The Astrophysical Journal, 989:114 (25pp), 2025 August 10 https://doi.org/10.3847/1538-4357/ade391

CLUES. III. Do User Choices Impact the Results of SED Fitting? Tests of "Off-the-shelf" Stellar Population and Dust Extinction Models in the CLUES Sample

Andrew Mizener[1], Daniela Calzetti[1], Angela Adamo[2], Aida Wofford[3,4], Matthew J. Hayes[2], John Chisholm[5], Michele Fumagalli[6,7], Svea Hernandez[8], Matteo Maria Messa[9], Linda J. Smith[10], Arjan Bik[11], Kathryn Grasha[12,13,14,15], and Mattia Sirressi[12]

[1] Department of Astronomy, University of Massachusetts Amherst, MA 01003, USA; amizener@umass.edu
[2] Department of Astronomy, Oskar Klein Centre, Stockholm University, AlbaNova University Centre, SE-106 91, Sweden
[3] Instituto de Astronomía, Universidad Nacional Autónoma de México, A.P. 106, Ensenada 22800, BC, México
[4] Department of Astronomy and Astrophysics, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA
[5] Department of Astronomy, University of Texas at Austin, 2515 Speedway, Austin, TX 78712, USA
[6] Dipartimento di Fisica "G. Occhialini," Università degli Studi di Milano-Bicocca, Piazza della Scienza 3, I-20126 Milano, Italy
[7] NAF—Osservatorio Astronomico di Trieste, Via G. B. Tiepolo 11, I-34143 Trieste, Italy
[8] AURA for ESA, Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA
[9] INAF–OAS, Osservatorio di Astrofisica e Scienza dello Spazio di Bologna, via Gobetti 93/3, I-40129 Bologna, Italy
[10] Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA
[11] Department of Astronomy, Stockholm University, Oscar Klein Centre, AlbaNova University Centre, 106 91 Stockholm, Sweden
[12] Research School of Astronomy and Astrophysics, Australian National University, Canberra, ACT 2611, Australia
[13] ARC Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), Australia
[14] Visiting Fellow, Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA
Received 2024 October 9; revised 2025 May 22; accepted 2025 June 2; published 2025 August 7

Abstract

The simple stellar population models produced by stellar population and spectral synthesis (SPS) codes are used as spectral templates in a variety of astrophysical contexts. In this paper, we test the predictions of four commonly used stellar population synthesis codes (YGGDRASIL, BPASS, FSPS, and a modified form of GALAXEV, which we call GALAXEVneb) by using them as spectral templates for photometric spectral energy distribution (SED) fitting with a sample of 18 young stellar clusters. All clusters have existing Hubble Space Telescope (HST) Cosmic Origins Spectrograph far-UV spectroscopy that provides constraints on their ages as well as broadband photometry from HST Advanced Camera for Surveys and Wide Field Camera 3. We use model spectra that account for both nebular and stellar emission, and additionally test four extinction curves at different values of R_V. We find that for individual clusters, choice of extinction curve and SPS model can introduce significant scatter into the results of SED fitting. Model choice can introduce scatter of 34.8 Myr in age, a factor of 9.5 in mass, and 0.40 mag in extinction. Extinction curve choice can introduce scatter of up to a factor of 32.3 Myr in age, a factor of 10.4 in mass, and 0.41 mag in extinction. We caution that because of this scatter, one-to-one comparisons between the properties of individual objects derived using different SED fitting setups may not be meaningful. However, our results also suggest that SPS model and extinction curve choice do not introduce major systematic differences into SED fitting results when the entire cluster population is considered. The distribution of cluster properties for a large enough sample is relatively robust to user choice of SPS code and extinction curve.

Unified Astronomy Thesaurus concepts: Stellar populations (1622); Spectral energy distribution (2129); Young star clusters (1833); Astronomy data modeling (1859); Interstellar dust extinction (837); Ultraviolet spectroscopy (2284); Broad band photometry (184)

Materials only available in the online version of record: machine-readable tables

1. Introduction

One of greatest successes of modern astrophysics is our fairly good understanding of the formation, evolution, and death of stars. Though naturally valuable in its own right, this knowledge is especially crucial to the study of galaxies and the complex stellar populations they contain. In this context, our understanding of stellar evolution is most often leveraged through the techniques of stellar population synthesis (SPS).

SPS in its modern form dates back to work done in 1970s by Beatrice Tinsley and her collaborators. Before Tinsley, attempts to interpret the stellar light emitted from galaxies typically involved constructing a linear combination of individual stellar spectra without physical constraints (known today as the trial-and-error approach; G. Bruzual & S. Charlot 2003). However, this technique suffered from strong degeneracies and was quickly abandoned. Instead, B. M. Tinsley (1968, 1973) and B. M. Tinsley & J. E. Gunn (1976) pioneered the evolutionary population synthesis approach, which computes the time evolution of a stellar population's spectrum by combining a stellar initial mass function (IMF) with stellar evolutionary tracks and modeled or observed stellar spectra (C. Maraston 2005). In the intervening years, a rich library of SPS codes have been produced using some form of the evolutionary population synthesis approach (e.g.,

[15] ARC DECRA Fellow.

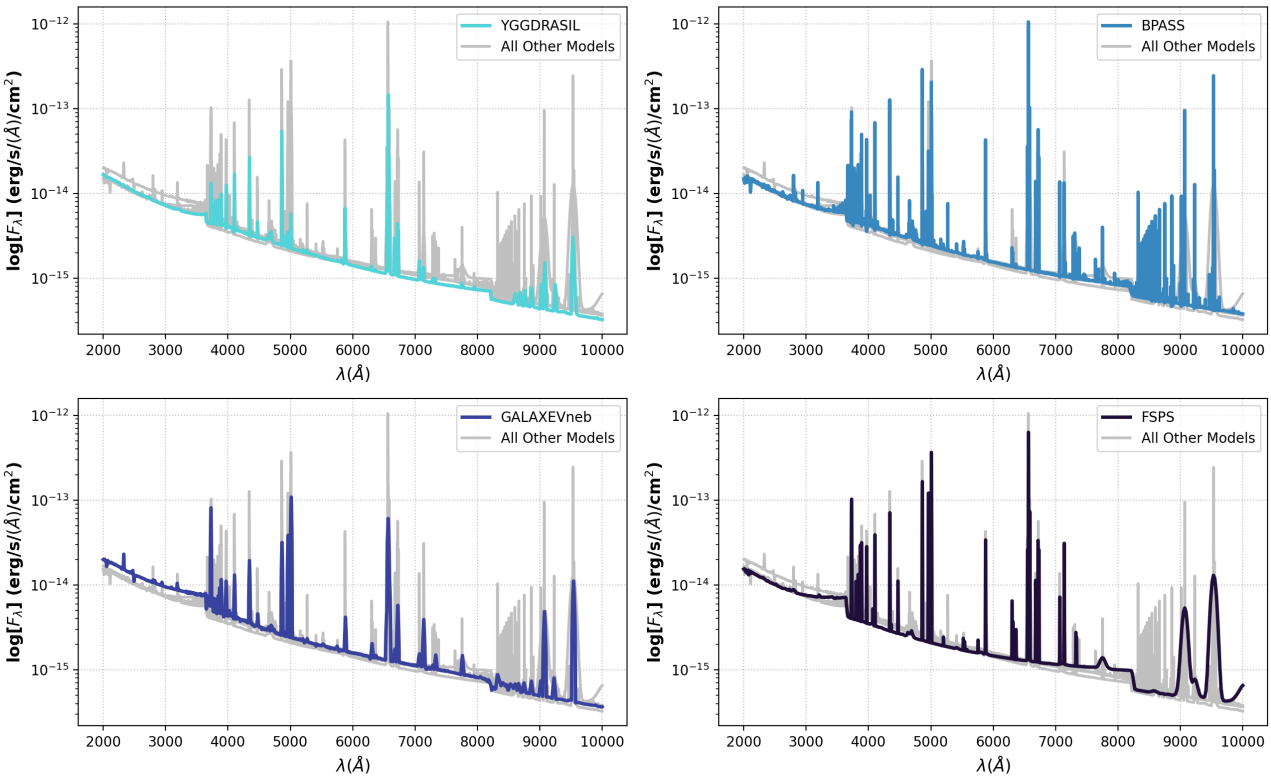


Figure 1. Different SPS codes can produce different spectra when evaluated using the same input parameters. This plot shows spectra for each of the SPS models investigated in this work: YGGDRASIL, FSPS, BPASS (binary), and our modified version of GALAXEV (GALAXEVneb). In each panel, the spectrum produced by the code of interest is shown as in color, while the spectra produced by the other codes are shown in gray for comparison. The spectra shown here are evaluated at log(Age) = 6.5, log(Mass) = 6.0, $E(B-V) = 0.2$, and a metallicity of 0.007. Despite the identical input parameters, differences in how each model handles stellar evolution, stellar atmospheres, and nebular emission produce differences in the output spectra. In this work, we test whether these differences are important to the results of broadband spectral energy distribution (SED) fitting.

C. Leitherer et al. 1999; G. Bruzual & S. Charlot 2003; C. Conroy et al. 2009; E. Zackrisson et al. 2011; J. J. Eldridge et al. 2017; R. Orozco-Duarte et al. 2022) and are in widespread use throughout the literature.

The fundamental product of SPS codes are single stellar populations (SSPs) and their spectra. An SSP represents an instantaneous burst stellar population with homogeneous chemical composition. By convolving SSPs with a star formation history and metallicity distribution, one may produce a model of the spectrum of a composite stellar population, making SSPs notable as a fundamental input to spectral energy distribution (SED) fitting and template-based photometric redshift estimation codes (C. Conroy 2013). Contemporary SED fitting tools like MAGPHYS (E. da Cunha et al. 2008), Prospector (B. D. Johnson et al. 2021), CIGALE (M. Boquien et al. 2019), SLUG (R. L. da Silva et al. 2012), and BAGPIPES (A. C. Carnall et al. 2018) and photometric redshift estimators such as EAZY (G. B. Brammer et al. 2008) vary in specific approach, but all use a combination of SSP spectra from various SPS models to represent the contribution from stellar light.

It is well known that existing SPS models are imperfect. Phenomena such as stellar rotation and binary effects make massive stars challenging to model. This in turn makes models somewhat more uncertain at young ages where these stars dominate the SED (C. Conroy & J. E. Gunn 2010; J. Walcher et al. 2011; C. Conroy 2013). There are also significant modeling challenges at low metallicity, as nearby metal-poor massive stars are rare (C. Conroy & J. E. Gunn 2010; C. Conroy 2013). Without sufficient empirical constraints, model grids that are complete in this region of parameter space are difficult to construct.

As SPS models are a fundamental input to the tools used to derive key galaxy parameters like stellar mass and redshift, it is important to verify that they behave in ways that we understand. Differences in how individual models handle, e.g., stellar evolutionary physics, stellar atmospheric physics, or nebular emission can cause different codes to predict different spectra for an SSP at the same age, mass, extinction, and metallicity (Figure 1). These differences can produce significant differences in intrinsic color. Further, different models are differently sensitive to metallicity effects. Over a similar range in metallicity and age, different regions of color space are accessible as a function of model (Figure 2). In turn, this means that the SSP choice could in principle systematically bias the derived cluster parameters.

Although work has been done to understand how these systematics propagate through to final derived galaxy parameters, the picture is not yet totally clear. It is evident that fits to the same data using different SPS models can produce discrepant results (e.g., A. Muzzin et al. 2009; X. Y. Chen et al. 2010; L. Whitler et al. 2023; J. Tang et al. 2024; B. Wang et al. 2024). However, it remains to be seen under exactly what circumstances these differences are significant, especially when convolved with observational uncertainties. Furthermore, SED fitting within large galaxy samples typically

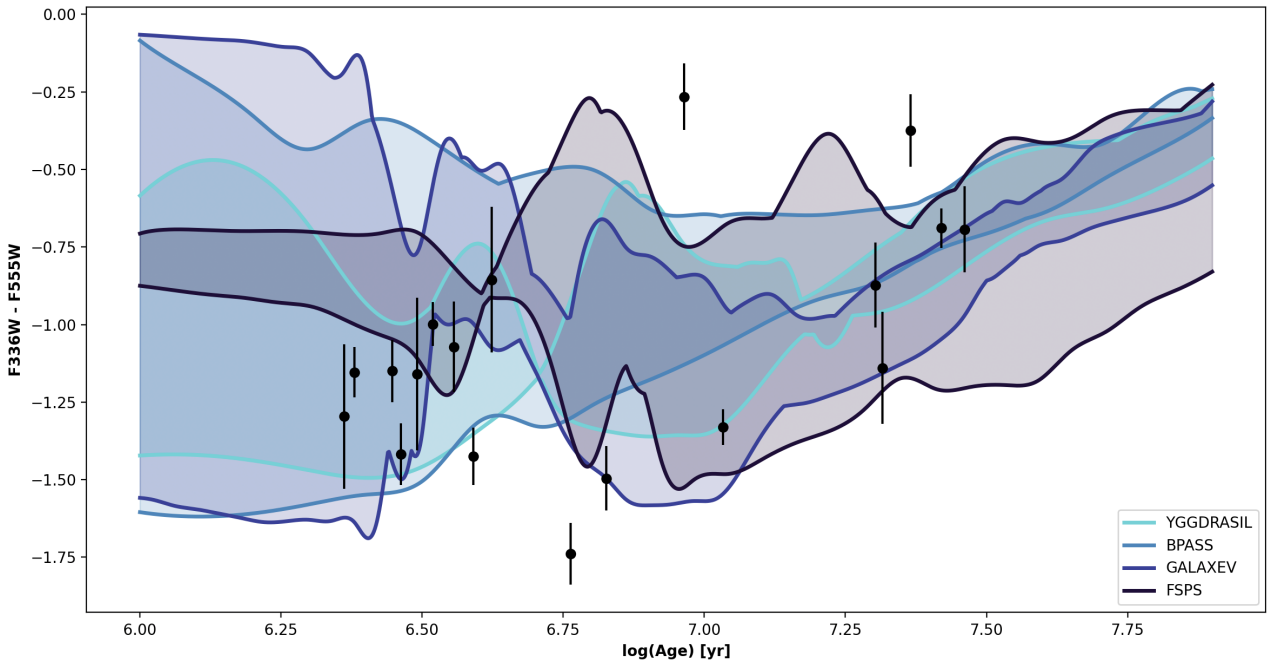


Figure 2. Different SPS models provide different dispersions in color as a function of metallicity and age. In this plot, we show the range in F336W–F555W color accessible by each of the models we tested over the range of metallicity covered by the model. For YGGDRASIL, the metallicity grid ranges from $Z = 0.0004$ to $Z = 0.02$. for BPASS, the metallicity grid ranges for $Z = 0.001$ to $Z = 0.040$. GALAXEVneb covers a range from $Z = 0.0001$ to $Z = 0.05$. We tested FSPS on a metallicity grid covering the same range used by GALAXEVneb. We show this metallicity-dependent color evolution as function of spectroscopic age from (M. Sirressi et al. 2022). We also plot the observed F336–F555W colors of the YSCs used in this work. If the observed data falls outside of the region in color space accessible to a model, the fit will either adjust age and extinction of the model until appropriate colors are achieved or the fitted model will simply be a poor representation of the data.

involves inference with sparse broadband photometry over a limited wavelength range. Since the results of such fitting are used to infer broader astrophysical trends (e.g., S. Tacchella et al. 2022; N. J. Adams et al. 2024; R. Endsley et al. 2023; C. Pacifici et al. 2023; P. G. Pérez-González et al. 2023), it is sensible to investigate if the results obtained are self-consistent regardless of model choice.

The most straightforward way to do this is through comparison to observations of young stellar clusters (YSCs). YSCs, which form nearly instantaneously from a single parent cloud (S. F. Portegies Zwart et al. 2010), are the best physical analogs to SSPs that exist in nature, allowing for very direct comparisons between observations and models. Using YSCs to benchmark SSPs and stellar evolutionary models is a well-established technique (e.g., A. Renzini & F. Fusi Pecci 1988; C. Maraston 2005; X. Y. Chen et al. 2010; A. Wofford et al. 2016).

Young clusters represent an excellent laboratory for benchmarking SPS codes not only because they are young, bright systems, but also because their far-UV (FUV) spectra are easily observed with the Cosmic Origins Spectrograph (COS) on the Hubble Space Telescope (HST). Fits to photospheric and P-Cygni-like lines in FUV spectra provide robust constraints on stellar ages (e.g., A. Wofford et al. 2013; J. Chisholm et al. 2019; M. Sirressi et al. 2022), which allows for the otherwise pernicious age/extinction degeneracy (e.g., P. Anders et al. 2004; A. Bridžius et al. 2008) to be broken. Additionally, using YSCs allows one to bypass the age/IMF degeneracy (A. Wofford et al. 2011).

It is also important to account for nebular gas, particularly when interpreting observations of young systems where the contribution to the SED by nebular emission can be very significant (D. E. Osterbrock 1974; R. Orozco-Duarte et al. 2022; B. Wang et al. 2024). The typical approach is to pass model stellar continuum spectra through a photoionization code like CLOUDY (G. J. Ferland et al. 2017) or MAPPINGS-III (B. A. Groves et al. 2004), which predict emitted spectra by calculating the full radiative transfer through a specified cloud composition and geometry. Because such an approach can be computationally expensive, it is also somewhat common to interpolate from tabulated nebular emission predictions as an alternative (N. Byler et al. 2017). We note that is no one-size-fits-all approach to nebular emission modeling; the important parameters here (e.g., covering fraction, ionization parameter, hydrogen density, and gas abundances) may be reasonable over a range of possible values, depending on the nature of the system at hand.

