

# First Light And Reionization Epoch Simulations (FLARES) VII: The star formation and metal enrichment histories of galaxies in the early Universe

Stephen M. Wilkins<sup>1,2\*</sup>, Aswin P. Vijayan<sup>1,3,4</sup>, Christopher C. Lovell<sup>1,5</sup>, William J. Roper<sup>1</sup>, Erik Zackrisson<sup>6</sup>, Dimitrios Irodotou<sup>1,7</sup>, Louise T. C. Seeyave<sup>1</sup>, Jussi K. Kuusisto<sup>1</sup>, Peter A. Thomas<sup>1</sup>, Joseph Caruana<sup>2,8</sup> and Christopher J. Conselice<sup>9</sup>

<sup>1</sup>Astronomy Centre, University of Sussex, Falmer, Brighton BN1 9QH, UK

<sup>2</sup>Institute of Space Sciences and Astronomy, University of Malta, Msida MSD 2080, Malta

<sup>3</sup>Cosmic Dawn Center (DAWN), Rådmandsgade 62, DK-2200 København, Denmark

<sup>4</sup>DTU-Space, Technical University of Denmark, Elektrovej 327, DK-2800 Kgs. Lyngby, Denmark

<sup>5</sup>Centre for Astrophysics Research, School of Physics, Engineering & Computer Science, University of Hertfordshire, Hatfield AL10 9AB, UK

<sup>6</sup>Observational Astrophysics, Department of Physics and Astronomy, Uppsala University, Box 516, SE-751 20 Uppsala, Sweden

<sup>7</sup>Department of Physics, University of Helsinki, Gustaf Hällströmin katu 2, FI-00014, Helsinki, Finland

<sup>8</sup>Department of Physics, University of Malta, Msida MSD 2080, Malta

<sup>9</sup>Jodrell Bank Centre for Astrophysics, University of Manchester, Oxford Road, Manchester, M13 9PL, UK

Accepted 2022 November 4. Received 2022 October 24; in original form 2022 August 30

## ABSTRACT

The star formation and metal enrichment histories of galaxies – at any epoch – constitute one of the key properties of galaxies, and their measurement is a core aim of observational extragalactic astronomy. The lack of deep rest-frame optical coverage at high redshift has made robust constraints elusive, but this is now changing thanks to *JWST*. In preparation for the constraints provided by *JWST*, we explore the star formation and metal enrichment histories of galaxies at  $z = 5$ –13 using the First Light And Reionization Epoch Simulations (FLARES) suite. Built on the EAGLE model, the unique strategy of FLARES allows us to simulate galaxies with a wide range of stellar masses (and luminosities) and environments. While we predict significant redshift evolution of average ages and specific star formation rates, our core result is mostly a flat relationship of age and specific star formation rate with stellar mass. We also find that galaxies in this epoch predominantly have strongly rising star formation histories, albeit with the normalization dropping with redshift and stellar mass. In terms of chemical enrichment, we predict a strong stellar mass–metallicity relation present at  $z = 10$  and beyond alongside significant  $\alpha$ -enhancement. Finally, we find no large-scale environmental dependence of the relationship between age, specific star formation rate, or metallicity with stellar mass.

**Key words:** methods: numerical – galaxies: evolution – galaxies: formation – galaxies: high-redshift – galaxies: extinction – infrared: galaxies.

## 1 INTRODUCTION

A key goal of observational extragalactic astrophysics is the measurement of star formation and metal enrichment histories of representative samples of galaxies stretching over a wide range of redshifts and stellar masses. Doing so provides critical insights into the physical processes responsible for galaxy formation and evolution (e.g. Schaye et al. 2010). These histories can also be used to measure the cosmic history star formation and metal enrichment history, and the build up of stellar mass throughout cosmic history. This includes probing star formation which is observationally inaccessible *in situ*, such as in small galaxies at very high redshift.

While much progress has been achieved at low and intermediate redshift, particularly thanks to surveys like SDSS, GAMA, COSMOS, UltraVISTA, VIDEO, and CANDELS (e.g. Adams et al. 2021; McLeod et al. 2021; Driver et al. 2022), constraints at the highest red-

shifts ( $z > 5$ ) remain highly uncertain for several reasons. First, most high-redshift galaxies, at least those identified so far, have intense recent star formation. Young stellar populations in these systems dominate the energy output, ‘outshining’ older stellar populations, even in the rest-frame optical. Secondly, prior to *James Webb Space Telescope* (*JWST*), our only access to the rest-frame optical at sensitivities unattainable by ground based telescopes was from *Spitzer*, and then only typically in two broad photometric bands at 3.6 and 4.5  $\mu\text{m}$ . The use of *Spitzer* not only limits us to the brightest sources but also limits *clean* access to essential diagnostics, including the Balmer break and the strong optical emission lines. This is further complicated by the increasing prevalence of strong nebular line emission (e.g. Stark et al. 2013; Wilkins et al. 2013; de Barros, Schaerer & Stark 2014; Smit et al. 2015; Wilkins et al. 2020) at high redshift, which can be easily confused for age sensitive features like the Balmer break, especially where only photometric redshifts are available. Despite these challenges, however, there has been some progress in constraining the stellar masses and star formation histories of galaxies at  $z > 5$  prior to *JWST*, (e.g. Eyles et al. 2007; Salmon et al. 2015; Song et al. 2016;

\* E-mail: [s.wilkins@sussex.ac.uk](mailto:s.wilkins@sussex.ac.uk)

Davidzon et al. 2017), with some constraints now available even at  $z \sim 10$  (Laporte et al. 2021; Tacchella et al. 2022). These have, so far, revealed a consistent picture of falling stellar mass densities and increasing specific star formation rates to higher redshift. However, finer details, such as the slope of the specific star formation rate–stellar mass relation, remain highly uncertain.

With the successful commissioning of the highly sensitive *JWST*, we can now address many open questions beyond *Spitzer*’s capabilities. *JWST* not only provides sensitive near and mid-infrared imaging in several bands, extending to  $5 \mu\text{m}$  and beyond with MIRI, but also near-infrared spectroscopy. Spectroscopy will provide unambiguous redshifts, allow us to constrain the contribution of nebular line emission, and enable the use of powerful rest-frame optical emission line diagnostics. At the time of writing, the first constraints on the star formation and metal enrichment histories of high-redshift galaxies from *JWST* are emerging (e.g. Adams et al. 2022; Carnall et al. 2022; Chen et al. 2022; Finkelstein et al. 2022a; Leethochawalit et al. 2022; Naidu et al. 2022; Trump et al. 2022).

Anticipating the power of these observations in this work, we present comprehensive predictions for the star formation and metal enrichment histories of galaxies simulated by FLARES: First Light And Reionization Epoch Simulations project (Lovell et al. 2021; Vijayan et al. 2021). FLARES combines the calibrated EAGLE physics model (Crain et al. 2015; Schaye et al. 2015) with a unique simulation strategy designed to efficiently simulate galaxies over a wide range of masses and environments at high redshift. This wider dynamic range is key to capturing the range of galaxy populations accessible to *JWST*.

This paper is structured as follows in Section 2, we introduce the FLARES suite of simulations. In Section 3, we then explore FLARES predictions for the star formation histories of galaxies. Here, we make predictions using several different metrics (Section 3.2), fit individual star formation histories by a range of simple parametrizations (Section 3.3), compare with existing observations (Section 3.5), and explore the environmental dependence (Section 3.4). In Section 4, we then focus on the metal enrichment of galaxies. We first present the evolution mass–metallicity relation (Section 4.2), including its environmental dependence (Section 4.2.3). We then explore how stellar metallicity is correlated with age (Section 4.3) and the distribution of metallicities within individual galaxies (Section 4.4). Finally in Section 5, we present our conclusions. In this work, distance measures preceded by ‘c’ are in comoving units while the ones with ‘p’ are in physical units. We assume a Planck year 1 cosmology ( $\Omega_m = 0.307$ ,  $\Omega_\Lambda = 0.693$ ,  $h = 0.6777$ , Planck Collaboration et al. 2014).

## 2 THE FIRST LIGHT AND REIONIZATION EPOCH SIMULATIONS

In this study, we make use of the FLARES simulation suite. FLARES is introduced in Lovell et al. (2021) and Vijayan et al. (2021), and we direct the reader to those papers for a detailed introduction. In brief, FLARES is a suite of 40 spherical  $14 h^{-1} \text{cMpc}$  radius re-simulations. Regions re-simulated by FLARES are selected from a large  $(3.2 \text{cGpc})^3$  parent dark matter only simulation, and span a range of environments: (at  $z \approx 4.7$ )  $\log_{10}(1 + \delta_{14}) = [-0.3, 0.3]^1$ , with over-representation of the extremes. With the knowledge of each region’s density contrast, combined distribution functions can

<sup>1</sup>Where  $\delta_{14}$  is the density contrast measured within the re-simulation volume size.

be constructed, which approximate a much larger volume than that simulated, allowing FLARES to predict distribution functions over a much larger dynamic range than that achievable with hydrodynamic periodic volumes.

### 2.1 Physics model

FLARES adopts the AGNdT9 variant of the EAGLE simulation project (Crain et al. 2015; Schaye et al. 2015), and utilizes identical resolution and cosmological parameters to the fiducial EAGLE reference simulation. The AGNdT9 variant produces similar mass functions to the reference model, but better reproduces the hot gas properties in groups and clusters leading to its utilization by the C-EAGLE project (Barnes et al. 2017). The key physics of the EAGLE model relevant to this work concerns star formation and chemical enrichment, and here we briefly summarize these deferring a thorough introduction to Schaye et al. (2015), Crain et al. (2015), and references therein.

#### 2.1.1 Star formation

Current large cosmological simulations, including EAGLE, do not have the resolution to simulate star formation from first principles. In EAGLE, the star formation recipe uses the observed Kennicutt–Schmidt star formation law (Kennicutt 1998), now rewritten as a metallicity dependent pressure law (equation 1 in Schaye et al. 2015). This is implemented as described in Schaye & Dalla Vecchia (2008), where gas particles which are cold enough (with temperatures  $\lesssim 10^4 \text{K}$ ) above the metallicity-dependent star formation threshold of Schaye (2004) are stochastically converted to star particles, using the pressure-dependent version of the observed Kennicutt–Schmidt law. The star particles formed are assumed to represent a simple stellar population (SSP) formed with a Chabrier (2003) initial mass function.

#### 2.1.2 Chemical enrichment

These SSPs lose mass through stellar winds arising from asymptotic giant branch (AGB) stars and massive stars as well as type Ia (SNIa) and type II (SNII) supernovae, following the prescription of Wiersma et al. (2009). Stellar particles lose mass dependent on their age and metallicity due to the main sequence lifetimes of the constituent stars (Portinari, Chiosi & Bressan 1998; Marigo 2001). This mass is distributed to neighbouring gas particles based on the SPH kernel. The elements H, He, C, N, O, Ne, Mg, Si, and Fe are tracked individually. At early times ( $\lesssim 100 \text{Myr}$ ), the mass-loss is mainly from SNII, leading to significant  $\alpha$ -enhancement (i.e.  $[\alpha/\text{Fe}] > 0$ ) in younger galaxies. Later, the mass-loss comes from AGB and SNIa.

