

Star formation at the edge of the Local Group: a rising star formation history in the isolated galaxy WLM

Saundra M. Albers¹, Daniel R. Weisz¹,¹★ Andrew A. Cole,² Andrew E. Dolphin,³ Evan D. Skillman,⁴ Benjamin F. Williams,⁵ Michael Boylan-Kolchin⁶, James S. Bullock,⁷ Julianne J. Dalcanton,⁵ Philip F. Hopkins⁸, Ryan Leaman,⁹ Alan W. McConnachie,¹⁰ Mark Vogelsberger¹¹ and Andrew Wetzel¹²

¹Department of Astronomy, University of California Berkeley, Berkeley, CA 94720, USA

²School of Natural Sciences, University of Tasmania, Private Bag 37, Hobart, Tasmania 7001, Australia

³Raytheon, 1151 E. Hermans Rd, Tucson, AZ 85756, USA

⁴Minnesota Institute for Astrophysics, University of Minnesota, Minneapolis, MN 55455, USA

⁵Department of Astronomy, Box 351580, University of Washington, Seattle, WA 98195, USA

⁶Department of Astronomy, The University of Texas at Austin, Austin, TX 78712, USA

⁷Center for Cosmology, Department of Physics & Astronomy, University of California, Irvine, 4129 Reines Hall, Irvine, CA 92697, USA

⁸TAPIR, Mailcode 350-17, California Institute of Technology, Pasadena, CA 91125, USA

⁹Max-Planck Institut für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany

¹⁰NRC Herzberg Astronomy and Astrophysics, 5071 West Saanich Road, Victoria, BC V9E 2E7, Canada

¹¹Kavli Institute for Astrophysics and Space Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

¹²Department of Physics, University of California, Davis, CA 95616, USA

Accepted 2019 October 1. Received 2019 September 15; in original form 2019 June 21

ABSTRACT

We present the star formation history (SFH) of the isolated ($D \sim 970$ kpc) Local Group dwarf galaxy Wolf–Lundmark–Melotte (WLM) measured from colour–magnitude diagrams (CMDs) constructed from deep *Hubble Space Telescope* imaging. Our observations include a central ($0.5 r_h$) and outer field ($0.7 r_h$) that reach below the oldest main-sequence turn-off. WLM has no early dominant episode of star formation: 20 per cent of its stellar mass formed by ~ 12.5 Gyr ago ($z \sim 5$). It also has an SFR that rises to the present with 50 per cent of the stellar mass within the most recent 5 Gyr ($z < 0.7$). There is evidence of a strong age gradient: the mean age of the outer field is 5 Gyr older than the inner field despite being only 0.4 kpc apart. Some models suggest such steep gradients are associated with strong stellar feedback and dark-matter core creation. The SFHs of real isolated dwarf galaxies and those from the Feedback in Realistic Environment suite are in good agreement for $M_*(z=0) \sim 10^7\text{--}10^9 M_\odot$, but in worse agreement at lower masses ($M_*(z=0) \sim 10^5\text{--}10^7 M_\odot$). These differences may be explainable by systematics in the models (e.g. reionization model) and/or observations (*HST* field placement). We suggest that a coordinated effort to get deep CMDs between *HST/JWST* (crowded central fields) and WFIRST (wide-area halo coverage) is the optimal path for measuring global SFHs of isolated dwarf galaxies.

Key words: galaxies: dwarf – galaxies: evolution – Local Group – galaxies: stellar content.

1 INTRODUCTION

1.1 Star formation histories of isolated dwarf galaxies

Isolated dwarf galaxies are key to understanding low-mass galaxy evolution. By virtue of their isolation, a balance between internal processes (e.g. feedback, galactic winds, and outflows) and

universal external drivers (e.g. gas accretion and cosmic reionization) alone govern their evolution. This picture stands in stark contrast to satellites, whose evolutionary trajectories and present-day properties have been strongly affected by their proximity to a massive galaxy, making it challenging to disentangle the relative contributions of environment (e.g. ram pressure and tidal effects) from secular processes.

Our empirical knowledge of low-mass galaxy evolution has increased markedly over the last decade. Ambitious investments in stellar spectroscopy (e.g. Leaman et al. 2013; Kirby et al. 2017) and

* E-mail: dan.weisz@berkeley.edu

imaging programmes such as local cosmology with isolated dwarf (LCID) programme (Gallart et al. 2015, and references therein), the ACS Nearby Galaxy Survey Treasury (Dalcanton et al. 2009), and archival efforts (e.g. Holtzman, Afonso & Dolphin 2006) have demonstrated a rich diversity in the star formation histories (SFHs) of isolated dwarf galaxies over cosmic time (e.g. Cole et al. 2007, 2014; Hidalgo et al. 2009; Monelli et al. 2010a,b; Weisz et al. 2011; Skillman et al. 2014; Weisz et al. 2014). These findings reinforce the long-standing theoretical concepts that the evolution of low-mass galaxies are sensitive to a variety of processes including inhomogeneous reionization, variable mass assembly histories, and stellar feedback.

Additionally, large gains in computing power and resolution have led to transformative improvements in numerical models of dwarf galaxy evolution, providing the opportunity to test these refined predictions observationally (e.g. El-Badry et al. 2016; Sawala et al. 2016; El-Badry, Weisz & Quataert 2017; Fitts et al. 2017; Robles et al. 2017; González-Samaniego et al. 2017; Fillingham et al. 2018; Fitts et al. 2018; Su et al. 2018; Garrison-Kimmel et al. 2019; and references therein).

The large diversity in the evolutionary histories of isolated dwarfs necessitates a larger sample of galaxies with deep colour-magnitude diagrams (CMDs) to establish ‘typical’ behaviour and its degree of variation. However, the desire to obtain more data is counterbalanced by the observational resources required to reach the faint oldest main-sequence turn-off (MSTO) – the ‘gold-standard’ CMD depth for measuring a well-constrained SFH over cosmic time (e.g. Gallart, Zoccali & Aparicio 2005; Dolphin 2012). Due to the high observational demands, only seven isolated dwarf galaxies to date have sufficiently deep observations for well-constrained SFHs, and the sample size has only been incrementally increasing.

Perhaps the most intriguing result from the existing studies is the possibility that isolated dwarf galaxies experience a significant delay in producing their stars relative to satellite galaxies. This was first hinted at from the *HST* observations of Leo A (Tolstoy et al. 1998; Cole et al. 2007), which indicated that over 90 per cent of its star formation occurred in the last 8 Gyr. This notion was supported by the Cole et al. (2014) study of the Aquarius dwarf irregular galaxy (DDO 210), where *HST* observations have shown that only 10 per cent of the stars formed earlier than 10 Gyr ago, with a strong increase in the star formation rate between 6 and 8 Gyr ago. Further tests of this hypothesis are clearly needed.

1.2 WLM

In this paper, we present the SFH of the isolated, gas-rich, dwarf galaxy Wolf–Lundmark–Melotte (WLM; also known as DDO 221) based on deep *HST* imaging. Table 1 lists the basic properties of WLM.

At ~ 970 kpc, WLM is one of the most distant and isolated dwarf galaxies with such deep *HST* imaging. *HST*’s spatial extent is sufficient to sample the inner and outer regions of WLM, providing a handle on both the global and radial SFH over cosmic time. Early resolved-star observations of WLM (Ferraro et al. 1989) indicated a relatively constant star formation rate over its lifetime. Minniti & Zijlstra (1996) and Minniti & Zijlstra (1997) demonstrated that the oldest stars were more widely distributed than intermediate age and young stars, indicating a clear radial gradient in the distributions of its stellar populations. SFHs measured from shallow WFPC2 data show that the central regions of WLM are predominantly young (Dolphin 2000b; Weisz et al. 2014).

Table 1. Properties of WLM.

Parameter	Value
Galactic coordinates (ℓ , b) (deg)	75.9, -73.6
Distance (kpc)	$968^{+5.41}_{-7.39}$
Distance modulus (mag)	$24.93^{+0.02, 0.07}_{-0.02, 0.07}$
Absolute magnitude (M_V)	-14.2 ± 0.01
Line of sight reddening $E(B - V)$	0.038
Half-light radius (arcmin)	8.62 ± 0.26
Stellar mass (M_\odot)	4.3×10^7
Current SFR [$\log(M_\odot \text{ yr}^{-1})$]	-2.24
Stellar metallicity (dex)	-1.28 ± 0.02

Note. All values are taken from McConnachie (2012) except for mean metallicity, which is taken from Leaman et al. (2013), and the SFR which is from Karachentsev, Makarov & Kaisina (2013). The distance is from this work. Uncertainties on the distance reflect the random and systematic errors.

Existing metallicity measurements within WLM can serve as consistency checks on the age–metallicity relationships that we derive from the stellar photometry. Leaman et al. (2013) presented a detailed study of 180 WLM red giant branch stars and found a mean value of $[\text{Fe}/\text{H}] = -1.28 \pm 0.02$ with a dispersion of 0.38 ± 0.04 dex. This relatively low average metallicity for the older stellar population is typical for a dwarf galaxy of this stellar mass. The young stellar and ISM metallicities are also relatively low and typical for galaxies of comparable stellar mass; Urbaneja et al. (2008) found an average metallicity of $[Z] = -0.87 \pm 0.06$ from 8 young supergiant stars in good agreement with the H II region oxygen abundances measured by Skillman, Terlevich & Melnick (1989), Hodge & Miller (1995), and Lee, Skillman & Venn (2005).

A few words about the degree of isolation of WLMs are appropriate here. WLM is currently slightly more distant from the Milky Way than from M31 ($D_{\text{M31}} = 836$ kpc; McConnachie 2012) and lies near the zero-velocity surface for the Local Group. From a comparison of positions and radial velocities, Teyssier, Johnston & Kuhlen (2012) conclude that there is a less than 1 per cent probability that WLM is associated with the Milky Way. To our knowledge, no similar analysis has been conducted for the WLM–M31 association, but the large current separation indicates a very low probability of past interaction.

WLM’s large distance also places it beyond the so-called ‘back-splash’ radius (e.g. Gill, Knebe & Gibson 2005). Galaxies beyond this distance, nominally twice the virial radius of the central system (i.e. 600 kpc for the MW), are unlikely to have been ejected from an early interaction with the central galaxy (though see D’Onghia et al. 2009 which suggests Cetus and Tucana may exist as the results of early interactions). Around the LG, only six galaxies more distant than the back-splash radius have SFHs measured from deep CMDs (e.g. Cole et al. 2014; Gallart et al. 2015).