There are two key effects investigated in this paper. First, we wish to quantify the systematic effects (if any) introduced into the results of photometric SED fitting via the SPS choice. In particular, we investigate four commonly used SPS codes: YGGDRASIL (which is built around Starburst99; E. Zackrisson et al. 2011), BPASS (J. J. Eldridge et al. 2017), FSPS (C. Conroy et al. 2009; C. Conroy & J. E. Gunn 2010), and a custom-modified version of GALAXEV (G. Bruzual & S. Charlot 2003, 2016 version). These codes are associated with publicly available model stellar continuum spectra that are usable "off the shelf" in many applications; as such, we wish to check that fits performed on broadband photometry using different SSPs produce consistent results. We want to test both whether the results produced are self-consistent in a model-to-model sense, as well as whether they are consistent

Table 1
Parent Cluster Sample

Cluster	R.A.	Decl.	Dist. (Mpc)	$E(B-V)_{\rm MW}$ (mag)	Metallicity (Z)	Age (LW) Log(yr)	Mass ($10^6\ M_\odot$)	Extinction (LW) (mag)
NGC 1313-1	49.597	−66.478	4.602	0.096	$0.007^{+0.000}_{-0.002}$	$6.826^{+0.019}_{-0.019}$	$0.15^{+0.03}_{-0.02}$	$0.125^{+0.002}_{-0.002}$
NGC 1313-2	49.659	−66.492	4.602	0.096	$0.004^{+0.008}_{-0.000}$	$7.461^{+0.03}_{-0.03}$	$0.06^{+0.2}_{-0.0}$	$0.197^{+0.004}_{-0.004}$
M74-1	24.157	15.752	9.691	0.062	$0.033^{+0.006}_{-0.001}$	$6.38^{+0.0}_{-0.0}$	$0.07^{+0.03}_{-0.0}$	$0.251^{+0.002}_{-0.002}$
M74-2	24.198	15.774	9.691	0.062	$0.039^{+0.001}_{-0.001}$	$6.362^{+0.0}_{-0.0}$	$0.04^{+0.0}_{-0.0}$	$0.205^{+0.005}_{-0.005}$
NGC 1512-1	60.948	−43.364	11.506	0.009	$0.038^{+0.001}_{-0.002}$	$6.491^{+0.014}_{-0.014}$	$0.03^{+0.0}_{-0.0}$	$0.126^{+0.007}_{-0.007}$
NGC 1512-2	60.976	−43.35	11.506	0.009	$0.040^{+0.000}_{-0.003}$	$6.462^{+0.0}_{-0.0}$	$0.03^{+0.01}_{-0.0}$	$0.101^{+0.002}_{-0.002}$
NGC 1566-2	65.012	−54.941	17.89	0.008	$0.037^{+0.001}_{-0.001}$	$7.316^{+0.023}_{-0.023}$	$0.25^{+0.63}_{-0.0}$	$0.163^{+0.004}_{-0.004}$
NGC 1566-1	64.98	−54.931	17.89	0.008	$0.005^{+0.035}_{-0.000}$	$6.623^{+0.01}_{-0.01}$	$0.8^{+0.93}_{-0.0}$	$0.096^{+0.002}_{-0.002}$
M95-1	160.99	11.704	9.658	0.025	$0.040^{+0.000}_{-0.000}$	$6.447^{+0.0}_{-0.0}$	$0.72^{+0.31}_{-0.0}$	$0.376^{+0.002}_{-0.002}$
NGC 4485-1	187.623	41.698	8.751	0.019	$0.008^{+0.000}_{-0.002}$	$6.763^{+0.03}_{-0.03}$	$0.05^{+0.0}_{-0.02}$	$0.111^{+0.002}_{-0.002}$
NGC 4485-2	187.618	41.69	8.751	0.019	$0.005^{+0.001}_{-0.000}$	$6.519^{+0.0}_{-0.0}$	$0.01^{+0.01}_{-0.0}$	$0.19^{+0.002}_{-0.002}$
M51-1	202.482	47.197	8.589	0.031	$0.040^{+0.000}_{-0.004}$	$6.556^{+0.012}_{-0.012}$	$0.07^{+0.05}_{-0.0}$	$0.242^{+0.002}_{-0.002}$
M51-2	202.507	47.23	8.589	0.031	$0.039^{+0.001}_{-0.010}$	$7.303^{+0.037}_{-0.037}$	$0.2^{+0.0}_{-0.14}$	$0.22^{+0.004}_{-0.004}$
NGC 7793-1	359.453	−32.582	3.403	0.017	$0.028^{+0.001}_{-0.000}$	$6.591^{+0.0}_{-0.0}$	$0.02^{+0.0}_{-0.0}$	$0.135^{+0.002}_{-0.002}$
NGC 7793-2	359.4	−32.595	3.403	0.017	$0.004^{+0.000}_{-0.000}$	$7.42^{+0.025}_{-0.025}$	$0.01^{+0.02}_{-0.0}$	$0.137^{+0.003}_{-0.003}$
NGC 4656-2	190.986	32.171	7.876	0.012	$0.001^{+0.000}_{-0.000}$	$7.033^{+0.052}_{-0.052}$	$0.45^{+0.0}_{-0.35}$	$0.239^{+0.003}_{-0.003}$
NGC 4656-1	190.99	32.17	7.876	0.012	$0.001^{+0.000}_{-0.000}$	$7.365^{+0.043}_{-0.043}$	$0.7^{+0.18}_{-0.26}$	$0.254^{+0.005}_{-0.005}$
NGC 4449-1	187.046	44.094	3.889	0.017	$0.001^{+0.000}_{-0.000}$	$6.964^{+0.052}_{-0.052}$	$0.31^{+0.07}_{-0.05}$	$0.266^{+0.003}_{-0.003}$

Note. This table shows the fundamental properties of our cluster sample. Ages, masses, extinctions, and metallicities are based on FUV spectroscopy and are taken from M. Sirressi et al. (2022). We report the light-weighted age and extinction alongside the mass of the dominant stellar population; for a complete catalog of the spectroscopic fit results, see M. Sirressi et al. (2022).

(This table is available in machine-readable form in the online article.)

with more reliable cluster properties (predominantly age) obtained via FUV photospheric line fits. We also want to determine what sort of constraints we can place on age, mass, and extinction when only sparse photometry is available.

Second, and similarly, we wish to investigate whether the adopted dust extinction curve (as parameterized by the specific-to-total extinction, R_V) has any systematic effect on the results of SED fitting. We use the SSPs produced by these four SPS codes, as well as a set of four extinction curves spanning a reasonable parameter space in R_V, as inputs for the SED fitting of 20 YSCs from the CLUES (M. Sirressi et al. 2022) sample.

We describe the CLUES sample and our photometric procedure in Section 2. We describe in detail the models investigated in Section 3. We describe our model post-processing and fitting procedure in Section 4. The results of these fits are shown in Section 5, and we conclude with a discussion of our results in Section 6.

2. Data and Photometry

2.1. The Clues Project

This work is part of the larger CLUES (CLusters In the Uv as EngineS; ID 15627, PI: Adamo) program, which has used FUV spectroscopy from HST's COS instrument to investigate feedback from YSCs (M. Sirressi et al. 2022, 2024).[16] A comprehensive description of the target selection criteria for CLUES may be found in M. Sirressi et al. (2022). The CLUES sample consists of 20 young, UV bright star clusters drawn from sources in the LEGUS (Legacy Extragalactic UV Survey; IDs 10402 and 13364, PI: Calzetti) footprint (D. Calzetti et al. 2015a). Clusters in CLUES were selected such that a wide range of host galaxy types (dwarfs, spirals, and interacting systems) were included. Individual clusters were drawn from within this larger sample with apparent UV magnitude brighter than <18.0 mag in F275W, color excess $E(B-V)<0.3$ mag, and age <3 Myr in preliminary SED fitting. This FUV flux cutoff acts as a rough selection on mass, nominally limiting the sample to clusters above $\sim 10^4 M_\odot$. These clusters are sufficiently massive to avoid stochastic sampling of the IMF (M. R. Krumholz et al. 2015; R. Orozco-Duarte et al. 2022), making deterministic models such as those tested here a reasonable assumption.

We note that because these clusters are selected to be very young, they are reasonable SSP analogs. Though including older globular clusters would allow us to test model performance at larger log(Age), older clusters are also more susceptible to contamination by stellar populations of multiple ages. Older clusters have been used to benchmark the performance of stellar evolutionary and populations synthesis models in the literature (e.g., P. Marigo et al. 2008), but their inclusion is out of the scope of this work. In this paper, we examine all YSCs in the CLUES sample except for those in NGC 5253. These clusters do not dominate the optical/near-IR flux within their respective apertures and so are not appropriate targets for our analysis.

2.2. FUV Spectroscopic Fits with COS

Fits to the FUV spectra for clusters in the CLUES sample were performed by M. Sirressi et al. (2022). We defer to their paper for a complete description of their methods but describe them in brief here. They begin by pre-processing the observed UV spectra to match the resolution of the SB 99

[16] https://archive.stsci.edu/hlsp/clues

Table 2
Photometry (All Clusters)

Cluster	F275W	σ_{F275W}	F336W	σ_{F336W}	F43(5/8)W	$\sigma_{F43(5/8)W}$	F(555/606)W	$\sigma_{F(555/606)W}$	F814W	σ_{F814W}
M95-1	14.699 (W)	0.072 (W)	15.124 (W)	0.07 (W)	15.672 (W)	0.065 (W)	16.273 (W)	0.072 (W)	17.692 (W)	0.084 (W)
NGC 4449-1	14.857 (W)	0.083 (W)	15.115 (W)	0.075 (W)	15.121 (A)	0.069 (A)	15.38 (A)	0.076 (A)	15.792 (A)	0.049 (A)
NGC 4485-2	16.383 (W)	0.064 (W)	16.865 (W)	0.062 (W)	17.477 (A)	0.033 (A)	17.864 (A606)	0.032 (A606)	19.126 (W)	0.064 (W)
NGC 4485-1	16.134 (W)	0.057 (W)	16.81 (W)	0.06 (W)	17.522 (A)	0.044 (A)	18.55 (A606)	0.08 (A606)	19.647 (W)	0.114 (W)
NGC 1512-2	17.063 (W)	0.072 (W)	17.656 (W)	0.07 (W)	18.47 (W)	0.065 (W)	19.074 (A)	0.072 (A)	20.16 (A)	0.084 (A)
NGC 1512-1	18.069 (W)	0.185 (W)	18.666 (W)	0.206 (W)	19.274 (W)	0.245 (W)	19.825 (A)	0.135 (A)	20.936 (A)	0.269 (A)
NGC 7793-1	15.7 (W)	0.047 (W)	16.34 (W)	0.051 (W)	17.076 (W)	0.061 (W)	17.765 (W)	0.076 (W)	19.172 (A)	0.149 (A)
M51-1	15.377 (W)	0.047 (W)	15.888 (W)	0.053 (W)	16.42 (A)	0.106 (A)	16.959 (A)	0.136 (A)	18.041 (A)	0.207 (A)
NGC 7793-2	16.38 (W)	0.062 (W)	16.842 (W)	0.057 (W)	16.947 (W)	0.043 (W)	17.531 (W)	0.029 (W)	18.342 (A)	0.03 (A)
M51-2	16.863 (W)	0.121 (W)	17.296 (W)	0.102 (W)	17.636 (A)	0.086 (A)	18.169 (A)	0.091 (A)	19.065 (A)	0.105 (A)
NGC 1313-1	15.055 (W)	0.065 (W)	15.594 (W)	0.06 (W)	16.413 (A)	0.074 (A)	17.089 (A)	0.085 (A)	18.308 (A)	0.114 (A)
NGC 1313-2	16.566 (W)	0.113 (W)	17.043 (W)	0.098 (W)	17.291 (A)	0.08 (A)	17.736 (A)	0.099 (A)	18.284 (A)	0.067 (A)
M74-2	17.812 (W)	0.161 (W)	18.277 (W)	0.162 (W)	19.032 (A)	0.126 (A)	19.573 (A)	0.168 (A)	20.751 (A)	0.2 (A)
NGC 4656-2	14.834 (W)	0.03 (W)	15.347 (W)	0.027 (W)	15.961 (W)	0.032 (W)	16.678 (W)	0.051 (W)	17.886 (W)	0.053 (W)
NGC 1566-1	17.152 (W)	0.139 (W)	17.652 (W)	0.152 (W)	17.953 (W)	0.169 (W)	18.507 (W)	0.179 (W)	19.346 (W)	0.253 (W)
M74-1	16.938 (W)	0.089 (W)	17.333 (W)	0.064 (W)	17.937 (A)	0.042 (A)	18.487 (A)	0.05 (A)	19.556 (A)	0.059 (A)
NGC 4656-1	16.527 (W)	0.095 (W)	16.909 (W)	0.09 (W)	16.927 (W)	0.082 (W)	17.283 (W)	0.076 (W)	17.798 (W)	0.062 (W)
NGC 1566-2	16.649 (W)	0.091 (W)	17.226 (W)	0.099 (W)	17.763 (W)	0.132 (W)	18.366 (W)	0.151 (W)	19.537 (W)	0.253 (W)

Note. This table reports the logarithm of the photometric flux in each band in STmag. The (W) or (A) after each entry denotes whether the image used was taken using Wide Field Camera 3 (WFC3) or the Advanced Camera for Surveys (ACS). For ACS F435W and WFC3 F438W, this letter also denotes which filter is used. We also use this notation to specify two clusters (in NGC 4485) for which ACS F606W measurements were used instead of WFC3 or ACS F555W.

(This table is available in machine-readable form in the online article.)

(C. Leitherer et al. 1999, 2014) models used in their fits and use the observed C III 1176 Å and C III 1247 Å line widths to determine the (kinematic Doppler) redshift and kinematic line broadening appropriate for each source. They used Geneva high-mass-loss stellar evolutionary tracks (G. Meynet et al. 1994), a Salpeter IMF between stellar masses 0.1 and 120 $M_{\odot}$, and no stellar rotation. An instantaneous burst star formation history was assumed.

Parameters estimated through fitting to stellar P-Cygni lines are relatively robust. To summarize from J. Chisholm et al. (2019), C IV P-Cygni absorption depends on metallicity, and the emission depends on age. The strength of the N V P-Cygni profile is also very sensitive to stellar age. Old populations have strong C III and Si III photospheric absorption features, while these features are undetected in younger populations. Joint measurement of multiple features in absorption and emission is thus a powerful probe of stellar population properties. Model-related systematics are also worth noting. D. A. Berg et al. (2024) used this method with both Sb99 and BPASS models. They found generally good agreement between the two, with typical age offsets $\sim$1.6 Myr and metallicity offsets around the 20% level. Naturally, calibration of the measured relationship between age or metallicity and P-cygni line properties will only be as good as the calibration of the underlying models.

The fitting procedure follows J. Chisholm et al. (2019). They fit out broad Lyα absorption with a Voigt profile and mask out bad regions of each spectrum (e.g., detector gaps) by eye. ISM absorption lines are also masked. Two fits were performed for each cluster: one using a single SSP, and another using two independent SSPs. Two SSPs were necessary to obtain good fits in situations where stellar populations separate from the target cluster contributed significantly to the measured spectrum. For each cluster, the results of each fit were evaluated with the Akaike information criterion estimator in order to determine which run (single or double-population) to adopt. For each cluster, the best-fit values of each component stellar population, as well as a light-weighted average, are reported. A table summarizing the location and host galaxy of each source is provided in Table 1. We also provide the metallicity of each cluster as reported by M. Sirressi et al. (2022).

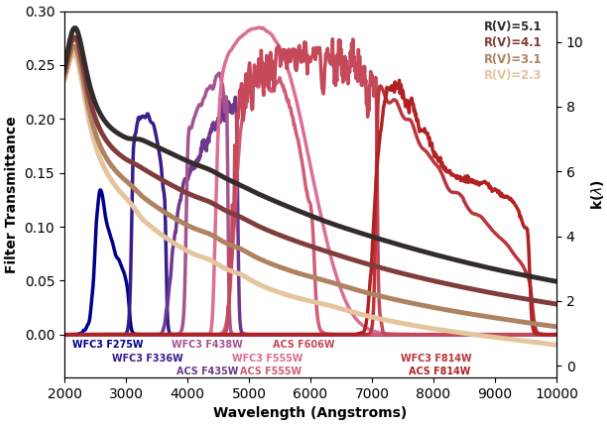
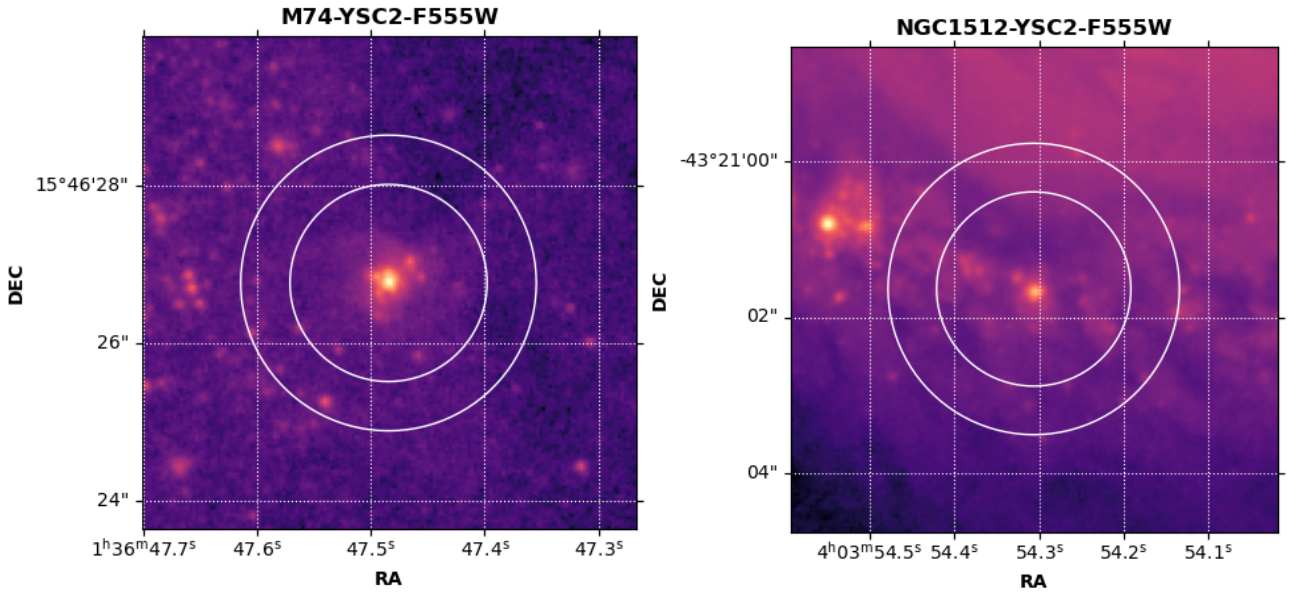


Figure 3. The filters and extinction curves used in this work. For each individual cluster, we have photometry in exactly five of the filters shown.

We use the results of these fits as loose baselines for model validation, and use the FUV-derived metallicities in our optical photometric fits. As such, it is reasonable to ask whether a direct comparison can truly be made between cluster parameters inferred via spectroscopic FUV-based fits and those from our optical photometric SED fits. Indeed, M. Sirressi et al. (2022) found poor agreement between their spectroscopic fits and preliminary photometric fits. In particular, they found that the spectroscopic and photometric ages were in agreement for only about 50% of the clusters,



(a) An aperture containing a relatively isolated cluster.

(b) An aperture containing a significant amount of diffuse emission from the background galaxy.

Figure 4. We use a large photometric aperture (necessary to match our results with COS FUV spectroscopy), which means that we must be careful to perform an appropriate background subtraction. Two example photometric apertures and annuli are shown, both in WFC3 F555W: one example where the contribution to the total aperture flux from the galaxy background is small (M74-YSC2) and one where the galaxy background is bright and spatially complex (NGC 1512-YSC2).

with the performance generally worsening with increasing age. They found metallicities that tend toward solar or supersolar values. In the cases where FUV fits infer a low metallicity, the measured values tend to be in agreement with nebular abundance measurements from the literature.

It is suggested that the general age offset between the two methods, as well as the greater age sensitivity of the FUV fits, are a result of differences in how relevant spectral features evolve with time. Optical photometric colors can be relatively insensitive to age at some epochs (e.g., >3 Myr), and this technique generally suffers from the well-known age-extinction degeneracy (e.g., M. Fouesneau et al. 2012). By comparison, features in the FUV spectrum such as massive-star P-Cygni lines can be strong and are strongly age sensitive. As such, age offsets between photometric SED fits and FUV spectroscopic fits does not necessarily signal any problem with the models used. Rather, it is a signpost for challenges fundamental to broadband SED fitting. Although changing the underlying models used for the spectral fits would change the best-fit YSC properties, it is unlikely that any choice of model would produce significantly better agreement with the broadband fits.

