DEVILS: Cosmic evolution of SED-derived metallicities and their connection to star-formation histories

Jessica E. Thorne,1* Aaron S. G. Robotham,1 Sabine Bellstedt,1 Luke J. M. Davies,1 Robin H. W. Cook,1 Luca Cortese,1 Benne Holwerda,2 Steven Phillipps,3 Malgorzata Siudek4,5
1 ICRAR, The University of Western Australia, 35 Stirling Highway, Crawley, WA 6009, Australia
2 Department of Physics and Astronomy, University of Louisville, Natural Science Building 102, 40292 KY, Louisville, USA
3 Astrophysics Group, School of Physics, University of Bróstol, Bristol BS8 1TL, UK
4 Institut de Física d’Altes Energies (IFAЕ), The Barcelona Institute of Science and Technology, 08193 Bellaterra (Barcelona), Spain
5 Institute of Space Sciences (ICE, CSIC), Campus UAB, Carrer de Magrans, 08193 Barcelona, Spain

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT
Gas-phase metallicities of galaxies are typically measured through auroral or nebular emission lines, but metallicity also leaves an imprint on the overall spectral energy distribution (SED) of a galaxy and can be estimated through SED fitting. We use the ProSPECT SED fitting code with a flexible parametric star formation history and an evolving metallicity history to self-consistently measure metallicities, stellar mass, and other galaxy properties for ~ 90,000 galaxies from the Deep Extragalactic VIisible Legacy Survey (DEVILS) and Galaxy and Mass Assembly (GAMA) survey. We use these to trace the evolution of the mass-metallicity relation (MZR) and show that the MZR only evolves in normalisation by ~ 0.1 dex at stellar mass $M_\star = 10^{10.5} M_\odot$. We find no difference in the MZR between galaxies with and without SED evidence of active galactic nuclei emission at low redshifts ($z < 0.3$). Our results suggest an anti-correlation between metallicity and star formation activity at fixed stellar mass for galaxies with $M_\star > 10^{10.5} M_\odot$ for $z < 0.3$. Using the star formation histories extracted using ProSPECT we explore higher-order correlations of the MZR with properties of the star formation history including age, width, and shape. We find that at a given stellar mass, galaxies with higher metallicities formed most of their mass over shorter timescales, and before their peak star formation rate. This work highlights the value of exploring the connection of a galaxy’s current gas-phase metallicity to its star formation history in order to understand the physical processes shaping the MZR.

Key words: galaxies: abundances – galaxies: general – galaxies: evolution – galaxies: star formation – galaxies: stellar content

1 INTRODUCTION
The relationship between the stellar mass and gas phase metallicity (MZR) of galaxies has been well studied and is known to be positively correlated (e.g. Lequeux et al. 1979; Tremonti et al. 2004). This correlation has been shown to hold to $z \approx 3.5$ (Maiolino et al. 2008; Mannucci et al. 2009; Sanders et al. 2021; Cullen et al. 2021), and shows evidence of flattening at higher stellar masses ($M_\star > 10^{10} M_\odot$, Lara-Lopez et al. 2010, 2013; Curti et al. 2020).

The MZR provides essential insight into galaxy evolution as metallicity is a tracer of the fraction of baryonic mass that has been converted into stars but is also sensitive to metal loss as a result of stellar winds, supernovae and active galactic nuclei (AGN) feedback (Larson & Dinerstein 1975; Brooks et al. 2007; de Rossi et al. 2007; Kobayashi et al. 2007; Davé et al. 2011; Tumlinson et al. 2017).

The MZR is thought to be shaped by two main physical processes: metal production and metal loss. To first order, the correlation between stellar mass and gas-phase metallicity can be explained using a closed box scenario where a galaxy initially consists of metal-poor gas which collapses to form stars and no gas enters or leaves the galaxy (see the review by Tinsley 1980). In this scenario, high-mass stars produce metals through nucleosynthesis and return these metals to the interstellar medium through stellar winds. With these assumptions, the gas-phase metallicity ($Z_{gal}$) increases as gas is processed through stars. This is often described as ‘chemical downsizing’ and suggests that low mass systems are still at an early evolutionary stage and have yet to convert most of their gas into stars (Maiolino et al. 2008; Calura et al. 2009; Vale Asari et al. 2009; Zahid et al. 2011; Somerville & Davé 2015; Spitoni et al. 2020; Zhou et al. 2022).

Metal loss refers to outflows generated by starburst winds (or potentially AGN feedback, see Camps-Fariña et al. 2021) which eject metal-enriched gas into the inter-galactic medium (Garnett 2002; Tremonti et al. 2004; Dalcanton 2007; Peeples & Shankar 2011; Chisholm et al. 2018). Outflows are often used to explain the lower metallicity in low-mass systems where gas can be efficiently expelled by winds and outflows due to a shallower gravitational potential (D’Eugenio et al. 2018). However, the importance of outflows might be lower at high redshift where gas fractions are significantly higher and metal dilution through gas accretion might play a larger role (see e.g. Scoville et al. 2017; Sanders et al. 2022).
The fact that the physical processes driving the MZR are poorly understood has prompted a number of authors to consider potential higher dimension relations including additional physical properties which could provide better constraints on the processes driving metallicity evolution. Of these, the three dimensional relationship between mass, metallicity and star formation rate (SFR) or specific SFR (sSFR) has been studied in the most detail (e.g. Ellison et al. 2008; Lara-López et al. 2010; Mannucci et al. 2010; Yates et al. 2012; Lara-López et al. 2013; Hunt et al. 2016; Hirschauer et al. 2018; Curti et al. 2020; Bellstedt et al. 2021). Ellison et al. (2008) first showed an anti-correlation between metallicity and sSFR for galaxies at fixed stellar mass. Mannucci et al. (2010) also observed a secondary dependence of the MZR on the SFR. Both Yates et al. (2012) and Lara-López et al. (2013) find a similar anti-correlation for low-mass galaxies, however, they demonstrate that for massive galaxies, the median metallicity is higher for galaxies with higher SFRs. Lara-López et al. (2013) attribute this to a combination of downsizing and differing amounts of neutral gas, while Yates et al. (2012) suggest that accretion may play a significant role in regulating the gas-phase metallicities of massive galaxies. Curti et al. (2020) also find a strong anti-correlation between metallicity and SFR at low masses, however, they find that the anti-correlation weakens with increasing mass until disappearing at high masses ($M_*>10^{10.5} M_\odot$).

Another key element to understanding the MZR is understanding how it has evolved with time. Studies of metallicities in high redshift galaxies have shown that the MZR has evolved in normalisation since earlier times, but has maintained a similar shape (e.g. Mannucci et al. 2010; Henry et al. 2013; Sanders et al. 2015; Ly et al. 2016; Kashino et al. 2017; Sanders et al. 2018; Kashino et al. 2019; Weldon et al. 2020; Sanders et al. 2020, 2021; Chartab et al. 2021; Topping et al. 2021; Gillman et al. 2022). Understanding how mass, metallicity, and star formation evolve with time and in relation to one another is critical to understand the physical processes that govern the efficiency and timing of star formation in galaxies.

Gas-phase metallicities of galaxies are typically measured through auroral or nebular emission lines. However, metallicity also leaves an imprint on the overall spectral energy distribution of a galaxy. Bellstedt et al. (2021) used metallicities derived using SED fitting for 4,500 galaxies with $z < 0.06$ from the Galaxy and Mass Assembly (GAMA) survey to extract the MZR at $z = 0$. Using 37,809 galaxies from the Deep Extragalactic VIsible Legacy Survey (DEVILS) and 50,147 galaxies from the GAMA survey, we extend this analysis to $z \approx 1$ to investigate the evolution of the MZR and the impact of star formation history on the scatter of the MZR.

The structure of this paper is as follows. We describe the DEVILS and GAMA projects, related data sets, and the SED fitting method employed for this work in Section 2. We present the recovered MZR for $z = 0$ – 1 in Section 3 including the MZR for AGN host galaxies and a comparison to simulations. In Section 4 we investigate the impact of star formation history on the scatter of the MZR. We summarise our results in Section 5. Throughout this work we use a Chabrier (2003) IMF and all magnitudes are quoted in the AB system. We adopt the Planck Collaboration et al. (2016) cosmology with $H_0 = 67.8 $ km s$^{-1}$ Mpc$^{-1}$, $\Omega_M = 0.308$ and $\Omega_{\Lambda} = 0.692$.

2 DATA

2.1 Deep Extragalactic VIsible Legacy Survey

For this work we use the Deep Extragalactic VIsible Legacy Survey (DEVILS; Davies et al. 2018). DEVILS is an optical spectroscopic redshift survey using the Anglo-Australian Telescope specifically designed to have high spectroscopic completeness over a large redshift range ($z < 1$) in three well-studied extragalactic fields: XMMSLSS/D02, ECDFS/D03, and COSMOS/D10 covering a total of 4.5 deg$^2$. DEVILS has measured spectroscopic redshifts to extend the redshift completeness limit by 1-2 mags in each of these regions to an 85% completeness limit of $Y_{\text{mag}} = 20/19.6/20.7$ in D02/D03/D10 respectively. This amounts to over 56,000 spectroscopic redshifts with high completeness, allowing for robust characterization of group and pair environments in the distant universe. For a full description of the survey science goals, survey design, target selection and spectroscopic observations see Davies et al. (2018).

In Thorne et al. (2021, 2022) we use the spectroscopic and photometric data from the D10-COSMOS field as it is the deepest field. We use the DEVILS photometry catalogue derived using the ProFound source extraction code (Robotham et al. 2018) and described in depth by Davies et al. (2021). ProFound is used for source finding and photometry extraction consistently across 22 bands spanning the FUV-FIR (1500 Å - 500 μm) and includes GALEX FUV NUV (Zamojski et al. 2007), CFHT u (Capak et al. 2007), Subaru HSC griz (Aihara et al. 2019), VISTA YJHK$_s$ (McCracken et al. 2012), Spitzer IRAC1 IRAC2 IRAC3 IRAC4 MIPS24 MIPS70 (Laigle et al. 2016; Sanders et al. 2007), and Herschel P100 P160 S250 S350 S500 (Lutz et al. 2011; Oliver et al. 2012) bands. Briefly, photometry was extracted in two phases to account for large differences in resolution and depth between the FUV-NIR and MIR-FIR regimes. Photometry for the FUV-IRAC4 bands was extracted using the segment mode in ProFound while the photometry for the MIPS24-S500 bands was extracted using the FitMagPSF mode in ProFound. Due to poor resolution and shallow imaging, FIR photometry was only extracted for optically selected objects with $Y < 21.2$ mag or that were detected in the MIPS24 imaging. If an object met the criteria for FIR photometry extraction it was passed through the FitMagPSF mode. Objects that met the FIR photometry criteria but were not detected in a FIR band have a flux measurement of zero and a flux error measured from the sky noise. If no attempt was made to measure FIR photometry for an object then it will have no value for the flux and flux error. 5 per cent of all optically-detected objects met the FIR photometry criteria and therefore have FIR photometry measurements.

