



**HAL**  
open science

# GALEX-SDSS-WISE LEGACY CATALOG (GSWLC): STAR FORMATION RATES, STAR MASSES, AND DUST ATTENUATIONS OF 700,000 LOW-REDSHIFT GALAXIES

Samir Salim, Janice C. Lee, Steven Janowiecki, Elisabete Da Cunha, Mark  
Dickinson, Médéric Boquien, Denis Burgarella, John J. Salzer, Stéphane  
Charlot

► **To cite this version:**

Samir Salim, Janice C. Lee, Steven Janowiecki, Elisabete Da Cunha, Mark Dickinson, et al.. GALEX-SDSS-WISE LEGACY CATALOG (GSWLC): STAR FORMATION RATES, STAR MASSES, AND DUST ATTENUATIONS OF 700,000 LOW-REDSHIFT GALAXIES. 2024. hal-01440100

**HAL Id: hal-01440100**

**<https://hal.science/hal-01440100>**

Preprint submitted on 20 Mar 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

GALEX–SDSS–WISE LEGACY CATALOG (GSWLC):  
STAR FORMATION RATES, STELLAR MASSES AND DUST ATTENUATIONS OF 700,000 LOW-REDSHIFT  
GALAXIES

SAMIR SALIM<sup>1</sup>, JANICE C. LEE<sup>2,3</sup>, STEVEN JANOWIECKI<sup>4</sup>, ELISABETE DA CUNHA<sup>5</sup>, MARK DICKINSON<sup>6</sup>, MÉDÉRIC BOQUIEN<sup>7</sup>, DENIS BURGARELLA<sup>8</sup>, JOHN J. SALZER<sup>1</sup>, STÉPHANE CHARLOT<sup>9</sup>

Draft

ABSTRACT

In this paper, we present GALEX–SDSS–WISE Legacy Catalog (GSWLC), a catalog of physical properties (stellar masses, dust attenuations and star formation rates (SFRs)) for  $\sim 700,000$  galaxies with SDSS redshifts below 0.3. GSWLC contains galaxies within the *GALEX* footprint, regardless of a UV detection, covering 90% of SDSS. The physical properties were obtained from UV/optical SED fitting following Bayesian methodology of [Salim et al. \(2007\)](#), with improvements such as blending corrections for low-resolution UV photometry, flexible dust attenuation laws, and emission line corrections. GSWLC includes mid-IR SFRs derived from IR templates based upon 22  $\mu\text{m}$  *WISE* observations. These estimates are independent of UV/optical SED fitting, in order to separate possible systematics. The paper argues that the comparison of specific SFRs (SSFRs) is more informative and physically motivated than the comparison of SFRs. SSFRs resulting from the UV/optical SED fitting are compared to the mid-IR SSFRs, and to SSFRs from three published catalogs. For “main sequence” galaxies with no AGN contribution all SSFRs are in very good agreement (within 0.1 dex on average). In particular, the widely-used aperture-corrected SFRs from MPA/JHU catalog show no systematic offsets, in contrast to some integral-field spectroscopy results. For galaxies below the main sequence ( $\log \text{SSFR} < -11$ ), mid-IR (S)SFRs based on fixed luminosity–SFR conversion are severely biased (up to 2 dex) because the dust is primarily heated by old stars. Furthermore, mid-IR (S)SFRs are overestimated by up to 0.6 dex for galaxies with AGN, presumably due to non-stellar dust heating. UV/optical (S)SFRs are thus preferred to IR-based (S)SFRs for quenched galaxies and those which host AGN.

*Keywords:* galaxies: fundamental parameters—galaxies: star formation

1. INTRODUCTION

The stellar mass and the current star formation rate (SFR) are two of the most fundamental physical properties of a galaxy. The SFR normalized by stellar mass, i.e., the specific star formation rate (SSFR, [Tully et al. 1982](#); [Bothun 1982](#)), in addition serves as a rough indicator of a galaxy’s star formation (SF) history. To infer properties such as stellar mass and SFR from observed quantities, stellar populations and the effects of dust attenuation are modeled and compared to galaxy spectra (e.g., [Cid Fernandes et al. 2005](#)) and/or the integrated photometry.

Pioneering work to derive physical properties of galaxies from their integrated light (e.g., [Faber 1972](#); [Searle et al. 1973](#); [Tinsley & Gunn 1976](#); [Larson & Tinsley 1978](#)) paved the way for the development of more advanced stellar population synthesis (SPS) models in the 1980s (e.g., [Bruzual A. 1983](#); [Renzini & Buzzoni 1986](#); [Guiderdoni & Rocca-](#)

[Volmerange 1987](#)). These conventional models were replaced by isochrone synthesis in the following decade (e.g., [Charlot & Bruzual 1991](#); [Bruzual A. & Charlot 1993](#)), culminating in modern, high-resolution models (e.g., ([Bruzual & Charlot 2003](#); [Maraston 2005](#); [Conroy et al. 2010b](#))).

SPS models are the basis for widely used “simple” relations for inferring the stellar mass from broad-band optical colors (e.g., [Bell & de Jong 2001](#); [Courteau et al. 2014](#)). These relations are by necessity averaged over galaxy populations and use simplified assumptions, so they cannot take into account variations in SF history, dust attenuation or stellar metallicity. Simple relations are also used for inferring the SFR from UV or  $H\alpha$  luminosity ([Kennicutt 1998](#); [Calzetti 2013](#)). In addition to again using average assumptions, these luminosities first need to be corrected for dust attenuation. For galaxies at  $z \gtrsim 0$ , the quantities that enter these relations must be explicitly corrected to rest-frame luminosities, which requires the spectral energy distribution (SED) of a galaxy to be assumed.

A more flexible and comprehensive approach is to derive physical parameters through SED fitting (see [Walcher et al. 2011](#); [Conroy 2013](#) for overview). SPS models are used to produce a library (a grid, in case of regular sampling intervals) of models in which the SF history and metallicity take a range of values. Model SEDs are further subjected to varying degrees of dust attenuation assuming some attenuation law. Redshifted model fluxes (broad-band magnitudes) are compared to observations (obviating a need for an explicit K-correction), and the best-fitting model is sought and its parameters are adopted as the parameters of the observed galaxy. Unlike with simple relations, in SED fitting every observed flux plays some role in constraining all derived parameters,

salims@indiana.edu

<sup>1</sup> Department of Astronomy, Indiana University, Bloomington, IN 47404, USA

<sup>2</sup> Space Telescope Science Institute, Baltimore, MD 21218, USA

<sup>3</sup> Visiting Astronomer, Spitzer Science Center, Caltech, Pasadena, CA 91125, USA

<sup>4</sup> International Centre for Radio Astronomy Research (ICRAR), The University of Western Australia, Perth, WA 6009, Australia

<sup>5</sup> Research School of Astronomy and Astrophysics, Australian National University, Canberra, ACT 2611, Australia

<sup>6</sup> National Optical Astronomy Observatory, Tucson, AZ 85719, USA

<sup>7</sup> Universidad de Antofagasta, Unidad de Astronomía, Antofagasta 1270300, Chile

<sup>8</sup> Laboratoire d’Astrophysique de Marseille, Aix-Marseille Université, CNRS, F-13013 Marseille, France

<sup>9</sup> Institut d’Astrophysique de Paris, CNRS, F-75014 Paris, France

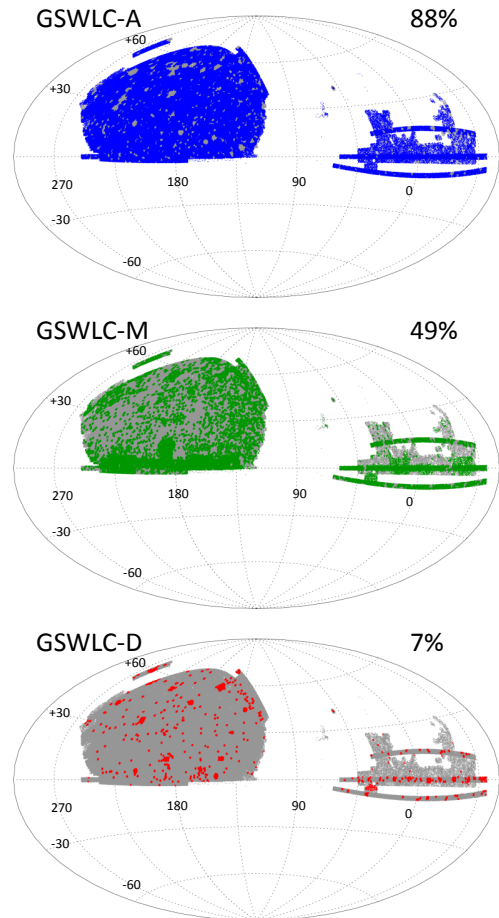
taking full advantage of the full set of observations. While in SED fitting the relationships between the observed and derived parameters are not explicit as with the simple relations, it should be noted that both approaches are ultimately rooted in SPS models.

For stellar SED fitting to produce useful constraints on the ongoing SFR, the rest-frame UV photometry must be included. SED fitting was initially employed to derive SFRs from broad-band observations of small samples of Lyman break galaxies at  $2 < z < 3$ , for which the UV was redshifted to easily observed optical range (Sawicki & Yee 1998; Papovich et al. 2001; Shapley et al. 2005). Subsequently, the UV surveys carried out by *Galaxy Evolution Explorer* (GALEX) satellite (Martin et al. 2005) placed low-redshift ( $z < 0.3$ ) galaxies within the domain of SED fitting. Salim et al. (2005, 2007) performed the SED fitting of low-redshift galaxies by combining the UV photometry from GALEX with the optical photometry from Sloan Digital Sky Survey (SDSS). Furthermore, they utilized a Bayesian approach to SED fitting, following the methods pioneered by the Max Planck Institute for Astrophysics / Johns Hopkins University (MPA/JHU) group (Kauffmann et al. 2003a; Brinchmann et al. 2004; Tremonti et al. 2004). Unlike the traditional best-fit (minimum  $\chi^2$ ) approach, Bayesian SED fitting determines the full probability distribution function of any parameter (or combination of parameters), yielding more robust parameter characterization, along with its uncertainty (e.g., Taylor et al. 2011). Subsequent efforts expanded Bayesian SED fitting to include thermal dust emission (da Cunha et al. 2008; Noll et al. 2009) and star formation histories derived from cosmological simulations (Pacifci et al. 2013).

Salim et al. (2007) (hereafter S07) performed a thorough investigation of SFRs in the local universe and showed that the UV/optical SED fitting is especially powerful for obtaining SFRs of galaxies with low SSFRs. However, S07 used an early GALEX data release covering only 10% of SDSS area. Because of its preliminary nature, the catalog of physical parameters from S07 was not released to the public.

Here we present GALEX–SDSS–WISE Legacy Catalog (GSWLC) of physical parameters, which builds on S07 UV/optical SED fitting efforts, but includes numerous methodological improvements, such as UV photometry corrections, flexible dust attenuation curves, and emission line corrections. Many of these improvements were made possible by the use of the Code Investigating GALaxy Emission (CIGALE).<sup>10</sup> (Noll et al. 2009, M. Boquien et al. in prep.), to calculate the models and perform the SED fitting. GSWLC contains SDSS galaxies within the GALEX footprint (regardless of the UV detection), covering 90% of SDSS area. In addition to the stellar masses and the SFRs from the UV/optical SED fitting, GSWLC also includes SFRs derived independently from *The Wide-field Infrared Survey Explorer* (WISE, Wright et al. 2010) mid-infrared (IR) observations. Community access to the catalog is provided online at <http://pages.iu.edu/~salims/gswlc> and from Mikulski Archive for Space Telescopes (MAST).

After defining the scope and the sample of GSWLC (Section 2), we describe the input data and the process of deriving the parameters (Sections 3, 5 and 4). The catalog is described in Section 6. In Section 7 we examine the contents of the catalog, while in Section 8 GSWLC is compared to several previously published catalogs of physical properties of SDSS



**Figure 1.** Distribution on the sky of galaxies included in GALEX–SDSS–WISE Legacy Catalog (GSWLC). Each panel shows SDSS target galaxies as shaded areas, covering  $\sim 9000$  sq. deg. The SDSS Main Galaxy Sample (MGS) legacy survey area consists of the contiguous area on the left (Northern Galactic Cap) and three horizontal stripes on the right. Non-stripe areas on the right are primarily from SDSS BOSS survey of passive galaxies. BOSS galaxies are also interspersed in the contiguous area on the left. Superimposed on each panel in color are SDSS galaxies that fall within the coverage of one of GALEX imaging surveys: shallow (all-sky) survey (AIS), medium-deep survey (MIS) and the deep survey (DIS), thus defining the coverage of GSWLC-A, M and D. Percentage of SDSS targets covered by each GALEX survey is indicated in the upper right corner of each panel.

galaxies.

## 2. SAMPLE

### 2.1. Scope and SDSS target sample

GSWLC is built around SDSS Main Galaxy Sample (MGS, Strauss et al. 2002), the most extensive spectroscopic survey of low-redshift ( $z < 0.3$ ) universe. More specifically, GSWLC includes all MGS-like SDSS galaxies that fall within GALEX footprint, covering  $\sim 90\%$  of SDSS area, or  $\sim 8000$  sq. deg. Galaxies are kept in the sample regardless of the detection in the UV, thus retaining the optical selection. SDSS and GALEX data were obtained through SDSS CasJobs SQL server.

The majority of low-redshift galaxies in SDSS were targeted spectroscopically as part of the magnitude-limited MGS, one of the three original (“legacy”) spectroscopic surveys (others being the luminous red galaxy, LRG, and quasar surveys). Legacy surveys were completed in 2008 with the release of DR7 (Abazajian et al. 2009), and cover a large contiguous area in the Northern Galactic cap, and three separate stripes in the Southern Galactic cap (Figure 1). These areas

<sup>10</sup> <http://cigale.lam.fr>

lie almost entirely in the Northern sky ( $\delta > -10^\circ$ ), and mostly at high Galactic latitude ( $|b| > 30^\circ$ ).

MGS selection algorithm targets non-stellar objects with well-measured photometry brighter than  $r_{\text{petro}} = 17.77$ , with some additional cuts based on surface brightness (Strauss et al. 2002). The magnitude limit of MGS yields a sample of galaxies peaking at  $z \sim 0.1$ , with very few galaxies above  $z = 0.3$ .

Selection of SDSS targets is performed on DR10 (Ahn et al. 2014), based on the `SpecPhoto` table, which combines primary (non-duplicate) photometric objects with primary spectroscopic observations. Nevertheless, in 25 cases more than one primary spectrum was found for the same photometric object, in which case we selected the spectrum that was closest to the photometric position, leaving 730,288 unique objects (objects with unique `ObjID`). These SDSS targets constitute shaded areas in Figure 1.

Our selection for inclusion in GSWLC consists of only two criteria:

$$\begin{aligned} r_{\text{petro}} < 18.0 \\ 0.01 < z < 0.30 \end{aligned} \quad (1)$$

The first criterion replicates MGS brightness cut, with some rounding to allow for targeting photometry fluctuations. The second cut removes stars and very nearby galaxies whose photometry is likely to be less accurate because of their large angular size (West et al. 2010), and whose redshifts are poorer indications of the distance (Tully et al. 2016). GSWLC therefore includes the entire MGS, complemented by MGS-like galaxies from other programs that fall within the brightness (and, effectively, the redshift) cut of MGS: 5.2% from BOSS (typically ellipticals), and the remaining 2.1% from other surveys (QSO, SEGUE). For studies that focus on general population of galaxies, it is strongly recommended to use only the MGS galaxies (`flag_mgs = 1` in GSWLC).

### 2.2. The final sample

GSWLC includes all SDSS targets within areas covered by *GALEX*, regardless of a UV detection. *GALEX* has observed the sky in 1.2 deg. wide circular fields (“tiles”), with a wide range of exposure times (Morrissey et al. 2007). The greatest UV sky coverage is provided by shallow ( $\sim 100$  s) observations, most of which were taken as part of the All-sky Imaging Survey (AIS). *GALEX* could not safely point in the direction of bright stars, so even the All-sky survey’s coverage contains holes. Observations of medium depth, corresponding to one *GALEX* orbit ( $\sim 1500$  s), are the basis for Medium Imaging Survey (MIS), which specifically targeted areas within SDSS footprint. For select fields, deeper observations were obtained by co-adding the observations from multiple orbits, to produce the Deep Imaging Survey (DIS), with nominal exposure time of 30,000 s.

Since the accuracy of UV photometry, and consequently of the derived physical properties, will depend on the depth of the observations, we produce three separate catalogs, each approximately corresponding to shallow (all-sky; A), medium (M), and deep (D) UV imaging surveys. The catalogs are designated as GSWLC-A, GSWLC-M, GSWLC-D. Each catalog is based only on *GALEX* imaging of certain depth, even if deeper exposures of an object exist:

$$\begin{aligned} \text{GSWLC-A: } & t_{\text{NUV}} \leq 650 \\ \text{GSWLC-M: } & 650 < t_{\text{NUV}} < 4000 \\ \text{GSWLC-D: } & t_{\text{NUV}} \geq 4000 \end{aligned} \quad (2)$$

where the exposure times in NUV band are given in seconds. FUV exposure times are typically identical to NUV times (the two bands were observed simultaneously), except in cases when the FUV image was missing due to camera malfunction. Note that in GSWLC-M we also include individual visits used for DIS co-adds, since the exposure times of individual visits fall in the range for GSWLC-M.

To define the final samples for inclusion in the catalogs, we take all SDSS targets from Section 2.1 that fall within 0.6 deg of *GALEX* tile centers of the specific UV survey (A, M or D). This gives the following final sample sizes:

$$\begin{aligned} \text{GSWLC-A: } & 640,659 \quad (88\%) \\ \text{GSWLC-M: } & 361,328 \quad (49\%) \\ \text{GSWLC-D: } & 48,401 \quad (7\%) \end{aligned} \quad (3)$$

where the percentage indicates the fraction of all SDSS DR10 targets. GSWLC-M contains  $7\times$  as many galaxies as the MIS sample used in S07. Since *GALEX* surveys are nested, objects in the deeper catalogs are mostly included in the shallower ones. The number of unique galaxies encompassed by the three catalogs, 658,911, is therefore only slightly larger than the number of objects in GSWLC-A, and corresponds to 90% of SDSS target sample.