2.3. Data

We obtained archival HST Advanced Camera for Surveys (ACS) and Wide Field Camera 3 (WFC3) data for each of the clusters in Table 1 via the Mikulski Archive for Space Telescopes (MAST). All data used in this paper are part of the LEGUS (D. Calzetti et al. 2015a) program. We retrieved images of our target clusters in the WFC3 F275W, WFC3 F336W, ACS F435W or WCF3 F438W, ACS F606W or WCF3 F555W, and ACS or WFC3 F814W. We show these filter profiles in Figure 3.

2.3.1. Photometry

We measured the flux of each source in each available filter via aperture photometry with a $2\farcs5$ diameter aperture, selected to match the COS aperture. COS suffers from nontrivial vignetting (B. L. James et al. 2022); in order to allow us to fairly compare the result of our photometric SED fits with the FUV spectral fits from M. Sirressi et al. (2022), this must be taken into account. We thus weight the flux of each pixel within the aperture by the COS vignetting function before taking any measurement (P. Goudfrooij et al. 2010). Background subtraction was performed via standard sigma-clipping methods with $\sigma = 3$ on an annulus with an outer radius of $1\farcs875$, starting at the outer edge of the aperture. Two example photometric apertures are shown in Figure 4.

Also important is an estimate of how the (potentially bright and complex) local galaxy background impacts our measured photometry. In order to estimate this contribution, we calculate the sigma-clipped standard deviation in an annulus centered on each cluster with a starting width of $0\farcs5$. The width of this annulus is increased by $0\farcs1$ until the standard deviation converges within 10% for three consecutive iterations. For two sources (M95-1 and NGC 1512-2), the size of the converged annulus includes significant surface brightness variation that artificially inflates the measured photometric uncertainty. To bring these estimates down into a more realistic range, we instead assign to these clusters a photometric uncertainty in each band equal to median relative uncertainty in that band averaged over the rest of the sample. Instrumental contributions to the photometric uncertainty are calculated using standard recipes.

Because the vignetting correction we apply to our photometry means that not all pixels in the aperture have equal weight, we correct the uncertainties in each band by the ratio of the vignetted aperture volume to the volume of an aperture with identical size but idealized uniform response (analogous

Table 3
Best-fit Parameters: All Clusters, All Models, All Curves

Model	R_V	$\mathrm{Age_{ML}}$ $\mathrm{Log_{10}(yr)}$	$\mathrm{Age_{MED}}$ $\mathrm{Log_{10}(yr)}$	$\mathrm{Mass_{ML}}$ $\mathrm{Log_{10}}(M_\odot)$	$\mathrm{Mass_{MED}}$ $\mathrm{Log_{10}}(M_\odot)$	$E(B-V)_{\mathrm{ML}}$ (mag)	$E(B-V)_{\mathrm{MED}}$ (mag)	χ^2/ν
M74-1								
YGGDRASIL	2.3	6.053	$6.138^{+0.139}_{-0.092}$	4.877	$4.865^{+0.029}_{-0.042}$	0.192	$0.205^{+0.024}_{-0.021}$	4.183
	3.1	6.093	$6.186^{+0.157}_{-0.12}$	4.945	$4.925^{+0.038}_{-0.055}$	0.202	$0.218^{+0.026}_{-0.024}$	7.19
	4.1	6.348	$6.258^{+0.159}_{-0.161}$	4.973	$5.001^{+0.047}_{-0.059}$	0.249	$0.233^{+0.027}_{-0.027}$	0.445
	5.1	6.5	$6.328^{+0.17}_{-0.194}$	5.024	$5.076^{+0.055}_{-0.057}$	0.258	$0.243^{+0.026}_{-0.029}$	0.754
BPASS	2.3	6.735	$6.705^{+0.046}_{-0.448}$	4.87	$4.832^{+0.047}_{-0.273}$	0.114	$0.144^{+0.069}_{-0.048}$	1.35
	3.1	6.724	$6.449^{+0.278}_{-0.153}$	4.91	$4.731^{+0.172}_{-0.082}$	0.128	$0.199^{+0.029}_{-0.077}$	1.071
	4.1	6.465	$6.48^{+0.215}_{-0.115}$	4.824	$4.827^{+0.105}_{-0.059}$	0.223	$0.207^{+0.028}_{-0.062}$	0.08
	5.1	6.522	$6.541^{+0.139}_{-0.094}$	4.897	$4.921^{+0.073}_{-0.058}$	0.206	$0.195^{+0.038}_{-0.044}$	0.315
GALAXEV	2.3	6.797	$6.812^{+0.029}_{-0.02}$	4.613	$4.64^{+0.038}_{-0.039}$	0.039	$0.035^{+0.019}_{-0.018}$	1.436
	3.1	6.797	$6.811^{+0.029}_{-0.019}$	4.625	$4.65^{+0.039}_{-0.04}$	0.04	$0.036^{+0.02}_{-0.019}$	1.463
	4.1	6.796	$6.808^{+0.029}_{-0.018}$	4.639	$4.662^{+0.042}_{-0.044}$	0.041	$0.037^{+0.021}_{-0.02}$	1.53
	5.1	6.795	$6.807^{+0.029}_{-0.017}$	4.653	$4.673^{+0.046}_{-0.046}$	0.041	$0.037^{+0.022}_{-0.02}$	1.645
FSPS	2.3	6.582	$6.585^{+0.017}_{-0.01}$	4.377	$4.366^{+0.049}_{-0.056}$	0.067	$0.061^{+0.027}_{-0.034}$	11.232 (*)
	3.1	6.581	$6.583^{+0.014}_{-0.01}$	4.408	$4.401^{+0.053}_{-0.064}$	0.073	$0.07^{+0.026}_{-0.031}$	10.846 (*)
	4.1	6.581	$6.582^{+0.011}_{-0.009}$	4.447	$4.444^{+0.057}_{-0.067}$	0.08	$0.079^{+0.024}_{-0.028}$	10.219 (*)
	5.1	6.582	$6.582^{+0.01}_{-0.008}$	4.485	$4.485^{+0.064}_{-0.069}$	0.085	$0.085^{+0.024}_{-0.025}$	9.504 (*)

Notes. This table reports the fitted parameters produced during this work. Each row represents a fit produced using a unique combination of cluster, extinction curve, and model SSP. Column 1 denotes which cluster and model were used in a given row. Column 2 shows the R_V used in a given row. Column 3 is the maximum-likelihood ("most probable") age for a given fit. Column 4 is the median ("most typical") age for a given fit, along with approximate uncertainties (calculated as the 16th and 84th percentiles within the posterior). Column 5 is the maximum-likelihood stellar mass. Column 6 is the median stellar mass and associated uncertainties. Column 7 is the maximum-likelihood extinction in magnitudes. Column 8 is the median extinction and associated uncertainties. Column 9 reports the reduced χ^2 associated with the maximum-likelihood fit in each row. Rejected fits are included in this table and are denoted with an asterisk.

(This table is available in its entirety in machine-readable form in the online article.)

to the JvM correction commonly performed with interferometric data; e.g., I. Czekala et al. 2021). We apply the same correction to our background subtraction in order to avoid over-subtraction. A typical correction factor is around ∼0.7. This vignette and corresponding correction is not wavelength dependent.

In M. Sirressi et al. (2022), no background subtraction was performed in the FUV. Instead, their SED fits were performed using a combination of two independent stellar populations. To first approximation, the dominant stellar population of the two can be considered to represent the cluster, with the other taken to represent background young stars. We are unable to perform an analogous two-population fit in this paper, as LEGUS only includes five filters and such an analysis would be associated with six degrees of freedom. The background still exists, however, and must be accounted for. We note that because older stellar populations are more dominant in our optical photometry than they are in the FUV, our backgrounds are likely stronger than those encountered by M. Sirressi et al. (2022). In the case of, for example, M95 YSC1, the ratio of background-subtracted to un-subtracted flux in our aperture is 0.96 in F150LP (analogous to the COS spectroscopic range) but just 0.23 in F814W. Our method is not a one-to-one equivalent to the procedure of M. Sirressi et al. (2022), but it should at least minimize the impact of background stellar populations on our derived cluster quantities.

All photometry was corrected for Milky Way foreground extinction following standard procedure (e.g., D. Calzetti et al. 2015b) using extinction maps from E. F. Schlafly & D. P. Finkbeiner (2011). The galaxies in which our clusters are located were bright enough to be masked during the creation of these extinction maps, and so the Milky Way extinction estimates should not be biased by their presence. Measurements were performed using the AstroPy (Astropy Collaboration et al. 2013) and PhotUtils (L. Bradley et al. 2016) packages. Typical photometric uncertainties are between 3% and 7%; however, a number of measurements are significantly more uncertain (>10% uncertainty). These are generally driven by our unusually large apertures; in these cases, there is generally a gradient or complex structure in the background making the background subtraction highly uncertain (Figure 2, right panel). We remind the reader that we use a large, fixed aperture in order to match the COS aperture. Our measurements are given in Table 2.

3. Models

In order to determine to what extent the choice of spectral synthesis code has an effect on the results of SED fitting, we performed SED fitting using four common spectral synthesis models: namely, YGGDRASIL, BPASS, FSPS, and a modified form of GALAXEV which we call GALAXEVneb. In the following sections, we briefly summarize the approach taken by each spectral synthesis code used.

3.1. YGGDRASIL

The YGGDRASIL (E. Zackrisson et al. 2011) spectral synthesis code was originally developed to model the SEDs of Population III galaxies. It is not a complete end-to-end isochrone synthesis code by itself; rather, it is designed to

handle arbitrary star formation histories by modifying the SSPs derived from various other population synthesis models. In practice, however, the precomputed Population I SSPs produced based on Starburst99 with Padova evolutionary tracks (C. Leitherer et al. 1999; G. A. Vázquez & C. Leitherer 2005a) and a Kroupa IMF (P. Kroupa 2001) are commonly used. The mass-loss rates used are described in detail in G. A. Vázquez & C. Leitherer (2005a) but are generally a modified form of those produced by C. de Jager et al. (1988). The stellar atmospheres of L. J. Smith et al. (2002) are used for hot stars, and the ATLAS9+Phoenix models are used for stars of spectral type B and below (T. Lejeune et al. 1997, 1998). YGGDRASIL has become a relatively common choice of spectral synthesis model for those studying star formation and young stellar populations (e.g., S. T. Linden et al. 2023), as it provides easily accessible SSPs including nebular emission predictions at a range of metallicities and gas covering fractions.

The contribution of nebular emission to the output SED is computed following the approach of E. Zackrisson et al. (2001). In brief, the input stellar continuum spectrum is passed through the photoionization code CLOUDY (G. J. Ferland et al. 2017) at each age step. A spherical, ionization-bounded, constant density nebula is assumed. Rather than specifying an ionization parameter U directly, it is instead derived from the state of the input stellar population. The ionizing photon production rate, Q(H), is set by the age of the stellar population, and the inner radius of the cloud is set to

$$R_{\rm in} = 100 R_{\odot} \left(\frac{L}{L_{\odot}} \right)^{1/2}.$$

A hydrogen density $n(H) = 100\,{\rm cm}^{-3}$ is used. Combined, these three parameters are sufficient to specify an ionization parameter. We use the version of YGGDRASIL that assumes a gas covering fraction $f_{\rm COV} = 0.5$. The nebular gas is assumed to be dust-free. For the purposes of this work, we used the standard YGGDRASIL data products described above, based on Starburst99 and a Kroupa IMF with an upper mass limit of 100 $M_{\odot}$ and a lower mass limit of 0.1 $M_{\odot}$.

3.2. BPASS

Most existing population synthesis models use evolutionary tracks/isochrones that follow the evolution of single, isolated stars. Up to 70% of massive stars will exchange mass with a companion at some point during their evolution (H. Sana et al. 2012), and this may have important consequences on the shape of stellar population SEDs and their time evolution.

The Binary Population and Spectral Synthesis (BPASS; J. J. Eldridge et al. 2017) suite of binary stellar evolution models and synthetic stellar populations was designed from the ground-up to improve upon existing population synthesis codes by including interacting binary effects. Because most existing isochrones do not include binary effects, the authors of BPASS calculate stellar evolutionary tracks themselves. The core of BPASS is its single-star stellar evolution code, the history of which can be traced back to the 1971 Cambridge STARS code (P. P. Eggleton 1971). It uses the standard methods of L. G. Henyey et al. (1964) to solve for the detailed stellar structure using an adaptive numerical mesh and time step. In all models, the authors apply the stellar wind mass-loss rates of C. de Jager et al. (1988) unless the star is of spectral type O or B, in which case the mass-loss rates of J. S. Vink et al. (2001) are used. All models are evolved from the zero-age main sequence (ZAMS) up until the end of core carbon burning, or neon ignition for the most-massive models.

The BPASS binary-star models are identical to single-star models in most respects, except for the fact that the authors allow for additional mass loss or gain via binary interactions. During the evolution of a binary, only one star is followed in detail at a time. After Roche lobe overflow, the code calculates a mass-loss rate following J. R. Hurley et al. (2002). If the mass transfer from a companion exceeds 5% of a star's initial mass, it is assumed to either rejuvenate the star or lead to quasi-homogeneous evolution (QHE). Rejuvenation is modeled by resetting the star to the ZAMS and evolving it normally thereafter. QHE is modeled by assuming the relevant stars are fully mixed throughout their main-sequence lifetimes, and is assumed to occur in low-metallicity stars with $M > 20 M_{\odot}$ after mass transfer. Spectral synthesis is performed by combining the aforementioned population synthesis results with the BaSeL v3.1 (P. Westera et al. 2002) stellar atmosphere library, with a few minor corrections. The fiducial BPASS models use a Kroupa-like IMF (P. Kroupa et al. 1993) with an upper mass limit of 300 $M_{\odot}$ and a lower mass limit of 0.1 $M_{\odot}$.

We use the 2.2.1 version of BPASS (2018 release) with nebular emission predictions by (L. Xiao et al. 2018). When calculating nebular emission, nebulae were assumed to be spherical, ionization-bounded, and at a constant hydrogen density. For our analysis, we use the model grid calculated using hydrogen density $\log({\rm cm}^{-1}) = 2.3$, a 1.0 covering fraction, and an ionization parameter at the Strömgen radius of -2. The nebular gas is assumed to be dust-free.

3.3. GALAXEV

GALAXEV is a library of evolutionary SPS models computed using the isochrone synthesis code of G. Bruzual & S. Charlot (2003). The earliest iterations of these models were notable at the time of their initial release (1993; A. Bruzual & G. Charlot 1993) for providing spectra at 3 Å resolution between 3200 and 9500 Å. Spectral evolution was also computed at a lower resolution between 91 Å and 160 μm across the same grids.

As these models have been continually updated with improved stellar physics, including detailed modeling of TP-AGB and post-AGB stars, they have remained quite popular and are used widely throughout the literature. For this work, we use the 2016 revision of the GALAXEV SSPs produced with a Kroupa IMF (P. Kroupa 2001) with an upper mass limit of 100 $M_{\odot}$ and a lower mass limit of 0.1 $M_{\odot}$, the STELIB stellar spectral library (J. F. Le Borgne et al. 2003), and Padova 1994 stellar evolutionary tracks (M. Alongi et al. 1993; A. Bressan et al. 1993; F. Fagotto et al. 1994a, 1994b). As the public versions of these models do not include nebular emission, we have performed our own nebular emission predictions as described in the next section.

3.4. FSPS

The flexible SPS (FSPS) code (C. Conroy et al. 2009; C. Conroy & J. E. Gunn 2010) was originally developed to address perceived deficiencies in how existing SPS models handled uncertainties in stellar evolutionary physics and the

IMF. Like most other models used in this paper, it uses the standard isochrone synthesis approach to generate SSPs. The original models used the Padova (L. Girardi et al. 2000) stellar evolutionary tracks and BaSeL (T. Lejeune et al. 1997; P. Westera et al. 2002) spectral libraries. From the end-user perspective, FSPS is particularly attractive because its outputs are accessible through a Python interface (D. Foreman-Mackey et al. 2014) and are easily customized. Users can, for example, select from a range of different stellar evolutionary tracks, spectral libraries, and IMFs.

They also use a model for nebular emission that computes continuum and line emission for arbitrary stellar populations without requiring the computational expense of a full radiative-transfer treatment like CLOUDY (N. Byler et al. 2017) at run-time. By default, a spherical nebula at the metallicity of the underlying stellar population with ionization parameter -2 and constant hydrogen density $100\,\mathrm{cm}^{-1}$ is assumed. They assume an ionizing photon escape fraction of zero (i.e., a covering fraction of 1.0). The only user-tunable parameters are the ionization parameter and gas-phase metallicity. We keep these at their defaults ($\log U = -2$ and a metallicity matched to the underlying stellar population). The nebular gas is assumed to be dust-free. We use the default Kroupa IMF with with an upper mass limit of 120 $M_\odot$ and a lower mass limit of 0.08 $M_\odot$ We also use the MIST (J. Choi et al. 2016) isochrones and MILES (A. Vazdekis et al. 2010) stellar spectral library, which are the FSPS defaults.

Following our general approach of using SPS models as they would be encountered by a typical user, we adopt the default evolutionary tracks (MIST; J. Choi et al. 2016) and default spectral library (MILES; A. Vazdekis et al. 2010) adopted by D. Foreman-Mackey et al. (2014), as well as a Kroupa IMF (P. Kroupa 2001).

4. Model Post-processing and Photometric SED Fitting

4.1. Model Post-processing

We cannot use raw stellar continuum spectra in our SED fits directly. We must first add nebular emission to GALAXEV and include dust extinction in all of our model spectra.

4.1.1. Nebular Emission Predictions for GALAXEV

Stars younger than about 10 Myr and more massive than about 15 $M_\odot$ produce significant ionizing flux, which in turn produces nebular continuum and line emission upon contact with the surrounding gas (D. E. Osterbrock 1974). Different ways of handling the nebular emission can produce significantly different predicted fluxes, even in broadband filters (E. Zackrisson et al. 2001; A. Wofford et al. 2016; N. Byler et al. 2017; R. Orozco-Duarte et al. 2022; B. Wang et al. 2024). Metallicity can have a strong impact on nebular emission, as the ionizing output of a stellar population is metallicity-dependent. When using clusters to benchmark SPS model performance, it is common to handle nebular emission uniformly such that only differences in the underlying stellar continuum spectra drive differences in inferred cluster properties (e.g., A. Wofford et al. 2016). This is a necessary approach, as it can (in principle) tease out how different assumptions about stellar evolutionary physics affect one's results. However, we do not adopt it in this paper. Instead, we are interested in comparing these models as they exist "off the shelf," as they would be encountered by a typical end user, in order to understand the systematics introduced by their use.

This includes systematics introduced by nebular emission predictions, which may not be consistent between models even if the assumptions made by any individual model are physically reasonable. We note that default model parameters are not always used and may not be appropriate for every problem; regardless, testing them remains a valuable exercise. Three of the four SPS codes we examine in this paper come with nebular emission predictions included; however, the GALAXEV models do not. Because we are studying young clusters in this analysis, we must produce nebular emission predictions for these models before proceeding.

