The main sequence of star-forming galaxies across cosmic times

P. Popesso,[1]★ A. Concas [ID],[1] G. Cresci [ID],[2] S. Belli [ID],[3] G. Rodighiero,[4] H. Inami [ID],[5] M. Dickinson,[6] O. Ilbert,[7] M. Pannella [ID][8] and D. Elbaz[9]

[1]*European Southern Observatory, Karl Schwarzschildstrasse 2, D-85748 Garching bei München, Germany*
[2]*INAF-Osservatorio Astronomico di Arcetri, Largo Enrico Fermi 5, I-50125 Firenze, Italy*
[3]*Dipartimento di Fisica e Astronomia "Augusto Righi", Università di Bologna, Viale Berti Pichat 6/2, Bologna, Italy*
[4]*Dipartimento di Fisica e Astronomia, Universitá di Padova, Vicolo dell' Osservatorio 3, I-35122 Padova, Italy*
[5]*Hiroshima Astrophysical Science Center, 1-3-2 Kagamiyama, Higashi-Hiroshima City, 739-8511 Hiroshima, Japan*
[6]*NOIRLab, 950 N. Cherry Ave. Tucson, AZ 85719, USA*
[7]*Laboratoire d'Astrophysique de Marseille, 38 rue Frederic Joliot Curie, F-13388 Marseille, France*
[8]*Dipartimento di Fisica e Astronomia, Universitá di Trieste, Via Valerio, 2, I-34127 Trieste, Italy*
[9]*Laboratoire AIM-Paris-Saclay, CEA/DRF/Irfu - CNRS - Université Paris Diderot, CEA-Saclay, F-91191 Gif-sur-Yvette, France*

ABSTRACT

By compiling a comprehensive census of literature studies, we investigate the evolution of the main sequence (MS) of star-forming galaxies (SFGs) in the widest range of redshift ($0 < z < 6$) and stellar mass ($10^{8.5}$–$10^{11.5}$ $M_\odot$) ever probed. We convert all observations to a common calibration and find a remarkable consensus on the variation of the MS shape and normalization across cosmic time. The relation exhibits a curvature towards the high stellar masses at all redshifts. The best functional form is governed by two parameters: the evolution of the normalization and the turnover mass ($M_0(t)$), which both evolve as a power law of the Universe age. The turn-over mass determines the MS shape. It marginally evolves with time, making the MS slightly steeper towards $z \sim 4$–6. At stellar masses below $M_0(t)$, SFGs have a constant specific SFR (sSFR), while above $M_0(t)$ the sSFR is suppressed. We find that the MS is dominated by central galaxies. This allows to turn $M_0(t)$ into the corresponding host halo mass. This evolves as the halo mass threshold between cold and hot accretion regimes, as predicted by the theory of accretion, where the central galaxy is fed or starved of cold gas supply, respectively. We, thus, argue that the progressive MS bending as a function of the Universe age is caused by the lower availability of cold gas in haloes entering the hot accretion phase, in addition to black hole feedback. We also find qualitatively the same trend in the largest sample of star-forming galaxies provided by the IllustrisTNG simulation. Nevertheless, we still note large quantitative discrepancies with respect to observations, in particular at the high-mass end. These can not be easily ascribed to biases or systematics in the observed SFRs and the derived MS.

Key words: galaxies: evolution – galaxies: high-redshift – galaxies: star formation.

1 INTRODUCTION

The main sequence (MS) of star-forming galaxies (SFGs) is considered one of the most useful tools in modern astrophysics in the field of galaxy evolution. This very tight relation between the galaxy star formation rate (SFR) and the stellar mass ($M_\star$) is in place from redshift ~ 0 up to ~ 6 (Brinchmann et al. 2004; Daddi et al. 2007; Elbaz et al. 2007; Noeske et al. 2007; Salim et al. 2007; Chen et al. 2009; Pannella et al. 2009; Santini et al. 2009; Magdis et al. 2010; Oliver et al. 2010; Elbaz et al. 2011; Karim et al. 2011; Rodighiero et al. 2011; Shim et al. 2011; Lee et al. 2012; Reddy et al. 2012; Salmi et al. 2012; Whitaker et al. 2012; Zahid et al. 2012; Kashino et al. 2013; Moustakas et al. 2013; Sobral et al. 2014; Speagle et al. 2014; Steinhardt et al. 2014; Whitaker et al. 2014; Lee et al. 2015; Schreiber et al. 2015; Tasca et al. 2015; Shivaei et al. 2015a; Erfanianfar et al. 2016; Kurczynski et al. 2016; Santini et al. 2017; Belfiore et al. 2018; Pearson et al. 2018; Popesso et al. 2019a,b; Leslie et al. 2020; Thorne

et al. 2021; Daddi et al. 2022; Leja et al. 2022). The evolution of its normalization, slope, and scatter have been largely studied in the past decade. It is now well-established that the normalization declines significantly but smoothly as a function of redshift, likely on mass-dependent time-scales (see also Speagle et al. 2014), rather than being driven by stochastic events like major mergers and starbursts (Oemler et al. 2017). More uncertain is the precise redshift dependence of such evolution, which is often expressed as $\propto (1 + z)^\gamma$, with γ varying from 1.9 to 3.7 (Ilbert et al. 2013; Speagle et al. 2014; Whitaker et al. 2014; Schreiber et al. 2015; Pearson et al. 2018; Leslie et al. 2020; Thorne et al. 2021; Leja et al. 2022). This is mainly due to the uncertainty in deriving the evolution of the exact shape of the relation, which is still matter of intense debate. Several studies point to a power law shape, SFR $\propto M_\star^\alpha$, both in the local Universe (Peng et al. 2010; Renzini & Peng 2015) and at high redshift (Rodighiero et al. 2014; Speagle et al. 2014; Kurczynski et al. 2016; Pearson et al. 2018). Other works suggest that the relation exhibits a curvature towards the high mass at low (Popesso et al. 2019b) and high redshift (Whitaker et al. 2014; Lee et al. 2015; Schreiber et al. 2015; Tasca

★ E-mail: paola.popesso@eso.org

et al. 2015; Tomczak et al. 2016; Popesso et al. 2019a; Leslie et al. 2020; Thorne et al. 2021; Leja et al. 2022). Also, the scatter around the relation is quite debated with very conflicting results from the literature. Some works report a quite constant scatter of 0.2–0.3 dex from low-to-moderate high masses (e.g. Elbaz et al. 2007; Noeske et al. 2007; Schreiber et al. 2015). Others, instead, report a decrease of the scatter from very low (10^8 M$_\odot$) to moderate (10^{10} M$_\odot$) stellar masses at different redshifts (Willett et al. 2015; Santini et al. 2017; Boogaard et al. 2018). Few others observe an increase as a function of the stellar mass from 0.3 dex at 10^{10} to 0.5 dex at $10^{11.5}$ M$_\odot$, from low to high redshift(e.g. Guo, Zheng & Fu 2013; Popesso et al. 2019b,a; Sherman et al. 2021).

Most of this discrepancy is ascribable to how SFGs are selected in the first place, on the SFR estimators and on the method of localization of the MS. Popesso et al. (2019a) show that colour–colour selection of SFGs leads to the exclusion of red dusty SFGs in particular at $z > 1.5$ and to a steep MS. Tomczak et al. (2016) show that without any selection, the MS is bending more significantly in the local than in the distant Universe (see also Katsianis, Tescari & Wyithe 2016). In addition, the use of different SFR indicators might lead to systematic biases. Part of the UV emission originating from the young star population is absorbed by dust and re-processed at infrared wavelengths. Such emission alone can provide a measure of the SFR only if corrected for this absorption. However, the measure of the dust attenuation is still uncertain because of the degeneracy between age and reddening, the assumption on galaxy metallicity and SF histories, and the parametrization of the extinction curve (e.g. Meurer, Heckman & Calzetti 1999; Dale et al. 2009; Bourne et al. 2017; Dunlop et al. 2017). With the launch of the *Spitzer* and *Herschel* satellites in 2003 and 2009, respectively, it became possible to measure the mid- and far-infrared (FIR) emission for statistical samples of galaxies up to $z \sim 3$ in the most studied deep fields (e.g. Lutz 2014). Such measurements allowed finally combining the unobscured UV emission and the reprocessed FIR component and calibrating them as a SFR indicator. Nevertheless, such an approach might also lead to systematics and uncertainty. For instance, the contamination by active galactic nucleus (AGN) and the overestimation of the SFR for starburst galaxies make of the mid-infrared emission, based e.g. on *Spitzer* MIPS 24 μm data, a less accurate indicator with respect to longer wavelengths (Nordon et al. 2010; Elbaz et al. 2011). None the less, such contamination can affect, to some extent, also the emission in the FIR, if the nuclear activity was particularly high in the last 100 *Myr* of a galaxy's life. More recently, the accuracy of the SFR based on the combination of UV and IR components has been questioned by the results based on sophisticated SED fitting codes that try to reconstruct the galaxy star formation histories (SFHs; e.g. Thorne et al. 2021; Leja et al. 2022). Navigating through all these differences makes it very difficult to understand if the literature has reached an overall consensus on the shape and evolution of the MS across cosmic times. Furthermore, the magnitude of these effects precludes robust interpretations of derived MS properties.

To overcome this problem, and so constrain the MS evolution and systematic errors, Speagle et al. (2014, hereafter S14) have compiled 64 MS observations from 25 studies published in the period 2007–2014, spanning $z \sim 0$–6, and converted them to the same absolute calibration. These MS estimates have been taken from a variety of fields, selected using different methodologies, including both stacked and non-stacked data, and estimated with a variety of SFR indicators. By calibrating consistently all data sets, S14 determine the MS best fit as a power law, with a slope marginally evolving with redshift. The purpose of such an experiment is not to provide the 'true' MS. Indeed,

there might still be biases and limitations related to the observational techniques that this approach is not capable of erasing or correcting (see for instance the discussion in Katsianis et al. 2020). Rather, it aims at understanding whether the MS estimates, reported in the literature, lead to a substantial consensus once they are brought to the same calibration and all systematics are considered. It also offers the advantage that the resulting MS can be easily re-calibrated, if necessary.

In this paper, we adopt the same approach of S14 to bring all estimates of the literature to a common calibration to check whether the community has reached or is far from reaching such a consensus. To this aim, we extend the collection of MS determination of S14 to the most recent estimates of the MS based on a variety of SFR indicators (from the combination of *Spitzer* mid-infrared and *Herschel* FIR data and UV emission, to SED fitting techniques) and methodology. These include additional 27 publications from 2014 to 2022, for a total of $\sim$120 MS measurements at $0 < z < 6$, which we convert to the same IMF and calibrate in a consistent framework. We collect a sample of about $\sim$1500 consistently calibrated data points, which express the location of the MS as a function of stellar mass and time, and check whether there is consensus in the evolution of the MS across cosmic time.

The paper is structured as follows. Section 1 describes the MS compilation. Section 2 presents our best-fitting procedure. Section 3 shows our results. Section 4 provides a comparison of our results with previous findings and theoretical predictions, while Section 5 lists our conclusions. We assume a ΛCDM cosmology with $\Omega_M = 0.3$, $\Omega_\Lambda = 0.7$, and $H_0 = 70$ km s^{-1} Mpc^{-1}, and a Kroupa IMF throughout the paper.

2 THE MS COLLECTION

In this work, we focus on MS estimates that have been published after 2014. To consider, in the analysis, the MS measurements that have been published in the period 2007–2014, we include here the MS relation determined by S14 as a result of a compilation of 64 MS estimates collected in 25 papers (see table 3 of S14). We consider, in particular, the fit n. 64, which S14 provide as their reference MS. To this, we add all the MS relations that satisfy the following criteria:

(i) Includes a published M$_\star$ – SFR or M$_\star$ – sSFR (sSFR = SFR/M$_\star$) relation (slope α and normalization β) or otherwise analogous quantities;

(ii) Fit(s) include more than two data points (if stacked) or 50 galaxies (if directly observed). This is required to avoid biases resulting from small number statistics;

(iii) Stacked points must provide mean or median of more than 25 points, to avoid large uncertainties due to low number statistics;

(iv) Published after 2014.

The 27 publications that are retrieved considering these criteria, in addition to S14, are listed in Table 1, together with the used IMF, SFR indicator, redshift range, cosmological parameters, SFGs selection method, and extinction curve. In Appendix A, we provide a full description of the individual publication data and of their calibration.

All the considered MS estimates are based on the deepest UV, optical, IR, and radio surveys ever realized on the Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey (CANDELS), Cosmic Evolution Survey (COSMOS), Extended Chandra Deep Field South (ECDFS) fields at intermediate and high redshift, and on the WISE, *GALEX* and optical spectroscopy in the SDSS area at $z \sim 0$. Of these, 10 publications (Heinis et al. 2014; Whitaker et al. 2014; Chang et al. 2015; Ilbert et al. 2015; Lee et al. 2015; Schreiber et al.

Table 1. Column 1: Reference. Column 2: Assumed stellar initial mass function. Column 3: SFR indicator. Column 4: Redshift range analysed. Column 5: Assumed cosmology. Column 6: Selection methods used for the parent samples (see Appendix A or individual papers for more details). Column 7: Type of data retrieved in the paper: *stacked* stands for stacked SFR data points as a function of stellar mass and redshift, *data-points* for average or median SFR, data-points as a function of stellar mass and redshift, and *best fit* stands for best-fitting parameters expressing the MS shape as a function of stellar mass over the retrieved stellar mass and redshift ranges. Column 8: Extinction curve reference: the details of the extinction curve implemented in MAGPHYS can be found in da Cunha, Charlot & Elbaz (2008); C00 refers to Calzetti et al. (2000), CF00 to Charlot & Fall (2000), KC13 to Kriek & Conroy (2013), and N09 to Noll et al. (2009). Column 9: indicates whether the data have been included in (✓symbol) or excluded from (− symbol) the current analysis. The reasons for excluding an individual data set from the analysis are given in Section 2 and in Appendix A with more details.

Paper	IMF	SFR indicator	z range	$(h, \Omega_m, \Omega_\Lambda)$	Selection	data type	Extinction curve	included
Speagle et al. (2014)	K	mixed	0.2–6	0.7, 0.3, 0.7	mixed	best fit	NA	✓
Rodighiero et al. (2014)	S	IR	1.4–2.5	0.7, 0.25, 0.75	BzK	stacked	NA	✓
Heinis et al. (2014)	C	SED+IR	1.5–4	0.7, 0.3, 0.7	UV	stacked	NA	✓
Whitaker et al. (2014)	C	NUV+IR	0.5–2.5	0.7, 0.3, 0.7	UVJ	stacked	NA	✓
Chang et al. (2015)	C	SED+IR	<0.1	0.7, 0.3, 0.7	mixed	best fit	MAGPHYS	✓
Lee et al. (2015)	C	NUV+IR	0.3–1.3	0.7, 0.28, 0.72	NUVRJ	stacked	NA	✓
Ilbert et al. (2015)	C	NUV+IR	0.2–1.4	0.7, 0.3, 0.7	NUVRJ	data-points	NA	✓
Tasca et al. (2015)	C	SED	0.4–5	0.7, 0.3, 0.7	LBG+$I_{AB} < 25$	data-points	C00	✓
Salmon et al. (2015)	C	SED	3.5 to −6.5	0.7, 0.3, 0.7	mixed	stacked	C00	✓
Renzini & Peng (2015)	C	H α	<0.085	0.7, 0.3, 0.7	mixed	best fit	CF00	✓
Schreiber et al. (2015)	S	FUV+IR	0.3–4	0.7, 0.3, 0.7	UVJ	stacked	NA	✓
De los Reyes et al. (2015)	S	H α	0.8	0.7, 0.3, 0.7	H α emission	stacked	CF00	✓
Erfanianfar et al. (2016)	C	IR	0.2–1.5	0.7, 0.3, 0.7	mixed	data-points	NA	✓
Tomczak et al. (2016)	C	NUV+IR	0.2–3	0.7, 0.3, 0.7	UVJ	stacked	NA	✓
Santini et al. (2017)	S	SED	1.3–6	0.7, 0.3, 0.7	mixed	best fit	C00	✓
Kurczynski et al. (2016)	S	SED	0.5–3	0.7, 0.3, 0.7	UVJ	best fit	C00	✓
Pearson et al. (2018)	C	SED+IR	0.2–6	0.704, 0.272, 0.728	UVJ	best fit	CF00	−
Belfiore et al. (2018)	C	H α	<0.1	0.7, 0.3, 0.7	mixed	best fit	CF00	✓
Davidzon et al. (2018)	C	GSMF	2–6	0.7, 0.3, 0.7	NUVRJ	stacked	C00	✓
Lee et al. (2018)	C	FUV+IR	1.2–4	0.7, 0.3, 0.7	UVJ	stacked	NA	✓
Iyer et al. (2018)	C	SED	0.5–6	0.7, 0.3, 0.7	mixed	stacked	C00	✓
Popesso et al. (2019a)	C	IR/H α	<0.085	0.7, 0.3, 0.7	mixed	data-points	NA	✓
Popesso et al. (2019b)	C	FUV+IR	0.2 to −2.5	0.7, 0.3, 0.7	mixed	data-points	NA	✓
Barro et al. (2019)	C	SED	0.5 to −3.	0.7, 0.3, 0.7	UVJ	best fit	C00	✓
Leslie et al. (2020)	C	radio	0.3 to −6	0.7, 0.3, 0.7	NUVRJ	stacked	NA	✓
Thorne et al. (2021)	C	SED	0 to −9	0.678, 0.308, 0.692	sSFR cut	best fit	CF00	✓
Sherman et al. (2021)	C	SED	1.5 to −3	0.7, 0.3, 0.7	mixed	best fit	KC13	✓
Leja et al. (2022)	C	SED	0.2 to −3.	0.698, 0.235, 0.765	ridged line	best fit	N09	✓