As described in Thorne et al. (2021), redshift catalogues for DEVILS spanning $0< z < 8$ have been compiled using photometric, grism, and spectroscopic redshifts. The redshift sources used for D10 are presented in table C1 of Thorne et al. (2021) and include 3,394 redshifts measured as part of the DEVILS program. We use these photometry and redshift measurements in this work. As in Thorne et al. (2021), we remove all objects classed as stars (starflag column), artefacts (artefactflag) or that are masked (mask). Each of these flags are described in detail by Davies et al. (2021), but briefly, stars and ambiguous objects are identified through size and colour, with cuts defined using the source type derived from the photometric redshift fitting code LePhare from the COSMOS2015 catalogue (Laigle et al. 2016). Potential artefacts are flagged where flux is not associated with an astronomical source if the source is only detected in one optical/NIR band, if the source’s Y-band R50 is less than half a pixel, or if the source has an r-z colour $< -0.75$ mag (an unphysical colour). Masking is performed to remove ghosting around bright stars as this is a significant problem in the DEVILS imaging (see figure 8 of Davies et al. 2021). This results in 49,084 objects, of which 24,099 have spectroscopic redshifts and 7,307 have grism redshifts.
2.2 Galaxy and Mass Assembly Survey

To supplement DEVILS at low-redshift, we also use the spectroscopic and photometric data from the 4th data release of the GAMA survey (Driver et al. 2022). GAMA was a large spectroscopic campaign on the Anglo-Australian Telescope targeting five fields (G02, G09, G12, G15, and G23), which gathered redshifts for ∼ 300,000 galaxies across a total sky area of 230 square degrees. GAMA targets were selected by size and colour above a magnitude limit of $r_{\text{mag}} \leq 19.8$ (or $i_{\text{mag}} \leq 19.0$ in G23).

We use the far-UV to far-IR photometry derived using the Pro-Found source-finding software described in detail by Bellstedt et al. (2020a). The photometric bands from this data release include GALEX FUV and NUV (Martin et al. 2005); VST u, g, r, i (de Jong et al. 2013a,b); VISTA Z, Y, J, H, Ks (Edge et al. 2013); WISE W1, W2, W3, W4 (Wright et al. 2010) and Herschel P100, P160, S250, S350, and S500 (Eales et al. 2010). The photometry for GAMA is derived in much the same way as for DEVILS, with minor differences due to variations in depth between the surveys.

For this work, we use the three equatorial fields (G09, G12, and G15) as well as G23. We select all objects with $z > 0$, a redshift quality flag $qZ \geq 3$ (where spectroscopic redshifts are reasonably certain, $P > 90$ per cent), and are classified as a galaxy based on size and colour (UBERCLASS=galaxy). This results in a sample of 233,762 galaxies all of which have spectroscopic redshifts.

2.3 SED Modelling

2.3.1 Impact of metallicity on a galaxy’s SED

The metallicity of stars in a galaxy impacts its SED in two ways. Firstly, an increase in metallicity in a stellar population results in lower effective temperatures, including a cooler main sequence and giant branch. Secondly, at fixed effective temperature, an increase in metallicity results in strong spectroscopic absorption features. Both of these effects result in an overall reddening of an SED with increasing metallicity (see Combes 2013 for more details).

To highlight the impact of metallicity on a galaxy’s SED, Figure 1 shows SEDs generated using ProSpect (Robotham et al. 2020) for two mock galaxies with ages 1 and 8 Gyr and stellar masses of $M_\star = 10^{10} M_\odot$. We assume that chemical enrichment is linearly mapped to the mass growth and show the resulting SED generated with a range of final metallicities spanning the limits of the Bruzual & Charlot (2003) stellar population models. It is clear that for both stellar ages, the change in metallicity has a noticeable impact on the SED fitting code (Robotham et al. 2020), with the Bruzual & Charlot (2003) stellar templates, Chabrier (2003) IMF and the Charlot & Fall (2000) dust attenuation and Dale et al. (2014) dust re-emission models. In our analysis we use the massfunc_massmap_lin parameterisation for the star formation history, which takes the form of a skewed Normal distribution, with the peak position ($mpeak$), peak SFR ($mSFR$), SFH width ($\text{mperiod}$), and SFH skewness ($\text{mskew}$) set as free parameters. The SFH is anchored to 0 at a lookback time of 13.4 Gyr, selected to be the age at which galaxies start forming (equivalent to $z = 11$, Oesch et al. 2016). Although a galaxy’s SED shows imprints of the metallicity of its stellar content, we model metallicity in ProSpect by considering the evolution of a galaxy’s gas phase metallicity across time. Our implementation maps the metallicity evolution of each galaxy linearly to the total stellar mass growth through the $Zfunc_massmap_\text{lin}$ function where the final metallicity for each galaxy is a free parameter, $Z_{\text{final}}$.

This ensures that chemical enrichment in a galaxy follows the assumed star formation rate, where increased star formation is associated with an increased rate of metal production. In this implementation, the value of the metallicity at any point in the history of the galaxy represents the gas-phase metallicity and thus the metallicity of the next generation of stars.

In the context of this work the physical meaning of the $Z_{\text{final}}$ parameter differs from the typically measured $Z_{\text{gas}}$ and stellar metallicity ($Z_\star$) from spectral line measurements. In the implementation of the evolving metallicity in ProSpect, $Z_{\text{final}}$ represents the metallicity of the gas from which the final stars formed. In this sense it is closest to the $Z_{\text{gas}}$ measured from spectral lines. However, if the gas content of a galaxy changed since the last epoch of star formation, for
example if it continued to accrete lower metallicity gas then the true $Z_{\text{gas}}$ of the galaxy would be lower than we measure with ProSPECT. Despite this, we will refer to our measured metallicity as $Z_{\text{gas}}$ for the remainder of this work.

As outlined in Bellstedt et al. (2020b) allowing for chemical enrichment over the lifetime of a galaxy directly impacts the derived star formation histories for each galaxy. Additionally, the metallicities recovered from SED fitting with ProSPECT have been shown to agree well with measurements directly from spectra (see figure 1 from Bellstedt et al. 2021), albeit with significant scatter.

In addition to the five free parameters specifying the star formation and metallicity histories, we include four free parameters to describe the contribution of dust to the SED. Within ProSPECT the dust is assumed to exist in two forms; in birth clouds formed around young stars (age $< 10^7$ yr), or distributed as a screen in the ISM. For each of these components we include two free parameters, describing the dust opacity ($\alpha_{\text{screen}}, \alpha_{\text{birth}}$), and the dust radiation field intensity ($\alpha_{\text{alpha_screen}}, \alpha_{\text{alpha_birth}}$). Figure 3 of Thorne et al. (2021) shows the impact of each parameter on a generated galaxy SED.

We also include an AGN component by incorporating the model outlined in Fritz et al. (2006) and Feltre et al. (2012). This models the primary source as a composition of power-laws, with different spectral indices as a function of the wavelength. To model the contribution from the torus, the Fritz et al. (2006) model uses a simple but realistic torus geometry, a flared disc, and a dust grain distribution function including a full range of grain sizes and assumes that the dust in the AGN torus is smoothly distributed. Within ProSPECT we model the AGN contribution by fitting the luminosity of the central source ($\text{AGN}_{\text{lum}}$), optical depth at 9.7$\mu$m ($\text{AGN}_{\alpha}$), and angle of observation ($\text{AGN}_{\text{ang}}$). We also re-attenuate the emission from the central source and dust torus through the general ISM screen (see figure 1 of Robotham et al. 2020).

2.4 Sample Selection

Due to the metallicity-dust-age degeneracy, and the subtle imprint left by metallicity on a galaxy’s SED (Figure 1), measuring metallicities from broad-band photometry can be difficult. If the signal to noise ratio of the optical-NIR photometry is too low, then subtle differences in the SED caused by differences in metallicity can be hidden in the uncertainties and not recovered when SED fitting. This means that we are limited to brighter objects with higher signal to noise ratio, generally low-redshift or high-mass galaxies. But in contrast to typical metallicity measurements, we are not limited to star-forming galaxies by the need to have emission lines.

In the context of this work, we are interested in understanding where SED-derived metallicity values are believable as a function of stellar mass and redshift. This is to ensure that we are not making cuts that will bias the resulting sample to particular galaxy types and to ensure that at each redshift and stellar mass we have a complete galaxy sample.

To determine the stellar mass cut for the D10 sample and how it evolves with redshift we investigate three techniques. The first technique identifies the stellar mass where the scatter in the MZR significantly increases, and cuts the sample below this mass. An artificial increase in the scatter of the MZR at low stellar masses can be caused by low signal to noise photometry being unable to constrain the metallicity of the galaxy. When this happens, the optimization routines preferentially select lower metallicities, often the
exact metallicity values of the Bruzual & Charlot (2003) templates rather than metallicities that require interpolated spectra. To calculate where this artificial increase in scatter occurs, we bin the MZR in 0.25 dex stellar mass bins and identify where the change in 1σ scatter between two adjacent mass bins is largest. These are changes of ~ 0.6 dex for the low redshift bins, and ~ 0.4 dex for the two highest redshift bins.

The second and third techniques investigate the relationship between the scatter in the MZR and the average error in the $i - J$ colour as uncertainties in the optical-NIR photometry result in a lower constraint on the metallicity. In selecting the colour choice for these comparisons, we found that using short wavelength photometric bands ($u, g, r$) introduced dependencies on SFH. Additionally, we decided to avoid the longest wavelength NIR bands ($H, K_s$) as in DEVILS these are shallower than the other NIR bands. We also found that using neighbouring bands provided little information on the metallicity, and the constraint on metallicity in SED fitting was driven by the overall slope of the optical-NIR SED. Because of this, we decided to use the $i - J$ colour as these bands are not adjacent, and not at too short or too long wavelengths.

To have the greatest constraint on metallicity from an SED fit we require high signal to noise photometry with low uncertainties ($\leq 0.1$ mag; see Robotham et al. 2020). At a given redshift the galaxies with the highest signal photometry will be the brightest and therefore most massive galaxies. If the photometry for a given galaxy has a larger associated uncertainty and therefore larger metallicity error, the scatter in the MZR will be artificially increased. In Figure 3 we show the scatter in the MZR as a function of the average $i - J$ colour error binned in stellar mass and redshift. We find that there is a correlation between scatter and average colour error for the highest mass bins, where a higher colour error is associated with larger scatter in the MZR moving from high to low masses. We do find that in each redshift bin, there is a point in stellar mass where additional uncertainty in the colour does not correspond to an increase of scatter. Based on these relationships, a cut in $\Delta$ MZR just below the flattening would be poorly motivated as it is apparent that the scatter has already reached the maximum value. We, therefore, use a cut at $\Delta$ MZR < 1.5
Figure 3. The scatter in the MZR as a function of average $i$-$J$ colour error in bins of stellar mass (colour bar) and redshift (different panels). We show the $\sigma_{\text{MZR}} < 1.5$ and $\sigma_{i-J}$ cuts as the dashed grey lines. The lower left region of each panel bounded by these cuts isolates the stellar mass regime where both the average $i$-$J$ colour error and scatter in the MZR are low. This indicates that the metallicity estimates from ProSper for these stellar masses at each redshift are reliable.