### 2.3. SDSS-GALEX matching

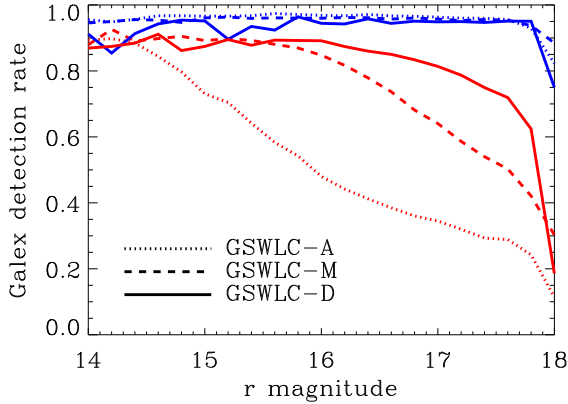
*GALEX* data are taken from the final data release (GR6/7)<sup>11</sup>. Matching of SDSS to *GALEX* is in general a non-trivial task (Budavári et al. 2009) because of the changes in galaxy morphology with wavelength, and different resolutions ( $5''$  for *GALEX* vs.  $1.''3$  for SDSS). Furthermore, the *GALEX* data release contains detections of same objects from multiple tiles. Thus, one needs to address the cases when multiple UV candidates exist for an optical source, and when the UV source may be a blend of two or more optical sources.

In our case, the problem of SDSS–*GALEX* matching is significantly alleviated by the fact that it involves relatively bright objects, having low sky density. In particular, we find that there are essentially no cases of genuine multiple UV candidates within  $5''$  search radius. The choice of search radius is based on *GALEX* positions of isolated SDSS targets, for which we find negligible bulk offset with respect to SDSS ( $0.''2$  in either RA or Decl.), and  $1\sigma$  1D positional uncertainty of  $1.''3$ . When multiple UV candidates are nominally present, they are either due to multiple observations (overlapping tiles), or from NUV and FUV detections of the same object that were erroneously left as separate sources in *GALEX* merged-band catalog. When there are multiple observations of the same object, we take the one from the tile with the longest FUV exposure (if no candidate has an available FUV image, the tile with the longest NUV exposure is selected).

While each SDSS object has an unambiguous *GALEX* match, the UV measurement may still be affected by blending of several objects independently detected in SDSS. We will address this issue and describe the implemented solution in Section 3.

UV detection rates of SDSS objects ( $3\sigma$  threshold, either UV band) are 54% for GSWLC-A, 74% for GSWLC-M and 84% for GSWLC-D. Besides the UV survey depth, the detection rates depend strongly on galaxy’s SSFR or color, as can be seen from Figure 2, where we plot the detection rates as a

<sup>11</sup> <http://galex.stsci.edu/GR6/>



**Figure 2.** UV detection rates of different GSWLC catalogs (defined by UV survey depth). Detection rate is shown as a function of  $r$  magnitude, separately for blue ( $g-r < 0.7$ , observer frame; blue lines) and red galaxies (red lines). The detection rate of red galaxies is strongly dependent on UV depth and magnitude.

function of  $r$ -band magnitude, separately for blue ( $g-r < 0.7$ ) and red galaxies.

#### 2.4. SDSS-WISE and 2MASS matching

GSWLC utilizes *WISE* observations at  $22 \mu\text{m}$  (*WISE* channel W4) to determine SFRs independently from the UV/optical SED fitting. The depth of *WISE* observations over the sky is not uniform ( $6 \pm 1$  mag Vega), but is still much more uniform than *GALEX* depth, and essentially covers the entire sky without gaps. FWHM of W4 PSF is  $12''$ . We use  $22 \mu\text{m}$  photometry from two independent reductions of *WISE* data. The first is the official AllWISE Source Catalog<sup>12</sup>, hosted at IPAC Infrared Science Archive (IRSA) at the Caltech/JPL, and the second is the “unWISE” reduction (Lang et al. 2016)<sup>13</sup>, where SDSS detections served as forced photometry priors.

To match SDSS to AllWISE, we take the closest candidate in a  $5''$  search radius. Note that AllWISE detections and astrometry are based on simultaneous PSF-matched fit to all channels. Thus, the shorter-wavelength channels ( $3.4$  (W1) and  $4.6 \mu\text{m}$  (W2)), which have smaller PSF widths ( $6''$ ) and higher S/N, will dominate the astrometric accuracy ( $< 0''.5$ ). This makes the adopted search radius sufficient to capture genuine matches, and justifies taking the closer candidate in case of multiple matches (3.2% of cases). The detection rate of SDSS targets at  $22 \mu\text{m}$  (based on  $2\sigma$  threshold of profile magnitudes) is 32%, which is lower than the detection rate of *GALEX*.

Since unWISE catalog is derived based on SDSS DR10 detections, no matching is required (sources already carry SDSS ObjID). Forced photometry increases the S/N of measured fluxes, so the  $2\sigma$  detection rate is 41%, a third higher than of the official catalog.

Near-IR photometry ( $JHK_s$ ) is taken from 2MASS Extended Source Catalog (XSC), also hosted at IRSA. We match SDSS to 2MASS XSC by selecting the closest candidate within  $5''$  search radius. The detection rate of SDSS targets is 48%.

### 3. DATA

The data used to construct GSWLC consist of broadband photometry (magnitudes or fluxes and their errors) from *GALEX*, SDSS and *WISE*, and redshifts from SDSS. Comparison datasets will also use line fluxes from SDSS spectra.

#### 3.1. Optical photometry

SDSS offers several choices for galaxy photometry: `modelMag` (magnitude extracted assuming either de Vaucouleurs or exponential profile), `cmodelMag` (weighted average of de Vaucouleurs and exponential magnitudes) and `petroMag` (surface-brightness dependent aperture magnitude). The usual practice in SED fitting is to use fluxes that yield most accurate colors. Of SDSS magnitudes, the `ModelMag` best fulfills that role. We confirm that this is the optimal choice by finding that the best-fitting models (based on optical fitting alone) have three times lower median  $\chi^2$  when `modelMag` are used as opposed to `cmodelMag` or `petroMag`. Also, SSFRs obtained from UV/optical SED fitting that uses `modelMag` magnitudes have a smaller scatter with respect to SSFRs from *WISE* and from  $H\alpha$ . While `ModelMag` produces stable colors, the degree to which it will estimate the total light will depend on galaxy size, morphology (Bernardi et al. 2010; Taylor et al. 2011) and color gradient. Simard et al. (2011); Meert et al. (2016) have derived magnitudes that were designed to better capture the total light in each band. Readers can use those catalogs to scale our estimates of the stellar mass and SFR, if that is required for their goals. Based on the comparison of `ModelMag` and Meert et al. (2016)  $r$ -band magnitude, the typical correction should nevertheless be relatively modest ( $+0.03 \pm 0.08$  dex).

#### 3.2. UV photometry and corrections

The *GALEX* source catalog offers only one measure of flux that is recommended for galaxies: `MAG_AUTO`, a Kron elliptical aperture magnitude derived by SExtractor (Bertin & Arnouts 1996). We investigate the differences in the methodology of this measurement compared to the one used for SDSS photometry, as well as the effects of lower *GALEX* resolution, by deriving `MAG_AUTO` magnitudes of 1000 randomly selected SDSS targets whose  $g$ -band images were degraded to match *GALEX* resolution. By comparison with isolated sources, we find that the differences between *GALEX*-resolution `MAG_AUTO` and SDSS resolution `modelMag` are dominated by systematic offsets (up to 0.4 mag) arising from close companions (especially at  $d < 10''$ ) which become blended in lower resolution, rather than the differences in the methodology of magnitude measurement ( $< 0.1$  mag difference).

To account for the systematics that arise from both the difference in *GALEX* and SDSS resolutions and the methods of flux measurement, we derive and implement several corrections to UV fluxes. Corrections are derived by comparing the UV magnitudes predicted from optical-only SED fitting of star-forming galaxies ( $g-r < 0.7$ ) with the actual UV magnitudes. While any individual predicted UV magnitude is crude ( $\sigma \approx 0.4$  mag in NUV), the corrections are accurate owing to a very large number of galaxies that define them.

First, we confirm that there are no zero-point offsets between the predicted and the actual UV magnitudes. Next, we find that a small edge-of-detector correction is required for NUV, but not FUV, photometry. NUV magnitudes of sources that appear in the outermost  $8'$  annulus of the field of view, which accounts for 40% of the detector area, are up to 0.1

<sup>12</sup> <http://wise2.ipac.caltech.edu/docs/release/allwise/>.

<sup>13</sup> [unwise.me](http://unwise.me).

mag too bright, possibly because of an inaccurate flat fielding. The amount of correction depends linearly on the radial position within the annulus:

$$\text{NUV}_{\text{corr}} = \text{NUV} + 0.78 \text{FOV} - 0.37, \quad \text{FOV} > 0.47 \quad (4)$$

where FOV is the distance from the center of the tile in degrees.

Next, we derive the centroid shift correction. The shift between optical and UV positions arises due to random errors, mostly from the lower accuracy of the *GALEX* astrometry, but also from the differences in UV versus optical morphology, which has a slight effect on the measured magnitudes. Again, using the difference between predicted and observed NUV magnitudes, we find the following correction:

$$\text{UV}_{\text{corr}} = \text{UV} + 0.054 \Delta x - 0.049, \quad \Delta x > 0.''7 \quad (5)$$

where  $\Delta x$  is the shift between *GALEX* and SDSS positions in arcsec. Correction is not applied when  $\Delta x \leq 0.''7$ , where random errors dominate. We apply the same correction to NUV and FUV magnitudes.

By far the most significant correction for UV magnitudes is due to blending. The degree to which a *GALEX* magnitude will be affected by blending will depend both on the relative brightness of the nearest companion (any SDSS photometric object) and its separation ( $d$ ) from the target. We construct corrections (same for NUV and FUV) as a function of the difference in  $g$  magnitude ( $\Delta g = g_{\text{comp}} - g_{\text{target}}$ ) for four ranges of separation:

$$\begin{aligned} \text{UV}_{\text{corr}} &= \text{UV} - 0.036 \Delta g + 0.37, & d < 4'' \\ \text{UV}_{\text{corr}} &= \text{UV} - 0.046 \Delta g + 0.35, & 4'' \leq d < 10'' \\ \text{UV}_{\text{corr}} &= \text{UV} - 0.019 \Delta g + 0.15, & 10'' \leq d < 15'' \\ \text{UV}_{\text{corr}} &= \text{UV} - 0.006 \Delta g + 0.04, & 15'' \leq d < 20'' \end{aligned} \quad (6)$$

We note that even though the corrections to UV magnitudes may be significant for individual objects (in particular those with a close blended companion), the SFRs for the majority of star-forming galaxies are not significantly affected by their application.

We use FUV fluxes measured at NUV positions (FUV\_NCAT) rather than from independent FUV detections, as they provide more robust UV color. Both NUV and FUV measurements require a  $3\sigma$  threshold in FLUX\_AUTO.

### 3.3. Galactic reddening and calibration errors

SDSS and *GALEX* photometry must be corrected for galactic reddening. We find that the combination of extinction coefficients from Peek & Schiminovich (2013) for UV bands and Yuan et al. (2013) for optical bands produces somewhat better fits ( $\sim 10\%$  smaller  $\chi^2$ ) than the corrections based on Schlegel et al. (1998), or on Yuan et al. (2013) alone. We reproduce the adopted corrections here.

$$\begin{aligned} A_{\text{FUV}} &= 10.47E(B-V) + 8.59E(B-V)^2 - 82.8E(B-V)^3 \\ A_{\text{NUV}} &= 8.36E(B-V) + 14.3E(B-V)^2 - 82.8E(B-V)^3 \\ A_u &= 4.39E(B-V) \\ A_g &= 3.30E(B-V) \\ A_r &= 2.31E(B-V) \\ A_i &= 1.71E(B-V) \\ A_z &= 1.29E(B-V) \end{aligned} \quad (7)$$

Photometry catalogs usually list only random flux errors, without systematic, or calibration errors. We find that the default CIGALE padding of formal photometry errors by the addition (in quadrature) of 0.1 mag to catalog magnitude errors to account for systematic errors in models and photometry leads to a significant loss of derived physical parameter accuracy, as evidenced by 50% larger scatter when derived SFRs are compared with independent SFR estimates. Instead, following S07, we add more modest calibration errors (determined from repeat observations) of (0.052, 0.026, 0.02, 0.01, 0.01, 0.01, 0.01) mag in (FUV, NUV,  $u$ ,  $g$ ,  $r$ ,  $i$ ,  $z$ ), plus the  $u$ -band red leak error of  $\sigma_{u,\text{RL}} = 0.0865(r-i)^2 + 0.0679(r-i)$ , derived in S07 based on the description of red leak in Abazajian et al. 2004. The validity of the adopted calibration errors is verified by constructing the distribution of magnitude residuals (fitted minus real magnitude divided by the total error), which follow unit Gaussians.

### 3.4. Mid-IR photometry and ancillary data

We derive what we refer to as “mid-IR SFRs” using two types of  $22 \mu\text{m}$  magnitude (flux) measurements. The first is what AllWISE Source Catalog calls a *profile-fitting* magnitude (w4mpro). The profile is simply the PSF, so these magnitudes are essentially PSF magnitudes, and are thus most appropriate for unresolved sources. Since the PSF FWHM at  $22 \mu\text{m}$  is  $12''$ , which is larger than 91% of galaxies ( $r_{90}$  size), these magnitudes are reasonably appropriate for SDSS galaxies. Yet, PSF magnitudes will systematically underestimate the flux in larger galaxies. This is largely remedied in forced photometry from unWISE catalog, which applies SDSS-measured galaxy profiles (convoluted with W4 PSF) as photometry priors.

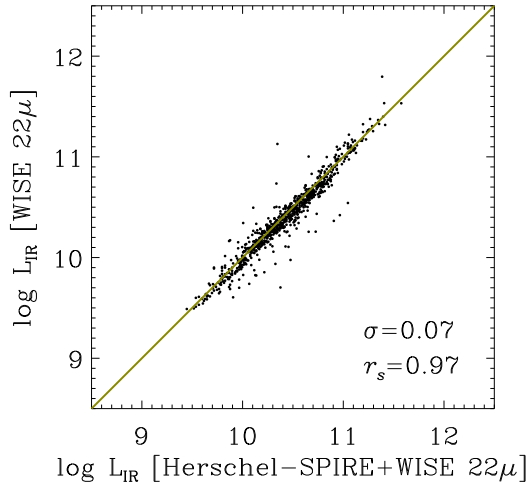
To derive emission-line SFRs (which we call “H $\alpha$  SFRs”), we use H $\alpha$  and H $\beta$  fluxes from the MPA/JHU catalog, based on DR7<sup>14</sup>. Measurement of emission line fluxes is described in Tremonti et al. (2004). We also use the emission-line (BPT diagram, Baldwin et al. 1981) classification of galaxies from the MPA/JHU catalog, derived as described in Brinchmann et al. (2004) (hereafter B04).

## 4. DERIVATION OF MID-IR AND EMISSION-LINE SFRS

Although CIGALE allows the IR dust emission to be fit in conjunction with the stellar emission, we derive the mid-IR SFRs separately, to avoid potential systematics. For example, CIGALE normalizes the IR SED template, so that the total dust emission (i.e., total IR luminosity,  $L_{\text{IR}}$ ) equals the stellar emission absorbed by the dust in the UV/optical/near-IR. However, the shape of the IR SED is a free parameter Noll et al. (2009), which means that without the far-IR SED, as in the case of *WISE* data, the IR luminosity will not be strongly constrained, and could thus be driven, through energy balance requirement, by any potential systematics in the estimate of the absorbed stellar luminosity, such as those that would arise from incorrect assumptions about the attenuation law.

Instead, accurate IR luminosities can be obtained from  $22 \mu\text{m}$  observations using *luminosity-dependent* IR templates, and these estimates can then be compared to the results of the SED fitting of the stellar emission (for actively star-forming galaxies, where dust heating by young stars dominates), to verify whether energy balance is satisfied.

<sup>14</sup> <http://www.mpa-garching.mpg.de/SDSS/DR7>  
<http://home.strw.leidenuniv.nl/~jarle/SDSS/>.



**Figure 3.** Comparison of IR luminosities derived from *WISE* 22  $\mu\text{m}$  flux, against *Herschel*+*WISE* IR luminosities from Rosario et al. (2016). Our *WISE* IR luminosity is based on a single flux point in the mid-IR, converted to total IR luminosity using the luminosity-dependent IR templates of Chary & Elbaz (2001), whereas the *Herschel*+*WISE* IR luminosity comes from 22  $\mu\text{m}$  and sub-mm (250–500  $\mu\text{m}$ ) measurements. The two estimates agree very well. Solid line is a 1:1 relation.

Another advantage of performing the stellar SED fitting separately is that in a joint fit to stellar and dust emission one also has to worry about the contamination of IR emission from dust-obscured Type 2 AGN (such as Seyfert 2, or high-excitation radio galaxies, Yan et al. 2013; Pace & Salim 2016). While the AGN contribution can be modeled in CIGALE, it introduces additional degrees of freedom, further weakening the usefulness of an energy balance requirement, in particular when only mid-IR photometry is available.