Most massive galaxies ( $M_*/M_\odot > 10^9$ ) at  $z > 5$  are compact with the majority of star formation taking place in dense cores (Roper et al. 2022). The low metallicity present at high redshift in tandem with the core density yield strong enrichment in these cores, while enrichment is inhibited by low densities elsewhere. None the less, it should be noted that despite the enhancement, stellar feedback is still inefficient at high redshift.

EAGLE tracks both the particle abundance (directly enriched by star particles) and the smoothed abundance (smoothed value obtained using the SPH kernel, see Wiersma et al. 2009). Since the simulation does not implement diffusion of metals between gas particles, it can lead to certain individual particles exhibiting extreme metallicities. All the quoted metallicity values in this work are

smoothed metallicities to mitigate some effects of this. The smoothed abundances were also used in computing the cooling rates of the gas, while the probability of star formation used the particle metallicities.

### 2.1.3 Variations on the resolution and physics model

The properties of galaxies predicted by hydrodynamical simulations like FLARES depend on modelling choices including the cosmology, the physics model, and resolution, expressed in-terms of the dark matter and gas particle masses. Varying these provides key insights into how properties such as distribution functions and scaling relations emerge, see e.g. Crain et al. (2015).

While in this work, we do not attempt a thorough exploration of how changes to the model and/or parameters impact our results, in a work in preparation we do explore how changes to the star formation, feedback, and reionization prescriptions using re-simulations of FLARES regions. However, we do briefly explore the impact of increasing the resolution using the higher-resolution ( $M_g = 2.26 \times 10^5 M_\odot$ ) EAGLE recal (ReCal) model presented in Schaye et al. (2015). It is however important to note that in addition to adopting different resolutions the ReCal and AGNdT9 models adopt different values for several key model parameters to ensure they are converged. It is also important to note that the volume simulated with the ReCal model is very small and thus lacks the dynamic range of FLARES limiting comparisons to the low-mass end and only at  $z = 5$ .

## 2.2 Galaxy identification

Galaxies in FLARES are first identified as groups via the Friends-Of-Friends (FOF, Davis et al. 1985) algorithm, and subsequently subdivided into bound objects with the SUBFIND (Springel et al. 2001; Dolag et al. 2009) algorithm. For a full description of the method and handling of pathological objects see McAlpine et al. (2016), whose methodology we follow.

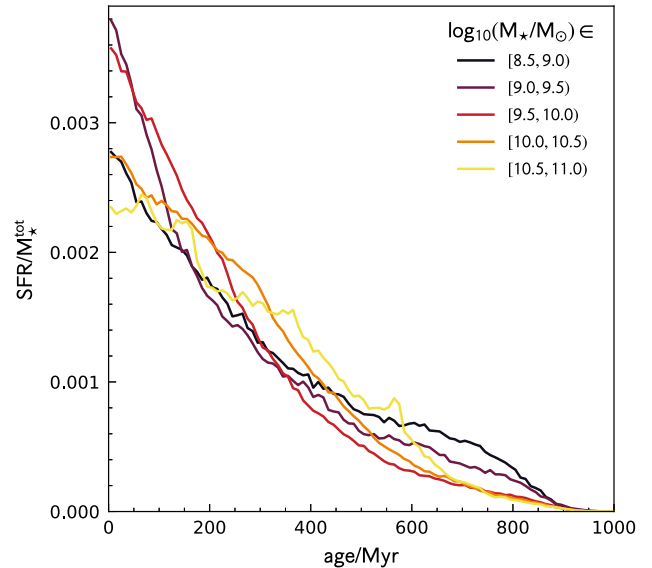
In this work, we restrict ourselves to galaxies with  $M_* > 10^{8.5} M_\odot$  corresponding to a minimum of a few hundred star particles. By default when measuring properties, we use 30 kpc radius apertures centred on the most bound particle of each subgroup (the particle with the highest gravitational potential). However, in Section A1, we explore the consequences of this assumption: while the choice of, aperture can have a significant impact on the total stellar mass and star formation, the impacts on the specific star formation rate (and other metrics of the star formation history) mostly cancel out.

## 3 STAR FORMATION HISTORIES

The star formation history (SFH) describes the evolution of a galaxy's star formation activity. In the case of hydrodynamical simulations like FLARES, the SFH is defined by the age (and initial mass) distribution of the star particles making up the galaxy. To describe the SFH of a galaxy, we can employ several different metrics, including: the specific star formation rate, age, and moments, amongst others. In this section, we begin by presenting stacked star formation histories before presenting predictions for these metrics, then go on to explore their environmental dependence, and compare with observations.

### 3.1 Stacked star formation histories

We begin, in Fig. 1, by presenting the average (stacked) star formation histories of galaxies at  $z = 5$  in bins of stellar mass. With the possible exception of the highest mass-bin, where we see tentative evidence



**Figure 1.** The weighted average (stacked) star formation histories of galaxies at  $z = 5$  in various stellar mass bins.

for a plateau, the average star formation history is rapidly increasing. In the subsequent section, we quantify the shapes of individual star formation histories using a set of metrics.

### 3.2 Metrics

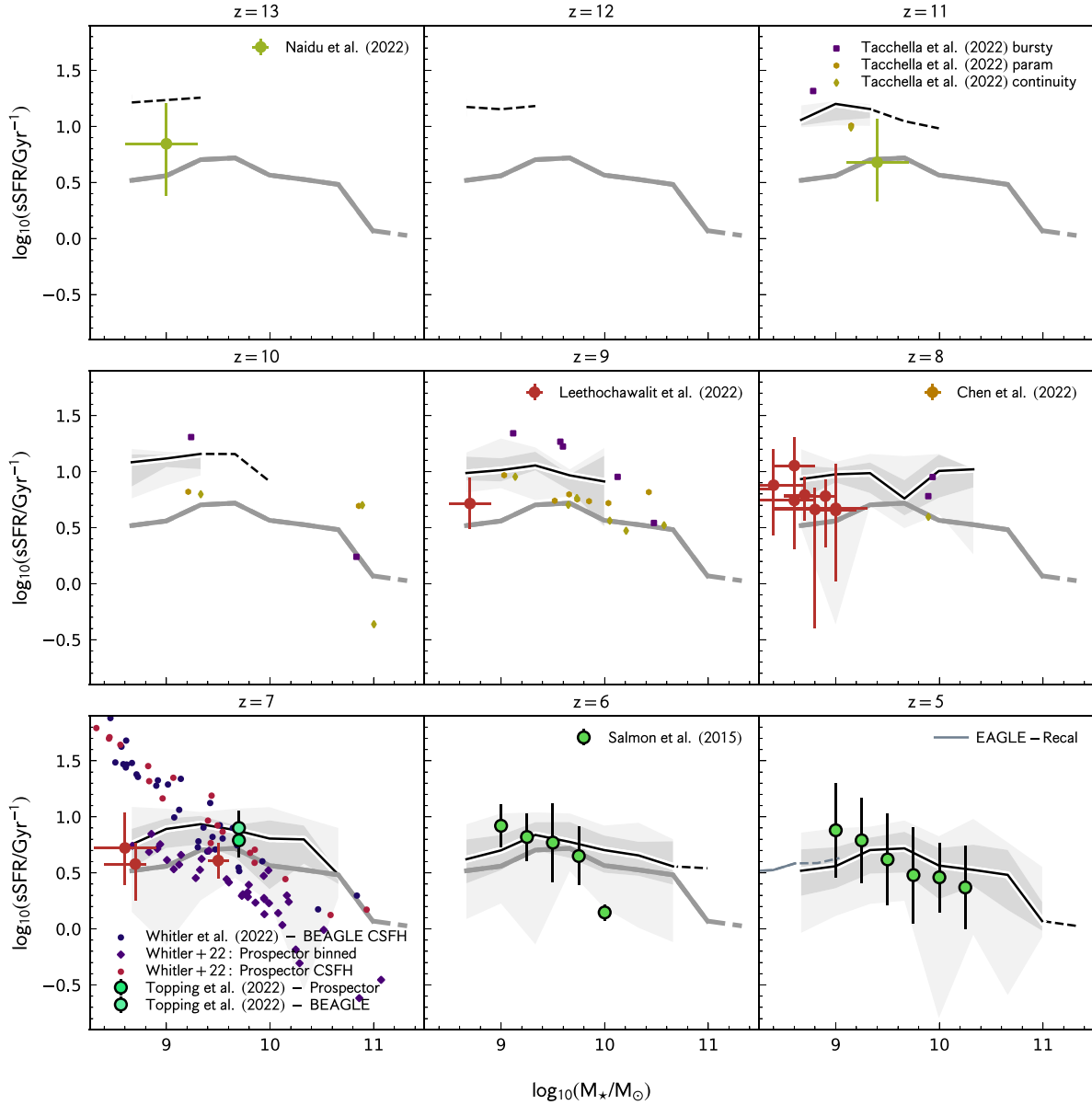
We now turn our attention to quantifying the star formation histories of individual galaxies using a variety of metrics. In most cases, we present the median of the population alongside both the central 68.4 per cent ( $P_{84.2} - P_{15.8}$ ) and 95.6 per cent ranges ( $P_{97.8} - P_{2.2}$ ). The median and these ranges are calculated by applying a weight to each galaxy dependent on its parent re-simulation to ensure the selected regions are representative of the full parent volume.

#### 3.2.1 Specific star formation rates

The most commonly utilized and observationally accessible metric of the SFH is the specific star formation rate (sSFR), the ratio of recent to integrated star formation activity normally expressed as the current star formation rate divided by the stellar mass. In this work, we employ the star formation rate averaged over 50 Myr. In Section A2 (and specifically Fig. A2), we explore the impact of this assumption finding a relatively small impact ( $<0.1$  dex).

The resulting relationship between sSFR and stellar mass – also known as the star forming main-sequence – is shown for  $z = 5-13$  in Fig. 2. This reveals a clear redshift evolution: from  $z = 13 \rightarrow 5$  the average sSFR drops by  $\sim 0.7$  dex. The evolution of both the specific star formation rate and age (see Section 3.2.2) is also presented in Fig. 3, but in terms of the age of the Universe. When expressed in this way, the redshift evolution is reduced though some still does remain, at least for the age.