This paper is organized as follows. In Section 2, we describe the observations, data reduction, distance determination, and method by which we measured the SFH. In Section 3, we present the SFH of the inner, and outer fields of WLM. In Section 4, we compare WLM’s SFH with other isolated dwarf galaxies and to simulated dwarf galaxies. In the appendices, we include tabulated SFHs for WLM and demonstrate the weak effect of stellar model choice on the resulting SFH of WLM.

This is the first paper in a series aimed at comparing the global and spatial properties of isolated dwarf galaxies to state-of-the-art models of low-mass galaxy formation. Accordingly, the scope of this first paper is to present the CMD and SFH of WLM, consider its global SFH relative to other isolated galaxies and select simulations,

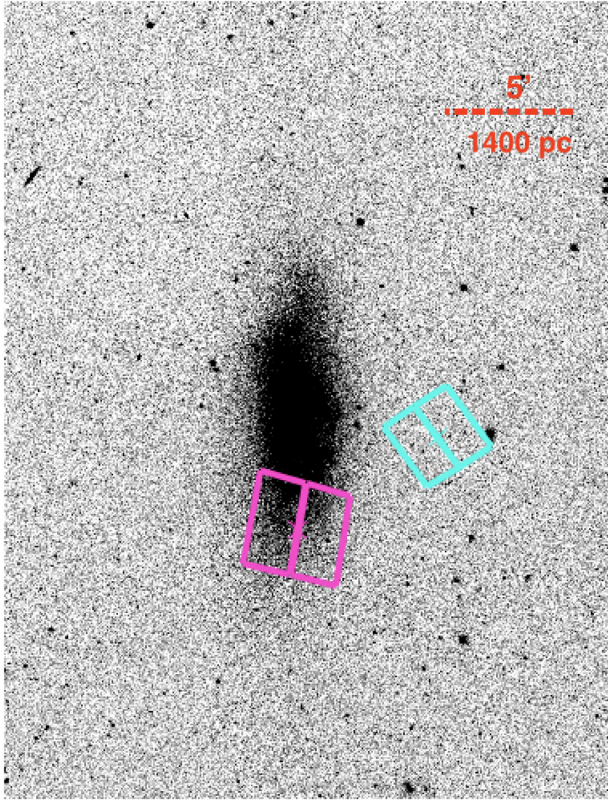


Figure 1. The two *HST* fields overlaid on a ground-based image of WLM. The central ACS field is shown in magenta and the outer UVIS field in cyan. The fields were placed to sample spatial gradients and provide a reasonable representation of WLM’s global SFH.

and briefly comment on radial trends, which will be explored more depth in a future paper in this series.

2 METHODOLOGY

2.1 Observations and photometry

The observations and data reduction for WLM mirror that of several previous *HST* programmes aimed at isolated dwarf galaxies. Here, we briefly summarize the data acquisition and reduction and refer the reader to more detailed descriptions in Monelli et al. (2010a), Cole et al. (2014), and Skillman et al. (2014).

WLM was observed with *HST* between 2015 July 17 and 19 as part of *HST*-GO-13768 (PI: D. Weisz). An inner ACS and outer UVIS field were observed for ~ 27 360s in F475W and ~ 34 050s in F814W. Fig. 1 shows the placement of the *HST* fields overlaid on a ground-based optical image. Two fields were selected to sample the population gradients in WLM. The inner ACS field is located 1 kpc ($0.5 r_h$) from the photometric centre, while the outer UVIS field is 1.4 kpc ($0.7 r_h$) from the centre. At the distance of WLM, the fields are 0.95 kpc (ACS) and 0.76 kpc (UVIS) across in linear size.

We measured photometry of stars in the ACS and UVIS fields using the point spread function fitting package *dolphot* (Dolphin 2000a), with specific ACS and UVIS modules, using the photometric reduction parameters recommended by Williams et al. (2014).

From the raw source catalogues, we culled the photometry to only include high-fidelity stellar sources. We required that stars

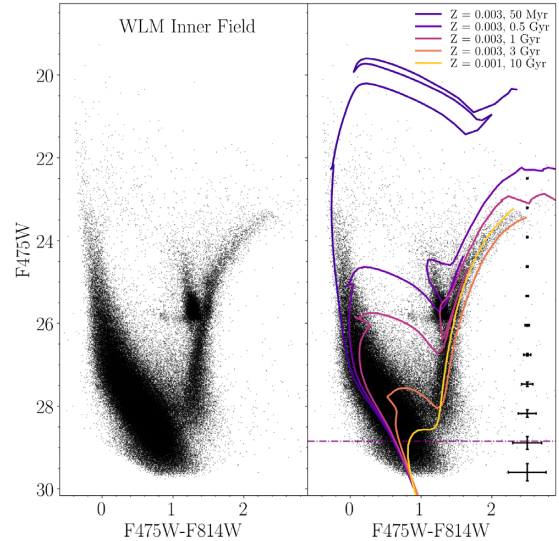


Figure 2. CMD of the inner ACS field of WLM. Virtually all regions of the CMD are well-populated, suggesting quasi-continuous star formation over the lifetime of WLM. In the right-hand panel, we overplot select Padova isochrones that confirm this scenario. The 50 per cent completeness limit (dashed line) in F475W of the ACS field is ~ 0.5 mag below the oldest MSTO, allowing for excellent leverage on the SFH at all look-back times.

have an $\text{SNR} > 5$ in both filters, $(\text{sharp}_{F475W}^2 + \text{sharp}_{F814W}^2) < 0.1$ and $\text{crowd}_{F475W} + \text{crowd}_{F814W} < 1.0$.

To quantify observational uncertainties and completeness, we ran $\gtrsim 10^5$ artificial star tests (ASTs) in each field. The 50 per cent completeness limits for the ACS field are $m_{F475W} = 28.85$, $m_{F814W} = 28.17$ and for the UVIS field are $m_{F475W} = 29.20$, $m_{F814W} = 28.15$.

2.2 Colour–magnitude diagrams

Figs 2 and 3 show the CMDs for the ACS and UVIS fields, respectively. The ACS field has 149 558 stars and all stellar sequences are well-populated and clearly visible. Visual inspection shows subgiant stars for every isochrone age and suggests that the inner ACS field has had continuous star formation at virtually all epochs. The presence of stars at oldest MSTO, as well on the horizontal branch (HB) in each field indicate ancient and intermediate-age star formation. The red HB is clearly visible in each CMD, while blue HB merges into the luminous MS, making it challenging to discern visually. The red HB and red clump indicate intermediate age star formation, and the luminous MS and blue and red core helium burning sequences (BHeBs, RHeBs) are signs of star formation within the most recent 1 Gyr.

We plot select Padova isochrones (Girardi et al. 2010) in the right-hand panel to guide the reader’s eye. For select ages (and metallicities) of 0.05, 0.5, 1, 3, and 10 Gyr, we find stars that populate each area of the isochrone, which is a sign of quasi-continuous star formation at these ages. It is challenging to visually interpret the relative densities of stars for each isochrone, i.e. which would indicate bursts of star formation, and we therefore defer detailed discussion of putative features until we present a quantitative SFH in Section 3.1.

The outer UVIS CMD (Fig. 3) contains 3362 stars and is not as well populated as the ACS field, by virtue of its location along WLM’s minor axis. However, even with fewer stars, many of the same features (e.g. oldest MSTO, HB, and red clump) are clearly

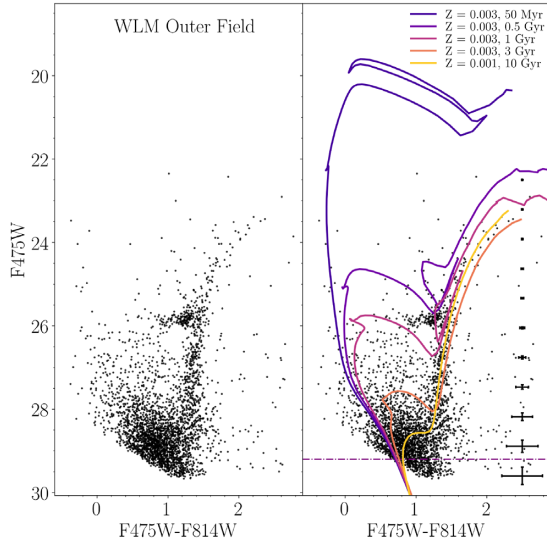


Figure 3. CMD of the outer UVIS field of WLM. Compared to the ACS field, this field is more sparsely populated and lacks stars on the upper main sequence suggestive of little to no recent star formation. As with the ACS field, the F475W 50 per cent completeness limit is ~ 0.5 mag below the oldest MSTO, allowing for excellent leverage on the SFH at all look-back times.

defined. Compared to the ACS CMD, the UVIS CMD has few, if any, stars younger than ~ 500 Myr. We overplot select Padova isochrones in the right-hand panel of Fig. 3. For ages > 3 Gyr, stars appear to populate all of the selected isochrones, suggesting continuous star formation for ages older than ~ 3 Gyr. For ages younger than ~ 3 Gyr, the interpretation is less clear. While stars do populate portions of the $\lesssim 1$ Gyr isochrones, the sparsity of objects suggests very little, or perhaps no, star formation at young ages. The visual translation of such sparse populations to a quantitative SFH is challenging in sparse regimes and we thus defer discussion of recent star formation to our measurement of the quantitative SFH in Section 3.2.

2.3 Distance determination

The well-populated ACS field allows us to measure a tip of the red giant branch (TRGB) distance to WLM. To compute this, we use an implementation of the maximum-likelihood technique first described in Makarov et al. (2006). Briefly, this code constructs a model power-law luminosity function, convolves it with photometric errors and completeness from the ASTs, and iterates over a grid of parameter values to find the apparent magnitude of the TRGB.

We find the magnitude of the TRGB to be $m_{F814W} = 20.91^{+0.02}_{-0.01}$, where the uncertainties reflect the narrowest 68 per cent confidence intervals that include the most likely value.

We then convert this to a distance modulus by first correcting the apparent magnitude for line-of-sight Galactic foreground extinction ($A_V = 0.104$) using the maps from Schlafly & Finkbeiner (2011). We then used the calibration of Rizzi et al. (2007)

$$M_{F814W}^{ACS} = -4.06 + 0.15 [(F555W - F814W) - 1.74] \quad (1)$$

to measure the distance modulus. Because Rizzi et al. (2007) do not provide a calibration for F475W, we use the colour approximation of $(F555W - F814W) = 0.675 (F475W - F814W)$. The mean colour of the TRGB region in both CMDs is $(F475W - F814W) \sim 2.4$.