The mid-IR SFRs that we report in GSWLC are computed as follows. First we calculate the total IR luminosity (8–1000  $\mu\text{m}$ ) by interpolating the luminosity-dependent IR templates of Chary & Elbaz (2001) so that they match the 22  $\mu\text{m}$  flux. We also tested using Dale & Helou (2002) templates, which do not have an associated luminosity, but the IR SED shape-luminosity dependence is imposed from empirically-calibrated relations of Marcillac et al. (2006). The two methods produce very similar IR luminosities (average difference 0.01 dex, scatter 0.02 dex), but Chary & Elbaz (2001)-based luminosities are marginally (< 2%) better correlated with IR luminosities from *Herschel*, so we adopt them for GSWLC.

We check the robustness of our 22  $\mu\text{m}$ -derived IR luminosities by comparing them to IR luminosities derived for SDSS Stripe82 galaxies by Rosario et al. (2016). Their IR luminosity is obtained by fitting IR templates to *WISE* 22  $\mu\text{m}$  flux and the sub-mm fluxes from *Herschel* SPIRE (250, 350 and 500  $\mu\text{m}$ ). Because they have multiple flux points, they derive IR luminosities from Dale & Helou (2002) IR templates that do not have an associated luminosity. The cross-matched sample consists of 817 galaxies. As can be seen in Figure 3, the agreement of IR luminosities is excellent, with no (< 0.01 dex) systematic offset over the range probed by *Herschel* ( $9.5 < \log L_{\text{IR}} < 11.4$ ), perfect linearity, and a scatter of only 0.07 dex. Comparison of the specific luminosities ( $L_{\text{IR}}/M_*$ ), which removes the distance dependence (Section 7), reveals a mild non-linearity (with a slope of 1.10), in the sense that *WISE* IR luminosities of galaxies with the lowest specific luminosity (or SSFR) are 0.1 dex lower compared to galaxies with the highest specific luminosity. While the agree-

ment between our and Rosario et al. IR luminosities is encouraging, there is a caveat that the latter are not fully independent from our estimates, because both use 22  $\mu\text{m}$  flux. The role of 22  $\mu\text{m}$  flux point in Rosario et al. estimate is likely significant given that their data do not sample the peak of the IR SED ( $\sim 100\mu\text{m}$ ).

The above comparison was carried out for 22  $\mu\text{m}$  IR luminosities derived using profile (PSF) fluxes from the official AllWISE catalog. If we instead used 22  $\mu\text{m}$  photometry from unWISE (Lang et al. 2016), the small scatter with respect to Rosario et al. (2016) *Herschel*+*WISE* IR luminosities of  $\sigma = 0.07$  dex (Figure 3) would grow to  $\sigma = 0.14$ . This does not necessarily imply that unWISE fluxes are noisier, because the *Herschel*+*WISE* uses the PSF flux from AllWISE catalog.

To obtain mid-IR SFRs from IR luminosity, we use a simple conversion given by (Kennicutt 1998), adjusted to Chabrier IMF using the 1.58 conversion factor (S07):

$$\log \text{SFR} = \log L_{\text{IR}} - 9.966 \quad (8)$$

where SFR is in  $M_{\odot}\text{yr}^{-1}$  and  $L_{\text{IR}}$  in  $L_{\odot}$ .

To derive  $\text{H}\alpha$  SFRs, which we use for comparison purposes, we take  $\text{H}\alpha$  fluxes and correct them using the Balmer decrement method, assuming Case B recombination and the Cardelli et al. (1989) extinction curve. Emission line fluxes are measured within 3'' SDSS fibers, which we find to capture, on average, 30% of the galaxy mass, depending on galaxy's distance, size and profile. Dust-corrected  $\text{H}\alpha$  fluxes are then converted to SFR using the Kennicutt (1998) relation, adjusted by  $-0.24$  dex (Muzzin et al. 2010) to convert from Salpeter to Chabrier IMF. Comparison of fiber  $\text{H}\alpha$  SFR with other, total SFR indicators, can be accomplished if all measurements are converted to SSFRs. The  $\text{H}\alpha$  SSFR is obtained by normalizing by the stellar mass present within the fiber, which are taken from the MPA/JHU catalog.<sup>15</sup>

## 5. DERIVATION OF PHYSICAL PARAMETERS FROM UV/OPTICAL SED FITTING

### 5.1. Methodology

Physical parameters from UV/optical SEDs are derived using CIGALE (Noll et al. 2009, M. Boquien et al. in prep.), a powerful code that produces libraries (grids) of model SEDs and performs the SED fitting. The code was originally written in Fortran and recently re-written in Python. It is continually being improved and expanded in capabilities. We use the most recent Python version 0.9, which offers full control over parameters that specify SF history, dust attenuation, and emission line fluxes. This flexibility has proved essential in order to derive robust results.

Model SEDs produced by CIGALE, and from which model photometry is extracted, consist of UV/optical/near-IR stellar emission ( $\lambda \lesssim 5\mu\text{m}$ ), and, optionally, the IR dust emission ( $\lambda \gtrsim 5\mu\text{m}$ ). Third, also optional component is the nebular emission (lines and continuum). For reasons discussed in Section 4, we will restrict the SED fitting to stellar emission, with the contribution of emission lines included.

CIGALE synthesizes stellar emission based on either Bruzual & Charlot (2003) (BC03) or Maraston (2005) (M05) SPS. The most important difference between them is in the treatment of the TP-AGB phase of stellar evolution (Maraston 2011; Marigo 2015). CIGALE cannot currently include

<sup>15</sup> Stellar masses and SFRs from the MPA/JHU catalog have been adjusted from Kroupa to Chabrier IMF by applying  $-0.025$  dex offset (S07).

the contribution of emission lines in conjunction with M05 models, which, as we will demonstrate, are important for obtaining robust measurements. Therefore we opt to use BC03 models. The use of BC03 models may anyhow be a more appropriate choice, given some evidence that BC03 models better reproduce optical and near-IR colors than M05 models (Kriek et al. 2010; Conroy & Gunn 2010; Zibetti et al. 2013).

CIGALE can also perform Bayesian SED fitting and report the physical parameters and their errors. Bayesian SED fitting consist of building probability distribution functions of physical parameters by assigning probabilities to each model spectrum (at a matching redshift) based on the goodness of fit between the model and observed broad-band SEDs. The methodology has been described in detail in S07, da Cunha et al. (2008) and other papers. Walcher et al. (2011) and Conroy (2013) provide comprehensive reviews of various aspects of SED fitting. While SED fitting represents an optimal way to extract information from photometry (and/or spectra), it will be limited by the uncertainties in stellar and dust attenuation models (Conroy et al. 2009; Conroy & Gunn 2010; Conroy et al. 2010b; Mitchell et al. 2013), and the choice of priors (e.g., appropriate SF histories and dust attenuation laws).

In order to specify the modeling assumptions (e.g., which dust model or parameterization of SF history to use) the usual practice is to select the parameters that yield the smallest relative differences between the observed and the best-fitting model photometry, i.e., that minimize some average  $\chi^2$  of the sample. However, in some cases different modeling assumptions may lead to marginal changes in  $\chi^2$  (suggesting that UV/optical SED is degenerate with respect to them). We will therefore require both the internal consistency (small  $\chi^2$ ) and the external one: maximum correlation of the SED fitting SFR with respect to two entirely independently derived SFR tracers (mid-IR and  $H\alpha$ ).

In our analysis we assume a Chabrier IMF (Chabrier 2003). CIGALE assumes a flat WMAP7 cosmology ( $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$ ,  $\Omega_m = 0.27$ ).

### 5.2. Near-IR photometry

We perform SED fitting on UV and optical photometry, (0.15–0.9  $\mu\text{m}$ ), omitting the longer-wavelength stellar emission (1.1–5.4  $\mu\text{m}$ ). The near-IR photometry is excluded because it increases the reduced  $\chi^2$  values and worsens (albeit slightly) the correlations with mid-IR and  $H\alpha$ -derived SFRs. Parameters from the SED fitting are reliable to the extent that the models are able to reproduce the observed colors. We find that the model photometry, either based on BC03 or M05, does not accurately reproduce the near-IR colors from 2MASS photometry. In particular, both models imply a strong correlation between  $J-K$  and  $i-J$ , which is not seen in the data (S. Salim, in prep.).

Stellar mass estimates based on near-IR photometry are often perceived as more accurate and/or more precise than those based on optical bands, or that, at the minimum, the near-IR photometry improves the precision of stellar mass estimates. The reasoning underlying this claim is that stellar emission peaks in the near-IR and that near-IR mass-to-light ratio is less sensitive to the stellar population age, and the effects of dust. While these arguments are correct, they do not take into account large modeling discrepancies in the near-IR mass-to-light ratios (McGaugh & Schombert 2014), arising due to the uncertainties in our understanding of the post-main sequence phases of stellar evolution, the phases which dominate the en-

ergy output in the near-IR (Conroy 2013). For example, van der Wel et al. (2006) find that the inclusion of near-IR photometry in SED fits leads to discrepancies with respect to dynamical masses, and that these discrepancies depend on the SPS model used.

Even if there existed no major uncertainties in the models and no discrepancies between the models and the data (thus, no concerns that the near-IR will affect the accuracy of mass estimates), the improvement in the precision of stellar masses achieved by adding the near-IR photometry to the optical photometry is much more modest than usually assumed. Taylor et al. (2011) have studied physical parameters derived from mock observations of GAMA galaxies when just the optical photometry, or optical plus UKIDSS-depth near-IR photometry was used to perform the SED fitting. They show that the improvement in the precision of stellar masses when the near-IR photometry is added to SED fitting is small (0.05 vs. 0.06 dex, their Fig. A2). The parameter whose precision improves the most is the stellar metallicity. Indeed, the sensitivity of near-IR luminosities, and even red optical bands ( $i$  or  $z$ ), to metallicity (e.g., Fig. 9 of Courteau et al. 2014; Fig. 10 of Taylor et al. 2011) is another reason why near-IR alone is a more problematic tracer of stellar mass than the multi-band optical light.

### 5.3. Star formation histories

In the following three sections we discuss the choice of modeling parameters, starting with SF histories and metallicities.

BC03 models are available for six stellar metallicities, of which we use the higher four, from  $0.2 Z_\odot$  to  $2.5 Z_\odot$ , which is an adequate range for most galaxies in SDSS (Gallazzi et al. 2005).

We have considered two different parametrizations of SF history offered by CIGALE: two-component exponential and delayed exponential. In the two-component exponential model, the SF history is a composite of two exponentially declining functions ( $\tau$  models), each with its own starting epoch (age) and e-folding time. The normalization of the younger population is specified by the mass fraction  $f$ , which can also be set to zero, resulting in a single exponential model. Each component starts at maximum value and then decreases monotonically. In contrast, the delayed exponential SF history is smooth (e.g., Gavazzi et al. (2002)): it starts from zero SFR, reaches a peak at some time and then declines. It is given by:

$$\text{SFR} \propto \frac{t}{\tau^2} e^{-t/\tau} \quad (9)$$

We find that as long as the parameters are chosen so that the model colors (and more specifically, model SSFRs) cover the range of colors or SSFRs present in observations (for delayed SFHs this requires allowing  $\tau < 0$  in Eq. 9), either SF history will produce similar stellar masses and SFRs. Specifically, there is no systematic difference for masses of non-main sequence, low-SSFR galaxies ( $\log \text{SSFR} < -11$ ), while the difference in masses for actively star-forming galaxies is typically 0.1 dex. The difference in SFRs is similar in degree (see also Boquien et al. 2014), and of the same sign, which means that the difference in SSFRs is less than either the difference in mass or SFR. While both parametrizations perform reasonably well, the two-component exponential SF histories yield better fits (geometric mean of reduced  $\chi^2$  0.7 vs. 1.0),



presumably because, by not being smooth, they are able to better match the bursty SF history of low-mass galaxies [Weisz et al. \(2011\)](#). We thus adopt the two-component exponential parametrization.

Delayed exponential models have been preferred over exponential models in some recent studies (e.g., [Simha et al. 2014](#)). However, there are important differences among the exponential models. If a single exponential model is assumed to have started in the early universe (e.g.,  $t_0 \sim 10^{10}$  yr), in order to acknowledge the fact that galaxies contain ancient stellar populations no matter how dominant the current episode of SF may be ([Aloisi et al. 2007](#)), then such model will fail to reproduce the SSFRs of many star-forming galaxies today, simply because  $\log \text{SSFR}_{\text{max}} = \log t_0^{-1} = -10$ . It is this naive implementation of the exponential model (e.g., [Simha et al. 2014](#)) that results in inferior performance. The problem of single exponential model not producing high SSFRs can be alleviated by allowing the starting epoch of the model to be more recent than the Big Bang. Though later epoch galaxy formation is obviously not realistic, it can instead be interpreted as the epoch of peak of SF activity. Variable starting time was assumed in essentially all work that base their SF histories on a single exponential, including S07, which in addition had stochastic bursts superimposed on exponential models.

Allowing recent starting times for the exponential model results in SFRs that are comparable to more sophisticated models.<sup>16</sup> However, this approach will neglect (“outshine”) any old population, resulting in somewhat underestimated masses ([Papovich et al. 2001](#); [Michałowski et al. 2014](#)). This problem is resolved with a two-component exponential model used here, where one component corresponds to high-redshift SF. While more realistic SF histories may have multiple bursts, or are in general quite variable ([Weisz et al. 2011](#)), modeling these features with high temporal precision is unimportant from the standpoint of broad-band SED fitting, where even the characteristics of the most recent burst cannot be constrained with any precision because of the burst age vs. burst amplitude degeneracy (e.g., [Smith & Hayward 2015](#)).

The parameters of the adopted two-component exponential model are as follows. The formation time of the old population is 10 Gyr before observation epoch, with e-folding times that span a range from 850 Myr (fast decline that defines the lower SSFR limit of  $\log \text{SSFR} = -13.8$ ) to nearly constant 20 Gyr (which has declined only 0.2 dex since formation). Formation times for the younger component span from 100 Myr (the shortest timescale to which UV observations are sensitive) to 5 Gyr. Their e-folding time is 20 Gyr, i.e., nearly constant. In order to produce the highest SSFRs observed today ( $\log \text{SSFR} \sim -8$ ), the mass fraction of the younger component must span up to  $f = 0.5$ . Thus our library of SF histories looks like old exponentials with various decay times, with a relatively flat burst superimposed.

#### 5.4. Dust attenuation laws

CIGALE allows the dust attenuation law to be specified as a simple power law ( $k(\lambda)_{\text{PL}} \propto \lambda^\alpha$ , following [Charlot &](#)

<sup>16</sup> [Pacifci et al. \(2015\)](#) report large (0.6 dex, on average) offsets at  $z > 1$  between SFRs derived using their classical models (which assume exponential histories with a range of formation times) and their more sophisticated models (SF history based on cosmological simulations). However, the two models also assume different attenuation laws (fixed slope vs. multiple slopes), which will affect the SFRs through the derived dust attenuation (see their Figure 6.), so it is not clear that the difference in SF histories is responsible for most of the offset.

[Fall 2000](#)), or according to the [Calzetti et al. \(2000\)](#) recipe ( $k(\lambda)_{\text{Cal}}$ ). The Calzetti curve can be modified so that the overall slope becomes more or less steep, by multiplying it with the power law of slope  $\delta$  ([Noll et al. 2009](#)):

$$k(\lambda)_{\text{mod}} = k(\lambda)_{\text{Cal}}(\lambda/5500\text{\AA})^\delta$$

Note that slope modification changes  $R_V$  from 4.05, the value it has for the Calzetti curve. The attenuation law can be further modified to include a UV bump of varying strength ([Stecher 1965](#); [Fitzpatrick & Massa 1986](#); [Conroy et al. 2010a](#)).

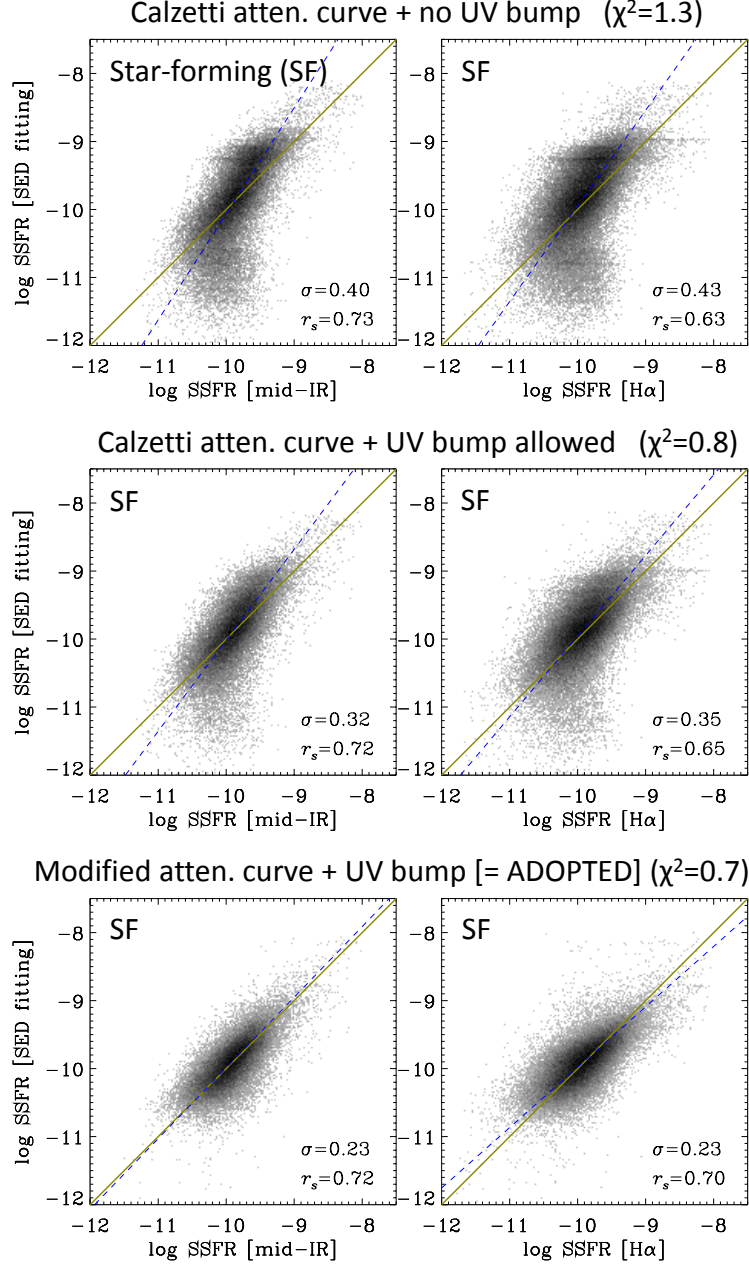
We systematically tested various modifications of the Calzetti attenuation law. We produce model grids where the range of color indices lies between  $E(B-V) = 0.1$  and 0.6 mag for the young ( $< 10$  Myr) population (i.e., the dust affecting the nebular lines; equivalent to  $A_{\text{H}\alpha} = 0.25\text{--}1.5$  mag). Attenuation that affects older stars, and therefore most of the stellar continuum, is taken to be smaller by a factor of 0.44, following [Calzetti et al. \(2000\)](#). The tests were carried out with galaxies classified in the BPT diagram as star-forming, for which mid-IR luminosity and dust-corrected  $\text{H}\alpha$  luminosity provide reliable estimates of SFR. The results are illustrated in Figure 4. The upper panels show the comparison between SSFR obtained from the SED fitting, assuming the standard Calzetti curve, which has no UV bump, against independently derived mid-IR and  $\text{H}\alpha$  SSFRs. Note that the comparison is carried out in terms of SSFR, which better reveals systematic trends than the usual SFR comparison (Section 7). Similar trends are seen vs. mid-IR and  $\text{H}\alpha$ . The differences increase with the increasing SSFR until  $\log \text{SSFR}_{\text{SED}} = -9$ , in the sense that the values derived from SED fitting assuming the Calzetti curve tend to be up to 0.4 dex higher. At still higher SSFRs however, there is reasonable agreement between the SSFRs, as expected, since the Calzetti curve was derived from starburst galaxies with high SSFRs.

