

<https://helda.helsinki.fi>

First Light and Reionisation Epoch Simulations (FLARES) - VI. The colour evolution of galaxies $z=5-15$

Wilkins, Stephen M.

2022-12

Wilkins , S M , Vijayan , A P , Lovell , C C , Roper , W J , Irodotou , D , Caruana , J ,
Seeyave , L T C , Kuusisto , J K & Thomas , P A 2022 , ' First Light and Reionisation Epoch
Simulations (FLARES) - VI. The colour evolution of galaxies $z=5-15$ ' , Monthly Notices of the
Royal Astronomical Society , vol. 517 , no. 3 , pp. 3227-3235 . <https://doi.org/10.1093/mnras/stac2548>

<http://hdl.handle.net/10138/351317>

<https://doi.org/10.1093/mnras/stac2548>

cc_by

publishedVersion

Downloaded from Helda, University of Helsinki institutional repository.

This is an electronic reprint of the original article.

This reprint may differ from the original in pagination and typographic detail.

Please cite the original version.

First Light and Reionisation Epoch Simulations (FLARES) – VI. The colour evolution of galaxies $z = 5–15$

Stephen M. Wilkins^{1,2*}, Aswin P. Vijayan^{3,4}, Christopher C. Lovell⁵, William J. Roper¹, Dimitrios Irodotou^{1,6}, Joseph Caruana^{2,7}, Louise T. C. Seeyave¹, Jussi K. Kuusisto¹ and Peter A. Thomas¹

¹Astronomy Centre, University of Sussex, Falmer, Brighton BN1 9QH, UK

²Institute of Space Sciences and Astronomy, University of Malta, Msida MSD 2080, Malta

³Cosmic Dawn Center (DAWN)

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

⁵Centre for Astrophysics Research, School of Physics, Engineering and Computer Science, University of Hertfordshire, Hatfield AL10 9AB, UK

⁶Department of Physics, University of Helsinki, Gustaf Hällströmin katu 2, FI-00014 Helsinki, Finland

⁷Department of Physics, University of Malta, Msida MSD 2080, Malta

Accepted 2022 August 30. Received 2022 August 30; in original form 2022 July 22

ABSTRACT

With its exquisite sensitivity, wavelength coverage, and spatial and spectral resolution, the *James Webb Space Telescope* (JWST) is poised to revolutionize our view of the distant, high-redshift ($z > 5$) Universe. While *Webb*'s spectroscopic observations will be transformative for the field, photometric observations play a key role in identifying distant objects and providing more comprehensive samples than accessible to spectroscopy alone. In addition to identifying objects, photometric observations can also be used to infer physical properties and thus be used to constrain galaxy formation models. However, inferred physical properties from broad-band photometric observations, particularly in the absence of spectroscopic redshifts, often have large uncertainties. With the development of new tools for forward modelling simulations, it is now routinely possible to predict observational quantities, enabling a direct comparison with observations. With this in mind, in this work, we make predictions for the colour evolution of galaxies at $z = 5–15$ using the First Light And Reionisation Epoch Simulations (FLARES) cosmological hydrodynamical simulation suite. We predict a complex evolution with time, driven predominantly by strong nebular line emission passing through individual bands. These predictions are in good agreement with existing constraints from *Hubble* and *Spitzer* as well as some of the first results from *Webb*. We also contrast our predictions with other models in the literature: While the general trends are similar, we find key differences, particularly in the strength of features associated with strong nebular line emission. This suggests photometric observations alone should provide useful discriminating power between different models and physical states of galaxies.

Key words: galaxies: evolution – galaxies: formation – galaxies: general – galaxies: high-redshift – galaxies: photometry.

1 INTRODUCTION

The study of the distant, high-redshift ($z > 5$) Universe stands on the cusp of a revolution, thanks to the *James Webb Space Telescope*. *Webb*'s combination of infrared coverage, sensitivity, and spectroscopic capabilities should ultimately enable the accurate identification of statistical samples of star-forming galaxies to $z > 10$ (Robertson 2021), and the measurement of many key properties including star formation rates, stellar masses, metallicities, and rest-frame optical morphologies.

A pillar of *Webb*'s exploration of the distant Universe will be its broad-band photometric observations obtained by NIRC*am*, MIRI, and NIRISS. In cycle 1 alone, *Webb* will acquire $> 1 \text{ deg}^2$ of NIRC*am*

imaging, with the deepest observations approaching 31 mag. These observations will enable precise measurement of the rest-frame UV luminosity function, particularly at the faint end, allowing the determination of a faint-end turnover. Photometric observations will also enable the measurement of key physical properties such as stellar masses, star formation rates, ages, and dust attenuation. However, the physical properties inferred from photometric observations alone yield large uncertainties (see, e.g. Whitler et al. 2022), reducing their usefulness in terms of constraining models. One critical cause of this uncertainty is the impact of nebular line emission (see e.g. Zackrisson, Bergvall & Leitet 2008; Schaerer & de Barros 2009; Stark et al. 2013; Wilkins et al. 2013b), which can shift predicted colours by up to 1 mag (e.g. Wilkins et al. 2020).

However, with the development of sophisticated forward modelling pipelines – which are used to create synthetic observations (e.g. Camps & Baes 2015; Narayanan et al. 2021) – it is increasingly

* E-mail: S.Wilkins@sussex.ac.uk

possible to directly compare observed and simulated individual galaxies and populations. While there remain uncertain elements of this modelling – in particular the modelling of dust and nebular emission – we are now at the point where this forward modelling process is comparable, or perhaps even simpler, than the reverse. Efforts to produce synthetic observations are now ubiquitous in the modelling community (e.g. Wilkins et al. 2013b; Trayford et al. 2015; Wilkins et al. 2016b; Vogelsberger et al. 2020; Vijayan et al. 2021) with the ability to make observational predictions across the electromagnetic spectrum incorporating a range of physical processes, including nebular emission from HII regions (e.g. Wilkins et al. 2013b; Orsi et al. 2014; Wilkins et al. 2020; Vijayan et al. 2021) and the interstellar medium (ISM; e.g. Lagache, Cousin & Chatzikos 2018; Katz et al. 2019; Popping et al. 2019; Leung et al. 2020; Kannan et al. 2022), and dust attenuation (e.g. Wilkins et al. 2018; Vogelsberger et al. 2020; Vijayan et al. 2021, 2022) and emission (e.g. Wilkins et al. 2018; Ma et al. 2019; Lovell et al. 2021b; Vijayan et al. 2022).

Beyond integrated photometry/spectroscopy it is now also possible to produce synthetic imaging (e.g. Snyder et al. 2015; Trayford et al. 2015; Ma et al. 2018; Marshall et al. 2022; Roper et al. 2022), enabling self-consistent comparisons of morphological metrics. Forward modelling has been applied to models covering a wide range of scales, from high-resolution simulations of individual haloes (e.g. Ma et al. 2019) to the construction of synthetic lightcones encompassing hundreds of square arcminutes (e.g. Williams et al. 2018; Davidzon et al. 2019; Laigle et al. 2019; Somerville et al. 2021; Drakos et al. 2022; Yung et al. 2022), enabling a direct comparison with observed galaxy populations.