We find that WLM has a TRGB distance modulus of $\mu = 24.93 \pm 0.02$ (random) ± 0.07 (systematic), making WLM the most distant galaxy for which the oldest MSTO has been observed. This places WLM at a distance of $D = 968^{+5.41}_{-7.39}$ kpc, the first error value representing random error, the second value representing the random and systematic error. This distance agrees with previous distance determinations of 24.93 ± 0.04 from a TRGB measurement (Rizzi et al. 2007), $24.92 \pm 0.04 \pm 0.04$ from Cepheid observations by Gieren et al. (2008), and 24.99 ± 0.10 using the flux-weighted gravity-luminosity relationship for A and B supergiants (Urbaneja et al. 2008). We adopt $\mu = 24.93$ for subsequent analysis in this paper.

2.4 Measuring the SFH

To measure the SFH of WLM, we model the CMDs of the ACS and UVIS fields using MATCH (Dolphin 2002). MATCH has been widely used to measure the SFHs of > 100 galaxies throughout the LG and Local Volume. In this analysis, we generally follow the MATCH usage as described in Weisz et al. (2014). Here, we briefly summarize how MATCH operates and how we applied it to the WLM CMDs.

For a given set of stellar evolution models, stellar initial mass function, binary fraction, distance modulus, extinction model, and SFH (i.e. star formation rate and metallicity history as a function of time that is simply the sum of simple stellar populations), MATCH constructs a synthetic CMD. The synthetic CMD is convolved with the error and completeness functions determined by the ASTs, and CMDs of foreground and background contaminants are linearly added to the galaxy synthetic CMD. This mock observed CMD is then compared to an observed CMD using a Poisson likelihood function. The process is repeated for various SFHs (i.e. different weights on each of the simple stellar populations) until a global, most likely SFH is determined.

For our analysis of WLM, we used the Padova (Girardi et al. 2010) stellar evolution models, a Kroupa (2001) IMF, a binary fraction of 0.35 with the primary-to-secondary mass ratios drawn from a uniform distribution, our TRGB distance of $\mu = 24.93$, and a foreground extinction value of $A_V = 0.104$ from Schlafly & Finkbeiner (2011). We use an age grid of $\log(t) = 6.6$ – 10.15 with a time resolution of 0.1 dex for $\log(t) \leq 9.0$ and 0.05 dex for $\log(t) > 9.0$. Our metallicity grid is $[M/H] = -2.3$ to -0.1 with a resolution of 0.1 dex. We model the entire CMD (i.e. do not exclude any regions such as the RGB).

Though more recent stellar models than Padova are now available (e.g. PARSEC, MIST, BaSTI; Bressan et al. 2012; Choi et al. 2016; Hidalgo et al. 2018), we adopt the Padova models for direct comparison to other isolated galaxies, which all have SFHs measured with these older models. We compare the results of fitting different models to the CMD in Appendix B. Future papers in this series will measure SFHs for all isolated galaxies using updated stellar models.

We also consider the effects of differential extinction on the SFH. Specifically, beyond the foreground dust value, MATCH reddens a uniform fraction of all model stars between 0 and 0.5 mag by steps of 0.1. We computed the SFH with this differential extinction model in both the ACS and UVIS field, but found that both fields were best fit with no differential extinction.

While differential extinction may not affect the majority of the CMD, it can affect only the younger stars (e.g. Dolphin et al. 2003) that may reside in dusty star formation regions. Following Dolphin et al. (2003), we adopt an age-dependent differential extinction

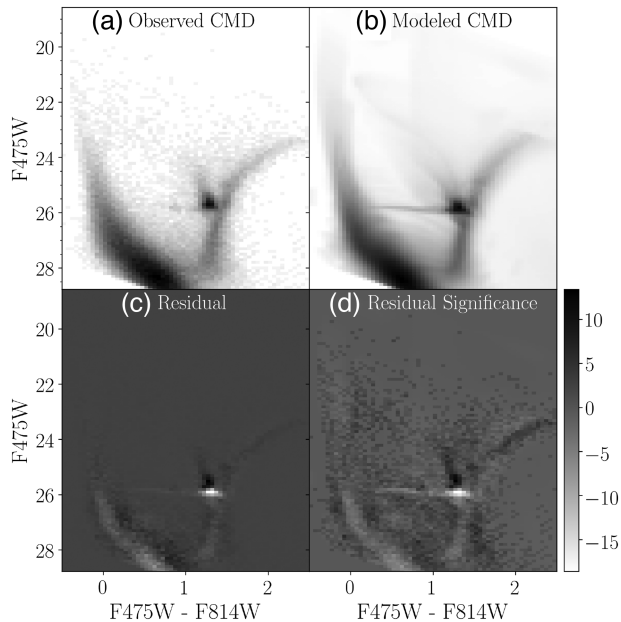


Figure 4. The residual CMDs for our best-fitting SFH of the ACS field. Panel (a) shows the observed Hess diagram; Panel (b) shows the best-fitting model CMD; Panel (c) is the residual (i.e. data-model); and Panel (d) is the residual significance, i.e. the residual weighted by the variance in each pixel. The scale bar is in units of standard deviation. The fit quality is typical of resolved isolated dwarf galaxies. The model is generally good, with known deficiencies in areas of evolved stars (e.g. lower blue loop, RGB, RC, and HB). The bulk of the residual mismatches are due to a spatially varying noise model as discussed in Section 2.4.

model for young stars such that stars younger than 40 Myr randomly get a value of $A_V = 0.5$ applied. The amount of dust applied to the model CMD decreases linearly from $A_V = 0.5$ to $A_V = 0.0$ between 40 and 100 Myr. Not including age-dependent young star dust results in a model upper main sequence that is too narrow compared to the data.

Using the best-fitting SFH as a starting point, we compute random and systematic uncertainties on the SFH. Random uncertainties capture plausible variations from the best-fitting SFH due to finite sampling on the CMD, and are determined using a Hybrid Markov chain Monte Carlo (HMC) algorithm as described in Duane et al. (1987) and implemented by Dolphin (2013). Throughout this paper, the reported random uncertainties represent the narrowest 68 per cent confidence interval around the best-fitting SFH.

Systematic uncertainties are designed to approximate variations in the SFH due to the choice of underlying stellar evolution model. We compute systematic uncertainties using 50 Monte Carlo realizations following the procedures described in Weisz et al. (2011), Dolphin (2012), and Weisz et al. (2014). The reported systematic uncertainties represent the narrowest 68 per cent confidence interval around the best-fitting SFH.

Figs 4 and 5 show the quality of our best-fitting CMDs for the ACS and UVIS fields, respectively. In both figures, the most important diagnostic is the residual significance (panel d), which is the residual (data-model) weighted by the variance in each pixel. A checkerboard pattern of grey indicates no major residuals, while areas of white or black indicate over/under predictions by the model. The residuals for the ACS and UVIS fields are typical for these types of observations. That is, the model for the ACS field, with more stars and a more complex SFH, is a good match with clear deficiencies.

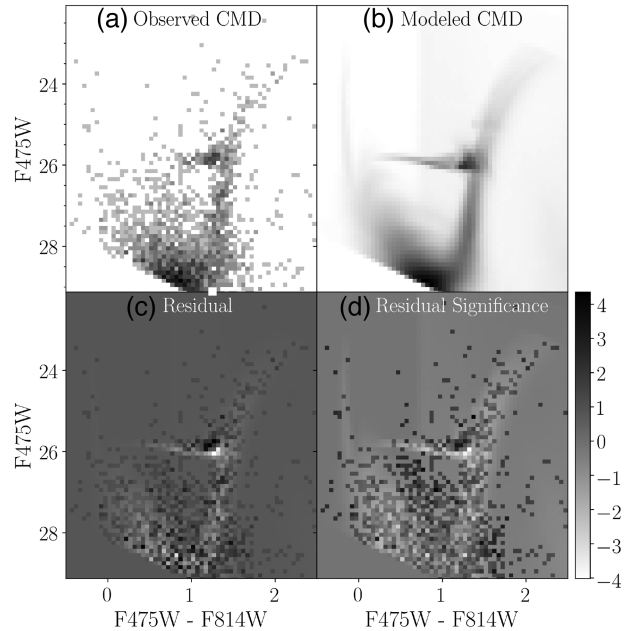


Figure 5. Same as Fig. 4, except for the UVIS field. The residual appears cleaner than the ACS field due to the simpler nature of the stellar population and having a noise model that does not strongly vary across the field due to no surface brightness gradient.

Notably, the areas around the HB and RC are not well modelled. In part, this is due to imposing a fixed distance. The SFH code would prefer a distance that maximizes agreement with the RC (i.e. it matches more data points). This would place the TRGB well above the observed TRGB location, thus violating a gold-standard distance measurement. Other areas of disagreement, such as on the RGB, are typical of the Padova models that tend to produce overly blue models (too hot) relative to the data (e.g. Gallart et al. 2005). Finally, the level of disagreement on the lower part of the luminous MS may be linked to poor matches of the RHeBs, which are known to be problematic for stellar models (e.g. McQuinn et al. 2011).

Beyond issues with stellar models, spatial variations in noise model, as determined from ASTs, can affect the quality of the fit. Given that the ACS field covers a large dynamic range in surface brightness (see Fig. 1), some of the residuals may be due to mismatch in the noise model.

To test this, we divided the ACS field in a 3×3 grid, and measured the SFH in each of the nine regions using photometry and ASTs limited to each region. For this exercise, we fixed the distance and foreground extinction to common values, but allowed the differential extinction to vary.

Visual inspection of residual CMDs for each of the nine regions reveal notable improvement in the residuals relative to one fit to the entire ACS field. Many of the large-scale systematic features that are apparent in the fit to the entire ACS field are absent in the individual regions. Fits to individual regions were consistent with no differential extinction, same as the SFH for the entire field.

Finally, we computed a total SFH from the nine individual regions by first running random and systematic uncertainties on each region and then by combining the nine best-fitting SFHs and their associated uncertainties. The resultant SFH from this exercise is within a few per cent (i.e. consistent within random uncertainties) of the SFH from analysing the entire field at once.

This exercise illustrates two points. First, the bulk of the poor residual areas in Fig. 4 are due to a spatial gradients in the noise model rather than issues with the physical model. Secondly, the consistency of the SFHs indicates that the entire field solution is robust.

Our model of the UVIS field (Fig. 5) is well matched to the data. There are no significant systematics in the residual CMDs. In part, the simpler stellar populations of this field are easier to model and suffer from less age–metallicity degeneracy than the ACS field. The residual for the UVIS field is better than the ACS field owing to a combination of a simpler SFH and lower variability in spatially dependent completeness.

As has been discussed extensively in the literature, the choice of underlying stellar model can affect the measurement of an SFH from a CMD, though variations are smallest for CMDs that reach below the oldest MSTO (e.g. Gallart et al. 2005; Hidalgo et al. 2009; Weisz et al. 2011; Dolphin 2012; Weisz et al. 2014; Skillman et al. 2017). To follow-up on this point, we measure the SFHs using multiple stellar libraries. As described in Appendix B, we find that while SFHs measured with different models vary in excess of random uncertainties, our reported systematic uncertainties accurately encompass this variation.