The middle row shows the comparisons when the Calzetti law is modified to include a UV bump of varying intensity (from no bump to  $4\times$  the MW value). The UV/optical SED fits become formally better (geometric mean of reduced  $\chi^2$ s goes from 1.3 to 0.8), and the correlation with mid-IR and  $\text{H}\alpha$  SSFRs improves (from  $\sigma = 0.40$  to 0.32 dex, for mid-IR comparison). However, the non-linearity (the slope of the correlation featuring logarithms of SSFRs not being unity) persists, as well as a tail of galaxies with unusually low SSFRs.

The bottom row shows the comparisons when the Calzetti attenuation law is further modified to make it steeper (with  $\delta = -0.5$  and  $-1.0$ ), in addition to allowing the UV bump. There is now a significant reduction in scatter with respect to mid-IR and  $\text{H}\alpha$  SSFRs ( $\sigma = 0.23$ ), and the relation is fairly linear. This is our adopted dust attenuation model. We note that achieving this level of agreement between SED fitting SSFRs and mid-IR SSFRs requires both the steepening of the attenuation curve and the addition of the UV bump. Just steepening the slope of the attenuation curve does remove most of the non-linearity with respect to mid-IR SSFR (plot not shown), but still yields relatively high reduced  $\chi^2$  of 1.0, regardless of the amount of steepening. In other words, the combination of the two modifications is necessary to achieve both the small  $\chi^2$  values and good agreement with other indicators.

The agreement between SED and mid-IR SFRs demonstrates that the energy balance is fulfilled when this modified attenuation law is used: the energy absorbed by the dust in the stellar SED matches the energy emitted in the IR.

We note that assuming an attenuation curve in a power-law

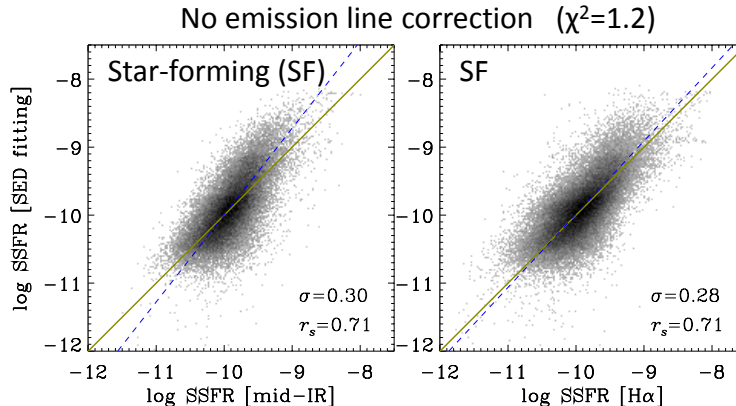


**Figure 4.** The choice of attenuation law to use in UV/optical SED fitting. We show comparison between specific SFRs (SSFR) obtained from the UV/optical SED fitting (assuming different attenuation laws) and SSFRs based on *WISE* 22  $\mu\text{m}$  observations (mid-IR SSFR; left panels) and from Balmer-decrement corrected  $\text{H}\alpha$  (right panels). The upper panels use the standard Calzetti attenuation curve. Strong trends against both SFR tracers are seen. They are reduced by allowing for a UV bump (the middle row), which also improves the formal  $\chi^2$  of SED fits. Further improvement, resulting in very good match with mid-IR and  $\text{H}\alpha$  SSFRs, is achieved by also making the attenuation curve steeper than the Calzetti one (bottom row). The changes in SSFRs are mostly driven by the changes in SFRs, rather than the stellar masses. Full line is a 1:1 relation, and the dashed line is the robust bisector linear fit (used throughout the paper). Standard deviation around the fit, as well as the Spearman correlation coefficient, are given in each panel. For these and subsequent figures (except where noted) the lower redshift bound is 0.025 in order to reduce AllWISE photometry systematics for galaxies having large angular size. In order to present a large number of data points without blotting, we use grayscale where the shade scales as number of galaxies in pixel to the power of 0.3. Sublinear exponent is chosen to better show the outliers. Shown are the data from GSWLC-M, but similar results are obtained with GSWLC-A or D.

form, having slopes  $\alpha = -1.0$  (preferred by the majority of galaxies) and  $\alpha = -1.5$ , and adding the UV bump to them, has a similar effect to modifying the standard Calzetti curve as described in the preceding paragraph. For comparison, the standard Calzetti curve can be approximated in the UV range by a power-law of slope  $\alpha = -0.5$ .

It must be stressed that the differences in SSFR obtained with various attenuation laws are largely driven by the changes in SFR, rather than  $M_*$ . The stellar masses of star-

forming galaxies obtained with the Calzetti dust attenuation law are on average only 0.06 dex lower than the stellar masses obtained with the modified law. However, the masses of a small number of individual galaxies, especially those with very high SSFR, can differ up to 0.4 dex in either direction, i.e., the dispersion of the two mass estimates increases with SSFR. For completeness, we report that for passive galaxies the use of the Calzetti dust law yields stellar masses that are 0.06 dex higher than the ones obtained with our modified at-



**Figure 5.** The effect on SSFRs of not correcting for nebular emission lines. This figure should be compared to the lower panels of Figure 4, where otherwise identical models include the contribution of emission lines. Not correcting for emission lines produces overestimates of SED (S)SFRs of high-SSFR galaxies, greater random errors (quantified as the standard deviation around the fit, blue dashed line) and also the increased mean reduced  $\chi^2$ .

attenuation law, and this difference is rather constant from one passive galaxy to another, i.e., it has a small dispersion of only 0.02 dex.

We conclude that while star-forming galaxies may be governed by a range of attenuation laws, they generally exhibit a steeper law than the standard Calzetti curve ( $\delta = -0.6$ , on average), and include a UV bump (1.1 times the MW value, on average). This average curve has  $R_V \approx 2.5$ , compared with  $R_V = 4.05$  for the Calzetti curve, and agrees well with the curve derived empirically by Conroy et al. (2010a): a MW-like curve with  $R_V = 2.0$  instead of  $R_V = 3.1$  (reducing  $R_V$  makes the curve steeper). Evidence for a steeper-than-Calzetti curve has been found in other studies as well (Charlot & Fall 2000; Buat et al. 2011; Wild et al. 2011; Hao et al. 2011; Salmon et al. 2015), though not as steep as the one found here and in Conroy et al. More detailed investigation of these results and their implications will be presented in a separate paper.

### 5.5. Correction for emission lines

Finally, we describe how we account for the flux from emission lines, which can have a significant effect on broadband fluxes and colors of galaxies with high equivalent widths (Papovich et al. 2001; Kauffmann et al. 2003a; Pacifici et al. 2015). CIGALE calculates the contribution of 124 lines, specified by metallicity (taken to equal the stellar metallicity of SPS models) and ionization parameter. We select the lowest ionization parameter available,  $\log U = -3$ , as it produces the biggest improvements and is closest to direct measurements (Dopita et al. 2000; Liang et al. 2006).

The formal quality of UV/optical fits is significantly improved by correcting for emission-line flux (geometric mean of  $\chi_{\text{red}}^2$  drops from 1.2 to 0.7). More importantly, accounting for emission lines significantly improves the correlation of SED SSFRs with other indicators. For galaxies with high SSFRs, not correcting for emission lines produces offsets of up to 0.5 dex, as can be seen by comparing Figure 5, which does not correct for emission lines, with the lower row of Figure 4. Emission line correction makes very little difference for stellar masses, which change by  $< 0.01$  dex on average.

While the addition of emission line to model fluxes results in overall improvement, the correction is not perfect. Nebular metallicity is fixed to stellar metallicity, which is not realistic. Also, the ionization parameter of many galaxies is lower than

the minimum ionization parameter available in CIGALE. We see some redshift-dependent trends in SFR, as lines straddle across the bandpasses. We account for this systematic offset ( $\sim 0.1$  dex on average) by deriving an average correction, in 0.01 wide bins of redshift, with respect to SFRs from B04 (using galaxies classified as star-forming; Section 8.2), and applying it to all SFRs from the SED fitting. We use B04 SFRs for this correction instead of mid-IR SFRs, because the latter are available for only 63% of star-forming galaxies. Nevertheless, the correction would have been essentially identical if mid-IR SFRs were used.

### 5.6. Derived parameters

In summary, for each 0.01-wide redshift bin, from 0.01 to 0.30, we calculate a grid of 342,720 models, a factor of three increase in the number of models and a factor of 5 increase in redshift resolution with respect to what was used in S07.

From the SED fitting we report the logarithm of the current stellar mass ( $M_*$ ), the logarithm of the SFR averaged over the last 100 Myr (the timescale for UV emission), and dust attenuations in FUV,  $B$  and  $V$  rest-frame bands. For each of these parameters CIGALE builds a probability distribution function (PDF). The nominal value of the reported parameter is the average of the PDF. We have performed mock fitting, in which true parameters are known, and find that the average of the PDF better retrieves the parameters than the median, and much better than the parameter corresponding to the best-fitting model. The best-fitting parameters are volatile and suffer from grid discreteness (e.g., Taylor et al. 2011). The formal error of the parameter is taken as the second moment (standard deviation) of the PDF. By default CIGALE applies a minimum error in the case the error from the PDF is lower than it, but we have disabled this adjustment. Instead, we caution the reader that the robustness of the error estimated from the PDF will depend on the goodness of the fit, i.e., it will be most accurate when  $\chi_r^2 \sim 1$ .

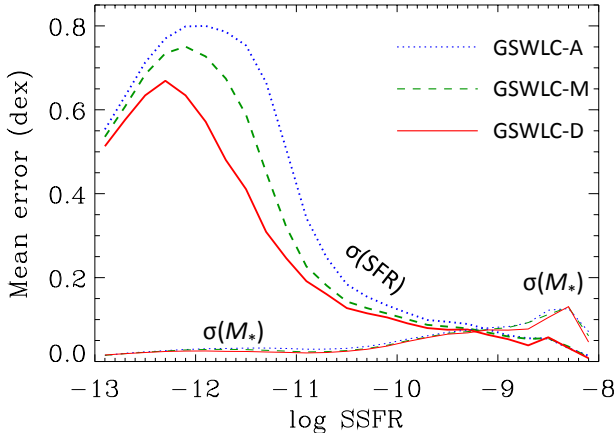
## 6. GSWLC

In Table 1, we describe the contents of the catalog which includes various IDs, coordinates, the SED fitting parameters (SFR,  $M_*$  and dust attenuations in several bands), mid-IR SFRs and flags. Flags describe cases when the SED parameters or the mid-IR SFRs are not listed, as we discuss further below.

**Table 1**  
Contents of *GALEX–SDSS–WISE* Legacy Catalog (GSWLC).

Column number	Column name	Units	Description
1	ObjID		SDSS photometric identification number
2	GLXID		<i>GALEX</i> photometric identification number
3	plate		SDSS spectroscopic plate number
4	MJD		SDSS spectroscopic plate date
5	fiber ID		SDSS spectroscopic fiber identification number
6	RA	deg	Right Ascension from SDSS
7	Decl.	deg	Declination from SDSS
8	$z$		Redshift from SDSS
9	$\chi_r^2$		Reduced goodness-of-fit value for the SED fitting
10	$\log M_*$	$M_\odot$	Stellar mass
11	$\sigma(\log M_*)$	$M_\odot$	Error of the stellar mass
12	$\log \text{SFR}_{\text{SED}}$	$M_\odot \text{yr}^{-1}$	UV/optical (SED) star formation rate
13	$\sigma(\log \text{SFR}_{\text{SED}})$	$M_\odot \text{yr}^{-1}$	Error of the SFR
14	$A_{\text{FUV}}$	mag	Dust attenuation in rest-frame FUV
15	$\sigma(A_{\text{FUV}})$	mag	Error of dust attenuation in FUV
16	$A_B$	mag	Dust attenuation in rest-frame $B$
17	$\sigma(A_B)$	mag	Error of dust attenuation in $B$
18	$A_V$	mag	Dust attenuation in rest-frame $V$
19	$\sigma(A_V)$	mag	Error of dust attenuation in $V$
20	flag_sed		SED fitting flag (0 = OK, 1 = broad-line spectrum, 2 = $\chi_r^2 > 30$ , 5 = missing SDSS photometry)
21	UV survey		1 = GSWLC-A, 2 = GSWLC-M, 3 = GSWLC-D
22	$\log \text{SFR}_{\text{mid-IR,AW}}$	$M_\odot \text{yr}^{-1}$	Mid-IR star formation rate from <i>WISE</i> (AllWISE catalog)
23	flag_wise		Mid-IR SFR (AllWISE) flag (0 = OK, 1 = no mid-IR SFR (low SSFR), 5 = no 22 $\mu\text{m}$ detection)
24	$\log \text{SFR}_{\text{mid-IR,uW}}$	$M_\odot \text{yr}^{-1}$	Mid-IR star formation rate from <i>WISE</i> (unWISE catalog)
25	flag_unwise		Mid-IR SFR (unWISE) flag (0 = OK, 1 = no mid-IR SFR (low SSFR), 5 = no 22 $\mu\text{m}$ detection)
26	flag_mgs		0 = not in SDSS Main Galaxy Sample (MGS), 1 = in MGS

Columns 10-19 originate from the SED fitting. If there are multiple reasons for setting the flag, the flag value will be the sum of individual flag values. When the SED (or un/wise) flag is set, the SED fitting parameters (or mid-IR SFR) are not given. Mid-SFRs based on unWISE are recommended over the AllWISE ones for  $z < 0.06$  samples, large ( $r > 10''$ ) galaxies, or studies that explore dependence of SFR on galaxy size or shape. SFRs and stellar masses are based on Chabrier IMF. Missing values are listed as -99.



**Figure 6.** Mean random error of SED fitting SFR (thick lines) and stellar mass (thin lines) as a function of SSFR and UV survey depth. The second moment of the parameter probability distribution function is adopted as the error of these parameters

To review, the criteria for inclusion in the catalog are for an SDSS object to be covered by *GALEX* observations of a certain depth, regardless of whether it was detected in the UV. Furthermore, galaxies need to lie in  $0.01 < z < 0.30$  redshift range and be above  $r_{\text{petro}} = 18.0$  magnitude threshold.

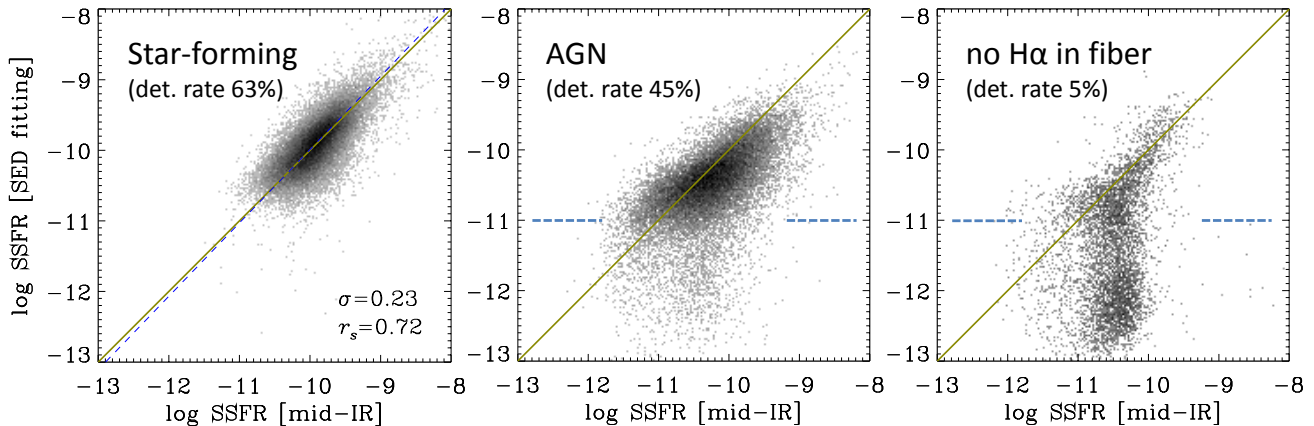
Approximately 1% of SDSS targets have spectra classified as “quasars” (CLASS = 1), based on the Balmer emission lines being broad. These galaxies are therefore likely to be Type 1 AGN, which includes Seyfert 1s, and, less frequently, true quasars. In Type 1 AGN the accretion disk is not entirely obscured and can contribute to the UV/optical continuum of

the host, which may significantly bias the derived SED parameters, especially the SFR. This is confirmed by the finding that the geometric mean of the reduced  $\chi^2$  values of the best-fitting model of broad-emission line galaxies is 4 times higher than of other galaxies. We flag such objects in the final catalog and do not report their SED-fitting parameters. Type 1 AGN are not obscured by dusty torus, which is why they mostly contribute to the UV/optical SED. We thus leave their mid-IR SFRs in the catalog, but they should still be used with caution in case there is some AGN emission in the mid IR.

