The Art of Measuring Physical Parameters in Galaxies: A Critical Assessment of Spectral Energy Distribution Fitting Techniques

Camilla Pacifici, Kartheik G. Iyer, Bahram Mobasher, Elisabete da Cunha, Viviana Acquaviva, Denis Burgarella, Gabriela Calistro Rivera, Adam C. Carnall, Yu-Yen Chang, Nima Chartab, et al.

The Astrophysical Journal, 944:141 (19pp), 2023 February 20 https://doi.org/10.3847/1538-4357/acacff

The Art of Measuring Physical Parameters in Galaxies: A Critical Assessment of Spectral Energy Distribution Fitting Techniques

Camilla Pacifici[1], Kartheik G. Iyer[2,3,57], Bahram Mobasher[4], Elisabete da Cunha[5,6,7], Viviana Acquaviva[8,9], Denis Burgarella[10], Gabriela Calistro Rivera[11], Adam C. Carnall[12], Yu-Yen Chang[13,14], Nima Chartab[4], Kevin C. Cooke[15], Ciaran Fairhurst[16], Jeyhan Kartaltepe[17], Joel Leja[18,19,20], Katarzyna Małek[10,21], Brett Salmon[1], Marianna Torelli[22], Alba Vidal-García[23,24], Médéric Boquien[25], Gabriel G. Brammer[26,27], Michael J. I. Brown[28], Peter L. Capak[27], Jacopo Chevallard[29], Chiara Circosta[30], Darren Croton[7,31], Iary Davidzon[27,32], Mark Dickinson[33], Kenneth J. Duncan[12], Sandra M. Faber[34], Harry C. Ferguson[1], Adriano Fontana[22], Yicheng Guo[35], Boris Haeussler[36], Shoubaneh Hemmati[37], Marziye Jafariyazani[4,38], Susan A. Kassin[1], Rebecca L. Larson[39], Bomee Lee[40], Kameswara Bharadwaj Mantha[41,42], Francesca Marchi[43], Hooshang Nayyeri[44], Jeffrey A. Newman[45,46], Viraj Pandya[2,34,47], Janine Pforr[48], Naveen Reddy[4], Ryan Sanders[49,*], Ekta Shah[49,50,58], Abtin Shahidi[4], Matthew L. Stevans[51], Dian Puspita Triani[7,31,52,53], Krystal D. Tyler[17], Brittany N. Vanderhoof[17], Alexander de la Vega[4,54], Weichen Wang[54,55], and Madalyn E. Weston[56]

[1] Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA; cpacifici@stsci.edu
[2] Dunlap Institute for Astronomy and Astrophysics, University of Toronto, 50 St George Street, Toronto, ON, M5S 3H4, Canada
[3] Columbia Astrophysics Laboratory, Columbia University, 550 West 120th Street, New York, NY 10027, USA
[4] Department of Physics and Astronomy, University of California, Riverside, 900 University Avenue, Riverside, CA 92521, USA
[5] International Centre for Radio Astronomy Research, University of Western Australia, 35 Stirling Hwy, Crawley, WA 6009, Australia
[6] Research School of Astronomy and Astrophysics, Australian National University, Canberra, ACT 2611, Australia
[7] ARC Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), Australia
[8] Physics Department, CUNY NYC College of Technology, Brooklyn, NY 11201, USA
[9] Center for Computational Astrophysics, Flatiron Institute, New York, NY 10010, USA
[10] Aix Marseille Univ, CNRS, CNES, LAM, Marseille, France
[11] European Southern Observatory, Karl-Schwarzschild-Str. 2, D-85748, Garching, Germany
[12] SUPA, Institute for Astronomy, University of Edinburgh, Royal Observatory, Edinburgh, EH9 3HJ, UK
[13] Department of Physics, National Chung Hsing University, Taichung, 40227, Taiwan
[14] Institute of Astronomy & Astrophysics, Academia Sinica, Taipei, 10617, Taiwan
[15] AAAS S&T Policy Fellow hosted at the National Science Foundation, 1200 New York Avenue, NW, Washington, DC 20005, USA
[16] Astronomy Centre, Department of Physics and Astronomy, University of Sussex, Brighton, BN1 9QH, UK
[17] School of Physics and Astronomy, Rochester Institute of Technology, Rochester, NY 14623, USA
[18] Department of Astronomy & Astrophysics, The Pennsylvania State University, University Park, PA 16802, USA
[19] Institute for Computational & Data Sciences, The Pennsylvania State University, University Park, PA 16802, USA
[20] Institute for Gravitation and the Cosmos, The Pennsylvania State University, University Park, PA 16802, USA
[21] National Centre for Nuclear Research, ul. Pasteura 7, 02-093 Warszawa, Poland
[22] INAF—Osservatorio Astronomico di Roma, via Frascati 33, I-00078 Monteporzio Catone, Italy
[23] Observatorio Astrónmico Nacional (IGN), C. Alfonso XII, 3, E-28014 Madrid, Spain
[24] Laboratoire de Physique de l'École Normale Supérieure, ENS, Université PSL, CNRS, Sorbonne Université, Université de Paris, F-75005, Paris, France
[25] Centro de Astronomía (CITEVA), Universidad de Antofagasta, Avenida Angamos 601, Antofagasta, Chile
[26] Niels Bohr Institute, University of Copenhagen, Jagtvej 128, Køobenhavn N, DK-2200, Denmark
[27] Cosmic Dawn Center (DAWN), Denmark
[28] School of Physics and Astronomy, Monash University, Clayton, VIC 3800, Australia
[29] CNRS, UMR7095, Institut d'Astrophysique de Paris, Sorbonne Université, F-75014, Paris, France
[30] Department of Physics & Astronomy, University College London, Gower Street, London, WC1E 6BT, UK
[31] Centre for Astrophysics & Supercomputing, Swinburne University of Technology, Hawthorn, VIC 3122, Australia
[32] Niels Bohr Institute, University of Copenhagen, Jagtvej 128, DK-2200 Copenhagen, Denmark
[33] Community Science and Data Center/NSF's NOIRLab, 950 N. Cherry Avenue, Tucson, AZ 85719, USA
[34] UCO/Lick Observatory, Department of Astronomy and Astrophysics, University of California, Santa Cruz, Santa Cruz, CA 95064, USA
[35] Department of Physics and Astronomy, University of Missouri, Columbia, MO 65211, USA
[36] European Southern Observatory, Alonso de Cordova 3107, Vitacura, Santiago, Chile
[37] IPAC, California Institute of Technology, 1200 E California Boulevard, Pasadena, CA 91125, USA
[38] IPAC/Caltech, 1200 E. California Boulevard, Pasadena, CA 91125, USA
[39] Department of Astronomy, The University of Texas at Austin, Austin, TX 78712, United States
[40] Korea Astronomy and Space Science Institute, Daedeokdae-ro 776, Yuseong-gu, Daejeon 34055, Republic of Korea
[41] Minnesota Institute for Astrophysics, University of Minnesota, 116 Church Street SE, Minneapolis, MN 55455, USA
[42] School of Physics and Astronomy, University of Minnesota, 116 Church Street SE, Minneapolis, MN 55455, USA
[43] INAF—Osservatorio Astronomico di Roma, Via di Frascati 33, I-00078, Monte Porzio Catone, Italy
[44] Department of Physics and Astronomy, University of California, Irvine, Irvine, CA 92697, USA
[45] Department of Physics and Astronomy, University of Pittsburgh, Pittsburgh, PA 15260, USA
[46] Pittsburgh Particle Physics, Astrophysics, and Cosmology Center (PITT PACC), University of Pittsburgh, Pittsburgh, PA 15260, USA
[47] Department of Astronomy, Columbia University, 550 West 120th Street, New York, NY 10027, USA
[48] Scientific Support Office, Directorate of Science and Robotic Exploration, European Space Research and Technology Centre (ESA/ESTEC), Keplerlaan 1, 2201 AZ Noordwijk, The Netherlands
[49] Department of Physics, University of California, Davis, One Shields Boulevard, Davis, CA 95616, USA
[50] School of Physics and Astronomy, Rochester Institute of Technology, 84 Lomb Memorial Drive, Rochester, NY 14623, USA
[51] Department of Astronomy, University of Texas at Austin, Austin, TX 78712, USA
[52] Research School of Astronomy and Astrophysics, Australian National University, Weston Creek, ACT 2611, Australia
[53] Center for Astrophysics Harvard & Smithshonian, 60 Garden Street, Cambridge, MA 02138, USA
[54] Department of Physics and Astronomy, Johns Hopkins University, Bloomberg Center, 3400 N. Charles Street, Baltimore, MD 21218, USA

[55] Dipartimento di Fisica 'G. Occhialini', Università degli Studi di Milano-Bicocca, Piazza della Scienza 3, I-20126 Milano, Italy
[56] Department of Physics and Astronomy, University of Missouri-Kansas City, Kansas City, MO 64110, USA
Received 2022 August 27; revised 2022 November 28; accepted 2022 November 30; published 2023 February 21

Abstract

The study of galaxy evolution hinges on our ability to interpret multiwavelength galaxy observations in terms of their physical properties. To do this, we rely on spectral energy distribution (SED) models, which allow us to infer physical parameters from spectrophotometric data. In recent years, thanks to wide and deep multiwave band galaxy surveys, the volume of high-quality data have significantly increased. Alongside the increased data, algorithms performing SED fitting have improved, including better modeling prescriptions, newer templates, and more extensive sampling in wavelength space. We present a comprehensive analysis of different SED-fitting codes including their methods and output with the aim of measuring the uncertainties caused by the modeling assumptions. We apply 14 of the most commonly used SED-fitting codes on samples from the CANDELS photometric catalogs at $z \sim 1$ and $z \sim 3$. We find agreement on the stellar mass, while we observe some discrepancies in the star formation rate (SFR) and dust-attenuation results. To explore the differences and biases among the codes, we explore the impact of the various modeling assumptions as they are set in the codes (e.g., star formation histories, nebular, dust and active galactic nucleus models) on the derived stellar masses, SFRs, and A_V values. We then assess the difference among the codes on the SFR–stellar mass relation and we measure the contribution to the uncertainties by the modeling choices (i.e., the modeling uncertainties) in stellar mass ($\sim$0.1 dex), SFR ($\sim$0.3 dex), and dust attenuation ($\sim$0.3 mag). Finally, we present some resources summarizing best practices in SED fitting.

Unified Astronomy Thesaurus concepts: Extragalactic astronomy (506); Spectral energy distribution (2129); Galaxies (573)

1. Introduction

Recent ground- and space-based galaxy surveys have generated rich data products that shed light on the physics of nearby and distant galaxies. Photometric surveys generally provide a large wavelength baseline at coarse resolution, e.g., COSMOS (Scoville et al. 2007; Weaver et al. 2022), CANDELS (Grogin et al. 2011; Koekemoer et al. 2011), KINGFISH (Kennicutt et al. 2011), UltraVISTA (McCracken et al. 2012; Muzzin et al. 2013), and ZFOURGE (Straatman et al. 2016). Spectroscopic surveys on the contrary provide generally a smaller wavelength range at higher resolution, e.g., the Sloan Digital Sky Survey (SDSS; York et al. 2000), GAMA (Driver et al. 2011), AGES (Kochanek et al. 2012), 3D-HST (Brammer et al. 2012; Momcheva et al. 2016), DEEP2 (Newman et al. 2013), MOSDEF (Kriek et al. 2015), LEGA-C (van der Wel et al. 2016), VANDELS (McLure et al. 2018; Pentericci et al. 2018), and VIPERS (Scodeggio et al. 2018). Some spectroscopic surveys are also accompanied by photometric catalogs using ancillary data and dedicated campaigns, e.g., SDSS, GAMA, and 3D-HST. While each of these surveys has been designed with specific science goals in mind, they share the same general set of tools necessary to translate observational measurements into physical parameters. Historically, single colors or luminosities were directly translated to physical parameters (e.g., Kennicutt 1989, 1998; Bell & de Jong 2001; Rujopakarn et al. 2013; Du & McGaugh 2020). When more data are available, the mapping from multiple colors to physical parameters is not as obvious. Full spectral energy distribution (SED) modeling, being physically motivated, is more widely applicable and does not rely on empirical calibrations that may only be valid for specific samples of galaxies.

SED fitting (see review by Conroy 2013) consists of comparing the observed SEDs of individual galaxies to existing templates based on various models to ultimately estimate the physical properties of such galaxies. This method has many advantages: it combines all information in a consistent way; it gives the flexibility to adapt the modeling to the data type; and it allows for customization of the ingredients that go into the creation of the model templates. SED fitting has also some limitations: it relies on assumptions that can be well established or still relatively unknown, and thus requires careful assessment of such assumptions (see, e.g., the introduction by Chevallard & Charlot 2016 on dust attenuation). To mitigate these limitations, we need to measure and account for the modeling uncertainties when deriving galaxy physical properties. Existing work in this area consists in varying modeling assumptions using individual codes and fitting simulated data. The conclusions of such studies are, however, hard to generalize because the results are normalized to a specific set of assumptions (i.e., the parameters assumed in the simulations). A systematic code comparison based on a real data set is instead not dependent on a specific set of assumptions and thus its results can quantify the modeling uncertainties without any specific normalization. This is the aim of this work.

SED fitting is used widely in the literature. In its early days, the targeted galaxies were at fairly low redshift and generally very luminous. Simple assumptions were enough to model the few data points acquired (e.g., models based on the evolutionary population synthesis technique; Spinrad & Taylor 1969; Bruzual A 1983; Guiderdoni & Rocca-Volmerange 1987; Bruzual A and Charlot 1993; Worthey 1994; Leitherer & Heckman 1995; Maraston 1998; Vazdekis 1999). Nowadays, data have become much more complex, sampling across redshift at high signal-to-noise ratios (S/Ns), for both massive and dwarf galaxies at both integrated and spatially

[57] NASA Hubble Fellow.
[58] LSSTC DSFP Fellow.

resolved scales. The spanned wavelength ranges have also increased with data sets that can cover the whole space from the far-ultraviolet (FUV) to the far-infrared (FIR) and beyond. These expansions in the data sets require improvements in the tools used to translate data points to physical parameters.

At the time of writing, there are many state-of-the-art SED-fitting tools. To give some examples among the most recently developed, we cite AGNfitter (Calistro Rivera et al. 2016), BEAGLE (Chevallard & Charlot 2016), Pipe3D (Sánchez et al. 2016), Prospector (Leja et al. 2017; Johnson et al. 2021), Dense Basis (Iyer & Gawiser 2017; Iyer et al. 2019), FIREFLY (Wilkinson et al. 2017), Lightning (Eufrasio 2017), Mr-Moose (Drouart & Falkendal 2018), BAGPIPES (Carnall et al. 2018), FortesFit (Rosario 2019), PEGASE.3 (Fioc & Rocca-Volmerange 2019), X-CIGALE (Yang et al. 2020), MCSED (Bowman et al. 2020), MIRKWOOD (Gilda et al. 2021), piXedfit (Abdurro'uf et al. 2021), ProSpect (Thorne et al. 2021), and Starduster (Qiu & Kang 2022). Machine-learning techniques are also advancing rapidly to provide redshift and physical parameter estimates for galaxies across a large redshift range, e.g., Davidzon et al. (2019) and Simet et al. (2021). All the above packages aim at covering the large parameter space spanned by multiwavelength observations of galaxies.