In this work, we make predictions for the redshift and luminosity evolution of *Hubble*, *Spitzer*, and *JWST* colours at $z \geq 5$ using the First Light And Reionisation Epoch Simulations (FLARES; Lovell et al. 2021a; Vijayan et al. 2021) cosmological hydrodynamical simulation suite. The current FLARES simulations combine the $z = 0$ validated EAGLE (Crain et al. 2015; Schaye et al. 2015) physics model with an innovative simulation strategy, resulting in a large effective volume, and thus dynamic range, of stellar masses and luminosities.

This paper is organized as follows: We begin, in Section 2, by exploring predictions for the evolution of NIRCcam and MIRI colours using a simple toy model. In Section 3, we then briefly describe the FLARES project before, in Section 4, presenting predictions including a comparison with existing observations (Section 4.1) and other models (Section 4.2). Finally, in Section 5, we present our conclusions.

2 THEORETICAL BACKGROUND

To obtain an understanding for the physical effects that drive the colour evolution of galaxies in the distant Universe, in this section, we explore predictions from a simple toy model utilizing simple star formation and metal enrichment histories and dust modelling.

In this toy model, composite spectral energy distributions (SEDs) are created by combining age/metallicity SED grids with a parametric star formation and metal enrichment history. Specifically, we employ the same stellar population synthesis model (SPS): version 2.2.1 of BPASS: Binary Population And Spectral Synthesis (BPASS; Stanway & Eldridge 2018), and initial mass function (IMF; Chabrier 2003), as used by FLARES. To account for nebular continuum and line emission, we process the pure stellar SED grids using the CLOUDY photoionization model (Ferland et al. 2017). This follows the same approach as Wilkins et al. (2020) and that utilized by FLARES. In short, each pure stellar SED is associated with a HII region with the

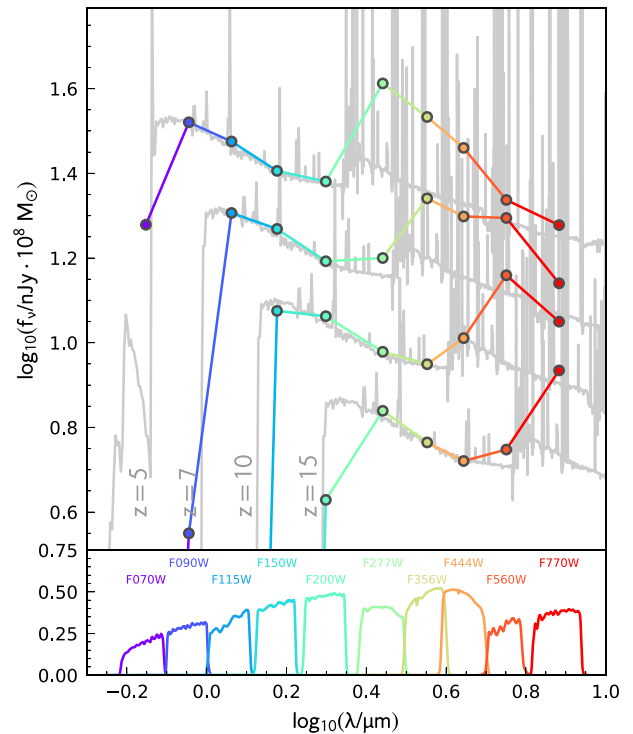


Figure 1. The observed SED of a star-forming galaxy at $z = 5\text{--}15$ alongside key *JWST*/NIRCam, and *JWST*/MIRI filter transmission functions. Coloured points denote the predicted fluxes in each of the NIRCam and MIRI bands, highlighting the impact of nebular emission in the rest-frame optical.

same metallicity with scale solar composition, a covering fraction of 1, and assuming a reference¹ ionization parameter of $\log_{10} U = -2$. A key difference between this modelling and FLARES, however, is the treatment of dust. In the absence of spatially resolved stellar populations and dust distributions in the toy model, we assume a simple screen model, while FLARES employs a line-of-sight model, in principle assigning a unique attenuation to every star particle.

We begin by using this model to generate spectra of an unobscured (i.e. $\tau_V = 0$) composite stellar population, with mass $M_* = 10^8 M_\odot$, a 100-Myr continuous star formation history, and metallicity $Z = 0.001$ at $z \in \{5, 7, 10, 15\}$, shown in Fig. 1. For this fiducial model, we assume $f_{\text{esc,LyC}} = 0$, i.e. maximizing the contribution of nebular emission. On this figure, we also add predicted broad-band fluxes for each of the NIRCam wide filters (*F070W*, *F090W*, *F115W*, *F150W*, *F200W*, *F277W*, *F356W*, *F444W*) in addition to the MIRI *F560W* and *F770W* bands. Immediately evident in this figure is the impact of the Lyman-limit/ α break, the largely smooth UV continuum, and the impact of strong nebular line emission, particularly from $[\text{O II}]\lambda\lambda 3726, 3729\text{\AA}$, $[\text{O III}]\lambda 5007\text{\AA}$, and $\text{H}\beta$. To further aid in the understanding of these predictions, in Fig. 2, we show the rest-frame wavelength probed by the same NIRCam and MIRI filters as a function of redshift, highlighting key emission lines and spectral features. This further reinforces that strong line emission will play an important role in driving NIRCam and MIRI colours at high redshift.

In the following sections, we discuss the implications of different model components on the colour evolution shown in Fig. 3.

¹This defines the ionization parameter at the reference age (1 Myr) and metallicity ($Z = 0.01$). The ionization parameter at other ages and metallicities are scaled according to the ionizing photon luminosity.

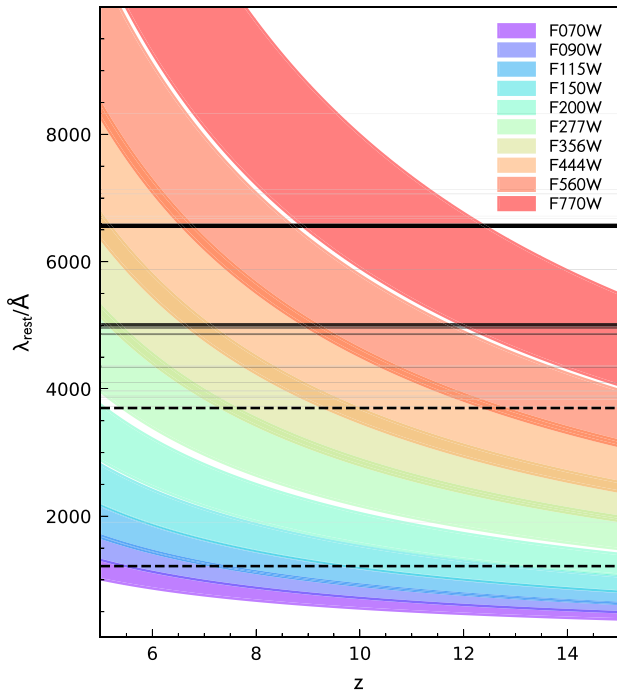


Figure 2. The rest-frame wavelength probed by selected NIRCam and MIRI filters at $z = 5$ –15. The two dashed horizontal lines denote the location of the Balmer (3646 Å) and Lyman α (1216 Å) break while solid lines denote strong nebular emission lines with line thickness and opacity indicating the equivalent width for a simple star-forming model.