Fig. 7 shows the absolute SFH zoomed-in on the more recent 500 Myr. Here, we see that the SFH in the ACS field is nearly constant over ~ 100 Myr intervals from 100 to 500 Myr ago, with most fluctuations of order unity. There is a notable increase in SFR within the last ~ 50 Myr, over time intervals of ~ 10 Myr. Fluctuations in the SFR over these short intervals are factors of ~ 4 –6.

3 RESULTS

The main results of this paper are the SFHs of the inner and outer fields in WLM. In the following section, we describe our findings, and in Section 4, we place our results into a broader context.

3.1 SFH of the inner field

We first consider the SFH of the inner, ACS field, which is shown in Fig. 6. The top panel presents the cumulative SFH, i.e. the fraction of stellar mass formed prior to a given time. The solid black line is the best-fitting SFH, the black error bars reflect random uncertainties, and the purple error bars represent the total uncertainties, i.e. random and systematic. From the cumulative SFH, we see that WLM had an initial burst of star formation, followed by lower level star formation that increases in activity in the last several Gyr. More quantitatively, ~ 20 per cent of the stellar mass in this field formed prior to ~ 12.5 Gyr ago and 50 per cent formed prior to ~ 5 Gyr ago. The fact that 50 per cent of the stellar mass formed within the most recent ~ 5 Gyr indicates that this is a fairly young region in WLM.

The middle panel of Fig. 6 shows the metallicity evolution of the ACS field. Our best-fitting metallicity exhibits linear enrichment in logarithmic metallicity, i.e. $\log(Z)$. The early populations of WLM were quite metal poor, $[M/H] \sim -2$. By the time 50 per cent of the stellar mass formed, the field enriched to $[M/H] \sim -1.2$. Our derived present-day metallicity for this field in WLM, $[M/H] \sim -0.8$, is in excellent agreement with the spectroscopic metallicities of the young stars and H II regions (see Section 1.2). Overall, our derived age–metallicity relationship is in great agreement with that derived from spectra of red giant branch stars by Leaman et al. (2013).

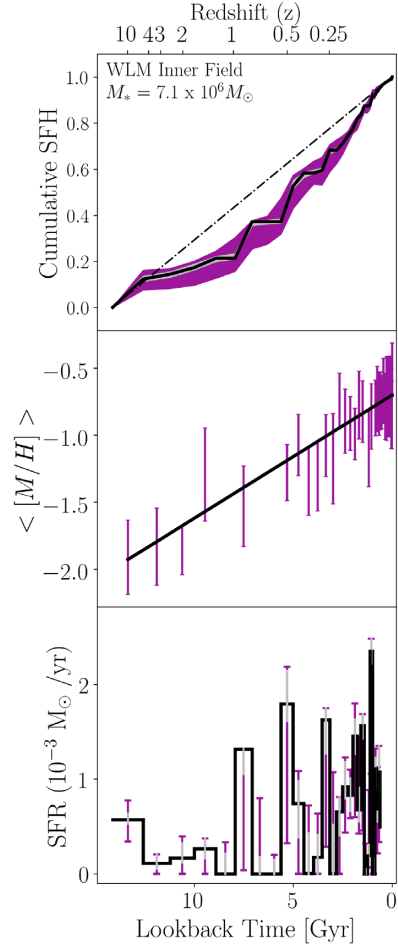


Figure 6. The SFH of the ACS field. *Top:* Cumulative SFH of Inner (ACS) Field versus Look-back Time. The black line reflects the best fit, the grey envelope reflects random uncertainties, and the purple envelope represents total uncertainties (systematic plus random). *Middle:* Mean metallicity ($\log(M/H)$) versus look-back time with random uncertainties in grey and total uncertainties (systematic plus random) in purple, as measured from fitting the CMD. Only metallicity points in which the best-fitting SFR is greater than zero are plotted. *Bottom:* SFR versus lookback time. About 20 per cent of the stars in the inner field formed prior to ~ 12 Gyr ago, and 50 per cent formed prior to ~ 5 Gyr ago. As in the panels above, the random errors are in grey and the total errors (systematic plus random) are in purple.

The bottom panel of Fig. 6 shows the absolute SFR versus look-back time. The patterns of star formation in the absolute SFR match the trends inferred from the cumulative SFH. Importantly, this plot provides a sense of how bursty the star formation in WLM has been. For most of its lifetime, fluctuations in the SFR rarely exceed factors of a few to several. Though the time resolution is only sufficient to probe short time-scale bursts in the most recent few hundred Myr, the lifetime SFH rules out sustained order-of-magnitude bursts. The strong bursts appear around ~ 5 –7 Gyr ago. While there is clearly increased activity during this time, the large amplitude of the uncertainties make a detailed deconstruction challenging. Additionally, the SFH of WLM shows no evidence of large gaps, i.e. periods of no star formation. One challenge with interpreting these bursts and gaps is that their relative timing and amplitude are coupled to the chosen stellar model. For example, in Appendix B, not all models show a burst of star formation from 5 to 7 Gyr ago,

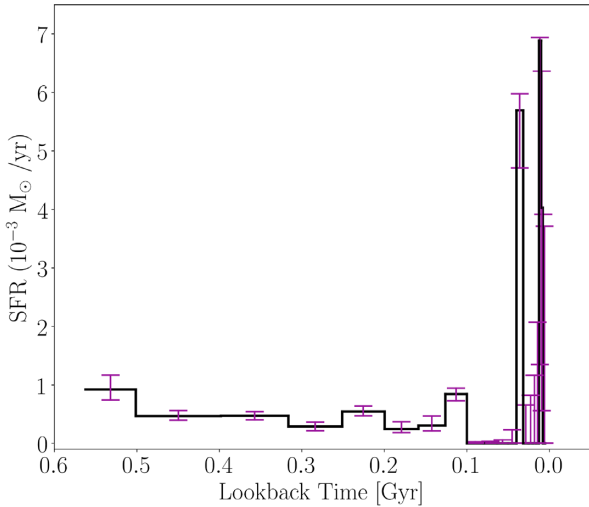


Figure 7. The absolute SFH of the ACS field over the most recent 500 Myr. In general, the SFH over the last 500 Myr has been fairly constant, with larger amplitude bursts within the more recent 50 Myr.

even though they show a similar overall shape. A detailed analysis of burstiness, and its dependence on the choice of stellar models requires a level of analysis that is beyond the scope of the this paper. We plan such a pursuit in a follow-up paper, in which we can explore trade-offs between real bursts of star formation versus features generated by artefacts in the stellar models.

3.2 SFH of the outer field

Fig. 8 shows the SFH of the UVIS field which is located along the minor axis of the galaxy. The SFH of this field is pre-dominantly old, with the bulk of star formation at earlier times: ~ 40 per cent of the stellar mass formed prior to 12.5 Gyr ago, 50 per cent by 10 Gyr ago, and 90 per cent by 7 Gyr ago. Overall, star formation in the UVIS field appears to decrease towards the present and resembles a declining τ -model SFH.

The chemical enrichment of this field follows a similar trend as the ACS field. The oldest populations are quite metal poor with $[M/H] \sim -1.7$, which is consistent with the ACS fields when factoring in uncertainties. By ~ 5 Gyr ago, this field enriched to its present-day metallicity, within errors, of $[M/H] \sim -0.8$.

The absolute SFH of the UVIS field is plotted in the bottom panel of Fig. 8. Note that the absolute scale is a factor of ~ 100 less than the ACS field. While there appear to be a number of bursts and gaps, the small number of stars results in large error bars, making them not statistically significant.

Fig. 9 shows the absolute SFH of the UVIS field plotted over the most recent 500 Myr. At all times shown, the SFH is consistent with zero, though with large error bars due to Poisson sampling statistics that are captured by the HMC runs.

4 DISCUSSION

4.1 Comparison with SFHs of other isolated dwarf galaxies

Over the past decade, there has been a concerted effort to measure SFHs of dwarf galaxies beyond the virial radii of the MW and M31. The LCID programme has been at the forefront of this effort, measuring SFHs for Leo A, Phoenix, LGS 3, Cetus, Tucana,

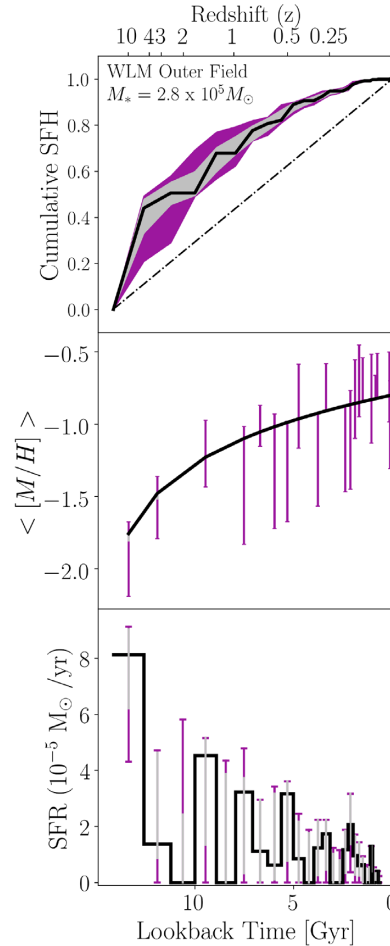


Figure 8. The same as Fig. 6 only for the SFH of the outer UVIS field. In comparison to the inner field, this outer field formed 50 per cent of its stars by ~ 12 Gyr ago and has had little star formation in the last few Gyrs.

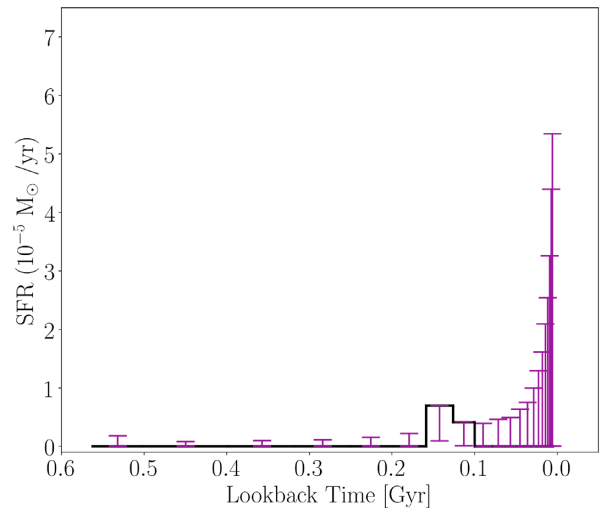


Figure 9. The absolute SFH of the UVIS field over the most recent 500 Myr. The SFH of this field is consistent with zero over the entire 500 Myr period.