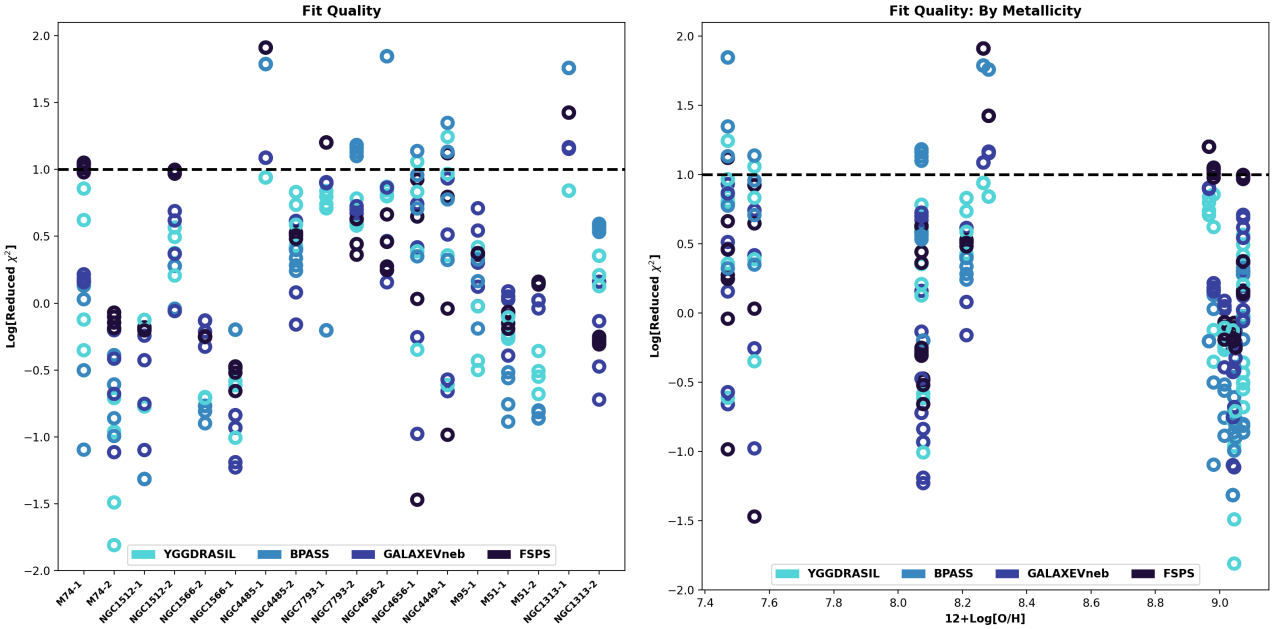
We used CLOUDY (G. J. Ferland et al. 2017) to produce nebular emission predictions for the GALAXEV SSPs, attempting to roughly compromise between the choices made by the other models we test. The key nebular parameters are a covering fraction of 1.0, hydrogen density of $100\,\mathrm{cm}^{-1}$, CLOUDY default H II region abundances, and a default spherical geometry. The stopping criteria were a temperature of 100 K or if the log of the ratio of electron to total hydrogen densities fell below -2. An ionization parameter of -2 was used to match the existing predictions for BPASS and FSPS. Though some work suggests a lower ionization parameter, around -3 is preferred (e.g., J. Gutkin et al. 2016). Particularly when trying to reproduce the properties of high-redshift systems, it has been found that the ionization parameter does not have a strong effect on fit quality within the range -2 to -3 (J. Chisholm et al. 2019).

We rescaled our CLOUDY output such that the total power of the nebular continuum and line emission was equal to the total power below 912A in the input stellar continuum spectra. This is a simplifying assumption that adopts a 100% absorption fraction of the ionizing photons. While some ionizing photons may escape the immediate surroundings of a YSC, this fraction is generally <50% (M. S. Oey et al. 2007). The nebular gas is assumed to be dust-free.

4.1.2. Dust Extinction

Dust extinction must be accounted for in order to make any meaningful comparison between model and observed SEDs. The shape of the extinction curve along any given line of sight is not known a priori; rather, a curve must be adopted as an assumption during SED fitting. Extinction curve shapes are strongly wavelength dependent and vary from sight line to sight line; in particular, extinction curves diverge strongly at and blueward of the 2175 Å feature, as well as at and redward of the 10 μm silicate feature. Between these bumps, extinction curves are nearly linear (e.g., E. L. Fitzpatrick & D. Massa 2007) but vary in slope and overall normalization, which can generally be described in tandem via the V-band selective-to-total extinction: $R_V = A(V)/E(B-V)$. The Milky Way average value of $R_V = 3.1$ is commonly adopted (e.g., E. L. Fitzpatrick 1999), but individual lines of sight vary greatly, with $R_V = 2.5$–5.5 typically reported (D. Massa et al. 2020; K. D. Gordon et al. 2023).

In order to explore the effects, if any, that different dust extinction curves have on the results of broadband optical SED fitting, we test a set of extinction curves from K. D. Gordon et al. (2023) selected to span this parameter space in R_V. Namely, we test four different values of R_V: 2.3 (probing the low end of the distribution), 3.1 (matching the Milky Way value), 4.1 (approximately matching the canonical D. Calzetti



(a) Reduced χ^2 as a function of cluster. The horizontal line denotes $\chi^2/\nu = 10.0$, our cutoff for inclusion.

(b) Reduced χ^2 as a function of metallicity, as measured in the FUV

Figure 5. This plot shows the χ^2 for each fit of each cluster. Sixteen fits are performed per cluster, one for each unique combination of SPS model and extinction curve. The left panel shows all fits as a function of cluster; the right panel shows all fits as a function of metallicity. We reject 45 of 288 fits due to unacceptable reduced χ^2. The horizontal dashed line represents the maximum χ^2 we consider acceptable. Poor fits are partially a consequence of our large apertures; if an aperture contains multiple bright star clusters, the assumption of an SSP breaks down. We also note that all models struggle to reproduce the measured photometry at low metallicity. The metallicities shown here are those measured by M. Sirressi et al. (2022), associated with the dominant stellar population.

et al. 2000 value), and 5.1 (probing the high end of the distribution). We follow the approach of D. Calzetti et al. (2015a) in our handling of dust by adopting a foreground screen of the form $F(\lambda)_{\rm out} = F(\lambda)_{\rm model} 10^{-0.4E(B-V)k(\lambda)}$ where $k(\lambda)$ is the extinction curve and $E(B-V)$ is the color excess in magnitudes. The four K. D. Gordon et al. (2023) extinction curves we test are shown in Figure 2. As provided, these curves were normalized such that $A(\lambda)/A(V) = 1.0$ regardless of their actual R_V; we thus multiply these input curves by their R_V such that they have the proper selective-to-total extinction before use. We show these extinction curves in Figure 3.

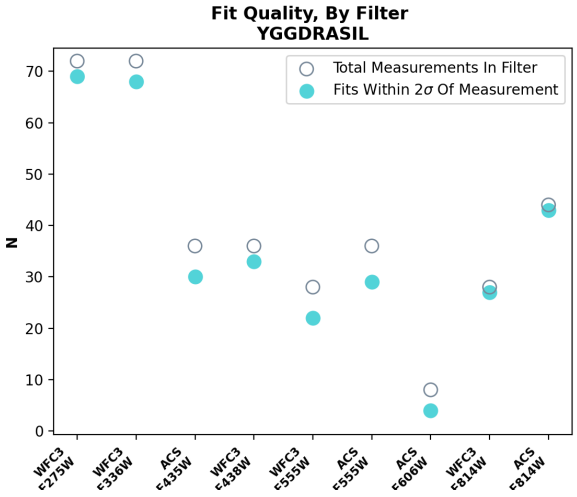
4.2. SED Fitting with Markov Chain Monte Carlo

The probability surfaces associated with SED fitting in YSCs can be complicated and are often multimodal (e.g., M. R. Krumholz et al. 2015), so we want to explore the posterior distributions of our fits and any degeneracies between parameters as completely as possible. We thus constrain model parameters (age, mass, $E(B-V)$) from the photometry with via a Markov Chain Monte Carlo (MCMC) approach with emcee (D. Foreman-Mackey et al. 2013). We perform a fit for each unique combination of cluster, SPS model, and extinction curve. Our fitting procedure is as follows:

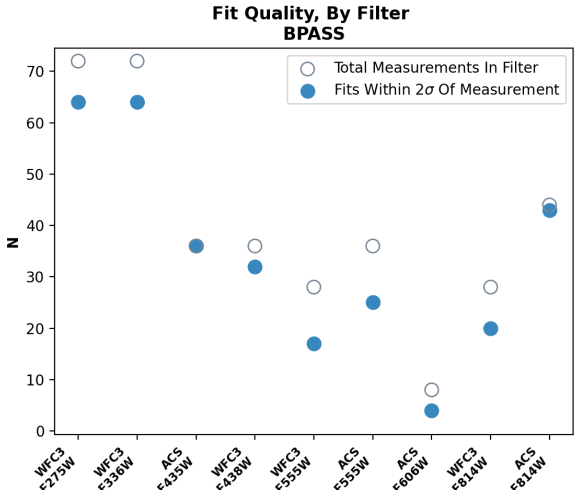
1. Maximum-likelihood estimators require an initial guess to function properly. Because the probability surfaces we are working with can be complex, it is important to have a reasonable initial guess. Thus, we begin by performing a brute-force initial fit. We evaluate over 200-step uniform grids from 0–1.0 in $E(B-V)$ and 6.0–7.9 in log(Age). We do not have enough data to constrain metallicity, so it is kept constant at the closest model grid point to the cluster's (FUV measured) metallicity. This is an important limitation of our analysis and means that our uncertainties are probably somewhat underestimated. However, we note that it is a common approach when fitting star clusters with sparse photometry (e.g., D. Calzetti et al. 2015b). We construct an effective mass grid by scaling model flux in the reddest available filter (ACS or WFC3 F814W) to the observed flux in that filter, and assembling a 300-step uniform grid from 0.5 times the scaled value to 5.0 times the scaled value. We then calculate the χ^2 at each grid point, selecting the solution corresponding to the minimum value.
2. Once an initial approximate solution has been identified, we calculate a refined maximum-likelihood solution using the SciPy Nelder-Mead solver. Instead of using a pre-calculated grid, we now interpolate over age (using a cubic spline), extinction, and mass such that the solver can move smoothly through the parameter space. We assume an uninformative uniform prior covering the same parameter space as the previous step. We apply the same overall bounds to this solver that we used to construct the grids in the previous step. For both this maximum-likelihood calculation and the MCMC analysis described below, we use the likelihood function:

$$\mathcal{L} = -\frac{1}{2}\sum_{i=1}^{5}\left(\frac{\log_{10}(\boldsymbol{d}_i) - \log_{10}(\boldsymbol{m}_i)}{\log_{10}(e) \times (\boldsymbol{\sigma}_i/\boldsymbol{d}_i)}\right)^2$$

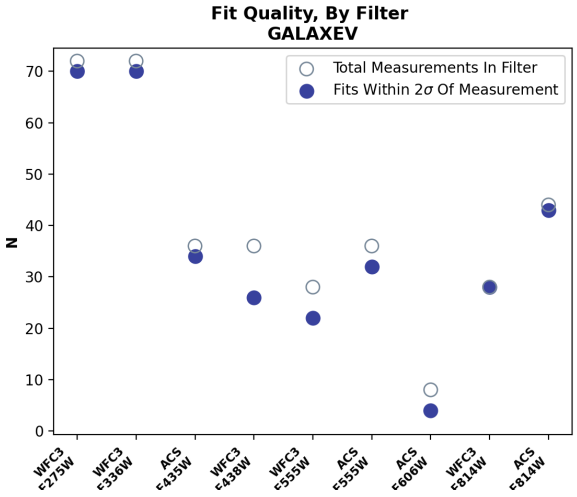
where $\boldsymbol{d}_i$ is the measured flux in filter i, $\boldsymbol{m}_i$ is the model photometry in filter i, and $\boldsymbol{\sigma}_i$ is the photometric uncertainty in filter i, summed over all five filters.
3. We then use MCMC to explore the posterior using emcee (D. Foreman-Mackey et al. 2013). Because multimodal



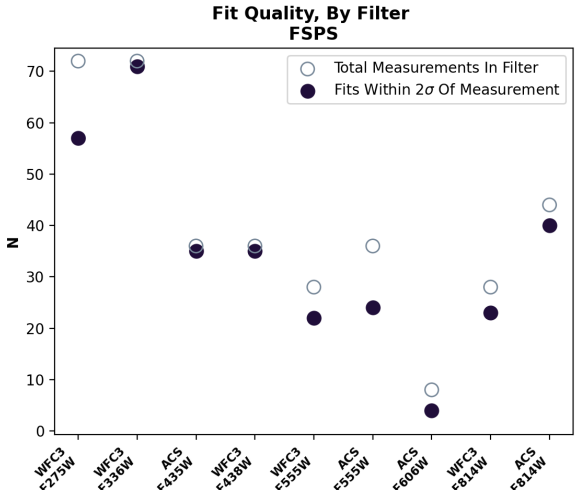
(a) Performance of YGGDRASIL as a function of band



(b) Performance of BPASS as a function of band



(c) Performance of GALAXEVneb as a function of band



(d) Performance of FSPS as a function of band

Figure 6. These plots show, for every accepted fit, the number of attempts to fit each band (open circles) and the number of fits that successfully reproduced the flux in that band (solid circles). All of the models we tested perform similarly as a function of band. That is, we do not find any cases in which, for example, a model performs significantly better in the red part of the spectrum.

posteriors are a possibility, we use a combination of differential evolution (B. Nelson et al. 2014) and Snooker differential evolution (C. J. F. ter Braak & J. A. Vrugt 2008), which move at 80% and 20% probability, respectively. We initialize 50 walkers in a tight Gaussian ball around the maximum-likelihood solution and use uninformative uniform priors corresponding to the same bounds described in earlier steps. In order to ensure sampling of the equilibrium distribution, we run 30,000 steps of burn-in before computing the chain for 7500 steps. This is sufficient to ensure that all chains have run for at least 50 autocorrelation times, as recommended by emcee. We verify the autocorrelation time of each chain hits this benchmark after fitting is complete. We save the entire computed chain (excluding burn-in) for further analysis.

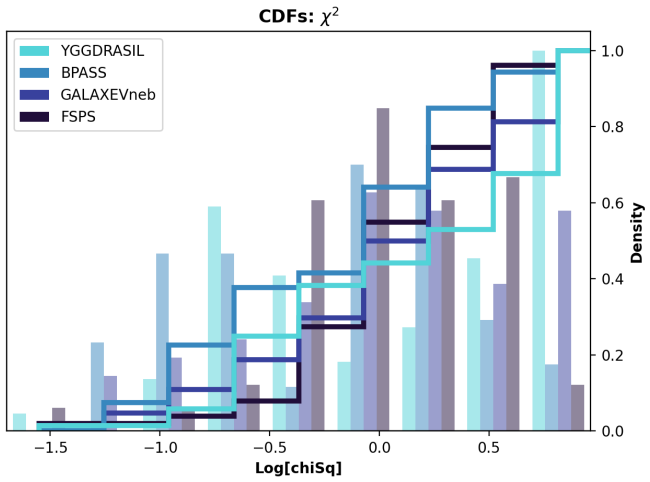
Finally, we generate the following products for each unique combination of cluster, SPS model, and extinction curve:

1. Maximum-likelihood solutions for age, $E(B-V)$, and stellar mass (i.e., the "most probable" values, following A. Wofford et al. 2016).
2. The full 3D posterior distribution of these parameters, given the SPS model and extinction curve.
3. The median (50th percentile) age, $E(B-V)$, and stellar mass in the posterior (i.e., the "most typical" values from A. Wofford et al. 2016).
4. The 1σ upper and lower errors (16th and 84th percentiles) on the median.

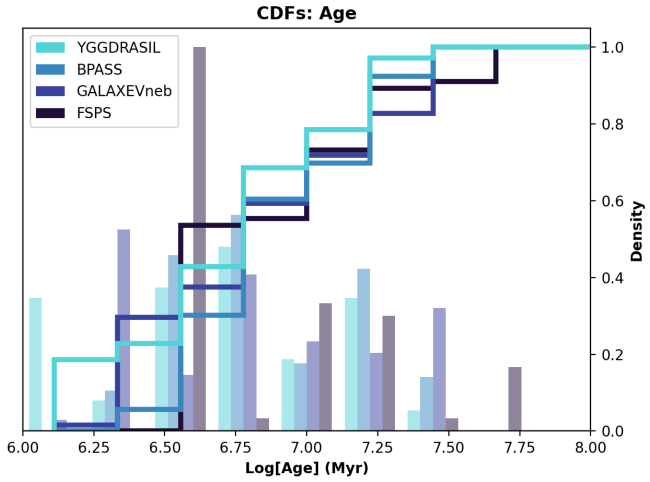
Note that in this paper, we define stellar mass as the total formed stellar mass rather than the surviving stellar mass at a given age.

5. Results

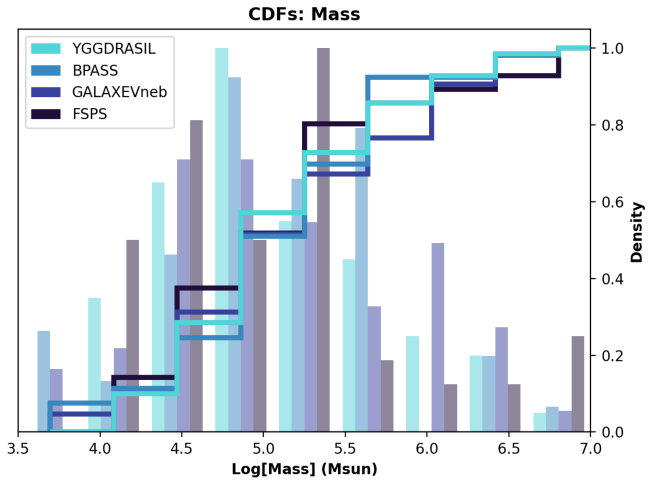
The results of our fits are provided in Table 3, where we list the most probable and most typical ages, masses, and



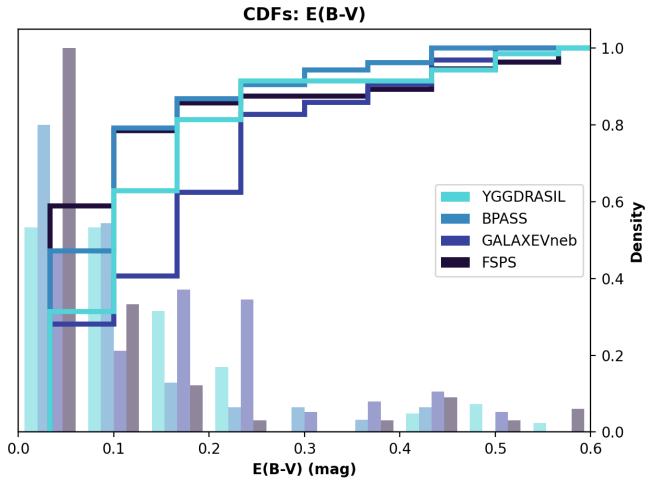
(a) CDFs and histograms for all accepted fits in χ^2, by SPS model.



(b) CDFs and histograms for all accepted fits in Log(Age), by SPS model



(c) CDFs and histograms for all accepted fits in Log(Mass), by SPS model



(d) CDFs and histograms for all accepted fits in E(B-V, by SPS model

Figure 7. This plot shows cumulative distribution functions (CDFs) and histograms of maximum-likelihood fitted parameters, as well as best-fit χ^2 values, divided by SPS model. In many cases, comparing the CDFs of fitted parameters on a model–model basis using the K-S test demonstrates that they were drawn from different parent distributions. This suggests that systematic differences between the results produced using different SPS models may exist. From this perspective, we find that age and color excess are most sensitive to model choice, while mass is relatively insensitive.