We produce three separate catalogs, GSWLC-A, M, and D, according to the UV depth, as described in Section 2.2. In addition, we produce a master catalog that combines the three catalogs, keeping for each galaxy the data from the deepest catalog. This catalog is designated GSWLC-X, and contains 658,911 objects. Readers are cautioned that non-uniform depth may lead to systematics in SFRs, especially for galaxies with  $\log \text{SSFR}_{\text{SED}} < -11$  where the UV detection rate and the resulting quality differ from survey to survey (Figure 6). On the other hand, the stellar masses will not be subject to any such biases. All catalogs have the same format as described in Table 1.

The catalogs are currently publicly available at <http://pages.iu.edu/~salims/gswlc>, which will document any changes implemented in subsequent versions. The catalog letter designation is followed by a version number. The analyses in this paper are all based on Version 1 catalogs (e.g., GSWLC-M1). GSWLC is also hosted at Mikulski Archive for Space Telescopes (MAST) and may be included in SDSS SciServer and/or hosted by NED/IPAC in the future.

In Figure 6 we show mean random errors of  $\log \text{SFR}_{\text{SED}}$  and



**Figure 7.** Comparison of UV/optical (SED fitting) and mid-IR SSFRs. Mid-IR SSFRs are derived from the  $22\ \mu\text{m}$  mid-IR photometry from *WISE* (AllWISE catalog). Galaxies are split into SF and AGN classes according to the position in the BPT emission line diagram. The AGN category includes all galaxies on the AGN branch (i.e., including the galaxies that B04 calls SF/AGN composites and LINERS). Galaxies with lines too weak to allow classification are presented separately in the right panel. The left panel plots SF galaxies and shows the best fitting line along with its scatter and correlation coefficient. Galaxies below dashed lines in middle and right panels have greatly overestimated mid-IR SFRs because their dust is primarily heated by old stars, invalidating the fixed  $L_{\text{IR}}$ -SFR conversion. Their mid-IR SFRs have been removed from GSWLC. Both SSFRs use the stellar mass from our SED fitting. Detection rate at  $22\ \mu\text{m}$  of a given class of galaxy is given in each panel. Data in this and subsequent figures is from GSWLC-M.

log  $M_*$  as a function of SSFR, for each of the three catalogs GSWLC-A, M, D. Errors in SFR depend very strongly on the SSFR, and for more passive galaxies also on the UV depth. The decrease of errors at lowest and highest SSFRs is an artifact of reaching the boundary of the model grid. For passive galaxies the formal SFR errors range between 0.65 and 0.80 dex. However, the SFR error of passive galaxies that have no current SF whatsoever ( $\log \text{SSFR} = -\infty$ ) is infinity (in log). SFR values for galaxies with  $\log \text{SSFR} < -11.7$  ( $< -11.5$  for GSWLC-A;  $< -12.0$  for GSWLC-D), should be considered upper limits. For actively star-forming galaxies the SFR errors are typically below 0.1 dex, a 50% improvement over S07. Errors on stellar mass are typically much lower than the SFR errors, with the opposite dependence on SSFR, and little difference between the UV surveys. They range from 0.03 dex for the passive galaxies, to 0.10 dex for the most active ones.

The geometric mean of the reduced  $\chi^2$  values for the GSWLC-M catalog is 0.7. However, the tail extends to  $\chi_r^2$  values in excess of 100. Instead of introducing an arbitrary cut to exclude poorly fit objects, we test the robustness of the derived SFRs and stellar masses as a function of the reduced  $\chi^2$ . We find that no systematic differences in excess of 0.1 dex (with respect to mid-IR SFRs and stellar masses from B04) are present when  $\chi_r^2 \leq 30$ . For those galaxies we retain the physical parameters from the SED fitting, and annul them if  $\chi_r^2 > 30$  (0.7%). Nevertheless, the readers are advised to treat galaxies with  $\chi_r^2 \geq 5$  with caution, especially for individual galaxies.

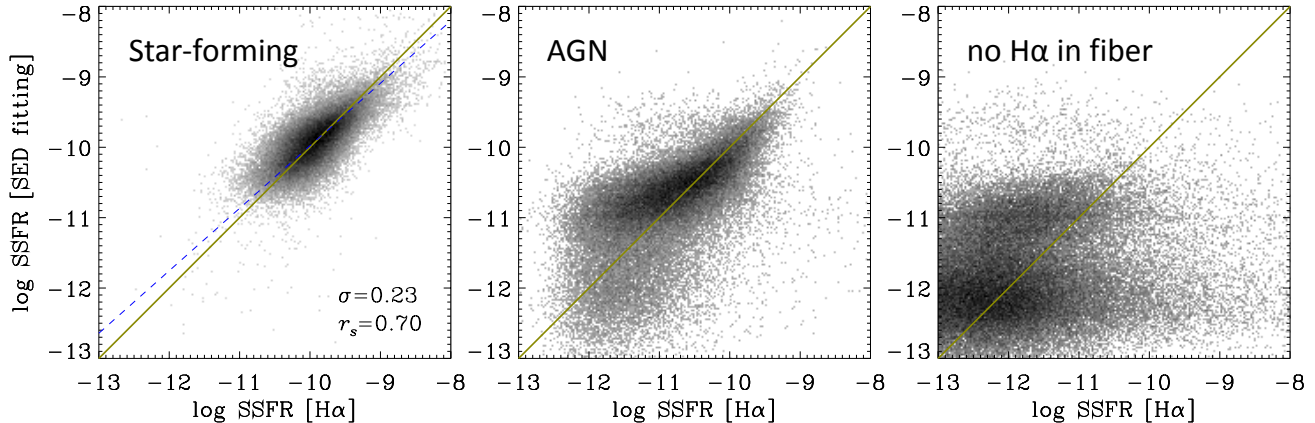
#### 7. COMPARISON BETWEEN SSFRs FROM THE SED FITTING, MID-IR, AND H $\alpha$ EMISSION

In this section we discuss comparisons between various SFRs derived in this work: specifically, the UV/optical SSFRs from the SED fitting on the one side, and mid-IR and H $\alpha$  SSFRs on the other. We remind the reader that the SED (S)SFRs are averaged over the preceding 100 Myr, the UV emission timescale.

The paper breaks away from the usual practice of comparing absolute SFRs, and instead performs the comparison in terms of the specific SFRs. The standard SFR comparison has its merits and may be more intuitive, however, when the

focus is on galaxies with similar stellar populations, subject to similar systematics, the SSFR comparison has its advantages: (1) SDSS probes a large dynamic range of distances, and therefore of luminosities, which means that SFRs (and stellar masses) will be correlated even if SFR/ $L$  ( $M_*/L$ ) are not. This “trivial” dependence is eliminated in SSFR. For example, the correlation coefficient between SED and mid-IR SFRs (of galaxies classified as star-forming), is  $r_s = 0.86$ , while it is  $r_s = 0.72$  for comparison involving SSFRs. Consequently, the SFR comparison is seemingly tighter. (2) The comparison of SSFRs is more informative because it contrasts galaxies which are physically similar. For example, a galaxy with  $\text{SFR} = 10M_\odot\text{yr}^{-1}$  can be a massive galaxy on the star-forming sequence, or a lower-mass galaxy experiencing a starburst. However, their SSFRs will be different. Or, a galaxy with  $\text{SFR} = 0.01M_\odot\text{yr}^{-1}$  can be a blue star-forming dwarf, or a massive red galaxy, but their SSFRs will differ by two orders of magnitude. Finally, (3) the comparison in terms of SSFR allows for a meaningful comparison between total and fiber SF, as long as the fiber SFR is normalized by the stellar mass within the fiber.

In Figure 7 we present the comparison of UV/optical SSFRs (i.e., SED fitting SSFRs) against mid-IR SSFRs from *WISE*  $22\ \mu\text{m}$  photometry from AllWISE. We show the comparisons using medium UV depth catalog (GSWLC-M), which balances the extensiveness of GSWLC-A with the depth of GSWLC-D. All of the findings hold for the more accurate GSWLC-D as well. We refrain from using GSWLC-X, because of its non-uniform UV depth. The comparisons are presented for galaxies split into three categories using the BPT diagram (*iclass* in MPA/JHU catalog). The star-forming class (*iclass* = 1) includes galaxies with well-measured BPT lines and lying below the Kauffmann et al. (2003b) empirical demarkation between galaxies with and without AGN contribution. As we have already seen in Section 5), once we have adopted an adequate attenuation law the comparison is quite linear (slope in log-log plot equals 1), with no significant systematics. Mid-IR SSFRs from *WISE* are available for 63% of galaxies in this class. The scatter around the bisector linear fit (linear in log) is 0.23 dex, to which the measurement errors of SED SFR and mid-IR SFR contribute approximately



**Figure 8.** Comparison of UV/optical (SED fitting) specific SFRs against SSFRs from dust-corrected  $H\alpha$  emission. SED fitting SSFRs are integrated (total), while  $H\alpha$  SSFRs are measured within SDSS fiber, covering on average 30% of galaxy mass. Galaxies are split using the BPT diagram as in Figure 7. The left panel plots SF galaxies and shows the best fitting line along with its scatter and correlation coefficient.  $H\alpha$  SSFR is normalized by stellar mass within the fiber.

equally (0.15 dex each). Note that in all of the figures where we show the linear fit, we use a robust (outlier-resistant) bisector least square fit, which treats the variables symmetrically, i.e., no variable is considered independent.

Had we used mid-IR SSFRs based on unWISE photometry, the comparison would look similar, but would have 5% greater scatter. We find a similar level of degradation with respect to SSFRs based on  $H\alpha$ . We have not positively identified the cause of the increased scatter of SSFRs based on unWISE compared to AllWISE photometry, but list two possible reasons. unWISE uses  $r$ -band profiles to extract photometry. Optical light in  $r$  band is dominated by the emission from older stellar populations, so it may not represent an optimal prior for  $22\ \mu\text{m}$  observations, where the bulge emission is suppressed, while the star-forming regions in the disk dominate. Another possibility is related to the fact that unWISE is based on `cmodeImag` profiles, whereas all other SFRs are based on `modelmag` SDSS photometry. Despite being somewhat noisier in comparisons with other (S)SFRs overall, unWISE IR luminosities and SFRs are less biased for large ( $> 10''$ ) galaxies (where the AllWISE PSF magnitudes can underestimate the SFR by up to 0.1 dex), so we retain them in GSWLC alongside SFRs from AllWISE catalog. Our recommendation is to use unWISE SFRs in studies that explore dependence on galaxy size or profile (e.g., Sersic index) and also in studies that focus on galaxies with large angular diameter. Otherwise, AllWISE SFRs are recommended, and we continue to focus on them in subsequent discussion.

The middle panel shows the comparison for AGN-hosting galaxies. This category includes all galaxies above the Kauffmann et al. (2003b) line, i.e., what B04 call SF/AGN composites, AGN (Seyferts) and LINERs (`iclass = 3, 4, 5`). AllWISE detects 45% of such galaxies. First, we notice that no correlation is present in the lower part of the plot, approximately when  $\log \text{SSFR}_{\text{SED}} < -11$ . These are nearly quiescent galaxies of LINER type (thus, potentially passive galaxies with non-AGN emission lines Stasińska et al. 2008). For such galaxies the IR emission, if detected, will be dominated by dust heating by relatively dusty old or intermediate-age (e.g. AGB) stellar populations (Bressan et al. 2001; Villaume et al. 2015), and will therefore not be indicative of the current SF (Salim et al. 2009; Cortese et al. 2008). When the IR luminosity of such galaxies is converted into SFRs using simple, fixed-coefficient formulae (Eq. 8), which assume dust is heated by young stars, the SFR will be overestimated

(Buat & Xu 1996; Kennicutt 1998; Boquien et al. 2016). For AGN galaxies with higher SSFRs, the mid-IR SSFRs tend to be up to 0.6 dex higher than the SED SSFRs, suggesting a non-negligible dust heating by AGN, that produces excess emission in the mid-IR. The excess appears to be greater for increasing SSFRs, which would be expected if the gas both fuels the SF and drives the AGN accretion (e.g., Kewley et al. 2006).

Classification of a galaxy as a star-former or AGN, as performed by B04, requires minimum adjusted SNR of 3 in  $H\alpha$  and other lines ( $\approx 7.4$  in raw  $H\alpha$  SNR). Galaxies that are too weak to allow classification, what we call “no  $H\alpha$ ” category (`iclass = -1`), are shown in the right panel. These galaxies are mostly quiescent, so it is not surprising that the AllWISE detection rate is only 5%. When detected, the majority of galaxies in this class have low SED SSFRs, as expected. However, mid-IR SSFRs can be too high by up to 2 dex, because the IR emission from old stars is interpreted as current SF. We conclude that mid-IR SFRs are not reliable for quiescent or nearly quiescent galaxies when obtained through simple recipes that have a fixed conversion factor between IR luminosity and SFR (Eq. 8), and therefore in GSWLC we remove mid-IR SFRs for galaxies for which  $\log \text{SSFR}_{\text{SED}} < -11$  (dashed line in middle and right panels of Figure 7).

Interestingly, there are some galaxies in the “no  $H\alpha$ ” class for which the SED fitting SSFRs are quite high, overlapping with the values of normal star-forming galaxies ( $\log \text{SSFR} > -10.5$ ). It may seem contradictory that a galaxy with essentially no  $H\alpha$  emission could have such high SSFR. There are two possible reasons for this: (1) Difference in SF timescales.  $H\alpha$  is not present because a galaxy is in a post-starburst phase – UV and mid-IR emission will persist after the O-stars that give rise to  $H\alpha$  emission have died off. In other words, high SED and mid-IR SSFRs reflect recent, but not instantaneous values. The post-starburst galaxies would form the tail of high SSFRs in Figure 7 (right;  $\log \text{SSFR} > -10$ ). Their SED and mid-IR SSFRs agree, with no mid-IR excess like the one seen for AGNs. This suggests that post-starburst galaxies, which are usually considered to be the results of mergers (e.g., Yang et al. 2008), nonetheless do not have a significant AGN emission. (2)  $H\alpha$  is not present within the SDSS fiber, but the SF (and presumably the  $H\alpha$  emission) is present outside of it. Salim et al. (2012) and Fang et al. (2012) have studied such population in detail with high-resolution UV, optical and  $H\alpha$  imaging, and have confirmed that these are typically lenticu-

lar (S0) galaxies with no SF in the bulge, but with low-level SF present in a ring outside of SDSS fiber. The SSFRs of such star-forming S0s would be in the range  $-12 < \log \text{SSFR} < -10$ , i.e., lower than post-starburst galaxies.

Figure 8 shows equivalent comparison of SED SSFRs with respect to  $H\alpha$  SSFRs.  $H\alpha$  SSFRs are measured within the  $3''$  SDSS fiber, which covers between 17% and 50% of the galaxy’s stellar mass. The comparison of star-forming galaxies (left panel) shows similar scatter as that seen with respect to the mid-IR. The relation is somewhat SSFR dependent. One expects the relation between total and fiber SSFR to be linear (slope in log-log plot equal to 1) only if there are no SSFR gradients. Somewhat increased nonlinearity for galaxies with  $\log \text{SSFR}_{H\alpha} \geq -9$  suggests either that the SF tends to be centrally concentrated so that it produces higher SSFRs within the fiber compared to the integrated SSFR, or, that the fiber mass from the MPA/JHU catalog, which we use to normalize the  $H\alpha$  SFR, is systematically underestimated for bursty galaxies (Section 8.1).

For galaxies with AGN/LINER contribution (middle panel), there are no systematic differences for hosts having high SF ( $\log \text{SSFR} > -10$ ). This is in contrast to mid-IR SSFRs that showed an excess in this regime.  $H\alpha$  SSFRs start to display systematic discrepancies at lower SSFRs, in the sense that  $H\alpha$  SSFRs are up to 0.7 dex too low. We find that many of such cases are galaxies with no optical signs of SF, either photometrically (optical colors are red) or spectroscopically (not much  $H\alpha$ ), and yet the UV is clearly indicative of SF—often the UV colors are noticeably blue and uniformly spread across the disk. These galaxies may be in a declining phase of SF where the UV emission is higher than the rapidly diminishing  $H\alpha$ .

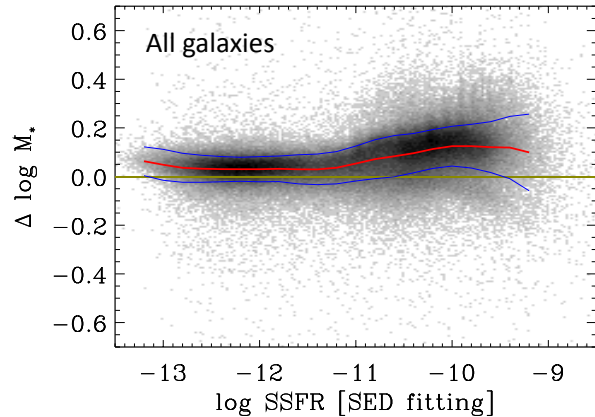
For galaxies with “no  $H\alpha$ ” (right panel), no correlation is present, as expected given the weakness of  $H\alpha$  that defines this category and making  $H\alpha$  SSFRs essentially meaningless. There is an overdensity of galaxies below  $\log \text{SSFR}_{\text{SED}} = -12$ , where SED SFRs estimates also become rather uncertain (Figure 6) and are more likely just the upper limits of galaxies with no current SF whatsoever, such as the majority of early-type galaxies.