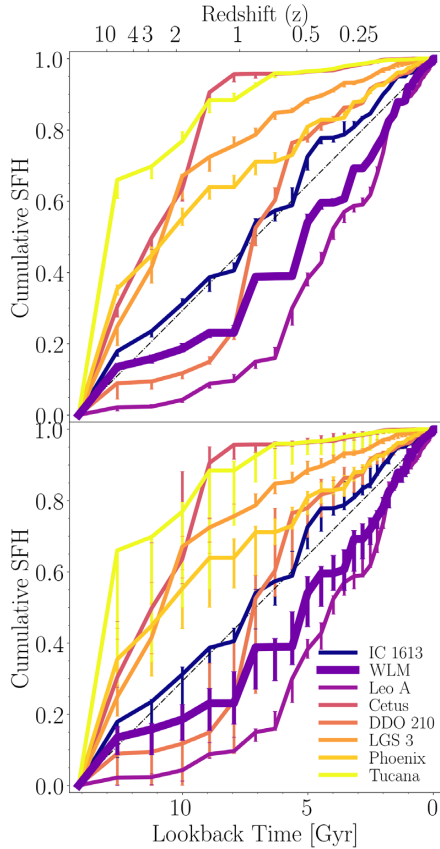


Figure 10. WLM in context. The cumulative SFHs of isolated LG dwarf galaxies, all measured with the Padova models. The upper panel shows best-fitting SFHs of each galaxy with random errors, while the lower panel includes random and systematic. Galaxy SFHs are colour coded by mass from highest mass (IC 1613) in navy to lowest mass (Tucana) in yellow. WLM, the thickest curve in purple, appears to have a rising SFH towards the present.

Phoenix, and IC 1613 (see Gallart et al. 2015 and references therein) from *HST*-based CMDs that extend below the oldest MSTO. A related programme measured the SFH of DDO 210 (Cole et al. 2014) from comparably deep *HST* imaging.

A main takeaway from LCID is that, even when removed from the effects of environment, dwarf galaxies exhibit a huge diversity in SFH. This primary finding is illustrated in Fig. 10, where we plot the cumulative SFHs of the LCID galaxies, DDO 210, and add the ACS SFH of WLM from this paper.¹ The lines are coloured by the present-day stellar mass of each systems from lowest mass (Tucana, yellow) to highest mass (IC 1613, purple).

The SFHs in Fig. 10 have all been measured with the same Padova stellar models and the same CMD-fitting code (Dolphin 2002). In the top panel, we plot the SFHs with only the random uncertainties from the CMD fits, which allows us to gauge the *relative* differences in the SFHs. Because the CMDs are so well populated, the random uncertainties are small and the SFHs of any two galaxies are inconsistent at a high level of statistical significance, as discussed in the series of LCID papers. Here, we focus on new information added by WLM.

¹ We use the ACS field as a proxy for WLM’s SFH because it forms 100 times more stellar mass than the UVIS field, meaning the UVIS field contribution to a ‘global’ SFH is negligible relative to ACS.

WLM is the second most massive galaxy in the sample, following IC 1613, making the two interesting to compare. Compared to IC 1613, WLM formed fewer stars at early times. WLM formed 50 per cent of its stellar mass by ~ 5 Gyr ago, whereas IC 1613 formed 50 per cent of its stellar mass by ~ 7.5 Gyr ago. At later times, the fraction of stars formed in WLM is rising with time, while that of IC 1613 is relatively constant. Thus, WLM formed 90 per cent of its stellar mass by ~ 1 Gyr ago, whereas IC 1613 formed 90 per cent of its stellar mass ~ 2 Gyr ago. Note that this comparison is limited by the placement and field size of *HST*. The FOV only covers ~ 10 – 20 per cent of the half-light radius in each of the galaxies, and the fields in WLM are all located at $r < 1r_h$, whereas in IC 1613, the field is located at $r \gtrsim 1r_h$. As discussed below, population gradients in both galaxies can affect this type of direct comparison.

Beyond from IC 1613, WLM’s SFH is not like any of the other isolated systems. The three lowest mass systems (Tucana, Phoenix, Cetus) formed the majority of their stellar mass prior to 10 Gyr ago and have had declining SFHs since then. DDO 210, Cetus, and Leo A are the next three most massive systems, yet their SFHs span nearly the entire range of possibility: Leo-A is mostly young, Cetus is mostly old, and DDO-210 is dominated by intermediate-age populations. In short, none are like WLM.

Gallart et al. (2015) suggest that SFHs can be separated into two classes of systems, ‘fast’ and ‘slow’, which are a manifestation of halo assembly bias (e.g. Gao, Springel & White 2005). In short, fast systems formed most of their stellar mass at early times because they were located in denser environments in the early Universe. In contrast, ‘slow’ systems were located in less dense environments and formed most of their stellar mass over longer periods of time. In this picture, WLM would be considered a ‘slow’ dwarf as it did not have an early dominant episode of star formation, suggesting it formed in a lower density region like other slow dwarfs such as IC 1613, Leo A, and DDO 210.

However, there remain several observational challenges to this interpretation. First is the small sample size. Secondly, we lack full-phase space information. This could be used to determine orbital histories, an important quantity in reconstructing the infall time into the LG and the density of environment at early times. Thirdly, the location of the *HST* fields vary from galaxy to galaxy. Placement of the *HST* field inside or outside the half-light radius could bias the median age of a system by ~ 1 Gyr relative to a true global SFH (Graus et al. 2019).

As a useful summary of the isolated dwarf galaxy SFHs, in Table 2, we list the look-back times when 50 per cent (τ_{50}) and 90 per cent (τ_{90}) of the stellar mass formed in each galaxy, along with the random and total uncertainties in those values.

4.2 Comparison to simulated dwarf galaxies

There are numerous theoretical models of isolated dwarf galaxy formation and evolution that predict SFHs as a function of various properties such as present-day stellar mass (e.g. Governato et al. 2010; Sawala et al. 2010; Hopkins et al. 2014; Shen et al. 2014; Vogelsberger et al. 2014; Christensen et al. 2016; El-Badry et al. 2016; Sawala et al. 2016; Fitts et al. 2017; Wheeler et al. 2018; Buck et al. 2019; Wright et al. 2019). While a detailed comparison of measured SFHs to all such models is of great interest, it is a large undertaking that is beyond the scope of this paper. Instead, we select a single model to illustrate one type of comparison that can be made.

Table 2. Summary statistics for isolated dwarf galaxy SFHs.

Galaxy	τ_{50} (log look-back age)	τ_{90} (log look-back age)
IC 1613	$9.87^{+0.00,0.00}_{-0.00,0.06}$	$9.30^{+0.00,0.09}_{-0.00,0.02}$
WLM ACS	$9.71^{+0.01,0.03}_{-0.00,0.07}$	$9.02^{+0.00,0.06}_{-0.00,0.01}$
WLM UVIS	$10.01^{+0.08,0.09}_{-0.01,0.02}$	$9.67^{+0.02,0.03}_{-0.07,0.10}$
Leo A	$9.61^{+0.01,0.01}_{-0.02,0.03}$	$8.93^{+0.02,0.06}_{-0.01,0.02}$
Cetus	$10.05^{+0.00,0.04}_{-0.01,0.03}$	$9.95^{+0.00,0.04}_{-0.02,0.05}$
DDO 210	$9.85^{+0.00,0.03}_{-0.01,0.02}$	$9.33^{+0.00,0.09}_{-0.01,0.02}$
LGS 3	$10.03^{+0.00,0.04}_{-0.01,0.01}$	$9.64^{+0.02,0.08}_{-0.01,0.01}$
Phoenix	$10.03^{+0.01,0.03}_{-0.01,0.08}$	$9.42^{+0.01,0.03}_{-0.02,0.02}$
Tucana	$10.11^{+0.00,0.00}_{-0.00,0.11}$	$9.88^{+0.03,0.04}_{-0.00,0.13}$

Note. τ_{50} is the look-back time when 50 per cent of the stellar mass formed and τ_{90} is the look-back time when 90 per cent of the stellar mass formed. The uncertainties reflect the random and total (random plus systematic) components.

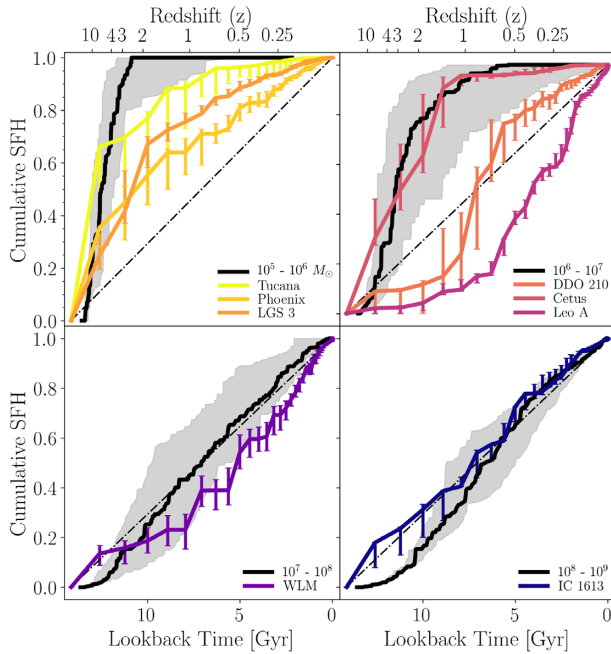


Figure 11. A comparison between the simulated dwarf galaxies from the FIRE suite (Garrison-Kimmel et al. 2019) and real SFHs measured from *HST* CMDs as a function of present-day stellar mass. The solid black lines in each panel represent the median of the simulations and the grey bands reflect the 68 per cent scatter. Uncertainties on the real galaxies include random and systematic components. The comparison is mixed with good general agreement at higher masses and worse agreement at lower masses.

Fig. 11 shows the cumulative SFHs of the real isolated dwarf galaxies from Fig. 10 and median SFHs from the 500 simulated dwarf galaxies analysed by Garrison-Kimmel et al. (2019) as part of the Feedback in Realistic Environments (FIRE; Hopkins et al. 2014, 2018) simulation suite. For the SFHs of real galaxies, we plot the systematic uncertainties, i.e. those in the bottom panel of Fig. 10.

Following fig. 4 in Garrison-Kimmel et al. (2019), we plot the median (black) and 68 per cent scatter (grey envelope) in the simulated SFHs. We have plotted the SFHs of ‘centrals in Local Groups,’ which are defined as dwarf galaxies between 0.3 and 2 Mpc from the central two galaxies (e.g. MW and M31 analogues) in FIRE

realizations of LG environments. As in Garrison-Kimmel et al. (2019), we group the real and simulated galaxies by stellar mass.