2015; Tomczak et al. 2016; Lee et al. 2018; Pearson et al. 2018; Popesso et al. 2019a) out of 27 are based on an SFR indicator given by the combination of UV and IR data. Additional three are based only on FIR PACS data (Rodighiero et al. 2014; Erfanianfar et al. 2016; Popesso et al. 2019a). Other three (Renzini & Peng 2015; Belfiore et al. 2018) are based on *H α*-derived SFR taken from the SDSS spectroscopic dataset (Brinchmann et al. 2004) or from 3D-HST data at higher redshift (de los Reyes et al. 2015). Additional nine include SFR derived through the SED-fitting technique based on different fitting methods and SFH reconstructions (Salmon et al. 2015; Tasca et al. 2015; Kurczynski et al. 2016; Santini et al. 2017; Iyer et al. 2018; Barro et al. 2019; Sherman et al. 2021; Thorne et al. 2021; Leja et al. 2022). Of the rest, one is based on the evolution of the galaxy stellar mass function of SFGs (Davidzon et al. 2018), and one is based on deep radio data (Leslie et al. 2020). S14 is based on a collection of different SFR estimators. According to the definition of S14, all of the considered publications but two are based on 'mixed' methods (see next section) for the selection of SFGs. These include colour-colour techniques (BzK, UVJ, and NUVRJ), 2σ clipping, and bimodality in the SFR–$M_\star$ plane. Only three publications have an MS obtained for blue and active galaxies selected in the UV or with the Lyman Break Galaxy (LBG) technique.

Differently from S14, we also include in our study the MS estimate at $z \sim 0$. S14 discuss the inability to distinguish a 'best' MS fit among the available $z \sim 0$ estimates obtained before 2014. Thus, they decide not to include the local MS estimates. Popesso et al. (2019b) discuss extensively all the data sets available in the local Universe, selection effects and different SFR estimators and the level of agreement between the different estimates. On the basis of these results, we conclude that the inter-publication scatter of the local MS estimates available in the period 2014–2022 is comparable to that found at higher redshifts. Indeed, no different level of agreement is found as a function of redshift, once all MS estimates are brought to the same calibration (see next section for more details).

From each publication, either we take the mean or median SFR data points or staked SFR data points at the observed stellar mass and redshift without any extrapolation or interpolation, or, if these are not available, we use the provided MS best-fitting parameters at a given redshift to estimate the MS in the provided stellar mass range in bins of 0.15 dex in stellar mass. The bin width is chosen to be a representative of the average stellar mass error of the considered papers. The mass ranges have either been taken directly from the paper in question or estimated based on the data included in the relevant fits, rounded to the nearest 0.1 dex (see Appendix A for a detailed description of the data taken from each publication). This

leads to ∼120 determinations of the MS for a sample of ∼1500 data points of MS SFR as a function of stellar mass and redshift or time ($SFR(M_*, z)$ or $SFR(M_*, t)$). The data collected here encompass the widest range in redshift ($0 < z < 6$), stellar mass ($10^{8.5}$–$10^{11.5}$ M$_\odot$), and SFR (0.01–500 M$_\odot$ yr^{-1}) available in the literature, and intend to give a census of most of the techniques and methods used to derived the MS location. Following the example of S14, we present what we hope is the broadest and most accurate census of MS observations to date.

2.1 MS calibration

As underlined in S14, several aspects need to be taken care of when comparing different MS estimates and before attempting any analysis:

- (i) Initial mass function;
- (ii) SFR estimator;
- (iii) SPS model;
- (iv) cosmology;
- (v) emission line effect in the estimate of SFR and M$_\star$;
- (vi) SFH;
- (vii) dust extinction curve;
- (viii) photo-z biases;
- (ix) SED fitting procedures;
- (x) Selection effects due to different SFG population selection.

A detailed discussion in S14 points out that only the points (i), (ii), and (iii) lead to relevant corrections when calibrating all the MS estimates to a common ground. Instead, different cosmologies (iv) have relatively negligible effects (<0.05 dex) at $0 < z < 6$, which is the same redshift range considered here. The effect of emission lines (v) in the estimate of M$_\star$ and SFR is relevant only when the multiwavelength information are limited to few photometric bands. They are, instead, negligible for data sets like CANDELS and COSMOS, as those considered here.

Similarly to S14, we choose not to adjust our results for differences in assumed SFH (vi), dust attenuation curves (vii), possible photo-z biases (viii), or differences in SED fitting procedures (ix). This might differ substantially (see the extensive discussion in Simha et al. 2014; Acquaviva, Raichoor & Gawiser 2015; Salmon et al. 2015; Ciesla, Elbaz & Fensch 2017; Iyer & Gawiser 2017; Carnall et al. 2019; Theios et al. 2019; Lower et al. 2020; Curtis-Lake et al. 2021; Thorne et al. 2021). Testing and correcting the effect due to such differences would require checking the galaxy SFR catalogs used to build the MS of the individual publications. In none of the considered cases, such information is publicly available. Only the stacked, averaged, or fitted MS is provided, which prevents performing any correction. In principle, the effect of the different assumptions and implementations could be predictable. There is a large recent literature of such attempts based on the combination of dust radiative transfer models and simulations (Baes et al. 2020; Lower et al. 2020; Trayford et al. 2020; Katsianis et al. 2021; Lovell et al. 2021; Narayanan et al. 2021). However, such corrections might introduce biases dependent on the different assumptions of the models. Thus, we choose not to apply such corrections. As shown at the end of the next section, neglecting the possible effects of the mentioned differences does not lead to an increase of the resulting inter-publication scatter.

We apply a correction only for two publications (Thorne et al. 2021; Leja et al. 2022), for which the authors explicitly indicate a systematic difference in the estimate of the stellar mass with respect to previous works of our collection. This correction is applied not because we consider these quantities wrong or inaccurate, but only

to bring all MS estimates to the same stellar mass and SFR scale (see next paragraph and Appendix A). All other publications based on the SED fitting technique do not report systematic differences in their stellar mass and SFR estimates. Thus, no further corrections are applied.

SFG selection methods (x) can lead to substantially different MS slopes. S14 distinguish between 'bluer', 'mixed', and 'non-selective' techniques: 'bluer' methods are based on a simple colour cut to select blue galaxies, 'mixed' ones involve a more sophisticated distinction between star-forming and quiescent galaxies, while 'non-selective' ones do not apply any selection. In particular, S14 find that 'bluer'-based ('non-selective'-based) MS slopes are biased towards values closer to unity (zero), with respect to 'mixed'-based slopes.

To bring every relation onto a common framework, we apply the following steps. We use the equations reported in Section 2.2 to convert the M$_\star$ values to a Kroupa IMF and the SFR estimates to the Kennicutt & Evans (2012, hereafter KE12) calibration (based on the Kroupa IMF). The choice of the KE12 calibration is dictated by the fact that it is the only one available in the literature that calibrates consistently all the SFR indicators considered here, with the exclusion of the SED fitting technique. Thus, it is the obvious choice to bring most of the SFR estimates to a common framework. The available SFRs based on the SED fitting technique included in the collection, are compared case by case with the KE12 calibration to check for consistency, as described in Appendix A. The KE12 calibration is based on data of star-forming regions in a nearby galaxy (Hao et al. 2011; Murphy et al. 2011). To check that the SFR indicators calibrated according to KE12 provide consistent estimates also at higher redshift, we perform several tests in Appendix B. The result of such tests is that SFRs derived through IR, $UV + IR$, $H\alpha$, and radio emission at 1.4 GHz, according to the KE12 calibration (see next section), are all consistent up to $z \sim 3$ with a scatter that varies from 0.2 to 0.25 dex.

We adjust for differences in cosmology using the $(h, \Omega_M, \Omega_\Lambda) = (0.7, 0.3, 0.7)$ WMAP concordance cosmology (Spergel et al. 2003). Furthermore, no additional dust correction is required, as the dust has been corrected for in the considered publications, and different extinction curves do not have a significant impact on the MS determination (see S14). Finally, as only 3 (Heinis et al. 2014; Tasca et al. 2015; Thorne et al. 2021) out of the 27 MS estimates considered here are consistent with a 'bluer' selection, we do not correct in this respect, but we discuss possible biases on the result.

As pointed out in Popesso et al. (2019b), another source of small discrepancy among different MS estimates is the method used to determine the MS location, e.g. as the mean or median SFR of the SFG population in the MS region. A correction can be made in this respect under the assumption that the SFR distribution at fixed M$_\star$ is log-normal in the MS region. This assumption is justified by several works in literature, which find a log-normal distribution at any stellar mass (Daddi et al. 2007; Rodighiero et al. 2011; Schreiber et al. 2015; Popesso et al. 2019b, a; Leja et al. 2022). In this case, the peak of the distribution coincides with the median SFR. The mean SFR is always larger than the median by an offset that depends only on the dispersion of the distribution. For a dispersion of ∼0.3 dex, the correction is ∼0.1 dex and it increases to ∼0.17 dex for a dispersion of ∼0.5 dex. As pointed out in the 'Introduction' section, many discrepant results in the literature do not allow us to reach a consensus on the value of the MS scatter as a function of the stellar mass. Thus, we assume an average dispersion of 0.3 dex and apply the corresponding correction in the further analysis to all MS estimates based on the median SFR to convert them into mean values.

2.2 The main corrections

We describe here the corrections applied to bring the different MS estimates to a common framework:

(i) **IMF correction**

We apply the IMF offsets to stellar masses as in S14 with the form:

$$M_{\star,K} = 1.06\, M_{\star,C} = 0.62\, M_{\star,S}, \tag{1}$$

with the subscripts referring to Kroupa, Chabrier, and Salpeter IMFs, respectively. These correspond to stellar mass offsets of 0.03 and -0.21 dex, for the Chabrier and the Salpeter IMF, respectively.

(ii) **SFR indicator correction**

All SFR estimates have been converted to the calibration of KE12, which is based on the Kroupa IMF. The KE12 prescriptions are different depending on whether νL_ν is estimated in the FUV ($\sim$1300–1700 Å) or NUV ($\sim$2300–2800 Å). Following tables 1 and 2 of KE12, we make use of the following calibrations:

$$SFR(FUV + IR) = 1.71 \cdot 10^{-10}(L_{FUV} + 0.46 * L_{IR}) \tag{2}$$

$$SFR(NUV + IR) = 2.06 \cdot 10^{-10}(L_{NUV} + 0.27 * L_{IR}) \tag{3}$$

$$SFR(H\alpha) = 2.06 \cdot 10^{-8}(L_{H\alpha}) \tag{4}$$

$$SFR(IR) = 1.49 \cdot 10^{-10}(L_{IR}) \tag{5}$$

$$SFR(1.4GHz) = 2.43 \cdot 10^{5}(L_{1.4GHz}) \tag{6}$$

where L_{IR}, L_{FUV}, L_{NUV}, and $L_{H\alpha}$ are the luminosities (in solar luminosity units) estimated in the IR (range 8–1000 μm), FUV and NUV, $H\alpha$ emission line, and radio emission (1.4 GHz), respectively. The SFRs based on SED modelling, instead, are corrected only for the IMF, according to the derivative of equation (1). The correction to a common SFR indicator varies from 0.05 to -0.2 dex. After applying the calibration, the local MS of Chang et al. (2015) is still systematically lower with respect to the other local MS estimates by 0.15 dex (see Popesso et al. 2019b for a detailed comparison with UV + IR based and H α-based SFR). Thus, we correct for this offset as indicated in Popesso et al. (2019b) before including the MS estimate of Chang et al. (2015) into our sample.

As already pointed out in Popesso et al. (2019b), Pearson et al. (2018) report a systematic offset of their SFR of $\sim$0.4 dex below all other SFR estimates considered here at the same redshift. As shown in Popesso et al. (2019b), their MS lies below all other determinations at more than 1σ at all redshift. Elbaz et al. (2011) show that SPIRE and PACS SFR estimators lead to consistent results. So we conclude that the discrepancy must be related to the deblending technique of the SPIRE detections and the SED fitting technique applied in Pearson et al. (2018, see their appendix C for an extensive discussion). Since the problem might be related to an over-deblending issue rather than to the SFR indicator, we decide not to correct the SFR of Pearson et al. (2018) and to exclude those MS estimates from the data set.

In addition, two of the considered publications (Thorne et al. 2021; Leja et al. 2022) based on two different SED fitting algorithms, *ProSpect* (Robotham et al. 2020) and *Prospector* (Leja et al. 2017), report lower values of SFR and larger values of stellar masses with respect to the rest of the MS estimates considered here. This discrepancy is due to a different reconstruction of the galaxy SFH with respect to other SED fitting methods. The main difference consists in the fact that both codes include the contribution of an extra component of stars older than 100 Myr, which would affect both the stellar mass and the SFR estimates. The effect on the stellar masses is obvious and it is reported to be 0.2 dex for Thorne et al.

(2021) and 0.3 dex for Leja et al. (2022). Leja et al. (2022), in particular, also report that the derived SFRs tend to be lower at fixed stellar mass with respect to *IR* or $UV + IR$-derived SFR in the redshift range $0.5 < z < 3$. They ascribe such a difference to the contribution of the extra component of old stars to the dust heating, which would increase the galaxy IR emission. They also find that such contribution is redshift-dependent, being larger at $0.5 < z < 3$ than at lower and higher redshifts. Leja et al. (2022) conclude that *IR* or $UV + IR$-derived SFRs are overestimated at $0.5 < z < 3$, because not all the IR emission is due to dust heating by young stars and, thus, to star formation.

The KE12 calibration does take into account that only a fraction of the *IR* emission is due to star formation. However, it assumes that such a fraction is the same at all redshifts. If this assumption is not valid, it implies that at higher redshift, the *IR* or $IR + UV$ SFR indicators based on the KE12 calibration might lead to over or underestimated SFRs with respect to the radio or H α-derived SFRs, based on the same calibration. This is because the contribution of an extra old star component might significantly affect the IR emission, but it has rather a negligible effect on the H α and radio emission, which are dominated by H II region nebular emission and non-thermal (synchrotron) emission due to Type II and Type Ia Supernovae, respectively. As shown in Appendix B, we do not observe such discrepancy at any redshift. Instead, the KE12 calibrated SFR derived through the radio, H α and $IR + UV$ emission are consistent from $z \sim 0$ up to $z \sim 3$. Thus, we do not find clear evidence of an evolution of the fraction of *IR* emission contributing to the SFR. In addition, we point out that Thorne et al. (2021) show that once the stellar mass discrepancy is taken into account, the SFRs derived with *ProSpect* are perfectly consistent with the radio emission-based SFRs of Leslie et al. (2020). We perform the same exercise between the *Prospector* MS estimates of Leja et al. (2022) and those of Whitaker et al. (2014) based on $IR + UV$ emission on the same 3DHST galaxy sample in the CANDELS fields (see Fig. A1 in Appendix A). Once the stellar masses are corrected for the discrepancy (0.3 dex), the MS estimates are consistent within 1–1.5σ at all stellar masses. Thus, we include the Thorne et al. (2021) and Leja et al. (2022) MS estimates by correcting the stellar masses by the reported discrepancy, and we leave unaltered the SFR estimates.

(iii) **Cosmology correction**: This is calculated as the ratios between luminosity distance, $d_L(z)$, derived from two different cosmologies, and, given the observed redshift range of a sample, applying a d_L^2 correction at the expected median z of galaxies in the sample. S14 also estimate first-order volume effects. However, they find that this account for a negligible effect in all cases. Thus, we do not take it into account.