While Fig. 2 reveals significant redshift evolution, there is no strong trend with stellar mass, at least at  $M_* = 10^{8.5-10.5} M_\odot$ . At  $z = 5$ , from a peak at  $M_* \approx 10^{9.5} M_\odot$ , the average specific star formation rate drops by only  $\approx 0.2$  dex by  $M_* \approx 10^{10.5} M_\odot$ . There is tentative evidence of a sharper drop by  $M_* = 10^{11} M_\odot$ , but FLARES contains only a handful galaxies at this mass. However, while there is no precipitous drop in the average sSFR, there is a sharply



**Figure 2.** The predicted relationship between the specific star formation rate (sSFR) and stellar mass for integer redshifts  $z \in \{5, 13\}$ . The black solid line denotes the weighted binned median. The two shaded regions show the central 68 and 95 per cent ranges. Where the number of galaxies in each bin falls below 10, the solid line is replaced by a dashed line, and we no longer present the range. The thick grey line in each panel denotes the median relation at  $z = 5$ . Also shown are various observational constraints including constraints on individual galaxies from Tacchella et al. (2022), Whittler et al. (2022) (based on the sample identified by Endsley et al. 2021), Carnall et al. (2022), Naidu et al. (2022), Leethochawalit et al. (2022), Finkelstein et al. (2022a), and Chen et al. (2022), and stacked results from Salmon et al. (2015) and Topping et al. (2022). Uncertainties on the stacked measurements of Salmon et al. (2015) and Topping et al. (2022) are the error on the median, not the distribution width. For the *JWST*-based results, we included the published uncertainties, however for the *Spitzer*-based results, we omit uncertainties for readability.

increasing fraction of galaxies that have much lower specific star formation rates. In Fig. 4, we plot the weighted fraction of massive  $M_* > 10^{10} M_\odot$  galaxies with specific star formation rates falling below three thresholds: the inverse age of the Universe,  $1 \text{ Gyr}^{-1}$ , and 10 per cent of the median sSFR. Irrespective of the choice of threshold, we see that the fraction rapidly increases from  $\approx 0$  at  $z \geq 8$  to  $\sim 0.05$  at  $z = 5$ . The origin of this rapid evolution will be discussed in a companion paper (Lovell et al.), though it appears that a reduction in the sSFR is strongly correlated with significant AGN activity.

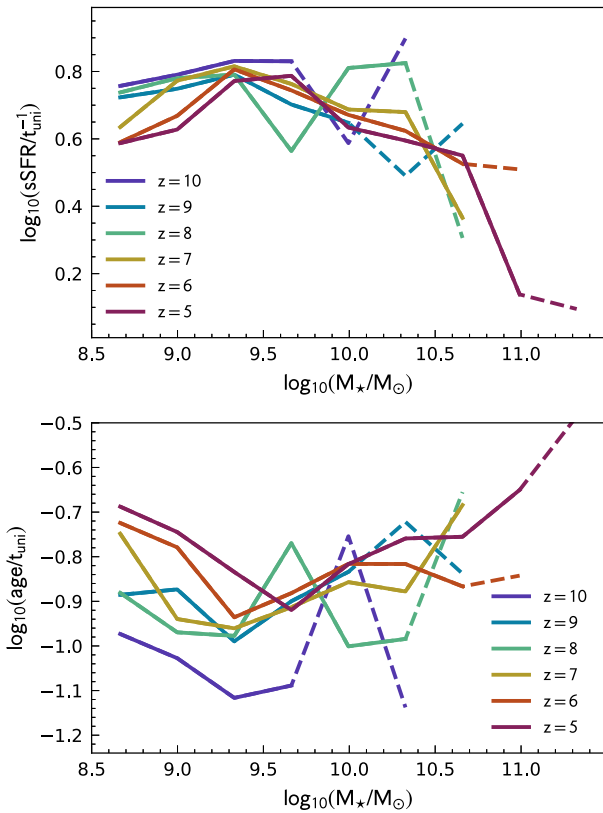
Finally, to test the sensitivity of our predictions to the resolution, we also include predictions from the higher-resolution EAGLE

ReCal model to the  $z = 5$  panel. While the overlap region is very small (due to the small volume simulated with the ReCal model), the resulting median specific star formation rates is very similar to that predicted by FLARES giving some confidence that we are not strongly sensitive to resolution effects at low masses.

### 3.2.2 Ages

Another common metric is the age, though this varies in definition. In this work, we define the age as the initial-mass weighted median





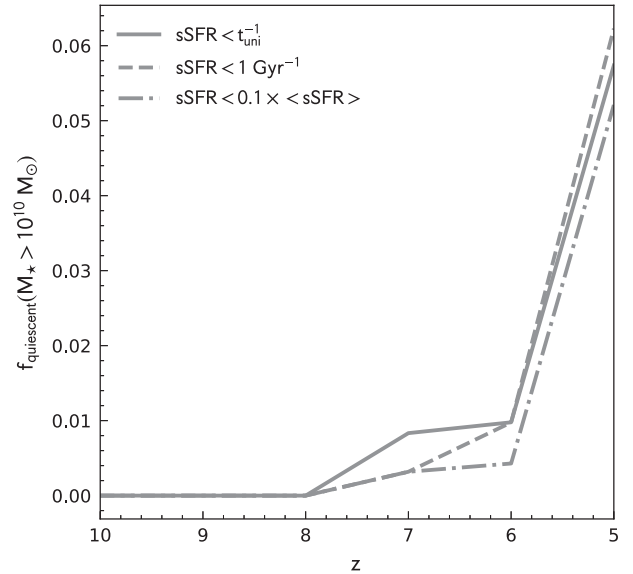
**Figure 3.** The redshift evolution of specific star formation rate (top) and age (bottom) expressed in terms of the age of the Universe.

of the stellar particle ages. Essentially, this is the time since the first 50 per cent of stars were formed. It is important to note, however, that the term age could also be used to describe the mean age; in the context of FLARES, the mean and median ages are similar – and exhibit similar trends – but are not identical (see Section 3.2.4). It is also worth noting that an additional literature definition of age is the duration since star formation began (e.g. Laporte et al. 2021), sometimes described as maximum age. In the context of simulations, however, this is expected to be resolution dependent and subject to large uncertainties, and for this reason we do not utilize it in this analysis.

The relationship between the predicted age and stellar mass is shown in Fig. 5. This relationship largely mirrors the trends seen for the sSFR – stellar mass – redshift plane: average ages decrease with increasing redshift while remaining largely flat with stellar mass.

### 3.2.3 Shape

While the specific star formation rate gives some indication of the shape of the recent star formation history a clearer picture is revealed by simply comparing the star formation activity averaged over two time-scales. We do exactly this in Fig. 6, where we show the star formation activity measured over the last 50 Myr compared to that measured over 200 Myr. The majority of galaxies at all redshifts have  $SFR_{50}/SFR_{200} > 1$ , meaning they have star formation histories that are rising over these time-scales. However, the fraction of galaxies with rising star formation histories declines from approximately unity at  $z = 10$  to  $\approx 70$  per cent at  $z = 5$ . Again, there is little trend with stellar mass though the most massive galaxies tend to have a lower fraction of galaxies with still increasing star formation histories.



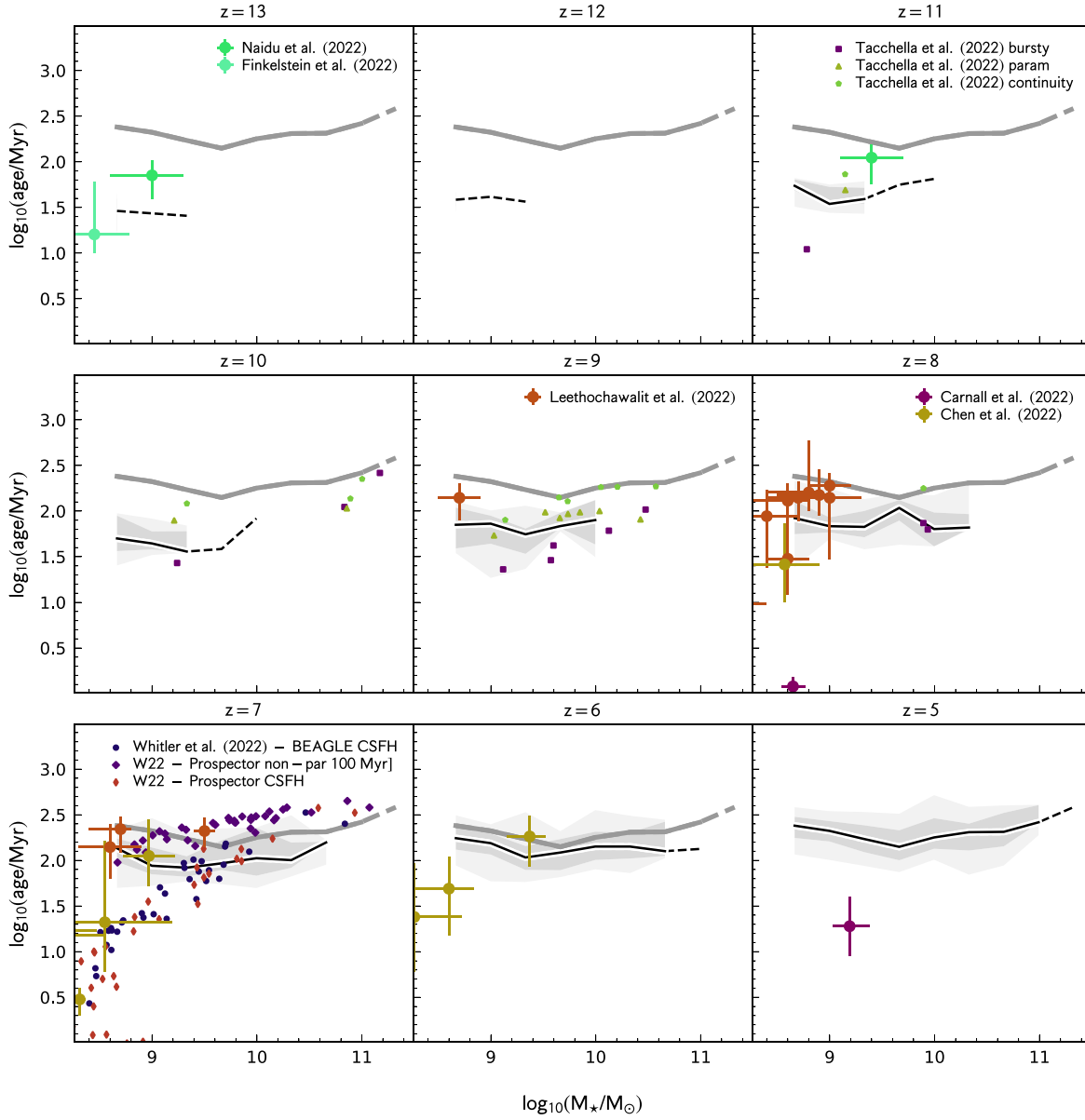
**Figure 4.** The weighted fraction of massive ( $M_* > 10^{10} M_{\odot}$ ) galaxies falling below a specific star formation rate threshold as a function of redshift  $z = 10 \rightarrow 5$ . The thresholds are the inverse age of the Universe (solid line),  $1 \text{ Gyr}^{-1}$  (dashed line), and 10 per cent of the median sSFR (dash-dotted line).