In order to cover a large parameter space, the SED-fitting tools need to include a certain number of ingredients with enough freedom to modify the modeling assumptions (see Figure 1 and Table 2 in Thorne et al. 2021). A first basic ingredient is the assumption of star formation and eventual metal enrichment histories (SFH and MEH). The most common ways to model the histories of galaxy growth are analytic and nonparametric functions (e.g., Carnall et al. 2019a; Leja et al. 2019). Lower et al. (2020) provide a comprehensive introduction on the various SFH models used in SED fitting (see also Lee et al. 2018). Along with the SFHs, accounting for metal formation histories is important, especially when interpreting photometric data sets spanning a large wavelength range or high-resolution spectroscopic data sets. For example, Bellstedt et al. (2020, 2021) show that SFHs derived from SED fits that account for closed-box metal enrichment can better match the cosmic star formation history (Madau & Dickinson 2014) compared to SFHs derived from SED fits with constant metallicity.

In addition to SFH and MEH components, simple stellar population (SSP) and nebular emission models are required to build the final model spectra. The papers by Baldwin et al. (2018) and Han & Han (2019) are examples of comparisons of different SSP models when fitting spectra of nearby galaxies. Additionally, the importance of including nebular emission in the spectra is highlighted by Pacifici et al. (2012, 2015), Smit et al. (2014), Salmon et al. (2015), among others. Dust attenuation and emission are ingredients necessary to interpret the SEDs of galaxies from UV-to-near-infrared (NIR) and to FIR, respectively (see review by Salim & Narayanan 2020). Multiple studies point to the need for flexible attenuation curves with varying parameters (e.g., Kriek & Conroy 2013; Buat et al. 2018; Salim et al. 2018; Wang et al. 2018; Barišić et al. 2020), while differences among IR models are reported by, for example, Hunt et al. (2019).

Finally, some SED-fitting tools also account for the contribution of a potential active galactic nucleus (AGN), to model the UV-to-IR emission and especially the mid-IR SED (e.g., Calistro Rivera et al. 2016, among others). The intergalactic medium (IGM) absorption on the bluest side of the SED is also very important especially for high-redshift sources (e.g., the Lyman-break or *dropout* technique to detect galaxies; Steidel et al. 1996).

With all the SED-fitting codes available, each with differences, it becomes difficult to compare results that use different codes, and also to quantify how much spread in physical parameters is introduced by the use of different SED-fitting approaches. All the studies mentioned above dive into the differences between specific models, generally focusing on a single ingredient.

Here, we provide a comprehensive comparative analysis of the output of 14 SED-fitting codes applied to the same observational photometric data set. We do not focus on any specific modeling ingredient and we assume that all modeling choices are reasonable and appropriate for the data set in question.[59] We thus can measure the contribution of the modeling assumptions to the uncertainty in the physical parameters. Knowing this contribution to the uncertainty is necessary when comparing results derived using different modeling approaches (see Goddard et al. 2017) and when comparing results to predictions from cosmological simulations. We stress that, in this work, we do not plan to identify any best code or best set of assumptions. The analysis presented here can be applied to any specific relation (e.g., the stellar mass function, the age–mass relation, the SFR density versus time). Here, we focus on the SFR–stellar mass relation (for star-forming galaxies, this is often referred to as the "star formation main sequence"; e.g., Brinchmann et al. 2004; Noeske et al. 2007; Whitaker et al. 2012), assess the different outputs by the different codes, and compute a "median" SFR–stellar mass relation, which accounts for the uncertainties in the modeling.

The paper is organized as follows. In Section 2 we provide some important working definitions in SED fitting; in Section 3 we present the three photometric data sets we use; in Section 4 we introduce the 14 SED-fitting codes used in this study; in Section 5 we compare the results of the codes and assess the effects of including/excluding IR measurements and AGN models; in Section 6 we assess the impact of the different modeling choices on the derived SFR–stellar mass relation and estimate modeling uncertainties for stellar mass, SFR, and dust attenuation; in Section 7 we list our lessons learned on SED fitting in general; finally, in Section 8 we provide our summary and conclusions.

Throughout this paper we use a cosmology with $\Omega_M = 0.3$, $\Lambda = 0.7$, and $h = H_0/100\ \mathrm{km\ s^{-1}\ Mpc^{-1}} = 0.7$. We present all magnitudes in the AB system.

2. The Spectral Energy Distribution Fitting Approach

The overall aim of the SED-fitting process is to constrain physical parameters, and the way this is achieved is by comparing the observed SEDs with a matrix of the template model SEDs spanning a wide range of physical parameters. How the best-matching models are selected varies from code to code. The code can return a single best-fitting model following a frequentist approach, or it can return a probability distribution following a Bayesian approach. Many codes nowadays adopt

[59] We note that some codes make similar assumptions and thus they can be biased in similar ways. In the same way, not all possible choices are explored in this work.

Table 1
Summary of the Possible Choices in the SED-fitting Tools

Code	Sampler	B/F	SFH	SSP	Neb. em.	Dust att.	Dust em.	AGN
AGNFitter[2]	MCMC	B	Exp. decl.	BC03	No	Single	Multiple	Yes
BAGPIPES[1,2,3]	Nested sam.	B	Flex.	Multiple	C17	Multiple	Single	No
BEAGLE[1,3]	Nested sam.	B	Flex. param.	BC03(16)	C13	Two comp.	No	No
CIGALE[1,2,3]	Grid	B	Flex.	Multiple	C13	Multiple	Multiple	Yes
Dense Basis[1,3]	Atlas	B	Nonparam.	FSPS	C	Multiple	Single	No
FITSED[1,3]	Grid	B	Exp. inc./decl.	BC03	C	Single	No	No
Interrogator[1,3]	MCMC	B	Flex. param.	Multiple	C17	Two comp.	No	No
LePhare[1,3]	Grid	F	Exp. decl.	BC03	No	Single	No	No
MAGPHYS[1,2,3]	Atlas	B	Flex. param.	BC03	No	Two comp.	Multiple	Yes
P12[1,3]	Atlas	B	Nonparam.	BC03(11)	C	Two comp.	No	No
Prospector[1,2,3]	Nested sam.	B	Nonparam.	FSPS	C13	Two comp.	Single	Yes
SED3FIT[1,2,3]	Grid	B	Exp. decl.	BC03	No	Two comp.	Single	Yes
SpeedyMC[1,3]	MCMC	B	Exp. inc./decl.	BC03	Empir.	Multiple	No	No
zPhot[1,3]	Grid	F	Exp. inc./decl.	BC03	Empir.	Multiple	No	No

Notes. The superscript in column [Code] identifies the sample from Section 3 to which the code has been applied. The [Sampler] identifies the algorithm used to explore the parameter space of the priors: by Markov Chain Monte Carlo ("MCMC") by nested sampling ("Nested sam.") on a "Grid," or on a precomputed atlas ("Atlas"). "B" signifies Bayesian and "F" signifies frequentist (i.e., a single best-fitting result). The [SFH] prior can include exponentially declining, increasing, or both functions ("Exp. decl.," "Exp. inc./decl."), multiple parametric functions ("Flex. param."), nonparametric functions only ("Nonparam"), or multiple parametric and nonparametric functions ("Flex."). The codes can select among different simple stellar population [SSP] models ("Multiple") or use a single SSP model (e.g., "BC03" or "FSPS"). The number in parenthesis for BC03 indicates the year the model was updated. The column [Neb. em.] stands for nebular emission and reports whether a code includes a nebular component calculated with the software Cloudy ("C" followed by the year of update), using empirical relations ("Empir.") or does not include a nebular component ("No"). The dust-attenuation model [Dust att.] can be in the form of a "Single" attenuation law, "Multiple" attenuation laws, or a two-component attenuation law ("Two comp."). The dust emission [Dust em.] modeling can be done using a "Single" or "Multiple" models, or can be skipped ("No"). An [AGN] component can be included or not. More information about the various models can be found in the text.

the latter. In Bayes' theorem, we must not only assess how well the model matches the data but also think critically about the reliability of the models themselves given all our prior knowledge (Kass & Raftery 1995). For example, from SED fitting we could find that a stellar population model of age 10 Gyr with a certain combination of dust attenuation and metallicity matches the observed data of a $z \sim 2$ galaxy with a high likelihood. However, we know that stars cannot be older than the age of the universe, which at $z \sim 2$ is ~3.2 Gyr. Given that knowledge, we ought to down-weight that model as a solution for these data. Therefore, *priors* on the models themselves are introduced to weight the likelihood. This is the basis of Bayesian statistics.

In technical terms, Bayes' theorem states the probability of an event (a model being a good representation of the data) given the prior knowledge of the conditions that might be related to that event (model). This is mathematically expressed as

$$\mathcal{L}(\Theta|D) = \frac{\mathcal{L}(D|\Theta) \times \mathcal{L}(\Theta)}{\mathcal{L}(D)}, \quad (1)$$

where $\mathcal{L}(\Theta|D)$ is the likelihood that the model Θ is true given data D. $\mathcal{L}(\Theta)$ contains the *prior information* on the model, the prior likelihood distribution. $\mathcal{L}(D|\Theta)$ is the likelihood that data D is expressed by model Θ, so-called the likelihood function, and $\mathcal{L}(D)$ is the unconditional marginal likelihood of the data.[60] The terms on the right are collectively referred to as the *posterior distribution*, and can be thought of as a likelihood function that has accounted for prior knowledge.

[60] $\mathcal{L}(D)$ is a constant, derived by integrating over all parameters Θ, and tells us the likelihood of all the models given the data. It is commonly called the "Bayesian evidence" and has many uses in Bayes factors and Bayesian inference criteria (Salmon et al. 2016).

The likelihood function is defined as the confidence for a given model Θ expressed in terms of some parameters, representing real data D. The likelihood for independent Gaussian errors is commonly expressed by the *chi-squared* metric, and for SED fitting is defined as

$$\chi^2 = \sum_{i=1}^{N} \frac{(f_\Theta(\lambda_i) - f(\lambda_i))^2}{\sigma_i^2}, \quad (2)$$

where each observed flux f is measured at some wavelength λ_i with an uncertainty σ_i, compared to the flux from the model f_Θ. The χ^2 is generally calculated using fluxes instead of magnitudes because the photometric errors are calculated on the flux and thus are symmetric in linear space and not in logarithmic space. For Gaussian error distributions, this is redefined as a likelihood via $\mathcal{L} = \exp(-\chi^2/2)$. If normalized such that it integrates to unity, this is called the probability density function (PDF). By this definition, at a low χ^2 there is a high probability that the model is a good representation of the data, given the optimized model parameters.

A "Bayesian approach" or "Bayesian framework" means to specifically design the model space with priors in mind, using prior knowledge to define which models to use or to weight them directly. The prior is generally defined as the distribution of values of the parameters that describe a specific physical model. For example, a flat distribution of $0 < \tau_V < 4$ is the prior for the parameter τ_V in a specific dust-attenuation physical model. However, we argue here that the choice of the specific physical model (or lack of such model) should also be considered a prior, even when the relative model parameters are fixed to a single value (e.g., the lack of emission-line modeling or a fixed initial mass function; see the various choices of physical models in Table 1, and the work by Curtis-Lake et al. 2021).

Within a Bayesian approach, the scientific use case determines whether a particular sampling technique is preferred over others while performing the SED fit. Sampling in this context refers to the way the SED-fitting code traverses the likelihood space $\mathcal{L}(D|\Theta)$, especially when the number of dimensions in Θ is large and cannot be covered easily with simple grid-search methods. Brute-force Bayesian methods (e.g., da Cunha et al. 2008; Pacifici et al. 2012; da Cunha et al. 2015; Iyer & Gawiser 2017; Abdurro'uf et al. 2021) that use a precomputed atlas of galaxies drawn from predefined prior distributions are preferred when scaling to large data sets from upcoming surveys using observatories such as the Nancy Grace Roman Space Telescope that will contain about 10^8 galaxies. These are particularly suited to large data sets since they trade space for time complexity, and are easily parallelizable. On the other hand, Markov Chain Monte Carlo (MCMC) and related techniques (Foreman-Mackey et al. 2013) are well suited for exploring the likelihood space of high S/N objects and accurately characterizing their uncertainties. Finally, nested-sampling-based techniques (Skilling 2004, 2006; Buchner 2016) are suited to problems where the likelihood space has pathological or highly multimodal features, or for model comparison where the Bayesian evidence needs to be estimated.

3. Data Sets

For our comparative analyses, we use three photometric data sets. These are chosen for two reasons: (1) the photometric bands cover a wavelength range from the UV to the NIR to sample the light from both old and young stars (the normalization of the NIR is necessary to infer the stellar mass, while the light in the UV is necessary to measure the current SFR of the galaxies) and also to have a large enough wavelength baseline to assess the effects of dust attenuation; (2) the additional IR photometry is measured at good-enough S/N ratios, for accurate measurements of the SFR. The three data sets are as follows:

1. A $z \sim 1$ sample with measured photometry spanning the rest-frame near-ultraviolet to the NIR (observed wavelength range between ∼0.35 and 4.5 μm).
2. A subset of sample 1 that includes targets observed with IR photometry (24–250 μm) with S/N in Multiband Imaging Photometer for Spitzer (MIPS)/24 μm larger than 3 and IR flags to avoid heavily blended targets.
3. A $z \sim 3$ sample with measured photometry spanning the rest-frame FUV to the NIR.

All samples are extracted from CANDELS (Grogin et al. 2011; Koekemoer et al. 2011) in the GOODS-South field (Guo et al. 2013). Photometric redshifts are provided by the CANDELS collaboration (Kodra et al. 2023) by combining photometric redshift estimates obtained using different codes. The full catalog of redshifts also includes publicly available spectroscopic redshifts and grism redshifts from the 3D-HST survey (Momcheva et al. 2016). A *best redshift* ($z_{\rm best}$) is selected for every galaxy using, in the following order of priority, an available spectroscopic redshift, or an available grism redshift, or the photometric redshift. In order to make the comparison of the derived absolute quantities (stellar mass and SFR) more meaningful, the tools assume $z_{\rm best}$ as the true redshift when fitting. We do not look for potential AGNs in any way, thus the sample may or may not include AGNs.

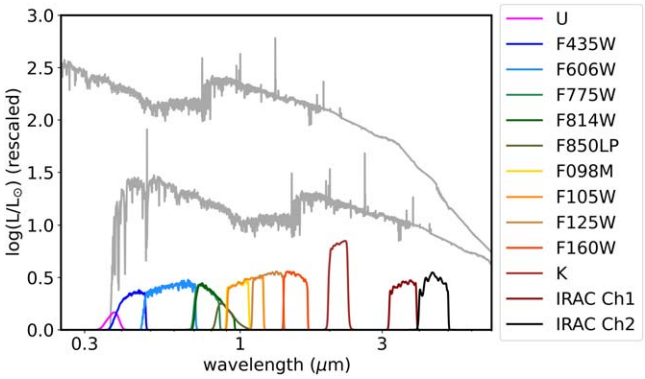

Figure 1. A model SED plotted at $z = 1$ (upper) and $z = 3$ (lower), along with the throughput curves of the filters in the photometric catalog (color code in the legend).

The $z \sim 1$ sample includes 339 galaxies selected at $1.0 < z < 1.1$, with good photometric flags (no contamination by bright stars and targets not on the edge of the detector), and $H < 24$ mag. We matched this catalog with IR measurements from the catalog by I. Hanae et al. 2022, (in preparation), following the methodology by Magnelli et al. (2013; see their Section 4) and flags by Barro et al. (2019). The fiducial IR sample includes 107 galaxies with S/N larger than 3 in MIPS 24 μm, accounting for flags to avoid heavily blended targets. Out of these 107 galaxies, 30 have photometry at 70 μm, 66 have photometry at 100 μm, 55 have photometry at 160 μm, and 28 have photometry at 250 μm.