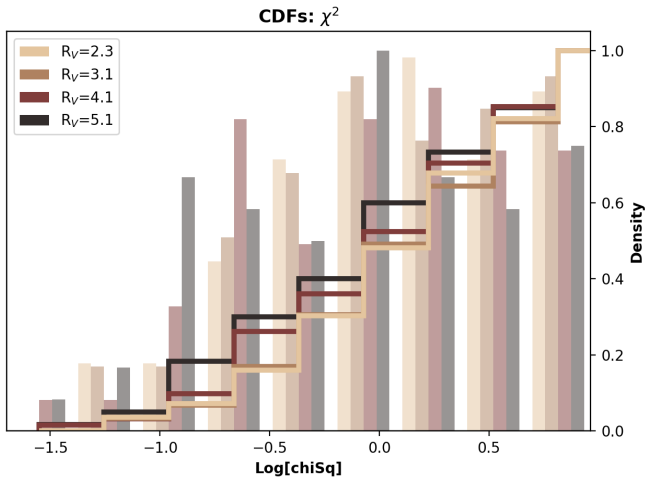
extinctions for each fit. We also provide the reduced χ^2 as described below.

5.1. Fit Quality

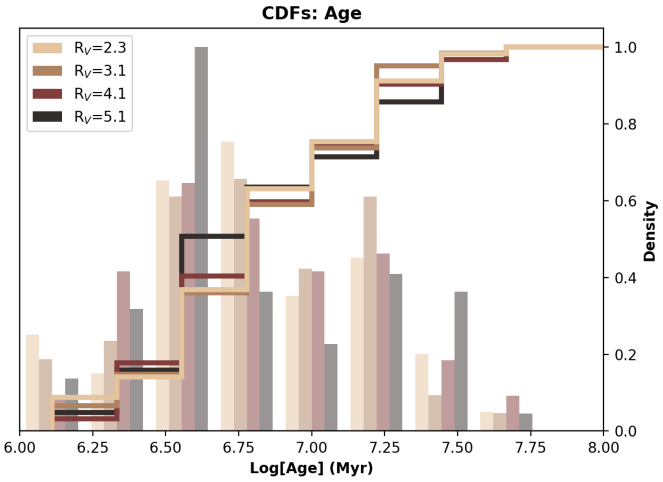
Fits with a large reduced χ^2 are not physically meaningful. This occurs when the model is an inappropriate description of the data: in our sample, two phenomena are likely responsible for these poor fits. First, this can occur when multiple bright star clusters contribute meaningfully to the flux within the aperture, breaking the assumption of an SSP. Second, and perhaps more importantly, it appears that all models perform poorly at low metallicity (as measured in the FUV). We note that we were unable to attempt fits using multiple stellar populations or free metallicity, as we only have photometry in five bands. We reject all fits with $\chi^2/\nu > 10$ before proceeding with any further analysis. This threshold allows us to remove the worst fits from further consideration and roughly symmetrizes our χ^2 distribution. We plot the reduced χ^2 of all fits performed in Figure 5.

Out of 288 fits performed, we reject 45 ($\sim$16%) on this basis. Dividing these rejected fits by the SPS model, we find that 19 of the 45 used BPASS, 16 used FSPS, two used YGGDRASIL, and eight used GALAXEVneb. Dividing them by R_V, we find a more uniform picture, with 15 rejected fits having an R_V of 2.3, 11 having an R_V of 3.1, 10 having an R_V of 4.1, and nine having an R_V of 5.1. We plot all fits performed as a part of this work in Figure 5 as a function of cluster and SPS model. We note that fitting performance is generally poor at lower metallicities. We also note that the rejected fits are not necessarily associated with the clusters with the lowest photometric uncertainty.

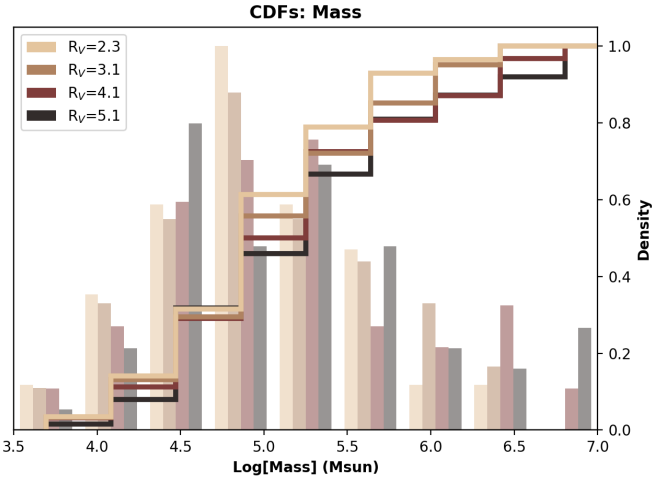
We also want to understand how well each model performs as a function of filter. We want to ensure that our fits are



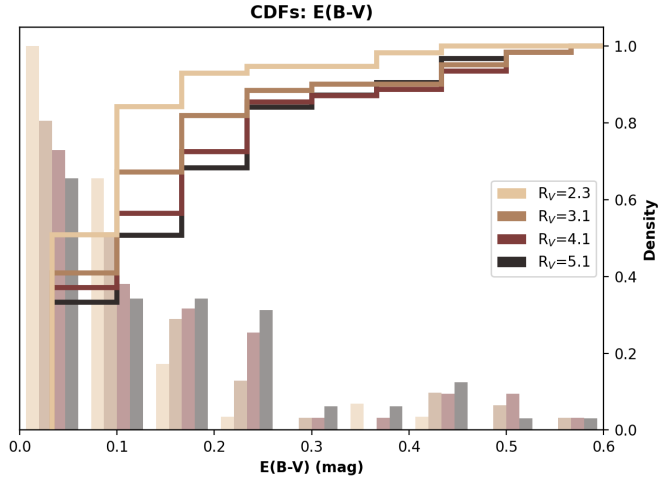
(a) CDFs and histograms for all accepted fits in χ^2, by extinction curve



(b) CDFs and histograms for all accepted fits in Log(Age), by extinction curve



(c) CDFs and histograms for all accepted fits in Log(Mass), by extinction curve



(d) CDFs and histograms for all accepted fits in E(V-B), by extinction curve

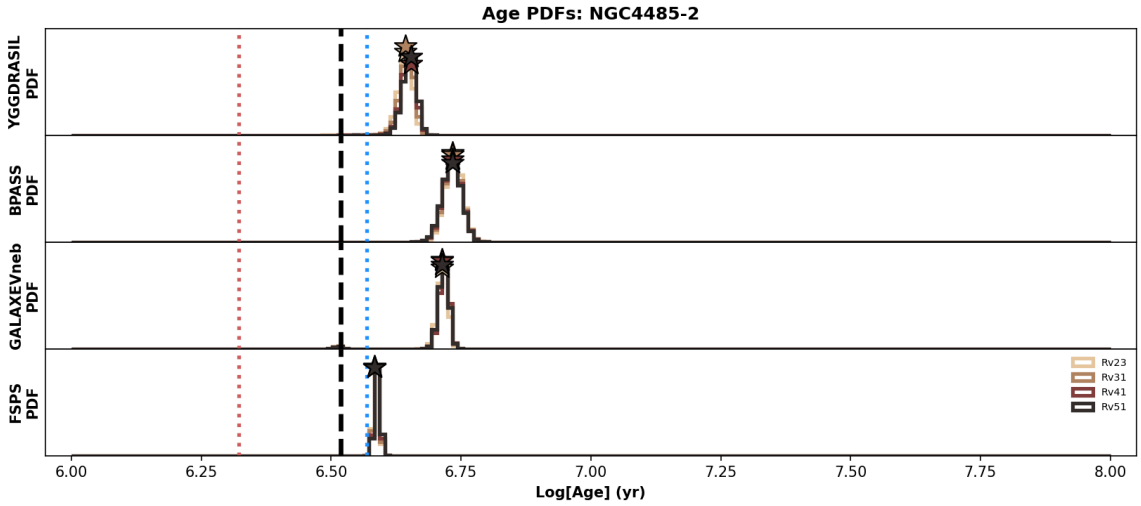
Figure 8. Unlike SPS model choice, the choice of extinction curve R_V does not appear to strongly affect derived cluster parameters outside of total extinction. Here, we show CDFs and histograms of maximum-likelihood fitted parameters, as well as best-fit χ^2 values, divided by extinction curve. As a function of R_v, the overall sample of fitted parameters is not consistent with being drawn from different parent distributions on an R_V–R_V basis. From the perspective of the K-S test, the only potentially significant R_V–R_V differences are between the distributions of $E(B-V)$.

assigning each filter approximately equal by weight—if, for example, the fits in bluer filters were in agreement with the data significantly more often than the redder filter, there would be cause for concern. To test this, we plot the total number of observations in each band and the number of (acceptable) fits that successfully reproduced the flux in that band as a function of the SPS model in Figure 6. We find no strong evidence of color bias—each model performs similarly well across the whole spectral range. We confirm this by checking the reduced chi-square as a function of filter, as described in the Appendix.

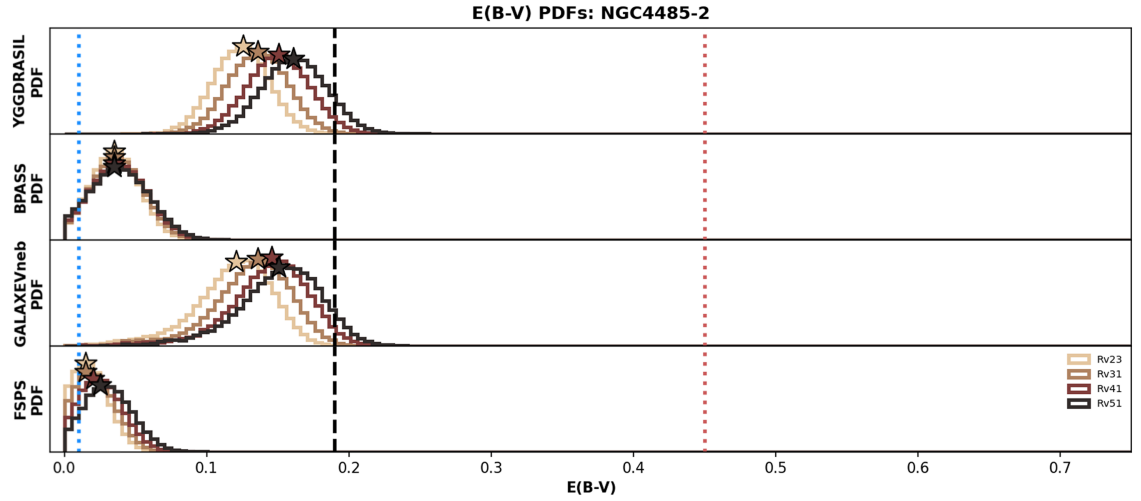
5.2. Model–Model Comparison across All Clusters

Before analyzing our results on a cluster-by-cluster basis, we first examine our fitted results as a population. We want to test whether our samples of fitted parameters produced using each SPS model (across all clusters) are drawn from different parent distributions. If true, this would suggest that there is a systematic offset between the results inferred using any given pair of models—the models would essentially be "seeing" different cluster populations.

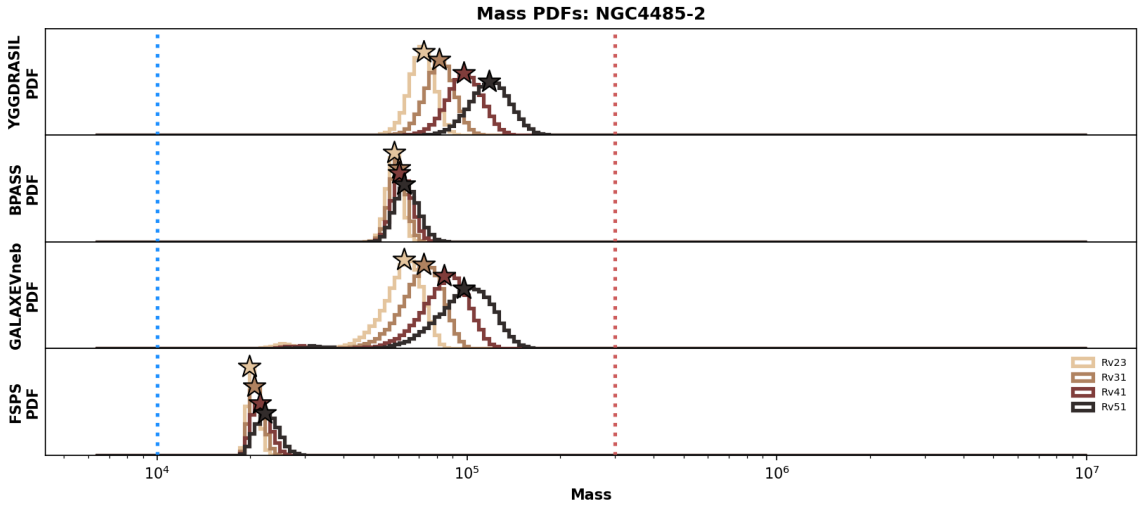
To test this, we calculate the two-sample Kolmogorov–Smirnov statistic between the samples of maximum-likelihood χ^2, age, mass, and $E(B-V)$ for each unique pair of SPS models. Only fits with a reduced $\chi^2 < 10$ are included in this analysis. We note that the Kolmogorov–Smirnov (K-S) test can only demonstrate inconsistency; alternatively stated, just because we cannot reject the null hypothesis does not necessarily mean there is no difference between the results of a given model–model or R_V–R_V pair. The corresponding p-values are shown in Table 3. We further show the cumulative distribution functions (CDFs) and histograms for reduced χ^2, age, mass, and $E(B-V)$ as a function of SPS model in Figure 7. In these plots (as well as those in Figure 7), we use



(a) 1-D PDFs in Log(Age)

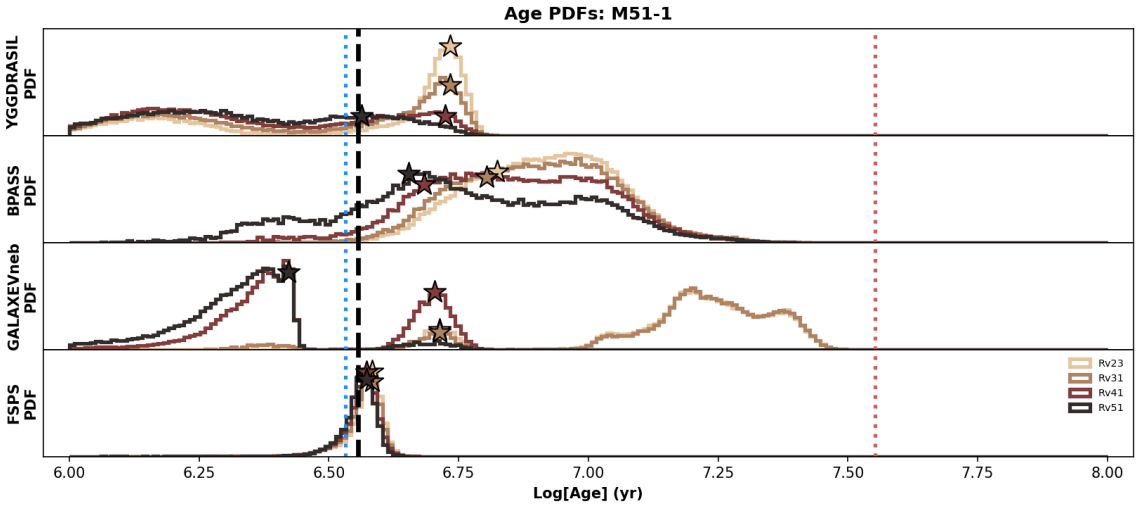


(b) 1-D PDFs in E(V-B)

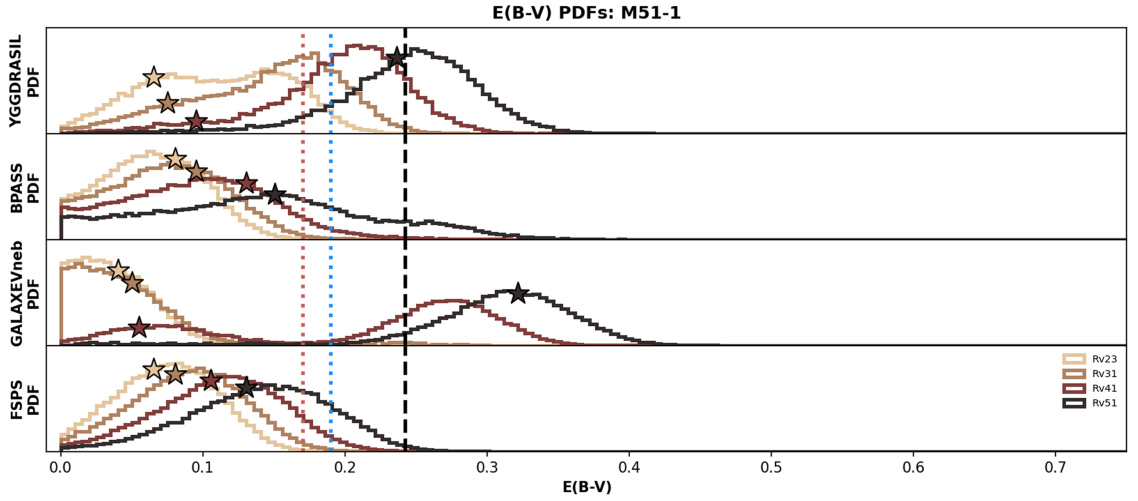


(c) 1-D PDFs in Log(Mass)

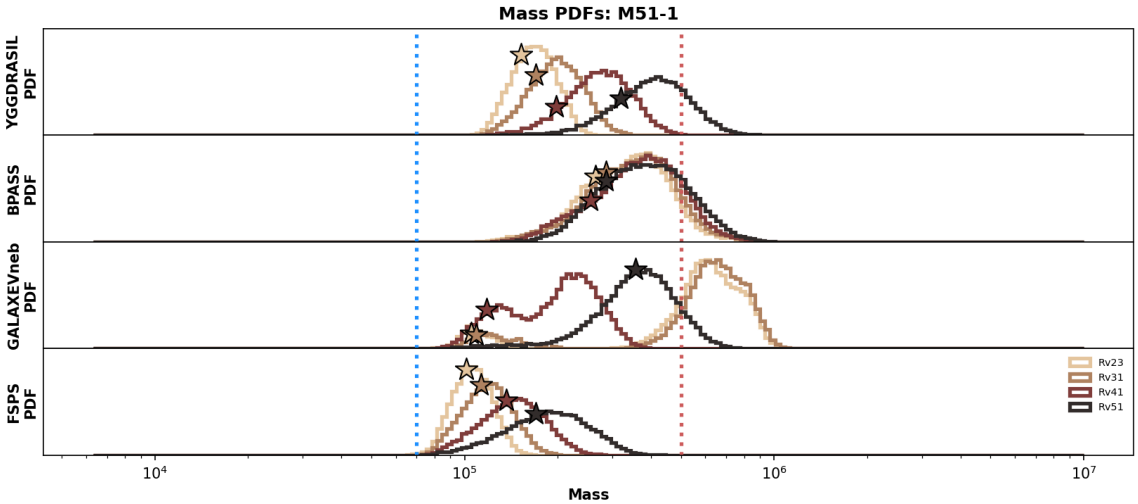
Figure 9. NGC 4485-2: a relatively well-behaved example. In this plot, the black dashed line is the M. Sirressi et al. (2022) "light-weighted" solution, the dotted blue line is the major population in the "two-population" solution, and the dotted red line is the minor population in that solution. Here, the choice of extinction curve affects derived cluster properties only minimally. Cluster properties as a function of SPS model are also broadly similar.