### 3.2.4 Moments

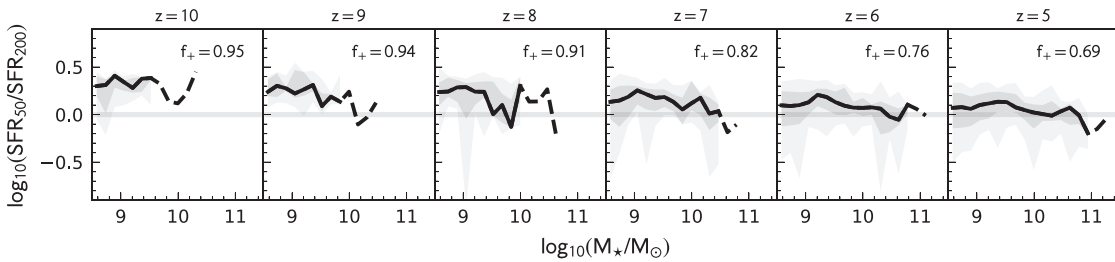
An alternative set of metrics are the moments of the (mass-weighted) stellar age distribution. These are useful as they can be compared directly to various distributions e.g. normal, exponential, and half-normal for which some of the moments have fixed values. The first four moments are presented in Fig. 7. Since the first moment is the mean this is similar, but not identical, to our definition of the age (the median). To avoid confusion with the age we present the inverse of the mean and label it  $\lambda$ . The second moment is the variance; but instead of presenting this directly we present the quantity  $\text{mean}/\sqrt{\text{var}}$ , which has fixed values of unity and  $\approx 1.32$  for an exponential and half-normal distribution, respectively. This immediately reveals that galaxy star formation histories are not statistically well described by an exponential distribution, particularly at the high-mass end for  $z < 9$ . This is also reflected in the third and fourth moments of the distribution, the skew and excess kurtosis, respectively; in both cases, the measured values lie below those expected (2 and 6 respectively) for a pure exponential distribution. On the other hand, the predicted median value of the  $\text{mean}/\sqrt{\text{var}}$ , skew, and excess kurtosis all closely match that expected for a half-normal distribution peaking at the observation epoch suggesting that this provides a useful single-parameter description of galaxy star formation histories in FLARES.

## 3.3 Star formation history parametrization

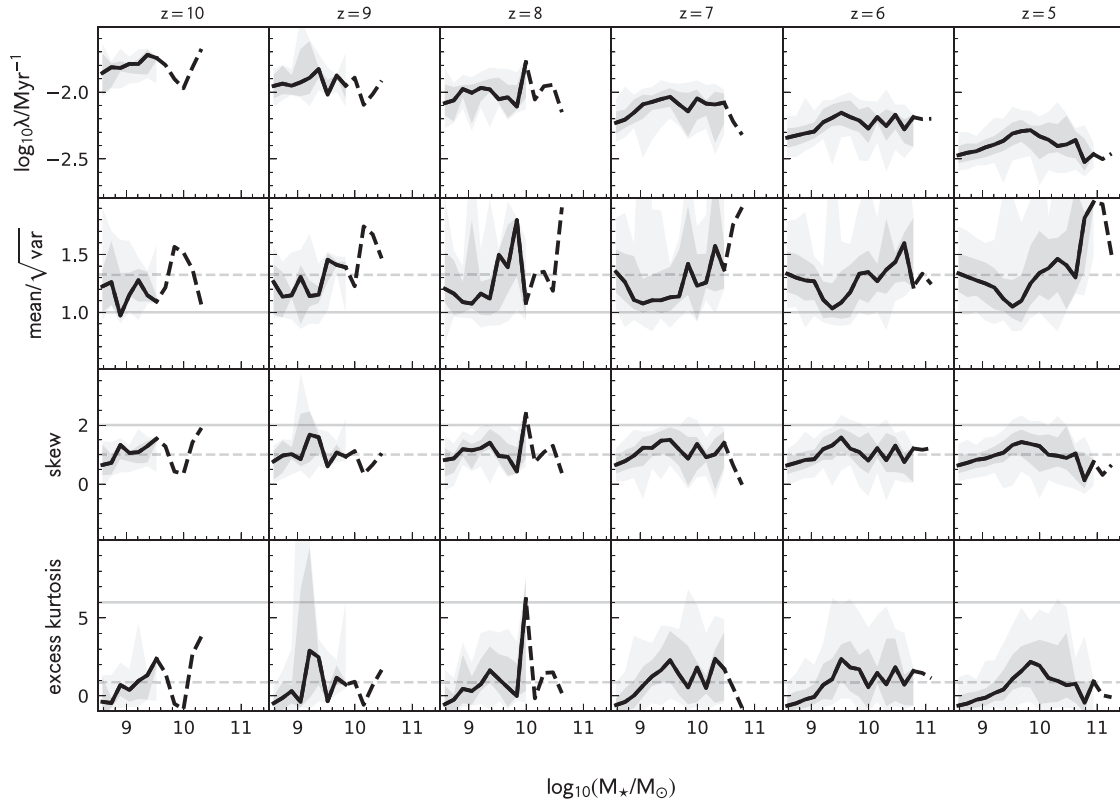
While a half-normal distribution reproduces the skew and variance, it assumes we are always observing galaxies at the peak of their star formation histories, something which is clearly not possible. Such a description also fails to account for the increasing numbers of galaxies, revealed in Fig. 6, which have declining star formation histories. As an alternative to the single parameter half-normal distribution, we also consider the two parameter truncated normal and truncated lognormal distributions, which have been shown to achieve good fits to simulated and observed SFHs at low redshift (Diemer et al. 2017). We fit the star formation histories of every galaxy by these



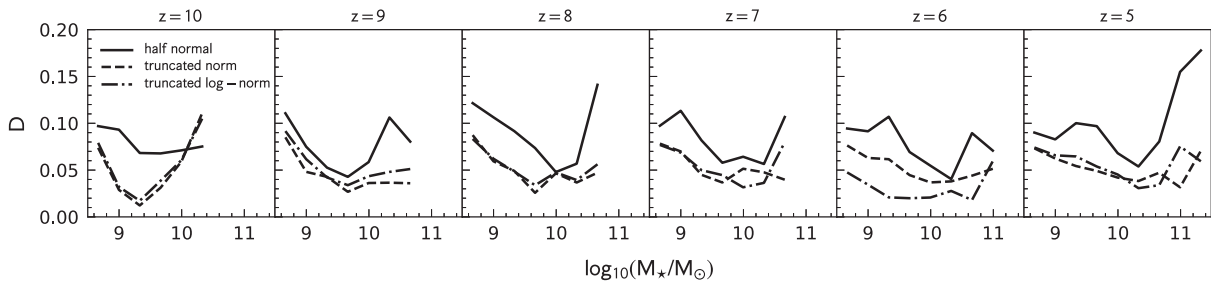
**Figure 5.** The same as Fig. 2 but for the mass-weighted median age.



**Figure 6.** The ratio of the SFR averaged over 50 and 200 Myr. The solid horizontal line denotes  $\text{SFR}_{50}/\text{SFR}_{200} = 1$  – galaxies on this line have constant star formation histories. The label denotes the fraction of galaxies that have rising star formation histories measured over this time-scale.



**Figure 7.** The first four moments of the mass-weighted stellar age distribution for galaxies in FLARES with  $\log_{10}(M_*/M_\odot) > 8.5$  at  $z = 10 \rightarrow 5$ . To avoid confusion with the median age, we present the inverse of the first moment. For an exponential distribution, this is equivalent to the parameter  $\lambda$ . For the second moment, we instead present the quantity  $\text{mean}/\sqrt{\text{var}}$ , which for an exponential distribution is simply unity and for a half-normal is  $\approx 1.32$ . We also add lines for the expected skew and excess kurtosis of an exponential (solid line) and half-normal distribution (dashed line).



**Figure 8.** The median value of the Kolmogorov–Smirnov test statistic  $D$  for half-normal, truncated normal, and truncated lognormal star formation history.

distributions, and computed the Kolmogorov–Smirnov test statistic  $D$  with the results presented in Fig. 8. This analysis reveals that both truncated normal and lognormal distributions yield similar level of improvement over a half-normal distribution. For parametric spectral energy distribution fitting, we then advocate one of these distributions as the most suitable parametrization.

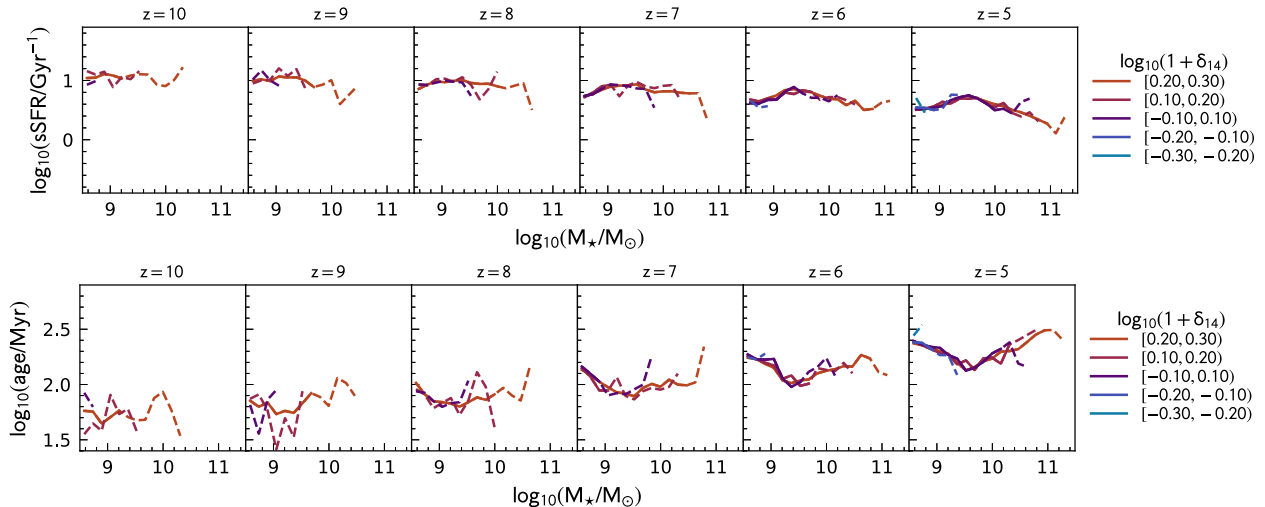
### 3.4 Environmental dependence

A unique feature in FLARES is the re-simulation of a wide range of environments from  $\log_{10}(\delta_{14}) = -0.3 \rightarrow 0.3$ . This allows us to study the relatively large-scale environmental dependence of the shape of the SFH. In Fig. 9, we present the relationship between stellar mass and sSFR and age as a function of galaxy environment. Matching that found in Lovell et al. (2021), we see no environmental dependence of

these properties on the scale of our individual simulations. However, it is important to note that this is not to say that environment doesn't play a strong role in galaxy formation at high-redshift: there is a strong environmental bias, such that underdense regions contain proportionally many fewer galaxies, and vice versa. In addition, there is an implicit dependence of galaxy mass on environment: only the most overdense regions produce massive galaxies at high-redshift.

### 3.5 Comparison with observational constraints

At the time of writing, we are now seeing the first constraints on the star formation histories of galaxies from *JWST* (Carnall et al. 2022; Chen et al. 2022; Finkelstein et al. 2022a; Leethochawalit et al. 2022; Naidu et al. 2022). Prior to *JWST* however, it was possible, by combining ground based or *Hubble* optical and near-IR observations



**Figure 9.** The relationship between the specific star formation rate (top) and age (bottom) as a function of stellar mass as split by the density contrast of the simulation.

with *Spitzer*, to constrain the rest-frame UV – optical emission of  $z > 5$  galaxies. Alongside the FLARES predictions in Figs 2 and 5, we show current observational constraints on the specific star formation rate and age, including early results from *JWST* and previous results from *Hubble*+*Spitzer*. However, it is important to note that these observed samples are unlikely to be complete in  $M_*$ , and in many cases will be biased to high-specific star formation rates and low ages since these lead to higher luminosities. Since all of these samples are rest-frame UV selected, it is also possible they are missing heavily dust obscured systems and/or systems with strongly quenched star formation histories.