The $z \sim 3$ sample includes 127 galaxies selected at $3.0 < z < 3.1$, with good photometric flags (the same as for the $z \sim 1$ sample), and $H < 26$. In all samples, objects with stellarity larger than 0.8 have been excluded.

Figure 1 shows a model SED plotted at $z = 1$ and $z = 3$ along with the response curves of the filters available in the catalogs. This shows the rest-frame wavelength range covered by the filters (excluding the IR) at the two redshifts. In both cases, UV, optical, and NIR are anchored, offering a good baseline to constrain both the old stellar populations (NIR) and the young and unobscured stellar populations (UV).

The catalogs can be found on GitHub.[61] By visual inspection of the photometry, we flagged and excluded 31 galaxies in the $z \sim 1$ sample and 34 galaxies in the $z \sim 3$ sample because of bad photometric measurements (e.g., photometry in a particular band is inconsistent with the photometry in adjacent bands). Flags files are published at the same URL.

4. The Spectral Energy Distribution Fitting Tools

We present here, in alphabetical order, the codes that take part in this study. Table 1 summarizes their main characteristics. All codes except one (AGNFitter is optimized to work only with IR data points, hence it is run only on data set 2) are run on data sets 1 and 3. The codes that can process IR data points (BAGPIPES, CIGALE, MAGPHYS, Prospector, and SED3FIT) are also run on data set 2. All fits were performed with the latest versions of each code available between 2018 and 2019.

[61] https://github.com/camipacifici/art_sedfitting

4.1. AGNFitter

AGNfitter (Calistro Rivera et al. 2016) is a Python package to fit the SEDs of AGNs and galaxies from the UV to the submillimeter with a fully Bayesian MCMC method. AGNFitter uses exponentially declining SFHs at constant metallicity and combines them with Bruzual & Charlot (2003; BC03) stellar population syntesis models. It does not include nebular emission. Dust attenuation is calculated with the Calzetti (2001) attenuation curve. The AGNfitter dust and AGN models consist of three physical emission components: an accretion disk, a torus of AGN-heated dust, and cold dust in star-forming regions.

4.2. BAGPIPES

Bayesian Analysis of Galaxies for Physical Inference and Parameter EStimation, or BAGPIPES (Carnall et al. 2018), is a Python package for modeling galaxy spectra and fitting spectroscopic and photometric observations in the FUV-to-FIR wavelength range. The code incorporates highly customizable models for emission from the stellar, nebular and dust components of galaxies, and an efficient Bayesian-fitting approach using nested sampling. BAGPIPES includes the option to fit flexible empirical models for systematic uncertainties in spectroscopic data, as described in Carnall et al. (2019b). The documentation can be found at https://bagpipes.readthedocs.io.

4.3. BEAGLE

BayEsian Analysis of gaLaxy sEds, or BEAGLE (Chevallard & Charlot 2016), enables the modeling of any combination of photometric and spectroscopic galaxy observable (e.g., full spectra, spectral indices, emission-line fluxes) with a flexible and fully self-consistent physical model from the UV to the NIR. It is used in combination with the MULTINEST algorithm to derive probabilistic (Bayesian) constraints on several galaxy physical parameters, and to create mock galaxy catalogs (e.g., Chevallard et al. 2019; Williams et al. 2018; Curtis-Lake et al. 2021). BEAGLE incorporates population synthesis models and models describing the emission of gas photoionized by young stars (lines and continuum). Several prescriptions to account for dust attenuation and IGM absorption are available. Galaxy star formation and chemical enrichment history are included in a flexible parametric way. Documentation and instructions can be found at https://www.iap.fr/beagle.

4.4. CIGALE

Code Investigating GALaxy Emission, or CIGALE (Burgarella et al. 2005; Boquien et al. 2019) is a Python-based code developed to interpret galaxy SEDs from the X-rays to the radio. SFHs can be parametric or ingested from a file. The code can use different stellar population models including ionized gas (continuum and emission lines) and a variety of dust models. An AGN component can be included and the newest version of the code can also accept X-ray data (Yang et al. 2020, 2022). The results are derived using a Bayesian approach on a grid of model SEDs and can take upper limits into account. In addition to fluxes, it can also fit intensive and extensive physical properties (e.g., a line equivalent width or the dust luminosity). CIGALE can be found at https://cigale.lam.fr.

4.5. Dense Basis Spectral Energy Distribution Fitting

The Dense Basis SED-fitting code (Iyer & Gawiser 2017; Iyer et al. 2019) is a Python-based code that creates and uses an atlas of galaxy SEDs from a physically motivated basis of flexible, nonparametric SFHs. The code uses flexible stellar population synthesis (FSPS; Conroy et al. 2009) to generate spectra adopting a variety of metallicities and dust-attenuation values. Gaussian-process regression is used to compute the SFHs and the lookback times at which a galaxy formed different fractions of its observed mass. These SFHs allow us to estimate quantities previously inaccessible through SED fitting, such as the number and duration of star formation episodes in a galaxy's past. The code can be found at https://github.com/kartheikiyer/dense_basis.

4.6. FITSED

FITSED (Papovich et al. 2001; Salmon et al. 2015) is an IDL-based code to infer the physical properties of galaxies from broadband photometry using a Bayesian framework. The code is set up to accept parametric SFHs, Bruzual & Charlot (2003) stellar population synthesis models, emission-line models, and variable dust-attenuation laws (Salmon et al. 2016). FITSED operates in two steps: it first creates a look-up-table grid of photometry from SED models and then calculates the likelihood between the data and every model in the grid.

4.7. Interrogator

Interrogator (Fairhurst, private communication; http://users.sussex.ac.uk/~sw376/Interrogator/) is a Python-based SED-fitting code exploiting the power of MCMC and Bayesian statistics to fit galaxy photometry and spectroscopy. The code takes a modular approach to modeling the stars (including Bruzual & Charlot 2003 and Eldridge & Stanway 2016), gas, and dust of a galaxy. In addition to data, users provide a choice of model for these various constituents, and can overwrite any of the default uninformative priors with arbitrary priors of their choice. The current version does not include models in the IR and so it interprets SEDs from the UV to the NIR.

4.8. LePhare

LePhare (Arnouts et al. 1999; Ilbert et al. 2006) consists of a set of Fortran commands to compute photometric redshifts and to perform SED fitting. It works with a grid of model templates created with PEGASE (Fioc & Rocca-Volmerange 2019) and Bruzual & Charlot (2003) population synthesis models, exponentially declining SFHs, and the Calzetti attenuation curve (Calzetti 2001). Since the physical parameters are estimated on a grid, uncertainties are derived by bootstrapping the data.

4.9. MAGPHYS

Multi-wavelength Analysis of galaxy Physical Properties, or MAGPHYS (da Cunha et al. 2008, 2015), is a Fortran-based code to model the emission by stellar populations and dust in galaxies and infer galaxy properties using a Bayesian approach. It uses a variety of exponentially declining and delayed exponentials SFHs along with superimposed random bursts, to account for stochasticity in star formation. The dust attenuation is modeled using the two-component model of Charlot & Fall (2000), and dust emission is self-consistently modeled using an

energy-balance technique. Recent updates include a radio component, an AGN component (see Chang et al. 2017a, 2017b), the inclusion of a UV bump in the dust-attenuation curve (Battisti et al. 2020), and a photometric redshift extension (Battisti et al. 2019). The code simultaneously fits the UV to radio emission by comparing each observed SED with all the models. MAGPHYS can be found at www.iap.fr/magphys.

4.10. P12

P12 (Pacifici et al. 2012) is a Fortran-based code to model and interpret any set of photometric and spectroscopic observations from the FUV to the NIR with a Bayesian algorithm. Model galaxies are generated assuming SFHs and MEHs from a semi-analytical model of galaxy formation. The model SEDs are created by combining consistently the emission by the stars from SSP models (the 2011 version of Bruzual & Charlot 2003) and the emission by the gas using a photoionization code (Charlot & Longhetti 2001). Dust attenuation is included with a two-component model, where multiple parameters are allowed to vary (similar to Charlot & Fall 2000).

4.11. Prospector

Prospector (Leja et al. 2017; Johnson et al. 2021) is a Python-based SED-fitting code to interpret photometric and/or spectroscopic data across the full FUV to FIR wavelength range. Sampling can be performed with simple minimization, MCMC, or nested-sampling techniques. The SED models are built using either parametric or nonparametric SFHs, the FSPS code (Conroy et al. 2009), and various dust-attenuation models. Prospector can also forward-model data calibration parameters such as spectrophotometric calibrations and wavelength solutions and incorporates uncertainties in these components in the final parameter uncertainties. Installation instructions and Jupyter notebooks are available at https://github.com/bd-j/prospector.

4.12. SED3FIT

SED3FIT (Berta et al. 2013) is a publicly available SED-fitting package that is a modification to the original MAGPHYS fitting routine. SED3FIT adds an AGN component to both the optical and FIR fitting steps. The AGN library in the original release of SED3FIT includes 10 AGN SED models which span viewing angles 0°–90°. The stellar, dust, and AGN components are moderated by energy balance at all times during the fitting process.

4.13. SpeedyMC

SpeedyMC (Acquaviva et al. 2012) is a MCMC algorithm for SED fitting, which builds upon the GalMC (Acquaviva et al. 2011) Bayesian framework for SED fitting, but is optimized for fitting large galaxy catalogs. The computational speed-up in SpeedyMC comes by precomputing galaxy spectra on a grid of locations exploring the entire parameter space, and then running the MCMC using multilinear interpolation between the precomputed spectra. It assumes exponentially increasing and declining SFHs, stellar population models (Bruzual & Charlot 2003), and various dust-attenuation curves.

4.14. zPhot

The zPhot code was born as a tool to measure galaxy photometric redshifts (Fontana et al. 2000) and has later evolved to also measure galaxy physical properties (see, e.g., Merlin et al. 2019). It uses a grid of template SEDs generated with exponentially declining, exponentially rising or constant SFHs, coupled with Bruzual & Charlot (2003) stellar models, nebular emission models, and a selection of dust-attenuation laws. The fitting range covers the UV to the NIR. Since the physical parameters are estimates on a grid, uncertainties are derived by bootstrapping the data.

5. Performance of the Codes on the Different Galaxy Samples

In this section, we present the physical parameters obtained from the various codes when run on the data sets described in Section 3. Each code has been run by their users with sufficient tuning in order to produce robust results for the data sets in question. Leaving the users free to choose their code's setup is representative of what normally happens in the scientific community and thus allows us to address that component of the uncertainty that is due to the modeling choices (i.e., the modeling uncertainty). In this exercise, we do not aim at identifying the "best" SED-fitting code, but we aim at assessing the biases and uncertainties caused by the modeling assumptions. The modeling uncertainties need to be taken into account when comparing results to other studies or to theoretical simulations. The outputs of the codes have been checked (e.g., evaluating the distribution of the χ^2 or the convergence of the fits) and considered acceptable.

5.1. Code Comparison on Individual Galaxies

We present here a comparison of the code outputs on three example individual galaxies in the parameter space of stellar mass, SFR (averaged over 100 Myr), and A_V (Figure 2). The best-estimate values are conventionally defined as the medians of the PDFs (see Section 7.3 for a discussion on other ways in which the best-estimate values can be defined). We select a star-forming galaxy (top row), a quiescent galaxy (middle row), and a galaxy where we find significant disagreement between the different codes, especially in their estimates of SFR and A_V (bottom row). All galaxies are taken from the $z \sim 1$ sample and the results are from fits to the UV-to-NIR photometry. All measurements are returned with uncertainties. For the Bayesian codes, this is done by calculating the full PDF and extracting either the half 16th-to-84th percentile interval, or the asymmetric 16th-to-50th and 50th-to-84th percentile intervals, depending on the code. For the codes that return only a single best-fitting model, uncertainties are calculated by fitting the data set multiple times after bootstrapping it about the photometric uncertainties and then measuring the dispersion in the derived parameters. The yellow shade marks an illustrative linear relation at constant specific SFR (sSFR; SFR divided by the stellar mass) of 1 Gyr^{-1} and 0.8 dex width similar to the observed SFR–stellar mass relation (e.g., Brinchmann et al. 2004; Whitaker et al. 2012; Speagle et al. 2014).

For the star-forming and quiescent galaxies, the results from the individual codes agree within their uncertainties (i.e., the observational uncertainties propagated to the physical parameters). Although the derived dust-attenuation values have

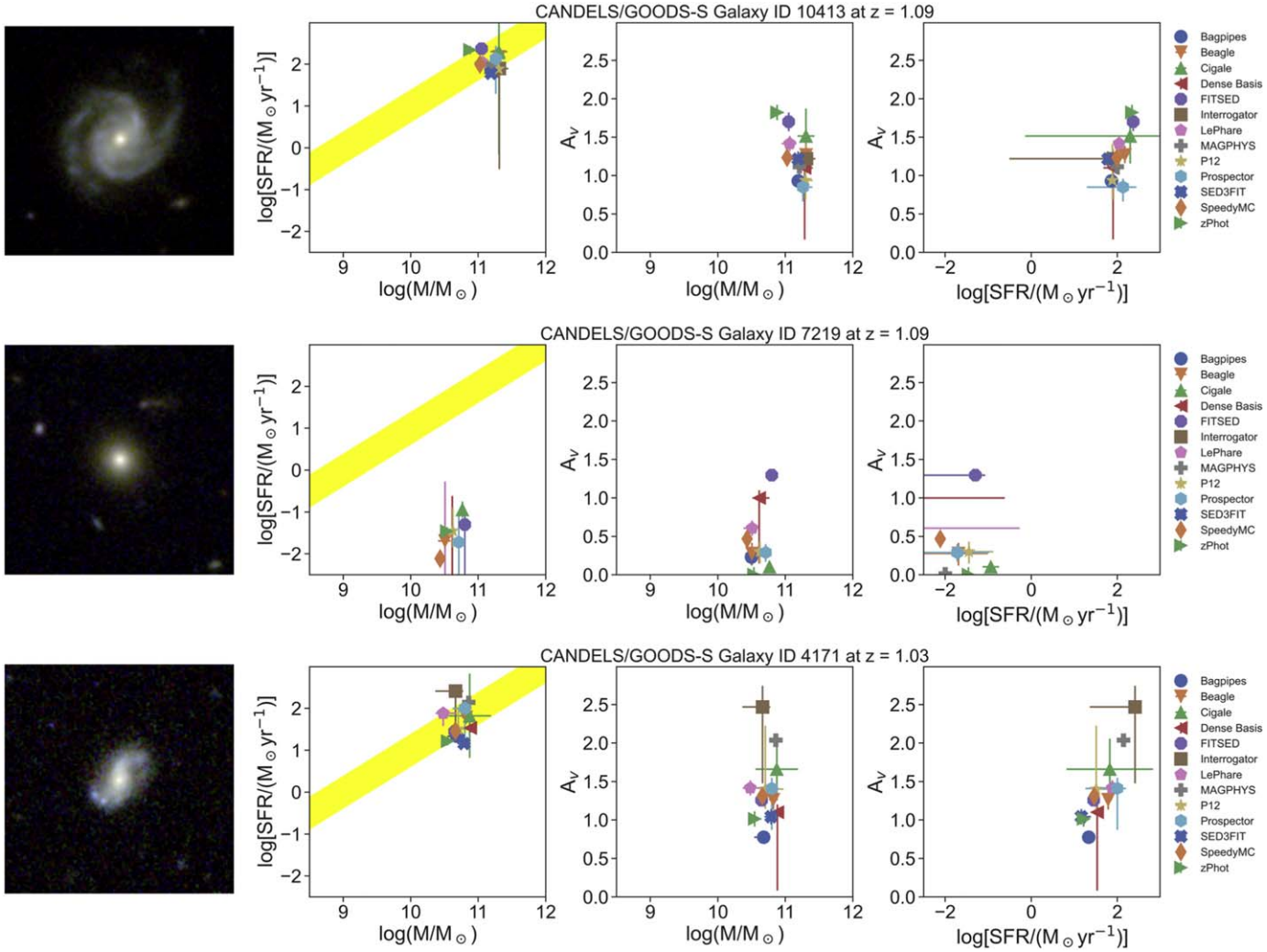

Figure 2. Individual measurements from the 13 codes for three galaxies in the $z \sim 1$ sample. We select a typical star-forming galaxy (top row), a typical quiescent galaxy (middle row), and a galaxy where we find significant disagreement between the different codes, especially in their estimates of SFR and A_V (bottom row). On the left, RGB images of the three galaxies at the same spacial scale. The three relations are SFR vs. stellar mass (the yellow shade marks an illustrative linear relation 0.8 dex wide; see text for details), A_V vs. stellar mass, and A_V vs. SFR.

small uncertainties in the majority of the codes (suggesting that the parameter is well constrained), they show the largest difference in the estimates because of the different assumed attenuation laws. The third galaxy shows some discrepancies and the differences between the estimates can be as high as one order of magnitude, while the uncertainties are generally smaller. The most evident correlation is between SFR and A_V (Figure 2, right panel), where the higher the inferred dust attenuation, the higher the SFR. This happens when there is a "degeneracy," i.e., when more than one parameter has the same effect on the observables and thus the SED is matched to different sets of parameters with the same likelihood. One way to break these degeneracies is to include new observables that depend on physical parameters in different ways. For example, adding IR observations can help break the degeneracy between SFR and A_V (see, e.g., da Cunha et al. 2008, 2015).