To conclude, for actively star-forming galaxies all three indicators (SED fitting, mid-IR, and  $H\alpha$ ) provide robust measurements of (S)SFR. For mid-IR, the depth of *WISE* allows only 2/3 of SDSS galaxies to be detected.  $H\alpha$  (S)SFRs pertain only to the fiber. Mid-IR luminosity breaks down as a SFR indicator for galaxies with  $\log \text{SSFR}_{\text{SED}} < -11$ , where dust heating is dominated by old populations. This corresponds to green-valley and quiescent galaxies. Furthermore, the mid-IR SFR appears to be affected by AGN emission, limiting its usefulness. SED SSFRs are measurements of choice in low-SSFR regime, but even they become essentially upper limits when  $\log \text{SSFR}_{\text{SED}} < -11.7$  (for GSWLC-M;  $-12.0$  for GSWLC-D;  $-11.5$  for GSWLC-A).

## 8. COMPARISON OF GSWLC WITH PREVIOUS CATALOGS

In this section we compare GSWLC stellar masses and SED fitting SFRs (primarily through SSFRs) with those from previously published catalogs. The focus will be on (S)SFR comparisons, which are more sensitive than stellar mass measurements. We continue to present the comparisons using GSWLC-M, but the results hold with GSWLC-A and GSWLC-D.

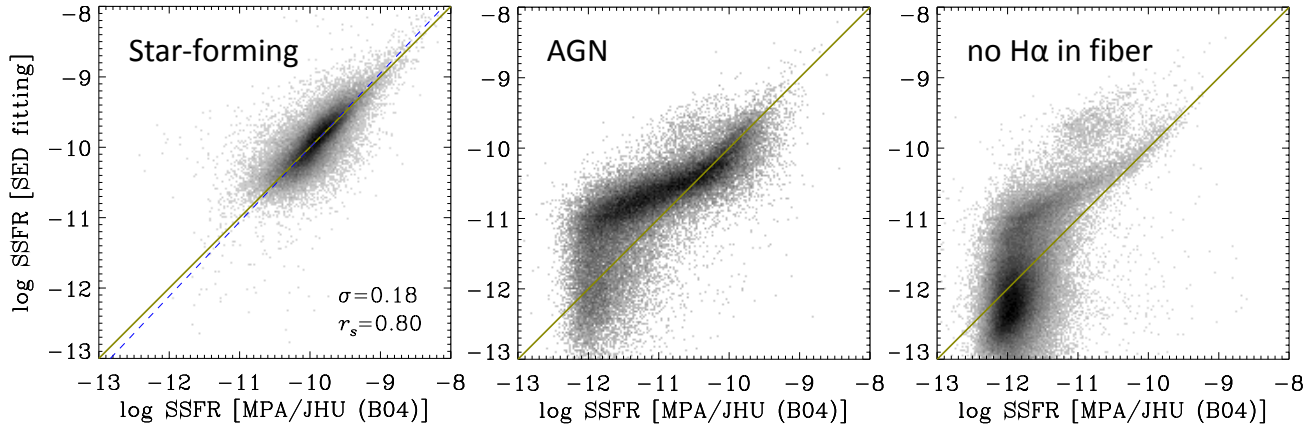
### 8.1. Comparison with MPA/JHU stellar masses



**Figure 9.** Comparison of stellar masses from GSWLC-M and MPA/JHU (DR7) catalogs, presented as the difference in masses as a function of SSFR. Red curve are the averages of the difference in 0.2 dex wide bins, and the blue curves are the averages  $\pm 1\sigma$ . GSWLC masses are on average higher than MPA/JHU catalog masses because of the differences in the assumed SF histories, but the difference is quite small (0.03–0.13 dex), and is SSFR dependent. The standard deviation of the mass difference is typically 0.07 dex.

The MPA/JHU catalog is the most widely used source of stellar masses for SDSS galaxies. Prior to DR7, the most recent version of this catalog, the MPA/JHU catalog listed stellar masses derived in Kauffmann et al. (2003a), using a method that combines photometric information to assess dust attenuation, and spectroscopic indices to constrain the SF history and mass-to-light ratio. Spectroscopic indices were measured within the fiber, so the method assumed no  $M/L$  gradients. In the DR7 version of MPA/JHU catalog, the method from Kauffmann et al. was replaced with the Bayesian SED fitting method, using only the optical broad-band photometry from SDSS (`modelMag`) and models described in S07. Here we will only present comparisons with respect to DR7 version of the MPA/JHU catalog.

As described in Section 5, GSWLC uses different specification of SF histories and dust extinction from those used in S07 and the MPA/JHU catalog. Furthermore, unlike the MPA/JHU catalog (but, like S07), GSWLC uses the constraints offered by UV photometry. Figure 9 shows the difference between GSWLC and MPA/JHU stellar masses as a function of SSFR. Again, we prefer the presentation of the results in this, relative, distance-independent way than the usual mass vs. mass comparison. On average, GSWLC stellar masses are somewhat higher than the MPA/JHU ones, with typical difference being 0.03 dex for passive galaxies, and  $\sim 0.13$  dex for the active ones. The scatter of the difference in mass, 0.07 dex, is consistent with the formal estimates of the mass error (Figure 6). The inclusion of the UV photometry is responsible for 0.04 dex increase for active galaxies. The remaining 0.09 dex increase for active galaxies (and 0.03 dex increase for passive ones) is mostly due to our use of two-component exponential SF histories. We confirm that the difference would not have been present if we had used the *delayed* exponential histories instead (Eq. 9). The critical difference between the delayed exponential and our implementation of the two-component exponential SF history is that in the latter the old component starts in the early universe, whereas the delayed exponential, having a single peak, will be shifted towards later epochs. Recently, Sorba & Sawicki (2015) have



**Figure 10.** Comparison of SSFRs from the SED fitting against aperture-corrected (total) SSFRs from the MPA/JHU DR7 catalog, derived based on a modified Brinchmann et al. (2004) method. Galaxies are classified using the BPT diagram, as in Figure 7. SF class shows the best fitting line and its scatter and correlation coefficient. Agreement is excellent for SF class, but systematics are present for low SSFR AGN and “no H $\alpha$ ” galaxies for which MPA/JHU SFRs are based on indirect methods that are not sensitive to low levels of SF.

reported that the masses of nearby high-SSFR galaxies obtained by summing up the masses in individual “pixels” are up to 0.1 dex higher than the masses from integrated light. They proposed that the difference arose from “outshining” of the old, fainter populations in galaxies with younger populations (high SSFR). It appears that our use of the two-component exponential SF history, in which the old component is set to have started in the early universe, may have recovered this deficit.

More recently, stellar masses for SDSS galaxies were published by Mendel et al. (2014) and Chang et al. (2015). The differences of GSWLC masses with respect to these masses follow the trends we have shown with respect to masses from the MPA/JHU catalog: little difference for passive galaxies, and up to 0.2 dex difference for galaxies with high SSFR.

Comparison of our masses obtained with and without the UV photometry reveals that in the case when there is another photometric source within 2–3”, the masses obtained from joint UV and optical photometry will be biased upward by 0.05 dex, presumably because of the unaccounted blending in the UV. This affects only a few percent of all sources.

### 8.2. Comparison with MPA/JHU star formation rates

The MPA/JHU catalog also provides SFRs, and until recently it was the only publicly available source of SFRs for SDSS galaxies. SFRs derived in MPA/JHU follow the method of B04, with modifications introduced in the most recent (DR7) version of the catalog. We first describe both the original method and its modifications, but the comparison will be presented only for the DR7 version of MPA/JHU catalog.

MPA/JHU catalog (and the original B04) SFRs are often described as H $\alpha$ , or emission-line SFRs. This is accurate only for the portion of a galaxy contained within the fiber, and only for galaxies classified as star-forming using the BPT diagram, for which the AGN contribution to emission lines should be negligible. For galaxies with AGN contribution or having weak lines (altogether 78% of SDSS galaxies), B04 derive SFR in the fiber based on a relation between the emission-line SSFR and D4000 index, constructed from star-forming galaxies. Next, the SFR estimate within the fiber is aperture-corrected to produce the total (integrated) SFR. B04 performs this correction by first establishing the relationship between fiber SSFR and the fiber broad-band colors of star-forming galaxies, and then applying these relations in a

Bayesian fashion to the light outside of the fibers to arrive at the out-of-fiber SFR. The total SFR is then obtained as the sum of fiber and out-of-fiber SFRs. The DR7 MPA/JHU catalog modifies the procedure for deriving out-of-fiber SFRs, by instead performing the SED fitting to *ugriz* photometry, using the models and methods described in S07. Altogether, the SFR method of MPA/JHU catalog (and original B04) is an emission-line/D4000/SED fitting hybrid. The temporal sensitivity of such SFRs will be between the  $\sim 10$  Myr timescales traced by the emission lines and  $\sim 1$  Gyr timescale for *u*-band light.

S07 presented detailed comparison of their SED SFRs and the original SFRs from B04. GSWLC contains many improvements over the S07 methodology, as does the MPA/JHU catalog with respect to B04. The comparison of GSWLC-M and DR7 MPA/JHU total SSFRs is given in Figure 10. It uses the same division of galaxies into SF, AGN and “no H $\alpha$ ” categories as employed earlier. In order to decouple SFR and mass systematics, both SSFRs are obtained by normalizing by the same mass (from GSWLC). Note that for 3% of galaxies in MPA/JHU catalog the total SFRs are present even when masses are not. We find that in those cases the reported total SFR is incorrect, and assumes the value of *fiber* SFR. We exclude those values from the comparison.

SSFRs for star-forming galaxies (left panel) agree very well, with a scatter of  $\sigma = 0.18$ , which is smaller than the scatter of SED SSFRs with respect to mid-IR or H $\alpha$  SSFRs (Figure 7 and 8). The comparison for AGN hosts reveals some systematic discrepancies, especially for intermediate SSFRs, which can in some cases reach  $\sim 1$  dex. We have already seen similar, but smaller offsets in comparison with H $\alpha$  *fiber* SSFRs, which we attributed to optically (spectroscopically and photometrically) inconspicuous SF. If we intentionally leave out the UV bands from our SED fitting, the resulting (ill-constrained) SSFRs are drawn to lower values (because the majority of models with red optical colors has low SSFRs), in better agreement with B04 values, which derive the greater part of their SFR (the out-of-fiber portion) from similar optical-only SED fitting.

Similar trends are present in galaxies with little or no H $\alpha$  emission in the fiber (right panel), except that the majority of galaxies have low SSFRs ( $\log \text{SSFR}_{\text{SED}} < -11$ ). In this class of galaxies there exists a peculiar feature: a cloud of galaxies lying  $\sim 1$  dex above the 1:1 line, having high SED SSFRs ( $\log$



$\text{SSFR}_{\text{SED}} > -10$ ), but much lower MPA/JHU SSFRs. We confirm that these galaxies have robust SED fits and are UV detected. In Section 7 we mentioned that “no  $\text{H}\alpha$ ” galaxies with high SSFRs are probably post-starburst galaxies. Such galaxies are traditionally identified as E+A galaxies, based on their Balmer absorption features (Dressler & Gunn 1983). Now we inspect the spectra of  $\sim 200$  of these outliers and confirm the presence of Balmer absorption lines. Visually (*gri* composites), these galaxies appear like red early-type galaxies, often with white centers, suggesting a central (post) starburst. We match our sample to Goto (2007) catalog of E+A galaxies (online version updated with SDSS DR7 data) and confirm that the majority of their E+As are found in this region of the plot. Why do the two methods yield discrepant SFRs for these galaxies? We find that the SED fitting (S)SFRs of E+A galaxies are very sensitive to the assumed dust attenuation law, more so than the normal star-forming galaxies. The (S)SFRs that we obtain for E+As assuming the modified attenuation law are an order of magnitude higher than what would be derived using the nominal Calzetti law. The reduced  $\chi^2$ s are 5 times lower when the modified attenuation law is used. From this we conclude that the high recent (S)SFRs of E+As that we derive in GSWLC are more likely to represent true levels of SF averaged over the last 100 Myr.

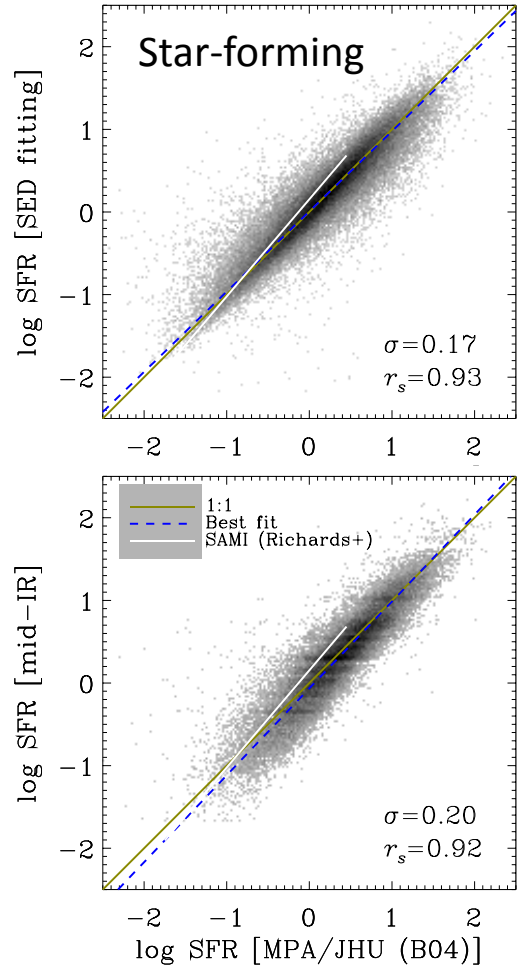
### 8.2.1. On the systematics of MPA/JHU SFRs reported by SAMI Galaxy Survey

Recent campaigns to obtain resolved spectra with integral-field spectroscopy have reported systematic differences with respect to SFRs from the MPA/JHU catalog. In particular, Richards et al. (2016), using preliminary data from the SAMI Galaxy Survey (Allen et al. 2015), find a non-linear, i.e. SFR-dependent, relation between their and MPA/JHU SFRs (which they refer to as B04 SFRs), in the sense that galaxies with high SAMI SFRs have underestimated MPA/JHU SFRs. The discrepancy is already  $\sim 0.3$  dex at  $\log \text{SFR}_{\text{SAMI}} \approx 0.6$ , the highest SFRs in their sample (we convert all SFRs from Richards et al. to Chabrier IMF). Richards et al. attribute the discrepancy to possible biases in B04 methodology for deriving aperture corrections.

In our analysis so far we have shown that GSWLC *specific* SFRs of star-forming galaxies have no significant systematics with respect to either mid-IR SSFRs or SSFRs from the MPA/JHU catalog. This implies, and we confirm it to be true, that MPA/JHU and mid-IR SSFRs agree between themselves as well. While this seems to imply that MPA/JHU measurements are not biased, it is necessary to verify if such results hold for SFRs, and not just the SSFRs.

Figure 11 shows SED fitting SFRs against MPA/JHU SFRs in the upper panel, and mid-IR SFRs against MPA/JHU SFRs in the lower panel. In both comparisons, the linear fits (blue dashed lines) follow closely the 1:1 relation. White solid lines show the relation between SAMI and B04 SFRs (adjusted to Chabrier IMF), over the range of SFRs covered in Richards et al. (2016).<sup>17</sup> The discrepancies suggested by Richards et al. relation are clearly excluded with either the SED fitting or the mid-IR SFR. Results are unchanged when the redshift range of galaxies is restricted to match the redshift range of SAMI Galaxy Survey. Apparently, more work is needed to under-

<sup>17</sup> The highest SFRs in Richards et al. (2016) are an order of magnitude lower than the highest SFRs in SDSS, probably because their sample is drawn from much smaller volume compared to that of SDSS (smaller area, plus  $z < 0.06$ ).

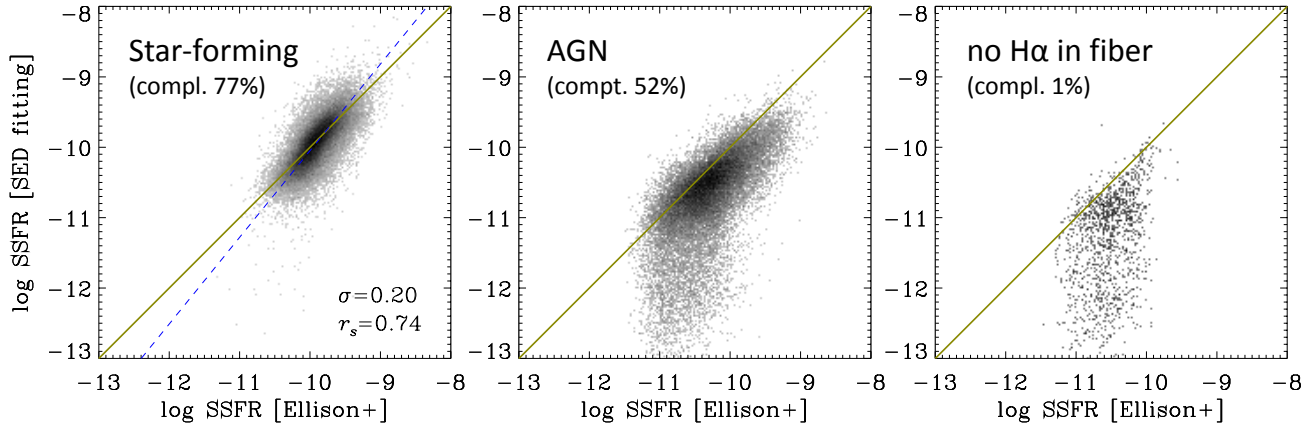


**Figure 11.** Comparison of SFRs from the SED fitting (upper panel) and from *WISE* mid-IR (lower panel), with respect to SFRs from the MPA/JHU catalog (Brinchmann et al. 2004). Shown are the galaxies from GSWLC-M classified as star-forming on the BPT diagram. There are no systematic differences or significant non-linearities. This is in contrast with the relation derived from integral-field spectroscopy  $\text{H}\alpha$  measurements from SAMI Galaxy Survey (Richards et al. 2016), shown as the white solid line (SAMI vs. MPA/JHU). In this figure we lower the redshift limit to  $z_{\text{min}} = 0.01$  to increase the contribution of low-SFR galaxies. Slight discretization in mid-IR SFRs arises from sampling of Chary & Elbaz (2001) IR templates.