Salmon et al. (2015) studied the evolution and slope of the relationship between star formation activity and stellar mass in a sample of galaxies  $3.5 \leq z \leq 6.5$  using multi-wavelength photometry in GOODS-S from the Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey (CANDELS; Grogin et al. 2011; Koekemoer et al. 2011) and *Spitzer* Extended Deep Survey (Ashby et al. 2013). As shown in Fig. 2, the individual binned median values in most cases match our predictions within the uncertainties. However, while our results show little variation with stellar mass these tend to decrease as a function of stellar mass.

Endsley et al. (2021) utilized a selection combining narrow-band and overlapping broad-band filters yielding precise photometric redshifts, essential for constraining the contribution of strong [OIII] and H $\beta$  line emission. Using this sample, Whitter et al. (2022) measured ages and specific star formation rates utilizing both the *Prospector* (Leja et al. 2017; Johnson et al. 2021) and BEAGLE (Chevallard & Charlot 2016) spectral energy distribution (SED) fitting codes. For the BEAGLE analysis, a constant SFH model while the *Prospector* analysis explored both constant and non-parametric star formation histories. As seen in Fig. 2, both constant SFH implementations yielded specific star formation rates at  $M_* \sim 10^{9.5} M_\odot$  similar to the FLARES predictions. However, both produced steeply declining specific star formation rates as a function of stellar mass, contrary to the mostly flat relationship predicted by FLARES. Specific star formation rates measured using *Prospector* assuming non-parametric SFHs yielded a flatter, but still declining, relationship and lower overall normalization, falling below the FLARES predictions. The observed trends in the age stellar mass relation (Fig. 5) mirror the trends in specific star

formation rate: the average ages at  $M_* \approx 10^{9.5} M_\odot$  are similar to the FLARES predictions but diverge at lower and higher masses. The non-parametric analysis resulted in larger ages and a much shallower trend with stellar mass, albeit with ages offset to larger values than predicted by FLARES. Beyond a model issue, one possibility for the difference in the slope of observations relative to the FLARES predictions is the impact of dust. While the SED modelling takes account of dust, it is possible that the obscured star formation in the most massive galaxies has been underestimated resulting in systematically low-stellar masses.

Tacchella et al. (2022) studied the stellar populations of a sample of 11 bright galaxies at  $z = 9–11$  selected from the CANDELS fields (Finkelstein et al. 2022b). These galaxies were analysed using the *Prospector*, using a range of non-parametric and parametric star formation history priors. Our predictions are bracketed by the Tacchella et al. (2022) measurements with those assuming a continuity or parametric prior falling below our predictions and those with a bursty prior above. In addition to specific star formation rates, Tacchella et al. (2022) also measured the ages of their sample of galaxies. Unsurprisingly, these yield a similar trend to what was found for the specific star formation rates: FLARES predictions are bracketed by measurements made assuming various priors.

Topping et al. (2022) measured the specific star formation rates of specific of 40 UV-bright galaxies at  $z \sim 7–8$  combining far-IR continuum and [C II] constraints from ALMA with observations from *Hubble* and *Spitzer*. Topping et al. (2022) also utilized both *Prospector* and BEAGLE, though recovered similar average specific star formation rates. Both measurements are consistent with FLARES within their uncertainties.

Carnall et al. (2022) studied five spectroscopically confirmed ( $z = 5–9$ ) lensed (lensing factor = 1.5–10) in the SMACS0723 Early Release Observation (ERO) NIRCcam imaging. Carnall et al. (2022) fit the observed NIRCcam  $F090W$ ,  $F150W$ ,  $F200W$ ,  $F277W$ ,  $F356W$ , and  $F444W$  fluxes using the *Bagpipes* SED fitting code (Carnall et al. 2018) obtaining very young ages: four objects have mean stellar ages  $\leq 2$  Myr and one has  $\approx 20$  Myr. Because these objects are both relatively faint and lensed, 3/5 have stellar masses fall outside our predictions. However, the two that do fall within our range are, unlike most other observational measurements, discrepant



with our predictions. This may suggest FLARES fails to fully capture the full range of possible star formation histories or reflects an observational/modelling issue.

Naidu et al. (2022) identified two bright  $z > 10$  candidates in the Grism Lens-Amplified Survey from Space (GLASS) Early Release Science (ERS) imaging data based on a search of both GLASS and CEERS NIRCcam observations. The SEDs of these objects were then fit with `Prospector` code assuming a continuity prior. The resulting specific star formation rates fall slightly below our predictions at their respective redshifts, though are consistent within the observational uncertainties and predicted scatter.

Leethochawalit et al. (2022) identified and studied a sample of 14 galaxies at  $7 < z < 9$  in GLASS Early Release Science (ERS) NIRCcam imaging and measured their star formation histories using `Bagpipes`. Most of the inferred ages and specific star formation rates are, generally, in good agreement with our predictions, albeit with a slight (0.2 dex) offset to higher ages and lower specific star formation rates.

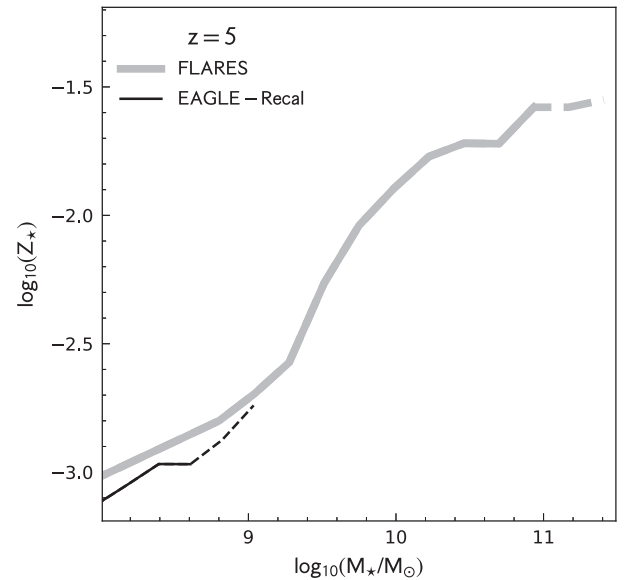
Finkelstein et al. (2022a) identified a sole  $z > 13$  candidate, ‘Maisie’s Galaxy’ at  $z \approx 14.3$  in CEERS NIRCcam imaging. Finkelstein et al. (2022a) measure stellar masses, ages, and specific star formation rates using `Prospector` finding an age of  $16_{-5}^{+45}$  Myr, consistent with our predictions.

Chen et al. (2022) study a sample of 12 galaxies at  $6 < z < 8$  using CEERS NIRCcam imaging. These sources were previously identified from *Hubble* observations of the EGS field. Chen et al. (2022) measured the star formation histories of these galaxies using `Beagle`, where the stellar masses of the Chen et al. (2022) sample overlap with our predictions. There is good agreement, however, the galaxies at lower stellar masses have smaller ages than expected from an extrapolation of the FLARES trend. Like the Carnall et al. (2022) study, this may mean that FLARES fails to predict the full range of possible star formation history scenarios.

These comparisons reveal a mixed picture with some tentative evidence for differences in the slope and normalization of the specific star formation rate – stellar mass relation. Beyond a model issue, one potential observational solution, at least at high-masses, is the effect of dust; all of these observational constraints are based on UV selected sources and, except for the case of Topping et al. (2022), utilize only rest-frame UV and optical observations. This may not only mean that dusty intensely star forming galaxies are missing but that the contribution from dust obscured star formation in the most massive galaxies is underestimated. Furthermore, it is clear from both the Tacchella et al. (2022) and Whitler et al. (2022) analyses, in addition to other (e.g. Carnall et al. 2019; Leja et al. 2019) studies, that these constraints are strongly sensitive to modelling assumptions, such as the choice of star formation history parametrization and the priors applied. On the modelling side, the relatively low resolution of FLARES limits the low-mass galaxy populations we can model at very high redshifts  $z > 10$ . The EAGLE model also ignores the effects of radiative feedback and reionization, which may have a particularly large effect on low-mass haloes. It is promising, however, that a model calibrated at  $z = 0$  shows such promising agreement with a range of observational constraints in this high-redshift domain.

## 4 METAL ENRICHMENT

We now turn our attention to the predictions for the mass-weighted stellar metallicity  $Z_*$  of galaxies. In this work, we focus on stellar metallicities  $Z_*$  deferring predictions for gas-phase metallicities, including predictions for observable line ratios etc. to a future work.



**Figure 10.** Comparison between the stellar mass–metallicity relation predicted by FLARES and the EAGLE-ReCal model at  $z = 5$ .

### 4.1 Earlier work and limitations

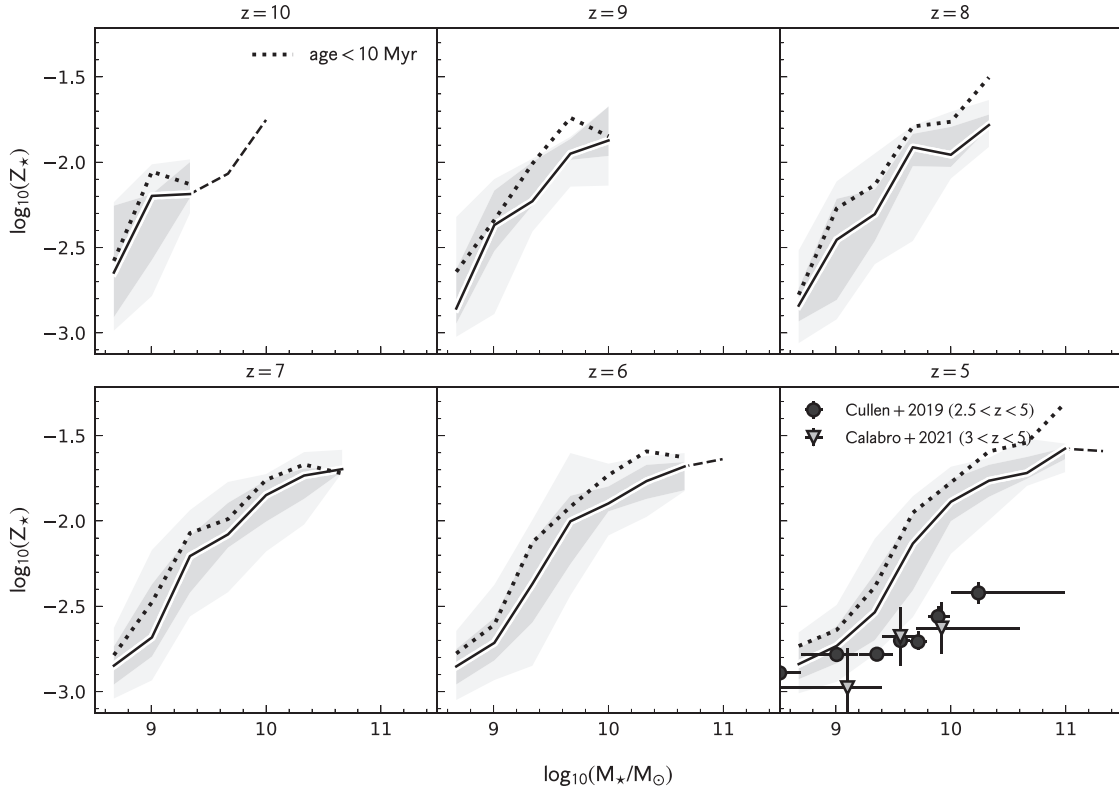
The stellar mass–metallicity relation predicted by the EAGLE model at  $z = 0$  was presented in section 6.3 of Schaye et al. (2015), with this analysis, subsequently, extended by De Rossi et al. (2017) to  $z = 3$ . In particular, Schaye et al. (2015) explored predictions for the reference (Ref), AGNdt9, and higher-resolution recal (ReCal) model. While there is a good agreement for high-mass ( $M_* > 10^{10} M_\odot$ ) galaxies, stellar metallicities diverge at lower stellar masses with the discrepancy maximized at  $M_* = 10^{8-9} M_\odot$  such that galaxies in the Ref/AGNdt9 variant simulations have stellar metallicities 0.2–0.3 dex larger than those in the ReCal simulation. To understand whether this offset still exists at  $z = 5$ , in Fig. 10, we compare the ReCal and FLARES stellar mass–metallicity relations at  $z = 5$ . While the small volume of the ReCal simulation compared to FLARES precludes a comparison at  $M_* > 10^9 M_\odot$ , at  $M_* = 10^{8-9} M_\odot$ , the offset is only  $\approx 0.1$  dex, giving some confidence that our predictions at higher masses will be robust.

