our search of the HST archive is far from complete, and it is likely that several additional objects have already been observed that would lend themselves to similar analysis. Furthermore, the ongoing Panchromatic Hubble Andromeda Treasury survey (Dalcanton & PHAT Collaboration 2011) will include a vast number of young clusters that may also be useful for addressing the issues discussed here (although most of them significantly less massive than those included here). Finally, the photometry is available online at CDS (Tables 4–10). It is our hope that the data presented here will prove useful as benchmarks for developers of stellar models.

References

Anders, P., de Grijs, R., Fritze-v. Alvensleben, U., & Bissantz, N. 2004, MNRAS, 347, 17
Barmina, R., Girardi, L., & Chiosi, C. 2002, A&A, 385, 847
Bastian, N., & de Mink, S. E. 2009, MNRAS, 398, L11
Bekki, K. 2011, MNRAS, 412, 2241
Bertelli, G., Nasi, E., Girardi, L., & Marigo, P. 2009, A&A, 508, 355
Bono, G., Marconi, M., Cassisi, S., et al. 2005, ApJ, 621, 966
Bresolin, F., & Kennicutt, J. 2002, ApJ, 572, 838
Brott, I., de Mink, S. E., Cantiello, M., et al. 2011a, A&A, 530, A115
Brott, I., Evans, C. J., Hunter, I., et al. 2011b, A&A, 530, A116
Carignan, C. 1985, ApJS, 58, 107
Chiosi, C. 1998, in Stellar Astrophysics for the Local Group: a First Step to the Universe, ed. A. Aparicio, A. Herrero, & F. Sanchez (Cambridge: Cambridge University Press), 1
Cignoni, M., Tosi, M., Sabbi, E., Nota, A., & Gallagher, J. S. 2011, AJ, 141, 31
Claeys, J. S. W., de Mink, S. E., Pols, O. R., Eldridge, J. J., & Baes, M. 2011, A&A, 528, A131
Conroy, C., & Spergel, D. N. 2011, ApJ, 726, 36
Dalcanton, J., & PHAT Collaboration. 2011, Am. Astron. Soc., 43, 222.01
D'Antona, F., & Ventura, P. 2007, MNRAS, 379, 1431
Davies, B., Figer, D. F., Kudritzki, R., et al. 2007, ApJ, 671, 781
De Marchi, G., Panagia, N., Romaniello, M., et al. 2011, ApJ, accepted [arXiv:1104.4494]
Dohm-Palmer, R. C., & Skillman, E. D. 2002, AJ, 123, 1433
Dohm-Palmer, R. C., Skillman, E. D., Saha, A., et al. 1997, AJ, 114, 2514

Edmunds, M. G., & Pagel, B. E. J. 1984, MNRAS, 211, 507
Efremov, Y. N. 2003, Astron. Rep., 47, 1000
Eggenberger, P., Meynet, G., & Maeder, A. 2002, A&A, 386, 576
Eldridge, J. J., & Relaño, M. 2011, MNRAS, 411, 235
Eldridge, J. J., & Stanway, E. R. 2009, MNRAS, 400, 1019
Eldridge, J. J., Izzard, R. G., & Tout, C. A. 2008, MNRAS, 384, 1109
Elson, R. A. W., Fall, S. M., & Freeman, K. C. 1987, ApJ, 323, 54
Fischer, P., Welch, D. L., Cote, P., Mateo, M., & Madore, B. F. 1992, AJ, 103, 857
Fischer, P., Welch, D. L., & Mateo, M. 1993, AJ, 105, 938
Fouesneau, M., & Lançon, A. 2010, A&A, 521, A22
Gallart, C., Zoccali, M., & Aparicio, A. 2005, ARA&A, 43, 387
Girardi, L., Chiosi, C., Bertelli, G., & Bressan, A. 1995, A&A, 298, 87
Girardi, L., Rubele, S., & Kerber, L. 2009, MNRAS, 394, L74
Girardi, L., Eggenberger, P., & Miglio, A. 2011, MNRAS, 412, L103
Grevesse, N., & Sauval, A. J. 1998, Space Sci. Rev., 85, 161
Grocholski, A. J., Aloisi, A., van Der Marel, R. P., et al. 2008, ApJ, 686, L79
Hennekemper, E., Gouliermis, D. A., Henning, T., Brandner, W., & Dolphin, A. E. 2008, ApJ, 672, 914
Hillebrandt, W., Höflich, P., Truran, J. W., & Weiss, A. 1987, Nature, 327, 597
Hirschi, R., Meynet, G., & Maeder, A. 2004, A&A, 425, 649
Israel, F. P. 1988, A&A, 194, 24
Kiminki, D. C., Kobulnicky, H. A., Gilbert, I., Bird, S., & Chunev, G. 2009, AJ, 137, 4608
Kobulnicky, H. A., & Fryer, C. L. 2007, ApJ, 670, 747
Kraft, R. P. 1979, ARA&A, 17, 309
Kroupa, P. 2002, Science, 295, 82
Langer, N., & Maeder, A. 1995, A&A, 295, 685
Larsen, S. S. 1999, A&AS, 139, 393
Larsen, S. S. 2004, A&A, 416, 537
Larsen, S. S., Origlia, L., Brodie, J., & Gallagher, J. S. 2008, MNRAS, 383, 263
Lee, H., & Skillman, E. D. 2004, ApJ, 614, 698
Lejeune, T., & Schaerer, D. 2001, A&A, 366, 538
Lyubimkov, L. S., Lambert, D. L., Rostopchin, S. I., Rachkovskaya, T. M., & Poklad, D. B. 2010, MNRAS, 402, 1369
Mackey, A. D., Broby Nielsen, P., Ferguson, A. M. N., & Richardson, J. C. 2008, ApJ, 681, L17
Maeder, A., & Meynet, G. 2000, ARA&A, 38, 143
Maeder, A., & Meynet, G. 2001, A&A, 373, 555
Maraston, C. 1998, MNRAS, 300, 872
Mason, B. D., Hartkopf, W. I., Gies, D. R., Henry, T. J., & Helsel, J. W. 2009, AJ, 137, 3358
Matraka, B., Wassermann, C., & Weigert, A. 1982, A&A, 107, 283
Maund, J. R., Smartt, S. J., Kudritzki, R. P., Podsiadlowski, P., & Gilmore, G. F. 2004, Nature, 427, 129