2.1 Impact of nebular emission

Firstly, as denoted by the grey line, we show our fiducial model (100-Myr constant star formation, $Z = 0.001$) but with no reprocessing by dust (i.e. $\tau_V = 0$) or gas (i.e. $f_{\text{esc,LyC}} = 1$) – that is, pure stellar emission. The resulting colours evolve smoothly remaining relatively blue ($A - B \approx 0$), except when encompassing the Lyman or Balmer breaks (e.g. $F277W - F356W$ at $z \approx 7.5$) with the latter shifting the colour by up to ≈ 0.5 mag. We next show our fiducial model (as previous, but with $f_{\text{esc}} = 0$) as the solid black line. The result is a complex colour evolution with rapid changes, coinciding with strong line emission falling within one of the bands. In this model colours shift by up to 0.5 mag relative to pure stellar colours. This can also lead to colours changing by up to 0.7 mag across small redshift intervals (e.g. $F444W - F560W$ at $z \approx 9$). As we will see in Section 4, these shifts are even more pronounced when medium and wide filters are combined (e.g. $F430W - F444W$) with shifts up to 1 mag predicted for FLARES galaxies. Where the colour is probing the UV continuum, the result is also a shift to redder colours, caused by nebular continuum emission (Wilkins et al. 2013a).

2.2 Star formation and metal enrichment history

We next consider models with different star formation histories: a short 10-Myr episode of continuous star formation and a maximal-aged burst.² For the short burst, the result is to further enhance the impact of nebular line emission, due to the increased ratio of ionizing

²That is, a model in which the stellar age is equivalent to the age of the Universe at that redshift.

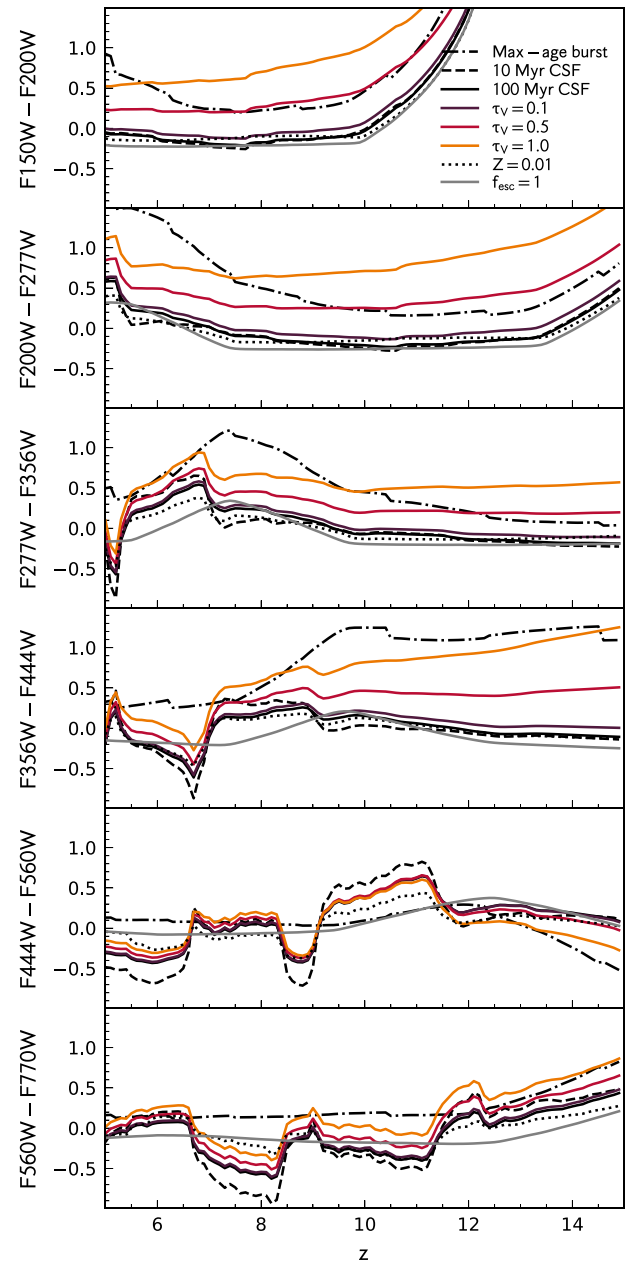


Figure 3. Predicted colours for a range of simple models. Our default model, denoted by the solid black line assumes 100 Myr constant star formation, nebular emissions assuming $f_{\text{esc}} = 0$, and no dust. Coloured solid lines show the same model but with increasing amount of dust attenuation. The solid grey line instead assumes no nebular emission ($f_{\text{esc}} = 1$). The dashed line assumes 10-Myr constant star formation. The dotted line assumes $Z = 0.01$. The dot dashed line denotes a maximal-aged burst of star formation, effectively the reddest intrinsic colour possible.

to optical photons for younger stellar populations. In this model colours now change by up to 1 mag over short redshift intervals. For the maximal-aged burst the resulting colours are consistently redder, due to the (increasing) lack of massive hot short-lived stars. As the ionizing photon luminosity drops rapidly in the first few million years (Wilkins et al. 2020), the contribution of nebular emission to these models is extremely limited resulting in little rapid evolution of the colour. In the rest-frame UV the shift from our fiducial model is