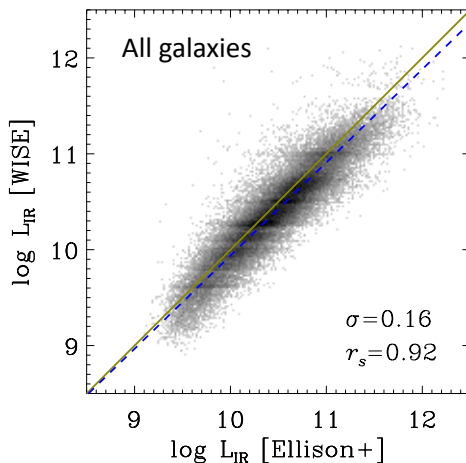
stand the source of differences. For now, we conclude that the total MPA/JHU SFRs of actively star-forming galaxies do not appear to be biased when compared with two independent measures of integrated SFR.

### 8.3. Comparison with Ellison et al. (2016) star formation rates

All-sky far-IR observations that would enable direct measurement of the total IR luminosity across SDSS are only available from relatively shallow *IRAS* and *AKARI* surveys, which preferentially more luminous galaxies (Ellison et al. 2016). In order to produce IR luminosity estimates for a significant portion of SDSS spectroscopic sample, Ellison et al. (2016) apply an artificial neural network (ANN) technique, using IR luminosities from Rosario et al. (2016) as the training set (these IR luminosities were discussed in Section 4). The idea behind the ANN is to establish intrinsic correlations between observable quantities of the training set and the target quantity (in this case the IR luminosity) and then apply



**Figure 12.** Comparison of SSFRs from the SED fitting (from GSWLC-M) with SSFRs from the IR luminosity catalog of Ellison et al. (2016), derived using artificial neural network (ANN) method trained on *Herschel*+*WISE* data. Galaxies are classified using the BPT diagram as in Figure 7. SF class shows the best fitting line and its scatter and correlation coefficient. Both SSFRs are normalized by the total stellar mass from GSWLC SED fitting. Comparison is similar to one involving mid-IR SSFRs (Figure 7), with some non-linearity for star-forming galaxies (left panel). Completeness of Ellison et al. (2016) catalog for a given class is indicated in each panel



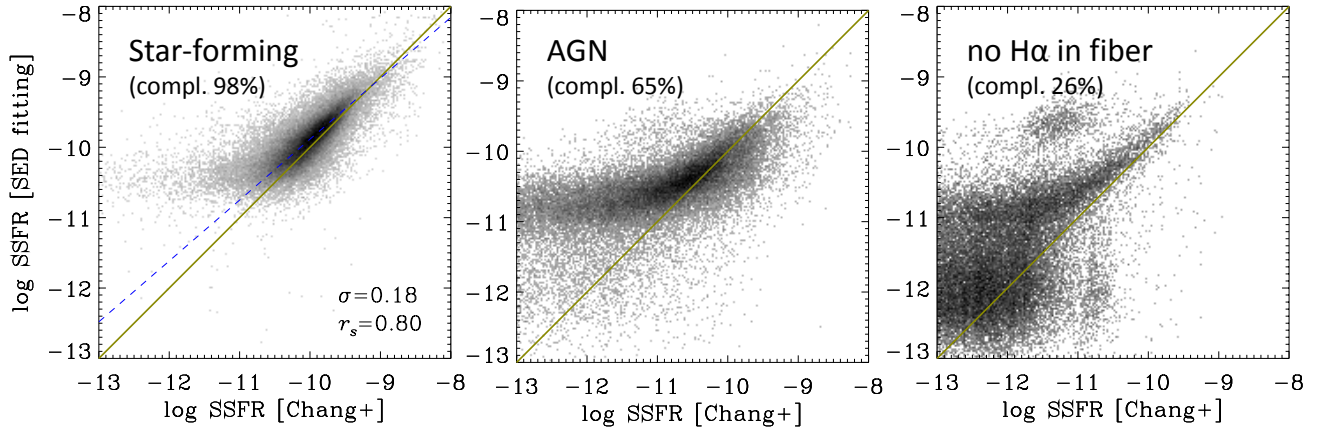
**Figure 13.** Comparison of total IR luminosities derived in GSWLC using *WISE* against the total IR luminosities derived with ANN in Ellison et al. (2016). *WISE* measurement is direct, but is based on extrapolating a single flux point at rest-frame  $\sim 20\mu\text{m}$ . ANN, on the other hand, uses various optical photometric and spectroscopic measurements to predict the IR luminosity based on a training set consisting of IR luminosities from *Herschel*+*WISE* (three sub-mm flux points from SPIRE and  $22\mu\text{m}$  point from *WISE*). The comparison is good notwithstanding some systematics for the most luminous galaxies. Luminosities are given in solar units.

these correlations to estimate (“predict”) the target quantity in the full dataset. The robustness of ANN estimates should be tested by comparing them with the independent measurements of the target quantity for an unbiased subsample drawn from the full dataset. In the case of Ellison et al. (2016), the parameters employed to estimate the IR luminosities using ANN include, in addition to the redshift, the photometric (magnitudes, colors) and spectroscopic measurements (emission lines strengths and D4000 break), as well as the stellar mass. The connection between the fiber and total quantities is established by also including the mass in the fiber and the  $r$ -band fiber covering fraction. The input parameters come from MPA/JHU DR7 catalog. The requirement to have available all of the parameters in the target dataset limits the application to 45% of SDSS galaxies, typically the ones with stronger emission lines. IR luminosities of the training set are recovered with the typical accuracy of  $\sigma \sim 0.1$  dex, with no major systematics (Ellison et al. 2016).

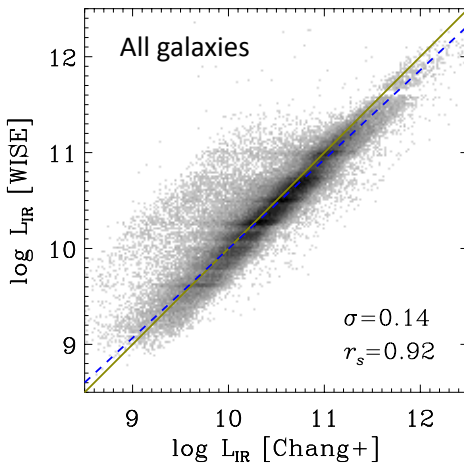
In Figure 12 we show the comparison of GSWLC SSFRs from the SED fitting against SSFRs converted from Ellison et al. (2016) IR luminosities using Eq. 8. The comparison is limited to galaxies with  $\sigma_{\text{ANN}} < 0.1$ , a cut recommended in Ellison et al. (2016) to remove the galaxies whose estimated IR luminosities may be uncertain.  $\sigma_{\text{ANN}}$  essentially measures the degree to which a target galaxy is represented in the training set. The larger the value, the less likely the target galaxy is represented in the training set and its IR luminosity may therefore be uncertain. The application of the cut on  $\sigma_{\text{ANN}}$  decreases the number of galaxies with ANN IR luminosities from 330,000 to 250,000, or 1/3 of SDSS. The comparison of SSFRs of star-forming galaxies (left panel) shows good general agreement, with mild non-linearity (slope of 1.23), which can be traced to the differences between *WISE* and *Herschel*-*WISE* specific IR luminosities (Section 4). If the galaxies with  $\sigma_{\text{ANN}} > 0.1$  were included in the comparison, the scatter at high SSFRs would increase, presumably because such galaxies are rare in the training set. ANN IR luminosities are available for 77% of galaxies in this class (after the application of  $\sigma_{\text{ANN}}$  cut), which is higher than the *WISE* detection rate of 63%.

For galaxies classified as AGN (the middle panel), the correlation is present when  $\log \text{SSFR}_{\text{SED}} > -11$ , but with ANN SSFRs tending to be higher, especially for galaxies with high SSFR. Below  $\log \text{SSFR}_{\text{SED}} = -11$  there is no correlation. Both of these behaviors mimic the comparison of SED fitting SSFRs with mid-IR SSFRs (Section 7 and Figure 7, middle panel). Excess SSFR for AGNs with higher SSFRs, which in the case of mid-IR SSFRs we attributed to AGN dust heating, is somewhat surprising, because it would be expected that the AGN contribution, which drops above  $40\mu\text{m}$  (Mullaney et al. 2011), will be significantly diminished in the total IR luminosity. However, we remind the reader that ANN luminosities are trained on IR luminosities derived from a combination of  $22\mu\text{m}$  flux (which is subject to AGN contamination) and three sub-mm flux points that lie well beyond the IR SED peak (and are therefore not as sensitive to the current SF, and more to the cold dust mass). More detailed analysis on the contribution of Type 2 AGN to IR SED lies outside of the scope of this paper.

After the application of  $\sigma_{\text{ANN}}$  cut, the IR luminosities from Ellison et al. (2016) are available for only 1% of galaxies having weak or no  $\text{H}\alpha$  (right panel). For a handful of such objects



**Figure 14.** Comparison of SSFRs from the SED fitting against the SSFRs from [Chang et al. \(2015\)](#), derived from optical+mid-IR SED fitting using MAGPHYS ([da Cunha et al. 2008](#)). Galaxies are classified using the BPT diagram as in Figure 7. SF class shows the best fitting line and its scatter and correlation coefficient. Both SSFRs are normalized by the total stellar mass from GSWLC SED fitting. Low SSFRs tend to be underestimated in [Chang et al. \(2015\)](#) catalog, but otherwise the correlation is good with a small scatter. Completeness of [Chang et al. \(2015\)](#) catalog for a given class is indicated in each panel.



**Figure 15.** Comparison of total IR luminosities derived in GSWLC from *WISE* to total IR luminosities derived from optical/mid-IR SED fitting from [Chang et al. \(2015\)](#). The comparison is generally quite good. A tail to the left is due to a small number of passive galaxies for which [Chang et al.](#) luminosities are too high.

with high SSFRs (including confirmed E+As) the match between the SED fitting and ANN SSFRs is good, but for the great majority, the ANN IR luminosities, when interpreted as SFRs, tend to be too high because of the dust heating from older stars, as already discussed in Section 7.

Considering the very different methods of estimating the IR luminosities, it is interesting to see how the ANN IR luminosities from [Ellison et al. \(2016\)](#) compare to the ones we obtain from *WISE*. This is shown in Figure 13, for all galaxies for which the two measurements are available (and with  $\sigma_{\text{ANN}} < 0.1$  cut applied). The comparison, extending over three orders of magnitude, is fairly good, both in terms of offset and the scatter ( $\sigma = 0.17$ ). Some systematic differences are present, especially at  $\log L_{\text{IR}} > 11$ , which is not surprising considering that the training set peters out at those luminosities. Overall, the IR luminosities from [Ellison et al. \(2016\)](#) represent remarkably good estimates considering that they were determined from optical properties alone, but their use as SFRs is subject to the same caveats (AGN dust heating and breakdown for passive galaxies) as the SFR obtained from *WISE* mid-IR data.

#### 8.4. Comparison with [Chang et al. \(2015\)](#) star formation rates

In Section 4 we mention that new SED fitting codes allow the modeling of the SED to extend into the IR, by including the dust emission. This approach is used in MAGPHYS ([da Cunha et al. 2008](#)), which models the IR SED as a sum of various SED components, the relative contribution of which is mildly related to galaxy’s SSFR. CIGALE allows the IR SED to be modeled according to one of the four published template sets, without constraints on the shape. In both cases the dust luminosity (i.e., the total IR luminosity) is normalized to match the stellar emission absorbed in the UV/optical/near-IR. Such IR luminosity will therefore include dust heating from stars of all ages.

Stellar plus dust emission modeling is applied in [Chang et al. \(2015\)](#), who use MAGPHYS to perform the SED fitting simultaneously on optical (*ugriz*) photometry from SDSS and mid-IR photometry from *WISE* (3.4, 4.6, 12 and 22  $\mu\text{m}$ ). Of these 9 bands, all but the longest two *WISE* bands will be dominated by the stellar emission. Emission at 12 and 22  $\mu\text{m}$ , if detected, will help constrain the dust luminosity and therefore the SFR (if the dust is primarily heated by young stars).

We follow the analysis established in previous sections, and in Figure 14 present comparisons between our SED fitting SSFRs and SSFRs from [Chang et al. \(2015\)](#), again split by galaxy type<sup>18</sup>. Note that we normalize both SSFRs using our stellar mass. Galaxies classified as star-forming on the BPT diagram (left panel) compare well, but show a tail of anomalously low values of [Chang et al.](#) SSFRs, which causes the best fit line to deviate from unity. However, if the tail is excluded, the non-linearity disappears, and a small bulk offset of 0.09 dex remains. We find that the offset is entirely due to  $\sim 1/3$  galaxies that are not detected at 22  $\mu\text{m}$  but only at 12  $\mu\text{m}$ , for which [Chang et al.](#) (S)SFRs are on average 0.17 dex lower than our SED fitting (S)SFRs. For galaxies with 22  $\mu\text{m}$  detection, the scatter with respect to SED SSFRs is 0.17 dex, compared to 0.23 dex between SED SSFRs and SSFRs from *WISE* 22  $\mu\text{m}$  alone (Figure 7, left panel). This reduction of scatter demonstrates that the 12  $\mu\text{m}$  photometry (a wave-

<sup>18</sup> We find that restricting the comparison to galaxies which [Chang et al. \(2015\)](#) flag as having good determinations has negligible effect on the results, so we use all galaxies regardless of that flag.

length at which *WISE* is significantly more sensitive than to 22  $\mu\text{m}$  helps constrain the SFRs compared to when 22  $\mu\text{m}$  is used alone. However, when 12  $\mu\text{m}$  is used to obtain the IR luminosity (or SFR) without the 22  $\mu\text{m}$  measurements, it leads to a systematic underestimate.

For AGN hosts (the middle panel), the tail of anomalously low SSFR values is more extensive than for the star-forming galaxies. We confirm that these galaxies have UV detections, and that the UV emission drives SSFR estimate in our SED fitting to higher values than when the UV is omitted from the SED fitting, as in the case of [Chang et al. \(2015\)](#) SED fitting. UV is not expected to be contaminated by non-stellar emission in these, Type 2, AGN. Furthermore, the UV emission is extended, suggestive of SF. Optical colors alone are not sensitive to such low levels of SF ([Kauffmann et al. 2007](#)), and are equally red for truly quiescent galaxies and those with intermediate SSFRs ([Fang et al. 2012](#)). When the SED fitting is performed without UV constraints, the SSFR defaults to very low values of the majority of optically red models

For AGN hosts with high SSFRs, the excess seen in other IR-based SSFRs is now smaller. This is true whether the galaxy was detected at both 12 and 22  $\mu\text{m}$ , or just at 12  $\mu\text{m}$ , suggesting that 12  $\mu\text{m}$  flux is much less affected by AGN-heated hot dust than the 22  $\mu\text{m}$ , so that when included in the SFR estimate it mitigates the excess. That 22  $\mu\text{m}$  suffers more AGN contamination is corroborated by the fact that out of 12  $\mu\text{m}$ -detected AGN, 80% are also detected at 22  $\mu\text{m}$ , whereas this fraction was 60% for star-forming galaxies. Previously, [Donoso et al. \(2012\)](#) have found that the 12  $\mu\text{m}$  luminosity is not affected by AGN contribution in all but the handful of most luminous sources (but they did not perform analogous assessment of 22  $\mu\text{m}$  emission).

For galaxies with weak or no  $\text{H}\alpha$  (right panel), [Chang et al. \(2015\)](#) values agree well for some of the galaxies with high SSFRs, with a small offset that we attribute to the preponderance (90%) of 12  $\mu\text{m}$ -only detections in this category, which we have shown to have somewhat underestimated SFRs in [Chang et al.](#) As in the case of comparison with MPA/JHU catalog SFRs, the E+As, or more generally, the post-starburst galaxies, again form a cloud of points offset from the 1:1 relation. As discussed in Section 8.2, the offset in (S)SFRs is attributable to the differences in the assumed dust attenuation laws, to which the E+As appear to be particularly sensitive. Comparison with entirely independent IR SFRs as well as better quality of SED fits, suggest that the high (S)SFRs obtained with the modified attenuation law in GSWLC are more realistic. Finally, for galaxies in “no  $\text{H}\alpha$ ” class with  $\log \text{SSFR} < -11$  the estimates largely agree, but are quite uncertain.

The general conclusion is that [Chang et al. \(2015\)](#) SFRs are reliable for galaxies with  $\log \text{SSFR}_{\text{C15}} > -10.7$ , regardless of the galaxy type, and especially when a galaxy is detected at 22  $\mu\text{m}$ . When SSFRs are low, [Chang et al. \(2015\)](#) correctly attribute the mid-IR emission to old populations and not the ongoing SF. However, in doing so, the weak signal from the actual SF, detectable in the UV, is in some cases lost, leading to anomalously low values of SSFR.