(a) 1-D PDFs in Log(Age)



(b) 1-D PDFs in E(V-B)



(c) 1-D PDFs in Log(Mass)

Figure 10. M51-1: a relatively poorly behaved example. Here, the choice of extinction curve strongly affects derived cluster properties, and most models produce a multimodal posterior. Model–model differences are also strong. In this plot, the black dashed line is the M. Sirressi et al. (2022) "light-weighted" solution, the dotted blue line is the major population in the "two-population" solution, and the dotted red line is the minor population in that solution. Some fits in GALAXEVneb and FSPS were rejected due to unacceptable χ^2 for this cluster despite being close to the spectroscopic results in age.

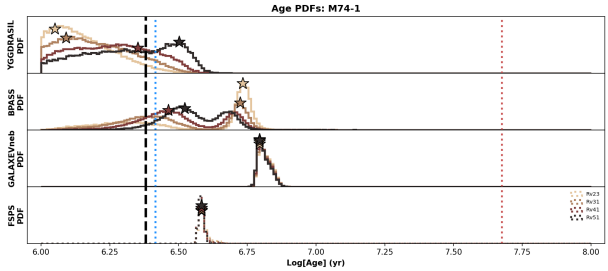


Figure 11. Sometimes, fitting the same data with different models produces results that differ not only in the sense of absolute offsets but also in precision. Here, we show the 1D posteriors in age for M74 YSC1. Age is poorly constrained for YGGDRASIL and BPASS but well constrained for GALAXEVneb and FSPS, even though the input data are identical.

bins of uniform size in χ^2, mass, age, and $E(B-V)$. Note that this binning is performed for visualization purposes only; it has no impact on the K-S statistic or median properties across the sample of clusters. We reject the null hypothesis wherever $p < 0.05$ (i.e., at the 2σ level).

It is clear from Figure 7(a) that some models perform better in χ^2 terms than others, with YGGDRASIL and GALAXEVneb in particular under-performing. BPASS is associated with more fits at low χ^2 than the other models, despite being associated with the greatest number of rejected fits overall. The fact that the K-S tests suggest the χ^2 samples are drawn from different parent distributions in three out of six cases reinforces this conclusion—on average, these models are not equally successful at reproducing the measured photometry.

In many cases, the samples of fitted parameters inferred using different SPS models were drawn from different parent distributions. As demonstrated in Figure 7(c), mass is the most robust choice against the SPS model. No model–model pairs have mass samples drawn from different parent distributions. This is true despite the fact that, at young ages, FSPS tends to find lower stellar masses than the other models by a factor of $\sim$0.2 dex (see Section 6.1 and Figure 13). It's also evident that the extinction distributions are quite different (Figure 7(d)).

By comparison, age and extinction are less robust against model choice. The fitted parameter samples differ significantly in four out of six possible model–model comparisons for age and for six out of six in extinction. Although the age samples inferred by BPASS and GALAXEV were drawn from different parent distributions, the medians of these age samples are very similar (6.74 and 6.72, respectively). We note that different models have different accumulation points in age; for example, GALAXEVneb finds only a few systems younger than logAge $\simeq$ 6.4, while FSPS finds no systems younger than logAge $\simeq$ 6.55. This is our first hint that SPS model choice can introduce meaningful scatter into the results of SED fitting.

We also compare the median fitted parameters across all clusters (i.e., the medians of the distributions shown in Figure 7), which are shown in Table 4. We find that model choice can introduce offsets in median age up to 1.6 Myr, in mass up to a factor of $\sim$1.3, and extinction of up to 0.13 mag. Though these offsets are modest, their existence when combined with evidence of model–model inconsistency from the K-S test demonstrates that the choice of SPS model is not necessarily trivial. Using different off-the-shelf SPS models to fit the same objects may produce statistically distinct results. However, as we will discuss in Section 6.1, the model–model differences are more important for individual objects than they are for the sample as a whole.

Table 4
Kolmogorov–Smirnov p-Values: Model–Model

Param.	Y-B	Y-G	Y-F	B-G	B-F	G-F
χ^2	**4.61e-4**	0.149	0.053	0.39	**4.52e-3**	**0.035**
Age	0.1309	0.079	**0.006**	**6.06e-4**	**0.015**	**0.027**
Mass	0.213	0.507	0.360	0.123	0.113	0.294
$E(B-V)$	**0.022**	**0.005**	**6.99e-4**	**4.65e-6**	**0.019**	**9.93e-5**

Note. Kolmogorov–Smirnov (K-S) test p-values comparing on a model–model basis the overall distributions of χ^2, age, mass, and color excess while marginalizing over R_V. Situations where the null hypothesis is rejected are bolded. In the column headers, "Y" refers to YGGDRASIL, "B" refers to BPASS, "G" refers to GALAXEVneb, and "F" refers to FSPS.

5.3. R_V–R_V Comparison across All Clusters

We compare the results inferred using extinction curves with different values of R_V via the same method we used to compare SPS models. We find no evidence that any R_V–R_V pair produced results drawn from different parent distributions in terms of fit quality, age, or cluster stellar mass. From the perspective of the K-S test, the only parameter that appears sensitive to extinction curve choice is color excess. We find that in two out of six cases (2.3 versus 4.1; 2.3 versus 5.1) the samples of maximum-likelihood color excesses were likely drawn from different parent distributions. As discussed below, these differences can be explained by differences in the median fitted values, which themselves differ by construction due to the selection of different values of R_V. The R_V–R_V K-S p-values are shown in Table 5. We show the CDFs and histograms for χ^2, age, mass, and $E(B-V)$ as a function of extinction curve in Figure 8.

As with the model–model comparisons, we compare the medians of the total age, mass, and color excess distributions on an R_V–R_V basis in Table 6. We find that in terms of these medians, R_V–R_V differences are small as a function of cluster age, around 0.9 Myr. Curve choice can introduce offsets up to a factor of $\sim$1.7 in mass and 0.06 mag in extinction. Inferred mass and extinction consistently rise as one pushes toward higher values of R_V, while age tends to decrease. This is not surprising—after all, R_V is just the amount of attenuation in V per magnitude of color change in $B-V$. If R_V increases, one infers a higher A_V at fixed color change, which translates to higher masses and extinctions.

5.4. One-dimensional Posterior Distributions

Although examining the overall distribution of fitted parameters can provide insight into whether model–model and R_V–R_V differences exist, it is difficult to understand the nature of these differences without examining these posteriors on a cluster-by-cluster basis. We plot the 1D posteriors in age, mass, and $E(B-V)$ for each cluster as a function of SPS model and R_V. A "good" example (NGC 4485-YSC2), where there appears to be decent agreement in terms of cluster properties regardless of extinction curve or SPS model used, is shown in Figure 9. A "bad" example (M51-YSC1), where there appears to be a significant systematic offset in these properties as a function of R_V and model choice, is shown in Figure 10. In these plots, rows denote model choice, and curve colors denote R_V. The black dashed line is the M. Sirressi et al.

(2022) "light-weighted" solution, the dotted blue line is the major population in the "two-population" solution, and the dotted red line is the minor population in that solution. M. Sirressi et al. (2022) did not provide a "light-weighted" mass solution. In these plots, fits rejected due to an unacceptable reduced χ^2 are shown using a dotted line rather than a solid one.

A few interesting trends are visible in these plots. First, it is clear that the posterior distributions are not always well described by a Gaussian and are often multimodal. This is reflective of the complex probability surfaces associated with SED fitting. We note that in some cases, as demonstrated in M51-1, the choice of extinction curve can change which peak of a multimodal distribution is associated with the maximum-likelihood solution.

Second, different models can provide constraints on cluster properties of different precision. Take, for example, the age solutions for M74 YSC1 (Figure 11). BPASS and YGGDRASIL constrain the cluster age weakly, with typical one-sided uncertainties across R_V in Log[Age] $\sim$0.1 to $\sim$0.4. GALAXEVneb and FSPS, however, provide much stronger constraints on the order of $\sim$0.01 to $\sim$0.02. This should serve as a stark reminder that the uncertainties associated with SED fitting include both propagated observational uncertainty as well as a contribution from the underlying models themselves, representing regions of the parameter space between which a given model cannot confidently distinguish.

It is also clear from these plots that on a model and R_V basis, there is sometimes significant disagreement between fits. On a model-to-model basis, posterior shapes can vary wildly. R_V choice can also strongly impact the posterior shape. On average, however, using an extinction curve with a higher R_V tends to produce solutions that are associated with an older age, more stellar mass, and more total extinction. We discuss this result further in Section 6.2.

5.5. How Well Do Broadband SED Fits Agree with FUV Spectral Fits?

The 1D posteriors are also a good benchmark for the overall performance of broadband SED fitting compared to FUV spectral fits. We consider three different ways to define agreement between our broadband fits and the M. Sirressi et al. (2022) FUV fits, which we denote as Case A, Case B, and Case C:

(A) The median of the broadband posterior is within 1σ of the FUV light-weighted solution. This comparison is ill-defined for mass, as the FUV fits provide no light-weighted stellar mass.
(B) The median of the broadband posterior is within 1σ of the major population of the FUV 2-population solution.
(C) The broadband maximum-likelihood solution lies somewhere between the major and minor population solutions of the FUV fit. This comparison is ill-defined in situ ations where only a single SSP was required to properly fit the FUV spectrum.

We calculate the number of fits meeting each criteria, both overall and on a model-by-model basis. As expected, we find that there is relatively poor agreement between cluster properties as estimated by broadband SED fitting and by FUV photospheric line fits. Assuming Case A, we find that 79/243 (32.5%) of all fits are in agreement with the FUV in terms of age, and 73/212 (30.0%) are in agreement in terms of extinction. Looking for Case B agreement produces a similar picture; we find agreement in 78/243 (32.1%) of fits for age, 50/243 (20.59%) for mass, and 45/243 (18.5%) of fits for extinction. Case C is the least conservative metric we tested; from this perspective, 107/167 (64.1%) of fits (ignoring clusters without a valid two-population FUV solution) are in age agreement alongside 107/132 (64.1%) in mass and 83/167 (49.7%) in extinction. On average, our photometric analysis agreed with the spectroscopic analysis less often than in M. Sirressi et al. (2022), who found agreement for $\sim$50% of clusters. The difference is likely due to a combination of factors—our large photometric apertures, more diverse set of models, and our more conservative definition of FUV/broadband consistency may all contribute here.

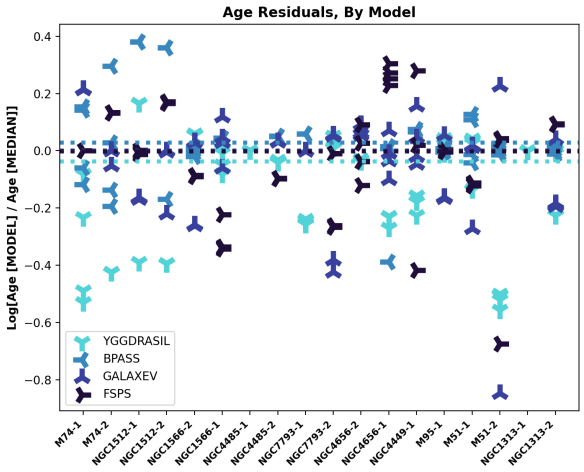
We also note that there is no strong age or mass bias in terms of FUV/broadband agreement. Our broadband fits returned ages older than the FUV fits in 126/243 (51.9%) of cases and masses larger than the dominant FUV population mass in 134/243 (55.1%) of cases. Broadband fits tended to find less total extinction than the FUV fits, with the broadband value exceeding the FUV value in 54/243 (22.2%) cases. The results on a model-by-model basis are summarized below.

5.5.1. Agreement with FUV Fits: YGGDRASIL

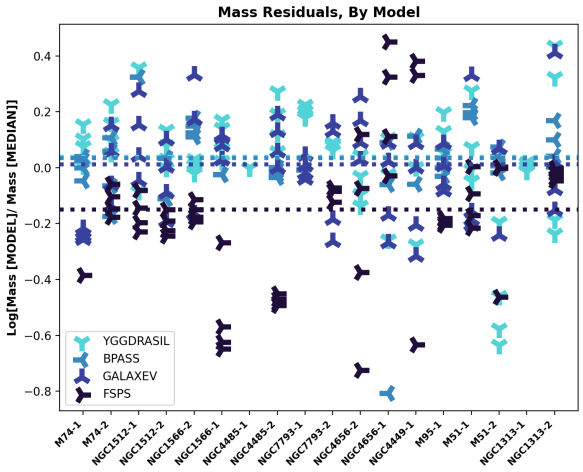
We find that YGGDRASIL has average performance among models we tested in terms of agreement with the FUV fits. Although both YGGDRASIL and the M. Sirressi et al. (2022) fits are based on the Starburst99 (C. Leitherer et al. 1999; G. A. Leitherer 2005b) models, the fact that one set of fits is based on broadband NUV-to-I photometry and the other is based on FUV spectroscopy means that agreement is not guaranteed. We also note that YGGDRASIL used Starburst99 with Padova stellar libraries, while the FUV fits used Starburst99 with Geneva stellar libraries. Particularly at young ages, the differences between these evolutionary tracks are significant. Out of 70 accepted fits performed using YGGDRASIL, we find Case A agreement for 20/70 (28.6%) in age and 29/70 (41.4%) in extinction. The Case B picture is similar, with age agreement in 23/70 (32.9%) of cases, mass agreement in just 11/70 (15.7%), and extinction agreement in 13/70 (18.6%). As expected, Case C presents the most optimistic picture. Fifty YGGDRASIL fits remain after filtering out clusters with only single-population FUV fits. Out of these, 30/50 (60.0%) fulfill Case C in age, 33/50 (66.0%) fulfill it in mass, and 27/50 (54.0%) fulfill it in extinction. Comparing these results to the average performance across all models, YGGDRASIL tends to underperform in terms of its ability to predict cluster ages but is decent at predicting masses and extinctions.

5.5.2. Agreement with FUV Fits: BPASS

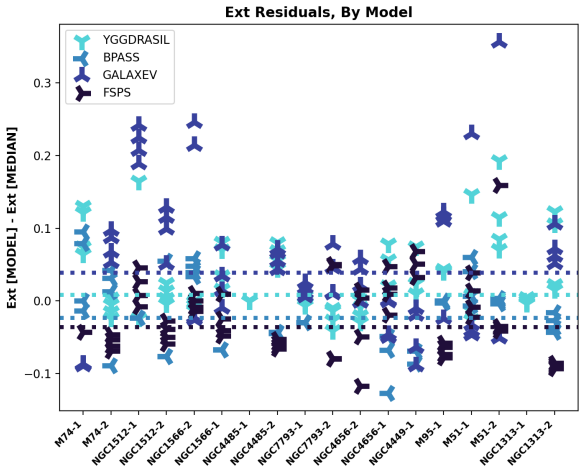
We find that results derived with BPASS are in reasonable agreement with the FUV fits, and it is perhaps the most accurate model we tested by this (limited) metric. Out of 53 accepted fits performed using BPASS, we find Case A agreement for 23/53 (43.4%) in age and 11/53 (20.8%) in extinction. As with YGGDRASIL, Case B paints a similar picture to Case A, with age agreement in 18/53 (34.0%), mass agreement in 23/53 (43.4%), and extinction agreement in



(a) Cluster ages relative to the median, for all clusters



(b) Cluster masses relative to the median, for all clusters



(c) Cluster E(B-V)s relative to the median, for all clusters

Figure 12. The difference between cluster parameters inferred using different SPS codes can be large for individual clusters. However, the systematic offset between models is small. Panel (a) shows cluster ages for individual fits relative to the median age determined for each cluster, marginalized over R_V. Similarly, panel (b) shows cluster masses relative to the median mass determined for each cluster. Panel (c) is the same, this time in terms of extinction. Only accepted fits are shown. Horizontal dotted lines represent the median ratio between fits performed using a given model and the median result for a given cluster.

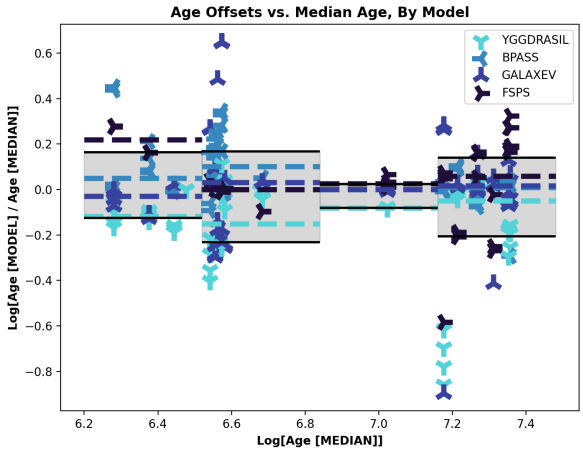
11/53 (20.8%) of fits. After filtering out FUV single-population clusters to enable Case C comparison, we are left with 33 accepted BPASS fits. Out of these, 24/33 (72.7%) fulfill Case C in age, 21/33 (63.6%) fulfill it in mass, and 15/33 (45.5%) do in extinction. Comparing these results to those obtained across all models, BPASS performs better than average in age and mass terms but sometimes struggles to recover the extinction.

5.5.3. Agreement with FUV Fits: GALAXEVneb

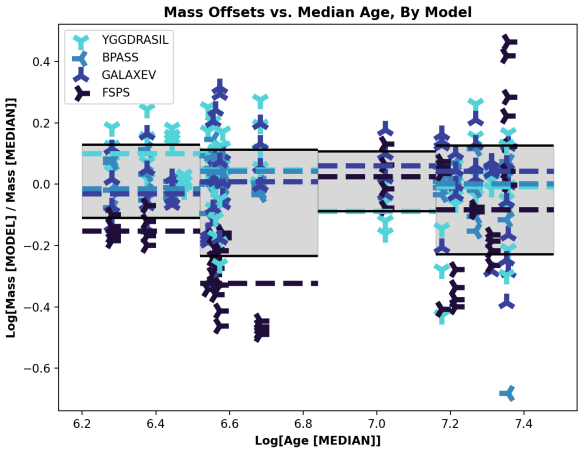
We find that our modified version of GALAXEVneb has an overall accuracy generally comparable to that of YGGDRASIL. We find Case A agreement here in 19/64 (29.7%) of fits for age and 22/64 (34.4%) for extinction. GALAXEVneb is in Case B agreement with the FUV in 18/64 (28.1%) fits in age terms, 10/64 (15.6%) in mass terms, and 14/64 (21.9%) in extinction terms. We once again filter out FUV single-population clusters, leaving us with 44 usable fits for Case C comparison. Here, 25/44 (56.8%) agree with the FUV in age, alongside 30/44 (68.2%) in mass and 21/44 (47.7%) in extinction. Comparing the overall acceptance fractions, we find that GALAXEVneb is better than average at recovering extinctions but suffers in terms of accurately recovering ages.