For the specific case of WLM (lower left-hand panel of Fig. 11), we find reasonable agreement between its SFH and predictions from the FIRE simulations. That is, WLM is statistically consistent (when systematics are included) with FIRE simulations in this mass range. The most notable difference is at the earliest epoch, when WLM formed 10–15 per cent of its stellar mass by 12.5 Gyr ago ($z \sim 5$) and the simulations formed 2–6 per cent. A similar discrepancy is observed at higher masses. By 12.5 Gyr ago, IC 1613 formed 15–20 per cent of its stellar mass and the simulations formed only 1–2 per cent of its stellar mass. There are a number of selection effects that could explain this difference (e.g. field placement and small number statistics). However, as discussed in Weisz et al. (2014), a similar underprediction in stellar mass formation is observed in some empirical models that include low-mass galaxies (e.g. Behroozi, Wechsler & Conroy 2013). While a detailed exploration of the earliest epochs of star formation in real and simulated dwarf galaxies is beyond the scope of this paper, understanding such discrepancies is important because ancestors of galaxies like WLM and IC 1613 played an important role in driving cosmic reionization (e.g. Boylan-Kolchin et al. 2015).

However, there are tensions between the models and data at lower stellar masses. For example, at intermediate masses ($M_* \sim 10^6\text{--}10^7 M_\odot$), Leo A and DDO 210 are inconsistent with the simulations even when uncertainties are included. The level of disagreement is worse at lower masses, where Phoenix, Tucana, and LGS 3 all exhibit much larger amounts of intermediate- and late-time star formation than the simulations predict.

There are several reasons for these tensions on both the observational and simulation sides. Foremost, we are in a regime of small number statistics. Only a handful of isolated dwarf galaxies have ‘gold standard’ SFHs. Similarly, the simulations are only drawn from two realizations of the LG, meaning only ~ 10 simulated galaxies are in each mass range. Moreover, because the simulations include the entire LG, the mass resolution in the lowest mass dwarfs is coarser compared to zoom-in realizations of individual dwarfs (e.g. Fitts et al. 2017). Numerical tests (e.g. Hopkins et al. 2018; Garrison-Kimmel et al. 2019) suggest that dwarf galaxies resolved with fewer than a few 100s of star particles are likely to have their star formation truncated too early. Additionally, as discussed in Garrison-Kimmel et al. (2019), the UV background model used (Faucher-Giguère et al. 2009) peaks too early compared to current constraints. An updated model (Faucher-Giguère 2019) is currently being implemented but results are not yet available for this analysis.

Overall, comparisons with the FIRE simulations have generally positive results, with some tensions that warrant further exploration. On one hand, WLM and IC 1613 suggest that simulations for more massive systems are reasonably accurate, modulo our understanding of the very first epochs. This is encouraging, as Skillman et al. (2014) noted that many simulations still struggled to reproduce roughly constant SFHs dwarf galaxies. On the other hand, the disagreement at lower masses may indicate that simulations are not accurately reproducing the lowest mass systems, we do not fully understand our selection effects/systematics, or both. However, before we can drawn any strong conclusions, several competing effects (e.g. the adopted UV background model, small number statistics) must be explored in more detail.

In terms of selection effects, a galaxy’s isolation is a particularly challenging selection effect to quantify. For example, Cetus and Tucana are quenched galaxies located large distances from massive host galaxies, a rarity in the local Universe (e.g. Geha et al. 2012).

Teyssier et al. (2012) suggest that these two systems likely interacted with the Milky Way in the early Universe, implying that present day isolation does not entirely mitigate the effects of environment. Moreover, galaxies such as Phoenix and LGS 3 are located at intermediate distances from the Milky Way and M31, respectively, and could be considered ‘backsplash’ galaxies, i.e. their evolution may have been affected by weak interactions with their host galaxy (e.g. Gill et al. 2005). Even if such interactions do not substantially affect a galaxy’s stellar populations (e.g. via tidal stripping; Knebe et al. 2011), it can strip gas, thereby reducing fuel for future star formation.

In principle, such effects are included in the simulated SFHs from the FIRE suite we have plotted, as they are selected to be from LG-like environments. However, this also means that the simulations must correctly model galaxy-scale interactions as well as internal processes. It may be that the underlying physics is close but not entirely correct. The use of data sets, such as the detailed SFHs of dwarf galaxies, can be useful to further refine models.

Additionally, it is important to consider *HST* field location and size in comparing measured and simulated SFHs. As highlighted by Graus et al. (2019), fields that are located within $1r_h$ are typically biased towards younger ages, relative to the true global SFH, while field located outside $1r_h$ are biased old. For Cetus, Tucana, LGS 3, and Leo A, ACS covers a large fraction of the area around the half-light radius making the bias in mean age fairly modest. However, in galaxies with larger angular sizes such as IC 1613, Phoenix, and WLM, *HST* covers smaller portions of the galaxies primarily within the half-light radius, potentially biasing the mean age young by ~ 1 Gyr (Graus et al. 2019). Accordingly, comparisons of SFHs in real and simulated dwarf galaxies must be treated with appropriate caution.

4.3 Radial gradients in WLM

Comparing SFHs of the ACS and UVIS field reveals a clear age gradient. As summarized in Table 2, the outer and inner fields of WLM formed 50 per cent of their stellar mass by ~ 10 Gyr ago ($z \sim 2$) and ~ 5 Gyr ago, respectively. This implies a very steep age gradient since the centres of the fields are only ~ 0.4 kpc apart. Several of the LCID galaxies appear to have gradual and smooth age gradients (e.g. Hidalgo et al. 2009; Monelli et al. 2010a,b), in contrast to what we find for WLM. However, WLM is a more massive system with a disc-like morphology.

In principle, WLM represents a perfect test bed for the effects of stellar feedback on low-mass galaxy formation. For example, FIRE simulations analysed by El-Badry et al. (2016) find that stellar feedback can take stars initially formed in the central regions and drive them to large radii over cosmic time-scales. Interestingly, steeper age gradients are associated with stronger feedback, and in turn, dark-matter haloes with larger cores in their central regions.

Using a combination of spectroscopy, photometry, H I observations and dynamic modelling Leaman et al. (2012) found a similar story: WLM’s evolution appears to be strongly influenced by internal feedback. Interestingly, they found a very weak metallicity gradient of -0.04 dex kpc^{-1} , which agrees with the metallicities recovered from our CMD fits. Thus, it seems that feedback could induce a steep age gradient, but only a weak metallicity gradient.

As pointed out in Graus et al. (2019), there are biases of up to ~ 1 Gyr in the values of τ_{50} based on the relative radii of the *HST* fields. However, folding this biases into our analysis is challenging, given that the Graus et al. (2019) work uses azimuthally averaged values. Our outer field is along WLM’s minor axis, and it remains

unclear how SFHs differ along the major and minor axes at a fixed radius. The implications of spatial gradients in WLM is a topic of great interest, and one that we will follow up with in another paper in this series.

4.4 Towards globally representative SFHs of isolated dwarf galaxies

One challenge in interpreting the SFHs of isolated dwarf galaxies with *HST* is the effect of a field of view (FOV) that is much smaller than the full extent of a galaxy. In almost all cases, *HST* has targeted fields within or near the half-light radius of isolated dwarfs, largely to provide enough stars for a robust SFH measurement. This placement can introduce biases into the recovered SFHs relative to the true global SFHs (e.g. Graus et al. 2019). For example, SFHs fields located exclusively within the half-light radius may be biased young by up for a few Gyr.

Going forward, mitigating this bias will require coordinated effort among major facilities. The high angular resolution capabilities of *HST*, and soon *JWST*, are needed to overcome crowding limitations to reach the oldest MSTO in a galaxy’s central region. However, the small fields of these systems are not ideal for the outer regions, which are sparsely populated and subtend large angular areas.

WFIRST is an excellent solution for measuring SFHs of the outer regions in isolated dwarfs. Though its angular resolution is twice as coarse as *HST* in the optical and *JWST* in the near-IR, its large field of view and high throughput are perfect for obtaining deep CMDs of the extended and less crowded halo populations (e.g. Williams et al. 2019). Fully leveraging the complementary nature of these telescopes will require a strategic and coordinated effort across missions.

5 CONCLUSIONS

We have measured and analysed the SFH of isolated LG dwarf galaxy WLM based on *HST* imaging of two fields that reach the oldest MSTO. We find that

- (i) The SFH of the inner ACS field ($0.5r_h$) is constant or slightly rising toward the present. Prior to 12.5 Gyr ago, this field formed 20 per cent of its stellar mass, with 50 per cent formed within the most recent 5 Gyr.
- (ii) The SFH of the outer UVIS field ($0.7r_h$) resembles a declining τ -model. This field formed 40 per cent of its stellar mass by 12.5 Gyr ago, 50 per cent by 10 Gyr ago, and 90 per cent by 7 Gyr ago.
- (iii) The chemical enrichment of WLM from the CMDs appears consistent with stellar metallicities of RGB stars reported in Leaman et al. (2013).
- (iv) We compare the SFHs of WLM and other isolated dwarf galaxies with SFHs of isolated FIRE dwarf galaxy simulations. We find good agreement for galaxies with $M_* > 10^7 M_\odot$ and not as good agreement for galaxies with $M_* < 10^7 M_\odot$. Small numbers of real galaxies and selection effects make it challenging to draw definitive conclusions.
- (v) The SFHs of the inner and outer field show that WLM has a steep age gradient. The median age changes by 5 Gyr over a distance of only 0.4 kpc. This may imply strong feedback and/or the presence of a prominent dark-matter core in WLM. However, because the outer field is along the minor axis, it is unclear if it is representative of the true SFH of WLM at that fixed radius.

(vi) The placement and size of the *HST* field of view can bias the measured SFHs relative to the true global SFH. Mitigating such biases will require a coordinated effort between *HST/JWST* to cover the crowded central regions and WFIRST to coverage the sparser, more expansive outer regions.