Figure 4. Stellar mass limits derived using (a) the largest change in the MZR scatter (circles), (b) the lowest stellar mass where the MZR scatter is less than 1.5 dex (squares), and (c) the lowest stellar mass where the $i$-$J$ error is less than 0.01 (triangles). The black line shows the smooth spline fit which is used to define the lower stellar mass limit for the D10 sample at each redshift.

which lies more than 0.5 dex below the flattening at all redshifts. The third technique uses the same relationship but uses a cut in colour error of $i - J < 0.01$ as in most redshift bins, this value occurs at masses above the flattening in scatter.

Figure 4 shows each of the stellar mass cuts from the three techniques as a function of the scale factor ($1/(1+z)$). Although these stellar mass limits were determined using three different methods, we find that they are in relatively good agreement in each redshift bin. To combine the three techniques together and determine a single stellar mass cut for the D10 sample we fit all the points with a smooth spline and use this to calculate a stellar mass cut at each redshift bin. We stress that this selection is unique to the combination and depth of photometry available as well as the quality and depth of the redshifts used. Even in the context of the DEVILS survey, this selection will be unique to the D10 field due to differences in imaging depths and photometric redshift quality between fields.

Although GAMA is a shallower survey, the limits of the photometry are not being pushed with the use of photometric redshifts. In the case of GAMA the imaging in the $r$-band is five magnitudes deeper than the spectroscopic completeness limit.\footnote{The photometry in the $Y$-band for DEVILS D10 is also five magnitudes below the spectroscopic completeness limit, however in the full D10 sample we push beyond the $5\sigma$ depth by using all available photometric redshifts.} For this reason, we use the spectroscopic completeness limit defined by Robotham et al. (2014) as the lower stellar mass range for GAMA in this work. This results in a sample of 50 147 galaxies. Additionally, the stellar mass cut derived for the D10 sample is above the completeness cut derived by Thorne et al. (2021) meaning that for both samples we can guarantee we are sampling all galaxy types present at a given stellar mass and are not biased to bright, star-forming galaxies.

Figure 5 shows the stellar mass distribution as a function of redshift of the GAMA and DEVILS samples. The stellar mass cut derived...
Figure 5. The stellar mass distribution as a function of redshift for the DEVILS (left) and GAMA (right) samples. In each panel the solid black line shows the mass selection function used in this work where, at each redshift, we select objects with masses above the black line. For DEVILS this black line is as per Figure 4 and for GAMA this is the mass completeness cut from Robotham et al. (2014).

Figure 6. The distribution of stellar masses (left), SFRs (middle), and metallicities (right) for DEVILS (top panels) and GAMA (bottom panels). In each panel we show the full sample (lightest colour), the sample after the mass cut has been applied (middle colour), and the sample of ProSpect-selected AGN (darkest colour).
above for DEVILS is shown as the black line on the left panel while the completeness cut from Robotham et al. (2014) for GAMA is shown as the black line on the right panel. In each case, we use only galaxies with stellar masses above this limit in our analysis.

2.4.1 AGN Selection

As our implementation of ProSpect includes a flexible AGN component we can use our SED fits alone to identify galaxies hosting an AGN. To identify and isolate AGN host galaxies we select objects that have $f_{\text{AGN}} > 0.1$ where $f_{\text{AGN}}$ is the fraction of flux contributed by the AGN component between 5-20 $\mu$m (as per Thorne et al. 2022). For DEVILS galaxies, to identify an AGN, we also require that the galaxy has $Y_{\text{mag}} < 21.2$ or was identified in the MIPS24 imaging and was therefore passed to the FIR photometry phase (as described in Section 2, also see Davies et al. 2021). By ensuring that these objects have attempted FIR measurements we have a better constraint on the AGN component than without. This requirement is not necessary for GAMA as the redshift completeness limit is much shallower than DEVILS, and the FIR imaging in the GAMA fields has sufficient resolution and depth that FIR photometry can be extracted for all optically detected sources. In Thorne et al. (2022) we show that the inclusion of an AGN component in the SED fitting process has no systematic impact on the recovered masses or metallicities for galaxies with $f_{\text{AGN}} > 0.1$ and should not impact these results.

Selecting all GAMA galaxies with $f_{\text{AGN}} > 0.1$ results in a sample of 14,882/50,147 galaxies identified as AGN. The additional FIR photometry requirement for DEVILS imposes a harsher cut and results in 5,916/23,784 galaxies identified as AGN. The resulting stellar mass, SFR, and metallicity distributions of the original sample, mass-selected sample, and sample of ProSpect-selected AGN are shown in Figure 6. Before the mass cut is applied, the DEVILS sample has a relatively uniform metallicity distribution with noticeable peaks at the values of the Bruzual & Charlot (2003) metallicity templates. However, applying the mass selection cut systematically removes more low-metallicity galaxies as is expected from the MZR. The same is true for GAMA. However, for both GAMA and DEVILS the distribution of metallicities are very similar for the mass-selected sample and AGN sample, however in GAMA we find that the AGN sample do not follow the same SFR-distribution as the mass-selected sample due to the lower SFRs associated with ProSpect-selected AGN.

3 MASS – METALLICITY RELATION

Bellstedt et al. (2021) previously demonstrated that the recovered MZR at $z \approx 0$ from SED-fitting using ProSpect was consistent with the MZR obtained through spectral measurements. We extend this work to a larger sample of galaxies spanning a larger redshift range, as shown in Figure 7. The GAMA and DEVILS galaxies are shown in two 2D histograms with GAMA shown in warm colours (lower colour bar) and DEVILS shown in cool colours (upper colour bar). We have removed SED-selected AGN from this Figure to ensure they are not contaminating the recovered MZR (the MZR derived for AGN host galaxies is discussed in Section 3.1) but do not make any other cuts at this stage.

The median trend of the MZR is calculated in stellar mass bins of width 0.25 dex for bins with at least 200 galaxies and is shown as the solid black line. We show the error on the median by perturbing the metallicity of each galaxy by sampling from a normal distribution centered on the best fit value and using the metallicity uncertainties as the standard deviation. We do this 100 times and show each iteration as a grey line. This method does not account for the uncertainty in stellar mass, but typically this is much smaller than the uncertainty in metallicity (the metallicity and stellar mass uncertainties are typically 0.4/0.2 dex and 0.1/0.05 dex for DEVILS/GAMA respectively). The 1σ scatter in the MZR is shown as the dashed lines.

As is well known in the literature (Lequeux et al. 1979; Tremonti et al. 2004), we find a correlation between mass and metallicity where higher mass objects have higher metallicities. Some previous work noted a flattening or saturation of metallicity at the highest stellar masses (Lara-López et al. 2013; Bellstedt et al. 2021), which we can see some evidence of in our lowest redshift bin. However, at higher redshifts, we do not see this as clearly as we are mostly limited to a ~1 dex range in stellar masses.

As mentioned in Section 2.4, the scatter in the recovered MZR increases at low stellar mass at all redshifts, but this scatter reduces at higher stellar masses, particularly at $z \approx 0$. This is seen consistently across redshift bins despite the fact that we are significantly limited in mass range at higher redshifts due to the imposed mass cut.

A clear artefact in these panels is the upper limit in the range of metallicities at $Z = 0.05$. This is particularly noticeable at $z > 0.3$ where the upper region of our 1σ range is at this limit. This hard limit is the highest metallicity template in the Bruzual & Charlot (2003) stellar population templates, and therefore the upper limit for the ZFinal parameter in ProSpect. Hence, our application of ProSpect is not sensitive to gas-phase metallicity values larger than $Z_{\text{gas}} = 0.05$. The pile-up of objects at the upper metallicity limit could represent cases where a higher final metallicity value would be preferred, or cases where a more rapid chemical enrichment needs to be accounted for to better produce realistic $Z_{\text{gas}}$ values. We show the impact of removing sources with best fitting metallicities of $Z_{\text{gas}} = 0.05$ as the dashed red line in Figure 7. In our three lowest redshift bins ($z < 0.296$), we see no change in the recovered median MZR when removing these objects. At these redshifts, the medians are driven by the sample of GAMA galaxies which do not hit the upper metallicity limit as frequently as the sample of DEVILS galaxies (see Figure 6). Removing galaxies at the upper limit results in an offset to lower metallicities which expands with increasing redshift. Despite this, the gradient and shape of the median MZR are consistent when these high metallicity sources are removed. However, we still expect that the majority of these sources truly have high metallicities and therefore the running median calculated including these sources is still a reasonable approximation for the whole sample. We present the maximal impact of these sources by showing the median when excluding them, however, we reiterate that this is an extreme solution.

To compare our results to previous measurements of the MZR, we include measurements made for GAMA galaxies by Lara-López et al. (2013) as well as results from Savaglio et al. (2005); Henry et al. (2013); Ly et al. (2016); Huang et al. (2019); Weldon et al. (2020); Gillman et al. (2021); Helton et al. (2022). Comparing to spectroscopically derived values is difficult due to significant biases and implicit selection effects associated with spectroscopic measurements. Spectroscopic measurements rely on subsets of emission lines, strong line parameters, and calibration methods which all vary between data sets. Two commonly used methods for deriving metallicities from spectroscopic data are strong line calibrations and electron temperature-based techniques. So-called $T_e$ methods are more directly linked to the physical processes governing ionised nebulae, and therefore more closely related with metallicity. However, the required emission lines are weak, especially in metal-rich objects. Strong-line derivations were proposed to measure abundances in faint, distant, and high metallicity galaxies where other techniques
Figure 7. The mass–metallicity relation at 1 Gyr intervals for \( z = 0 \) to \( z = 1 \) using the ProSpect-derived stellar masses and metallicities for galaxies with no significant AGN component from DEVILS and GAMA. We show the redshift range of each panel in the lower left. At each interval, the running median is indicated with a solid black line, and the 1\( \sigma \) range in the scatter is shown as the dashed lines. The dashed red line shows the running median re-calculated removing sources that have best-fitting metallicities equal to the highest metallicity in the Bruzual & Charlot (2003) templates. We also include the \( z = 0 \) median in each panel as the white line, and the inferred mass-metallicity relation from Bellstedt et al. (2021) in yellow. Where possible we have included other observational measurements at the relevant epochs as a comparison. These studies include those of Savaglio et al. (2005); Henry et al. (2013); Lara-López et al. (2013); Ly et al. (2016); Huang et al. (2019); Weldon et al. (2020); Gillman et al. (2021); Helton et al. (2022).

could not be used (Alloin et al. 1979; Pagel et al. 1979). Strong-line derivations rely on ratios of collisionally excited lines and Balmer lines which have been shown to have a dependence on metallicity. However, these techniques must be calibrated either empirically, for samples in which \( T_e \) based abundances have been derived (Pettini & Pagel 2004; Pilyugin & Thuan 2005; Pilyugin et al. 2012, 2010; Marino et al. 2013; Pilyugin & Grebel 2016), or theoretically, using oxygen abundances that have been inferred via photoionisation models (e.g. Kewley & Dopita 2002; Kobulnicky & Kewley 2004; Tremonti et al. 2004; Dopita et al. 2013, 2016; Curti et al. 2017). Unfortunately, comparisons of metallicities estimated using different techniques and calibrations are highly discrepant, even for the same sample of galaxies. Kewley & Ellison (2008) demonstrated that by using different calibrators, the resulting MZR can vary in normalization by up to 0.5 dex but also have dramatically varying slopes. Additionally, spectroscopic metallicities are measured using different element bases – stellar metallicities are often measured using iron, while gas phase metallicities are measured using oxygen (see Fraser-McKelvie et al. 2022 for a comparison of stellar and gas metallicities). This differs to the metallicity value from ProSpect, which represents the mass in all elements heavier than hydrogen or helium, and is not centered on a particular element. The metallicity indicators and calibrations used for each of these literature measurements are outlined in table 2 of Bellstedt et al. (2021).