### 4.2 Mass–metallicity relation

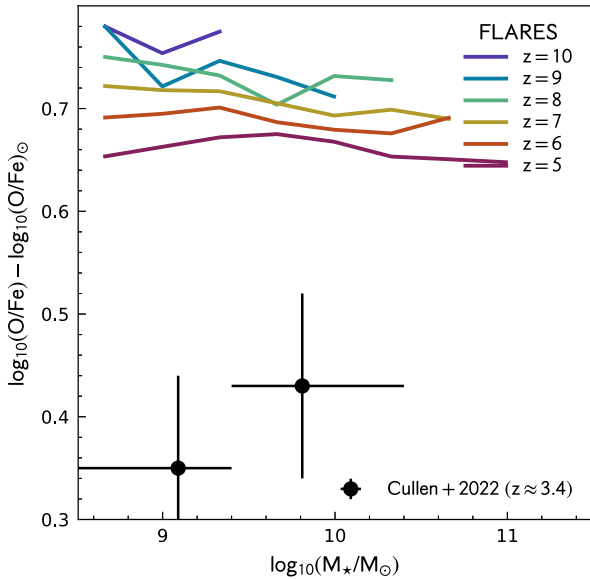
We next present the FLARES stellar mass–metallicity relation at  $z = 5$ –10 in Fig. 11. Unlike the age, stellar metallicities exhibit a strong dependence on stellar mass, increasing by a factor of 10 over the mass range  $10^9 \rightarrow 10^{11} M_\odot$ . Fig. 11 also includes the median stellar metallicity for the youngest (age  $< 10$  Myr) star particles within each galaxy. The metallicity of young stars follows the same trend with stellar mass as the general stellar mass–metallicity relation but is offset to higher metallicity by  $\approx 0.1$ –0.2 dex, reflecting the correlation between metallicity and age explored in more detail in Section 4.3.

#### 4.2.1 $\alpha$ -enhancement

In young, star forming galaxies, such as those at high-redshift, the dominant mode of chemical enrichment will be via type II (core-collapse) supernovae (SNII). Consequently, early galaxies should be ‘ $\alpha$ -enhanced’ relative to the galaxies present in the later Universe. In Fig. 12, we show the  $\alpha$ -enhancement, quantified as  $[O/Fe] =$



**Figure 11.** The same as Fig. 2 but for the total stellar metallicity. The dotted line shows the median stellar metallicity of only young (<10 Myr) star particles. In the  $z = 5$  panel, we show observational constraints, at  $2.5 < z < 5$ , from Cullen et al. (2019) and Calabrò et al. (2021).

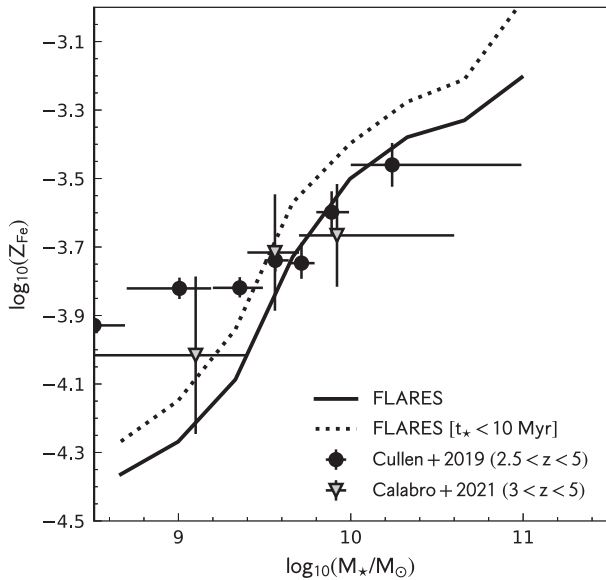


**Figure 12.** The predicted redshift evolution of  $\alpha$ -enhancement of galaxies in FLARES. Also included are recent constraints from Cullen et al. (2021) at  $z \approx 3.4$ .

$\log_{10}(\text{O}/\text{Fe}) - \log_{10}(\text{O}/\text{Fe})_{\odot}$ , as a function of stellar mass and redshift. At  $z = 10$ , we predict  $[\text{O}/\text{Fe}] \approx 0.75$ , falling to  $\approx 0.65$  at  $z = 5$ . The relationship with stellar mass is mostly flat, mirroring the trends seen in Section 3 for the age and specific star formation rate.

#### 4.2.2 Comparison to observational constraints

At present, most galaxy metallicity constraints come from observations of optical emission lines integrated across galaxies. These are sensitive to the gas-phase oxygen abundance  $(\text{O}/\text{H})_{\text{g}}$ . Prior to *JWST*, measurements were limited to  $z \approx 4$  (e.g. Sanders et al. 2021), however, thanks to *JWST*'s near-infrared spectroscopic capabilities, the first constraints are emerging to  $z \sim 8$  (e.g. Trump et al. 2022). It is, however, possible to constrain stellar metallicities from rest-frame observations of the UV continuum, where features sensitive to the stellar photospheric iron abundance  $(\text{Fe}/\text{H})$  exist (Leitherer et al. 2010). Recent efforts (e.g. Cullen et al. 2019; Calabrò et al. 2021) have pushed these measurements to  $z > 2$  using deep ground-based near-infrared spectroscopy. Constraints on the stellar mass-metallicity relation from Cullen et al. (2019) and Calabrò et al. (2021) are presented in Fig. 11 alongside our predictions. At  $M_{*} \sim 10^{10} M_{\odot}$ , our predictions are up to  $5 \times$  higher than both Cullen et al. (2019) and Calabrò et al. (2021). However, both Cullen et al. (2019) and Calabrò et al. (2021) are predominantly sensitive to the iron abundance  $(\text{Fe}/\text{H})$ , not the total stellar metallicity. In both cases, the total stellar metallicity was inferred by extrapolation using a solar abundance pattern. As noted above, however, due to the relative lack of enrichment from Type Ia supernovae, our abundances are extremely  $\alpha$ -enhanced. If we compare the iron mass fractions directly – using the solar abundances of Asplund et al. (2009) to convert the Cullen et al. (2019) and Calabrò et al. (2021) measurements to an iron mass fraction – we find greatly improved agreement with the FLARES predictions, at least at the high-mass end (see Fig. 13). In Fig. 12, we also compare our  $\alpha$ -enhancement predictions with Cullen et al. (2021) who combined stellar and gas-phase metallicity measurements of the same galaxies at  $z \approx 3.4$  to constrain  $[\text{Fe}/\text{O}]$ .



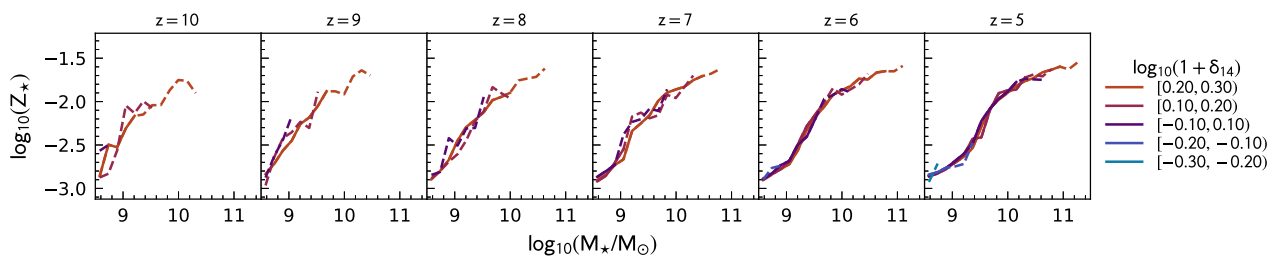
**Figure 13.** The predicted mass fraction of iron in stars ( $Z_{*,\text{Fe}}$ ) as a function of stellar mass from FLARES and the observations of Cullen et al. (2019) and Calabrò et al. (2021). The dotted line shows only the mass fraction of only young (<10 Myr) star particles.

These results are  $\approx 0.25$  dex below our predictions, though this is likely to be at least partially reconcilable with subsequent redshift evolution.

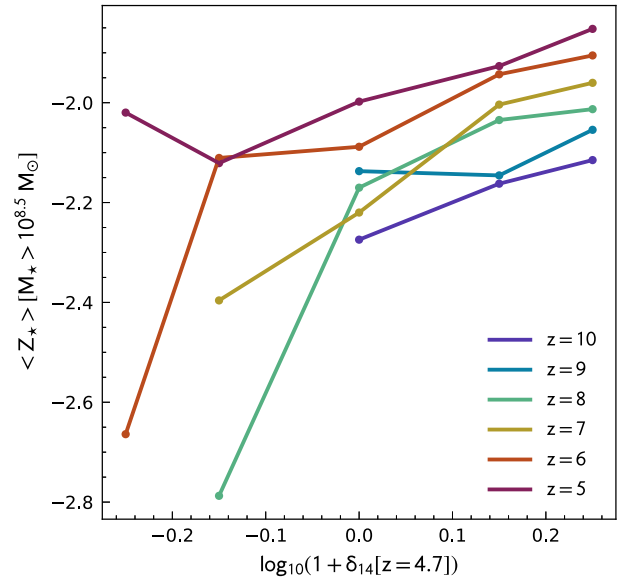
With *JWST*, now, beginning to collect deep near-infrared spectroscopy of high-redshift galaxies observation constraints on both  $Z_*$  and  $Z_g$  will soon dramatically improve. This in turn will begin to provide critical constraints to galaxy formation models.

#### 4.2.3 Environmental dependence

As demonstrated in Fig. 14 and as seen for the age and specific star formation rate (see Section 3.4, Fig. 9), we see no evidence of a large-scale environmental dependence on the stellar mass–metallicity relationship. This is not to say however that higher-density environments are not more metal enriched, as demonstrated in Fig. 15 they are. This figure shows a clear trend such that the average stellar metallicity is higher in more dense environments. This figure also shows clear redshift evolution such that average metallicity increases. The reconciliation of this and Fig. 14 is that higher-density environments, and lower-redshifts, contain a larger fraction of their mass in more massive, metal-rich galaxies.