After applying the calibration, we obtain an inter-publication scatter of 0.08 dex per bin of time and stellar mass. We also observe that the scatter is not much dependent on the SFR indicator or methodology used in the different publications. If we limit the analysis only to publications based on *IR* or $UV + IR$-based SFRs or only to those based on the SED fitting technique, we do not observe any scatter variation.

2.3 Selection effects

While efficient at selecting SFGs, most selection techniques differ from each other and do not all select the same population. S14 discuss extensively that the selection of blue objects preferentially select actively star-forming, non-dusty galaxies and exclude a large percentage of galaxies that are classified as SFGs via other selection

mechanisms (e.g. colour-colour selection). This leads on average to larger MS slopes than the ones retrieved with other selection methods. One of the methods to retrieve blue sources is the Lyman break technique (Steidel et al. 1999; Stark et al. 2009; Bouwens et al. 2011), which used to select high-*z* LBGs. In the list of publications considered here, only Tasca et al. (2015) apply the LBGs selection technique. Tomczak et al. (2016) compare the results of Tasca et al. (2015) with those based on the 'mixed' selection method (see also discussion below) and find very good agreement. We conclude that the results of Tasca et al. (2015) seem less affected by the bias of the bluer selection discussed by S14. Similarly, Heinis et al. (2014) apply a UV selection to identify distant SFGs. Nevertheless, their estimates are not scattering significantly with respect to the other relations. Thus, we conclude that in this case also, the UV selection does not affect significantly the slope of the relation. Thorne et al. (2021) applies a sSFR cut, which might similarly select blue galaxies. However, the authors report a very good agreement with the MS estimates of Leslie et al. (2020), which is based on the radio selection. Thus, we conclude that in this case also the bias is negligible.

All other MS estimates included in our analysis are based on methods that S14 classify as 'mixed'. Such techniques include redder objects in the selection, thus also considering a large portion of the SFG population dominated by dust. These methods are considered to provide a more physical distinction between SFGs and quiescent galaxies (Ilbert et al. 2013; Schreiber et al. 2015). Among these 'mixed' methods, we include the colour-colour selection based on the rest-frame (U-V) − (V-J) and (M_{NUV}-M_R) − (M_R-M_J) absolute colours, the 2σ clipping, the bimodality between SFGs, and quiescent galaxies in the SFR-$M_\star$ plane. While these different selection methods do not seem to affect the average observed SFRs across different publications, as pointed out by S14, they do seem to influence the derived slopes and the intrinsic scatter of the MS. Popesso et al. (2019a) show that all these methods agree very well over most of the stellar mass range considered. However, small discrepancies can be observed due to little selection biases at very high stellar masses towards high redshift. Namely, selections as the (U-V) − (V-J) colour selection tend to exclude part of the high mass SFGs at relatively lower SFR with respect to the MS location. This would lead to a steeper MS at higher redshift with respect to the low-redshift relation and thus a more significant evolution. The effect is an offset of ∼0.15–0.2 dex at $M_\star$ of 10^{11}–$10^{11.5}$ M$_\odot$. Instead, the selection based on the (M_{NUV}-M_R) − (M_R-M_J) absolute colours is less prone to this selection bias, as it allows to select all galaxies populating the log-normal distribution around the MS location (Ilbert et al. 2015; Popesso et al. 2019a).

We point out that the MS estimates affected by this bias are the stacked points based on the CANDELS field data set (Whitaker et al. 2014; Schreiber et al. 2015; Tomczak et al. 2016). Due to the very small volume sampled by the CANDELS fields, these include less than 15 galaxies per stacked point at stellar masses larger than 10^{11} M$_\odot$. Since this is below the limit required in Section 2, those stacked points are anyhow not included in our analysis.

3 FITTING THE MAIN SEQUENCE

In this work, we adopt two approaches to fit the MS relation. The first one follows S14 and consists in looking for the functional form that best describes the MS relation. As an alternative approach, we investigate the functional form proposed by Lee et al. (2015), which has the advantage of being expressed with parameters that have a physical description. In the following procedure, after correcting stellar masses and SFRs with the calibration described above, we

take into account the redshift and mass ranges of each study, and we include in the fitting procedure only objects or stacked points actually observed at given mass and redshift without any extrapolation or interpolation. These mass ranges have either been taken directly from the paper in question or estimated based on the data included in the relevant fits, rounded to the nearest 0.1 dex after excluding outlying points (see Appendix A for a detailed description of the data taken from each publication).

3.1 S14 approach

In this section, we describe how we fit the MS and retrieve the functional form that best expresses the evolution of the slope and normalization as a function of time. We follow a revised version of the S14 approach. We proceed to fit the evolution of the SFR provided by each individual MS as a function of time:

$$logSFR(t) = \alpha_i t + \beta_i, \qquad (7)$$

The choice of fitting as a function of time rather than redshift is mainly practical, as a straightforward linear fit works very well for the time variable, while a more complicated functional form is required as a function of redshift.

By fitting α_i's and β_i's for a grid of $M_\star$, as shown in Fig. 1, we can derive a function of the form:

$$logSFR(t, logM_\star) = \alpha(logM_\star)t + \beta(logM_\star), \qquad (8)$$

assuming a given parametrization for $\alpha(logM_\star)$ and $\beta(logM_\star)$.

As shown in the right-hand panel of Fig. 1, the slope $\alpha(logM_\star)$ depends linearly on $logM_\star$, consistently with S14. Conversely, the best-fitting form for $\beta(logM_\star)$ is not a simple linear dependence but it requires a quadratic form. This is because, differently for nearly all the MS compiled by S14, most of the MS estimates included here find a bending of the MS towards large stellar masses. Thus, the best-fitting functions for $\alpha(logM_\star)$ and $\beta(logM_\star)$ are, respectively:

$$\alpha(logM_\star) = a_0 + a_1 logM_\star$$
$$\beta(logM_\star) = b_0 + b_1 logM_\star + b_2 log^2 M_\star \qquad (9)$$

which gives:

$$logSFR(t, logM_\star) = (a_1 t + b_1)logM_\star$$
$$+ b_2 log^2 M_\star + (b_0 + a_0 t). \qquad (10)$$

Equation (10) differs from the best-fitting function of S14 for the quadratic term $b_2 log^2 M_\star$, which is time-independent. The slope of the MS is driven by the combination of the quadratic term and the linear term that sets the faint-end slope, $(a_1 t + b_1)logM_\star$, which evolves with time. The normalization of the relation depends linearly on time through the term $(b_0 + a_0 t)$.

We use the functional form retrieved in equation (10) to fit the whole MS data set without binning in stellar masses and time. We limit, however, the fitting procedure to the $10^{8.7}$–$10^{11.3}$ M$_\odot$ range, where the MS included in our collection were declared to have high completeness in stellar mass. The best-fitting parameters of equation (9) for $\alpha(logM_\star)$ and $\beta(logM_\star)$ are used as the first guess for the fitting procedure. The final best-fitting parameters are given in Table 2, and the best-fitting MS is shown as a function of time in Fig. 2. The bottom panel shows the residual distribution of the MS estimates with respect to the best fit at a given time and stellar mass. The scatter around the best fit is 0.09 dex. The MS is found to bend towards large stellar masses at all times due to the time-independent quadratic term. The time variation of the faint-end slope makes the relation steeper at early epochs. Nevertheless, the evolution of the
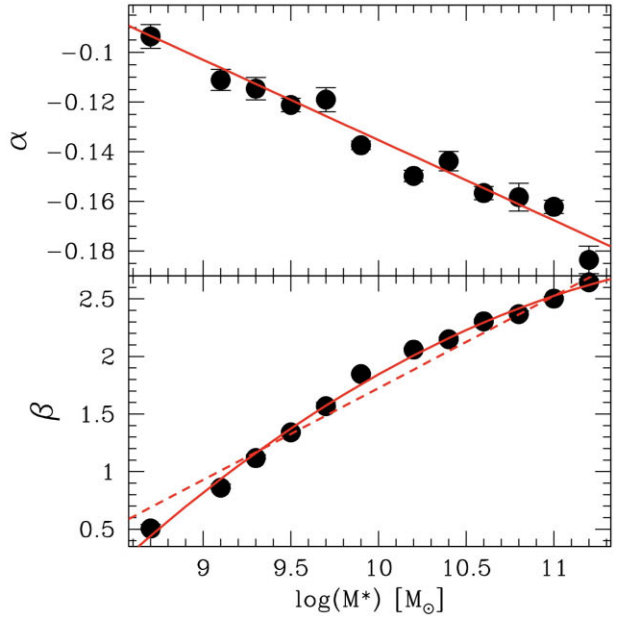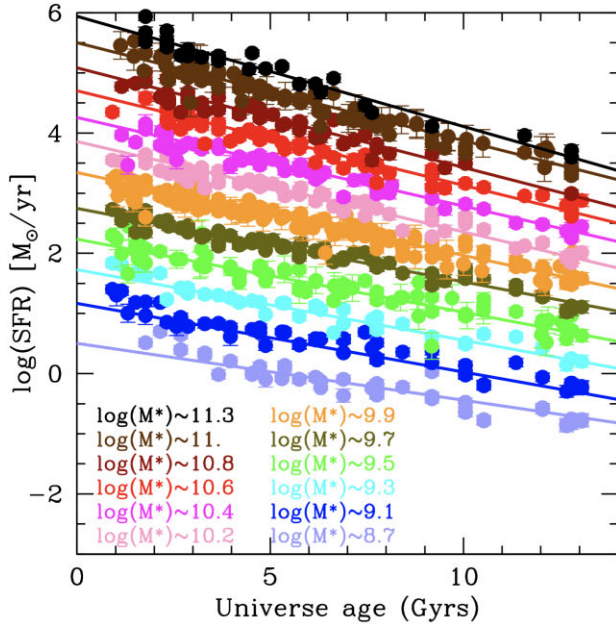
Figure 1. *Left-hand panel:* logSFR versus the Universe age in several bins of stellar masses. The data points indicate the SFR based on the MS estimates collected in this work. The solid lines show the best linear fit as in equation (9). Data points and lines are colour-coded as a function of the stellar mass bin as indicated in the figure. For clarity, the relations are artificially displaced by 0.4 dex from one another. *Right-hand panel:* $\alpha(logM_\star)$ (upper panel) and $\beta(logM_\star)$ (lower panel) as a function of $logM_\star$. The red solid lines in both panels indicate the best-fitting relations of equation (9). The dashed line in the bottom panel shows the liner fit approximation as proposed in S14.

Table 2. The table lists the best-fitting parameters of equations (10) and (14) in the first two columns. The last column lists the best-fitting parameters of equation (15), which corresponds to equation (14) with the slope $a4 = 1$ (γ in equation 11).

	Equation (10)		Equation (14)		Equation (15)
a0	0.20 ± 0.02	a0	2.693 ± 0.012	a0	2.71 ± 0.01
a1	-0.034 ± 0.002	a1	-0.186 ± 0.009	a1	-0.186 ± 0.007
b0	-26.134 ± 0.015	a2	10.85 ± 0.05	a2	10.86 ± 0.03
b1	4.722 ± 0.012	a3	-0.0729 ± 0.0024	a3	-0.0729 ± 0.0016
b2	-0.1925 ± 0.0011	a4	0.99 ± 0.01	a4	1

slope is marginal after the first ∼4 Gyrs, consistently with the results of S14, and as found in Popesso et al. (2019a). The normalization of the relation evolves consistently with the results of S14. We point out that it evolves roughly as $(1 + z)^3$. However, this is only an approximation, because the SFR(z) can not be expressed accurately as a power law of $(1 + z)$. The observed evolution is consistent with previous results in the literature, where there is general consensus on the MS normalization evolving as $(1 + z)^{2.8 - 3}$ at least up to z ∼ 4 (e.g. Sargent et al. 2012; Schreiber et al. 2015; Davidzon et al. 2017). Such evolution has been interpreted as reflecting the evolution of the molecular gas mass density and of the consequent availability of gas supply for the galaxy star formation process (Daddi et al. 2008, 2010; Tacconi et al. 2010, 2018).

3.2 The MS turn-over

The approach of S14 is very effective in providing a good functional form for the MS. However, it is not trivial to physically interpret equation (10). For this reason, we explore a different approach by using the fitting function of Lee et al. (2015):

$$SFR(M_\star) = SFR_{max}/(1 + (M_\star/M_0)^{-\gamma}) \qquad (11)$$
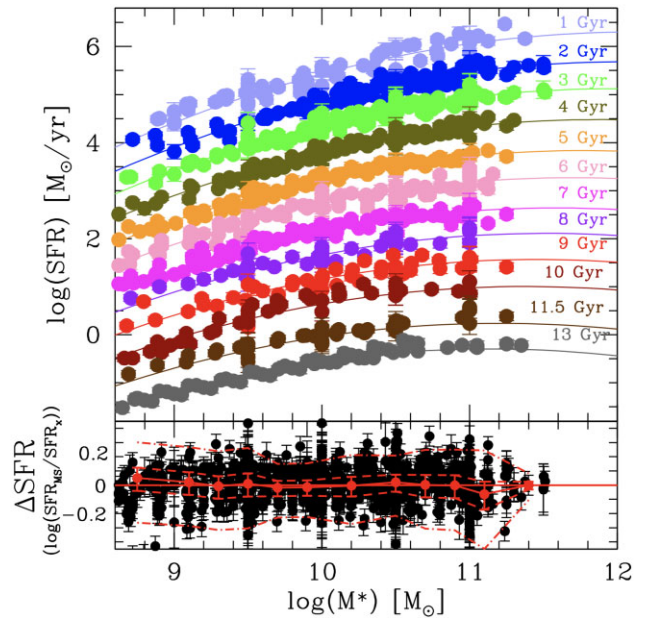
Figure 2. The upper panel shows the MS shape as a function of the Universe age from 1 to 13 Gyrs. The data points show the MS estimates collected and calibrated in this work. The solid lines indicate the best-fitting relation obtained with equation (10). Points and lines are colour-coded as a function of time, as indicated in the panel. For clarity, the MS relations are artificially displaced by 0.4 dex from one another. The bottom panel shows the residual distribution as a function of stellar mass. The dashed red line shows the 0 level, corresponding to the best-fitting value. The red points indicate the mean value of the residuals, while the dotted and dashed-dotted lines indicate the 1σ and 3σ regions, respectively.
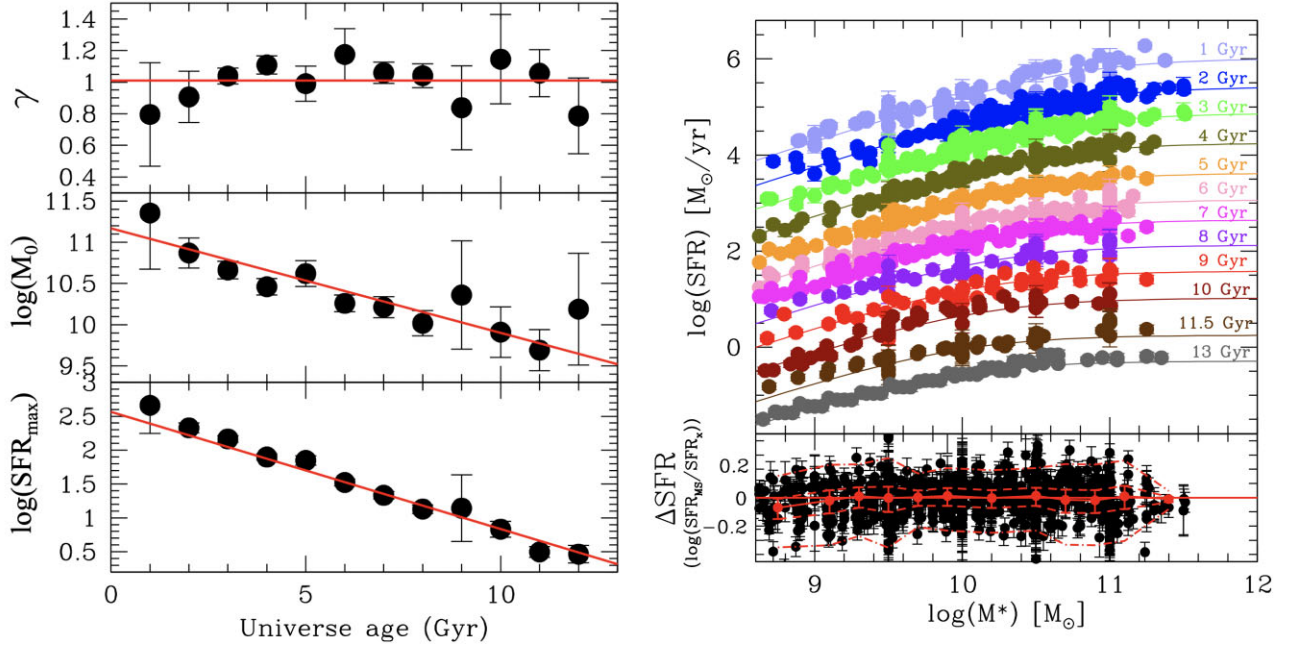
Figure 3. *Left-hand panel:* evolution of γ (upper panel), $log(M_0)$ (central panel) and $log(SFR_{max})$ (bottom panel) as a function of time. The red solid lines in each panel indicate the best-fitting relations: a constant equal to 1 for γ and the best fit of equations (12) and (13) for M_0 and SFR_{max}, respectively. *Right-hand panel:* the upper panel shows the MS shape as a function of the Universe age from 1 to 13 Gyrs. The data points show the MS estimates collected and calibrated in this work. The solid lines indicate the best-fitting relation obtained with equation (14). Points and lines are colour-coded as a function of time, as indicated in the panel. For clarity, the MS relations are artificially displaced by 0.4 dex from one another. The bottom panel shows the residual distribution as a function of stellar mass. The dashed red line shows the 0 level, corresponding to the best-fitting value. The red points indicate the mean value of the residuals, while the dotted and dashed-dotted lines indicate the 1σ and 3σ regions, respectively.