Additionally, spectroscopic measurements are derived from oxygen abundances (12 + log(O/H)) and require conversions to metallicity values using assumed oxygen and solar abundance values. To convert literature oxygen abundances to metallicities for comparison, we use the prescription from Kobulnicky & Kewley (2004)\(^3\):

\[
Z = 29 \times 10^{[12+\log(O/H)]-12}.
\]  
(1)

However, additional uncertainty is associated with these conversions as the oxygen abundance and metallicity of the Sun are still uncertain. Changes in both the assumed oxygen abundance and metal mass fraction of the Sun will change the normalisation of the conversion. Consequently, scatter between observed and our modelled metallicities could arise from the adopted scaling between oxygen abundances and total metallicities.

Additionally, it is very difficult to measure gas-phase metallicities from spectra for galaxies with low-to-no star formation as the required

\(^3\) This is equivalent to \( Z_{\text{gas}} = Z_\odot \times 10^{[12+\log(O/H)]-12+\log(O/H)_\odot} \), where \( Z_\odot = 0.0142 \) and \( [12 + \log(O/H)]_\odot = 8.69 \) (Asplund et al. 2009).
emission lines are not detectable. Metallicities are also difficult to measure for AGN host galaxies as the associated broad and narrow emission lines can contaminate measurements of lines required for metallicity estimates. Therefore, the previous measurements of the MZR using emission line techniques are predominantly based on star-forming, galaxies with no AGN component.

At \( z < 0.3 \), we recover an MZR systematically lower than that derived by Lara-López et al. (2013). In our highest redshift bins (\( z = 0.5 - 1.0 \)), we recover a mass-metallicity relation in close agreement with the MZR derived by Huang et al. (2019). At these redshifts, our derived MZR is also consistent with the individual galaxy measurements from Savaglio et al. (2005) and Helton et al. (2022). However, we measure more low-metallicity values than the observations shown in Figure 7. This could be due to a number of factors. Recovering metallicity estimates through broad-band SED fitting allows us to work with larger samples and potentially trace the lower tail of metallicity values. Additionally, using SED-derived metallicities allows us to probe different areas of the SFR-\( M_\star \) parameter space where the emission lines might not be detectable. Alternatively, emission line surveys may not be sensitive to low metallicities, or we may be recovering artificially low values in some cases.

We also compare to results from Bellstedt et al. (2021) who used ProSpect to trace back in time the inferred MZR from the star formation and metallicity histories of a sample of \( \sim 4,500 \) galaxies with \( z < 0.06 \). We refer to this as a forensic determination of the MZR and show the forensic MZR recovered in each redshift bin as the yellow line in each panel. The masses and metallicities from Bellstedt et al. (2021) were derived using a very similar implementation of ProSpect as used in this work, with the major difference being the inclusion of the AGN model in this work and the use of a different optimization routine (Appendix C explores the impact of some of these differences). Bellstedt et al. (2021) also used GAMA galaxies for their analysis, but limited their analysis to 4,500 galaxies with \( z < 0.06 \) and secure metallicity measurements in the G09, G12, and G15 fields (not including G23, which is included in this work). Interestingly, we find good agreement with the backwards modelled results from Bellstedt et al. (2021) at higher redshifts (\( z > 0.3 \)) where our results are determined from the DEVILS galaxies while the Bellstedt et al. (2021) results are extrapolated backwards from GAMA galaxies at \( z \approx 0 \).

Despite the fact that the median MZR at \( z = 0 \) from Bellstedt et al. (2021) is derived using the same SED fitting code on the same parent sample of galaxies, there is an offset in the recovered MZR at \( z = 0 \). There are a number of differences between the implementations that could cause the offset including the difference in sample selection (due to field choice, metallicity uncertainty cuts, and redshift range used), inclusion of the AGN component in SED fitting, choice of optimization routines and selection of parameter values. See Appendix C for a demonstration of the impact of each implementation difference. Even after accounting for these differences, there are other reasons why we might expect the forensic MZR to differ from our measured MZR at higher redshifts. First, the forensic technique cannot take into account any merger activity, inflows or outflows which might change the metallicity after the final stars have formed. Second, any agreement with the forensic technique is dependent on the implemented enrichment model being correct – if it were not, we would expect a larger discrepancy. Additionally, Bellstedt et al. (2020b) found rough agreement between the measured cosmic metal density using these forensic metallicity histories from ProSpect with those from observations suggesting that the evolving metallicity implementation in ProSpect does a reasonable job modelling the true chemical enrichment of galaxies.

Typically the MZR is measured for star-forming galaxies, because gas-phase metallicity is typically measured from nebular emission lines. By extracting metallicities via SED fitting, we are not limited to star-forming objects and as such include passive systems in the MZR presented in Figure 7. However, the concept of a gas-phase metallicity is unclear for a galaxy with minimal gas as is typical for high mass, passive galaxies. To test the impact of including passive galaxies, we trialled removing objects with a specific SFR (SSFR) \( \text{SSFR} = \frac{SFR}{M_\star} < 10^{-13} \text{yr}^{-1} \) as motivated by the cut in Thorne et al. (2021). This cut had no impact on the median MZR, but slightly (\( < 0.05 \) dex) reduced the scatter in the highest mass bin of each redshift.

Figure 8 shows the evolution of the median MZR in the redshift bins from Figure 7. We calculate and show the MZR using just the GAMA sample (left), DEVILS sample (center), and the combination of the two (right) and show mass bins with less than 200 galaxies as the dashed lines. When examining the results from GAMA and DEVILS separately, there is a slight evolution present where, at fixed stellar mass, the normalisation of the MZR decreases with increasing redshift. When combining the two samples together however, there is a discontinuity in the evolution where we move from having GAMA and DEVILS to redshift bins with only DEVILS. Overall, this evolution is less than 0.1 dex in the normalisation at \( M_\star = 10^{10.75} M_\odot \) over the last 8 Gyrs. The discontinuity between redshift bins demonstrates the potential impact of sample selection and highlights the importance of using cohesive data sets when exploring trends with redshift. Limited evolution in the MZR at high stellar masses (\( M_\star > 10^{10} M_\odot \)) since \( z = 1 \) is also found by Bellstedt et al. (2021). They find that most of the evolution in the MZR at \( M_\star > 10^{10} M_\odot \) occurs in the first ~ 5 Gyrs of cosmic time. These results are also supported by Cresci et al. (2012) who find no evolution in the MZR over the last 7 billion years, however Lara-López et al. (2010) find a ~ 0.2 dex increase in metallicity from \( z = 0.35 \) to \( z = 0.07 \).

### 3.1 Mass – Metallicity Relation for AGN Host Galaxies

Typically, the MZR has been measured only for star-forming galaxies with no AGN. This is primarily because the nebular line metallicity diagnostics have mostly been calibrated on star-forming galaxies classified as such based on the Baldwin, Phillips & Terlevich (1981) diagrams. These diagrams are a useful way to broadly divide galaxies by the excitation mechanism of the ISM into those with ongoing star formation, and the population of non-star-forming galaxies including LINERs and AGN. Extending the MZR to galaxies classified as LINERs and AGN is difficult due to a lack of suitable metallicity calibrators for estimating the gas-phase metallicity of such systems.

The benefit of measuring galaxy metallicities through SED fitting is that we are not limited by existing metallicity calibrators or the presence of emission lines produced by star formation, allowing us to not only extend the MZR relation to quiescent or quenched galaxies, but also to galaxies with a significant AGN component.

Figure 9 shows the recovered MZR for AGN galaxies. Due to the AGN selection criteria, we have significantly fewer galaxies at higher redshifts (\( z > 0.3 \)) and so we limit the redshift range here to \( z < 0.3 \). For comparison we show the median MZR for non-AGN galaxies as the black line. We find that at \( z = 0 \) the recovered MZRs for non-AGN and AGN galaxies are practically indistinguishable for \( M_\star = 10^{10} - 10^{11} M_\odot \). Between \( z = 0.2 \) and \( z = 0.3 \) we find that for galaxies with \( M_\star > 10^{10} M_\odot \) galaxies that host an AGN have marginally higher metallicities than galaxies with no AGN. Below these masses we do not have a large enough sample to comment on the differences between selections.

These results are in agreement with Cai et al. (2020) who found...
Figure 8. The evolution of the median MZR computed using only GAMA (left), only DEVILS (center), and the combination of GAMA and DEVILS (right). Each line shows a different redshift bin from Figure 7 and is shown as a solid (dashed) line if the mass bin contains more (less) than 200 galaxies. The polygons show the error on the median. In the inset panels, we show a zoom of the median relations without the polygons for clarity.

Figure 9. The mass-metallicity relation for galaxies with a significant AGN component ($f_{\text{AGN}} > 0.1$) at 1 Gyr intervals for $z = 0$ to $z = 0.3$. The running median and 1σ range is shown in solid and dashed blue respectively. We also show the error on the median as the light blue shaded region. We show the median MZR and error on the median from the equivalent panels of Figure 7 (non AGN only) as the solid black line and grey shaded region for comparison.

that 136 dwarf galaxies with BPT-selected AGN follow a similar mass-metallicity relation to normal star-forming galaxies, indicating that AGN have little impact on the chemical evolution of the host galaxy. Netzer & Trakhtenbrot (2007) also find significant correlation between the normalised accretion rate ($L/L_{\text{edd}}$) and the Fe II / Hβ line ratio. It is also well known that most AGN exist in galaxies with solar or supersolar metallicities (Storchi-Bergmann et al. 1998; Hamann et al. 2002), although due to the high masses of AGN host galaxies this is not surprising.

3.2 Comparison with Simulations

Historically, the MZR has been very difficult for simulations and semi-analytic models (SAMs) to reproduce. Recently, there has been an increased reporting of theoretical models producing MZR trends similar to observations at $z = 0$. In Figure 10 we compare our derived MZR at $z = 0$ and $z = 1$ to those produced by leading simulations/SAMs. We include the MZRs derived by cosmological hydrodynamic simulations Illustris, IllustrisTNG, MUFASA, SIMBA, EAGLE and the semi-analytic models GAEA and SHARK.