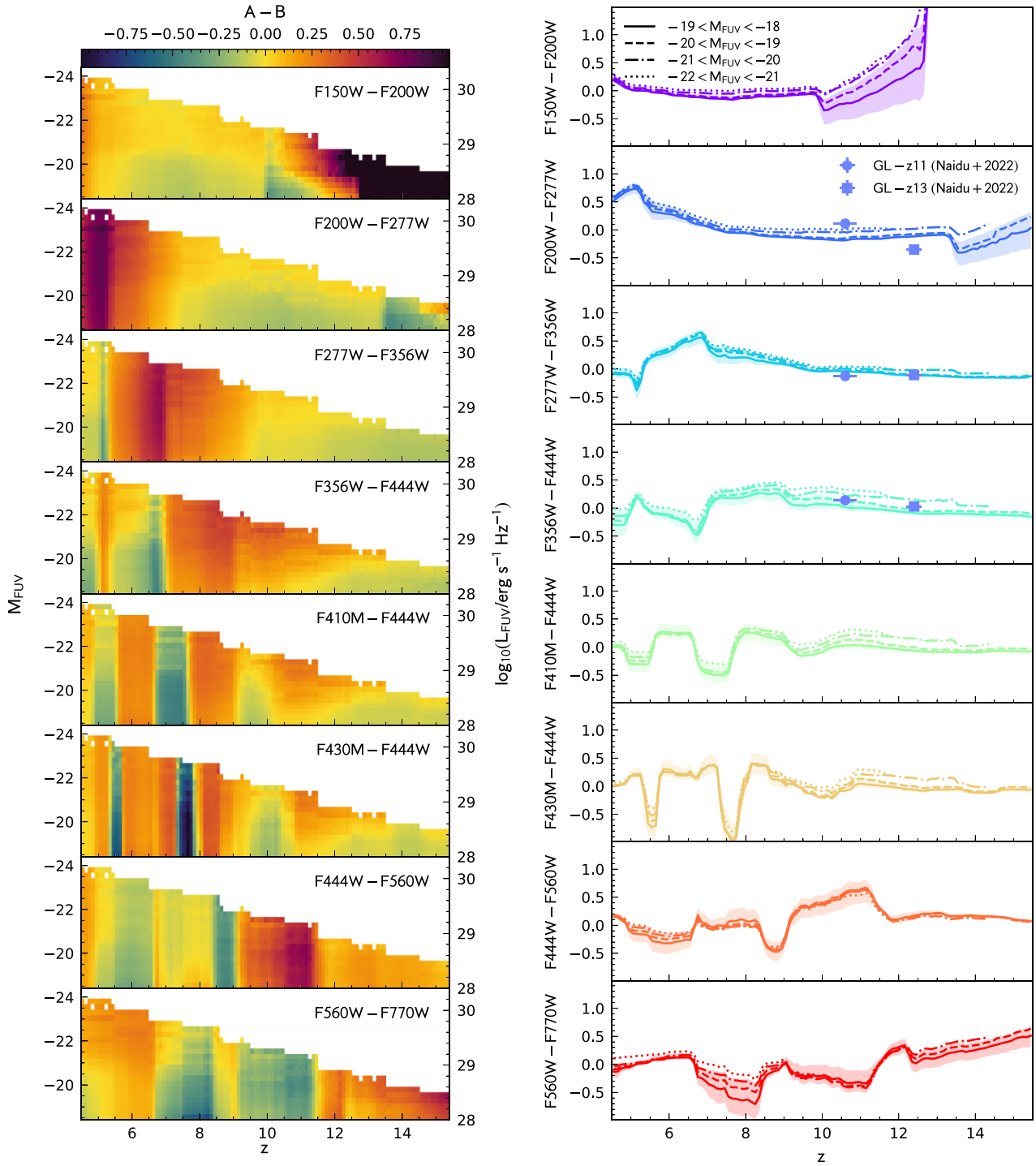


Figure 4. FLARES predictions for the evolution of NIRCam and MIRI colours of galaxies at $z = 5-15$. The left-hand side plot shows the average colour in bins of redshift and rest-frame far-UV luminosity. The right-hand side plot shows both the median colour for galaxies in four absolute magnitude bins and central 68 per cent range for the entire sample. Also included in the right-hand side figure are recent observational constraints from Naidu et al. (2022).

≈ 0.5 mag, growing to ≈ 1 mag when encompassing the age sensitive Balmer break feature.

As galaxies in FLARES span a range of metallicities (Wilkins et al. 2022b), we also explore the impact of increasing the metallicity of our fiducial model to $Z = 0.01$. Increasing the metallicity both makes the pure stellar SED redder and also reduces the ionizing

photon luminosity, reducing the overall contribution of nebular emission. Changing the metallicity will also impact individual line ratios leading to a complex impact on colours, especially when one or more strong line is present. However, for the most part, the effect of metallicity is fairly subtle, shifting colours by < 0.1 mag.

2.3 Reprocessing by dust

Finally, we consider the impact of dust, applying a screen model assuming a simple λ^{-1} attenuation law parametrized using the optical depth in the V band: τ_V . Because this simple model leaves nebular equivalent widths unchanged, where the colours are dominated by nebular emission (e.g. $F444W - F560W$ at $z = 5-10$), the result is only a weak shift to redder colours. In the rest-frame UV, however, the shift from the fiducial model can be more dramatic, with the colours increasing by ≈ 1 mag for a model with $\tau_V = 1$.

Together, the results in this section highlight the significant impact of many modelling assumptions on the colour evolution of galaxies in the epoch of reionization, and demonstrates how colours can be used as a key constraint on these assumptions.

3 THE FIRST LIGHT AND REIONISATION EPOCH SIMULATIONS

In this rest of this study, we make use of the core suite of simulations from the first phase of FLARES. The core suite and its initial processing are described in Lovell et al. (2021a) and Vijayan et al. (2021), while predictions at the redshift frontier ($z > 10$) are presented in Wilkins et al. (2022a), and we refer the reader to those articles for a detailed introduction. In short, the core FLARES suite is a set of 40 spherical re-simulations, $14 h^{-1} \text{cMpc}$ in radius, of regions selected from a large $(3.2 \text{cGpc})^3$ dark matter only simulation. The regions selected to re-simulate span a range of environments: (at $z \approx 4.7$) $\log_{10}(1 + \delta_{14}) = [-0.3, 0.3]^3$ with overrepresentation of the extremes of the density contrast distribution. As demonstrated in Lovell et al. (2021a), this strategy allows us to efficiently simulate a much larger dynamic range in mass (or luminosity) than a traditional periodic box for the same computational resources. We adopt the AGNdT9 variant of the EAGLE simulation project (Crain et al. 2015; Schaye et al. 2015) with identical resolution and cosmology to the fiducial EAGLE simulation. This allows us to resolve galaxies with stellar masses $M_* > 10^8 M_\odot$ corresponding to intrinsic rest-frame far-UV absolute magnitudes of $M_{\text{UV}} \lesssim -18.4$ ($L_{\text{FUV}} \gtrsim 10^{28} \text{erg s}^{-1} \text{Hz}^{-1}$) and $m \lesssim 29.1$ at $z = 10$.

3.1 SED modelling

To produce galaxy observables we process the outputs with a custom pipeline with the approach described in Vijayan et al. (2021), broadly following the approach developed by Wilkins et al. (2013b, 2016b, 2018, 2020), with modifications to the dust treatment. In short, we begin by associating each star particle with a *pure stellar* SED using v2.2.1 of the BPASS (Stanway & Eldridge 2018) stellar population synthesis model assuming a Chabrier (2003) IMF according to its age and metallicity. We then associate each star particle with an H II region giving rise to nebular continuum and line emission following the approach detailed in Wilkins et al. (2020). We then account for the effect of dust, both in the birth clouds of young stellar populations (with age less than 10 Myr, following Charlot & Fall 2000, that birth clouds disperse along these time-scales) and the wider ISM. For the latter, we employ a simple line-of-sight attenuation model, similar to that described in Wilkins et al. (2018), but using the fitting function for the dust-to-metal ratio presented in Vijayan et al. (2019) (equation 15 in their paper, which parametrizes the

³Where δ_{14} is the density contrast measured within the re-simulation volume size.