5.5.4. Agreement with FUV Fits: FSPS

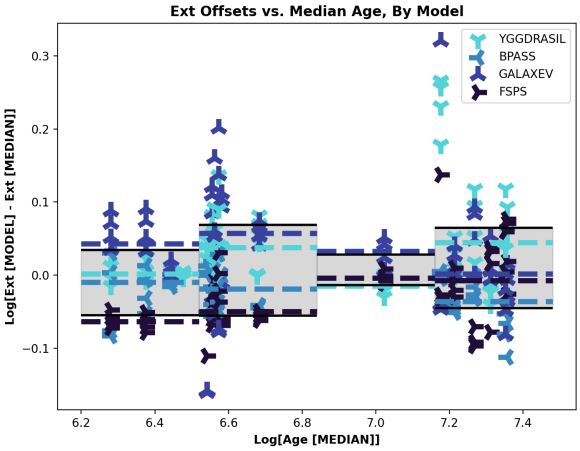
We find that FSPS performed reasonably well in terms of agreement with the FUV fits. The accuracy picture drawn by Case A is average, with agreement in 17/56 (30.4%) of accepted FSPS fits in age terms, alongside 11/56 (19.6%) in extinction. The Case B results are unimpressive; 19/56 (33.9%) of fits agree in age, alongside a mere 6/56 (10.7%) in mass and 7/56 (12.5 %) in extinction. FSPS is associated with 40 fits appropriate for Case C comparison, of which 28/40(70.0%) agree in age, 23/40 (57.49%) agree in mass, and 20/40 (50.0%) agree in extinction. Comparing these results to the fits across all models, FSPS performed particularly well in terms of recovering ages, inferior in this respect only to BPASS.



(a) Age Residuals vs. Age



(b) Mass Residuals vs. Age



(c) Extinction Residuals vs. Age

Figure 13. With the possible exception of FSPS, each model we tested appears to perform similarly as a function of cluster age. Here, we show residuals in age, mass, and extinction as a function of cluster median age for each model. The running medians are shown as dashed lines, while the regions spanned by the 16th and 84th percentiles are shown as the shaded regions. FSPS seems to systematically find lower stellar masses than the other models, particularly at younger ages and high metallicity.

Table 5
Population Median Fitted Parameters: By Model

Param.	YGGDRASIL	GALAXEVneb	BPASS	FSPS
Age	6.72	6.72	6.74	6.59
Mass	4.91	5.03	5.04	5.04
$E(B-V)$	0.12	0.17	0.08	0.04

Note. We find that median fitted parameters within our cluster sample differ somewhat as a function of SPS model. The median age across all clusters varies by up to a factor of $\sim$1.4 on a model-to-model basis, the mass varies by up to a factor of $\sim$3.2, and $E(B-V)$ can vary by up to 0.15 mag. Age and mass are reported in log units.

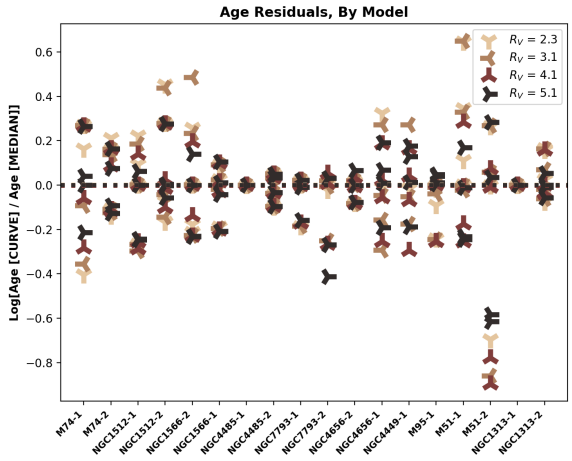
6. Discussion

6.1. Model–Model Comparison: Cluster-by-cluster

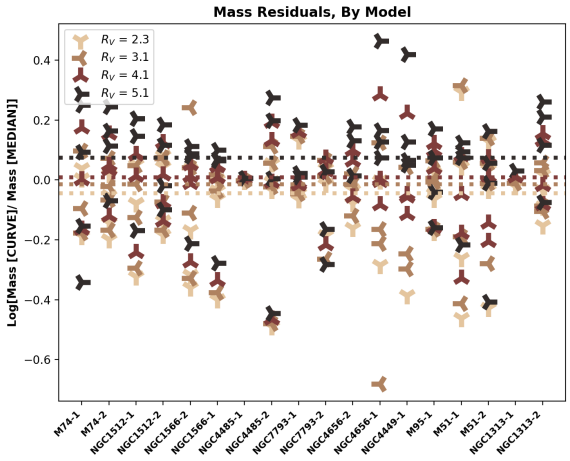
In this work, we demonstrate that there is sometimes a statistically significant difference between the parent distributions of cluster properties derived using different SPS models, which implies that different SPS models may "see" different cluster populations given the same data. We also show that different SPS models perform differently in terms of how well their results agree with fits of the same objects in the FUV. Taken together, these suggest that SPS model choice is not trivial. We caution that making comparisons between the results of SED fitting performed using different input SPS models may not be meaningful, at least for individual objects. With this in mind, a few questions still remain:

1. Are the model–model differences in derived cluster properties random, or are some SPS models consistently biased (e.g., toward older ages or higher extinctions)?
2. Do any models show biases as a function of cluster properties (e.g., do some models have issues that only occur in a certain age range)?
3. If no strong biases exist, how extreme can the scatter introduced by model choice be?

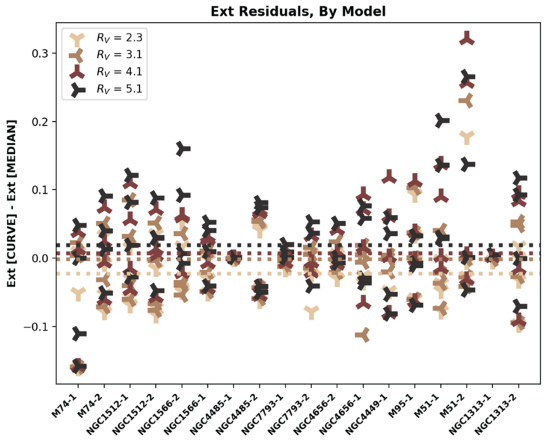
To answer these questions, we plot fitted parameters on a cluster-by-cluster basis for every combination of SPS model and extinction curve (Figure 12), normalized to the median



(a) Cluster ages relative to ages derived with $R_V = 2.3$, for all clusters



(b) Cluster masses relative to masses derived with $R_V = 2.3$, for all clusters



(c) Cluster E(B-V)s relative to extinctions derived with $R_V = 2.3$, for all clusters

Figure 14. Panel (a) shows cluster ages (posterior median) relative to the cluster age determined using an extinction curve with $R_V = 2.3$, marginalized over choice of SPS model. Similarly, panel (b) shows cluster masses (posterior median) relative to the cluster mass determined with $R_V = 2.3$. Panel (c) is the same, this time in terms of extinction.

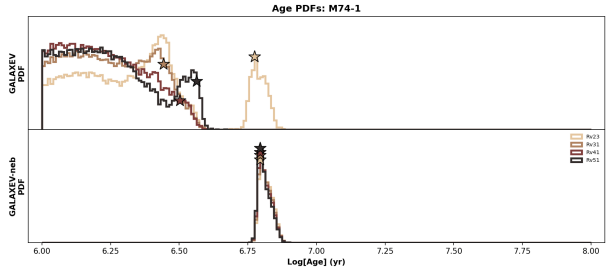
value of that parameter over all models and curves in a given cluster. That is, each point on this plot represents the offset between the results of an individual fit and the median result for that cluster, averaged over every possible extinction curve and SPS model. We show the median of these offsets (as a function of SPS model) over the entire cluster sample as the horizontal dotted lines. To search for systematics related to cluster age, we also plot age, mass, and extinction offsets as a function of median age (Figure 13).

We find a picture similar to that found in Section 5.2—there are, on average, modest offsets as a function of the SPS model. Importantly, we note that the lines of median offset shown here are not always consistent with the population medians shown in Table 5 in terms of their relative magnitude. For example, we can see in Figure 12(b) that FSPS is associated with slightly lower masses relative to the other models, while in Table 5, YGGDRASIL had the lowest median mass. This is evidence that no strong model-dependent bias exists for these parameters—if such a bias existed, it would be visible both in the population medians as well as the individual offsets. Instead, this suggests any systematic bias introduced by model choice is relatively weak, and the model–model inconsistencies we find are primarily the result of scatter between

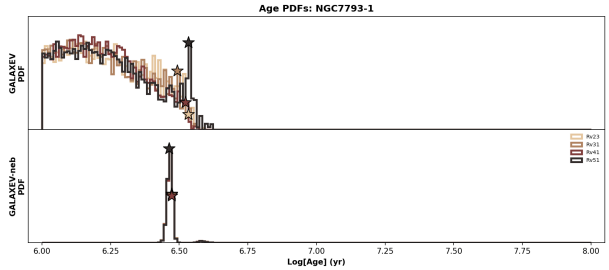
Table 6
Kolmogorov–Smirnov p-Values: R_V–R_V

Param.	2.3–3.1	2.3–4.1	2.3–5.1	5.1–4.1	5.1–3.1	3.1–4.1
χ^2	0.979	0.555	0.426	0.672	0.373	0.814
Age	0.999	0.978	0.688	0.938	0.650	0.988
Mass	0.926	0.398	0.203	0.848	0.376	0.875
$E(B-V)$	0.388	**0.002**	**0.001**	0.941	0.102	0.311

Note. K-S test p-values comparing on an R_V–R_V basis the overall distributions of χ^2, age, mass, and color excess while marginalizing over choice of SPS model. As before, situations where the null hypothesis is rejected are bolded. We find that the distributions of fitted parameters are consistent within 3σ for each possible pair of R_V values tested.



(a) 1d-posterior comparison for M74-1



(b) 1d-posterior comparison for NGC7793-1

Figure 15. One-dimensional posterior comparisons for two example clusters: M74-1 and NGC 7793-1. The top subpanel of each panel shows the age posterior obtained if nebular emission is not included; the bottom subpanel of each panel shows the result if nebular emission is included. In both cases, the lack of nebular emission in the runs using unmodified GALAXEV causes the code to (erroneously?) infer that these systems are younger on average than what we find when including nebular emission. Additionally, various young ages become difficult to distinguish without nebular emission in the models.

Table 7
Population Median Fitted Parameters: By Curve

Param.	$R_V = 2.3$	$R_V = 3.1$	$R_V = 4.1$	$R_V = 5.1$
Age	6.74	6.73	6.73	6.66
Mass	4.87	4.95	5.06	5.11
$E(B-V)$	0.07	0.08	0.13	0.13

Note. Median age, mass, and $E(B-V)$ over all accepted fits as a function of extinction curve R_V. From the perspective of the overall distribution of fitted parameters, curve choice can introduce up to a factor of ∼66% in mass and ∼3 in extinction.

individual fits of the same object. We discuss this apparent mass offset in more detail below.

We find no strong evidence of age-dependent bias for most parameters, though we note that many of the FSPS solutions with low mass relative to other models were found in systems with log[Age] < 7.0 (Figure 13(b)). It is clear that model–model differences can be quite strong for individual clusters. When performing SED fits of the same cluster and using the same extinction curve, SPS model choice can introduce offsets of up to 34.8 Myr in age, a factor of 9.5 in mass, and 0.40 mag in extinction (Figure 12). Averaged over all models, the median offset between any individual fit and the median result for a given cluster was 2.98 Myr in age, a factor of 2.1 in mass, and 0.13 mag in extinction.

6.2. R_V–R_V Comparison: Cluster-by-cluster

Analogously to the previous section, we plot the cluster properties determined using each R_V relative to the per-cluster

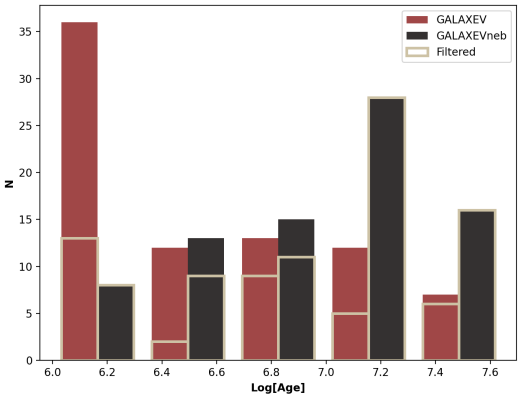


Figure 16. This plot shows age histograms for our entire cluster sample, using our modified version of GALAXEV that includes nebular emission (black) as well as unmodified (i.e., stellar continuum only) GALAXEV (red). Fits that were otherwise rejected due to having reduced $\chi^2 > 10$ are included here as to better demonstrate this phenomenon. The gray open bars show the relative level in each bin if such fits are actually rejected. Note that although raw GALAXEV produced many fits with ages below $\sim 10^7$ yr, very few of these fits were good enough to be accepted.

medians in Figure 14. The effect of R_V on the results of SED fitting is of a similar magnitude to the effect of SPS model choice and is usually more predictable; however, it remains the case that extinction curve choice can introduce large offsets within individual clusters.

R_V does not strongly bias cluster age. As one pushes to higher R_V, however, one tends to recover higher masses and more total extinction. As discussed in Section 5.3, this is a consequence of how R_V is defined. Keeping model choice constant, we find that extinction curve choice can introduce an age offset of up to 32.3 Myr in age, a factor of 10.4 in mass, and 0.41 mag in extinction. This generally corresponds to cases where the underlying posterior distribution was multimodal and the choice of extinction curve changed which mode was associated with the maximum-likelihood solution. Median offsets over all curves were a factor of 0.9 Myr in age, a factor of 1.47 in mass, and 0.05 mag in extinction. Although the curve-curve differences can be extreme in rare cases, the extinction curve choice introduces less scatter on average than the choice of SPS model.

6.3. How Important Is Nebular Emission?

In this work, we have gauged the performance of a modified version of GALAXEV (which we call GALAXEVneb), that has been altered to include contributions to the SED from nebular emission. Given that we necessarily also have the unaltered (i.e., stellar-continuum-only) GALAXEV spectra, it is sensible to ask: how much does the inclusion (or lack thereof) of nebular emission alter the results of broadband SED fitting? Can reasonable fits be obtained for young systems without the inclusion of nebular emission?

We find that nebular emission must be included if one wishes to measure stellar population properties with accuracy. In particular, nebular emission predictions are required to accurately recover stellar population ages. The presence of nebular continuum and line emission causes optical colors (with the exception of $U-B$) to become redder at young ages (C. Leitherer & T. M. Heckman 1995). This means that SED fitting codes will underestimate the age of old systems and

overestimate the dust content of young systems if using templates that do not include this emission. In these cases, the age-extinction degeneracy is important and the fit cannot distinguish between any various solutions at young ages (Figure 15). This can significantly impact the average age of a given cluster population (Figure 16). If nebular emission is not accounted for, it becomes difficult to discern between the youngest systems. In turn, this tends to push average age within a given sample down relative to otherwise identical fits performed with nebular emission included.

This is all to say that if you are fitting a stellar population that might include a component young enough to produce nebular emission, nebular emission should be included in your spectral templates. For example, it has been found that omitting nebular emission can bias galaxy ages high if the intrinsic optical colors are somewhat red (S. de Barros et al. 2014)—the fits must push to old ages if they cannot account for this behavior via nebular lines. Although the precise impact that nebular emission has on SED fitting varies as a function of the systems under examination, it absolutely must be considered.

6.4. Comparison with Previous Results

We are by no means the first to benchmark the performance of SPS models using YSCs. The best and most recent outstanding example of such an analysis is A. Wofford et al. (2016), who tested seven different SPS models against eight clusters in the LEGUS sample with a focus on understanding the impact of massive-star prescriptions. Through a similar analysis to our own, they found that typical maximum differences in properties derived using different models for an individual cluster are 0.09 mag in $E(B-V)$, a factor of 2.8 in mass, and a factor of 2.5 in age. These are generally smaller maximum offsets than we find, perhaps due to contamination in our photometry driven by our large apertures. The key advances provided by our work are the comparison with FUV spectroscopy and the analysis of how the choice of extinction curve affects SED fitting results as a function of R_V. We also use a larger cluster sample than existing studies.

FUV spectra and extinction curves aside, our approach differs from A. Wofford et al. (2016) primarily in how we think about the SPS models themselves. A. Wofford et al. (2016) was interested in performing precise tests on how different ways of handling massive-star evolution effect the results of SED fitting. They, for example, tested how using different input stellar tracks affects fits produced using SB 99. We, on the other hand, are interested in testing SPS models that are publicly available and (with the exception of GALAXEVneb) are usable "off-the-shelf." Our philosophy is that these codes are used commonly throughout the literature, and it is uncommon for teams to test how adopting a different SPS code would alter their results.

6.5. How Should These Systematics Be Managed When Performing and Interpreting SED Fits?

We have demonstrated that the choice of both SPS model and extinction curve can introduce scatter into the results of SED fitting. This scatter is mostly random; model–model and curve-curve differences are more modest when averaged over our whole cluster sample, and there is generally little evidence of ordered bias. However, we note that the choice of extinction curve introduces a subtle systematic such that increasing R_V tends to increase the mass and total extinction one derives.

Given that this scatter can be quite strong—over an order of magnitude in some cases—we suggest that the most conservative approach is to run multiple SED fits for each object of interest using various reasonable combinations of underlying SSP and extinction curve. Though more computationally expensive than a single run, this allows one to identify situations where the idiosyncrasies of model or data produce large scatter. This also allows one to propagate the scatter introduced by model and extinction curve choice throughout the rest of one's analysis, if desired. We also suggest that great care be taken when comparing the properties of different objects if those properties were derived using different SPS models. After all, if the apparent difference between two objects could feasibly be attributed to scatter, it is difficult to convincingly interpret that difference as an actual result.

6.6. How Broadly Applicable Are These Results?

Our analysis here has some limitations that should be kept in mind. First, our sample is limited to young ($<\sim$30 Myr), relatively unobscured, stellar clusters. It is not clear how our results would change if extended to clusters with older ages and higher levels of dust attenuation, where we might expect the age-extinction degeneracy to play a more significant role. It is also unclear how our results would change were an analogous analysis to be performed for entire galaxies rather than individual star clusters. Such an experiment would likely be impacted by factors such as a more complicated star formation history. Still, in broad terms, we are able to show that careful SED fitting should account for the possibility of scatter introduced by SPS model and extinction curve choice.

7. Summary and Conclusions

In this paper, we perform a comprehensive test of SPS models and extinction curves via SED fitting of YSCs in the CLUES sample. Our intention is to identify and quantify any systematic effects introduced into the results of SED fitting through SPS model or extinction curve choice. In particular, we test four commonly used SPS codes in "off-the-shelf" forms that would likely be encountered by a typical end user: YGGDRASIL (E. Zackrisson et al. 2011), BPASS (J. J. Eldridge et al. 2017), FSPS (C. Conroy et al. 2009; C. Conroy & J. E. Gunn 2010), and a modified version of GALAXEV (G. Bruzual & S. Charlot 2003, 2016 version). Additionally, we test the extinction curve parameterized by K. D. Gordon et al. (2023) at four different values of R_V spanning a reasonable range: 2.3, 3.1, 4.1, and 5.1.