Unlike polynomial fits, as equation (10), the parameters of this model allow to quantify the interesting characteristics of the relation between stellar mass and SFR: (i) γ, the power-law slope at low stellar masses, (ii) M_0, the turnover mass, and (iii) $logSFR_{max}$, the maximum value of $logSFR$ that the function asymptotically approaches at high stellar masses.

To capture the evolution of the MS through equation (11), we first fit the MS in bins of time to estimate the time dependence. This allows to obtain $S_0(t)$, $M_0(t)$, and $\gamma(t)$ to find the best functional form, which includes the time dependence in equation (11). The left-hand panel of Fig. 3 shows the results of the best-fitting parameters as a function of time. We find that the exponent γ does not show any time dependence and it is consistent with the value -1. Both $logSFR_{max}$ and $logM_0$ depend linearly on time. Thus, they can be expressed as:

$$logSFR_{max}(t) = a_0 + a_1 t \qquad (12)$$

and

$$logM_0(t) = a_2 + a_3 t \qquad (13)$$

By taking into account such evolution, we can write equation (11) as a function of time as:

$$logSFR(M_\star, t) = a_0 + a_1 t - log(1 + (M_\star/10^{a_2+a_3 t})^{-a_4}) \qquad (14)$$

We use this functional form to fit the MS estimate data set as a function of mass and time without any binning. The results of the best-fitting procedure are indicated in Table 2. The right-hand panel of Fig. 3 shows the MS data points as a function of time with the best-fitting curves. The scatter around the best fit is 0.09 dex, indicating that equation (14) also provides an excellent fitting function for the evolution of the MS. Consistently with the results obtained with the functional form given by equation (10), the MS bends at the high mass

end with a turn-over mass that is evolving with time. The exponent γ does not evolve with time and it is consistent with the value 1. This implies that the MS is well represented by the functional form:

$$SFR(M_\star) = SFR_{max}(t)/(1 + (M_0(t)/M_\star)) \qquad (15)$$

which is regulated by only two parameters, $M_0(t)$ and $SFR_{max}(t)$. We also estimate the best-fitting parameters for this functional form. The results are included in Table 2. The evolution of the MS shape is regulated by the change of the turn-over mass $M_0(t)$. This change of only 25 per cent over the past 9–10 Gyrs, and is a factor of 2 larger in the first 3–4 Gyrs. The normalization of the relation, set by $SFR_{max}(t)$ parameter, exhibits the strongest evolution and it is consistent within 1σ with the value found for the normalization in equation (10).

Fig. 4 shows the comparison of the best fit obtained with equations (10) and (14). The two fits exhibit a very consistent evolution in normalization as a function of time. We do observe a discrepancy only at the very low mass and high mass ends, in particular at early epochs, where the uncertainty and scatter of the data points are the highest. Nevertheless, the two fits are consistent with each other within the 1σ uncertainty.

3.3 Towards a physical explanation of the MS shape evolution
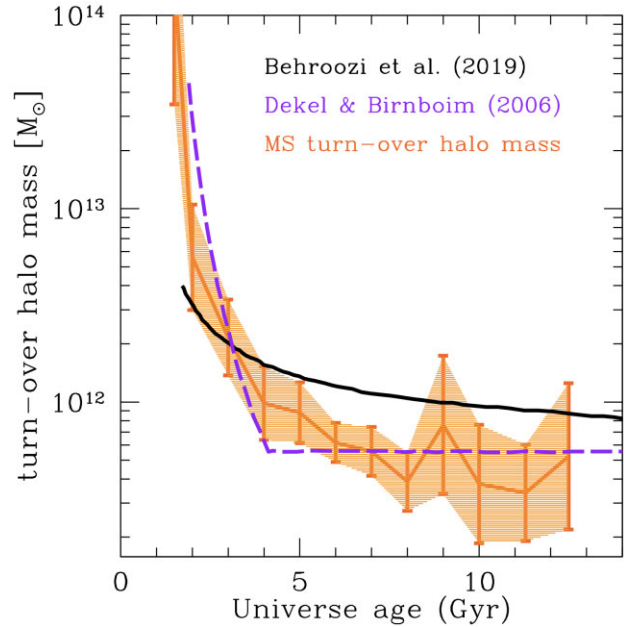
Given equation (15), $M0(t)$ can be interpreted as the mass thresholds between two regimes. At $M_\star << M_0(t)$, the sSFR of galaxies is nearly constant as a function of stellar mass and equal to $\sim SFR_{max}(t)/M_0(t)$. At $M_\star >> M_0(t)$, the sSFR is progressively suppressed as it is approximated by $SFR_{max}(t)/M_\star$. Thus, the turn-over stellar mass separates a regime of constant SFR per unit of stellar mass from a regime of SFR suppression.
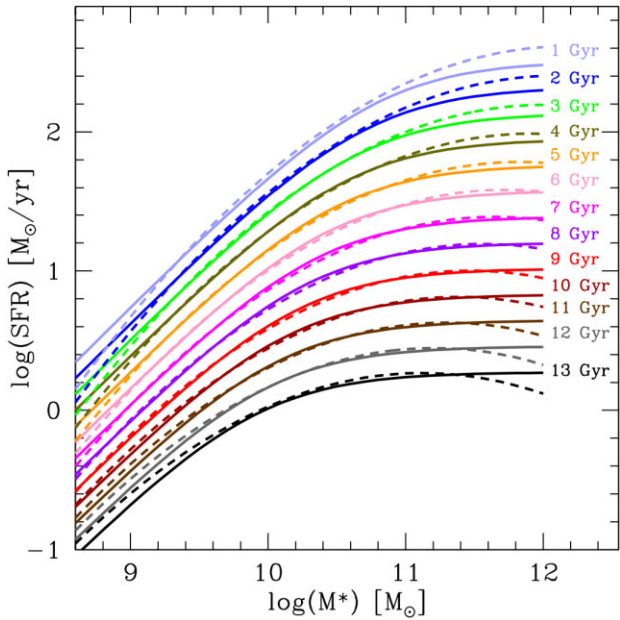
Figure 4. Comparison of the two MS best fits as a function of time obtained through equation (10; dotted lines) and equation (14; solid lines). All lines are colour-coded as a function of the Universe age as indicated in the figure from 1 to 13 Gyrs.

Figure 5. Evolution of the host halo turn-over mass (orange line) as a function of the Universe Age. The shaded orange region indicates the 1σ uncertainty. This is estimated by combining the error on $M_0(t)$ and the uncertainty in the central galaxy stellar mass and host halo mass correlation of UNIVERSEMACHINE (Behroozi et al. 2019). The solid black line indicates the evolution of the quenching halo mass of Behroozi et al. (2019). The dashed magenta line shows the evolution of the halo mass threshold between cold and hot accretion regime of Dekel & Birnboim (2006).

Popesso et al. (2019b) point out, exploiting the halo mass catalog of Yang et al. (2007), that at $z \sim 0$ the region of the MS (within 3σ from the relation) is completely dominated by central galaxies. Due to the rather tight correlation between the central galaxy stellar mass and the host halo mass (Mh), this implies that the SFG mean host halo mass is increasing along the MS with M$_\star$. Up to $z \sim 1.3$, we are able to check if this holds by exploiting the cosmic web catalog of Darvish et al. (2017) in the COSMOS field. This is based on the accurate photometric redshifts of the COSMOS15 catalog (Laigle et al. 2016) to identify clusters, groups, and filaments and assign a membership probability to each galaxy up to $z \sim 1.3$ and down to stellar masses of 10^9 M$_\odot$. For galaxies with a high probability to be in groups and clusters, the catalog also provides a classification in central and satellite galaxies, by identifying as central the most massive system. We use the IR-selected galaxy catalog of Popesso et al. (2019a) in the COSMOS field, based on the combination of *Herschel* and *Spitzer* MIPS data, to check the central galaxy fraction in the MS region in the redshift window explored by Darvish et al. (2017). The SFR of each galaxy is given by the combination of the UV and IR contribution (Popesso et al. 2019b). The MS is identified as the region within 3σ from the MS relation given by equation (14) in several redshift bins. We assume $\sigma = 0.3$ dex. At all redshifts up to $z = 1.3$, the MS region turns out to be dominated by central galaxies, which account for 70 per cent of the galaxy population. To check whether the MS region is dominated by central galaxies also at $z > 1.3$, we use the predictions from the Illustris TNG hydro-dynamical simulation (Pillepich et al. 2018a). In this case, we use as reference the MS determined as in Donnari et al. (2019) on the same data. Also, in this case, as expected, central galaxies account for 70–80 per cent of the galaxy population at $z > 1.3$. Thus, is it plausible to assume that the MS is dominated by central galaxies also in the distant Universe.

This aspect is crucial because it allows us to convert the turn-over stellar mass of the SFG MS into a *turn-over host halo mass*,

thanks to the correlation between central galaxy stellar mass and host halo mass (Yang et al. 2007; Behroozi, Wechsler & Conroy 2013; Behroozi et al. 2019). In particular, we use the results of the empirical model UNIVERSEMACHINE of Behroozi et al. (2019) to convert $M_0(t)$ into host halo turn-over mass $M_{h0}(t)$. In Fig. 5, we plot the evolution of $M_{h0}(t)$ as a function of time together with the halo mass quenching threshold derived in Behroozi et al. (2019), and the evolution of the transition mass between cold and hot accretion predicted by the theory of mass accretion as in Dekel & Birnboim (2006). The halo mass quenching threshold is defined as the halo mass above which the fraction of quenched galaxies is larger than 50 per cent. The hot/cold transition, instead, is defined as the halo mass at which the cold gas streams coming from the cosmic web filaments are no longer able to penetrate the halo and feed the central galaxy. The curve of Dekel & Birnboim (2006) depends on the halo temperature, but it predicts that at higher redshift ($z > 2.5 - 3$) cold gas streams are still able to penetrate massive hot haloes. Thus, it represents the threshold between hot and cold gas accretion onto the central galaxy. Behroozi et al. (2019) discuss that the disagreement between the empirical and the theoretical predictions might originate from the non-inclusion of the effects of black hole feedback in the treatment of accretion. This, indeed, might play an important role in affecting the thermodynamical conditions of the circumgalactic-medium of the central galaxy, mainly by injecting large quantities of energy into it (Weinberger et al. 2017; Nelson et al. 2019).

The evolution of M_{h0} with time is remarkably in agreement with the theoretical model of Dekel & Birnboim (2006). The evolution of $M_{h0}(t)$ decreases steeply in the first 4–5 Gyrs of the Universe and it reaches a plateau afterwards. This would suggest that the turn-over mass of the MS might be indicative of the transition between an environment that efficiently sustains the star formation process of

the central galaxies, e.g. through cold gas streams, to one which is hostile to the star formation process due to the suppression of the feeding mechanism of the central galaxy (see also Tacchella et al. 2016; Daddi et al. 2022). This suppression is likely maintained in the hot environments, as suggested by nearly all the most sophisticated hydrodynamical simulations, by the interplay between the hot gas in massive halos and central black hole feedback (Voit et al. 2015, 2017; Weinberger et al. 2017; Nelson et al. 2019).

4 COMPARISON WITH PREVIOUS RESULTS

In this section, we compare our results on the evolution of the MS with previous observational and theoretical results.

We first compare our results to the ones of S14, keeping in mind that the analysis of S14 is limited to the stellar mass range $10^{9.7}$–10^{11} M$_\odot$. In Fig. 6, we plot in magenta the homogenized collection of MS relations in seven different redshift bins. The green solid line marks our best-fitting estimates at the given redshift, as expressed in equation (10), while the green line indicates the best fit of equation (14). The MS relation of S14 is perfectly overlapping our relation in their limited stellar mass range. S14 claim that they do not find a bending of the MS at the high mass end. However, we point out that this could be due to the limited stellar mass range considered in the fitting procedure. Indeed, while a single power law might be a good approximation over this range, the extrapolation to larger masses would largely disagree with the data above 10^{11} M$_\odot$.

The bending of the MS has been largely discussed in the literature and the redshift and stellar mass at which the relation bends vary largely among the different publications included in our collection, from an MS bending only at $z < 2$ (e.g. Lee et al. 2018), to a correlation that becomes a power law at $z > 2$ (Schreiber et al. 2015; Tomczak et al. 2016), or an MS constant in shape and evolving only in normalization (Whitaker et al. 2012, 2014; Lee et al. 2015). In this work, we find that the MS bends at all redshifts. However, the bending happens above very large stellar masses at early epochs ($z > 3$). Works limited to lower stellar masses or to relatively small deep fields cannot capture this feature. In addition, some of the previous works focus on different aspects, from the analysis of the low mass slope of the relation in a few cases (e.g. Whitaker et al. 2012, 2014), to the shape and normalization at the high mass end in others (e.g. Sherman et al. 2021). We point out, though, that the bending observed in the MS estimated in this work is the result of the combination of all these different estimates, which all tend to be in agreement within a relatively small scatter (0.08 dex), when brought to a common framework. This, perhaps, might suggest that the reported discrepancies are mainly due to possible biases introduced by a limited stellar mass range, the limited volume of the studied deep fields, possible low number statistics, systematics due to SFR indicator, or a different fitting function. Thus, the approach suggested by S14 and implemented here has the potential of overcoming the limitations of the individual analysis and offers a broader and more complete view of the evolution of the MS over a much larger time interval and stellar mass range.

Fig. 6 also shows the comparison between our results and the predictions of the Illustris TNG300 simulation (Pillepich et al. 2018a). TNG300 is the largest volume simulated in the suite of Illustris TNG hydrodynamical simulations. The choice of TNG300 is driven by the need to sample a sufficiently large volume to capture the rare giant SFGs at the high mass end of the MS. The SFRs are averaged over 200 Myr and measured within a physical aperture of $2R_{star}$, where R_{star} is the stellar half mass radius (see also Donnari et al. 2019). The shaded region in Fig. 6 shows the distribution of the
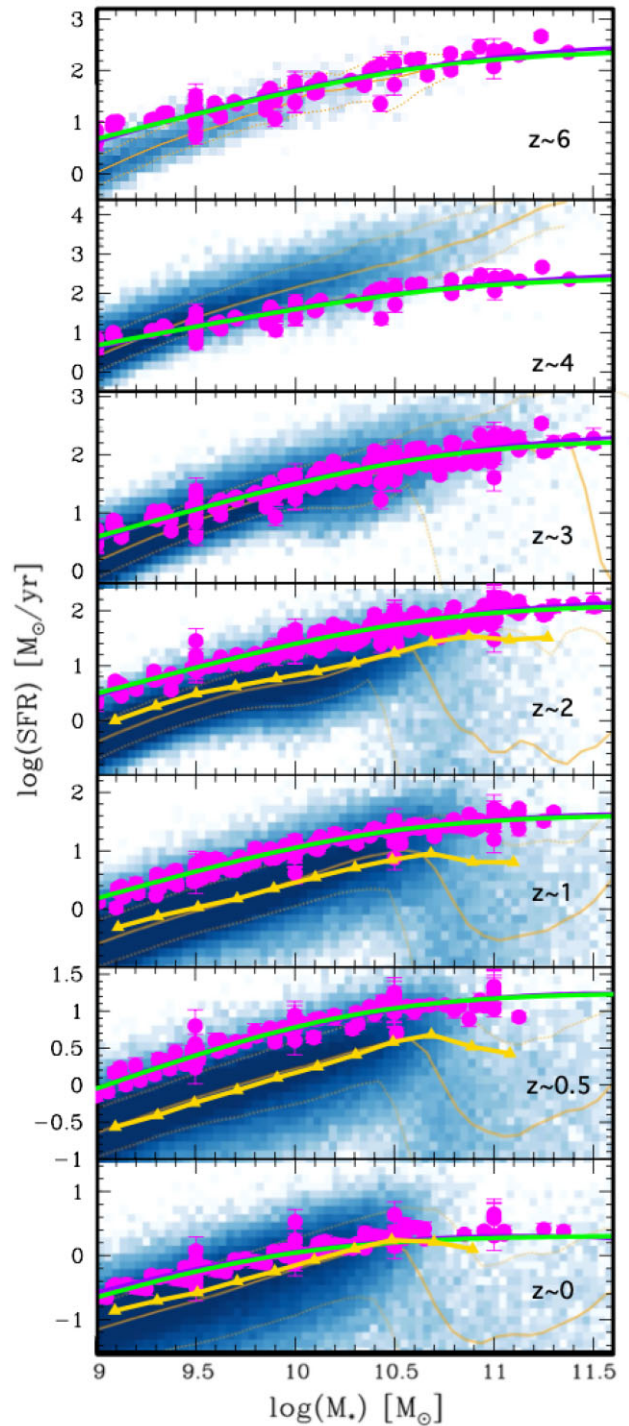
Figure 6. MS of SFGs in several redshift bins from $z \sim 6$ to $z \sim 0$. The shaded blue region in each panel indicates the distribution of simulated Illustris TNG galaxies in the SFR–stellar mass plane. The blue scale is according to the number density of galaxies. The orange line shows the running median in bins of stellar mass for the simulated galaxies, while the connected yellow triangles show the relation of Donnari et al. (2019) limited to the UVJ selected galaxies in Illustris TNG. The magenta points show the data collected in this paper. The green and purple solid lines indicate our best fit according to equations (10) and (14).