5.2. Distribution of Results for the z ∼ 1 Sample

We now explore the distributions of the best-estimate results (defined as the median of the PDF) for the three physical parameters of interest (stellar mass, SFR, and dust attenuation) from the various codes. The way the distributions of the physical parameters vary code by code given the same set of observations depends entirely on the assumptions in the SED-fitting tools and sampling of their parameter space. The stellar mass distributions, as shown on the left-hand side in Figure 3, are very similar among codes. This suggests that the different modeling assumptions do not have a large impact on this parameter (see also Santini et al. 2015) and thus it could be considered the most robust result that can be obtained with SED fitting. The differences are mainly in the normalizations of the distributions. Generally, systematic offsets in stellar mass (i.e., offsets that affect the sample on the entire range) can be caused by different IMF assumptions, differences in the stellar evolution models, or differences in the dust-attenuation models. For this exercise, the IMF was fixed to Chabrier (2003), thus the systematic differences are imputable to the different stellar evolution models (see, e.g., Baldwin et al. 2018) and to assumptions for the dust-attenuation law (see, e.g., Małek et al. 2018). We will not explore the details of the different stellar population models in this work. The SFH choice can also affect the stellar mass estimates, and generally affects low-mass and high-mass galaxies in different ways. For example, low-mass

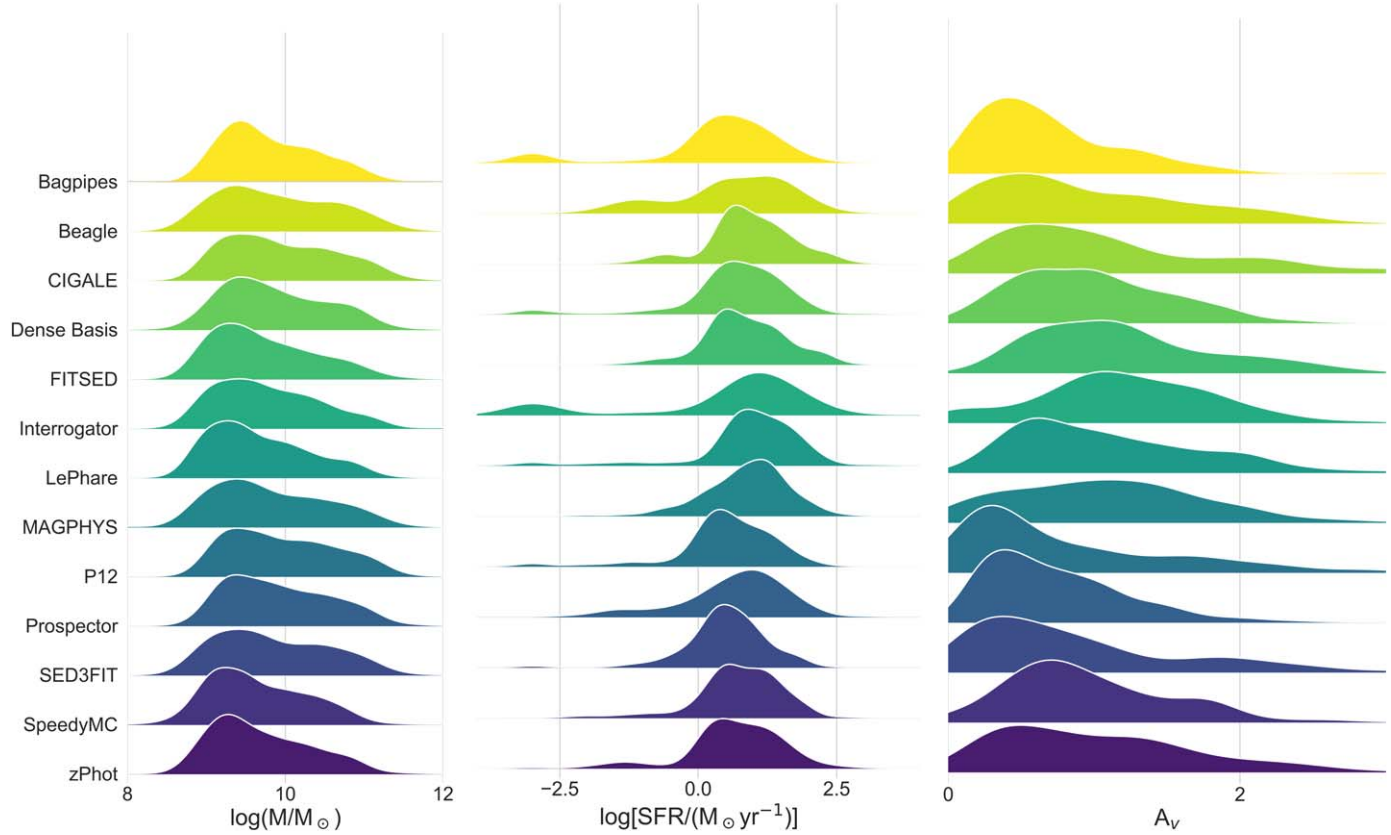

Figure 3. Each shade represents the distribution of the best-estimate results (defined as the median of the probability density function) of a specific parameter for all the galaxies in the $z \sim 1$ sample. Columns show stellar mass, SFR, and A_V. Each row represents the results from an individual code. The color code is purely graphical to ease the distinction from one row to the next. The differences in the distributions depend entirely on the assumptions in the SED-fitting tools and sampling of their parameter space.

or high-redshift galaxies are better represented by rising SFHs, while high-mass or low-redshift galaxies can be approximated by declining SFH functions (see, e.g., Pacifici et al. 2013). Thus, it is important to adopt appropriate SFHs for the sample under analysis.

The SFR distributions show more variety than the stellar mass ones and thus the constraint of this parameter is generally less robust than stellar mass. This is because the SFR estimates are affected by more modeling assumptions. For example, along with the assumed SFHs (and especially the sSFR prior) and stellar evolution models, the SFR estimates can be largely affected by the dust-attenuation models. Usually, large dust-attenuation measurements are coupled with large SFRs. Figure 3 shows that when the distribution of A_V is broad or peaks around $A_V = 1$ (e.g., for CIGALE, FITSED, Interrogator, LePhare, and MAGPHYS) the distribution of the SFR is also broad or peaks at large values. By contrast, the A_V distributions that peak at $A_V < 1$ are coupled with SFR distributions that peak at values close to $\log(\mathrm{SFR}/(M_\odot\ \mathrm{yr}^{-1}) = 0$. The location of the quiescent galaxies ($\log[\mathrm{SFR}/(M_\odot\ \mathrm{yr}^{-1})] < -1$) varies from code to code with BAGPIPES, Dense Basis, and Interrogator, showing a clear peak at $\log[\mathrm{SFR}/(M_\odot\ \mathrm{yr}^{-1})] \sim -3$, while the others show a tail instead.

5.3. Distribution of Results for the z ∼ 1 Sample with IR

Here, we discuss how crucial is the inclusion of IR data when measuring the physical parameters of galaxies, and outline some important caveats. The IR is especially important to constrain dust emission, and thus dust attenuation and SFRs (see, e.g., Buat et al. 2005; da Cunha et al. 2008; Davies et al. 2017; and the detailed introduction by Pappalardo et al. 2021). However, the large point-spread functions of some IR facilities (e.g., Spitzer and Herschel) can hamper our ability to properly separate individual sources and can easily cause biases in the photometric measurements at $z > 1$. Although precious work has been done to deblend IR data in order to produce reliable galaxy catalogs (e.g., Liu et al. 2018), and although similar problems are being resolved with the JWST, it is important to be aware of possible issues and make decisions to identify the best data set for the specific science questions.

Six codes among the 14 we have presented can process IR data along with the UV-to-NIR wavelength range. Using these codes, we can compare the outputs obtained including or not including the IR portion of the SED in the fit.

We present in Figure 4 the distribution of the results obtained running the different SED-fitting codes that model the IR emission on our sample number 2 (see Section 3). The sample includes 107 galaxies with a S/N in MIPS/24 μm larger than 3 and covers the mid-IR and IR with seven bands (5.8, 8.0, 24, 70, 100, 160, and 250 μm). We show the results when including and not including the IR emission in the fit. The distributions for the fits without IR are different than those presented in Figure 3 because sample 2 is a subset of sample 1. AGNFitter, CIGALE, MAGPHYS, and Prospector can include the contribution by AGN light. We note that AGNFitter is specifically designed to fit the AGN contribution in the IR and

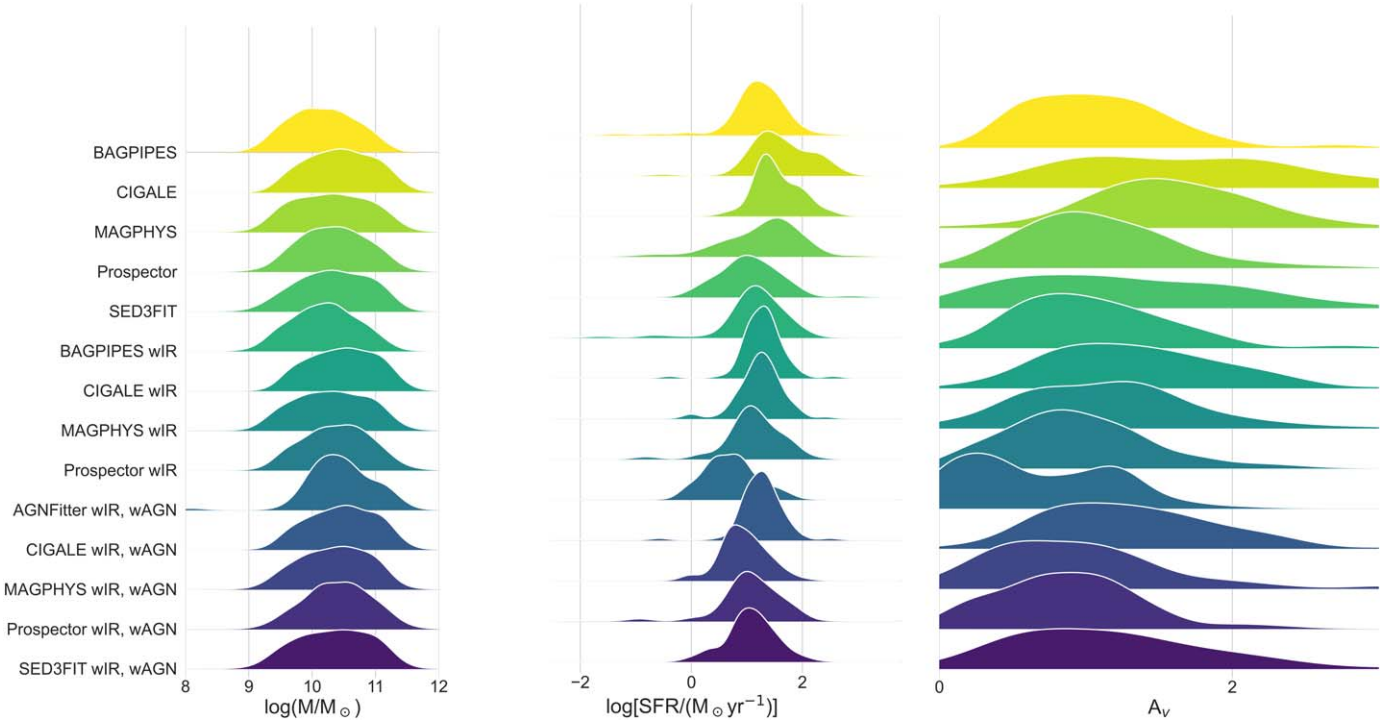

Figure 4. Distribution of the results obtained by all the codes (one code per row with different configurations) when fitting the $z \sim 1$ with IR measurements. The name of the code alone means the fits do not include IR measurements. "wIR" means the fits include IR measurements. "wAGN" means the code allows for an AGN component in the fit. Columns show stellar mass, SFR, and A_V. The color code is purely graphical to ease the distinction from one row to the next.

is not optimal to fit the UV-to-NIR alone, thus its default output is with IR and with AGNs.

We observe a trend in SFR and A_V: fits without the IR photometry seem to return larger values of these parameters (by about a factor of 2) compared to fits that include the IR photometry, suggesting that the correction for dust attenuation was overestimated when fitting the UV to NIR alone. This is the case for galaxies with moderate SFRs and moderate dust attenuation (i.e., not luminous or ultra-luminous infrared galaxies, LIRGs or ULIRGs) where the dust attenuation prior can easily bias the results toward large dust-attenuation values rather than toward small dust-attenuation values. Such effect can be mitigated by choosing an exponential prior for the dust attenuation (large weight at low A_V values and small weight at high A_V values) instead of a flat prior (same weight for all A_V values). The codes that return more similar SFR values with and without fitting the IR bands are those with more flexible SFHs (BAGPIPES and Prospector). In these two cases, the flexible SFHs allow the fit to find solutions with old stellar populations rather than adding more dust to match a *red* SED. Thus, more flexible SFHs also help mitigate the effects caused by a lack of IR coverage. The inclusion of AGN models further lowers the estimates, which is expected since part of the light is now interpreted as dust emission heated by AGNs rather than young stars.