Using Illustris, Torrey et al. (2014) demonstrate that the adopted

4 Directly comparing observed metallicities to simulated metallicities is difficult as an assumption has to be made about stellar atmospheres either to turn simulations into observables (generate SEDs/spectra) or to invert measured SEDs/spectra to recover estimates for the metallicity. Figure 30 of Robotham et al. (2020) demonstrates that ProSpect can recover input stellar masses and metallicities for simulated galaxies. Additionally, Bellstedt et al. (2021, appendix A) demonstrate the validity of proportionally evolving metallicity histories as implemented in ProSpect using simulated galaxies.
feedback models (both stellar and AGN) have a dramatic influence on the resulting MZR at all redshifts with no feedback resulting in much higher metallicities, whilst a strong wind implementation reduces the normalisation of the MZR. Note, that in the fiducial Illustris simulation, the normalisation of the MZR was used to set the metal content of ejected wind material. At both $z = 0$ and $z = 1$, the Illustris simulations do not find a flattening of the MZR at high stellar masses when including feedback, but the no-feedback run does recover this flattening. However, the no-feedback run of Illustris recovers far too many galaxies at all stellar masses, and a higher cosmic star formation rate density at all times. This suggests that the flattening at high stellar masses is driven by a saturation in metallicity through high SFRs and not driven by metal loss.

Unlike Illustris, where little bending was observed, the MZR recovered from IllustrisTNG (Torrey et al. 2019) does recover a saturation in metallicity at $z = 0$. However, both the stellar mass and metallicity where this bending occurs is much lower than that recovered by ProSpect. At $z = 1$, the MZR recovered from IllustrisTNG has a negative slope for $M_\star > 10^{10.5} M_\odot$ which is not found in our results or other spectroscopic observations.

SIMBA, and its predecessor MUFASA, recover systematically lower metallicities at all stellar masses than our results. Additionally, the shape of the MZR from SIMBA has a different shape to those from other simulations. A dip in metallicity can be seen at $M_\star \sim 10^{10} M_\odot$ which is not evident in MUFASA. The normalisation and shape of the MZR recovered from MUFASA are closer to our derived MZR than SIMBA, despite the fact that Davé et al. (2019) describe the MZR from MUFASA as being too steep.

We also include the MZR from the recalibrated EAGLE simulations presented in De Rossi et al. (2017). On average, EAGLE
galaxies have slightly higher metallicities than other simulations but are in reasonable agreement with our results at \( M_\star > 10^{10} M_\odot \). However, EAGLE recovers a flatter MZR with little increase in metallicity with increasing stellar mass.

In addition to results from cosmological hydrodynamic simulations, Figure 10 also shows the MZR derived from the GAEA (De Lucia et al. 2020) and SHARK (Lagos et al. 2018) semi-analytic models. Of all the simulation results shown, the GAEA MZR agrees closest with our results in both normalisation and shape at \( z = 0 \) over the full mass range covered by GAMA and the more sparsely sampled DEVILS mass range (dashed black line). The GAEA MZR recovers the expected lower metallicities for low-mass galaxies and a saturation of metallicity for high-mass galaxies. De Lucia et al. (2020) describe the predicted scatter as larger than observed particularly for galaxies with \( M_\star < 10^{10} M_\odot \), however their scatter is smaller than the \( 1 \sigma \) range that we derive. At \( z = 1 \) we show the MZR recovered from GAEA and presented in Hirschmann et al. (2016). It matches the normalisation of our MZR at \( M_\star = 10^{11} M_\odot \) and flattens at the same metallicity, but this flattening occurs at higher stellar masses than we find.

The MZR from SHARK is relatively consistent with our results for \( 10^9 < M_\star / M_\odot < 10^{10} \) but produces significantly higher metallicities for \( M_\star > 10^{10.5} M_\odot \) at both \( z = 0 \) and \( z = 1 \).

### 3.2.1 The impact of AGN feedback on the MZR

One of the physical processes often used to explain the shape of the MZR is galactic winds and feedback removing metals from galaxy potential wells (Tremonti et al. 2004; Kobayashi et al. 2007). For galaxies with low stellar masses, supernova explosions are effective at ejecting metals from the ISM via outflows (e.g. Kobayashi et al. 2007; Blanc et al. 2019). Although dying stars produce metals, the force of a supernova explosion can be sufficient to eject this enriched gas from the ISM, especially for lower mass galaxies with a lower gravitational potential (see D’Eugenio et al. 2018 for a discussion on the relationship between metallicity and gravitational potential). Galaxies with a higher gravitational potential might be less affected by supernova feedback, but they are more likely to host AGN that can also remove gas from a galaxy through mechanical feedback (Fabian 2012). Mechanical feedback occurs when most of the energy of an AGN is released in kinetic form via radio-emitting jets, which can expel large amounts of gas from galaxies (Morganti 2017). This form of AGN feedback can readily eject enriched material but can also prevent accretion onto a galaxy, further impacting the metallicity of the ISM (Silk & Rees 1998). These forms of AGN feedback are also implemented in simulations in order to match the bright end of the luminosity function, while stellar feedback drives the agreement at the faint end. However, the physical mechanisms and impact of AGN feedback on host galaxies are still poorly understood (Shangguan et al. 2020; Laha et al. 2021; Koudmani et al. 2022; Nobels et al. 2022).

Altering the physics implemented in theoretical simulations can allow us to understand the impact of AGN feedback on resulting galaxy properties. Taylor & Kobayashi (2015) find that including AGN feedback in their simulations has no impact on the derived stellar or gas-phase mass-metallicity relations. They attribute this to the fact that AGN feedback quenches star formation most efficiently at low redshift, after the peak of star formation and chemical enrichment. However, in the context of the EAGLE simulations, De Rossi et al. (2017) find that the flattening of the MZR at \( M_\star > 10^{10} M_\odot \) is mainly regulated by AGN feedback. By comparing a run with no AGN feedback to one with AGN feedback, while keeping the resolution and box size the same, they find a metallicity offset of \( \approx 0.3 \) dex at \( M_\star \sim 10^{11} M_\odot \). They comment that, despite the fact that galaxies with evidence of current AGN activity are removed from observed MZR studies, black hole feedback may have previously occurred influencing the chemical enrichment of high mass galaxies. Using three hydrodynamical simulations of a single galaxy, Eisenreich et al. (2017) find that mechanical AGN feedback flattens the metallicity gradient across the galaxy, but the inclusion of radiation feedback decreases the overall metallicity.

Our results suggest little impact on the MZR from galaxies with current signatures of AGN. However, as highlighted by De Rossi et al. (2017), black hole activity could occur in cyclic episodes and galaxies which do not currently display signs of activity could have been impacted by past episodes.

### 4 THE INFLUENCE OF STAR FORMATION HISTORY ON THE MZR

In this section, we explore additional higher dimensional correlations with the MZR to provide better constraints on the processes that regulate metallicity. Due to the nature of our SED fitting technique, we can unpack not only the relationship with current SFR or sSFR, but also with the overall star formation history.

For this section, we limit our sample to just galaxies from the GAMA survey at \( z < 0.3 \) to maximize the number of galaxies in each stellar mass and metallicity bin. Additionally, the uncertainties on the stellar masses, metallicities, and SFRs estimates are lower for GAMA galaxies than for DEVILS galaxies. We also remove galaxies with \( f_{\text{AGN}} > 0.1 \) to limit to only galaxies with no sign of current AGN activity.

#### 4.1 Star Formation Rate

Higher-order correlations of the MZR are most commonly explored using the SFR or sSFR as an indicator of the star formation history of the galaxy. The existence or not of a correlation of the MZR with SFR is still under debate (e.g. Sánchez et al. 2013, 2017, 2019; Salim et al. 2014; Barrera-Ballesteros et al. 2017; Cresci et al. 2019). Previous work by Mannucci et al. (2010) also showed an inverse correlation between SFR and metallicity at a given stellar mass for objects with \( M_\star < 10^{11} M_\odot \) (see also Dayal et al. 2013; Curti et al. 2020). Salim et al. (2014) also found an anti-correlation between metallicity and sSFR but only for \( M_\star < 10^{10.5} M_\odot \). Above this mass, they find no correlation between metallicity and sSFR. Similarly, at higher redshifts (\( z \sim 1.6 \)), Zahid et al. (2014) find an anti-correlation between metallicity and SFR for galaxies at fixed stellar mass (see also Stott et al. 2013; Kashino et al. 2017, 2019; Cresci et al. 2019). This is opposed to the findings of Lara-López et al. (2013) who find that at higher stellar masses (\( M_\star > 10^{10} M_\odot \)) a higher SFR/sSFR is associated with a higher metallicity at a given stellar mass. However, Sánchez et al. (2013, 2017, 2019) find that there is no dependence of the MZR on SFR or sSFR and find no evidence for a SFR-M_\star-Z fundamental plane.

Figure 11 shows the recovered MZR when binning by SFR, sSFR, and by location relative to the star-formation main sequence (ASFMR, Noeske et al. 2007; Speagle et al. 2014) for the GAMA sample. In each case, we only show bins with more than 20 galaxies.

When considering the dependence of the MZR on SFR, we recover similar relations across all SFRs, especially at \( z = 0 \). At \( z = 0 \) and \( M_\star \approx 10^{11} M_\odot \), there is evidence of a trend within the highest
Figure 11. The dependence of the MZR on SFR (top), sSFR (middle) and location relative to the star-formation main sequence (bottom). In all panels we show the median $Z_{\text{gas}}$ in $M_\star$ bins for each bin in SFR or sSFR, or the location relative to the main-sequence. The underlying 2D histogram shows the total distribution of galaxies in each redshift bin.

SFR bins, where a higher SFR is associated with lower metallicity. However, at all stellar masses, the range in median metallicities is $\approx 0.13$ dex while the SFR bins span 2.5 dex. At lower masses ($M_\star \approx 10^{10} M_\odot$), the trend in metallicity is reversed for galaxies with relatively high star formation rates (blue, teal, and green lines), where a higher metallicity is associated with a higher star formation rate. In the higher redshift bins, it is also evident that the highest star formation rates are offset to lower metallicity values, however the spread in median metallicities is greater than at $z = 0$. At all redshifts and mass ranges explored here, we do not find significant offsets to higher metallicities for galaxies with very low SFRs (yellow, orange, and red lines) which might be expected if a galaxy was undergoing strangulation (e.g. the stellar metallicity results from Peng et al. 2015).

We also show the MZR binned in sSFR intervals to remove mass dependence in each bin. There are similar trends between sSFR and metallicity as described above with SFR, where higher sSFRs are associated with lower metallicities. This is most extreme in our highest redshift bin where the spread in median metallicity is $\approx 0.2$ dex. These trends are in agreement with the findings of Mannucci et al. (2010), but opposite to the findings of Lara-López et al. (2013).

We also present the MZR recovered when binned in distance from the star-formation main sequence in Figure 11. The star-formation main sequence (MS) refers to the tight relation between stellar mass and star formation rate which holds over more than five orders of magnitude in stellar mass and has been shown to hold out to $z \approx 5$ (Thorne et al. 2021). The reason for binning by distance from the MS is to isolate galaxies that have similar relative star formation activity across the full range of stellar masses (e.g. starburst, main-sequence, green-valley, or passive galaxies). By binning in distance from the
Figure 12. An example SFH with the descriptive parameters from Section 4.2 labelled. Note that the Δ log SFR200 parameter is defined to be the difference of the log10 SFR, but the SFR is shown here in linear space for ease. The $f_{\text{SFH early}}$ is defined as the integral of the orange shaded region divided by the integral of the whole SFH, $t_{50}$ is the half mass age or the age at which the galaxy has formed half of its total stellar mass and the FWHM is the width of the SFH in Gyrs at half of the maximum SFR.