IllustrisTNG galaxies, colour-coded according to the galaxy number density in bins of SFR and $M_\star$. The IllustrisTNG MS is estimated as a running mean as a function of $M_\star$ and it is indicated by the orange-yellow line. We also show in yellow the MS of Donnari et al. (2019), which is estimated up to $z \sim 2$ by mocking the UVJ selection of Whitaker et al. (2012). The two estimates agree remarkably well up to stellar masses of $\sim 10^{10.5}$ $M_\odot$ and diverge at larger stellar masses, where the UVJ selection of Donnari et al. (2019) exclude most of the low-SFR systems.

It is interesting to notice that a bending of the predicted MS is observed also in Illustris TNG at least up to $z \sim 3$. This was already pointed out by Donnari et al. (2019) at $z < 2$. At higher redshift, the relation is steeper than the observed MS, but this could be due to a lack of massive SFGs in the simulations, as pointed out below.

We notice a good agreement between the IllustrisTNG predictions and our results at early epochs ($z \sim 6$). The observed MS lies over the simulated relation. However, no giant SFGs are observed at all in the TNG300 volume at $z \sim 6$ above $10^{10.5}$ $M_\odot$. The observations, instead, indicate that such galaxies populate the high mass end of the relation, in agreement with the results of the Galaxy Stellar Mass Function (GSMF) of active galaxies of Davidzon et al. (2017). This is also confirmed by the recent *JWST* discoveries of very massive blue galaxies in the distant Universe (e.g. Castellano et al. 2022; Donnan et al. 2022; Finkelstein & Bagley 2022; Santini et al. 2022).

Massive SFGs appear in the TNG simulation only between $z \sim 4$ and ~ 3, with a larger SFR compared to observations, in particular at $z \sim 4$. At this epoch, the predicted and observed low mass slope (below stellar masses of $10^{10.5}$ $M_\odot$) are in agreement within 1σ. At later epochs ($z < 2$), the predicted MS is systematically below the observations by ~ 0.5 dex, according to the results of Donnari et al. (2019). Observations and predictions are again in agreement at $z \sim 0$ up to $\sim 5 \times 10^{10}$ $M_\odot$. Above this mass, no more SFGs are detected in the simulations at odds with observations.

The tension between observations and simulations in the evolution of the MS is not new and has been extensively discussed in the literature (see for instance, Katsianis et al. 2020; Nelson et al. 2021). From the observational point of view, the use of sophisticated SED fitting codes, such as the *Prospector* or its incarnation of *Prospector-α* used in Leja et al. (2022), might apparently solve this issue. Indeed, adding an extra population of old stars might justify an overestimation of the previous *Spitzer-* and *Herschel*-based SFR estimates and reconcile ad hoc the observed and simulated MS in specific redshift bins. However, it is worth pointing out that such an approach might simply move the tension somewhere else. Indeed, the stellar masses provided by *Prospector* increase considerably with respect to previous works. This increase would shift the stellar mass function leading to a substantial disagreement with the predicted ones. These, indeed, are overall in agreement with previous measurements, as shown, for instance, in Pillepich et al. (2018b) for Illustris TNG and in the detailed comparison of Thorne et al. (2021).

It is, perhaps, more interesting to point out that the largest disagreement between the observed and the simulated MS is in the distribution of the SFG population along the MS. Fig. 6 shows quite clearly that, in the observations, SFGs with stellar masses above $\sim 5 \times 10^{10}$ $M_\odot$ exhibit a much slower evolution than predicted by the simulations. Indeed, they should have formed by $z \sim 4$, while in Illustris TNG they do not appear at all. Furthermore, the high mass end of the MS, above this mass threshold, is almost completely evacuated by $z \sim 0$ in the simulation. Instead, many works in the literature, based on the SDSS galaxy spectroscopic sample or the WISE survey, show clearly that this region of the SFR-stellar
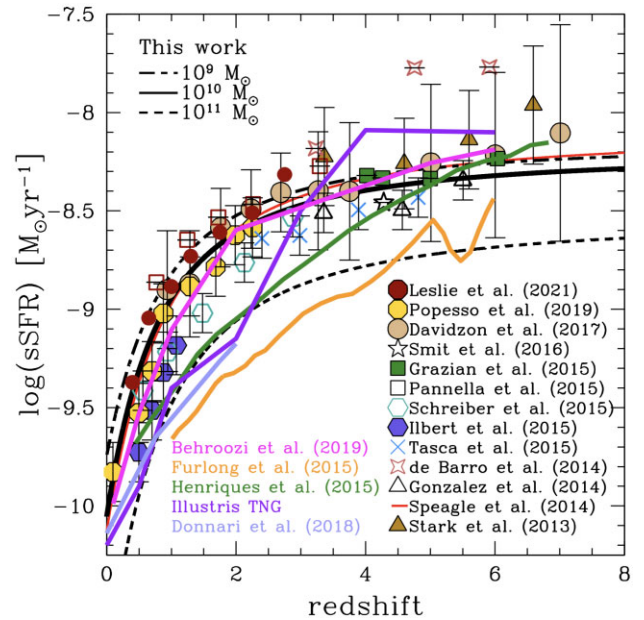
Figure 7. Evolution of sSFR as a function of redshift. The black lines show the sSFR evolution given by the best fit of equation (14) at three stellar mass values: 10^9 (dot–dashed line), 10^{10} (solid line) and 10^{11} $M_\odot$ (dashed line). We provide the same relation for the S14 best fit at 10^{10} (red line). We plot for comparison the data of several works in literature with symbols indicated in the figure. The yellow, green, purple, and light purple solid lines show the predictions of EAGLE simulation of Furlong et al. (2015), The MUNICH simulation of Henriques et al. (2015), the Illustris TNG as estimated here, and the one of Donnari et al. (2019), respectively. The magenta solid line shows the best-fitting model of Behroozi et al. (2019).

mass plane is highly populated up to stellar masses of 3×10^{11} $M_\odot$ (see for instance, Popesso et al. 2019b, for a collection of the available data sets in the local Universe). Such discrepancy cannot be ascribed to an overestimation of the observed SFR, as recently proposed (Leja et al. 2022). Indeed, this would simply lead to a discrepancy of the MS normalization rather than to the lack of the massive SFG subpopulation. Thus, it is, perhaps, more likely that the predicted fast evolution of massive SFG is due to an over-efficient SF quenching process in simulations. We point out that the use of the larger stellar masses estimated by *Prospector* would increase the observed discrepancy at the high mass end, not only in the local Universe but also at higher redshift ($0.5 < z < 2$).

We provide in Fig. 7 a comparison of the sSFR evolution with simulations, as well as other works in literature. We plot the evolution of the sSFR derived from the best fit of equation (14) computed in three different stellar mass bins: 10^9 (black dot–dashed line), 10^{10} (black solid line), and 10^{11} $M_\odot$ (black dashed line). The red line shows the result of S14 at 10^{10} $M_\odot$, which is in agreement with our result at the same stellar mass. The prediction of IllustrisTNG, up to $z \sim 6$, those from Donnari et al. (2019), up to $z \sim 2$, as well as the best fit of the empirical model UNIVERSEMACHINE (Behroozi et al. 2019), all computed for $M_\star = 10^{10}$ $M_\odot$, are shown in purple, lavender, and magenta, respectively. The simulated data of Illustris TNG are in agreement with observations at high redshift and in the local Universe, but the slope of the predicted sSFR evolution is steeper than observations, as pointed out in the previous paragraph. The redshift evolution of sSFR obtained from Behroozi et al. (2019) is perfectly overlapping with observations at the same stellar mass. Finally, we show in Fig. 7 the redshift evolution of the sSFR obtained from

the EAGLE hydrodynamical simulation (Furlong et al. 2015) and the Munich simulation of Henriques et al. (2015). Both predictions lie below the observations and exhibit a steeper relation. However, we point out that, as discussed by Davidzon et al. (2018), such comparison is complicated by the fact that simulated galaxies are not selected to be SFGs. Thus, the slope and normalization of the sSFR-redshift relation is biased by quiescent galaxies, which might have a faster evolution.

We conclude that the star formation activity of simulated galaxies is declining much faster than in observations between redshift $\sim$3 and $\sim$0. This leads to an MS normalization too low with respect to observations in the same redshift window. The disagreement is more significant for the most massive systems above 10^{11} M$_\odot$, which completely disappear from the simulated MS in the local Universe, at odds with observations. We do not find evidence for a clear overestimation or underestimation of the SFRs in specific redshift bins, which might solve the tension. As discussed in Donnari et al. (2019), the observed discrepancy might point to some fundamental limitations in our understanding of the processes governing the star formation evolution in galaxies, such as the role of cold and hot accretion and of supernovae and black hole feedback.

5 SUMMARY AND CONCLUSION

We compile a collection of the most important publications regarding the evolution of the MS of SFGs in the widest range of redshift ($0 < z < 6$), stellar mass ($10^{8.5}$–$10^{11.5}$ M$_\odot$), and SFR ($0 - 500$ M$_\odot yr^{-1}$) ever probed. We convert all observations to a common calibration to check for consistency in the literature estimates and to study the evolution of the relation with different approaches. We find a remarkably good agreement between the different estimates at any stellar mass and time. The resulting MS exhibits a curvature towards the high stellar masses, which is slowly evolving with time. We provide two functional forms, which take into account the time evolution of the MS normalization and slope. Following the approach of S14, we estimate the best polynomial fitting form in the $\log$SFR$-\log$M$_\star$ space, by studying the evolution with time of the $\log$SFR at fixed stellar mass. A second-order polynomial form is well representing the relation, with a non-evolving quadratic term. The normalization is evolving as a power law of the Universe age. The slow evolution of the linear term reproduces the steepening of the relation towards the first 3–4 Gyrs of the Universe. We provide, as an alternative fitting form, one of Lee et al. (2015), which has the advantage of being expressed by physical parameters. These are the low mass slope, the normalization, and the turn-over mass ($M_0(t)$). While the slope does not evolve with time, normalization and turn-over mass evolve as a power law of the Universe age. The turn-over mass, in particular, determines the MS shape. It marginally evolves with time, and it is responsible for the steepening of the relation towards $z \sim 4$–6. At stellar masses below $M_0(t)$, SFGs have a constant sSFR, while above $M_0(t)$, the sSFR is suppressed. As the MS region is dominated by central galaxies, we use the relation between central galaxy stellar mass and host halo mass to convert $M_0(t)$ into a 'turn-over host halo mass'. We find that its evolution is remarkably consistent with the one of the halo mass threshold between cold and hot accretion regimes predicted by Dekel & Birnboim (2006). This might indicate that $M_0(t)$ defines the transition between an environment able to sustain the SF process of the central galaxy, for instance through cold gas streams as predicted by Dekel & Birnboim (2006), to a regime hostile to the same process. The latter might be the result of the interplay between the hot gas in massive halos and the black hole feedback generated by the central galaxy itself.

The comparison of our results with the state-of-the-art hydrodynamical simulations shows that the simulated MS shape is qualitatively consistent with the observations. However, the normalization of the simulated relation is systematically lower by at least 0.2–0.5 dex with respect to observations at $0.5 < z < 3$. As a consequence, the sSFR of SFGs evolves more rapidly in simulated galaxies with respect to the observed ones. We do not find clear evidence for an overestimation of the observed SFR at $0.5 < z < 3$ with respect to simulated SFR. This might suggest that the feedback implemented in the simulations is likely too efficient in suppressing the star formation activity of galactic systems, in particular, above 5×10^{10} M$_\odot$ (see also Katsianis et al. 2021; Corcho-Caballero, Ascasibar & Scannapieco 2021).

DATA AVAILABILITY

The data underlying this article are publicly available in the articles listed in Table 1 and in their online supplementary material.

REFERENCES

Acquaviva V., Raichoor A., Gawiser E., 2015, ApJ, 804, 8
Baes M. et al., 2020, MNRAS, 494, 2912
Barro G. et al., 2019, ApJS, 243, 22
Behroozi P. S., Wechsler R. H., Conroy C., 2013, ApJ, 770, 57
Behroozi P., Wechsler R. H., Hearin A. P., Conroy C., 2019, MNRAS, 488, 3143
Belfiore F. et al., 2018, MNRAS, 477, 3014
Boogaard L. A. et al., 2018, A&A, 619, A27
Boquien M., Burgarella D., Roehlly Y., Buat V., Ciesla L., Corre D., Inoue A. K., Salas H., 2019, A&A, 622, A103
Bourne N. et al., 2017, MNRAS, 467, 1360
Bouwens R. J. et al., 2011, ApJ, 737, 90
Brinchmann J., Charlot S., White S. D. M., Tremonti C., Kauffmann G., Heckman T., Brinkmann J., 2004, MNRAS, 351, 1151
Bruzual G., Charlot S., 2003, MNRAS, 344, 1000
Bundy K. et al., 2015, ApJ, 798, 7
Byler N., Dalcanton J. J., Conroy C., Johnson B. D., 2017, ApJ, 840, 44
Calzetti D., Armus L., Bohlin R. C., Kinney A. L., Koornneef J., Storchi-Bergmann T., 2000, ApJ, 533, 682
Camps P. et al., 2018, ApJS, 234, 20
Carnall A. C., Leja J., Johnson B. D., McLure R. J., Dunlop J. S., Conroy C., 2019, ApJ, 873, 44
Castellano M. et al., 2022, ApJ, 938, L15
Chang Y.-Y., van der Wel A., da Cunha E., Rix H.-W., 2015, ApJS, 219, 8
Charlot S., Fall S. M., 2000, ApJ, 539, 718
Chen Y.-M., Wild V., Kauffmann G., Blaizot J., Davis M., Noeske K., Wang J.-M., Willmer C., 2009, MNRAS, 393, 406
Ciesla L., Elbaz D., Fensch J., 2017, A&A, 608, A41
Concas A., Popesso P., Brusa M., Mainieri V., Erfanianfar G., Morselli L., 2017, A&A, 606, A36
Conroy C., Gunn J. E., 2010, ApJ, 712, 833
Conroy C., Gunn J. E., White M., 2009, ApJ, 699, 486
Corcho-Caballero P., Ascasibar Y., Scannapieco C., 2021, MNRAS, 506, 5108

Curtis-Lake E., Chevallard J., Charlot S., Sandles L., 2021, MNRAS, 503, 4855

da Cunha E., Charlot S., Elbaz D., 2008, MNRAS, 388, 1595

Daddi E. et al., 2007, ApJ, 670, 156

Daddi E. et al., 2010, ApJ, 713, 686

Daddi E. et al., 2022, A&A, 661, L7

Daddi E., Dannerbauer H., Elbaz D., Dickinson M., Morrison G., Stern D., Ravindranath S., 2008, ApJ, 673, L21