Figure 5 shows, for each galaxy, the measurement of SFR obtained by each individual code including the IR photometry versus the median of the distribution of the SFR best estimates (i.e., medians of the PDFs) obtained with all codes, fitting only the UV-to-NIR wavelength range. We find a slight bias toward higher SFRs when the IR data are not included in the fit. This is due to an overestimation of the dust attenuation, which is otherwise better constrained when IR measurements are included. This bias is of the order of 0.2 dex, and in the case of CIGALE it is a strong function of the SFR and A_V. This bias can be reduced by allowing for more freedom in the dust-attenuation curve (e.g., allowing for variations in the slope of the attenuation law) and by including enough freedom in the SFHs to make the spectrum "red" by adding old stars instead of by increasing the dust attenuation. In this context, SFHs with a rising-and-falling shape and freedom in the time when the peak of star formation happens are preferred. The codes that allow these kinds of variations (BAGPIPES, Dense Basis, P12, and Prospector) return small values of A_V without fitting the IR wavelength range (see Figure 3).

Figure 6 shows the AGN fraction as a function of *H*-band magnitude for the five codes that can include an AGN component. The AGN fraction is defined as the luminosity due to the AGN over the total luminosity in the wavelength range between 8 μm and 1000 μm. Five galaxies are identified as AGNs according to their IRAC colors (Lacy et al. 2004, 2007; orange marks). The AGN fractions derived by the five different codes are fairly comparable and generally smaller than 25%. The few galaxies for which each code infers AGN fractions larger than the bulk of the sample also show AGN signatures in their IRAC colors. The codes that predict large AGN fractions measure slightly lower SFRs compared to the other codes. This demonstrates the degeneracy between AGNs and star formation contribution in the mid-IR. The number of IR bands fitted and their S/Ns also affect the measured AGN fractions and their accuracy. More in-depth discussions about the effects of choosing specific AGN templates or of the quality of the IR photometry are outside the scope of this work, but we can conclude that the inclusion of AGN components in SED fits is generally beneficial and enables more robust studies of galaxies with strong AGN features, as opposed to including them solely for the purpose of constraining AGN contamination fractions.

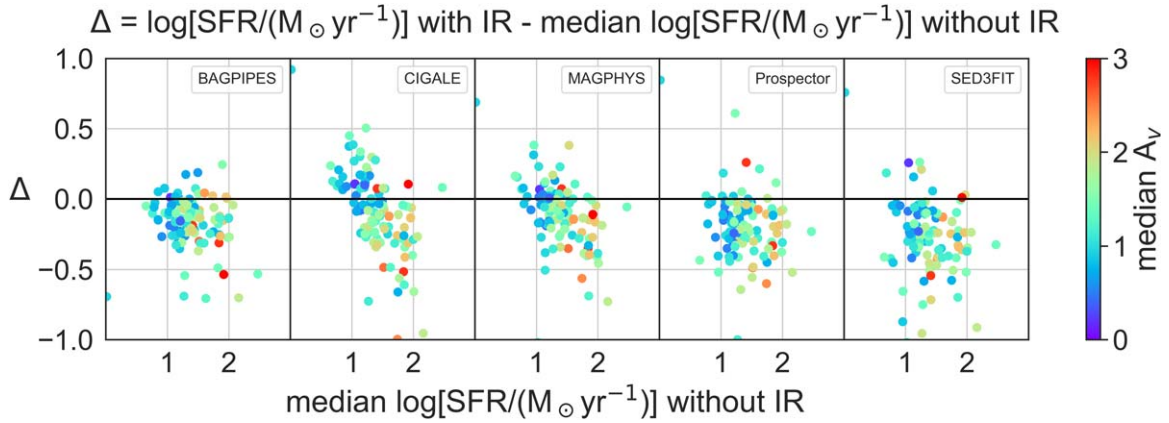

Figure 5. For the five codes that can fit IR measurements, we plot the difference between the SFR estimate including IR measurements and the median of the distribution of the SFR best estimates obtained with all codes without including IR measurements vs. the median of the distribution of the SFR best estimates obtained with all codes without including IR measurements. Points show individual galaxies in the $z \sim 1$ sample with IR measurements and are color-coded by the median of the distribution of the A_V best estimates obtained with all codes without including IR measurements. The black line marks zero, i.e., no difference between the two SFRs.

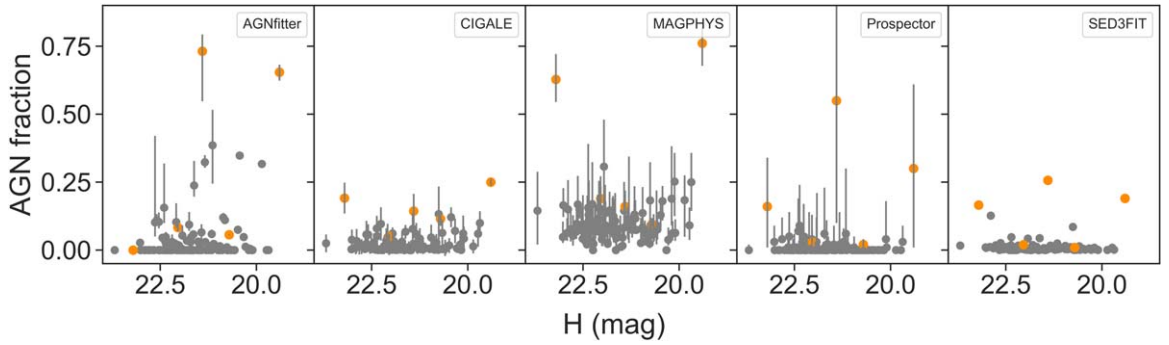

Figure 6. For the five codes that can include an AGN component in the fit, we plot the AGN fraction vs. the *H*-band magnitude. Points are individual galaxies in the $z \sim 1$ sample with IR measurements. Orange circles mark galaxies identified as AGNs according to their IRAC colors (Lacy et al. 2004, 2007). The AGN fraction is defined as the luminosity due to the AGN over the total luminosity in the wavelength range between 8 μm and 1000 μm.

5.4. Distribution of Results for the z ∼ 3 Sample

In Figure 7, we present the comparison of the different codes run on the $z \sim 3$ sample. The trends are similar to those we pointed out for the $z \sim 1$ sample: the mass distributions are all very similar, and the codes that return larger SFRs also return larger A_V values. Although at $z \sim 3$ the rest-frame wavelength range covered by the photometry only extends to $\sim$1.2 μm (while at $z \sim 1$ it extends to 2.2 μm), the universe is much younger. The emission by young stellar populations is less degenerate in age than the emission by old stellar populations, thus it is easier to disentangle the contributions by different stellar populations (i.e., the SFH) to the SED. This reduces the degeneracies among the parameters and with that the possible discrepancies among the different codes. We note that all codes point to a low level of dust attenuation, as is expected for *H*-band-detected galaxies at $z \sim 3$, because the *H* band traces the blue side of the optical, which becomes too faint for very dusty galaxies.

6. Impact on Science Results

6.1. Star Formation Rate–Stellar Mass Relation

The SFR–stellar mass (SFR–M_*) correlation shows the current growth rate of galaxies (Brinchmann et al. 2004; Noeske et al. 2007; Speagle et al. 2014; Whitaker et al. 2014; Leja et al. 2022). It is generally described by its intercept, slope, and scatter (intrinsic or observed) assuming a linear correlation. The intercept marks the level of star formation activity at a given mass and epoch and is observed to increase with redshift (Whitaker et al. 2012; Schreiber et al. 2015). The slope is used to measure the differential rates at which higher-mass galaxies form stars compared to lower-mass galaxies. It is generally linked to the different feedback mechanisms that regulate star formation across different populations of galaxies (see, e.g., Schreiber et al. 2015; Popesso et al. 2019a, 2019b; Davies et al. 2021). The scatter is associated with the burstiness of star formation on short timescales, because galaxies can scatter above and below the relation on timescales of 5–100 Myr. A large scatter can also be caused by a spread of galaxy ages on long timescales in the sense that galaxies can keep high or low SFR levels for their entire life and thus assemble their stellar mass faster or slower compared to the average population (Tacchella et al. 2016; Ciesla et al. 2017; Pandya et al. 2017; Boogaard et al. 2018; Matthee & Schaye 2018; Katsianis et al. 2019; Berti et al. 2021; Curtis-Lake et al. 2021). A small scatter would instead be consistent with all galaxies having very similar SFHs. These three measurements are often used to constrain growth and feedback mechanisms in simulations of galaxy evolution. It is thus important to understand the potential effects of different modeling and prior assumptions on the slope, intercept, and scatter of measurements of the SFR–M_* relation.

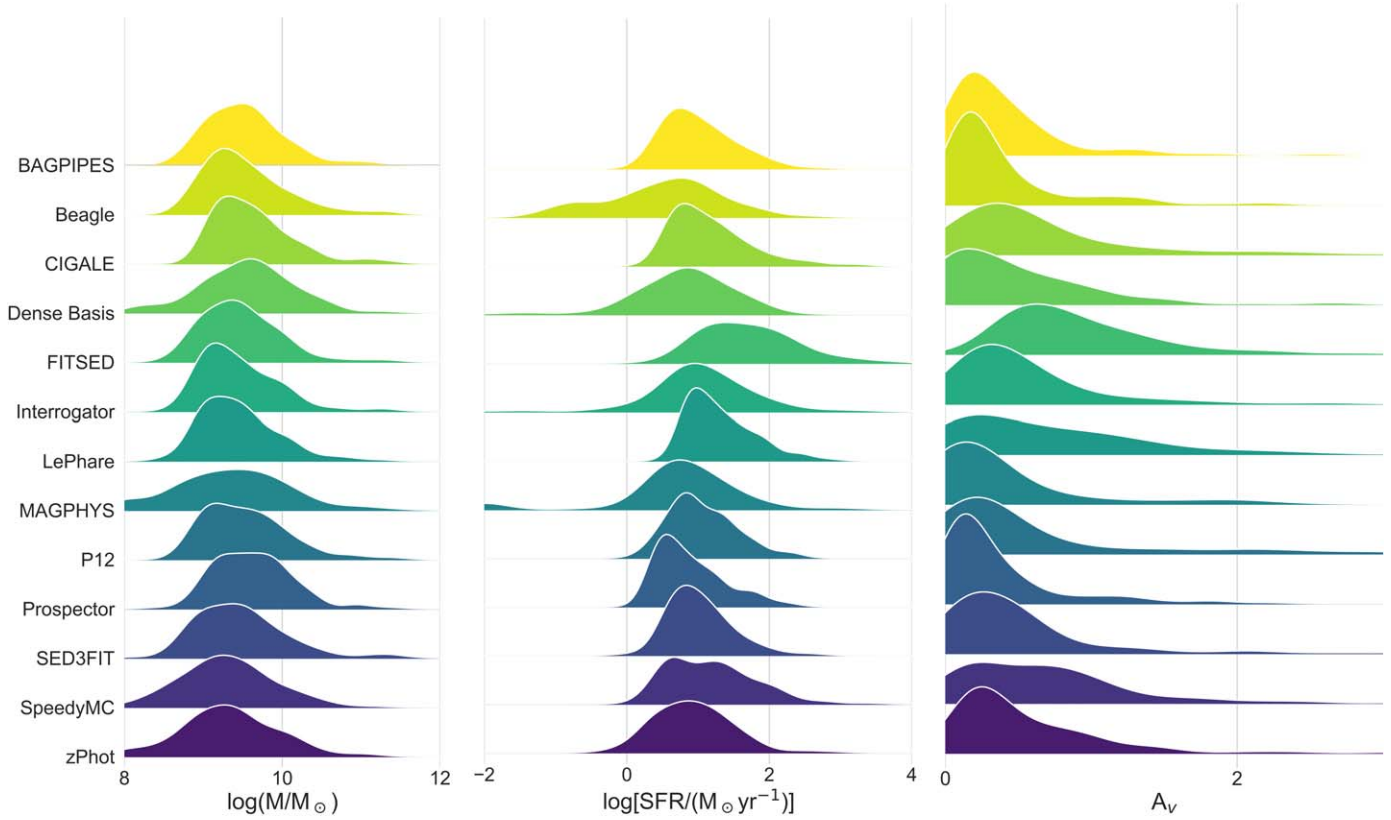

Figure 7. Distribution of the results obtained by all the codes (one code per row) when fitting the $z \sim 3$ sample. Columns show stellar mass, SFR, and A_V. The color code is purely graphical to ease the distinction from one row to the next.

Figure 8 shows the relations at $z \sim 1$ derived by the individual codes. The color code is set by the dust attenuation derived by each code. The black lines mark a sSFR of 1 Gyr^{-1} for comparison purposes. Globally, all codes measure a correlation between SFR and stellar mass. All codes measure more dust attenuation at higher masses and little to no dust attenuation for the quiescent galaxies that fall well below the relation.

Conversely, the normalization of the relation as well as the measured scatter can differ by up to 0.5 dex. Also, the locus of the galaxies that fall below the relation can vary considerably. We note that intermediate- and low-SFR galaxies generally show the highest uncertainties in this measurement (see Section 6.2) because the signal in the SED by SFRs lower than $10^{-1} M_{\odot}\,\mathrm{yr}^{-1}$ is not strong enough to be detected with high confidence.

The similarities and differences we observe can be attributed to the specific choices of modeling assumptions and priors in the individual codes. The choice of stellar population model, the prescription for the amount of mass returned to the interstellar medium in aging stellar populations, and the assumptions for the dust-attenuation law can change the stellar mass, shifting galaxies horizontally in Figure 8. The choice of SFH can affect the derived ages and thus the derived SFRs, causing, to first order, shifts up and down in the plots. Massive galaxies are generally older than low-mass galaxies, thus the shift can affect differently low- and high-mass galaxies. In addition, SFH priors can also change the extent to which the oldest stellar populations contribute to the stellar mass, by up to a factor of 0.2 dex (Leja et al. 2019).

The prior on the SFHs affects the prior on the sSFRs, which can cause artificial boundaries and affect the measurement of the scatter in the relation. We note here that the definition of SFR (essentially the timescale on which it is calculated) and the data used to measure it can affect systematically the derived SFR values and, thus, to make meaningful comparisons it is imperative to keep the definition and the data sets consistent. For example, Caplar & Tacchella (2019) shows there can be a nearly 0.3 dex difference between the Hα and UV measured scatter for different types of SFHs. In this paper, the data set is the same for all codes and the SFR is defined as the average SFR in the last 100 Myr, which is appropriate for photometric measurements. This is because the light in the rest-frame UV is produced by O and B stars whose lifetimes range between $\sim$10 and $\sim$300 Myr.

In this exercise, all codes are programmed with reasonable assumptions for the majority of the priors given this specific data set, and we encouraged the participants to keep track of the "goodness" of the fits and deliver results only if "good." The "goodness" of a fit can be assessed in many ways (i.e., by checking the χ^2 values, by comparing the posteriors to the priors, etc.) and this task was left in the hands of the participants. Comparing the goodness of fits among codes is beyond the scope of this exercise. With these assumptions, it is not possible for us to decide which SFR–M_* relation is the *true* one, but it is fair to say that different scientific conclusions could be derived on the modes of galaxy formation from any individual SFR–M_* relation. This is a first hint of how important it is to account for modeling uncertainties.