To calculate the distance from the MS for each galaxy in our sample we use the redshift dependent MS fit presented in Thorne et al. (2021). The Thorne et al. (2021) MS was derived using the same parent sample as this work and used a double power law to model the flattening of the MS at higher stellar masses. For each galaxy, we calculate the vertical distance in dex from the MS ($\Delta SFR_{MS}$), where positive (negative) values represent galaxies that lie above (below) the main sequence at their given stellar mass. Binning by $\Delta SFR_{MS}$ results in similar trends across the MZR as with SFR and sSFR where high star formation activity correlates with lower metallicities at high masses ($M_\star > 10^{10.5} M_\odot$) at all redshifts. However any trend between $\Delta SFR_{MS}$ and metallicity at lower masses ($M_\star < 10^{10.5} M_\odot$) can be excluded at a $\sim 0.1$ dex level, but not at levels below 0.1 dex.

Overall, we find evidence of an anti-correlation between SFR, sSFR, and $\Delta SFR_{MS}$ with $Z_{\text{gas}}$ at $M_\star > 10^{11} M_\odot$ across all redshifts explored. Of the three parameters, the association between $\Delta SFR_{MS}$ and metallicity is clearest especially in our two highest redshift bins. However, in some cases the range in median metallicities across different star formation bins is only $\sim 0.1$ dex which could be within errors from total or bin selection effects, or could be within uncertainties due to the SED fitting technique. Furthermore, due to the nature of SED fitting the recovered SFRs and metallicities are more intertwined than when measured from emission lines which could impact the direction and strength of the trends recovered.

However, these trends are consistent with results from Salim et al. (2014), who find an anti-correlation between metallicity and SFR, sSFR, and $\Delta SFR_{MS}$ at fixed stellar mass using a sample of low redshift galaxies from the Sloan Digital Sky Survey (SDSS; Abazajian et al. 2004, 2009).

4.2 Star Formation History

Although the SFR of galaxies provides insight into the current star formation activity, it is the overall star formation history (SFH) and chemical evolution history of the galaxy that influences the observed metallicity. In this section we use the SFHs derived using ProSPECT for the sample of GAMA galaxies to explore higher-order correlations with the MZR. As luminous AGN components can produce emission very similar to that generated by star formation, we remove galaxies with $f_{\text{AGN}} > 0.1$ to ensure that the star formation histories recovered by ProSPECT are not biased by this degeneracy.

Although we parameterise the SFH within ProSPECT using four parameters to describe the skewed Normal shape, these can often be degenerate and produce similar SFHs despite having very different
values. Because of this we define and use four different parameters that can better describe the features of a galaxy’s SFH without being degenerate. These are:

- \( t_{50} \) - the half mass age, i.e. the age at which the galaxy has formed half of its total stellar mass (Gyrs)
- FWHM - the full width half maximum, defined to be the width of the SFH at half the height of the maximum SFR (Gyrs)
- \( f_{\text{mass early}} \) - the fraction of mass formed before the peak of star formation (e.g. \( f_{\text{mass early}} = 0.5 \) corresponds to a galaxy with a perfectly Normal SFH where half the mass is formed before the peak, and half the mass is formed after the peak.)
- \( \Delta \log \text{SFR}_{200} = \log(\text{SFR}_{0\text{Myrs}}) - \log(\text{SFR}_{200\text{Myrs}}) \), the change in SFR in dex over the last 200 Myrs. Negative values correspond to declining SFHs and larger absolute values correspond to a steeper change over the last 200 Myrs.

For clarity we show an example SFH in Figure 12 with each of the descriptive parameters labelled. Figure 13 shows the distribution of each of these descriptive parameters as a function of redshift. We show only GAMA objects above the mass completeness limit and with \( f_{\text{AGN}} < 0.1 \). As the \( t_{50} \) and FWHM parameters are dependent on the age of the Universe at the redshift of each object, the upper right corner of both parameter spaces is not a possible combination (indicated by the grey shading).

Figure 14 shows the MZR where each cell is coloured by the median value of each of these parameters. We also show the MZR coloured by the median SFR for comparison. As discussed in Section 4.1, at a given stellar mass there is no clear trend of SFR with metallicity at \( z \approx 0 \). There is a slight trend with SFR at \( z \approx 0.25 \) where higher star formation rates are associated with a lower metallicity at \( M_* = 10^{11} M_\odot \). However, the most apparent trend is a manifestation of the star-forming main sequence (see e.g. Noeske et al. 2007; Lee et al. 2015; Thorne et al. 2021) where higher masses are associated with higher star formation rates, except at the highest stellar masses where we see a significant decrease in average SFR due to a large population of quenched/passive galaxies.

When considering the dependence on half mass age (\( t_{50} \)), it is immediately evident that we are recovering downsizing in our results through the trend between stellar mass and age, where higher mass galaxies have older stellar populations than low mass galaxies. However, at lower masses (\( M_* \sim 10^{10} M_\odot \)) we also find that galaxies with a higher metallicity at a given stellar mass are younger.

The third row of Figure 14 colours the MZR by the median full-width half maximum of the SFH in a given mass and metallicity bin. This parameter is a proxy for how quickly the majority of star formation happened in a galaxy, where smaller FWHM values represent quicker star formation episodes. As with the median age, these panels are in agreement with cosmic downsizing where higher mass galaxies formed earlier with shorter star formation timescales while lower mass galaxies formed later on longer timescales. At \( z = 0 \), there is also a slight vertical trend at the lowest stellar masses examined here (\( M_* < 10^{10} M_\odot \)), where lower metallicities are associated with longer star formation timescales.

The fourth row of Figure 14 shows each cell coloured by the median fraction of mass formed before the peak of star formation (\( f_{\text{SFH early}} \)). A higher value corresponds to a galaxy with a slow increase in star formation in the early Universe, while a value close to zero corresponds to a galaxy in which star formation ramped up very quickly. We find clear evidence that objects with low masses (\( M_* \sim 10^{10} M_\odot \)) and high metallicities formed a larger fraction of mass before the peak of star formation than lower metallicity galaxies at the same stellar mass. At high stellar masses (\( M_* > 10^{11} M_\odot \)), it is clear that, on average, the fraction of mass formed before the maximum star formation rate is low, consistent with a picture where these galaxies rapidly ramped up star formation in the early Universe, and formed quickly over short timescales.

The final row shows the MZR coloured by the median change in SFR in dex of the galaxy over the last 200 Myrs (\( \Delta \log \text{SFR}_{200} \)). For the GAMA sample, most galaxies are experiencing declining SFHs over the last 200 Myrs consistent with the overall star formation rate density trend. We find that higher mass galaxies are, on average, experiencing faster declines in SFR than low mass galaxies. However, at low masses we see that lower-metallicity galaxies have flatter recent SFHs than higher-metallicity galaxies at the same stellar mass which are declining at a faster rate.

Figure 15 shows the range of SFHs found in different locations on the MZR at \( z = 0 \). We show each individual SFH for galaxies in each region as the coloured lines, normalised to the same total stellar mass to highlight similarities/differences in shape. Each line is coloured by the galaxy’s vertical distance from the MS (\( \Delta \text{SFR}_{\text{MS}} \)) as per Section 4.1 where purple represents starburst galaxies while red represents passive/quenched galaxies.

For each region, we show the median SFH, which is calculated by taking the median inferred SFR of all galaxies at each lookback time (dashed black line). As these median SFHs are independent at each lookback time they are not forced to take the form of a skewed-Normal. We also show a toy-model SFH (solid black line) generated by transforming the median parameters in the left column of Figure 14 (\( t_{50} \), FWHM, \( f_{\text{mass early}} \), and \( \Delta \log \text{SFR} \)) to the required inputs of the massfunc_snorm_trunc. These are generated with the following parameters:

(a) \( m_{\text{peak}} = 3 \), \( m_{\text{period}} = 3 \), \( m_{\text{skew}} = -0.3 \)
(b) \( m_{\text{peak}} = 8 \), \( m_{\text{period}} = 2 \), \( m_{\text{skew}} = 0.3 \)
(c) \( m_{\text{peak}} = 11 \), \( m_{\text{period}} = 1.5 \), \( m_{\text{skew}} = 0.3 \)
(d) \( m_{\text{peak}} = 4 \), \( m_{\text{period}} = 3 \), \( m_{\text{skew}} = 0.5 \)
(e) \( m_{\text{peak}} = 11 \), \( m_{\text{period}} = 3 \), \( m_{\text{skew}} = 0.5 \)
(f) \( m_{\text{peak}} = 8.5 \), \( m_{\text{period}} = 2 \), \( m_{\text{skew}} = 0.5 \)

The regions were selected to examine the SFHs in ’extreme’ locations on the MZR to explore how changes in SFH can explain the scatter in the MZR. In each region there is a diversity in individual SFHs and cases where the individual SFHs are very different from the toy model. This is to be expected as the toy model is based on the median parameter distribution and there will always be outlier SFHs. Additionally, there is nothing built into ProSpect to link a particular SFH to a final metallicity value, however some SFH and \( Z_{\text{gas}} \) combinations could be the consequence of degeneracies, which we discuss below.

In the low mass, intermediate metallicity (panel d) and intermediate mass, high metallicity (panel b) regions, we find that two distinct shapes of SFH are equally prevalent - an old, and a young one. These two populations correlate well with the current position of each galaxy on the MS where the older population are currently passive while the younger population are still actively forming stars. Meanwhile, in the high mass, high metallicity case (panel c) we recover SFH shapes that are consistently similar to the toy model.