Dale D. A. et al., 2009, ApJ, 703, 517

Dale D. A., Helou G., Magdis G. E., Armus L., Díaz-Santos T., Shi Y., 2014, ApJ, 784, 83

Darvish B., Mobasher B., Martin D. C., Sobral D., Scoville N., Stroe A., Hemmati S., Kartaltepe J., 2017, ApJ, 837, 16

Davidzon I. et al., 2017, A&A, 605, A70

Davidzon I., Ilbert O., Faisst A. L., Sparre M., Capak P. L., 2018, ApJ, 852, 107

Davies L. J. M. et al., 2016, MNRAS, 461, 458

Davies L. J. M. et al., 2018, MNRAS, 480, 768

de los Reyes M. A. et al., 2015, AJ, 149, 79

Dekel A., Birnboim Y., 2006, MNRAS, 368, 2

Delvecchio I. et al., 2017, A&A, 602, A3

Domínguez A. et al., 2013, ApJ, 763, 145

Donnan C. T. et al., 2022, MNRAS, preprint (arXiv:2207.12356)

Donnari M. et al., 2019, MNRAS, 485, 4817

Dunlop J. S. et al., 2017, MNRAS, 466, 861

Elbaz D. et al., 2007, A&A, 468, 33

Elbaz D. et al., 2011, A&A, 533, A119

Erfanianfar G. et al., 2016, MNRAS, 455, 2839

Finkelstein S. L., Bagley M. B., 2022, ApJ, 938, 25

Furlong M. et al., 2015, MNRAS, 450, 4486

Guo K., Zheng X. Z., Fu H., 2013, ApJ, 778, 23

Hao C.-N., Kennicutt R. C., Johnson B. D., Calzetti D., Dale D. A., Moustakas J., 2011, ApJ, 741, 124

Hayward C. C. et al., 2014, MNRAS, 445, 1598

Heinis S. et al., 2014, MNRAS, 437, 1268

Henriques B. M. B., White S. D. M., Thomas P. A., Angulo R., Guo Q., Lemson G., Springel V., Overzier R., 2015, MNRAS, 451, 2663

Ilbert O. et al., 2013, A&A, 556, A55

Ilbert O. et al., 2015, A&A, 579, A2

Iyer K. et al., 2018, ApJ, 866, 120

Iyer K., Gawiser E., 2017, ApJ, 838, 127

Karim A. et al., 2011, ApJ, 730, 61

Kashino D. et al., 2013, ApJ, 777, L8

Katsianis A. et al., 2020, MNRAS, 492, 5592

Katsianis A. et al., 2021, MNRAS, 500, 2036

Katsianis A., Tescari E., Wyithe J. S. B., 2016, PASA, 33, e029

Kennicutt R. C. Jr, 1998, ApJ, 498, 541

Kennicutt R. C., Evans N. J., 2012, ARA&A, 50, 531 (KE12)

Kewley L. J., Dopita M. A., Leitherer C., Davé R., Yuan T., Allen M., Groves B., Sutherland R., 2013, ApJ, 774, 100

Kriek M., Conroy C., 2013, ApJ, 775, L16

Kriek M., van Dokkum P. G., Labbé I., Franx M., Illingworth G. D., Marchesini D., Quadri R. F., 2009, ApJ, 700, 221

Kurczynski P. et al., 2016, ApJ, 820, L1

Laigle C. et al., 2016, ApJS, 224, 24

Le Fèvre O. et al., 2015, A&A, 576, A79

Lee B. et al., 2018, ApJ, 853, 131

Lee K.-S. et al., 2012, ApJ, 752, 66

Lee N. et al., 2015, ApJ, 801, 80

Leja J. et al., 2019, ApJ, 877, 140

Leja J. et al., 2022, ApJ, 936, 165

Leja J., Johnson B. D., Conroy C., van Dokkum P. G., Byler N., 2017, ApJ, 837, 170

Leslie S. K. et al., 2020, ApJ, 899, 58

Lovell C. C., Geach J. E., Davé R., Narayanan D., Li Q., 2021, MNRAS, 502, 772

Lower S., Narayanan D., Leja J., Johnson B. D., Conroy C., Davé R., 2020, ApJ, 904, 33

Lutz D., 2014, ARA&A, 52, 373

Madau P., Ferguson H. C., Dickinson M. E., Giavalisco M., Steidel C. C., Fruchter A., 1996, MNRAS, 283, 1388

Magdis G. E., Elbaz D., Daddi E., Morrison G. E., Dickinson M., Rigopoulou D., Gobat R., Hwang H. S., 2010, ApJ, 714, 1740

Martis N. S., Marchesini D. M., Muzzin A., Stefanon M., Brammer G., da Cunha E., Sajina A., Labbe I., 2019, ApJ, 882, 65

Meurer G. R., Heckman T. M., Calzetti D., 1999, ApJ, 521, 64

Momcheva I. G. et al., 2016, ApJS, 225, 27

Moustakas J. et al., 2013, ApJ, 767, 50

Murphy E. J. et al., 2011, ApJ, 737, 67

Narayanan D. et al., 2021, ApJS, 252, 12

Nelson D. et al., 2019, MNRAS, 490, 3234

Nelson E. J. et al., 2021, MNRAS, 508, 219

Nersesian A. et al., 2019, A&A, 624, A80

Noeske K. G. et al., 2007, ApJ, 660, L47

Noll S., Burgarella D., Giovannoli E., Buat V., Marcillac D., Muñoz-Mateos J. C., 2009, A&A, 507, 1793

Nordon R. et al., 2010, A&A, 518, L24

Oemler, Augustus J., Abramson L. E., Gladders M. D., Dressler A., Poggianti B. M., Vulcani B., 2017, ApJ, 844, 45

Oliver S. et al., 2010, MNRAS, 405, 2279

Oliver S. J. et al., 2012, MNRAS, 424, 1614

Pannella M. et al., 2009, ApJ, 698, L116

Pearson W. J. et al., 2018, A&A, 615, A146

Peng Y.-j. et al., 2010, ApJ, 721, 193

Pillepich A. et al., 2018a, MNRAS, 473, 4077

Pillepich A. et al., 2018b, MNRAS, 475, 648

Popesso P. et al., 2019a, MNRAS, 483, 3213

Popesso P. et al., 2019b, MNRAS, 490, 5285

Reddy N. et al., 2012, ApJ, 744, 154

Renzini A., Peng Y.-j., 2015, ApJ, 801, L29

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

Robotham A. S. G., Davies L. J. M., Driver S. P., Koushan S., Taranu D. S., Casura S., Liske J., 2018, MNRAS, 476, 3137

Rodighiero G. et al., 2011, ApJ, 739, L40

Rodighiero G. et al., 2014, MNRAS, 443, 19

Salim S. et al., 2007, ApJS, 173, 267

Salim S. et al., 2016, ApJS, 227, 2

Salmi F., Daddi E., Elbaz D., Sargent M., Dickinson M., Renzini A., Bethermin M., Le Borgne D., 2012, ApJ, 754, L14

Salmon B. et al., 2015, ApJ, 799, 183

Santini P. et al., 2009, A&A, 504, 751

Santini P. et al., 2017, ApJ, 847, 76

Santini P. et al., 2022, ApJL, preprint (arXiv:2207.11379)

Sargent M. T., Béthermin M., Daddi E., Elbaz D., 2012, ApJ, 747, L31

Schreiber C. et al., 2015, A&A, 575, A74

Sherman S. et al., 2021, MNRAS, 505, 947

Shim H., Chary R.-R., Dickinson M., Lin L., Spinrad H., Stern D., Yan C.-H., 2011, ApJ, 738, 69

Shivaei I., Reddy N. A., Steidel C. C., Shapley A. E., 2015a, ApJ, 804, 149

Shivaei I., Reddy N. A., Steidel C. C., Shapley A. E., 2015b, ApJ, 804, 149

Simha V., Weinberg D. H., Conroy C., Dave R., Fardal M., Katz N., Oppenheimer B. D., 2014, preprint (arXiv:1404.0402)

Skelton R. E. et al., 2014, ApJS, 214, 24

Smolčić V. et al., 2017, A&A, 602, A1

Sobral D., Best P. N., Smail I., Mobasher B., Stott J., Nisbet D., 2014, MNRAS, 437, 3516

Speagle J. S., Steinhardt C. L., Capak P. L., Silverman J. D., 2014, ApJS, 214, 15 (S14)

Spergel D. N. et al., 2003, ApJS, 148, 175

Stark D. P., Ellis R. S., Bunker A., Bundy K., Targett T., Benson A., Lacy M., 2009, ApJ, 697, 1493

Steidel C. C., Adelberger K. L., Giavalisco M., Dickinson M., Pettini M., 1999, ApJ, 519, 1

Steinhardt C. L. et al., 2014, ApJ, 791, L25

Tacchella S., Dekel A., Carollo C. M., Ceverino D., DeGraf C., Lapiner S., Mandelker N., Primack Joel R., 2016, MNRAS, 457, 2790

Tacconi L. J. et al., 2010, Nature, 463, 781

Tacconi L. J. et al., 2018, ApJ, 853, 179

Tasca L. A. M. et al., 2015, A&A, 581, A54

Theios R. L., Steidel C. C., Strom A. L., Rudie G. C., Trainor R. F., Reddy N. A., 2019, ApJ, 871, 128

Thorne J. E. et al., 2021, MNRAS, 505, 540

Tomczak A. R. et al., 2016, ApJ, 817, 118

Trayford J. W., Lagos C. d. P., Robotham A. S. G., Obreschkow D., 2020, MNRAS, 491, 3937

Valiante E. et al., 2016, MNRAS, 462, 3146

Viaene S. et al., 2017, A&A, 599, A64

Voit G. M., Bryan G. L., O'Shea B. W., Donahue M., 2015, ApJ, 808, L30

Voit G. M., Meece G., Li Y., O'Shea B. W., Bryan G. L., Donahue M., 2017, ApJ, 845, 80

Weinberger R. et al., 2017, MNRAS, 465, 3291

Whitaker K. E. et al., 2014, ApJ, 795, 104

Whitaker K. E., van Dokkum P. G., Brammer G., Franx M., 2012, ApJ, 754, L29

Willett K. W. et al., 2015, MNRAS, 449, 820

Yang G. et al., 2020, MNRAS, 491, 740

Yang X., Mo H. J., van den Bosch F. C., Pasquali A., Li C., Barden M., 2007, ApJ, 671, 153

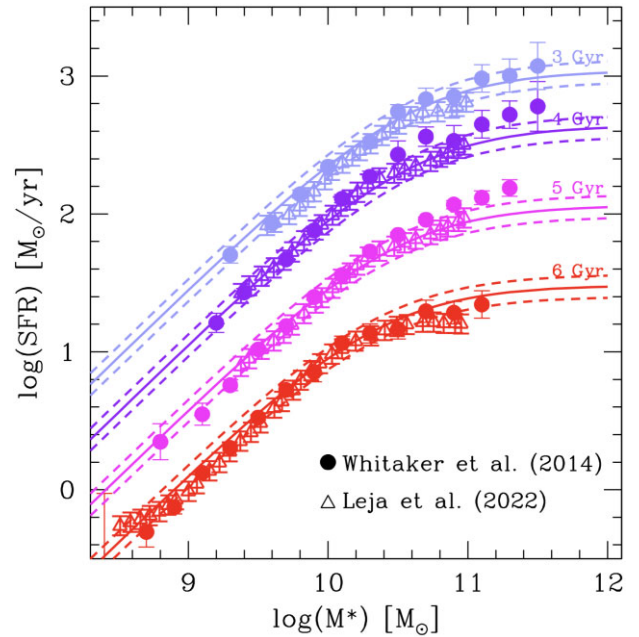Zahid H. J., Bresolin F., Kewley L. J., Coil A. L., Davé R., 2012, ApJ, 750, 120

APPENDIX A: THE SELECTED MS

Here, we list and describe the MS estimates included in the analysis and the data they are based on:

(i) Speagle et al. (2014) do not provide the collection of data used for calibrating the average MS estimates but only the best fits. A large number of best fits are provided based on different subsamples, depending on the SFG selection method, and on the range of stellar masses and cosmic time included in the fit. We use in this paper the fit n. 64 based on a 'mixed selection', which S14 consider as more inclusive of the SFG population (see Section 2.3 for a detailed discussion). The fit is restricted to stellar masses between $10^{9.7}$ and 10^{11} $M_\odot$ and it excludes the first 2 Gyrs of the Universe Age. In order to properly populate the S14 MS given by the fit n. 64, we use the data of Fig. 4, which shows the value of the SFR as a function of time in four stellar mass bins. We distribute randomly the data within the stellar mass bins and use the time information to estimate the SFR as a function of stellar mass and time in the range $10^{9.7}$–10^{11} $M_\odot$ and 2–12 Gyrs of the Universe age. The SFRs obtained in this way are calibrated to the KE12 calibration and a Kroupa IMF.

(ii) Rodighiero et al. (2014) provide the MS estimate of BzK selected galaxies in the redshift bin $1.4 < z < 2.5$ in the COSMOS field, on the basis of various SFR indicators, including UV emission, $H\alpha$ emission, mid-infrared (MIR) and far-infrared (FIR) emission. In this analysis, we use the four data points obtained through the stacking analysis of the BzK sample in the COSMOS PACS maps in the stellar mass range 10^{10}–$10^{11.5}$ $M_\odot$. The SFR and stellar masses are obtained with a Salpeter IMF and are corrected to the KE12 calibration and a Kroupa IMF.

(iii) Heinis et al. (2014) provide the MS based on the stacking UV selected SFGs in bins of FUV luminosity in the HerMES maps (Oliver et al. 2012). The derived SFR is based on the combination of NUV and IR luminosities, with a Chabrier IMF. The MS is estimated between $10^{9.5}$ and $10^{11.3}$ $M_\odot$ at $z \sim 1.5$, $z \sim 3$, and $z \sim 4$. We include in our analysis the stacked points at the observed stellar masses

and redshift given in the paper after correcting them to the KE12 calibration.

(iv) Whitaker et al. (2014) provide the MS based on stacking of UVJ selected SFGs on *Spitzer* MIPS 24 μm data in the 3D-HST CANDELS fields. The derived SFR is based on the combination of NUV and IR luminosities, with a Chabrier IMF. The considered redshift range is 0.5-2.5 and the MS is estimated in the following stellar mass and redshift ranges: $10^{8.5}$–$10^{11.2}$ $M_\odot$ at $0.5 < z < 1$, 10^9–$10^{11.3}$ $M_\odot$ at $1 < z < 1.5$, $10^{9.2}$–$10^{11.5}$ $M_\odot$ at $1.5 < z < 2$, and $10^{9.3}$–$10^{11.5}$ $M_\odot$ at $2 < z < 2.5$. We include in our analysis the stacked points at the observed stellar masses and redshift given in the paper, recalibrated to KE12.

(v) Chang et al. (2015) provide SFR based on SED fitting results of MAGHPHYS code, from *GALEX* to WISE data in the local Universe. As already reported by Popesso et al. (2019b), the catalog provides SFRs underestimated with respect to the $H\alpha$-based and IR-based SFR estimates. Salim et al. (2016) discuss that this is likely due to the fact that the MAGHPHYS SED fitting results are mostly driven by the higher SNR 12 μm WISE data point than the low SNR 22 μm WISE data, leading to artificially low SFR. Chang et al. (2015) report that their MS at $z < 0.1$ is systematically lower (0.15 dex) with respect to other $H\alpha$- and IR-based MS. We estimate the MS by using the best-fitting relation in stellar mass bins of 0.15 dex to take into account stellar mass uncertainties and after correcting the SFRs for the 0.15 dex offset, which brings them in agreement to the Ke12 calibration.
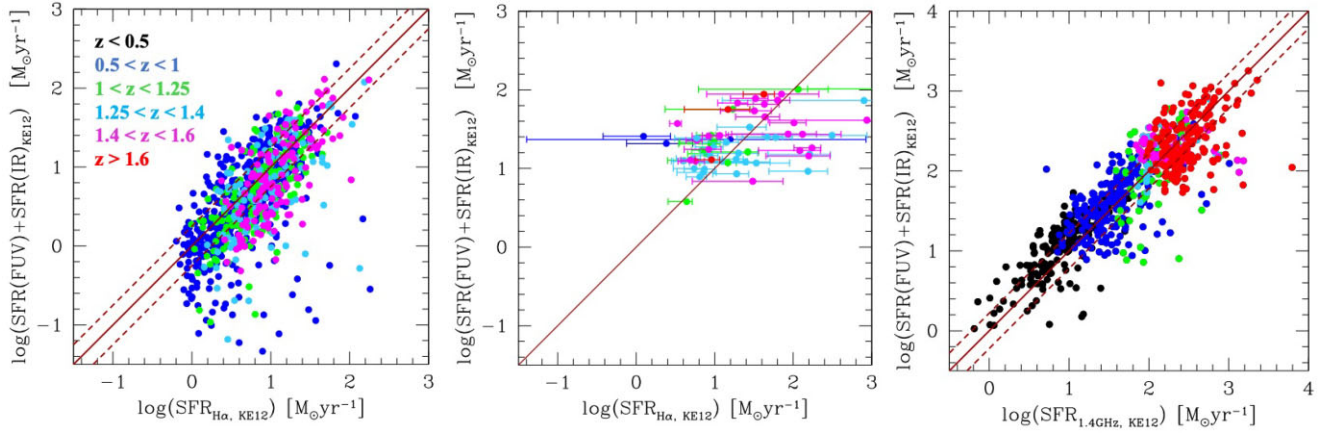

Figure A1. Comparison of the MS estimates of Leja et al. (2022) based on the SED fitting code *Prospector* (empty triangles) and those of Whitaker et al. (2014) based on the combination of $IR + UV$ in the same deep fields (filled points). The stellar masses of Leja et al. (2022) are corrected by 0.3 dex as indicated in Leja et al. (2019), for consistency with the stellar masses of Whitaker et al. (2014), based on FAST. No correction is applied to the SFR estimates. The MS are colour coded as a function of the Universe age as indicated in the figure and are displaced by 0.4 dex from one another for clarity. The solid line indicates the best fit of the MS as a function of stellar mass and time based on equation (14), while the dashed lines indicate the best-fitting 1σ error.