[Chang et al. \(2015\)](#) provide an estimate of the dust luminosity, i.e., the total IR luminosity. In Figure 15, we compare it to the IR luminosity that we derive from *WISE* 22  $\mu\text{m}$ . There is a good overall agreement and the scatter is smaller than in the comparison involving [Ellison et al. \(2016\)](#)  $L_{\text{IR}}$ , no doubt because *WISE* and [Chang et al.](#) IR luminosities are somewhat correlated through the use of the same 22  $\mu\text{m}$  photom-

etry. There is, however, a small fraction (1.8%) of galaxies for which there is a  $\sim 1$  dex discrepancy in  $L_{\text{IR}}$ . These galaxies produce a bump in the distribution of specific luminosities ( $\log(L_{\text{IR}}/M_*)$ ) when using [Chang et al. \(2015\)](#) IR luminosities, but not with *WISE* values. Most of these galaxies do not have  $L_{\text{IR}}$  from [Ellison et al. \(2016\)](#) catalog, but when they do, their agree with our IR luminosity from *WISE*. Finally, the direct comparison of [Chang et al.](#) IR luminosities with IR luminosities from *Herschel* (plot not shown) confirms that [Chang et al.](#) IR luminosities are underestimated when they are based on 12  $\mu\text{m}$  detection without the 22  $\mu\text{m}$  detection.

## 9. SUMMARY

The paper presents GSWLC, an extensive catalog of physical parameters (SFRs, dust attenuations and stellar masses) of  $\sim 700,000$  SDSS galaxies covered by *GALEX* (90% of SDSS). SFRs and stellar masses were derived using state-of-the-art SED fitting of UV and optical fluxes. SFRs derived independently from 22  $\mu\text{m}$  *WISE* photometry are also included in the catalog.

The construction of the catalog, the internal checks and the comparisons with previously published catalogs have produced a number of results, which we summarize here:

1. The principal source of bias in *GALEX* UV photometry is from blending of unresolved sources. We provide empirical recipe to reduce its effect.
2. Total IR luminosities obtained from *WISE* 22  $\mu\text{m}$  observations via [Chary & Elbaz \(2001\)](#) luminosity-dependent templates agree remarkably well ( $\Delta < 0.01$  dex, 0.07 dex of scatter) with IR luminosities obtained from *WISE* 22  $\mu\text{m}$  and *Herschel*-SPIRE sub-mm bands.
3. The comparison of SED SSFRs with SSFRs from the mid-IR and, separately, with  $\text{H}\alpha$  SSFRs, suggests that the majority of galaxies require dust attenuation curve that is significantly steeper than the [Calzetti et al. \(2000\)](#) curve, and is on average similar to the [Conroy et al. \(2010a\)](#) curve. Allowing this steep attenuation curve to include a UV bump further improves the quality of UV/optical SED fits.
4. Not accounting for the emission lines in the SED modeling of broadband UV-optical photometry produces significant biases in the derived (S)SFRs: up to 0.5 dex, on average, for high-SSFR galaxies ( $\log \text{SSFR} > -9.5$ ). Stellar masses are not affected by emission lines.
5. SFRs and stellar masses are not very different ( $< 0.1$  dex) when assuming a smooth, delayed SF history as opposed to the two-component (old and new) exponentially declining SF history, adopted for GSWLC. Fixing the old component to have started in the early universe ( $t_0 = 10$  Gyr), as done here, yields 0.1 dex higher masses for galaxies with active SF, possibly removing the “outshining” bias.
6. For actively star-forming (“main”) sequence galaxies ( $\log \text{SSFR} > -10.5$ ), there is a good general agreement between SED fitting and mid-IR (S)SFR from GSWLC and also with (S)SFRs from the literature.
7. Aperture-corrected (total) SFRs from the MPA/JHU catalog ([Brinchmann et al. 2004](#)) have no systematics

at high SFRs when compared with any other SFR, in contrast to systematic offsets reported in integral-field spectroscopy studies. GSWLC, by offering two independent SFRs can potentially be used in the forthcoming integral-field spectroscopy studies to elucidate the source of the discrepancy.

8. IR luminosities of galaxies that host AGN (as identified from the BPT diagram) and have relatively high SSFRs, appear to have an excess IR emission, presumably due to AGN dust heating affecting the  $22\ \mu\text{m}$  flux. If interpreted as SFR, this excess IR luminosity leads to SSFR overestimates of 0.2–0.6 dex, with greater effect at higher SSFR.
9. For galaxies that lie below the star-forming main sequence ( $12 < \log \text{SSFR}_{\text{SED}} < -11$ ), which includes quenching, nearly quiescent, or rejuvenated galaxies, SSFRs are low and their determination is challenging by any method. Using simple (fixed factor) conversions of IR luminosity to SFR (intended for use with actively star-forming galaxies) produces greatly exaggerated (S)SFRs (up to 2 dex). (S)SFRs from the UV/optical SED fitting tend to retain sensitivity in this regime. Below  $\log \text{SSFR} = -11.7$  (for GSWLC-M;  $= -11.5$  for GSWLC-A,  $-12.0$  for GSWLC-D), even the UV/optical SSFRs should be considered as upper limits, as these are, for all practical purposes, truly quiescent galaxies.

SDSS is the workhorse dataset for many galaxy evolution studies at low redshift, but without surveys in other wavelengths it has been limited in terms of characterizing the SF. We hope that GSWLC, by combining SDSS with *GALEX* and *WISE*, will fill this gap and serve as a resource for many new discoveries.

The construction of GSWLC was funded through NASA ADAP award NNX12AE06G. We thank Sara Ellison for valuable comments. SS thanks the members of *GALEX* science team for their contributions over the years, especially Tim Heckman, Chris Martin, R. Michael Rich and Mark Seibert. Funding for SDSS-III has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, and the U.S. Department of Energy Office of Science. The SDSS-III web site is <http://www.sdss3.org/>. Based on observations made with the NASA Galaxy Evolution Explorer. *GALEX* is operated for NASA by the California Institute of Technology under NASA contract NAS5-98034. This publication makes use of data products from the Wide-field Infrared Survey Explorer, which is a joint project of the University of California, Los Angeles, and the Jet Propulsion Laboratory/California Institute of Technology, funded by the National Aeronautics and Space Administration.

## REFERENCES

- Abazajian, K., Adelman-McCarthy, J. K., Agüeros, M. A., et al. 2004, *AJ*, 128, 502
- Abazajian, K. N., Adelman-McCarthy, J. K., Agüeros, M. A., et al. 2009, *ApJS*, 182, 543–558
- Ahn, C. P., Alexandroff, R., Allende Prieto, C., et al. 2014, *ApJS*, 211, 17
- Allen, J. T., Croom, S. M., Konstantopoulos, I. S., et al. 2015, *MNRAS*, 446, 1567
- Aloisi, A., Clementini, G., Tosi, M., et al. 2007, *ApJL*, 667, L151
- Baldwin, J. A., Phillips, M. M., & Terlevich, R. 1981, *PASP*, 93, 5
- Bell, E. F., & de Jong, R. S. 2001, *ApJ*, 550, 212
- Bernardi, M., Shankar, F., Hyde, J. B., et al. 2010, *MNRAS*, 404, 2087
- Bertin, E., & Arnouts, S. 1996, *A&AS*, 117, 393
- Boquien, M., Buat, V., & Perret, V. 2014, *A&A*, 571, A72
- Boquien, M., Kennicutt, R., Calzetti, D., et al. 2016, *A&A*, 591, A6
- Bothun, G. D. 1982, *ApJS*, 50, 39
- Bressan, A., Aussen, H., Granato, G. L., et al. 2001, *Astrophysics and Space Science Supplement*, 277, 251
- Brinchmann, J., Charlot, S., White, S. D. M., Tremonti, C., Kauffmann, G., Heckman, T., & Brinkmann, J. 2004, *MNRAS*, 351, 115
- Bruzual A., G. 1983, *ApJ*, 273, 105
- Bruzual A., G., & Charlot, S. 1993, *ApJ*, 405, 538
- Bruzual, G., & Charlot, S. 2003, *MNRAS*, 344, 1000
- Buat, V., Giovannoli, E., Takeuchi, T. T., et al. 2011, *A&A*, 529, A22
- Buat, V., & Xu, C. 1996, *A&A*, 306, 61
- Budavári, T., Heinis, S., Szalay, A. S., et al. 2009, *ApJ*, 694, 1281
- Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, *ApJ*, 533, 682
- Calzetti, D. 2013, *Secular Evolution of Galaxies*, 419, Cambridge, UK: Cambridge University Press (arXiv:1208.2997)
- Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, *ApJ*, 345, 245
- Chabrier, G. 2003, *PASP*, 115, 763
- Chang, Y.-Y., van der Wel, A., da Cunha, E., & Rix, H.-W. 2015, *ApJS*, 219, 8
- Charlot, S., & Bruzual, A. G. 1991, *ApJ*, 367, 126
- Charlot, S., & Fall, S. M. 2000, *ApJ*, 539, 718
- Chary, R., & Elbaz, D. 2001, *ApJ*, 556, 562
- Cid Fernandes, R., Mateus, A., Sodré, L., Stasińska, G., & Gomes, J. M. 2005, *MNRAS*, 358, 363
- Conroy, C., Gunn, J. E., & White, M. 2009, *ApJ*, 699, 486
- Conroy, C., & Gunn, J. E. 2010, *ApJ*, 712, 833
- Conroy, C., Schiminovich, D., & Blanton, M. R. 2010, *ApJ*, 718, 184
- Conroy, C., White, M., & Gunn, J. E. 2010, *ApJ*, 708, 58
- Conroy, C. 2013, *ARA&A*, 51, 393
- Cortese, L., Boselli, A., Franzetti, P., et al. 2008, *MNRAS*, 386, 1157
- Courteau, S., Cappellari, M., de Jong, R. S., et al. 2014, *Reviews of Modern Physics*, 86, 47
- da Cunha, E., Charlot, S., & Elbaz, D. 2008, *MNRAS*, 388, 1595
- Dale, D. A., & Helou, G. 2002, *ApJ*, 576, 159
- Donoso, E., Yan, L., Tsai, C., et al. 2012, *ApJ*, 748, 80
- Dopita, M. A., Kewley, L. J., Heisler, C. A., & Sutherland, R. S. 2000, *ApJ*, 542, 224
- Dressler, A., & Gunn, J. E. 1983, *ApJ*, 270, 7
- Ellison, S. L., Teimoorinia, H., Rosario, D. J., & Mendel, J. T. 2016, *MNRAS*, 455, 370
- Faber, S. M. 1972, *A&A*, 20, 361
- Fang, J. J., Faber, S. M., Salim, S., Graves, G. J., & Rich, R. M. 2012, *ApJ*, 761, 23
- Fitzpatrick, E. L., & Massa, D. 1986, *ApJ*, 307, 286
- Gallazzi, A., Charlot, S., Brinchmann, J., White, S. D. M., & Tremonti, C. A. 2005, *MNRAS*, 362, 41
- Gavazzi, G., Bonfanti, C., Sanvito, G., Boselli, A., & Scodreggio, M. 2002, *ApJ*, 576, 135
- Goto, T. 2007, *MNRAS*, 381, 187
- Guiderdoni, B., & Rocca-Volmerange, B. 1987, *A&A*, 186, 1
- Hao, C.-N., Kennicutt, R. C., Johnson, B. D., et al. 2011, *ApJ*, 741, 124
- Kauffmann, G., Heckman, T. M., White, S. D. M., et al. 2003, *MNRAS*, 341, 33
- Kauffmann, G., Heckman, T. M., Tremonti, C., et al. 2003, *MNRAS*, 346, 1055
- Kauffmann, G., Heckman, T. M., Budavári, T., et al. 2007, *ApJS*, 173, 357
- Kennicutt, R. C., Jr. 1998, *ARA&A*, 36, 189
- Kewley, L. J., Groves, B., Kauffmann, G., & Heckman, T. 2006, *MNRAS*, 372, 961
- Kriek, M., Labbé, I., Conroy, C., et al. 2010, *ApJL*, 722, L64
- Lang, D., Hogg, D. W., & Schlegel, D. J. 2016, *AJ*, 151, 36
- Larson, R. B., & Tinsley, B. M. 1978, *ApJ*, 219, 46
- Liang, Y. C., Yin, S. Y., Hammer, F., et al. 2006, *ApJ*, 652, 257
- Maraston, C. 2005, *MNRAS*, 362, 799
- Maraston, C. 2011, *Why Galaxies Care about AGB Stars II: Shining Examples and Common Inhabitants*, 445, 391
- Marcillac, D., Elbaz, D., Chary, R. R., Dickinson, M., Galliano, F., & Morrison, G. 2006, *A&A*, 451, 57
- Marigo, P. 2015, *Why Galaxies Care about AGB Stars III: A Closer Look in Space and Time*, 497, 229
- Martin, D. C., Seibert, M., Buat, V., et al. 2005, *ApJL*, 619, L59

- McGaugh, S. S., & Schombert, J. M. 2014, *AJ*, 148, 77
- Meert, A., Vikram, V., & Bernardi, M. 2016, *MNRAS*, 455, 2440
- Mendel, J. T., Simard, L., Palmer, M., Ellison, S. L., & Patton, D. R. 2014, *ApJS*, 210, 3
- Michałowski, M. J., Hayward, C. C., Dunlop, J. S., et al. 2014, *A&A*, 571, A75
- Mitchell, P. D., Lacey, C. G., Baugh, C. M., & Cole, S. 2013, *MNRAS*, 435, 87
- Morrissey, P., Conrow, T., Barlow, T. A., et al. 2007, *ApJS*, 173, 682
- Mullaney, J. R., Alexander, D. M., Goulding, A. D., & Hickox, R. C. 2011, *MNRAS*, 414, 1082
- Muzzin, A., van Dokkum, P., Kriek, M., et al. 2010, *ApJ*, 725, 742
- Noll, S., Burgarella, D., Giovannoli, E., et al. 2009, *A&A*, 507, 1793
- Pace, C., & Salim, S. 2016, *ApJ*, 818, 65
- 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
- Papovich, C., Dickinson, M., & Ferguson, H. C. 2001, *ApJ*, 559, 620
- Peek, J. E. G., & Schiminovich, D. 2013, *ApJ*, 771, 68
- Renzini, A., & Buzzoni, A. 1986, *Spectral Evolution of Galaxies*, 122, 195
- Richards, S. N., Bryant, J. J., Croom, S. M., et al. 2016, *MNRAS*, 455, 2826
- Rosario, D. J., Mendel, J. T., Ellison, S. L., Lutz, D., & Trump, J. R. 2016, *MNRAS*, 457, 2703
- Salim, S., Charlot, S., Rich, R. M., et al. 2005, *ApJL*, 619, L39
- Salim, S., Rich, R. M., Charlot, S., et al. 2007, *ApJS*, 173, 267
- Salim, S., Dickinson, M., Michael Rich, R., et al. 2009, *ApJ*, 700, 161
- Salim, S., Fang, J. J., Rich, R. M., Faber, S. M., & Thilker, D. A. 2012, *ApJ*, 755, 105
- Salmon, B., Papovich, C., Long, J., et al. 2015, *arXiv:1512.05396*
- Sanders, D. B., & Mirabel, I. F. 1996, *ARA&A*, 34, 749
- Sawicki, M., & Yee, H. K. C. 1998, *AJ*, 115, 1329
- Shapley, A. E., Steidel, C. C., Erb, D. K., et al. 2005, *ApJ*, 626, 698
- Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, *ApJ*, 500, 525
- Searle, L., Sargent, W. L. W., & Bagnuolo, W. G. 1973, *ApJ*, 179, 427
- Simard, L., Mendel, J. T., Patton, D. R., Ellison, S. L., & McConnachie, A. W. 2011, *ApJS*, 196, 11
- Simha, V., Weinberg, D. H., Conroy, C., et al. 2014, *arXiv:1404.0402*
- Smith, D. J. B., & Hayward, C. C. 2015, *MNRAS*, 453, 1597
- Sorba, R., & Sawicki, M. 2015, *MNRAS*, 452, 235
- Stasińska, G., Vale Asari, N., Cid Fernandes, R., et al. 2008, *MNRAS*, 391, L29
- Stecher, T. P. 1965, *ApJ*, 142, 1683
- Strauss, M. A., Weinberg, D. H., Lupton, R. H., et al. 2002, *AJ*, 124, 1810
- Taylor, E. N., Hopkins, A. M., Baldry, I. K., et al. 2011, *MNRAS*, 418, 1587
- Tinsley, B. M., & Gunn, J. E. 1976, *ApJ*, 203, 52
- Tremonti, C. A., Heckman, T. M., Kauffmann, G., et al. 2004, *ApJ*, 613, 898
- Tully, R. B., Mould, J. R., & Aaronson, M. 1982, *ApJ*, 257, 527
- Tully, R. B., Courtois, H. M., & Sorce, J. G. 2016, *arXiv:1605.01765*
- van der Wel, A., Franx, M., Wuyts, S., et al. 2006, *ApJ*, 652, 97
- Viero, M. P., Asboth, V., Roseboom, I. G., et al. 2014, *ApJS*, 210, 22
- Villaume, A., Conroy, C., & Johnson, B. D. 2015, *ApJ*, 806, 82
- Walcher, J., Groves, B., Budavári, T., & Dale, D. 2011, *Ap&SS*, 331, 1
- Weisz, D. R., Dalcanton, J. J., Williams, B. F., et al. 2011, *ApJ*, 739, 5
- West, A. A., Garcia-Appadoo, D. A., Dalcanton, J. J., et al. 2010, *AJ*, 139, 315
- Wild, V., Charlot, S., Brinchmann, J., et al. 2011, *MNRAS*, 417, 1760
- Wright, E. L., Eisenhardt, P. R. M., Mainzer, A. K., et al. 2010, *AJ*, 140, 1868-1881
- Yan, L., Donoso, E., Tsai, C.-W., et al. 2013, *AJ*, 145, 55
- Yang, Y., Zabludoff, A. I., Zaritsky, D., & Mihos, J. C. 2008, *ApJ*, 688, 945-971
- Yuan, H. B., Liu, X. W., & Xiang, M. S. 2013, *MNRAS*, 430, 2188
- Zibetti, S., Gallazzi, A., Charlot, S., Pierini, D., & Pasquali, A. 2013, *MNRAS*, 428, 1479