For every galaxy in the sample, there is a distribution of estimates from all the codes and therefore a median and a

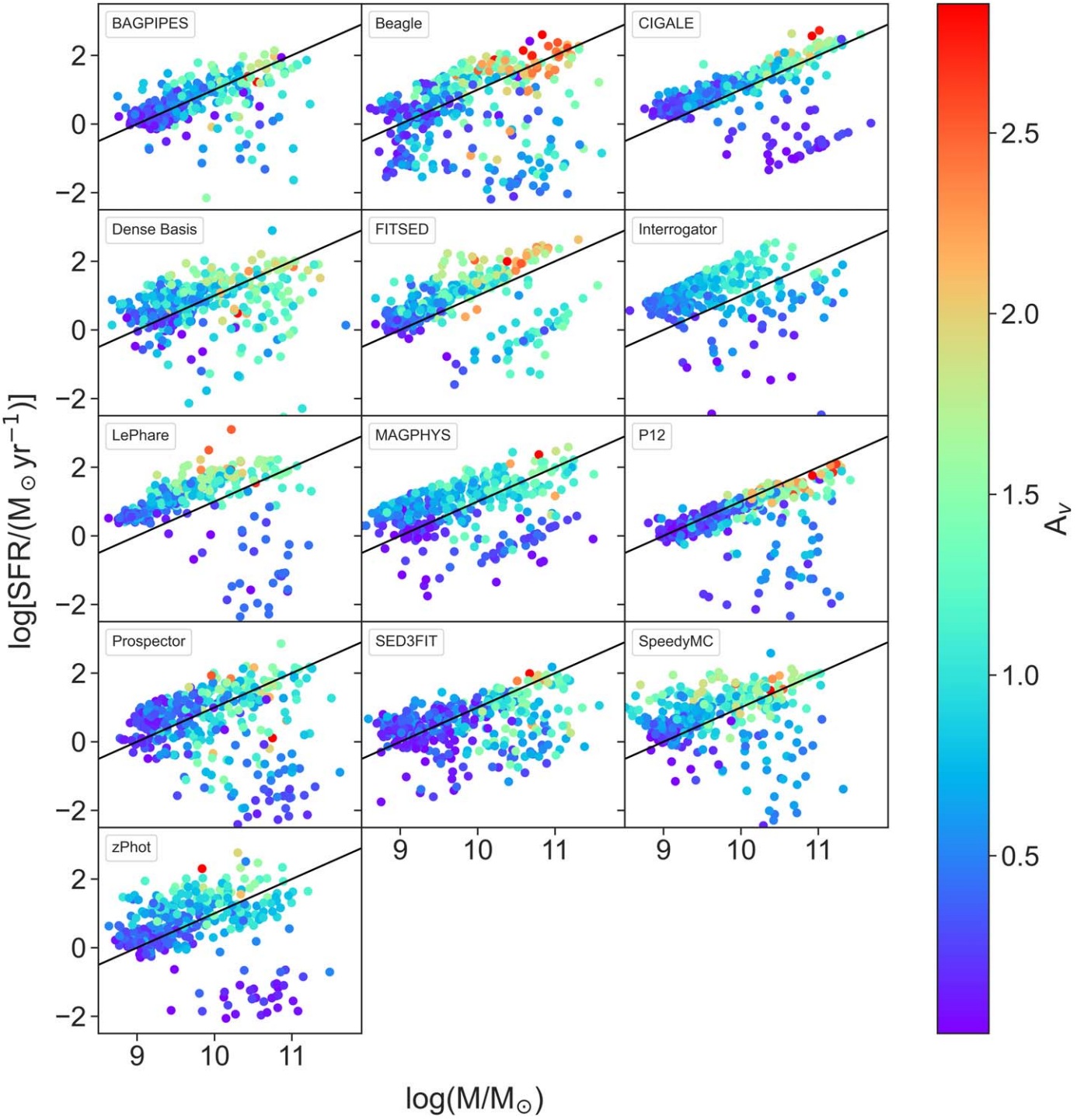

Figure 8. SFR–stellar mass relation for all the codes when fitting the $z \sim 1$ sample without IR measurements. Points are individual galaxies and the color code is A_V. The black line marks a constant specific SFR of 1 Gyr^{-1} in every plot and it is not a fit to the points.

dispersion of such distribution can be calculated. Figure 9 shows the SFR–M_* relation at $z \sim 1$ and $z \sim 3$ using the median values of stellar mass, SFR, and A_V (for the color code) for each galaxy (we will address the dispersion on these measurements in the next section). The median estimates of stellar mass and SFR take into account the difference in priors and modeling assumptions in deriving these quantities. As such, these estimates (to some extent) marginalize over these different assumptions and therefore provide more robust estimates of the mass and SFR.

6.2. Observational and Modeling Uncertainties of the Physical Parameters

With the available measurements, we can quantify more realistic uncertainties that include both *observational* (i.e., propagated from the uncertainties in the photometric measurements) and *modeling* (i.e., caused by the specific choices in the models) contributions. To calculate the total uncertainty for each galaxy, the observational uncertainty needs to be added in quadrature to the modeling uncertainty. Table 2 summarizes the

Table 2
Median Observational Uncertainties from All Codes for the Galaxies in the $z \sim 1$ and $z \sim 3$ Samples, Fitting the UV-to-NIR SEDs

	$z \sim 1$			$z \sim 3$		
	$\log(M_*/M_\odot)$ (dex)	$\log(\mathrm{SFR}/(M_\odot\ \mathrm{yr}^{-1})$ (dex)	A_V (mag)	$\log(M_*/M_\odot)$ (dex)	$\log(\mathrm{SFR}/(M_\odot\ \mathrm{yr}^{-1})$ (dex)	A_V (mag)
BAGPIPES	0.09	0.12	0.08	0.14	0.11	0.08
BEAGLE	0.07	0.15	0.14	0.08	0.33	0.09
Cigale	0.17	1.07	0.27	0.14	0.36	0.12
Dense Basis	0.10	0.11	0.15	0.10	0.18	0.15
FITSED	0.07	0.13	0.14	0.22	0.22	0.16
Interrogator	0.04	0.24	0.25	0.11	0.25	0.27
LePhare	0.21	0.42	NA	0.39	0.43	NA
MAGPHYS	0.02	0.05	0.08	0.00	0.00	0.00
P12	0.06	0.14	0.16	0.08	0.16	0.14
Prospector	0.09	0.38	0.13	0.17	0.18	0.15
SED3FIT	0.10	0.10	0.01	0.06	0.09	0.10
SpeedyMC	0.04	0.10	0.09	0.00	0.14	0.11
zPhot	0.10	0.10	NA	0.10	0.10	NA

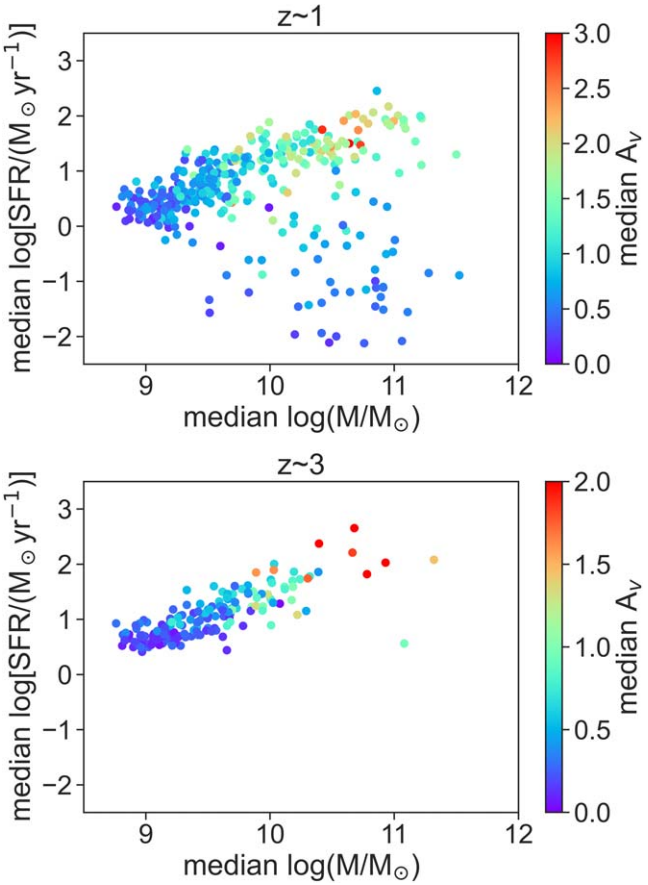

Figure 9. All plots show the SFR–M_* relation using the median stellar mass and median SFR from all codes for every galaxy in the $z \sim 1$ sample (top) and $z \sim 3$ sample (bottom). Points are color-coded by the median A_V from all the codes.

median observational uncertainties returned by the codes for the three parameters of interest, for the sample at $z \sim 1$ and $z \sim 3$, fitting the UV-to-NIR SEDs.

For the $z \sim 1$ sample, Figure 10 shows the variability of stellar masses, star formation rates, and A_V, along with a median value for all points and the locally weighted scatterplot smoothing, a nonparametric smoothing algorithm, to the data. To compute this we use the estimates of the given parameter for individual galaxies from each code and estimate the dispersion, here quantified as $(X_{84}–X_{16})/2$, where X is the distribution of either $\log(M_*/M_\odot)$ or $\log(\mathrm{SFR}/(M_\odot \mathrm{yr}^{-1})$ or A_V and the subscripts denote the 16th and 84th percentiles, respectively. The total uncertainty (including the observational and modeling components) should thus be the observational uncertainty each individual code returns added in quadrature to the modeling uncertainty (i.e., dispersion) measured here.

For the stellar mass, we find that the median of the dispersion is 0.12 dex at $z \sim 1$ and is largely independent of stellar mass itself, but can be mildly correlated with the SFR, such that the modeling uncertainty is larger for high SFRs. At $z \sim 3$, the median of the dispersion is 0.16 dex and is similarly not a function of stellar mass itself.

For the SFR, we find that the amount of variability depends on the inferred SFR value itself, and thus it depends on the flux in the rest-frame UV (these fits do not include IR measurements). Galaxies with low SFRs show faint UV luminosity with low S/Ns. This hamper our ability to measure accurate SFRs, in the sense that it is generally possible to infer that the SFR is low (or effectively consistent with no star formation), but it is hard to constrain the exact value. We measure the median of the dispersion to be on average 0.27 dex at $z \sim 1$ and 0.28 dex at $z \sim 3$, with low-SFR galaxies with dispersions as high as 1 dex.

The modeling uncertainty in A_V is a strong function of the value itself. Highly attenuated galaxies are, not surprisingly, the harder to interpret. At $z \sim 1$, the median of the dispersion is 0.27 mag and it is larger than 0.4 mag for $A_V > 2$. At $z \sim 3$, we find a similar correlation between the dispersion and the A_V value itself. The median of the dispersion is 0.22 mag with values as high as 1 mag for large A_V values.

7. Lessons Learned

In this section, we will provide some suggestions on how to approach SED fitting and understanding the power and limitations of this technique. We will list a series of good practices that should be obvious to experienced users, but might be of help for those who are just starting working with SED fitting. As we have seen, it is important to be aware of the tool characteristics and make a few appropriate decisions (e.g., type of tool, available ingredients, flexibility of the priors, etc).

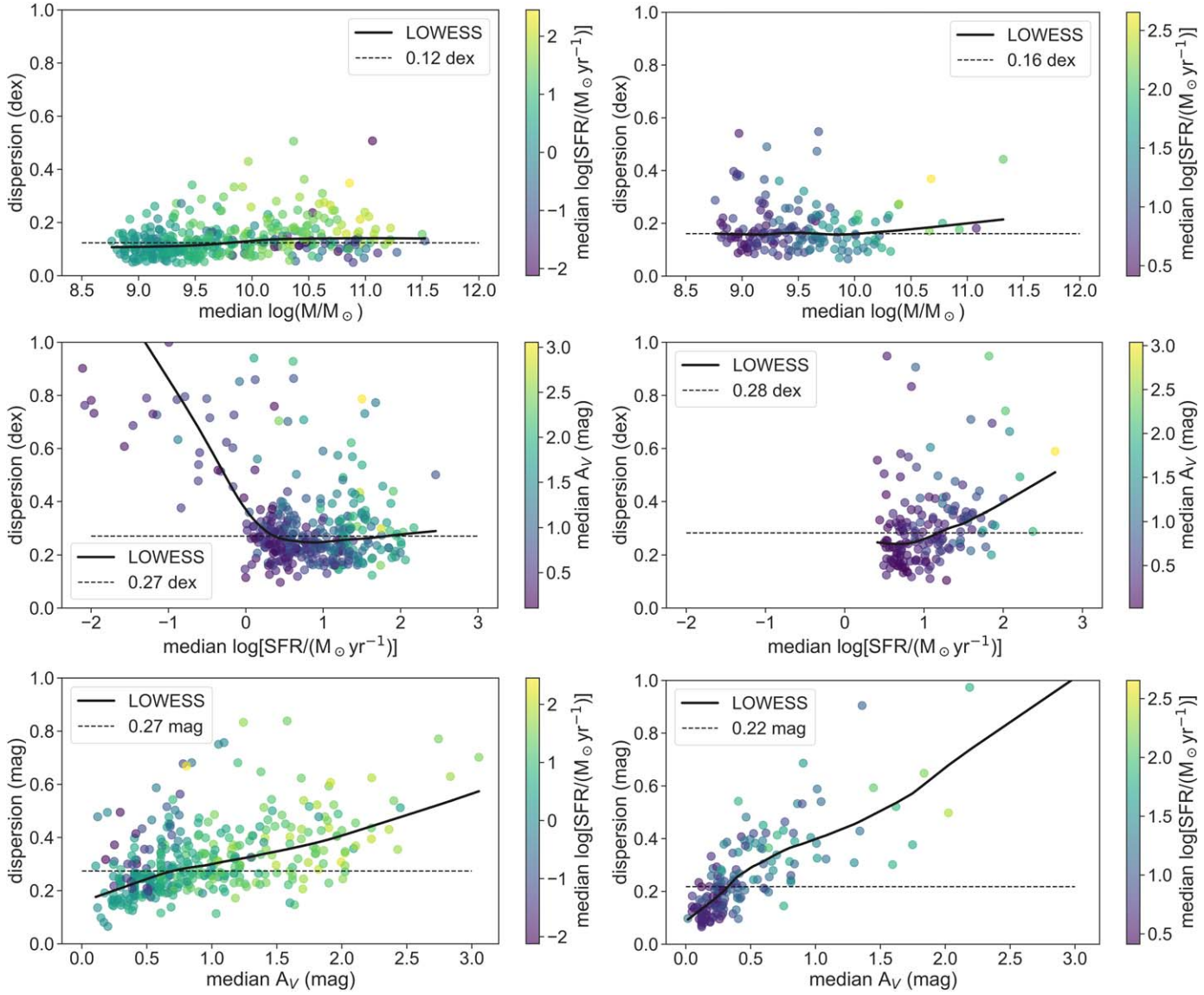

Figure 10. In all plots, each point represents the dispersion in the estimate of the physical parameter from all the codes vs. the median parameter from all the codes. The dispersion is defined as the half difference between the 16th and 84th percentile of the distribution of the parameter derived from all the codes. Stellar mass is plotted on top, SFR in the middle, and A_V at the bottom. The left side uses measurements from the $z \sim 1$ sample, while the right side uses measurements from the $z \sim 3$ sample. In each plot, the dotted line marks the median value for all the points, while the solid line marks the locally weighted scatterplot smoothing (LOWESS) to the data.

We will start this section with a flowchart to check whether the tool of choice is appropriate for the selected sample. We will then parse a list of "knobs" that can potentially be adjusted in a SED-fitting tool. Finally, we will talk about the outputs of the tools and the definition of "results."