Figure A2. *Left-hand panel*: Comparison of the *UV + IR*-based SFRs versus the SFRs based on H α emission. The H α emission is corrected for dust attenuation on the basis of an average value of the Balmer decrement, calibrated as a function of the galaxy H α luminosity and stellar mass in Domínguez et al. (2013). All SFR estimates are obtained according to the KE12 calibration. The points are colour-coded as a function of the galaxy redshift. The red solid line shows the one-to-one relation and the dashed lines show the dispersion around the relation. *Central panel*: Same as in the left-hand panel for a subsample of the 3D-HST galaxies where both H α and H β exhibit SNR >5 and the Balmer decrement is estimated individually for each system. *Right-hand panel*: Comparison of the *UV + IR*-based SFRs versus the SFRs based on the radio emission at 1.4 GHz. The colour code is the same as in the left-hand panel. The red solid line indicates the one-to-one relation and the dashed lines show the dispersion around the relation.

(vi) Lee et al. (2015) use a ladder of SFR indicators to study the MS in the COSMOS field at $0.3 < z < 1.3$. Due to the flux limit of the *Spitzer* MIPS 24 μm and *Herschel* PACS and SPIRE catalogs, the combination of NUV and IR luminosities provides the SFR only for highly star-forming objects (starburst and massive SFGs). For less active or less dusty objects they use dust-corrected NUV luminosities (see fig. 3 of the cited paper for the contribution of any SFR indicator as a function of stellar mass). SFGs are selected according to their $(M_{NUV}\text{-}M_R) - (M_R\text{-}M_J)$ colours with a Chabrier IMF. The MS is estimated as the median SFR value in equally populated bins of stellar mass in the following ranges: $10^{8.5}$–$10^{11.}$ M$_\odot$ at $z = 0.36$, 10^9–$10^{11.}$ M$_\odot$ at $z = 0.55$, 10^9–$10^{11.}$ M$_\odot$ at $z = 0.70$, and between $10^{9.3}$ and $10^{11.2}$ M$_\odot$ at $z = 0.85$, $z = 0.99$, and $z = 1.19$. We include in our analysis the median points at the observed stellar masses and redshift, after correcting the median into the mean value and to the KE12 calibration.

(vii) Ilbert et al. (2015) use the flux-limited sample of *Spitzer* MIPS 24 μm and *Herschel* PACS and SPIRE of the COSMOS field to study the distribution of galaxies in the SFR-stellar mass plane. Galaxies are further selected according to their $(M_{NUV}\text{-}M_R) - (M_R\text{-}M_J)$ colours. The SFR is estimated with a combination of NUV and IR luminosities. The MS is identified as the peak of the SFR distribution at several stellar mass bins at stellar masses above 10^{10} M$_\odot$ and in the redshift range $0.2 < z < 1.4$. This corresponds to the median SFR in a log-normal distribution. We include in our analysis the median points at the observed stellar masses and redshift given in the paper, after correcting the median to the mean and to the KE12 calibration.

(viii) Tasca et al. (2015) use the spectroscopically selected sample of the VIMOS Ultra Deep Survey (Le Fèvre et al. 2015). Galaxies are selected from a combination of photometric redshifts, as well as from colour selection criteria like LBG, combined with a flux limit $22.5 < i_{AB} < 25$. The SFR is derived via SED fitting techniques with *Le Phare*, assuming the BC03 SPS model, a Chabrier IMF, and a Calzetti et al. (2000) extinction law. The MS is provided in the following stellar mass and redshift ranges: $10^{7.5}$–10^{10} M$_\odot$ at $0 < z < 0.7$, $10^{8.5}$–10^{11} M$_\odot$ at $0.7 < z < 1.5$, 10^9–10^{11} M$_\odot$ at $1.5 < z < 2.5$ and

at $2.5 < z < 3.5$, $10^{9.3}$–$10^{10.6}$ M$_\odot$ at $3.5 < z < 4.5$, and 10^9–10^{11} M$_\odot$ at $z > 4.5$. We include in our analysis the mean points at the observed stellar masses and redshift, corrected to the KE12 calibration.

(ix) Salmon et al. (2015) study the evolution of the slope and scatter of the SFR-stellar mass relation for galaxies at $3.5 < z < 6.5$ using multiwavelength photometry in GOODS-N from the CANDELS and Spitzer Extended Deep Survey. They use an updated, Bayesian spectral-energy distribution fitting method that incorporates effects of nebular line emission, SFHs that are constant or rising with time, and different dust attenuation prescriptions (starburst and Small Magellanic Cloud). They use a modified version of Bruzual & Charlot (2003) stellar population synthesis models, and a Salpeter IMF.

(x) Renzini & Peng (2015) provide a fit of the local MS at $z < 0.085$ in the stellar mass range 10^9–$10^{10.5}$ M$_\odot$. The SFR is based on dust-corrected SDSS H α fluxed for systems classified as star forming in the BPT diagram and through the D4000 break for non-active systems or AGN hosts (Brinchmann et al. 2004). SFR and stellar masses are estimated with a Chabrier IMF. The MS is identified as the peak of the SFR distribution at fixed stellar mass in the SFR–stellar mass plane. This corresponds to the median SFR in a log-normal distribution. We estimate the MS by using the best-fitting relation in stellar mass bins of 0.15 dex to take into account stellar mass uncertainties. The MS estimates are included after correcting the median into the mean value and to the KE12 calibration.

(xi) Schreiber et al. (2015) perform a stacking analysis of UVJ selected galaxies in the deep *Herschel* PACS maps of the CANDELS fields. The mean IR luminosity derived for the stacks is combined with the mean FUV luminosity to derive the SFR. SFR and stellar masses are estimated with a Salpeter IMF. The MS is given in the following stellar mass and redshift ranges: 10^9–10^{11} M$_\odot$ at $z \sim 0.5$, $z \sim 1$, and $z \sim 1.6$ and $10^{9.8}$–10^{11} M$_\odot$ at $z \sim 2$ and $z \sim 3$. We include in our analysis the stacked points at the observed stellar masses and redshift, corrected to the KE12 calibration.

(xii) de los Reyes et al. (2015) use $\sim$300 H α-selected galaxies at $z \sim 0.8$ to study the MS. They use deep optical spectra obtained with the IMACS spectrograph at the *Magellan* telescope to measure

strong oxygen lines. They combine spectral information with rest-frame UV-to-optical imaging, which allows them to determine stellar masses and dust attenuation corrections, and H α narrow-band imaging, which provides a robust measurement of the instantaneous SFR. The SEDs are fit with a library of Bruzual & Charlot ([2003](#)) stellar population synthesis models. The model libraries are built with a wide range of SFHs and metallicities, as described in Salim et al. ([2007](#)), and updated da Cunha et al. ([2008](#)). Each model is attenuated according to the prescription of Charlot & Fall ([2000](#)). The dust attenuation in the SED fitting is mainly constrained by the UV slope.

Their sample spans stellar masses of $\sim 10^9 - 6 \times 10^{11}$ M$_\odot$. They find consistency with previous results about the MS at similar redshift. We do not implement any correction in stellar mass and SFR.

(xiii) Erfanianfar et al. ([2016](#)) provide the MS at $z < 1.1$. The SFR is based on the flux-limited sample of *Spitzer* MIPS 24 μm and *Herschel* PACS and SPIRE available in the ECDFS and COSMOS fields. The SFR is derived from the far-IR flux with a Chabrier IMF. The MS is retrieved via σ-clipping. It is estimated in the stellar mass range 10^9–$10^{11.5}$ at $0.15 < z < 0.5$ and $0.5 < z < 1.1$. We include in our analysis the stacked points at the observed stellar masses and redshift, corrected to the KE12 calibration.

(xiv) Tomczak et al. ([2016](#)) perform a stacking analysis of UVJ selected galaxies in the deep *Herschel* maps of the CANDELS fields. The mean IR luminosity derived for the stacks is combined with the mean NUV luminosity to derive the SFR. SFR and stellar masses are estimated with a Chabrier IMF. The MS is given in the following stellar mass and redshift ranges: $10^{8.5}$–$10^{11.2}$ M$_\odot$ at $z \sim 0.6$, $z \sim 0.9$, $z \sim 1.1$, and $z \sim 1.4$, $10^{8.7}$–$10^{11.2}$ M$_\odot$ at $z \sim 1.7$, 10^9–$10^{11.2}$ M$_\odot$ at $z \sim 2.25$ and $z \sim 2.75$, and $10^{9.5}$–$10^{11.5}$ M$_\odot$ at $z \sim 3.5$. We include in our analysis the stacked points at the observed stellar masses and redshift, corrected to the KE12 calibration.

(xv) Santini et al. ([2017](#)) derive the MS in the ultra-deep *Hubble Space Telescope* Frontier fields. The MS is derived via SED fitting with a Salpeter IMF, BC03 SPS model, and a Calzetti et al. ([2000](#)) extinction law. The MS location is derived via σ-clipping at the following stellar masses and redshift ranges: 10^8–$10^{10.6}$ M$_\odot$ at $1.3 < z < 2$, 10^8–10^{11} M$_\odot$ at $2 < z < 3$ and $3 < z < 4$, 10^8–10^{11} at $4 < z < 5$, and $10^{8.6}$–10^{11} M$_\odot$ at $5 < z < 6$. We estimate the MS by using the best-fitting relations in stellar mass bins of 0.25 dex to take into account stellar mass uncertainties. We correct the MS to the Kroupa IMF.

(xvi) Kurczynski et al. ([2016](#)) utilize photometry in the Hubble Ultradeep Field (HUDF12) and Ultraviolet Ultra Deep Field campaigns and CANDELS/GOODS-S to estimate the SFR via SED fitting. They use a Salpeter IMF, BC03 SPS model and a Calzetti et al. ([2000](#)) extinction law. The MS is identified via σ-clipping between 10^7 and 10^{11} M$_\odot$ in the following redshift ranges: 0.5–1.0, 1.0–1.5, 1.5–2.0, 2.0–2.5, and 2.5–3.0. We point out that beyond $z \sim 1.5$, the stellar mass range $10^{10.2}$–10^{11} is populated by less than 20–25 systems. Thus, we limit the use of the MS estimates to the 10^7–$10^{10.2}$ M$_\odot$ stellar mass range. We estimate the MS by using the best-fitting relations in stellar mass bins of 0.25 dex to take into account stellar mass uncertainties. We correct the MS to the Kroupa IMF.

(xvii) Pearson et al. ([2018](#)) use the SED modelling and fitting tool CIGALE to generate flux density priors in the Herschel SPIRE bands in the COSMOS field. These priors are fed into a deblending tool, called *XID* +, to extract flux densities from the SPIRE maps. As the last step, multiwavelength data are combined with the extracted SPIRE flux densities to constrain SEDs and provide stellar masses and SFRs. These are used to populate the SFR-M^* plane over the redshift range $0.2 < z < 6$. The SED fitting is performed with a Chabrier IMF, BC03 SPS model, and a Charlot & Fall ([2000](#)) extinction law. SFGs are selected through the UVJ selection, as in Whitaker et al. ([2014](#)). The MS is estimated in the following stellar mass and redshift ranges: between 10^9 and 10^{11} M$_\odot$ at $0.2 < z < 0.5$ and $0.5 < z < 0.8$, between $10^{9.5}$ and 10^{11} M$_\odot$ at $0.8 < z < 1.1$, $1.1 < z < 1.4$ and $1.4 < z < 1.8$, between 10^{10} and 10^{11} M$_\odot$ at $1.8 < z < 2.3$ and $2.3 < z < 2.9$, between $10^{10.5}$ and 10^{11} M$_\odot$ at $2.9 < z < 3.9$ and $3.9 < z < 4.9$, and above 10^{11} M$_\odot$ at $4.9 < z < 6$. We estimate the MS by using the best-fitting relations in stellar mass bins of 0.25 dex to take into account stellar mass uncertainties. We correct the MS to the Kroupa IMF.

(xviii) Belfiore et al. ([2018](#)) provide a measure of the local MS at $z < 0.05$ in the MaNGA sample (Bundy et al. [2015](#)). SFGs are selected according to the BPT diagram in any spaxel of the MaNGA map. The SFR is estimated through the integrated, extinction-corrected H α emission over the galaxy region, converted via Kennicutt ([1998](#)) law and with a Chabrier IMF. The MS is estimated as the mean SFR in bins of stellar masses in the range 10^9–$10^{11.5}$ M$_\odot$. We include in our analysis the mean points at the observed stellar masses and redshift and correct to the Kroupa IMF.

(xix) Davidzon et al. ([2018](#)) estimate the evolution of the specific SFR at a fixed mass of $10^{10.3}$ M$_\odot$. The SFR is derived by the evolution of the galaxy stellar mass function of active galaxies of Davidzon et al. ([2017](#)), under the assumption that the stellar mass growth of a galaxy is mainly driven by the integral of the SFR history, minus the quantity of stellar mass returned to the interstellar medium. The GSMF of active galaxies, and thus, the derived SFR evolution is based on the $(M_{NUV}\text{-}M_R) - (M_R\text{-}M_J)$ colours. The SED fitting technique used in this analysis is based on a Chabrier IMF. The MS is estimated at $z \sim 2.22$, 2.75, 2.25, 3.75, 5, 6, and 7. We correct the MS to the Kroupa IMF.

(xx) Lee et al. ([2018](#)) use the deep CANDELS observations in the GOODS North and South fields to revisit the correlations between stellar mass and SFR in galaxies at $1.2 < z < 4$. The quantities are estimated via SED fitting with a Chabrier IMF and a BC03 SPS model. SFGs are selected through the UVJ selection. The MS is estimated at stellar masses above 10^9 M$_\odot$ and in the following redshift bins: 1.2–1.5, 1.5–2, 2–2.8, and 2.8–4. We correct the MS to the Kroupa IMF.

(xxi) Iyer et al. ([2018](#)) use SFHs reconstructed via the Dense Basis method of Iyer & Gawiser ([2017](#)) at $0.5 < z < 6$ in the CANDELS GOODS-S field to study the nature and evolution of MS. They use the reconstructed SFHs as trajectories in SFR–stellar mass plane. This allows them to study galaxies at epochs earlier than observed by propagating them backward in time along these trajectories. To generate spectra corresponding to a galaxy with a given basis SFH, they use the Flexible Stellar Population Synthesis (FSPS) mode (Conroy, Gunn & White [2009](#); Conroy & Gunn [2010](#)). They use a Chabrier IMF, a Calzetti et al. ([2000](#)) attenuation law, and IGM absorption according to Madau et al. ([1996](#)) prescription. They study the MS at z = 1, 2, 3, 4, 5, and 6 using both direct fits to galaxies observed at those epochs and $SFR - M^*$ trajectories of galaxies observed at lower redshifts. In order to exclude quiescent galaxies, they impose a cutoff, excluding systems that are at a distance larger than 0.4 dex from the best-fitting MS correlation. We correct the MS to the Kroupa IMF.