ACKNOWLEDGEMENTS

This study is based on observations made with the NASA/ESA *Hubble Space Telescope*, obtained at the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS5-26555. These observations are associated with programme 13768. The authors thank Shea Garrison-Kimmel and Andrew Graus for sharing results from their dwarf galaxy simulations and Piero Madau and Peter Behroozi for helpful comments during the editing process. Support for this work was provided by NASA through grants HST-GO-13768, HST-GO-15006, and JWST-ERS-1334 from the Space Telescope Science Institute, which is operated by AURA, Inc., under NASA contract NAS5-26555. SMA is supported by the National Science Foundation Graduate Research Fellowship under Grant DGE 1752814. DRW acknowledges support from an Alfred P. Sloan Fellowship and an Alexander von Humboldt Fellowship. MBK acknowledges support from NSF grant AST-1517226 and CAREER grant AST-1752913 and from NASA grants NNX17AG29G and HST-AR-14282, HST-AR-14554, and HST-GO-14191 from the Space Telescope Science Institute, which is operated by AURA, Inc., under NASA contract NAS5-26555. This research has made use of NASA's Astrophysics Data System Bibliographic Services and the NASA/IPAC Extragalactic Database (NED), which is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

REFERENCES

- Aparicio A., Hidalgo S. L., 2009, *AJ*, 138, 558
- Behroozi P. S., Wechsler R. H., Conroy C., 2013, *ApJ*, 770, 57
- Boylan-Kolchin M., Weisz D. R., Johnson B. D., Bullock J. S., Conroy C., Fitts A., 2015, *MNRAS*, 453, 1503
- Bressan A., Marigo P., Girardi L., Salasnich B., Dal Cero C., Rubele S., Nanni A., 2012, *MNRAS*, 427, 127
- Buck T., Macciò A. V., Dutton A. A., Obreja A., Frings J., 2019, *MNRAS*, 483, 1314
- Choi J., Dotter A., Conroy C., Cantiello M., Paxton B., Johnson B. D., 2016, *ApJ*, 823, 102
- Christensen C. R., Davé R., Governato F., Pontzen A., Brooks A., Munshi F., Quinn T., Wadsley J., 2016, *ApJ*, 824, 57
- Cole A. A. et al., 2007, *ApJ*, 659, L17
- Cole A. A., Weisz D. R., Dolphin A. E., Skillman E. D., McConnachie A. W., Brooks A. M., Leaman R., 2014, *ApJ*, 795, 54
- D'Onghia E., Besla G., Cox T. J., Hernquist L., 2009, *Nature*, 460, 605
- Dalcanton J. J. et al., 2009, *ApJS*, 183, 67
- Dolphin A. E., 2000a, *PASP*, 112, 1383
- Dolphin A. E., 2000b, *ApJ*, 531, 804
- Dolphin A. E., 2002, *MNRAS*, 332, 91
- Dolphin A. E., 2012, *ApJ*, 751, 60
- Dolphin A. E., 2013, *ApJ*, 775, 76
- Dolphin A. E. et al., 2003, *AJ*, 126, 187
- Duane S., Kennedy A. D., Pendleton B. J., Roweth D., 1987, *Phys. Lett. B*, 195, 216
- El-Badry K., Wetzel A., Geha M., Hopkins P. F., Kereš D., Chan T. K., Faucher-Giguère C.-A., 2016, *ApJ*, 820, 131
- El-Badry K., Weisz D. R., Quataert E., 2017, *MNRAS*, 468, 319
- Faucher-Giguère C.-A., 2019, preprint ([arXiv:1903.08657](https://arxiv.org/abs/1903.08657))
- Faucher-Giguère C.-A., Lidz A., Zaldarriaga M., Hernquist L., 2009, *ApJ*, 703, 1416
- Ferraro F. R., Fusi Pecci F., Tosi M., Buonoanno R., 1989, *MNRAS*, 241, 433
- Fillingham S. P., Cooper M. C., Boylan-Kolchin M., Bullock J. S., Garrison-Kimmel S., Wheeler C., 2018, *MNRAS*, 477, 4491
- Fitts A. et al., 2017, *MNRAS*, 471, 3547
- Fitts A. et al., 2018, *MNRAS*, 479, 319
- Gallart C., Zoccali M., Aparicio A., 2005, *ARA&A*, 43, 387
- Gallart C. et al., 2015, *ApJ*, 811, L18
- Gao L., Springel V., White S. D. M., 2005, *MNRAS*, 363, L66
- Garrison-Kimmel S. et al., 2019, *MNRAS*, 489, 4574
- Geha M., Blanton M. R., Yan R., Tinker J. L., 2012, *ApJ*, 757, 85
- Gieren W. et al., 2008, *ApJ*, 683, 611
- Gill S. P. D., Knebe A., Gibson B. K., 2005, *MNRAS*, 356, 1327
- Girardi L. et al., 2010, *ApJ*, 724, 1030
- González-Samaniego A., Bullock J. S., Boylan-Kolchin M., Fitts A., Elbert O. D., Hopkins P. F., Kereš D., Faucher-Giguère C.-A., 2017, *MNRAS*, 472, 4786
- Governato F. et al., 2010, *Nature*, 463, 203
- Graus A. S. et al., 2019, *MNRAS*, 490, 1186
- Hidalgo S. L., Aparicio A., Martínez-Delgado D., Gallart C., 2009, *ApJ*, 705, 704
- Hidalgo S. L. et al., 2018, *ApJ*, 856, 125
- Hodge P., Miller B. W., 1995, *ApJ*, 451, 176
- Holtzman J. A., Afonso C., Dolphin A., 2006, *ApJS*, 166, 534
- Hopkins P. F., Kereš D., Oñorbe J., Faucher-Giguère C.-A., Quataert E., Murray N., Bullock J. S., 2014, *MNRAS*, 445, 581
- Hopkins P. F. et al., 2018, *MNRAS*, 480, 800
- Karachentsev I. D., Makarov D. I., Kaisina E. I., 2013, *AJ*, 145, 101
- Kirby E. N., Cohen J. G., Simon J. D., Guhathakurta P., Thygesen A. O., Duggan G. E., 2017, *ApJ*, 838, 83
- Knebe A., Libeskind N. I., Knollmann S. R., Martínez-Vaquero L. A., Yepes G., Gottlöber S., Hoffman Y., 2011, *MNRAS*, 412, 529
- Kroupa P., 2001, *MNRAS*, 322, 231
- Leaman R. et al., 2012, *ApJ*, 750, 33
- Leaman R. et al., 2013, *ApJ*, 767, 131
- Lee H., Skillman E. D., Venn K. A., 2005, *ApJ*, 620, 223
- Makarov D., Makarova L., Rizzi L., Tully R. B., Dolphin A. E., Sakai S., Shaya E. J., 2006, *AJ*, 132, 2729
- McConnachie A. W., 2012, *AJ*, 144, 4
- McQuinn K. B. W., Skillman E. D., Dalcanton J. J., Dolphin A. E., Holtzman J., Weisz D. R., Williams B. F., 2011, *ApJ*, 740, 48
- Minniti D., Zijlstra A. A., 1996, *ApJ*, 467, L13
- Minniti D., Zijlstra A. A., 1997, *AJ*, 114, 147
- Monelli M. et al., 2010a, *ApJ*, 720, 1225
- Monelli M. et al., 2010b, *ApJ*, 722, 1864
- Rizzi L., Tully R. B., Makarov D., Makarova L., Dolphin A. E., Sakai S., Shaya E. J., 2007, *ApJ*, 661, 815
- Robles V. H. et al., 2017, *MNRAS*, 472, 2945
- Rosenfield P., Marigo P., Girardi L., Dalcanton J. J., Bressan A., Williams B. F., Dolphin A., 2016, *ApJ*, 822, 73
- Rosenfield P. et al., 2017, *ApJ*, 841, 69
- Sawala T., Scannapieco C., Maio U., White S., 2010, *MNRAS*, 402, 1599
- Sawala T. et al., 2016, *MNRAS*, 457, 1931
- Schlafly E. F., Finkbeiner D. P., 2011, *ApJ*, 737, 103
- Shen S., Madau P., Conroy C., Governato F., Mayer L., 2014, *ApJ*, 792, 99
- Skillman E. D., Terlevich R., Melnick J., 1989, *MNRAS*, 240, 563
- Skillman E. D. et al., 2014, *ApJ*, 786, 44
- Skillman E. D. et al., 2017, *ApJ*, 837, 102
- Su K.-Y. et al., 2018, *MNRAS*, 480, 1666
- Teyssier M., Johnston K. V., Kuhlen M., 2012, *MNRAS*, 426, 1808
- Tolstoy E. et al., 1998, *AJ*, 116, 1244
- Urbaneja M. A., Kudritzki R.-P., Bresolin F., Przybilla N., Gieren W., Pietrzyński G., 2008, *ApJ*, 684, 118
- Vogelsberger M., Zavala J., Simpson C., Jenkins A., 2014, *MNRAS*, 444, 3684
- Weisz D. R. et al., 2011, *ApJ*, 739, 5

- Weisz D. R., Dolphin A. E., Skillman E. D., Holtzman J., Gilbert K. M., Dalcanton J. J., Williams B. F., 2014, *ApJ*, 789, 147
Wheeler C. et al., 2018, preprint (arXiv:1812.02749)
Williams B. et al., 2019, Bulletin of the American Astronomical Society. p. 301
Williams B. F. et al., 2014, *ApJS*, 215, 9
Wright A. C., Brooks A. M., Weisz D. R., Christensen C. R., 2019, *MNRAS*, 482, 1176

APPENDIX A: TABULATED SFHS

Tables A1 and A2 list the cumulative and absolute SFHs, along with metallicity enrichment, for the ACS and UVIS fields, respectively. The first error listed is the random error, the second error is the total error (random plus systematic). 10 lines are listed in these tables. The full tables are available online.

Table A1. Column (1) log youngest look back time in bin; column (2) log oldest look back time in bin; column (3) cumulative SFH; column (4) absolute SFH; column (5) mean metallicity. Note that the value of [M/H] is set to 0 when there is no star formation.

$\log(t_1)$ (yr ago) (1)	$\log(t_2)$ (yr ago) (2)	cSFH (3)	SFR ($10^{-3} M_{\odot} \text{ yr}^{-1}$) (4)	[M/H] (dex) (5)
6.6	6.7	$1.00^{+0.00,0.00}_{-0.00,0.00}$	$0.00^{+3.66,3.66}_{-0.00,0.00}$	$0.00^{+0.00,0.00}_{-0.70,0.70}$
6.7	6.8	$1.00^{+0.00,0.00}_{-0.00,0.00}$	$0.00^{+3.71,3.71}_{-0.00,0.00}$	$0.00^{+0.00,0.00}_{-0.70,0.70}$
6.8	6.9	$1.00^{+0.00,0.00}_{-0.00,0.00}$	$0.00^{+3.91,3.91}_{-0.00,0.00}$	$0.00^{+0.00,0.00}_{-0.70,0.70}$
6.9	7.0	$1.00^{+0.00,0.00}_{-0.00,0.00}$	$4.03^{+2.33,2.33}_{-3.47,4.03}$	$-0.70^{+0.00,0.39}_{-0.00,0.00}$
7.0	7.1	$1.00^{+0.00,0.00}_{-0.00,0.00}$	$6.89^{+0.05,1.82}_{-5.54,5.61}$	$-0.70^{+0.00,0.10}_{-0.00,0.40}$
7.1	7.2	$1.00^{+0.00,0.00}_{-0.00,0.00}$	$0.00^{+2.07,2.07}_{-0.00,0.00}$	$0.00^{+0.00,0.00}_{-0.70,1.30}$
7.2	7.3	$1.00^{+0.00,0.00}_{-0.00,0.00}$	$0.00^{+1.16,1.16}_{-0.00,0.00}$	$0.00^{+0.00,0.00}_{-0.70,0.70}$
7.3	7.4	$1.00^{+0.00,0.00}_{-0.00,0.00}$	$0.00^{+0.82,0.82}_{-0.00,0.00}$	$0.00^{+0.00,0.00}_{-0.70,0.70}$
7.4	7.5	$1.00^{+0.00,0.00}_{-0.00,0.00}$	$0.00^{+0.65,0.65}_{-0.00,0.00}$	$0.00^{+0.00,0.00}_{-0.70,0.70}$
7.5	7.6	$1.00^{+0.00,0.00}_{-0.00,0.00}$	$5.69^{+0.28,3.53}_{-0.99,1.00}$	$-0.70^{+0.00,0.20}_{-0.00,0.40}$

Table A2. Same as Table A1 only for the UVIS field.