We begin by performing aperture photometry of 18 YSCs in the CLUES sample using HST ACS and WFC3 data from the LEGUS survey, being careful to use an aperture precisely matching that of HST's COS instrument. This is because the CLUES clusters are associated with COS FUV spectroscopy and spectral fits, which are useful as a baseline for comparison with our work in the broadband photometry.

We then perform SED fitting via an MCMC approach, which gives us a reliable estimate of the posterior distributions on each physical parameter of interest (age, mass, and extinction). Sixteen fits are performed for each cluster—one per unique combination of SPS model and extinction curve—for a total of 288. We reject any fits with a reduced $\chi^2 > 10$,

leaving us with a final sample of 243 fits suitable for further analysis.

Averaged across all clusters, we find (via the K-S test) that the samples of best-fit age and extinction are often drawn from different parent distributions on a model–model basis. This suggests that different models "see" different cluster populations given the same data. From this perspective, the effect of extinction curve is less important, with only the derived extinctions often drawn from different parent distributions on an R_V–R_V basis.

We find that the agreement between broadband photometric SED fits and FUV-based spectral fits is generally poor. Depending on how exactly one defines "agreement," broadband photometry agrees with the FUV in age terms between 32.1% and 64.1% of the time in age, 20.6%–64.1% of the time in mass, and 18.5%–49.7% in extinction. We find that some different SPS models perform better than others in terms of how well they agree with the FUV. The reason for these offsets is unclear, but they are possibly due to the FUV and optical broadband data better tracing different underlying stellar populations. Differences between the model used to perform the FUV fits and the models tested here may also play a role.

Finally, we find that the scatter introduced by SPS model choice can be large, but model choice introduces minimal ordered bias. The model-derived scatter for an individual cluster can be quite large, up to a factor of 34.8 Myr in age, a factor of 9.5 in mass, and 0.4 mag in extinction. The median scatter introduced by model choice is much smaller but still significant; it is 2.9 Myr in age, a factor of 2.1 in mass, and 0.13 mag in extinction. We find a similar picture for extinction curves, though we note that as one pushes to higher R_V, one tends to systematically recover higher masses and more total extinction. Large scatter for single objects exists here, too: up to 32.3 Myr in age, a factor of 10.4 in mass, and 0.4 mag in extinction. Median scatter between extinction curves was 0.93 Myr in age, a factor of 1.5 in mass, and 0.05 mag in extinction.

For individual sources, SPS model and extinction curve choice can introduce large scatter. We thus make two strong recommendations: First, we suggest that when performing SED fits, one should explore various reasonable combinations of SPS model and extinction curve. This should make any outliers obvious and provide a measure of the systematic uncertainty involved. Second, we caution that because of the potentially significant scatter introduced by model and extinction curve choice, one-to-one comparisons between the properties of individual objects derived using different SED fitting setups may not be meaningful.

We also suggest that more work be done to understand the interplay between nebular emission modeling and the results of SED fitting. Because we wanted to test various SPS+nebular emission models as they exist off-the shelf, a detailed examination of nebular emission modeling was beyond the scope for this paper. However, we have demonstrated that nebular emission is clearly an important component of any realistic SED model. A careful study of which nebular emission parameters are important in SED fitting, and over what range these parameters are physically reasonable, would be of clear value.

Acknowledgments

A.A and M.S. acknowledge support from the Swedish National Space Agency (SNSA) through the grant Dnr158/19. A.W. acknowledges UNAM's DGAPA for support during her sabbatical at UCSD, USA, through program PASPA, and additional support through project PAPIIT-DGAPA IN106922. K.G. is supported by the Australian Research Council through the Discovery Early Career Researcher Award (DECRA) Fellowship (project No. DE220100766) funded by the Australian Government. K.G. is supported by the Australian Research Council Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), through project No. CE170100013. M.J.H. is fellow of the Knut & Alice Wallenberg Foundation. Calculations were performed on the University of Massachusetts' Green High Performance Computing Cluster (MGHPCC) and Unity Cluster. The CLUES data presented in this article were obtained from the Mikulski Archive for Space Telescopes (MAST) at the Space Telescope Science Institute. The specific observations analyzed can be accessed via doi:10.17909/2bay-1d16. We thank our anonymous referee for providing comments and suggestions that greatly improved the quality of this manuscript.

Appendix
Chi-square per Filter

It is possible for SPS models to perform better in some regions of the spectrum than others due to effects such as the presence (or lack) of strong spectral lines that may not be properly accounted for in the nebular emission models. This behavior may elide important information about the performance of the models if only the reduced Chi-Squared of the entire fit is considered. As such, we calculate the Chi-Square for each band in each (accepted) fit. The result is shown in Figure 17; although all fits tend to be somewhat worse in the central bands (roughly F555W through F606W), the models perform similarly well to one another. There are no strong biases as a function of filter, though all models are marginally worse at predicting the emission around ∼5000–7000 Å. The only significant outlier is YGGDRASIL, which seems to perform poorly in F606W. However, given that we only have two clusters (out of 18) with an F606W measurement, this may be the result of low number statistics and so is not a cause for concern.

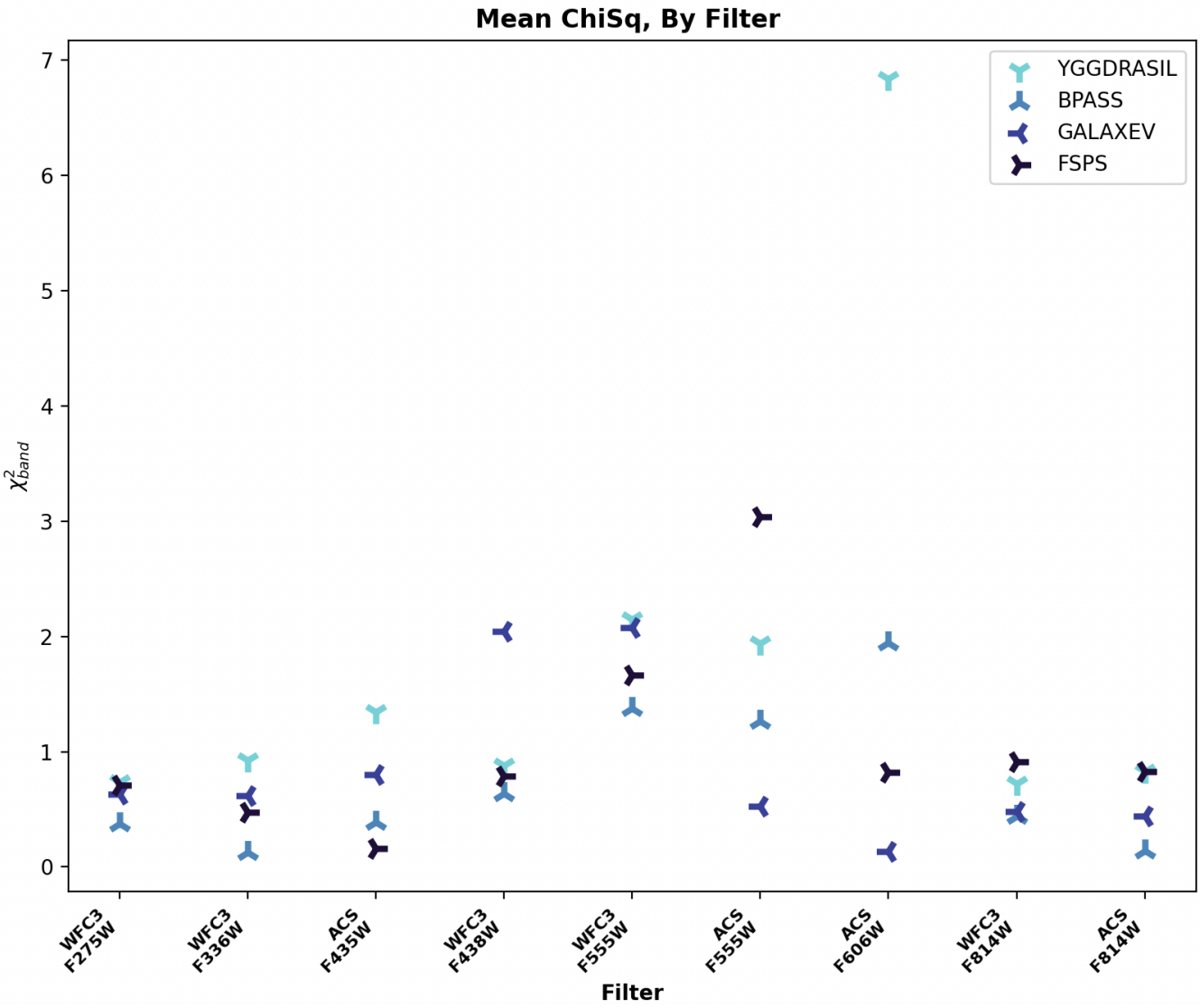


Figure 17. The mean chi-squared in each filter for each model, marginalized over all clusters and extinction curves. Only accepted fits are included here. The models appear to have comparable performance, with the performance of YGGDRASIL in F606W being the only potential outlier.

ORCID iDs

Andrew Mizener https://orcid.org/0000-0003-2722-6600
Daniela Calzetti https://orcid.org/0000-0002-5189-8004
Angela Adamo https://orcid.org/0000-0002-8192-8091
Aida Wofford https://orcid.org/0000-0001-8289-3428
Matthew J. Hayes https://orcid.org/0000-0001-8587-218X
John Chisholm https://orcid.org/0000-0002-0302-2577
Michele Fumagalli https://orcid.org/0000-0001-6676-3842
Svea Hernandez https://orcid.org/0000-0003-4857-8699
Matteo Maria Messa https://orcid.org/0000-0003-1427-2456
Linda J. Smith https://orcid.org/0000-0002-0806-168X
Arjan Bik https://orcid.org/0000-0001-8068-0891
Kathryn Grasha https://orcid.org/0000-0002-3247-5321

References

Adams, N. J., Conselice, C. J., Austin, D., et al. 2024, ApJ, 965, 169
Alongi, M., Bertelli, G., Bressan, A., et al. 1993, A&AS, 97, 851
Anders, P., Bissantz, N., Fritze-v. Alvensleben, U., & de Grijs, R. 2004, MNRAS, 347, 196
Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33
Berg, D. A., Skillman, E. D., Chisholm, J., et al. 2024, ApJ, 971, 87
Boquien, M., Burgarella, D., Roehlly, Y., et al. 2019, A&A, 622, A103
Bradley, L., Sipocz, B., Robitaille, T., et al., 2016 Photutils: Photometry Tools, Astrophysics Source Code Library, ascl:1609.011
Brammer, G. B., van Dokkum, P. G., & Coppi, P. 2008, ApJ, 686, 1503
Bressan, A., Fagotto, F., Bertelli, G., & Chiosi, C. 1993, A&AS, 100, 647
Bridžius, A., Narbutis, D., StonkutÄ, R., Deveikis, V., & Vansevičius, V. 2008, BaltA, 17, 337
Bruzual, A., & Charlot, G. 1993, ApJ, 405, 538
Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000
Byler, N., Dalcanton, J. J., Conroy, C., & Johnson, B. D. 2017, ApJ, 840, 44
Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, ApJ, 533, 682
Calzetti, D., Johnson, K. E., Adamo, A., et al. 2015b, ApJ, 811, 75
Calzetti, D., Lee, J. C., Sabbi, E., et al. 2015a, AJ, 149, 51
Carnall, A. C., McLure, R. J., Dunlop, J. S., & Davé, R. 2018, MNRAS, 480, 4379
Chen, X. Y., Liang, Y. C., Hammer, F., et al. 2010, A&A, 515, A101
Chisholm, J., Rigby, J. R., Bayliss, M., et al. 2019, ApJ, 882, 182
Choi, J., Dotter, A., Conroy, C., et al. 2016, ApJ, 823, 102
Conroy, C. 2013, ARA&A, 51, 393
Conroy, C., & Gunn, J. E. 2010, ApJ, 712, 833
Conroy, C., Gunn, J. E., & White, M. 2009, ApJ, 699, 486
Czekala, I., Loomis, R. A., Teague, R., et al. 2021, ApJS, 257, 2

de Barros, S., Schaerer, D., & Stark, D. P. 2014, A&A, 563, A81
da Cunha, E., Charlot, S., & Elbaz, D. 2008, MNRAS, 388, 1595
de Jager, C., Nieuwenhuijzen, H., & van der Hucht, K. A. 1988, A&AS, 72, 259
da Silva, R. L., Fumagalli, M., & Krumholz, M. 2012, ApJ, 745, 145
Eggleton, P. P. 1971, MNRAS, 151, 351
Eldridge, J. J., Stanway, E. R., Xiao, L., et al. 2017, PASA, 34, e058
Endsley, R., Stark, D. P., Whitler, L., et al. 2023, MNRAS, 524, 2312
Fagotto, F., Bressan, A., Bertelli, G., & Chiosi, C. 1994a, A&AS, 104, 365
Fagotto, F., Bressan, A., Bertelli, G., & Chiosi, C. 1994b, A&AS, 105, 29
Ferland, G. J., Chatzikos, M., Guzmán, F., et al. 2017, RMxAA, 53, 385
Fitzpatrick, E. L. 1999, PASP, 111, 63
Fitzpatrick, E. L., & Massa, D. 2007, ApJ, 663, 320
Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306
Foreman-Mackey, D., Sick, J., & Johnson, B. 2014, python-fsps: Python bindings to FSPS (v0.1.1), Zenodo, doi:10.5281/zenodo.12157
Fouesneau, M., Lançon, A., Chandar, R., & Whitmore, B. C. 2012, ApJ, 750, 60
Girardi, L., Bressan, A., Bertelli, G., & Chiosi, C. 2000, A&AS, 141, 371
Gordon, K. D., Clayton, G. C., Decleir, M., et al. 2023, ApJ, 950, 86
Goudfrooij, P., Burgh, E., Aloisi, A., Hartig, G., & Penton, S. 2010, SMOV: COS NUV Imaging Performance, Tech. Rep. COS 2010-10(v1), STSci
Groves, B. A., Dopita, M. A., & Sutherland, R. S. 2004, ApJS, 153, 9
Gutkin, J., Charlot, S., & Bruzual, G. 2016, MNRAS, 462, 1757
Henyey, L. G., Forbes, J. E., & Gould, N. L. 1964, ApJ, 139, 306
Hurley, J. R., Tout, C. A., & Pols, O. R. 2002, MNRAS, 329, 897
James, B. L., Berg, D. A., King, T., et al. 2022, ApJS, 262, 37
Johnson, B. D., Leja, J., Conroy, C., & Speagle, J. S. 2021, ApJS, 254, 22
Kroupa, P. 2001, MNRAS, 322, 231
Kroupa, P., Tout, C. A., & Gilmore, G. 1993, MNRAS, 262, 545
Krumholz, M. R., Adamo, A., Fumagalli, M., et al. 2015, ApJ, 812, 147
Le Borgne, J. F., Bruzual, G., Pelló, R., et al. 2003, A&A, 402, 433
Leitherer, C., Ekström, S., Meynet, G., et al. 2014, ApJS, 212, 14
Leitherer, C., & Heckman, T. M. 1995, ApJS, 96, 9
Leitherer, C., Schaerer, D., Goldader, J. D., et al. 1999, ApJS, 123, 3
Lejeune, T., Cuisinier, F., & Buser, R. 1997, A&AS, 125, 229
Lejeune, T., Cuisinier, F., & Buser, R. 1998, A&AS, 130, 65
Linden, S. T., Evans, A. S., Armus, L., et al. 2023, ApJL, 944, L55
Maraston, C. 2005, MNRAS, 362, 799
Marigo, P., Girardi, L., Bressan, A., et al. 2008, A&A, 482, 883
Massa, D., Fitzpatrick, E. L., & Gordon, K. D. 2020, ApJ, 891, 67
Meynet, G., Maeder, A., Schaller, G., Schaerer, D., & Charbonnel, C. 1994, A&AS, 103, 97
Muzzin, A., Marchesini, D., van Dokkum, P. G., et al. 2009, ApJ, 701, 1839
Nelson, B., Ford, E. B., & Payne, M. J. 2014, ApJS, 210, 11
Oey, M. S., Meurer, G. R,, Yelda, S., et al. 2007, ApJ, 661, 801
Orozco-Duarte, R., Wofford, A., Vidal-García, A., et al. 2022, MNRAS, 509, 522
Osterbrock, D. E. 1974, Astrophysics of Gaseous Nebulae (San Franciso, CA: Freeman)
Pacifici, C., Iyer, K. G., Mobasher, B., et al. 2023, ApJ, 944, 141
Pérez-González, P. G., Barro, G., Annunziatella, M., et al. 2023, ApJL, 946, L16
Portegies Zwart, S. F., McMillan, S. L. W., & Gieles, M. 2010, ARA&A, 48, 431
Renzini, A., & Fusi Pecci, F. 1988, ARA&A, 26, 199
Sana, H., de Mink, S. E., de Koter, A., et al. 2012, Sci, 337, 444
Schlafly, E. F., & Finkbeiner, D. P. 2011, ApJ, 737, 103
Sirressi, M., Adamo, A., Hayes, M., et al. 2022, AJ, 164, 208
Sirressi, M., Adamo, A., Hayes, M., et al. 2024, AJ, 167, 166
Smith, L. J., Norris, R. P. F., & Crowther, P. A. 2002, MNRAS, 337, 1309
Tacchella, S., Finkelstein, S. L., Bagley, M., et al. 2022, ApJ, 927, 170
Tang, J., Grasha, K., & Krumholz, M. R. 2024, MNRAS, 532, 4583
ter Braak, C. J. F., & Vrugt, J. A. 2008, Statistics and Computing, 18, 435
Tinsley, B. M. 1968, ApJ, 151, 547
Tinsley, B. M. 1973, A&A, 24, 89
Tinsley, B. M., & Gunn, J. E. 1976, ApJ, 203, 52
Vazdekis, A., Sánchez-Blázquez, P., Falcón-Barroso, J., et al. 2010, MNRAS, 404, 1639
Vázquez, G. A., & Leitherer, C. 2005a, ApJ, 621, 695
Leitherer, G. A. 2005b, ApJ, 621, 695
Vink, J. S., de Koter, A., & Lamers, H. J. G. L. M. 2001, A&A, 369, 574
Walcher, J., Groves, B., Budavári, T., & Dale, D. 2011, Ap&SS, 331, 1
Wang, B., Leja, J., Atek, H., et al. 2024, ApJ, 963, 74
Westera, P., Lejeune, T., Buser, R., Cuisinier, F., & Bruzual, G. 2002, A&A, 381, 524
Whitler, L., Stark, D. P., Endsley, R., et al. 2023, MNRAS, 519, 5859
Wofford, A., Charlot, S., Bruzual, G., et al. 2016, MNRAS, 457, 4296
Wofford, A., Leitherer, C., & Chandar, R. 2011, ApJ, 727, 100
Wofford, A., Leitherer, C., & Salzer, J. 2013, ApJ, 765, 118
Xiao, L., Stanway, E. R., & Eldridge, J. J. 2018, MNRAS, 477, 904
Zackrisson, E., Bergvall, N., Olofsson, K., & Siebert, A. 2001, A&A, 375, 814
Zackrisson, E., Rydberg, C.-E., Schaerer, D., Östlin, G., & Tuli, M. 2011, ApJ, 740, 13