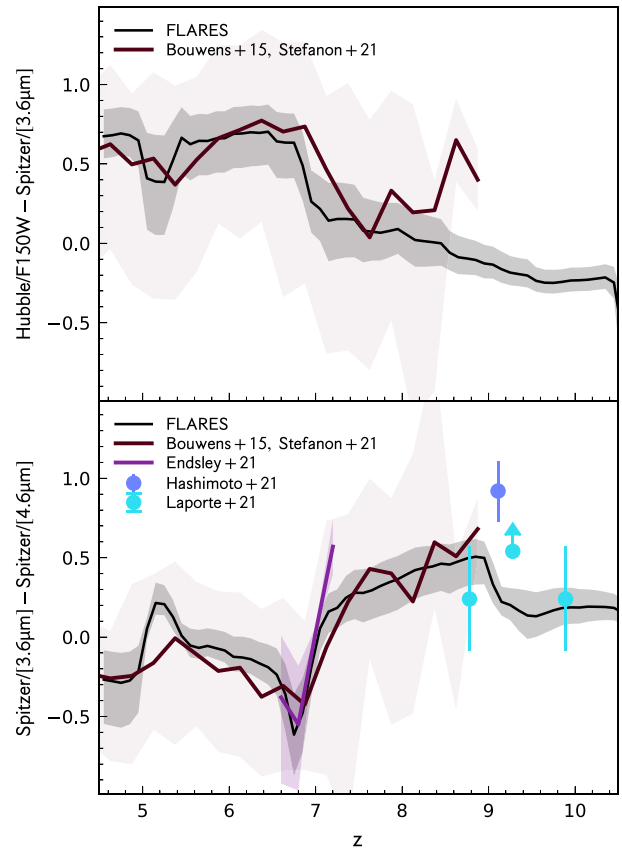


Figure 5. FLARES predictions, and observations, for the *Hubble*/WFC3/F160W – *Spitzer*/IRAC/[3.6] and *Spitzer*/IRAC/[4.5] – *Spitzer*/IRAC/[3.6] colour evolution of galaxies with $M_{\text{FUV}} < -18$. The shaded band shows the central 68 per cent range of all galaxies $M_{\text{FUV}} < -18$. The lines denote the median in the various M_{FUV} bins.

dust-to-metal ratio as a function of the mass-weighted age of the stellar population and the gas-phase metallicity). For the attenuation due to the birth cloud component, we scale it with the star particle metallicity, thus assuming a constant dust-to-metal ratio. For more details, see section 2.4 in Vijayan et al. (2021).

As demonstrated in Fig. 3, broad-band colours can evolve rapidly due to the presence of strong emission lines and continuum breaks. However, FLARES, like many hydrodynamical simulations, only produces outputs at discrete snapshots, in FLARES’ case being integer redshifts from $z = 15 \rightarrow 5$. To provide continuous redshift coverage we re-compute observed frame colours from the rest-frame SEDs using perturbed redshifts extending ± 0.5 around the snapshot redshift. For example, galaxies in the $z = 10$ snapshot are used to produce predicted colours over the redshift range $z = [9.5, 10.5]$. In practice, we re-sample every galaxy 10 times across the redshift interval, though it ensures that no galaxy appears more than once in each $\delta z = 0.1$ interval. The downside of this approach is that it assumes no evolution in the physical properties of galaxies between snapshots. However, if there was strong evolution between snapshots, this would lead to discontinuities at the boundaries between snapshot redshift ranges. As we will see in Fig. 4, while discontinuities exist, they are small.

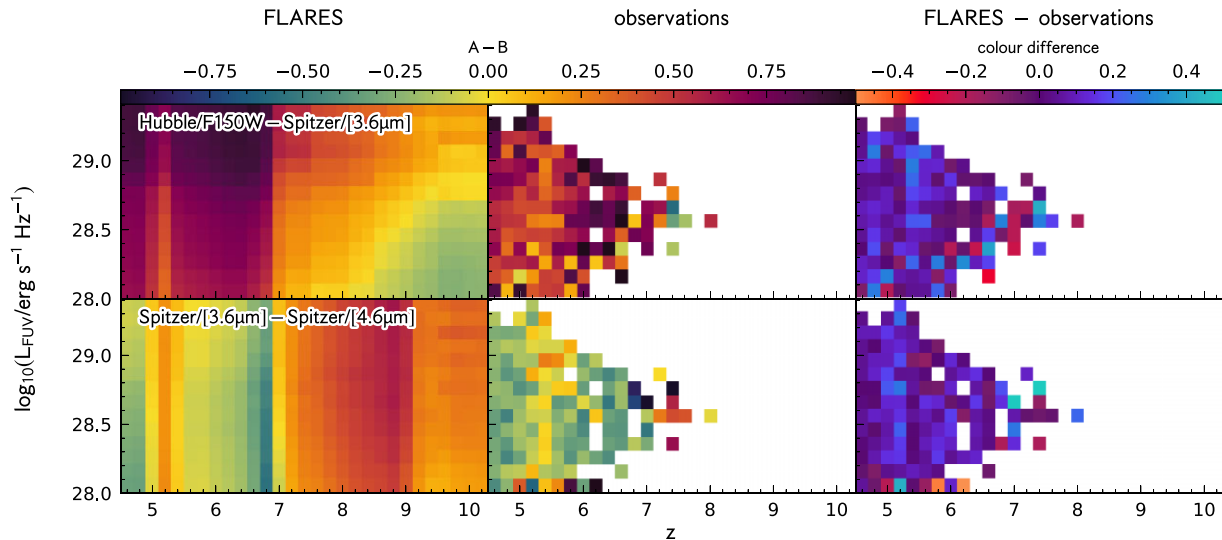


Figure 6. FLARES predictions for the colour evolution of galaxies with $M_{\text{FUV}} < -18$. The shaded band shows the central 68 per cent range of all galaxies $M_{\text{FUV}} < -18$. The lines denote the median in the various M_{FUV} bins.

4 PREDICTIONS

We begin by presenting the predictions for the average colour, in various NIRCcam and MIRI bands, as a function of redshift and rest-frame far-UV absolute magnitude in Fig. 4. Similar predictions for two *Hubble/Spitzer* colours are presented in Figs 5 and 6 in the context of the comparison with current observational constraints described in Section 4.1. As described in the Data Availability section, we make these results publicly available for comparison with observations or other models.

Most notable in Fig. 4 is the complex redshift evolution of most colours, with this variability driven by strong nebular line emission and break features (e.g. $F150W - F200W$ at $z > 10$). To see the impact of nebular line emission more clearly in Fig. 7, we also show the evolution of pure stellar and unattenuated colours. This reveals that the impact of nebular emission is to shift colours by up to 0.7 mag, with the largest shifts for $F430M - F444W$. As strong lines cross to adjacent filters this can also cause sharp redshift evolution of colours, with e.g. $F430W - F444W$ shifting by > 1 mag from $z = 7.5 \rightarrow 8.5$. This analysis also reveals that nebular emission is often important even in the absence of strong line emission. For example, in colours probing the rest-frame UV continuum (e.g. $F356W - F444W$ at $z > 12$), the impact of nebular emission reddens colours by up to 0.3 mag.

Fig. 4 also reveals that the most luminous galaxies are typically redder, albeit the shift is small relative to the variability introduced by nebular emission. As this trend largely disappears for intrinsic colours, we attribute this to the impact of dust attenuation. This is consistent with wider predictions from FLARES that broadly predicts increasing attenuation with observed UV absolute magnitude (see Vijayan et al. 2021)⁴

In Fig. 4, in addition to the median colour, we also show the central 68 per cent range of colours predicted over the full absolute magnitude range. This range is often very narrow, though does

⁴While there is a broad trend of increasing attenuation with observed UV luminosity, the most heavily attenuated galaxies have more modest UV luminosities due to the effects of attenuation dominating.

increase to up to 0.7 mag where nebular emission-line emission is important. This arises due to the strong sensitivity of the ionizing photon luminosity, and thus the contribution of line emission, to the recent star formation histories of galaxies.