$\log(t_1)$ (yr ago) (1)	$\log(t_2)$ (yr ago) (2)	cSFH (3)	SFR ($10^{-3} M_{\odot} \text{ yr}^{-1}$) (4)	[M/H] (dex) (5)
6.6	6.7	$1.00^{+0.00,0.00}_{-0.00,0.00}$	$0.00^{+0.07,0.07}_{-0.00,0.00}$	$0.00^{+0.00,0.00}_{-0.80,1.00}$
6.7	6.8	$1.00^{+0.00,0.00}_{-0.00,0.00}$	$0.00^{+0.05,0.05}_{-0.00,0.00}$	$0.00^{+0.00,0.00}_{-0.80,0.80}$
6.8	6.9	$1.00^{+0.00,0.00}_{-0.00,0.00}$	$0.00^{+0.04,0.04}_{-0.00,0.00}$	$0.00^{+0.00,0.00}_{-0.80,0.80}$
6.9	7.0	$1.00^{+0.00,0.00}_{-0.00,0.00}$	$0.00^{+0.03,0.03}_{-0.00,0.00}$	$0.00^{+0.00,0.00}_{-0.80,0.80}$
7.0	7.1	$1.00^{+0.00,0.00}_{-0.00,0.00}$	$0.00^{+0.03,0.03}_{-0.00,0.00}$	$0.00^{+0.00,0.00}_{-0.80,0.80}$
7.1	7.2	$1.00^{+0.00,0.00}_{-0.00,0.00}$	$0.00^{+0.02,0.02}_{-0.00,0.00}$	$0.00^{+0.00,0.00}_{-0.80,0.80}$
7.2	7.3	$1.00^{+0.00,0.00}_{-0.00,0.00}$	$0.00^{+0.02,0.02}_{-0.00,0.00}$	$0.00^{+0.00,0.00}_{-0.80,0.80}$
7.3	7.4	$1.00^{+0.00,0.00}_{-0.00,0.00}$	$0.00^{+0.01,0.01}_{-0.00,0.00}$	$0.00^{+0.00,0.00}_{-0.80,0.80}$
7.4	7.5	$1.00^{+0.00,0.00}_{-0.00,0.00}$	$0.00^{+0.01,0.01}_{-0.00,0.00}$	$0.00^{+0.00,0.00}_{-0.80,0.80}$
7.5	7.6	$1.00^{+0.00,0.00}_{-0.00,0.00}$	$0.00^{+0.01,0.01}_{-0.00,0.00}$	$0.00^{+0.00,0.00}_{-0.80,0.80}$

APPENDIX B: EFFECTS OF STELLAR MODELS ON THE SFH OF WLM

Choice of underlying stellar model has been a long-standing challenge in translating CMDs to SFHs. CMDs are so information rich that variations in the underlying stellar interior, atmosphere physics (e.g. choice in mixing length value, boundary conditions, nuclear reaction rates) and associated implementation can lead to differences in ages and SFHs. Gallart et al. (2005) present an in-depth discussion on how adopted physics in given stellar models affect the CMDs they predict. Choi et al. (2016) includes similar discussion with newer stellar models.

The effect of stellar models on SFHs measured from CMDs has been explored extensively in the past (e.g. Aparicio & Hidalgo 2009; Weisz et al. 2011; Dolphin 2012; Weisz et al. 2014; Skillman et al. 2017). In essence, when a CMD reaches the oldest MSTO with sufficient SNR ($\gtrsim 5-10$), the resulting SFH weakly depends on the choice in stellar models. The impact of stellar models on the SFH becomes larger as the CMD gets shallower. This is because SFHs from shallower CMDs are measured only from evolved stars (e.g. RGB, HB, AGB) whose age sensitivity is poorer than MSTO and sub-giants. Additionally, their exact location on the CMD is more sensitive to choice in underlying stellar model.

In the process of measuring the SFH of WLM from our deep *HST* data, we varied the choice in stellar model, holding all other parameters fixed (e.g. same IMF, age/metallicity binning, distance). The results for the ACS and UVIS fields are shown in Figs B1 and B2, respectively. The SFHs are colour coded by stellar model: Padova (navy; Girardi et al. 2010); MIST (purple; Choi et al. 2016); BaSTI (magenta; Hidalgo et al. 2018); and PARSEC (yellow; Bressan et al. 2012).

The upper panel in each plot shows the best-fitting cumulative SFHs for each model with random uncertainties, i.e. computed following Dolphin (2013). In both the ACS and UVIS fields, the SFHs are qualitatively quite similar, which is expected from such deep data. Because the random errors are quite small, they are often in tension at the several sigma level on a strictly statistical basis. This is perhaps useful for diagnosing challenges in the underlying physics (e.g. Rosenfield et al. 2016, 2017), but it poses a challenge for interpreting the SFH of a galaxy, i.e. which one is correct?

One solution proposed by Dolphin (2012) is to include an error term on the SFH that is meant to encompass plausible variations on the stellar models. More specifically, the procedure is to use a Monte Carlo process to sample variations in stellar models in the $M_{\text{bol}}-T_{\text{eff}}$ plane, and re-fit the CMD with a slightly perturbed set of models. Repeating this process for many iterations is designed to produce an error estimate that captures the effects of varying stellar models. Implementation and calibration details are discussed in Dolphin (2012).

The bottom panels of Figs B1 and B2 include both the random and systematic errors on each of the SFHs. These inflated uncertainties have brought the SFHs from different models into formal statistical agreement, i.e. $1-\sigma$. Thus, by virtue of a more comprehensive uncertainty treatment, the SFHs are now a good representation of the true underlying SFH of the CMD.

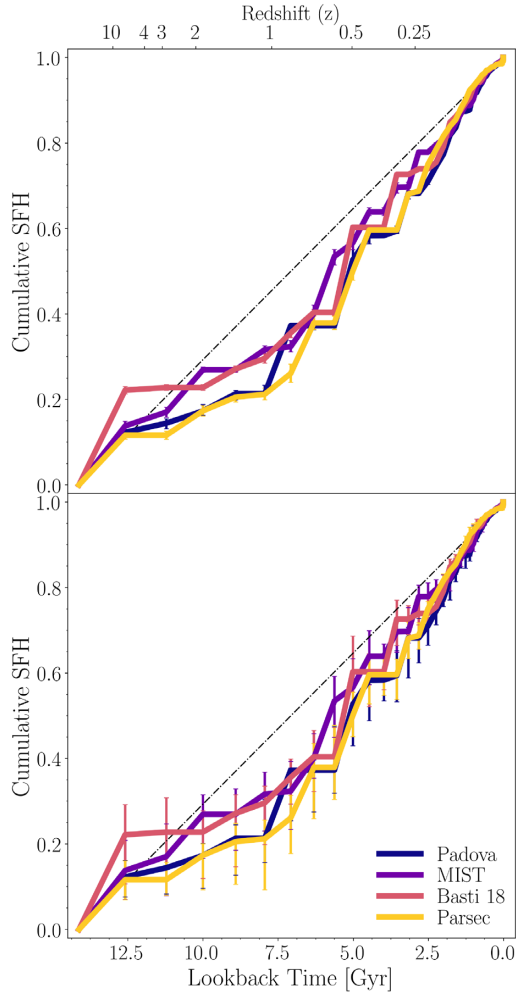


Figure B1. The above plot shows the SFH for the inner ACS field of WLM using different stellar models. The top panel displays only the random errors, while the bottom panel displays the total error (random plus systematic). Overall, there is good agreement between the different models.

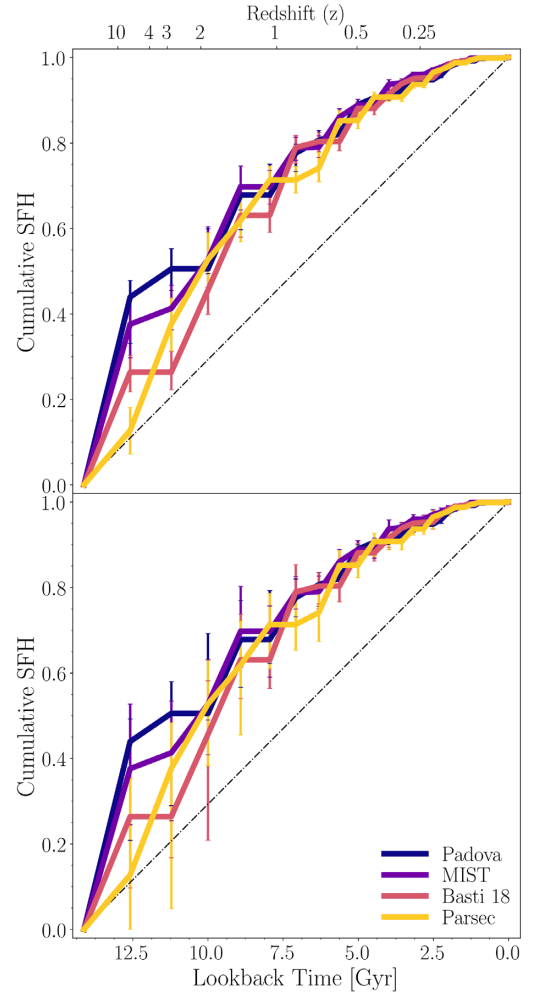


Figure B2. The above plot shows the SFH for the outer UVIS field of WLM using different stellar models. The top panel displays only the random errors, while the bottom panel displays the total error (random plus systematic). The overall agreement between the different models supports the robustness of our measured SFH.

This paper has been typeset from a \LaTeX file prepared by the author.