7.1. Spectral Energy Distribution Fitting Flowchart

Running a SED-fitting tool requires a few steps and necessary customization, as seen in Figure 11. First of all, it is critical to choose the appropriate SED-fitting tool for the selected data set. There needs to be a match in wavelength coverage and supported spectral resolution. The second critical step is the customization of the priors. This is rarely a one-time activity, and is often an iterative step. We note that over-customizing the tool can make the comparison to other results more challenging, thus it should be done only when necessary. A common technique to check whether the selected priors are appropriate for the sample is the comparison between observed and model-predicted colors (see, e.g., Pacifici et al. 2015). It is always more appropriate to bring the models to the observed redshift than to modify the observations converting them to rest frame. We will go into more details in Section 7.2. The last step is the analysis of the results and this too can be used to assess the appropriateness of the priors. We will talk more about results in Section 7.3.

7.2. Setting Up the Appropriate Models

A set of ingredients is needed to generate the SED templates that are used to extract physical parameters (photometric redshifts, stellar mass, star formation rates, age, extinction, etc.) from the multiwave band data. The priors on these ingredients are of crucial importance.

SED-fitting tools can be generally divided into two groups: those where a library of models is created beforehand spanning

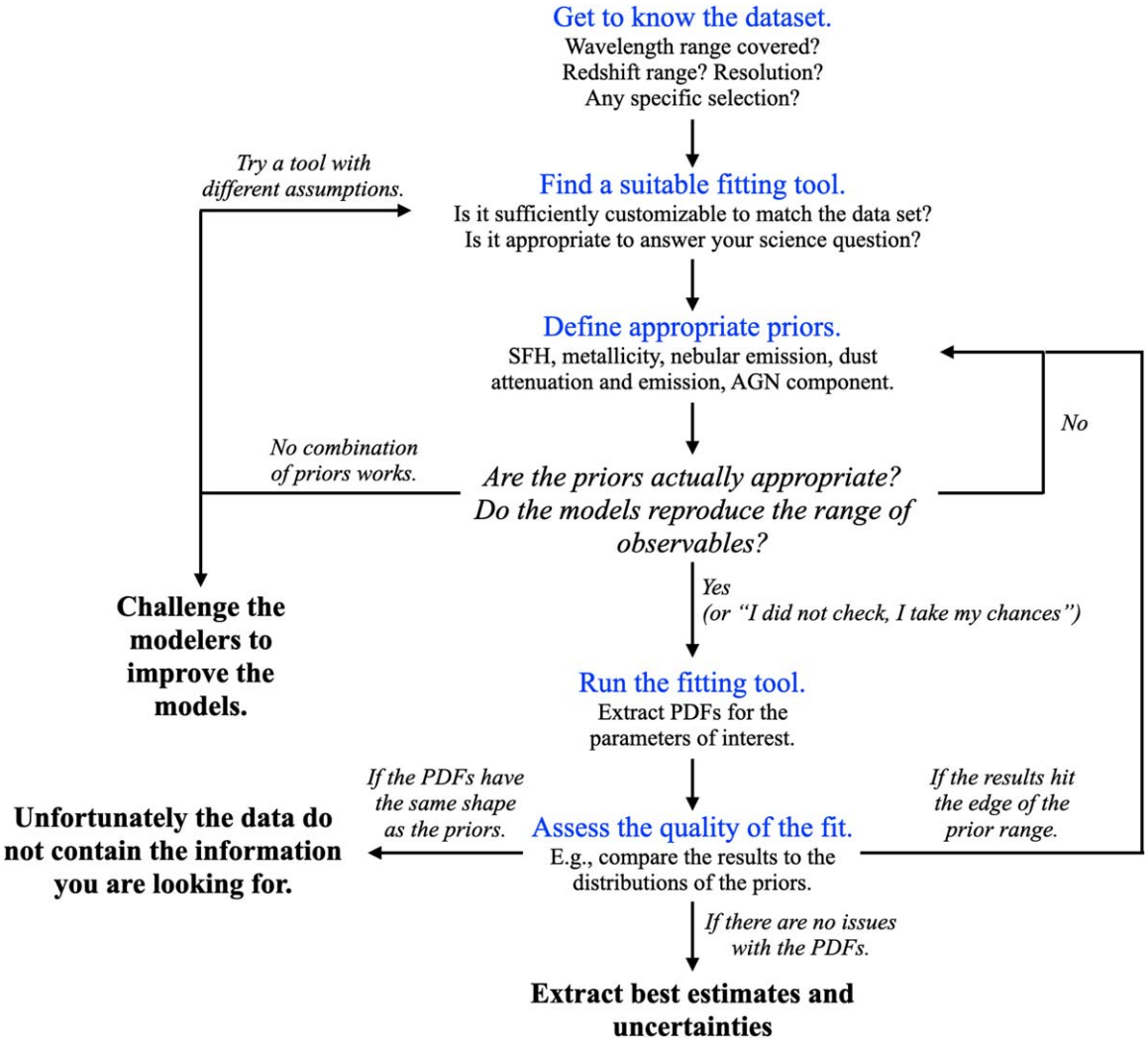

Figure 11. This chart shows the basic flow of the SED-fitting process, which requires understanding the data set, setting the priors, testing the assumptions, running the selected tool, and evaluating the results.

a certain parameter space and all models are compared to each observation (e.g., Dense Basis, CIGALE, FITSED, LePhare, MAGPHYS, P12, SED3FIT, zPhot), and those where the library of models is built on the fly exploring only the range of the parameter space that matches the data more closely (e.g., AGNFitter, BAGPIPES, BEAGLE, Interrogator, Prospector, SpeedyMC). In both cases, it is the user who chooses that parameter space, deciding the range over which the individual ingredients can vary (or not), e.g., IMF, type of SFH and MEH, stellar population models, nebular emission models, dust-attenuation and emission models, AGN models.

The user has to specify a physically reliable range for these parameters and test that the produced templates match the selected observed sample. Such a set of priors allows the user to rule out possible solutions that are nonphysical or simply very unlikely. For example:

1. Low-redshift, massive galaxies are likely in a declining SFR phase, while high-redshift galaxies are likely in a rising SFR phase (see, e.g., Madau & Dickinson 2014).
2. High-redshift galaxies tend to be more metal-poor than low-redshift ones (see, e.g., Maiolino et al. 2008).
3. The SFR seems to be correlated with the stellar mass at all redshifts (see, e.g., Whitaker et al. 2012; Schreiber et al. 2015).

If fitting a large sample of galaxies, it is important to account for all these possibilities because we do not necessarily know what types of galaxies we are looking at, and restricting the priors might lead to missing a very interesting population of outliers. If instead the aim is to refine the fits of a specific population of galaxies, it is appropriate to restrict the priors to account for existing knowledge. Any model can return a χ^2 and a likelihood, but it is in the hands of the user to set up the appropriate models for the specific observations.

There are at least two ways to validate the choice of the priors, and they should both be adopted. The first is to compare observable quantities (e.g., emission-line fluxes and ratios, spectral indices, photometric colors) between the model and the observations. An appropriate library of models will span a similar (if not larger) space in those observable quantities compared to the galaxies in the observed sample. The second way requires running the fitting tool first and analyzing the distribution of the results. We will talk about this in the next Section (7.3).

This brings us to the concept that SED fitting can often be an iterative process, where the user refines the priors to find the appropriate set for the specific set of observations. When the sample to be analyzed includes all sorts of galaxies, it is usually better to start with priors that allow for all possible combination of parameters and thus a very large parameter space. Alternatively, if the sample is selected in a specific way (e.g., spectroscopy data tell us that all the galaxies have low Balmer breaks), the parameter space can be tailored to match the observations (e.g., the dust-attenuation parameter space can be reasonably small, i.e., there is no need to include model galaxies with large A_V in the prior).

7.3. Outputs and Results

Many commonly used codes return as outputs the physical parameters associated with the best-fitting model. For physical parameters that are sensitive mainly to the normalization of the SED, this could be a sensible approximation. However, small variations in the fluxes generating the SEDs could result in very different colors and, hence, very different physical parameters (Ocvirk et al. 2006). In this case, models with different physical parameters can match the data to the same likelihood (same χ^2) resulting in degeneracies. This needs to be taken into account in the SED-fitting process.

To move away from the single best-fitting model, we can generate PDFs for each of the parameters of interest using the likelihood of each model to match the data. In case of a PDF being a Gaussian function, we can define a peak (most likely value for the parameter) and a width (the uncertainty in the parameter). In cases of skewed or nonsymmetric distributions, we define the 50th percentile as the "best estimate" and the distance between the 84th and 16th percentiles as the confidence interval.

The case of bimodal PDFs is tricky, and thus it is in the hands of the user to decide what to report as a result because neither a median nor a single percentile can be acceptable values. This means there is a strong degeneracy that needs to be understood (see, e.g., the redshift–dust attenuation degeneracy in the work by da Cunha et al. 2015). In such a case, we can suggest carrying the full PDF. The attention of the user is also required when the prior distribution and the PDF of a certain physical parameter have the same shape and extension. This should tell the user that the specific parameter is not constrained. This does not mean that the fit is useless. It simply means that the information relative to that specific parameter is not in the data. Alternatively, when the PDFs hit the edges of the allowed range, one needs to extend the parameter space sampled in the model library.

In summary, using a single best-fitting model can easily introduce unnecessary scatter in the correlations we wish to study. Exploring the full PDFs is the safest choice, and it can tell us whether the parameter we are trying to measure can actually be constrained by the data we have and whether the model library we are using is appropriate for our project.

8. Summary and Conclusion

In recent years, data have improved dramatically in quality and quantity, and SED-fitting techniques and codes have started to upgrade accordingly. There is a large number of publicly available SED-fitting codes. They are generally customizable and adaptable to different data sets. Yet, when used to interpret the same data set, they provide different results. Here, we explore the differences among the results (stellar mass, SFR, and dust attenuation) of 14 different SED-fitting codes, run on three data sets: galaxies at $z \sim 1$ with photometry spanning rest-frame UV-to-NIR wavelengths; a subset of the first data set including IR measurements out to 200 μm; and galaxies at $z \sim 3$ with photometry spanning rest-frame UV-to-NIR wavelengths.

The aim of this exercise is twofold: assessing the differences and eventual biases between codes to demonstrate the impact of different modeling choices on the results; and measuring the modeling uncertainties that need to be taken into account when comparing studies that use different assumptions and when comparing observational results with theoretical predictions. We stress that we do not attempt to identify any "best" code or set of assumptions. We use the SFR–M_* relation as a case example to determine the similarities and differences among the codes.

When assessing the differences between the codes, we find the following:

1. Qualitatively, all codes return similar distributions in terms of stellar mass. The distributions of the results for the SFR and dust-attenuation parameter (A_V) show differences in the location of the peaks among the codes.
2. The codes that tend to return larger SFR values also return larger A_V values, indicating a strong age–dust degeneracy for the used data sets and models.
3. The codes that can process IR measurements return smaller SFR values when the IR measurements are included in the fit compared to when the IR measurements are not included in the fit. This difference is smaller for the codes that assume flexible SFHs, suggesting that the fit is finding *older* solutions instead of *dustier* solutions to match the *red* SED.
4. All the codes that model also an AGN component return AGN fractions smaller than 20% for data set number 2.
5. All codes return identifiable SFR–M_* relations. The slope is very similar in all codes. The normalization can vary by about 0.1–0.2 dex between codes. The scatter varies the most along with the distribution of the galaxies that fall below the relation.

When measuring the modeling uncertainties, we find the following:

1. The median modeling uncertainty is around 0.12 dex for stellar mass, 0.27 dex for SFR (dominated by galaxies on the SFR–M_* relation), and 0.27 mag for A_V at $z \sim 1$. It is around 0.16 dex for stellar mass, 0.28 dex for SFR, and 0.22 mag for A_V at $z \sim 3$.
2. The modeling uncertainty in SFR can increase to 1 dex or more for galaxies with low SFR ($\log(\mathrm{SFR}/(M_\odot \mathrm{yr}^{-1}) < 0$).
3. The modeling uncertainty in A_V is a strong function of A_V itself and can be as larger than 0.4 mag for $A_V > 2$.

Toward the end of the paper, we also provide some examples and best practices to use any SED-fitting tool, not as a black box but being cognizant about the modeling assumptions and how those can shape the results. We advise the user to always test different modeling assumptions and measure the appropriate modeling uncertainties to be able to put results into context with other studies or theoretical models.

The data sets, the outputs from the different SED-fitting tools, and the Jupyter Notebooks used to produce the figures in this paper are published on GitHub,[62] and copies of these materials have been deposited to Zenodo doi:10.5281/zenodo.7396555.

We thank the anonymous referee for their very constructive report. We thank the University of California, Riverside for hosting the workshop where this work started. The workshop was supported by National Science Foundation funding. This paper does not reflect the views or opinions of the National Science Foundation or the American Association for the Advancement of Science (AAAS). We thank Audrey Galametz, Joel Primack, and Meaghann Stoelting for insightful conversations. C.P. was supported by the Canadian Space Agency under a contract with NRC Herzberg Astronomy and Astrophysics. Support for K.I. was provided by NASA through the NASA Hubble Fellowship grant No. HST-HF2-51508 awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under contract NAS5-26555. Support for V.P. was provided by NASA through the NASA Hubble Fellowship grant No. HST-HF2-51489 awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under contract NAS5-26555. M.B. acknowledges support from FONDECYT regular grant No. 1211000 and by the ANID BASAL project FB210003. K.M. is grateful for support from the Polish National Science Centre via grant No. UMO-2018/30/E/ST9/00082. For the purpose of open access, the author has applied a Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising from this submission.

Software: NumPy (Harris et al. 2020), Matplotlib (Hunter 2007), pandas (pandas development team 2020), Astropy (Astropy Collaboration et al. 2013, 2018), seaborn (Waskom 2021), joypy (https://github.com/leotac/joypy), and statsmodels (Seabold & Perktold 2010).