4.1 Comparison with observations

At the time of writing, the first constraints on galaxy colours at high-redshift have just become available (Naidu et al. 2022) based on observations from the GLASS and CEERS Early Release Surveys. Naidu et al. (2022) report the discovery of two promising high-redshift galaxy candidates at $z \approx 11$ (GL-z11) and ≈ 13 (GL-z13). The reported colours of these sources are included in Fig. 4. The $F277W - F356W$ and $F356W - F444W$ colours of both candidates are consistent with the FLARES predictions. For $F200W - F277W$, GL-z11 falls slightly above our predictions and GL-z13 falls slightly below.

The handful of observations available, however, do not provide statistically useful constraints. Instead, we compare with comprehensive observed samples *Hubble* and *Spitzer*. In Fig. 5, we compare FLARES predictions for the *Hubble*/WFC3/ $F160W - Spitzer$ /IRAC/[3.6 μm] and *Spitzer*/IRAC/[3.6 μm] - *Spitzer*/IRAC/[4.5 μm] colours with recent measurements from Stefanon et al. (2021) (based on the sample identified by Bouwens et al. (2015)), Endsley et al. (2021), Hashimoto et al. (2018), and Laporte et al. (2021).

Stefanon et al. (2021) consistently reduced *Spitzer*/IRAC imaging across the Great Observatories Origins Deep Survey (GOODS)-N and GOODS-S fields. This data set was then used to measure the IRAC fluxes of almost 10 000 galaxies at $3.5 < z < 10$ based on the catalogue of Bouwens et al. (2015). The Bouwens et al. (2015) sample was selected and assigned photometric redshifts using *Hubble* imaging. The resulting colour evolution of this sample, matched to have the same rest-frame UV luminosity limit ($M_{\text{UV}} < -18.4$) is shown in Fig. 5. This reveals a median observed colour providing a close match to the FLARES predictions. The impact of [O III] and H β exiting the IRAC/[3.6 μm] band and entering the [4.5 μm] band at $z = 6.5 - 7.5$ can be clearly discerned. While the median colour is well matched, the scatter in the observations is up to five times

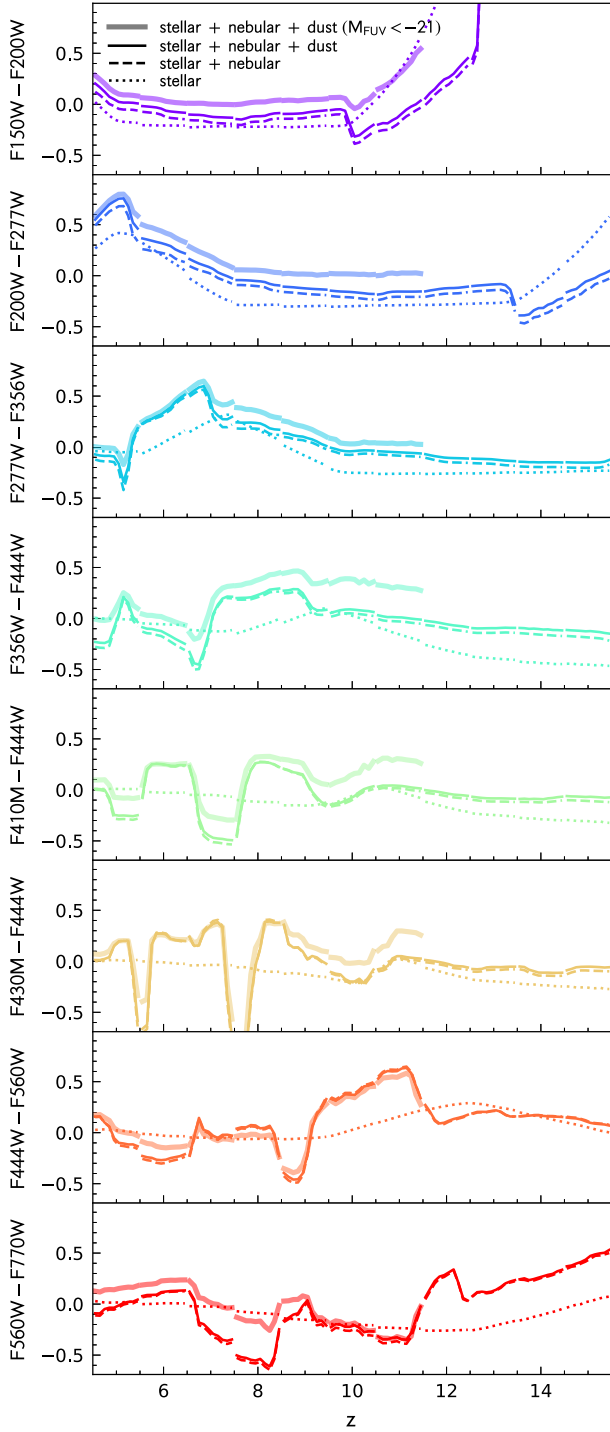


Figure 7. FLARES predictions for the pure stellar (dotted line), stellar + nebular (dashed line), and observed: stellar + nebular + dust (solid thin line) colour evolution of galaxies with $M_{\text{FUV}} < -18$, and in the latter scenario, also the brightest ($M_{\text{FUV}} < -21$) galaxies (solid thick line).

larger. While photometric scatter and redshift uncertainties will drive some of this difference it is possible this may reflect a real difference between the observations and FLARES. While we have applied a consistent luminosity limit to both samples, the Bouwens et al. (2015)/Stefanon et al. (2021) sample do not cover the same range of luminosities. To provide a fairer comparison in Fig. 6, we calculate the average colour in bins of luminosity and redshift and contrast

FLARES and the observations. This is inevitably more noisy but does not appear to reveal any systematic bias.

In Fig. 5, we also compare against the sample of Endsley et al. (2021). Endsley et al. (2021) select a sample of galaxies at $z \sim 6.5$ –7 from ground-based imaging of the COSMOS and XMM1 fields. They use a colour selection employing Subaru/Hyper Suprime-Cam NB921 narrow-band imaging to yield precise photometric redshifts. This, in turn, results in clean constraints on the [OIII] + H β equivalent widths of the sources from the *Spitzer*/IRAC photometry. The resulting $[3.6 \mu\text{m}] - [4.5 \mu\text{m}]$ colours at $z \sim 7$ closely match the FLARES predictions, in particular the very blue colours predicted at $z \approx 6.8$ and strong subsequent evolution to $z > 7$.