The low mass, high metallicity region (a) also shows a diversity in SFHs with two dominant populations, the population that formed early and have since quenched, and the, dominant, population of galaxies that have formed less than 4 Gyrs ago. When comparing to the other high metallicity regions at both intermediate and high masses, there are significantly fewer galaxies that formed recently. If this population of low mass, high metallicity galaxies were actually older and more metal-poor, then due to the age-metallicity degener-
Figure 14. The MZR in three redshift bins using only the GAMA sample coloured by the median (first row) SFR, (second row) half mass age, (third row) the width of the star formation history at half the height of the peak, (fourth row) the fraction of the total mass formed before the peak of star formation, and (fifth row) the change in the SFR over the last 200 Myrs in dex.
In the main panel we show the MZR at $z = 0$ using the GAMA sample as the grey 2D histogram. We also show the running median and 1σ range as the solid and dashed lines respectively. We highlight six regions of the MZR as the black rectangles and in the surrounding panels show the SFHs of all galaxies within that region of the MZR. The colouring of each SFH corresponds to each galaxy’s position relative to the star-formation main sequence from Thorne et al. (2021). We show the $z = 0$ main-sequence fit in the small inset panel as the solid black line and show each of the $\Delta \text{SFR}_{\text{MS}}$ bins as the shaded regions where purple corresponds to starburst galaxies while red represents passive galaxies. In each of the SFH panels, the dashed line shows the median inferred SFR of all SFHs at each age step. The thick black line in each panel shows a toy model SFH informed by the distribution of parameters shown in Figure 14. As we are only concerned about the shape and not the normalisation of the SFH we do not show values on the SFR axis and every SFH in each box is normalised to the same total mass value.
acy it is possible for ProSpect to falsely characterise these galaxies as young and metal rich. In this scenario, the galaxies should actually be in the intermediate metallicity, low mass region (panel d) of the MZR (a more typically populated region), whereas instead they have been placed in the low mass, high metallicity region (panel a). Although it is unclear the extent to which this is occurring, we can use the genbox function included in ProSpect to explore potential situations in which low mass, high metallicity galaxies might exist. The genbox function allows the user to specify star formation functions, inflow and outflow rates, and metallicity arguments to explore the evolution of gas and stellar mass, and metallicity of a system over time. Through investigation using these models, it is clear that in order to form a galaxy with a very high $Z_{\text{gas}} (Z \sim 0.05)$ at any stellar mass, the galaxy has to reach low gas fractions ($M_{\text{ISM}}/M_* < 0.1$). As the gas fraction of a galaxy decreases, supernova enrichment is more efficient at increasing the metallicity of the ISM due to the fact that there is much less gas to enrich. Observational measurements of the gas fraction of $z \approx 0$ galaxies in the xGASS sample (Catinella et al. 2018) shows that there are populations of galaxies with $M_* \sim 10^{10} M_\odot$ with the required gas fractions to reach high gas-phase metallicities as recovered in this work. Galaxies at $M_* \sim 10^{10} M_\odot$ with high metallicities ($Z_{\text{gas}} > 0.04$) are also found in large spectral line samples (e.g. Tremonti et al. 2004; Lara-López et al. 2010) and therefore the recovery of these properties in our work is not necessarily an artefact of measuring metallicities through SED fitting. The consequence of requiring low gas fractions in galaxies is that there is very little gas left from which to form stars and would suggest that these galaxies are likely quenching and therefore have low current SFRs. The combination of these arguments suggests that the reason behind the existence of two populations in region (a) is that there are galaxies with physical SFHs and metallicities (the older, passive galaxies), and galaxies that are likely an artefact the age-metallicity degeneracy (the younger, star forming galaxies).

Despite these caveats, recent results suggest that these trends in SFH, and therefore gas mass, across the MZR are physical. For example, Zhou et al. (2022) use a semi-analytic spectral fitting technique to recover the star formation history and chemical evolution history of galaxies in the MANGA sample. They find that both the SFHs and chemical evolution histories have strong mass dependence whereby massive galaxies accumulate their stellar masses and become enriched earlier. Their use of spectroscopic data means their results are less likely to be biased by the age-dust-metallicity degeneracy. Additionally, Brown et al. (2018) and Chen et al. (2022) find anti-correlations between H i gas mass and metallicity at a fixed stellar mass where at a fixed stellar mass, a higher metallicity corresponds to a lower gas mass (see also Bothwell et al. 2013; Hughes et al. 2013).

5 CONCLUSION

It has been known for decades that broadband optical-NIR colors can be used to separate the age and metallicity of composite populations. Despite this, metallicity is still treated as a nuisance parameter in most SED fitting codes. Almost all SED fitting codes assume that metallicity is constant over the lifetime of a galaxy, where the value is either set to the solar value, fitted as a free parameter, or constrained using the mass-metallicity as an informative prior. However, galaxies are known to evolve chemically over time and trend towards higher metallicities as the Universe evolves.

Using the SED fitting code ProSpect (Robotham et al. 2020), we have used a simple physically motivated prescription to model an evolving metallicity history for individual galaxies (also used in Bellstedt et al. 2020b, 2021; Thorne et al. 2021, 2022) to extract star formation and metallicity histories for $\sim 90,000$ galaxies from the GAMA and DEVILS surveys.

Measuring metallicities from SED fitting is difficult due to the age-dust-metallicity degeneracy even when making simple, physically motivated assumptions about the chemical enrichment history. It is even more challenging when trying to measure metallicities pushing near the depths of photometry or with poor quality redshifts. We stress that despite this, it is still important to model metallicity in a physically motivated way so that it does not impact other galaxy properties (Appendix A). If the photometry quality is too poor to accurately measure the metallicity we suggest at least allowing the metallicity to evolve but using the MZR to provide an informative prior on the metallicity value.

In this work we used the metallicity and stellar mass estimates from ProSpect to derive the mass-metallicity relation in eight redshift bins from $z = 0$ to $z = 1.1$ (Section 3 and Figure 7). While there was no fitting prior set on the resulting gas-phase metallicity values in our implementation of ProSpect, we recover a MZR that is consistent at all redshifts with previous literature measurements from both nebular emission lines and forensic predictions from Bellstedt et al. (2021). Our results suggest that the normalisation of the MZR has evolved by 0.1 dex since $z = 1$, consistent with results from Cresci et al. (2012) and Bellstedt et al. (2021) who find little evolution in the normalisation of the MZR over the last 7 billion years (Figure 8). However, as we find a discontinuity in the normalisation of the MZR when swapping from the GAMA and DEVILS combined sample to just DEVILS, the existence and strength of this evolution could be impacted by sample selection.

Using the Fritz et al. (2006) AGN model included in ProSpect we also measure the MZR for AGN host galaxies for $z < 0.3$ and can rule out differences in the normalisation or shape of the MZR to that measured for galaxies with no SED-detected AGN at a level of 0.1 dex, but not at levels below 0.1 dex (Section 3.1 and Figure 9).

When exploring higher order correlations of the MZR with star formation activity we find that for galaxies with $M_* > 10^{10.5} M_\odot$, higher levels of star formation are associated with lower metallicities (Section 4.1). This trend is most clear when binning by distance from the star-formation main-sequence ($\Delta \text{SFMR}_{\text{MS}}$). This trend is in agreement with results demonstrating the existence of an anti-correlation between metallicity and star-formation activity (either SFR, sSFR, or AsSFR) at fixed stellar mass (see e.g. Mannucci et al. 2010; Salim et al. 2014; Curti et al. 2020). We also find that the overall shape of a galaxy’s SFH is correlated with its location on the MZR (Section 4.2 and Figure 14). Most apparent is the effect of galaxy downsizing, where we find that more massive galaxies formed at earlier times. This is consistent with the resolved galaxy study from Zhou et al. (2022) who found that chemical evolution histories have a strong dependence on stellar mass. We find that for galaxies with $M_* \approx 10^{10} M_\odot$, lower metallicities are recovered for galaxies that formed half their mass earlier than galaxies with higher metallicities.

For galaxies at a given stellar mass, we also find that higher metallicity values are associated with shorter periods of star formation and
star formation histories where the majority of stellar mass was formed before the peak star formation rate (Figure 15). These correlations between age and metallicity could be due to the age–metallicity–dust degeneracy, however recent results from integral field spectroscopy studies suggest these results could be physical.

The analysis in this work, combined with the accurate derivation of the cosmic SFH (Bellstedt et al. 2020b) and stellar mass functions and star-forming main sequence (Thorne et al. 2021), highlights the importance of giving careful consideration to both the metallicity and star formation histories of galaxies and demonstrates that accurate galaxy properties can be derived using SED fitting on only broadband photometry.

ACKNOWLEDGEMENTS

We thank the anonymous referee for their detailed reading and comments which improved this paper. JET is supported by the Australian Government Research Training Program (RTP) Scholarship. ASGR and LJMD acknowledge support from the Australian Research Council’s Future Fellowship scheme (FT200100375 and FT200100055 respectively). SB acknowledges support from the Australian Research Council’s Discovery Project and Future Fellowship funding schemes (DP180103740, FT200100375). LC acknowledges support from the Australian Research Council Discovery Project and Future Fellowship funding schemes (DP210100337, FT180100066).

We acknowledge the traditional owners of the land on which this research was completed, the Whadjuk Noongar people, and the land on which the AAT stands, the Gamilaraay people, and pay our respects to elders past and present.

DEVILS is an Australian project based around a spectroscopic campaign using the Anglo-Australian Telescope. The DEVILS input catalogue is generated from data taken as part of the ESO VISTA-VIDEO (Jarvis et al. 2013) and UltraVISTA (McCracken et al. 2012) surveys. DEVILS is part funded via Discovery Program by the Australian Research Council and the participating institutions. The DEVILS website is https://devilsurvey.org. The DEVILS data is hosted and provided by AAO Data Central (https://datacentral.org.au/).

GAMA is a joint European-Australasian project based around a spectroscopic campaign using the Anglo-Australian Telescope. The GAMA input catalogue is based on data taken from the Sloan Digital Sky Survey and the UKIRT Infrared Deep Sky Survey. Complementary imaging of the GAMA regions is being obtained by the GALEX MIS, VST KiDS, VISTA VIKING, WISE, Herschel-ATLAS, GMRT and ASKAP providing UV to radio coverage. GAMA is funded by the STFC (UK), the ARC (Australia), the AAO, and the participating institutions. The GAMA website is http://www.gama-survey.org/.

This work was supported by resources provided by the Pawsey Supercomputing Centre with funding from the Australian Government and the Government of Western Australia. We gratefully acknowledge DUG Technology for their support and HPC services.

All of the work presented here was made possible by the free and open R software environment (R Core Team 2020). All figures in this paper were made using the R MAGICaxis package (Robotham 2016b). This work also makes use of the CELESTIAL package (Robotham 2016a).

DATA AVAILABILITY

The DEVILS data products used in this paper are from the internal DEVILS team data release and presented in Davies et al. (2021), Thorne et al. (2021) and Thorne et al. (2022). They will be made publicly available as part of the DEVILS first data release described in Davies et al. (in preparation). The GAMA data products used in this paper were released as part of GAMA DR4 (Driver et al. 2022) and are available on the GAMA website6.