ORCID iDs

Camilla Pacifici https://orcid.org/0000-0003-4196-0617
Kartheik G. Iyer https://orcid.org/0000-0001-9298-3523
Elisabete da Cunha https://orcid.org/0000-0001-9759-4797
Viviana Acquaviva https://orcid.org/0000-0002-6788-6315
Denis Burgarella https://orcid.org/0000-0002-4193-2539
Adam C. Carnall https://orcid.org/0000-0002-1482-5818
Yu-Yen Chang https://orcid.org/0000-0002-6720-8047
Kevin C. Cooke https://orcid.org/0000-0002-2200-9845
Jeyhan Kartaltepe https://orcid.org/0000-0001-9187-3605
Joel Leja https://orcid.org/0000-0001-6755-1315
Brett Salmon https://orcid.org/0000-0002-7453-7279
Alba Vidal-García https://orcid.org/0000-0001-6798-9202
Médéric Boquien https://orcid.org/0000-0003-0946-6176
Gabriel G. Brammer https://orcid.org/0000-0003-2680-005X
Michael J. I. Brown https://orcid.org/0000-0002-1207-9137
Jacopo Chevallard https://orcid.org/0000-0002-7636-0534
Darren Croton https://orcid.org/0000-0002-5009-512X
Iary Davidzon https://orcid.org/0000-0002-2951-7519
Mark Dickinson https://orcid.org/0000-0001-5414-5131
Harry C. Ferguson https://orcid.org/0000-0001-7113-2738
Adriano Fontana https://orcid.org/0000-0003-3820-2823
Yicheng Guo https://orcid.org/0000-0003-2775-2002
Shoubaneh Hemmati https://orcid.org/0000-0003-2226-5395
Susan A. Kassin https://orcid.org/0000-0002-3838-8093
Rebecca L. Larson https://orcid.org/0000-0003-2366-8858
Bomee Lee https://orcid.org/0000-0003-1954-5046
Hooshang Nayyeri https://orcid.org/0000-0001-8242-9983
Jeffrey A. Newman https://orcid.org/0000-0001-8684-2222
Viraj Pandya https://orcid.org/0000-0002-2499-9205
Naveen Reddy https://orcid.org/0000-0001-9687-4973
Ryan Sanders https://orcid.org/0000-0003-4792-9119
Ekta Shah https://orcid.org/0000-0001-7811-9042
Abtin Shahidi https://orcid.org/0000-0001-6975-6293
Matthew L. Stevans https://orcid.org/0000-0001-8379-7606
Alexander de la Vega https://orcid.org/0000-0002-6219-5558
Weichen Wang https://orcid.org/0000-0002-9593-8274

[62] https://github.com/camipacifici/art_sedfitting

References

Abdurro'uf, Lin, Y.-T., Wu, P.-F., & Akiyama, M. 2021, ApJS, 254, 15
Acquaviva, V., Gawiser, E., & Guaita, L. 2011, ApJ, 737, 47
Acquaviva, V., Gawiser, E., & Guaita, L. 2012, IAU Symp. 284, The Spectral Energy Distribution of Galaxies—SED 2011, ed. R. J. Tuffs & C. C. Popescu (Cambridge: Cambridge Univ. Press), 42
Arnouts, S., Cristiani, S., Moscardini, L., et al. 1999, MNRAS, 310, 540
Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, AJ, 156, 123
Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33
Baldwin, C., McDermid, R. M., Kuntschner, H., Maraston, C., & Conroy, C. 2018, MNRAS, 473, 4698
Barišić, I., Pacifici, C., van der Wel, A., et al. 2020, ApJ, 903, 146
Barro, G., Pérez-González, P. G., Cava, A., et al. 2019, ApJS, 243, 22
Battisti, A. J., Cunha, E. d., Shivaei, I., & Calzetti, D. 2020, ApJ, 888, 108
Battisti, A. J., da Cunha, E., Grasha, K., et al. 2019, ApJ, 882, 61
Bell, E. F., & de Jong, R. S. 2001, ApJ, 550, 212
Bellstedt, S., Robotham, A. S. G., Driver, S. P., et al. 2020, MNRAS, 498, 5581
Bellstedt, S., Robotham, A. S. G., Driver, S. P., et al. 2021, MNRAS, 503, 3309
Berta, S., Lutz, D., Santini, P., et al. 2013, A&A, 551, A100
Berti, A. M., Coil, A. L., Hearin, A. P., & Behroozi, P. S. 2021, AJ, 161, 49
Boogaard, L. A., Brinchmann, J., Bouché, N., et al. 2018, A&A, 619, A27
Boquien, M., Burgarella, D., Roehlly, Y., et al. 2019, A&A, 622, A103
Bowman, W. P., Zeimann, G. R., Nagaraj, G., et al. 2020, ApJ, 899, 7
Brammer, G. B., van Dokkum, P. G., Franx, M., et al. 2012, ApJS, 200, 13
Brinchmann, J., Charlot, S., White, S. D. M., et al. 2004, MNRAS, 351, 1151
Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000
Bruzual A, G. 1983, ApJ, 273, 105
Bruzual A, G., & Charlot, S. 1993, ApJ, 405, 538
Buat, V., Boquien, M., Małek, K., et al. 2018, A&A, 619, A135
Buat, V., Iglesias-Páramo, J., Seibert, M., et al. 2005, ApJL, 619, L51
Buchner, J. 2016, S&C, 26, 383
Burgarella, D., Buat, V., & Iglesias-Páramo, J. 2005, MNRAS, 360, 1413
Calistro Rivera, G., Lusso, E., Hennawi, J. F., & Hogg, D. W. 2016, ApJ, 833, 98
Calzetti, D. 2001, PASP, 113, 1449
Caplar, N., & Tacchella, S. 2019, MNRAS, 487, 3845
Carnall, A. C., Leja, J., Johnson, B. D., et al. 2019a, ApJ, 873, 44
Carnall, A. C., McLure, R. J., Dunlop, J. S., & Dave, R. 2018, MNRAS, 480, 4379
Carnall, A. C., McLure, R. J., Dunlop, J. S., et al. 2019b, MNRAS, 490, 417
Chabrier, G. 2003, PASP, 115, 763
Chang, Y.-Y., Le Floc'h, E., Juneau, S., et al. 2017a, MNRAS, 466, L103
Chang, Y.-Y., Le Floc'h, E., Juneau, S., et al. 2017b, ApJS, 233, 19
Charlot, S., & Fall, S. M. 2000, ApJ, 539, 718
Charlot, S., & Longhetti, M. 2001, MNRAS, 323, 887
Chevallard, J., & Charlot, S. 2016, MNRAS, 462, 1415
Chevallard, J., Curtis-Lake, E., Charlot, S., et al. 2019, MNRAS, 483, 2621

Ciesla, L., Elbaz, D., & Fensch, J. 2017, A&A, 608, A41
Conroy, C. 2013, ARA&A, 51, 393
Conroy, C., Gunn, J. E., & White, M. 2009, ApJ, 699, 486
Curtis-Lake, E., Chevallard, J., Charlot, S., & Sandles, L. 2021, MNRAS, 503, 4855
da Cunha, E., Charlot, S., & Elbaz, D. 2008, MNRAS, 388, 1595
da Cunha, E., Walter, F., Smail, I. R., et al. 2015, ApJ, 806, 110
Davidzon, I., Laigle, C., Capak, P. L., et al. 2019, MNRAS, 489, 4817
Davies, J. I., Baes, M., Bianchi, S., et al. 2017, PASP, 129, 044102
Davies, L. J. M., Thorne, J. E., Robotham, A. S. G., et al. 2021, MNRAS, 506, 256
Driver, S. P., Hill, D. T., Kelvin, L. S., et al. 2011, MNRAS, 413, 971
Drouart, G., & Falkendal, T. 2018, MNRAS, 477, 4981
Du, W., & McGaugh, S. S. 2020, AJ, 160, 122
Eldridge, J. J., & Stanway, E. R. 2016, MNRAS, 462, 3302
Eufrasio, R. T. 2017, Lightning: SED Fitting Package, Astrophysics Source Code Library, ascl:1711.009
Fioc, M., & Rocca-Volmerange, B. 2019, A&A, 623, A143
Fontana, A., D'Odorico, S., Poli, F., et al. 2000, AJ, 120, 2206
Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306
Gilda, S., Lower, S., & Narayanan, D. 2021, ApJ, 916, 43
Goddard, D., Thomas, D., Maraston, C., et al. 2017, MNRAS, 466, 4731
Grogin, N. A., Kocevski, D. D., Faber, S. M., et al. 2011, ApJS, 197, 35
Guiderdoni, B., & Rocca-Volmerange, B. 1987, A&A, 186, 1
Guo, Y., Ferguson, H. C., Giavalisco, M., et al. 2013, ApJS, 207, 24
Han, Y., & Han, Z. 2019, ApJS, 240, 3
Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, Natur, 585, 357
Hunt, L. K., De Looze, I., Boquien, M., et al. 2019, A&A, 621, A51
Hunter, J. D. 2007, CSE, 9, 90
Ilbert, O., Arnouts, S., McCracken, H. J., et al. 2006, A&A, 457, 841
Iyer, K., & Gawiser, E. 2017, ApJ, 838, 127
Iyer, K. G., Gawiser, E., Faber, S. M., et al. 2019, ApJ, 879, 116
Johnson, B. D., Leja, J., Conroy, C., & Speagle, J. S. 2021, ApJS, 254, 22
Kass, R. E., & Raftery, A. E. 1995, J. Am. Stat. Assoc., 90, 773
Katsianis, A., Zheng, X., Gonzalez, V., et al. 2019, ApJ, 879, 11
Kennicutt, R. C., Jr. 1989, ApJ, 344, 685
Kennicutt, R. C., Jr. 1998, ARA&A, 36, 189
Kennicutt, R. C., Calzetti, D., Aniano, G., et al. 2011, PASP, 123, 1347
Kochanek, C. S., Eisenstein, D. J., Cool, R. J., et al. 2012, ApJS, 200, 8
Kodra, D., Andrews, B. H., Newman, J. A., et al. 2023, ApJ, 942, 36
Koekemoer, A. M., Faber, S. M., Ferguson, H. C., et al. 2011, ApJS, 197, 36
Kriek, M., & Conroy, C. 2013, ApJL, 775, L16
Kriek, M., Shapley, A. E., Reddy, N. A., et al. 2015, ApJS, 218, 15
Lacy, M., Petric, A. O., Sajina, A., et al. 2007, AJ, 133, 186
Lacy, M., Storrie-Lombardi, L. J., Sajina, A., et al. 2004, ApJS, 154, 166
Lee, B., Giavalisco, M., Whitaker, K., et al. 2018, ApJ, 853, 131
Leitherer, C., & Heckman, T. M. 1995, ApJS, 96, 9
Leja, J., Carnall, A. C., Johnson, B. D., Conroy, C., & Speagle, J. S. 2019, ApJ, 876, 3
Leja, J., Johnson, B. D., Conroy, C., Dokkum, P. G. v., & Byler, N. 2017, ApJ, 837, 170
Leja, J., Speagle, J. S., Ting, Y.-S., et al. 2022, ApJ, 936, 165
Liu, D., Daddi, E., Dickinson, M., et al. 2018, ApJ, 853, 172
Lower, S., Narayanan, D., Leja, J., et al. 2020, ApJ, 904, 33
Madau, P., & Dickinson, M. 2014, ARA&A, 52, 415
Magnelli, B., Popesso, P., Berta, S., et al. 2013, A&A, 553, A132
Maiolino, R., Nagao, T., Grazian, A., et al. 2008, A&A, 488, 463
Małek, K., Buat, V., Roehlly, Y., et al. 2018, A&A, 620, A50
Maraston, C. 1998, MNRAS, 300, 872
Matthee, J., & Schaye, J. 2018, MNRAS, 479, L34
McCracken, H. J., Milvang-Jensen, B., Dunlop, J., et al. 2012, A&A, 544, A156
McLure, R. J., Pentericci, L., Cimatti, A., et al. 2018, MNRAS, 479, 25
Merlin, E., Fortuni, F., Torelli, M., et al. 2019, MNRAS, 490, 3309
Momcheva, I. G., Brammer, G. B., van Dokkum, P. G., et al. 2016, ApJS, 225, 27
Muzzin, A., Marchesini, D., Stefanon, M., et al. 2013, ApJS, 206, 8
Newman, J. A., Cooper, M. C., Davis, M., et al. 2013, ApJS, 208, 5
Noeske, K. G., Weiner, B. J., Faber, S. M., et al. 2007, ApJL, 660, L43
Ocvirk, P., Pichon, C., Lancon, A., & Thiebaut, E. 2006, MNRAS, 365, 46
Pacifici, C., Charlot, S., Blaizot, J., & Brinchmann, J. 2012, MNRAS, 421, 2002
Pacifici, C., Kassin, S. A., Weiner, B., Charlot, S., & Gardner, J. P. 2013, ApJL, 762, L15
Pacifici, C., da Cunha, E., Charlot, S., et al. 2015, MNRAS, 447, 786
pandas development team, T. 2020, pandas-dev/pandas: Pandas, latest, Zenodo, doi:10.5281/zenodo.3509134
Pandya, V., Brennan, R., Somerville, R. S., et al. 2017, MNRAS, 472, 2054
Papovich, C., Dickinson, M., & Ferguson, H. C. 2001, ApJ, 559, 620
Pappalardo, C., Bendo, G. J., Boquien, M., et al. 2021, A&A, 655, A104
Pentericci, L., McLure, R. J., Garilli, B., et al. 2018, A&A, 616, A174
Popesso, P., Concas, A., Morselli, L., et al. 2019a, MNRAS, 483, 3213
Popesso, P., Morselli, L., Concas, A., et al. 2019b, MNRAS, 490, 5285
Qiu, Y., & Kang, X. 2022, ApJ, 930, 66
Rosario, D. J. 2019, FortesFit: Flexible Spectral Energy Distribution Modelling with a Bayesian Backbone, Astrophysics Source Code Library, ascl:1904.011
Rujopakarn, W., Rieke, G. H., Weiner, B. J., et al. 2013, ApJ, 767, 73
Salim, S., Boquien, M., & Lee, J. C. 2018, ApJ, 859, 11
Salim, S., & Narayanan, D. 2020, ARA&A, 58, 529
Salmon, B., Papovich, C., Finkelstein, S. L., et al. 2015, ApJ, 799, 183
Salmon, B., Papovich, C., Long, J., et al. 2016, ApJ, 827, 20
Sánchez, S. F., Pérez, E., Sánchez-Blázquez, P., et al. 2016, RMxAA, 52, 21
Santini, P., Ferguson, H. C., Fontana, A., et al. 2015, ApJ, 801, 97
Schreiber, C., Pannella, M., Elbaz, D., et al. 2015, A&A, 575, A74
Scodeggio, M., Guzzo, L., Garilli, B., et al. 2018, A&A, 609, A84
Scoville, N., Aussel, H., Brusa, M., et al. 2007, ApJS, 172, 1
Seabold, S., & Perktold, J. 2010, in 9th Python in Science Conf., ed. S. van der Walt & J. Millman
Simet, M., Chartab, N., Lu, Y., & Mobasher, B. 2021, ApJ, 908, 47
Skilling, J. 2004, in AIP Conf. Proc. 735, Bayesian Inference and Maximum Entropy Methods in Science and Engineering: 24th Int. Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering, ed. R. Fischer, R. Preuss, & U. von Toussaint (Melville, NY: AIP), 395
Skilling, J. 2006, BayAn, 1, 833
Smit, R., Bouwens, R. J., Labbé, I., et al. 2014, ApJ, 784, 58
Speagle, J. S., Steinhardt, C. L., Capak, P. L., & Silverman, J. D. 2014, ApJS, 214, 15
Spinrad, H., & Taylor, B. J. 1969, ApJ, 157, 1279
Steidel, C. C., Giavalisco, M., Pettini, M., Dickinson, M., & Adelberger, K. L. 1996, ApJL, 462, L17
Straatman, C. M. S., Spitler, L. R., Quadri, R. F., et al. 2016, ApJ, 830, 51
Tacchella, S., Dekel, A., Carollo, C. M., et al. 2016, MNRAS, 457, 2790
Thorne, J. E., Robotham, A. S. G., Davies, L. J. M., et al. 2021, MNRAS, 505, 540
van der Wel, A., Noeske, K., Bezanson, R., et al. 2016, ApJS, 223, 29
Vazdekis, A. 1999, ApJ, 513, 224
Wang, W., Kassin, S. A., Pacifici, C., et al. 2018, ApJ, 869, 161
Waskom, M. 2021, JOSS, 6, 3021
Weaver, J. R., Kauffmann, O. B., Ilbert, O., et al. 2022, ApJS, 258, 11
Whitaker, K. E., Franx, M., Leja, J., et al. 2014, ApJ, 795, 104
Whitaker, K. E., van Dokkum, P. G., Brammer, G., & Franx, M. 2012, ApJL, 754, L29
Wilkinson, D. M., Maraston, C., Goddard, D., Thomas, D., & Parikh, T. 2017, MNRAS, 472, 4297
Williams, C. C., Curtis-Lake, E., Hainline, K. N., et al. 2018, ApJS, 236, 33
Worthey, G. 1994, ApJS, 95, 107
Yang, G., Boquien, M., Buat, V., et al. 2020, MNRAS, 491, 740
Yang, G., Boquien, M., Brandt, W. N., et al. 2022, ApJ, 927, 192
York, D. G., Adelman, J., Anderson, J. E., Jr., et al. 2000, AJ, 120, 1579