Méndez, B., Davis, M., Moustakas, J., et al. 2002, AJ, 124, 213
Milone, A. P., Bedin, L. R., Piotto, G., & Anderson, J. 2009, A&A, 497, 755
Milone, A. P., Piotto, G., Bedin, L. R., et al. 2010, in SF2A-2010: Proceedings of the Annual meeting of the French Society of Astronomy and Astrophysics, ed. S. Boissier, 319
Mora, M. D., Larsen, S. S., Kissler-Patig, M., Brodie, J. P., & Richtler, T. 2009, A&A, 501, 949
Norris, J. E. 2004, ApJ, 612, L25
O'Connell, R. W., Gallagher, J. S., & Hunter, D. A. 1994, ApJ, 433, 65
Origlia, L., Leitherer, C., Aloisi, A., Greggio, L., & Tosi, M. 2001, AJ, 122, 815
Palmieri, R., Piotto, G., Saviane, I., Girardi, L., & Castellani, V. 2002, A&A, 392, 115
Penny, L. R., & Gies, D. R. 2009, ApJ, 700, 844
Perina, S., Barmby, P., Beasley, M. A., et al. 2009, A&A, 494, 933
Pflamm-Altenburg, J., & Kroupa, P. 2009, MNRAS, 397, 488
Pinsonneault, M. H., & Stanek, K. Z. 2006, ApJ, 639, L67
Piotto, G. 2009 in The Ages of Stars, IAU Symp., 258, 233
Piotto, G., Villanova, S., Bedin, L. R., et al. 2005, ApJ, 621, 777
Pols, O. R., & Marinus, M. 1994, A&A, 288, 475
Renzini, A., & Buzzoni, A. 1986, in Spectral Evolution of Galaxies, ed. C. Chiosi, & A. Renzini, Astrophysics and Space Science Library, 122, 195
Renzini, A., & Fusi Pecci, F. 1988, ARA&A, 26, 199
Ritossa, C. 1996, MNRAS, 281, 970
Salaris, M., Held, E. V., Ortolani, S., Gullieuszik, M., & Momany, Y. 2007, A&A, 476, 243
Salasnich, B., Bressan, A., & Chiosi, C. 1999, A&A, 342, 131
Sana, H., & Evans, C. J. 2010 [arXiv:1009.4197]
Sana, H., Gosset, E., Nazé, Y., Rauw, G., & Linder, N. 2008, MNRAS, 386, 447
Sandage, A., Tammann, G. A., & Reindl, B. 2004, A&A, 424, 43
Sansom, A. E., Izzard, R. G., & Ocvirk, P. 2009, MNRAS, 399, 1012
Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, ApJ, 500, 525
Silva-Villa, E., & Larsen, S. S. 2010, A&A, 516, A10
Silva-Villa, E., & Larsen, S. S. 2011, A&A, 529, A25
Stetson, P. B. 1994, PASP, 106, 250
Taam, R. E., & Sandquist, E. L. 2000, ARA&A, 38, 113
Thim, F., Tammann, G., Saha, A., et al. 2003, ApJ, 590, 256
Tosi, M., Sabbi, E., Bellazzini, M., et al. 2001, AJ, 122, 1271
Vazquez, G. A., Leitherer, C., Heckman, T. M., et al. 2004, ApJ, 600, 162
Vink, J. S., Brott, I., Gräfener, G., et al. 2010, A&A, 512, L7
Vinkó, J., Sárneczky, K., Balog, Z., et al. 2009, ApJ, 695, 619
Walborn, N. R., Prevot, M. L., Prevot, L., et al. 1989, A&A, 219, 229
Walsh, J. R., & Roy, J.-R. 1997, MNRAS, 288, 726
Westmoquette, M. S., Exter, K. M., Smith, L. J., & Gallagher, J. S. 2007, MNRAS, 381, 894
Woosley, S. E. 1988, ApJ, 330, 218