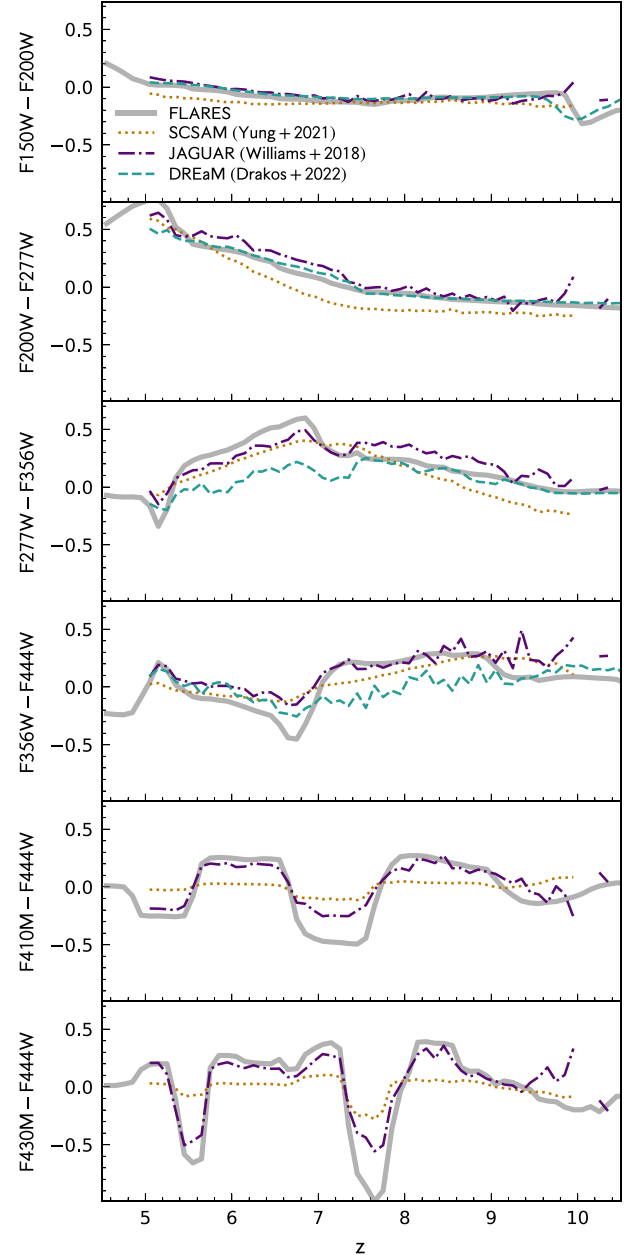


Figure 8. Predictions for the *JWST* colour evolution of galaxies with $M_{\text{FUV}} < -18$ from FLARES (thick grey line), the Santa Cruz SAM (Somerville et al. 2021; Yung et al. 2022), JAGUAR (Williams et al. 2018), and DREaM (Drakos et al. 2022).

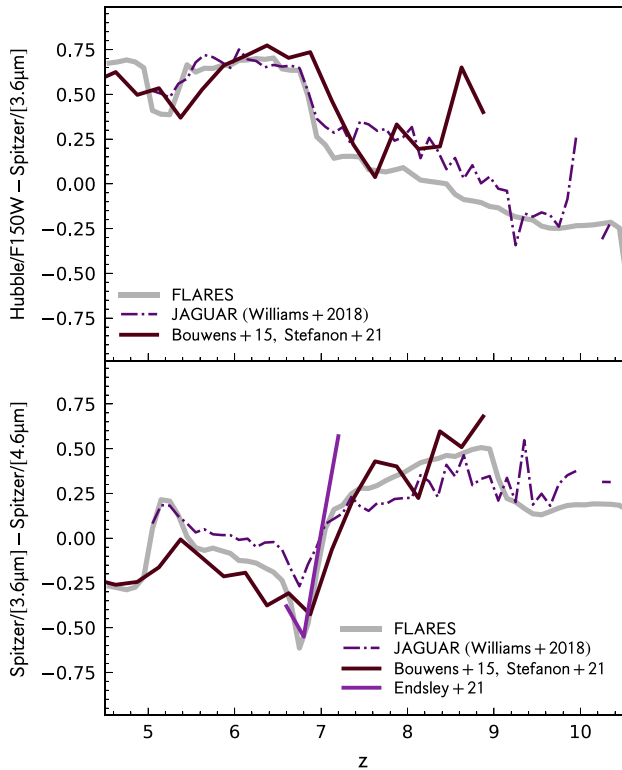


Figure 9. Predictions for the *Hubble* colour evolution of galaxies with $M_{\text{FUV}} < -18$ from FLARES (thick grey line) and JAGUAR (Williams et al. 2018). DREaM and the Santa Cruz SAM are omitted as they do not provide both *Hubble* and *Spitzer* photometry.

4.2 Comparison with other models

We now compare our predictions to mock catalogues from the phenomenological models of Williams et al. (2018) (also known as the JAGUAR: JAdes extraGalactic UltraDeep Artificial Realizations package) and DREaM: the Deep Realistic Extragalactic Model (Drakos et al. 2022), in addition to the Santa Cruz semi-analytical model (Somerville et al. 2021; Yung et al. 2022). The colour evolution of these models are contrasted with FLARES in Figs 8 and 9 for *JWST* and *Hubble* + *Spitzer* colours, respectively.

While the evolution of colours in all three sets of predictions is qualitatively similar there are some important differences, particularly around the impact of nebular emission. Specifically, FLARES consistently predicts a stronger contribution from nebular emission, resulting in more extreme variation of colours. While, at present, it is not possible to definitively claim one model provides better agreement with the observations, FLARES does appear to better reproduce the magnitude of the observed dip in the average $[3.6 \mu\text{m}] - [4.5 \mu\text{m}]$ colour at $z \approx 6.8$.

Due to the multitude of differences between the three models, the exact cause of this discrepancy is difficult to ascertain. Possibilities include the presence of more stochastic, young, or rapidly increasing star formation in FLARES or simply the choice of stellar population synthesis model, IMF, and/or photoionization modelling assumptions (see e.g. Wilkins et al. 2013b, 2016a, 2020). Critical to ascertaining the cause of this discrepancy is applying a consistent approach to the choice of SPS model, IMF, and photoionization assumptions, which should ultimately help diagnose the cause of this discrepancy. At the same time, results from *Webb* are now beginning to provide

the observational constraints to differentiate between different sets of predictions.

5 CONCLUSIONS

In this work, we have presented theoretical predictions for the colour evolution of galaxies at $z = 5-15$ from the FLARES. These predictions enable direct comparison with observational constraints, allowing both FLARES and the underlying model to be tested in a new regime. Our major findings are as follows:

(i) The predicted galaxy colours show complex evolution with rapid changes (up to ≈ 1 mag) in the average colours over short redshift intervals. This is due to the presence of strong nebular line emission moving through individual bands. In addition, our predicted colours show a trend, albeit modest, with luminosity, attributed to the increasing impact of dust in the most luminous galaxies.

(ii) FLARES predictions currently closely match, with the possible exception of the scatter, recent observational constraints at high redshift using *Hubble* and *Spitzer* from Stefanon et al. (2021) (based on the galaxy sample identified by Bouwens et al. 2015) and Endsley et al. (2021), in addition to early results from *Webb* (Naidu et al. 2022).

(iii) FLARES predictions qualitatively match other available predictions including the phenomenological models of Williams et al. (2018) and Drakos et al. (2022), and the Somerville et al. (2021) and Yung et al. (2022) semi-analytical model. However, FLARES predicts larger variations due to stronger nebular line emission. The exact cause of this difference may lie in the fundamental physical properties of galaxies in the models but may also reflect different modelling assumptions – e.g. choice of stellar population synthesis model and IMF – for the forward modelling.

With the imminent explosion of constraints from *Webb*, we will soon be in a position to differentiate between these and other models, providing insights into the physics driving the physical and observational properties of galaxies in the early Universe. Mock observations like those presented here, and in e.g. Williams et al. (2018), Somerville et al. (2021), Drakos et al. (2022), and Yung et al. (2022), provide a critical resource to test and refine photometric redshift selection techniques and/or the inference of physical properties from broad-band photometry.