(xxii) Popesso et al. ([2019b](#)) provide measures of the local MS at $z < 0.085$ and stellar masses $> 10^{10}$ M$_\odot$ based on several SFR indicators. These include SDSS dust-corrected H α emission (Brinchmann et al. [2004](#)), WISE 22 μm emission, dust-corrected UV emission (Salim et al. [2016](#)), and *Herschel* PACS and SPIRE emission at 100, 160, 250, 350, and 500 μm (Valiante et al. [2016](#)).

We include all of them in the analysis. The SFR are all based on a Chabrier IMF. The MS is estimated as the peak of the SFR distribution in several stellar mass bins in the SFR–stellar mass plane. We include in our analysis the mean points at the observed stellar masses and redshift. We correct the MS to the Kroupa IMF.

(xxiii) Popesso et al. (2019a) measure the evolution of the MS up to $z \sim 2.5$ in the stellar mass range 10^{10}–$10^{11.5}$ $M_\odot$. The SFR is derived by a combination of IR fluxes based on *Spitzer* MIPS 24 μm and *Herschel* PACS and SPIRE data and NUV emission available in the CANDELS and COSMOS fields. The SFR and stellar masses are based on a Chabrier IMF. The MS is estimated as the peak of the SFR distribution in several stellar mass and redshift bins at $0.3 < z < 0.5$, $0.5 < z < 0.8$, $0.8 < z < 1.2$, $1.2 < z < 1.6$, $1.6 < z < 2.2$, and $2.2 < z < 2.5$. This corresponds to the median SFR in a log-normal distribution. We correct the median into the mean.

(xxiv) Barro et al. (2019) use a WFC3 F160W (H-band) selected catalog in the CANDELS/GOODS-N field containing photometry from the UV to FIR, photometric redshifts, and stellar parameters derived from the analysis of the multiwavelength data. Stellar masses are estimated with FAST and are consistent with those of the 3D-HST survey (e.g. Whitaker et al. 2014). They use a ladder of SFR indicators. The SFR ladder consists of three steps that differ on the amount of SFR indicators that are available for each galaxy, namely, UV, mid-IR, and far-IR. The first step is to compute the SFR from the UV luminosity. In this respect, they estimated the total UV luminosity from SED fitting, and use the relation of Kennicutt (1998), applying a correction for dust attenuation. The UV attenuation is inferred directly from the best-fitting model to the overall SED, which assumes a Calzetti et al. (2000) attenuation law. For the second and third steps of the ladder, they obtain the total SFR by adding the contributions from the IR emission, either from mid-IR or from far-IR, when available, and the observed, unobscured UV component. They use the calibration of Kennicutt (1998) corrected to a Chabrier IMF. The authors state that stellar masses and SFRs are consistent with previous results. The MS is estimated in the redshift bins $0.5 < z < 1.0$, $1.0 < z < 1.8$, and $1.8 < z < 3.0$ over the stellar mass range $\sim 10^{8.6}$–$10^{11.3}$ $M_\odot$. It is measured as the running median of the SFGs identified with the UVJ criterion. We correct the median to the mean and to the Kroupa IMF.

(xxv) Leslie et al. (2020) measure the evolution of the MS in the COSMOS field by exploiting the VLA-COSMOS 3 GHz Large Project data set. The MS is estimated in the range $0.3 < z < 6$ in the stellar mass range 10^9–10^{11} $M_\odot$ with different completeness limits as a function of the redshift. The median SFR is estimated via stacking in radio data in several redshift and stellar mass bins. We consider here only the bins where the prior galaxy sample is complete in stellar mass, as indicated by the authors. The SFR is estimated with a Chabrier IMF. SFGs are selected as in Ilbert et al. (2013) through the NUVRJ colour-colour selection. We correct the median to the mean and to the Kroupa IMF.

(xxvi) Thorne et al. (2021) apply *ProSpect* (Robotham et al. 2020) in a parametric mode to multiwavelength photometry from the *Deep Extragalactic VIsible Legacy Survey* (DEVILS; Davies et al. 2018) in order to measure stellar and dust masses and SFR for galaxies in the COSMOS field in the $0 < z < 9$ redshift range. They use the new DEVILS *UV* to *FIR* photometry derived using *profound* (Robotham et al. 2018). *ProSpect* uses the Bruzual & Charlot (2003) stellar libraries and the Chabrier IMF to model the stellar components. To model dust attenuation in galaxies, *ProSpect* uses the Charlot & Fall (2000) model. *ProSpect* utilizes the Dale et al. (2014) templates to model the reradiation of photons absorbed by dust into the infrared. The authors point out that, with respect to other SED fitting code, e.g.

MAGPHYS (da Cunha et al. 2008), based on the same templates, the benefit of *ProSpect* lies in the fact that it is extremely flexible in how it processes SFHs and in the fact that it incorporates evolving galaxy metallicities. The author state also that such differences in the SED fitting approach lead to an increase of the stellar mass estimates by 0.2 dex. The comparison with the MAGPHYS stellar masses over the GAMA fields is shown in section 5 of Robotham et al. (2020). No discrepancies, instead, are reported with respect to the galaxy SFR. This is also shown in the comparison provided by Thorne et al. (2021) with the MS estimated by Leslie et al. (2020) based on radio emission. These are, in turn, in agreement with the $UV + IR$ based MS of Whitaker et al. (2014) and Popesso et al. (2019a) over the same area. Thus, we correct only the stellar masses by 0.2 dex to bring the MS estimates of Thorne et al. (2021) on the same calibration framework of the other publications. We correct also the median to the mean and to the Kroupa IMF.

(xxvii) Sherman et al. (2021) provide the MS of massive galaxies at stellar masses larger than 10^{11} $M_\odot$ at $1.5 < z < 3.0$ over an area of 17.5 deg^2. They use $EAZY - PY$ to perform SED fitting using 12 FSPS (Conroy et al. 2009; Conroy & Gunn 2010) templates in the non-negative linear combination. The $EAZY - PY$ FSPS templates are built with a Chabrier IMF, Kriek & Conroy (2013) dust law, solar metallicity, and SFHs, including bursty and slowly rising models. SFGs are identified by looking at the SFR distribution per stellar mass bin. The authors identify the minimum corresponding to the green valley galaxies to isolate the star-forming systems above this minimum SFR. The MS is estimated as the average SFR of the SFGs in four stellar mass and four redshift bins. We correct to the Kroupa IMF.

(xxviii) Leja et al. (2022) use the panchromatic SED-fitting code *Prospector* to measure the star-forming MS across $0.2 < z < 3.0$ using the COSMOS-2015 and 3D-HST UV-IR photometric catalogs. In particular, they measure the star-forming sequence using stellar population properties inferred by the Prospector-α model built in the Prospector SED-fitting code. The model has 14 free parameters, consisting of a seven-component non-parametric SFH using the 'continuity' prior that disfavours sharp changes in SFR(t) (Leja et al. 2019), a two-component dust attenuation model with a flexible dust attenuation curve (Noll et al. 2009), free gas-phase and stellar metallicity, and a mid-infrared AGN component with a free normalization and dust optical depth. Prospector includes dust emission powered by energy balance (da Cunha et al. 2008) with a SED and nebular emission self-consistently powered by the model stellar ionizing continuum (Byler et al. 2017). As shown in Leja et al. (2019), *Prospector*, or its incarnation in *Prospector-α*, provides larger stellar masses with respect to previous SED fitting code as CIGALE (Noll et al. 2009; Boquien et al. 2019; Yang et al. 2020), MAGPHYS (da Cunha et al. 2008), or FAST (Kriek et al. 2009), largely used in the other publications collected here (see Appendix B for a discussion on the cause of the discrepancy). The authors indicate that such stellar mass increase is on average 0.3 dex with respect to other codes. Leja et al. (2022) also indicate that the SFRs derived by *Prospector* are lower with respect to the SFRs based on the $IR + UV$ contribution, derived in Whitaker et al. (2014). However, as shown in Fig. A1, we find that after correcting the stellar masses by 0.3 dex to this quantity to the same calibration framework of all other publications, we do not observe any significant discrepancy with respect to the SFR estimated with different indicators, e.g., the $IR + UV$ indicator used in Whitaker et al. (2014). Indeed, the SFRs of Leja et al. (2022) are consistent with the MS best fit as a function of redshift and stellar mass, within the observed scatter of 0.08 dex. Thus, we do not apply any correction to the SFR if not to for the Kroupa IMF.

They use a flexible neural network known as a normalizing flow to identify the ridge line of the density distribution in the $log(SFR) - log(M^*)$ plane.

APPENDIX B: THE KE12 SFR CALIBRATION

KE12 offers a collection of different calibrations that leads to consistent SFRs based on different SFR indicators. In this work, we rely mostly on the calibration of Murphy et al. (2011) and Hao et al. (2011) for the *IR*, $UV + IR$, H α, and radio 1.4 GHz SFR indicators. Murphy et al. (2011), in particular, uses the free–free emission measured in the Ka-band (26–40 GHz) for 10 star-forming regions in the nearby galaxy NGC 6946, including its star-bursting nucleus, to compare a number of SFR diagnostics. These diagnostics include non-thermal radio (i.e. 1.4 GHz), total infrared (IR 8–1000 μm), and warm dust (i.e. 24 μm) emission, along with hybrid indicators that attempt to account for obscured and unobscured emission from star-forming regions including $UV + IR$ measurements. By construction, the SFR derived from the indicators calibrated in this way result to be consistent with each other. Nevertheless, the KE12 calibration is based on the data obtained at $z \sim 0$, and one might ask if this calibration still holds for high redshift galaxies.

There is a substantial disagreement between the observed SFRs based on the KE12 or similar calibrations, and the SFRs estimated in simulated galaxies, as provided in state-of-the-art hydro-dynamical simulations. The discrepancy is that the observed SFRs tend to be larger with respect to the predicted ones, in particular, at $0.5 < z < 3$. This questions the reliability of the observed SFRs (Hayward et al. 2014; Davies et al. 2016; Katsianis et al. 2016; Leja et al. 2019; Martis et al. 2019; Lower et al. 2020). Katsianis et al. (2020), for instance, argue that the SFR based on the combination of the UV emission and the *Spitzer* 24 μm or *Herschel* 70, 100, 160, 250, and 350 μm emissions are largely overestimated up to 1 dex with respect to the SFRs of simulated galaxies. This work compares the intrinsic SFRs of EAGLE simulated galaxies and those obtained through the KE12 or similar calibrations by analysing the synthetic SED generated for the same systems by Camps et al. (2018). Furthermore, Leja et al. (2022) argue that the galaxy SFHs reconstructed by the SED fitting code *Prospector* (Leja et al. 2017) lead to larger values of stellar masses and lower values of the galaxy SFRs at $0.5 < z < 3.0$, more in agreement with the SFRs predicted by simulations. This is due to the contribution of an extra component of stars older than 100 Myr, not accounted for in previous SED fitting codes. Indeed, a larger old stellar population would obviously increase a galaxy's stellar mass budget, and it might largely contribute to dust heating, thus, increasing its emission in the mid and FIR (see also Viaene et al. 2017; Leja et al. 2019; Nersesian et al. 2019). Leja et al. (2022) argue that neglecting such a contribution when calibrating the SFRs, derived from the *IR* or *IR + UV* SFR indicators, leads to a clear overestimation of the galaxy SFR as a function of redshift. For this reason, *Prospector*, which accounts for this contribution, provides lower SFR, bringing in agreement observed and predicted estimates at $0.5 < z < 3$. However, it is worth pointing out that Leja et al. (2022) shows that such contribution is particularly relevant for galaxies below the MS, as expected (Hayward et al. 2014; Nersesian et al. 2019). It is less significant in the MS region, where the larger old stellar population affects mostly the galaxy stellar mass rather than the SFR. This is the main reason why the MS estimated by Leja et al. (2022) are in agreement with previous estimates, based on KE12 calibration, once the stellar masses are corrected for this discrepancy, as shown in Fig. A1.

To further check if the KE12 calibration still holds at high redshift, we use the following approach. If there is an evolution in the fraction of IR emission due to star formation as a function of redshift, as proposed by Leja et al. (2022), the SFR derived through the *IR* or *IR + UV* emissions and other SFR indicators based on KE12, such as the non-thermal radio emission at 1.4 GHz and the nebular H α emission, should no longer be consistent at high redshift. Indeed, the contribution of a larger old stellar population might affect the dust heating and the galaxy IR emission, but it can hardly affect the radio and the H α emissions. The radio emission originating from SFGs is thought to be caused by Type II and Types Ib supernovae whose remnants are believed to accelerate most of the relativistic electrons in these galaxies. The same supernovae ionize the H II regions as well. Only stars more massive than ~ 8 M$_\odot$ produce Type II and Ib supernovae, and these have lifetimes of $\sim 3 \times 10^7$ yr, while the relativistic electrons probably live ~ 100 yr. Radio observations, therefore, probe very recent star formation in galaxies, and have the advantage that the contribution to radio emission of stellar populations older than 100 Myr is insignificant. Similarly, the H α emission in SFGs is dominated by nebular emission in the H II regions, ionized by the most massive stars that have a short lifetime (Shivaei et al. 2015b; Katsianis et al. 2016). The strongest possible contamination is usually due to the AGN emission, rather than old stars. Those might contribute through the emission generated by stellar winds, but this is estimated to be negligible as seen in the local SDSS galaxies. Indeed, it is usually seen in evolved green valley galaxies not classified as star forming in the BPT diagram (Concas et al. 2017).

We, thus, compare the SFRs derived from the *IR + UV* and those derived from the radio 1.4 GHz and H α indicators, calibrated with KE12, in high redshift galaxies. This is to check if they still provide consistent results or if the *IR + UV* leads to an overestimation at $0.5 < z < 3$. For this exercise, we use UV emission estimated at 2800 Å and the IR emission based on *Spitzer* 24 μm data provided in the 3D-HST CANDELS field (Whitaker et al. 2014). Those are combined to obtain the SFRs, according to equation (3). The stellar masses are those derived through FAST, as described in Skelton et al. (2014).

We use the H α emission measured in the public 3D-HST data as described in Momcheva et al. (2016) in the same CANDELS fields. Unfortunately, the SNR of H β, [O III] and [N II] emission lines are too low to allow any classification in the BPT diagram (Kewley et al. 2013). Thus, we select all galaxies with an H α emission SNR higher than 5, although we are aware that there might be some contamination by AGN. The H α emission must be corrected for dust attenuation before being converted into SFR through equation (4). The low SNR of the H β emission prevents to estimate the Balmer decrement if not in a limited subsample of ~ 80 galaxies where the SNR is higher than 5 in both H α and H β emission lines. Thus, we correct the H α emission with an average estimate of the Balmer decrement, as calibrated in Domínguez et al. (2013), as a function of the uncorrected H α luminosity and stellar mass.

We use the VLA-COSMOS 1.4 and 3GHz catalog (Smolčić et al. 2017) to retrieve the galaxy radio emission in the COSMOS field. This catalog provides the galaxy radio luminosity at 1.4 and 3 GHz and isolates a clean subsample of SFGs with no contamination from AGN identified in the X-rays, infrared, or through the radio excess (Delvecchio et al. 2017). The 1.4 luminosity is converted to SFR through equation (6).

The comparison of the $UV + IR$ based SFRs and the SFRs derived from the radio 1.4 GHZ and H α emissions is shown in Fig. A2. The comparison with the H α-based SFRs is shown in the left-hand panel.

The points are colour-coded as a function of the redshift, as indicated in the figure. We do not observe a clear overestimation of the KE12 $UV + IR$ based SFRs with respect to the H α-based SFRs. All data points are distributed around the one-to-one relation with a scatter of 0.25 dex, without any clear bias as a function of redshift. On the contrary, we observe outliers with larger values of the H α-based SFRs with respect to the $UV + IR$-based SFRs, which we interpret as possible AGN contaminants. The central panel shows the same comparison for the subsample of 3D-HST galaxies where the H β emission-line SNR allows measuring the Balmer Decrement. Also, in this case, we do not observe a clear overestimation of the $UV + IR$-based SFRs. Nevertheless, we point out that the uncertainties are large also because we do not attempt any correction for absorption and possible line-blending issues, as done in Domínguez et al. (2013). The right-hand panel shows the comparison with the SFR based on the radio emission calibrated according to KE12. All data points are distributed around the one-to-one relation with a scatter of 0.22 dex and without any clear over- or underestimation as a function of redshift. Few residual sources exhibit a radio excess with respect to the $UV + IR$-based SFRs, which we interpret as radio AGN contaminants not included in the catalog of Delvecchio et al. (2017).

We, thus, conclude that there is no clear evidence for an overestimation of the $UV + IR$-based SFRs at $0.5 < z < 3$ and that the KE12 calibration allows obtaining consistent SFR estimates up to $z \sim 3$ from the different SFR diagnostics considered here.

This paper has been typeset from a TEX/LATEX file prepared by the author.