**Figure 14.** The same as Fig. 9 but for the metallicity.



**Figure 15.** The average mass-weighted metallicity of galaxies with  $M_* > 10^{8.5} M_\odot$  as a function of environment.

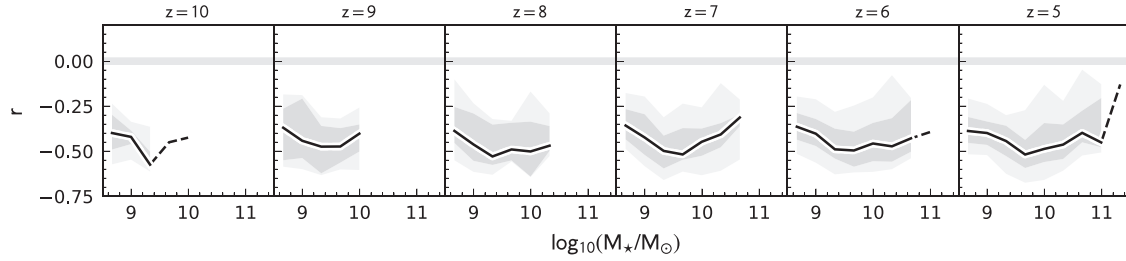
#### 4.3 Correlation with age

As preceding generations stars return enriched material to the interstellar medium, subsequent generations are, in general, expected to become increasingly enriched with metals. Consequently, we should expect a negative correlation between the stellar age and metallicity. To test this, we calculate the Pearson correlation coefficient  $r$ , finding (see Fig. 16) typical values of  $\approx -0.5$  suggesting a weak negative correlation.

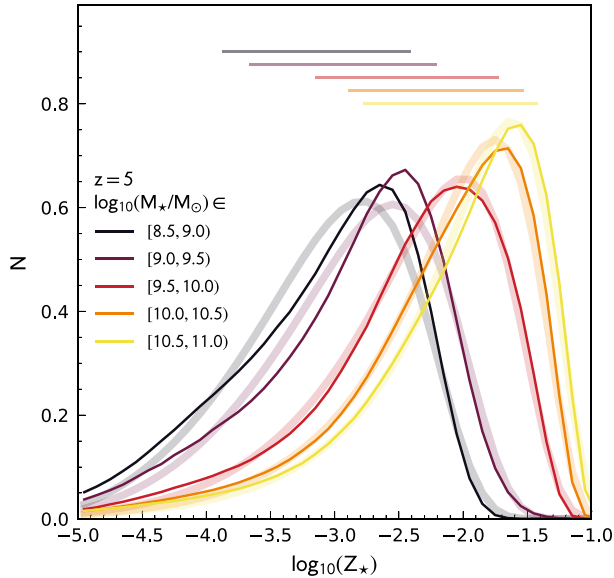
#### 4.4 Distribution

As with ages, stellar populations in galaxies will have a range of metallicities. Fig. 17 shows the initial mass weighted distribution of stellar metallicities for galaxies stacked by galactic stellar mass. This reveals that galaxies have a broad predicted range of metallicities with the central 68.4 per cent range of  $Z_*$  extending over more than 1 dex. This is also seen for individual galaxies as shown in Fig. 18, where the average range is 1–1.5 dex. This figure also reveals a trend with stellar mass such that the most massive galaxies have, on average, higher and narrower metallicity distributions.

Fig. 17 also reveals that the distribution of metallicities is asymmetric with a long tail to low metallicity. While this distribution can be fit adequately by a range of standard distributions, none have been found to give consistently good ( $D < 0.1$ ) results across the



**Figure 16.** The results of linear fitting to the age and  $\log_{10}(Z_*)$  of star particles in each galaxy. The top row shows the best-fitting slope, the middle row the age = 0 intercept, and the final row the Pearson correlation coefficient  $r$ .



**Figure 17.** The distribution of stellar metallicities for stacks of galaxies in stellar mass. The horizontal lines denote the central 68.4 per cent ( $P_{84.2} - P_{15.8}$ ) range of each distribution. The faint thick lines are parametric fits to predicted distribution and assume the (two parameter) exponential power distribution.

full range of stellar mass stacks and redshifts. Fig. 17 presents fits assuming the two parameter exponential power distribution,<sup>2</sup> while this yields good results in the larger stellar mass bins, the fit becomes progressively worse for the smaller stellar mass bins.

The fact that individual galaxies have stellar metallicities spanning a wide range, alongside a correlation with age, will have implications for the extent to which it is possible to constrain metallicities and other physical properties from observations using spectral energy distribution fitting. While some codes allow the metallicity to evolve, either through a parametric form (e.g. Robotham et al. 2020), or in independent bins (e.g. Tojeiro et al. 2007; Johnson et al. 2021), none that we are aware of explicitly allows the modelling of metallicity distributions, which may lead to significant biases.

## 5 CONCLUSIONS

In this work, we have presented predictions for the star formation and metal enrichment histories of galaxies at  $z = 5$ –

10 from the FLARES project. Our main conclusions are as follows:

(i) We find that specific star formation rates strongly evolve with redshift and exhibit a weak negative trend with stellar mass at  $M_* > 10^9 M_\odot$ . These predictions are in broadly good agreement with many current observational constraints, including early results from *JWST*. However, there is a tentative evidence of a difference in the slope of the relation and larger scatter present in the observations.

(ii) As redshift decreases, an increasing fraction of massive ( $M_* > 10^{10} M_\odot$ ) galaxies have relatively low-specific star formation rates. This is attributed to AGN feedback and is the focus of a companion work.

(iii) The individual star formation histories of galaxies in FLARES are generally rising. However, the fraction with rising SFHs declines with redshift and stellar mass. Star formation histories are well described by a (two parameter) truncated normal or truncated lognormal distribution.

(iv) There is a well defined stellar mass–metallicity relation in place at  $z = 10$  and beyond (for  $z > 10$  see Wilkins et al. 2022). Galaxies at these redshifts are also predicted to be very  $\alpha$ -enhanced with  $[\text{Fe}/\text{H}] \approx 0.65$  at  $z = 5$ . On the surface, these predictions are discrepant with recent metallicity measurements based on rest-frame observations of galaxies at  $z = 2.5$ –5 (Cullen et al. 2019; Calabrò et al. 2021). However, the inferred iron mass fraction, which these observations are sensitive to, are much similar, particularly at high-masses where we have the most confidence in the models and observations.

(v) Within individual galaxies, the metallicities of individual star particles are both inversely correlated with age and span a wide range, typically 1–1.5 dex.

(vi) We find no evidence of large-scale environmental dependence of the relationship between stellar mass and metallicity, star formation, or age. However, this is not to say that galaxy formation is not sensitive to environment: the densest regions contain a (much) larger fraction of their stellar mass in massive metal enriched galaxies making these environments more metal enriched than lower density regions.

At present our predictions are *mostly* consistent with existing observational constraints, including the handful of observations to emerge so far from *JWST*. However, at least in part, this current good agreement reflects the large observational uncertainties due to small samples, limited  $>2 \mu\text{m}$  observations (particularly spectroscopy), and observational modelling assumptions (e.g. the choice of stellar population synthesis model, or star formation history parametrization). New observations from *JWST*, and in particular deep rest-frame optical spectroscopy, may soon challenge these predictions.

<sup>2</sup><https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.exponpow.html>



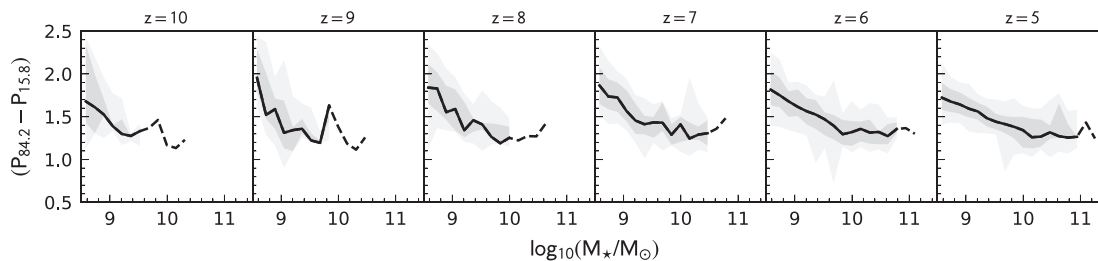


Figure 18. The same as Fig. 2 but for the central 68.4 per cent range of stellar metallicities.

## ACKNOWLEDGEMENTS

We thank the EAGLE team for their efforts in developing the EAGLE simulation code. This work used the DiRAC@Durham facility managed by the Institute for Computational Cosmology on behalf of the STFC DiRAC HPC Facility ([www.dirac.ac.uk](http://www.dirac.ac.uk)). The equipment was funded by BEIS capital funding via STFC capital grants ST/K00042X/1, ST/P002293/1, ST/R002371/1 and ST/S002502/1, Durham University and STFC operations grant ST/R000832/1. DiRAC is part of the National e-Infrastructure. The EAGLE simulations were performed using the DiRAC-2 facility at Durham, managed by the ICC, and the PRACE facility Curie based in France at TGCC, CEA, Bruyeres-le-Chatel. CCL acknowledges support from the Royal Society under grant RGF/EA/181016. DI acknowledges support by the European Research Council via ERC Consolidator Grant KETJU (no. 818930). The Cosmic Dawn Center (DAWN) is funded by the Danish National Research Foundation under grant No. 140. EZ acknowledges funding from the Swedish National Space Agency. We also wish to acknowledge the following open source software packages used in the analysis: NUMPY (Harris et al. 2020), SCIPY (Virtanen et al. 2020), and MATPLOTLIB (Hunter 2007). This research made use of ASTROPY <http://www.astropy.org> a community-developed core PYTHON package for Astronomy (Astropy Collaboration et al. 2013, 2018). Parts of the results in this work make use of the colormaps in the CMASHER package (van der Velden 2020).

## DATA AVAILABILITY

Binned data (medians and ranges) are available in the `astropy` Enhanced Character Separated Values (.ecsv) table format at [https://github.com/stephenmwilkins/flares\\_sfzh\\_data](https://github.com/stephenmwilkins/flares_sfzh_data) and as part of the wider First Light and Assembly of Galaxies model predictions repository available at [https://github.com/stephenmwilkins/flags\\_data](https://github.com/stephenmwilkins/flags_data). Data from the wider FLARES project is available at <https://flaresimulations.github.io/data.html>. If you use data from this paper, please also cite Lovell et al. (2021) and Vijayan et al. (2021). The data from the eagle simulations is publicly available at <http://icc.dur.ac.uk/Eagle/database.php> (see McAlpine et al. 2016; The EAGLE team 2017).