ACKNOWLEDGEMENTS

We dedicate this paper to healthcare and other essential workers, the teams involved in developing the vaccines, and to all the parents who found themselves having to home-school children while holding down full-time jobs. We thank Rychard Bouwens, Ryan Endsley, and Mauro Stefanon for providing machine readable catalogues of their observations for comparison. 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). 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. 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. 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 colourmaps in the CMASHER package (van der Velden 2020).

DATA AVAILABILITY STATEMENT

The 2.2, 15.8, 50, 84.2, and 97.8 percentiles of the colour distribution in redshift and rest-frame far-UV luminosity bins are available in the ASTROPY Enhanced Character Separated Values (.ecsv) table format at https://github.com/stephenmwilkins/flares_colours_data and as part of the wider First Light and Assembly of Galaxies model predictions repository available at https://github.com/stephenmwilkins/flags_data. Data from the wider FLARES project are available at <https://flaresimulations.github.io/data.html>. If you use data from this paper, please also cite Lovell et al. (2021a) and Vijayan et al. (2021).

REFERENCES

- Astropy Collaboration et al., 2013, *A&A*, 558, A33
 Astropy Collaboration et al., 2018, *AJ*, 156, 123
 Bouwens R. J. et al., 2015, *ApJ*, 803, 34
 Camps P., Baes M., 2015, *Astron. Comput.*, 9, 20
 Chabrier G., 2003, *PASP*, 115, 763
 Charlot S., Fall S. M., 2000, *ApJ*, 539, 718
 Crain R. A. et al., 2015, *MNRAS*, 450, 1937
 Davidzon I. et al., 2019, *MNRAS*, 489, 4817
 Drakos N. E. et al., 2022, *ApJ*, 926, 194
 Endsley R., Stark D. P., Chevallard J., Charlot S., 2021, *MNRAS*, 500, 5229
 Ferland G. J. et al., 2017, *Rev. Mex. Astron. Astrofis.*, 53, 385
 Harris C. R. et al., 2020, *Nature*, 585, 357
 Hashimoto T. et al., 2018, *Nature*, 557, 392
 Hunter J. D., 2007, *Comput. Sci. Eng.*, 9, 90
 Kannan R., Smith A., Garaldi E., Shen X., Vogelsberger M., Pakmor R., Springel V., Hernquist L., 2022, *MNRAS*, 514, 3857
 Katz H. et al., 2019, *MNRAS*, 487, 5902
 Lagache G., Cousin M., Chatzikos M., 2018, *A&A*, 609, A130
 Laigle C. et al., 2019, *MNRAS*, 486, 5104
 Laporte N., Meyer R. A., Ellis R. S., Robertson B. E., Chisholm J., Roberts-Borsani G. W., 2021, *MNRAS*, 505, 3336
 Leung T. K. D., Olsen K. P., Somerville R. S., Davé R., Greve T. R., Hayward C. C., Narayanan D., Popping G., 2020, *ApJ*, 905, 102
 Lovell C. C., Vijayan A. P., Thomas P. A., Wilkins S. M., Barnes D. J., Irodotou D., Roper W., 2021a, *MNRAS*, 500, 2127
 Lovell C. C., Geach J. E., Davé R., Narayanan D., Li Q., 2021b, *MNRAS*, 502, 772
 Ma X. et al., 2018, *MNRAS*, 477, 219
 Ma X. et al., 2019, *MNRAS*, 487, 1844
 Marshall M. A. et al., 2022, *MNRAS*, 516, 1047
 Naidu R. P. et al., 2022, preprint ([arXiv:2207.09434](https://arxiv.org/abs/2207.09434))
 Narayanan D. et al., 2021, *ApJS*, 252, 12
 Orsi Á., Padilla N., Groves B., Cora S., Tecce T., Gargiulo I., Ruiz A., 2014, *MNRAS*, 443, 799
 Popping G., Narayanan D., Somerville R. S., Faisst A. L., Krumholz M. R., 2019, *MNRAS*, 482, 4906
 Robertson B. E., 2021, preprint ([arXiv:2110.13160](https://arxiv.org/abs/2110.13160)) [astro-ph]
 Roper W. J., Lovell C. C., Vijayan A. P., Marshall M. A., Irodotou D., Kuusisto J. K., Thomas P. A., Wilkins S. M., 2022, *MNRAS*, 514, 1921
 Schaerer D., de Barros S., 2009, *A&A*, 502, 423
 Schaye J. et al., 2015, *MNRAS*, 446, 521
 Snyder G. F. et al., 2015, *MNRAS*, 454, 1886
 Somerville R. S. et al., 2021, *MNRAS*, 502, 4858
 Stanway E. R., Eldridge J. J., 2018, *MNRAS*, 479, 75
 Stark D. P., Schenker M. A., Ellis R., Robertson B., McLure R., Dunlop J., 2013, *ApJ*, 763, 129
 Stefanon M. et al., 2021, *ApJS*, 257, 68
 Trayford J. W. et al., 2015, *MNRAS*, 452, 2879
 van der Velden E., 2020, *J. Open Source Softw.*, 5, 2004
 Vijayan A. P., Clay S. J., Thomas P. A., Yates R. M., Wilkins S. M., Henriques B. M., 2019, *MNRAS*, 489, 4072
 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
 Vijayan A. P. et al., 2022, *MNRAS*, 511, 4999
 Virtanen P. et al., 2020, *Nat. Methods*, 17, 261
 Vogelsberger M. et al., 2020, *MNRAS*, 492, 516
 Whittle L., Stark D. P., Endsley R., Leja J., Charlot S., Chevallard J., 2022, preprint ([arXiv:2206.05315](https://arxiv.org/abs/2206.05315))
 Wilkins S. M., Bunker A., Coulton W., Croft R., di Matteo T., Khandai N., Feng Y., 2013a, *MNRAS*, 430, 2885
 Wilkins S. M. et al., 2013b, *MNRAS*, 435, 2885
 Wilkins S. M., Feng Y., Di-Matteo T., Croft R., Stanway E. R., Bouwens R. J., Thomas P., 2016a, *MNRAS*, 458, L6
 Wilkins S. M., Feng Y., Di-Matteo T., Croft R., Stanway E. R., Bunker A., Waters D., Lovell C., 2016b, *MNRAS*, 460, 3170
 Wilkins S. M., Feng Y., Di Matteo T., Croft R., Lovell C. C., Thomas P., 2018, *MNRAS*, 473, 5363
 Wilkins S. M. et al., 2020, *MNRAS*, 493, 6079
 Wilkins S. M. et al., 2022a, preprint ([arXiv:2204.09431](https://arxiv.org/abs/2204.09431))
 Wilkins S. M. et al., 2022b, preprint ([arXiv:2208.00976](https://arxiv.org/abs/2208.00976))
 Williams C. C. et al., 2018, *ApJS*, 236, 33
 Yung L. Y. A. et al., 2022, *MNRAS*, 515, 5416
 Zackrisson E., Bergvall N., Leitert E., 2008, *ApJ*, 676, L9

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