REFERENCES

Abazajian K., et al., 2004, AJ, 128, 502
Abazajian K. N., et al., 2009, ApJS, 182, 543
Aihara H., et al., 2019, PASJ, 71, 114
Alloin D., Collin-Souffrin S., Joly M., Vigroux L., 1979, A&A, 78, 200
Asplund M., Grevesse N., Sauval A. J., Scott P., 2009, ARA&A, 47, 481
Baldwin J. A., Phillips M. M., Terlevich R., 1981, PASP, 93, 5
Barrera-Ballesteros J. K., Sánchez S. F., Heckman T., Blanc G. A., MaNGA Team 2017, ApJ, 844, 80
Bellstedt S., et al., 2020a, MNRAS, 496, 3235
Bellstedt S., et al., 2020b, MNRAS, 498, 5581
Bellstedt S., et al., 2021, MNRAS, 503, 3309
Blanc G. A., La Y., Benson A., Katsianis A., Barraza M., 2019, ApJ, 877, 6
Boquien M., Burgarella D., Roehly Y., Buat V., Ciesla L., Corre D., Inoue A. K., Salas H., 2019, A&A, 622, A103
Bothwell M. S., Maiolino R., Kennicutt R., Cresci G., Mannucci F., Marconi A., Cicone C., 2013, MNRAS, 433, 1425
Brooks A. M., Governato F., Booth C. M., Willman B., Gardner J. P., Wadsley J., Stinson G., Quinn T., 2007, ApJ, 655, L17
Brown T., Cortese L., Catinella B., Kilborn V., 2018, MNRAS, 473, 1868
Bruzual G., Charlot S., 2003, MNRAS, 344, 1000
Cai W., Zhao Y., Zhang H.-X., Bai J.-M., Liu H.-T., 2020, ApJ, 903, 58
Catura F., Pipino A., Chiappini C., Matteucci F., Maiolino R., 2009, A&A, 504, 373
Camps-Fariña A., Sánchez S. F., Carigi L., Lacerda E. A. D., García-Benito R., Mast D., Galbany L., Ballesteros-Ballesteros J. K., 2021, ApJL, 922, L20
Capak P., et al., 2007, ApJ, 172, 99
Carnall A. C., McLure R. J., Dunlop J. S., Davé R., 2018, MNRAS, 480, 4379
Catinella B., et al., 2018, MNRAS, 476, 875
Chabrier G., 2003, PASP, 115, 763
Charlot S., Fall S. M., 2000, ApJ, 539, 718
Charta N., et al., 2021, ApJ, 908, 120
Chen X., Wang J., Kong X., 2022, ApJ, 933, 39
Chisholm J., Tremonti C., Leitherer C., 2018, MNRAS, 481, 1690
Conroy C., 2013, ARA&A, 51, 393
Cresci G., Mannucci F., Sommariva V., Maiolino R., Marconi A., Brusa M., 2012, MNRAS, 421, 262
Cresci G., Mannucci F., Curti M., 2019, A&A, 627, A42
Cullen F., et al., 2021, MNRAS, 505, 903
Curti M., Cresci G., Mannucci F., Marconi A., Maiolino R., Esposito S., 2017, MNRAS, 465, 1384
Curti M., Mannucci F., Cresci G., Maiolino R., 2020, MNRAS, 491, 944
D’Eugenio F., Colless M., Groves B., Bian F., Barone T. M., 2018, MNRAS, 479, 1807
Dalcanton J. J., 2007, ApJ, 658, 941
Dale D. A., Helou G., Magdis G. E., Armus L., Díaz-Santos T., Shi Y., 2014, ApJ, 784, 83 Davé R., Finlator K., Oppenheimer B. D., 2011, MNRAS, 416, 1354
Davé R., Thompson R., Hopkins P. F., 2016, MNRAS, 462, 3265
Davé R., Rafieferantsoa M. H., Thompson R. J., Hopkins P. F., 2017, MNRAS, 467, 115

6 http://www.gama-survey.org/dr4/
APPENDIX A: IMPACT OF METALLICITY IMPLEMENTATION ON DERIVED GALAXY PROPERTIES

To demonstrate the impact of different metallicity implementations on derived galaxy properties, we use the sample of ~7,000 low-redshift (z < 0.06) GAMA galaxies studied in Bellstedt et al. (2020b). We fit this sample using same implementation of ProSPECT as described in Bellstedt et al. (2020b) including SFH (massfunc_snorm_trunc), band dependent error floors, and dust priors. However, we fit the sample using three different metallicity prescriptions;

(i) using the linear mapping of metallicity growth to mass growth used in this work (Zfunc_massmap_11n - ‘Zevolve’),
(ii) assuming that all stars in a galaxy formed with the same metallicity but allowing this metallicity to be modelled as a free parameter (‘Zconstant’),
(iii) or assuming that all stars formed at solar metallicity (Z = 0.02, as adopted by Boquien et al. 2019 - ‘Zsolar’).

Figure A1 shows comparisons of the stellar mass, SFR, half-mass age, and dust opacity (tau_screen) when changing the metallicity prescription. It is immediately obvious that fixing the metallicity to Z = 0.02 results in systematic offsets in all properties when compared to the ‘Zevolve’ and ‘Zconstant’ cases. Fixing the metallicity to Z = 0.02 has the largest impact on the derived dust opacity, which can be offset by up to ~ 3 dex, however, the derived age is also significantly impacted with systematic offsets of ~ 0.5 dex at intermediate ages. The stellar masses, and SFRs are not systematically offset when allowing the metallicity value to change with either a constant or evolving metallicity history. However, age can be underestimated by ~ 0.1 dex at high ages when assuming a constant metallicity history as opposed to an evolving one due to the age-metallicity degeneracy (Worthey 1994; Papovich et al. 2001).

APPENDIX B: IMPACT OF CHOSEN STELLAR POPULATION MODELS ON DERIVED GALAXY PROPERTIES

One of the choices required when fitting the SEDs of galaxies is the set of templates, known as stellar population synthesis (SPS) models, used to model the expected emission from stars of certain ages and metallicities. The Bruzual & Charlot (2003) SPS model is the most commonly employed SPS model in broad-band SED fitting (see Figure 1 or Table A1 of Thorne et al. 2021 for an outline of the SPS models employed by different SED fitting codes), while the E-MILES SPS models are typically used for full spectral fitting techniques due to the finer wavelength resolution. However, E-MILES has limited wavelength coverage and does not cover the far ultraviolet for all age templates, and does not cover the near-infrared for the youngest age models (age < 450 Myr).

The creation of SPS models is conceptually straightforward, but current SPS templates are limited by incomplete isochrone tables, incomplete stellar libraries, and poorly calibrated physics (see Conroy 2013 for more details). Some of the challenges of SPS models include the potential importance of thermally pulsating asymptotic giant branch stars (Maraston et al. 2006), and the importance of binary evolution (Eldridge & Stanway 2012; Eldridge et al. 2017).

As Bruzual & Charlot (2003, BC03) and Eldridge et al. (2017, BPASS) are both implemented in ProSPECT, we can compare the impact on derived galaxy properties when changing the employed SPS model7. Using the sample of ~ 7,000 low redshift (z < 0.06) GAMA galaxies used in Bellstedt et al. (2020b, 2021) and the methods described in Bellstedt et al. (2020b) we re-fit the galaxies using the BPASS. The BPASS models include 13 metallicities spanning Z = 10^{-5} – 0.04, while BC03 only has 5 metallicity templates spanning Z = 10^{-4} – 0.05. However, BC03 includes a significantly finer sampling of ages with 196 age bins with ages less than 14 Gyrs while BPASS only has 42 bins with ages less than 14 Gyrs. BPASS has very high time resolution for young stars (< 100 Myrs) with age bins of millions of years, and very coarse resolution for older populations with some bins separated by 3 Gyrs.

Figure B1 shows comparisons between general galaxy properties derived using BPASS to those derived using BC03. The stellar masses show good agreement, however the stellar masses derived using BPASS are offset at high stellar masses at a level of approximately 0.1 dex. The SFRs show a systematic offset where BPASS recovers systematically lower SFRs at the level of ~ 0.15 dex across the entire SFR range. To quantify the ages of galaxies we use the half mass age, which is the time at which the galaxy had formed half of its stellar mass. We find that BPASS recovers older ages for young galaxies, and younger ages for old galaxies than BC03. We also show the difference in recovered metallicity when using BPASS, where the different choice of SPS model results in significantly different metallicity values than when using BC03. The straight line artefacts in both panels are a result of galaxies hitting the upper and lower limits of the metallicity templates.

Figure B2 shows the impact of the offset between the BC03, and

---

7 Vazdekis et al. (2016, E-MILES) is also incorporated in ProSPECT, but due to the limited wavelength coverage, there are many additional caveats that would make these comparisons far more difficult.
Figure A1. Impact of changing metallicity implementation on derived stellar masses (top left), SFRs (top right), ISM dust opacities (tau_screen, lower left), and half-mass ages (lower right). The different metallicity implementations are assuming constant solar ($Z = 0.02$; $Z_{\text{solar}}$), constant but the value of the metallicity is free ($Z_{\text{constant}}$), and using the evolving metallicity implementation described in Section 2.3.2 ($Z_{\text{evolve}}$). The running median and $1\sigma$ ranges are shown as the red solid and dashed lines respectively.

BPASS fits on the SFMS. As BPASS predicts systematically lower SFRs, the SFMS recovered using BPASS is below that recovered using BC03. Despite the difference in normalisation, both SPS models recover the same overall bending trend.

Figure B3 presents the recovered MZR using BC03 and BPASS. Due to the large offset in metallicities discussed previously, the MZR recovered using BPASS lies significantly below the MZR recovered using BC03. Despite the offset in metallicity values between the two SPS codes, the derived MZR follow the same trend with low mass galaxies having lower gas metallicities than high mass galaxies. The scatter on the MZR relation is higher for the BPASS derived sample however. As the same trends exist for the SFMS and MZR when fitting with both BC03 and BPASS, we do not expect that changing SPS model will have an impact on the general trends we find in this work. However, there would be differences in metallicity normalisation due to the differences discussed above.

APPENDIX C: DIFFERENCES BETWEEN THE $Z=0$ MZR FROM THIS WORK AND BELLSTEDT ET AL. 2021

Figure C1 reproduces the $z \approx 0$ panel of Figure 7 but also includes a recalculation of the MZR using the catalogue of SED fits from
Figure B1. Comparison of galaxy properties derived using BPASS to those derived using BC03. We show the difference in stellar mass, star formation rate, half mass age, and metallicity. In each panel the running median and 1σ range are shown as the red solid and dashed lines respectively.

Figure B2. Comparison of the SFMS derived using BC03 and BPASS at $z = 0$. We show the underlying distribution of galaxies using the BPASS SPS models in the gray 2D histogram with the running median and 1σ ranges in solid and dashed red respectively. We also show the running median and 1σ range from the BC03 version in blue.

Figure B3. Comparison of the MZR derived using BC03 (blue) and BPASS (red). As per Figure B2, the underlying 2D histogram shows the distribution of galaxies when using BPASS. For comparison we also show values from Lara-López et al. (2013) and Ly et al. (2016) both of which are measured through spectra.
Figure C1. The lowest redshift panel from Figure 7 with the addition of the median MZR derived using the SED fits presented in Bellstedt et al. (2021, B21) using the same sample of galaxies used in this work. We recalculate the median MZR using the median and the maximum-likelihood (best fit) parameter values provided by Bellstedt et al. (2021) which are shown as the blue and red lines respectively.

Bellstedt et al. (2021) (released with GAMA DR48) but with various changes. First, we extend the sample from Bellstedt et al. (2021) to match the sample used in this work. Specifically, this includes the addition of the G23 field, not using a ΔZ cut, and extending the redshift range from \( z < 0.06 \) to \( z < 0.112 \). The Bellstedt et al. (2021) results were derived using the median parameter value of both the mass and metallicity, whereas our parameter values are based on the maximum-likelihood (best fit) step of the posterior. As both the median and best fit parameter values are provided by Bellstedt et al. (2021) we recalculate the MZR using both to highlight the difference. These are shown as the blue and red lines respectively. The difference between the blue and yellow lines highlights the impact from changing the sample selection to encompass a larger sample of galaxies as these are both extracted using the median parameter values. The difference between the blue and red lines highlights the effect of how the galaxy parameter values were chosen, where the best fit values recover a MZR relation with systematically lower metallicities at a given stellar mass. And finally, the difference between the red and black lines demonstrates that the incorporation of the AGN component in the SED fitting, and the change from a full MCMC chain to the HIGHLANDER optimization routine does not significantly impact the derived MZR. Additionally the Bellstedt et al. (2021) MZR shown in the first panel of Figure 7 is extracted using the metallicity and star formation histories and has been traced back to a lookback time of 1 Gyr rather than using the \( Z_{\text{gas}} \) and stellar mass at \( t = 0 \).

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

http://www.gama-survey.org/dr4/