## REFERENCES

Adams N. J., Bowler R. A. A., Jarvis M. J., Häußler B., Lagos C. D. P., 2021, *MNRAS*, 506, 4933  
 Adams N. J. et al., 2022, *MNRAS*, tmp.3152A  
 Ashby M. L. N. et al., 2013, *ApJ*, 769, 80  
 Asplund M., Grevesse N., Sauval A. J., Scott P., 2009, *ARA&A*, 47, 481  
 Astropy Collaboration et al., 2013, *A&A*, 558, A33  
 Astropy Collaboration et al., 2018, *AJ*, 156, 123

Barnes D. J. et al., 2017, *MNRAS*, 471, 1088  
 Calabrò A. et al., 2021, *A&A*, 646, A39  
 Carnall A. C., McLure R. J., Dunlop J. S., Davé R., 2018, *MNRAS*, 480, 4379  
 Carnall A. C. et al., 2019, *MNRAS*, 490, 417  
 Carnall A. C. et al., 2022, *MNRAS*, 518, L45  
 Chabrier G., 2003, *PASP*, 115, 763  
 Chen Z., Stark D. P., Endsley R., Topping M., Whittler L., Charlot S., 2022, preprint ([arXiv:2207.12657](https://arxiv.org/abs/2207.12657))  
 Chevallard J., Charlot S., 2016, *MNRAS*, 462, 1415  
 Crain R. A. et al., 2015, *MNRAS*, 450, 1937  
 Cullen F. et al., 2019, *MNRAS*, 487, 2038  
 Cullen F. et al., 2021, *MNRAS*, 505, 903  
 Davidzon I. et al., 2017, *A&A*, 605, A70  
 Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, *ApJ*, 292, 371  
 de Barros S., Schaerer D., Stark D. P., 2014, *A&A*, 563, A81  
 De Rossi M. E., Bower R. G., Font A. S., Schaye J., Theuns T., 2017, *MNRAS*, 472, 3354  
 Diemer B., Sparre M., Abramson L. E., Torrey P., 2017, *ApJ*, 839, 26  
 Dolag K., Borgani S., Murante G., Springel V., 2009, *MNRAS*, 399, 497  
 Driver S. P. et al., 2022, *MNRAS*, 513, 439  
 Endsley R., Stark D. P., Chevallard J., Charlot S., 2021, *MNRAS*, 500, 5229  
 Eyles L. P., Bunker A. J., Ellis R. S., Lacy M., Stanway E. R., Stark D. P., Chiu K., 2007, *MNRAS*, 374, 910  
 Finkelstein S. L. et al., 2022a, preprint ([arXiv:2207.12474](https://arxiv.org/abs/2207.12474))  
 Finkelstein S. L. et al., 2022b, *ApJ*, 928, 52  
 Grogin N. A. et al., 2011, *ApJS*, 197, 35  
 Harris C. R. et al., 2020, *Nature*, 585, 357  
 Hunter J. D., 2007, *Comput. Sci. Engineer.*, 9, 90  
 Johnson B. D., Leja J., Conroy C., Speagle J. S., 2021, *ApJS*, 254, 22  
 Kennicutt R. C. Jr, 1998, *ApJ*, 498, 541  
 Koekemoer A. M. et al., 2011, *ApJS*, 197, 36  
 Laporte N., Meyer R. A., Ellis R. S., Robertson B. E., Chisholm J., Roberts-Borsani G. W., 2021, *MNRAS*, 505, 3336  
 Leethochawalit N. et al., 2022, preprint ([arXiv:2207.11135](https://arxiv.org/abs/2207.11135))  
 Leitherer C., Ortiz Otálvaro P. A., Bresolin F., Kudritzki R.-P., Lo Faro B., Pauldrach A. W. A., Pettini M., Rix S. A., 2010, *ApJS*, 189, 309  
 Leja J., Johnson B. D., Conroy C., van Dokkum P. G., Byler N., 2017, *ApJ*, 837, 170  
 Leja J., Carnall A. C., Johnson B. D., Conroy C., Speagle J. S., 2019, *ApJ*, 876, 3  
 Lovell C. C., Vijayan A. P., Thomas P. A., Wilkins S. M., Barnes D. J., Irodoutou D., Roper W., 2021, *MNRAS*, 500, 2127  
 Marigo P., 2001, *A&A*, 370, 194  
 McAlpine S. et al., 2016, *Astron. Comput.*, 15, 72  
 McLeod D. J., McLure R. J., Dunlop J. S., Cullen F., Carnall A. C., Duncan K., 2021, *MNRAS*, 503, 4413  
 Naidu R. P. et al., 2022, *ApJL*, 940, L14  
 Planck Collaboration et al., 2014, *A&A*, 571, A1  
 Portinari L., Chiosi C., Bressan A., 1998, *A&A*, 334, 505  
 Robotham A. S. G., Bellstedt S., Lagos C. d. P., Thorne J. E., Davies L. J., Driver S. P., Bravo M., 2020, *MNRAS*, 495, 905  
 Roper W. J., Lovell C. C., Vijayan A. P., Marshall M. A., Irodoutou D., Kuusisto J. K., Thomas P. A., Wilkins S. M., 2022, *MNRAS*, 514, 1921

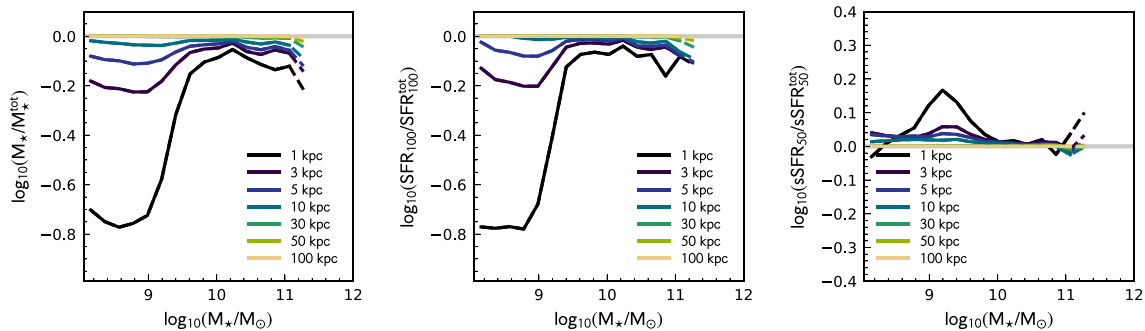


- Salmon B. et al., 2015, *ApJ*, 799, 183  
 Sanders R. L. et al., 2021, *ApJ*, 914, 19  
 Schaye J., 2004, *ApJ*, 609, 667  
 Schaye J., Dalla Vecchia C., 2008, *MNRAS*, 383, 1210  
 Schaye J. et al., 2010, *MNRAS*, 402, 1536  
 Schaye J. et al., 2015, *MNRAS*, 446, 521  
 Smit R. et al., 2015, *ApJ*, 801, 122  
 Song M. et al., 2016, *ApJ*, 825, 5  
 Springel V., White S. D. M., Tormen G., Kauffmann G., 2001, *MNRAS*, 328, 726  
 Stark D. P., Schenker M. A., Ellis R., Robertson B., McLure R., Dunlop J., 2013, *ApJ*, 763, 129  
 Tacchella S. et al., 2022, *ApJ*, 927, 170  
 The EAGLE team, 2017, preprint (arXiv:1706.09899)  
 Tojeiro R., Heavens A. F., Jimenez R., Panter B., 2007, *MNRAS*, 381, 1252  
 Topping M. W. et al., 2022, *MNRAS*, 516, 975  
 Trump J. R. et al., 2022, preprint (arXiv:2207.12388)  
 van der Velden E., 2020, *J. Open Source Softw.*, 5, 2004  
 Vijayan A. P., Lovell C. C., Wilkins S. M., Thomas P. A., Barnes D. J., Irodotou D., Kuusisto J., Roper W. J., 2021, *MNRAS*, 501, 3289  
 Virtanen P. et al., 2020, *Nat. Methods*, 17, 261  
 Whittler L., Stark D. P., Endsley R., Leja J., Charlot S., Chevillard J., 2022, preprint (arXiv:2206.05315)  
 Wiersma R. P. C., Schaye J., Theuns T., Dalla Vecchia C., Tornatore L., 2009, *MNRAS*, 399, 574  
 Wilkins S. M. et al., 2013, *MNRAS*, 435, 2885  
 Wilkins S. M. et al., 2020, *MNRAS*, 493, 6079  
 Wilkins S. M. et al., 2022, preprint (arXiv:2204.09431)

## APPENDIX A: MEASURING STELLAR MASSES AND STAR FORMATION RATES

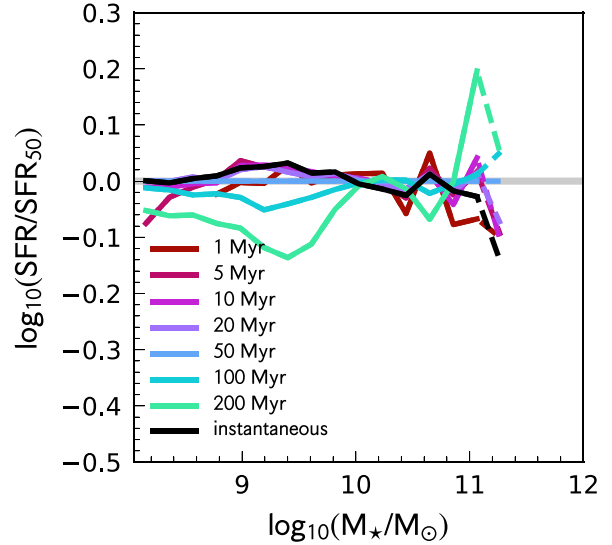
### A1 Apertures

Galaxy formation simulations identify objects as bound structures including haloes and subhaloes. However, to better match observational methodologies, it is common to measure galaxy properties in a spherical aperture. In both the core EAGLE and FLARES analyses 30 kpc diameter apertures were used. In Fig. A1, we briefly explore the consequences of this assumption on the stellar mass, star formation rate, and specific star formation rate of galaxies at  $z = 5$ . While our chosen aperture captures virtually all stellar mass and star formation, smaller apertures tend to miss both in the lowest and highest mass galaxies. However, despite this sensitivity the measured specific star formation rate is mostly insensitive to the choice of aperture.



**Figure A1.** The sensitivity of stellar mass (left), star formation rate (centre), and specific star formation rate (right) to the choice of aperture used to define the galaxy.

This paper has been typeset from a  $\text{\TeX}/\text{\LaTeX}$  file prepared by the author.



**Figure A2.** The impact of the star formation averaging time-scale on star formation rates.

### A2 Star formation measures

Within hydrodynamical simulations like FLARES there are two approaches for measuring star formation rates (SFRs): 1) from the properties of the gas particles themselves – the *instantaneous* SFR and 2) from the star particles by averaging over a particular time-scale. Short time-scales  $< 10$  Myr are limited by sampling effects: in low-SFR galaxies there may only be handful, or even no, star particles formed. While short time-scales are likely to match SF diagnostics based on reprocessed Lyman continuum emission (e.g.  $\text{H}\alpha$ ) the same is *possibly* not true for the FUV, and by extension far-IR, which have a contribution from stars up to several hundred million years old. In Fig. A2, we explore how the choice of averaging time-scale affects the SFR. At high-mass ( $M^* > 10^{10} M_{\odot}$ ) the choice of time-scale has little effect, the implication being that star formation histories in these galaxies are approximately constant. At lower masses however SFRs derived from longer time-scales are typically smaller suggesting rising star formation histories. As expected instantaneous SFRs derived from gas particles closely match short time-scales. Unless otherwise stated in this work, we choose to average on a 50 Myr time-scale.