



The ALFALFA-SDSS Galaxy Catalog

Adriana Durbala¹ , Rose A. Finn² , Mary Crone Odekon³ , Martha P. Haynes⁴ , Rebecca A. Koopmann⁵ , and Aileen A. O'Donoghue⁶ 

¹Department of Physics and Astronomy, University of Wisconsin - Stevens Point, Stevens Point, WI 54481, USA; adurbala@uwsp.edu

²Department of Physics and Astronomy, Siena College, 515 Loudon Road, Loudonville, NY 12211, USA; rfinn@siena.edu

³Department of Physics, Skidmore College, Saratoga Springs, NY 12866, USA; mcrone@skidmore.edu

⁴Cornell Center for Astrophysics and Planetary Science, Cornell University, Ithaca, NY 14853, USA; haynes@astro.cornell.edu

⁵Department of Physics and Astronomy, Union College, Schenectady, NY 12308, USA; koopman@union.edu

⁶Department of Physics, St. Lawrence University, 23 Romoda Drive, Canton, NY 13617, USA; aodonoghue@stlawu.edu

Received 2020 July 30; revised 2020 September 28; accepted 2020 October 7; published 2020 November 18

Abstract

We present a HI optical catalog of $\sim 30,000$ galaxies based on the 100% complete Arecibo Legacy Fast Arecibo L -band Feed Array (ALFALFA) survey combined with data from the Sloan Digital Sky Survey (SDSS). Our goal is to facilitate the public use of the completed ALFALFA catalog by providing carefully determined matches to SDSS counterparts, including matches for $\sim 10,000$ galaxies that do not have SDSS spectra. These identifications can provide a basis for further crossmatching with other surveys using SDSS photometric IDs as a reference point. We derive absolute magnitudes and stellar masses for each galaxy using optical colors combined with an internal reddening correction designed for small- and intermediate-mass galaxies with active star formation. We also provide measures of stellar masses and star formation rates based on infrared and/or ultraviolet photometry for galaxies that are detected by the Wide-field Infrared Survey Explorer and/or the Galaxy Evolution Explorer. Finally, we compare the galaxy population in the ALFALFA-SDSS sample with the populations in several other publicly available galaxy catalogs and confirm that ALFALFA galaxies typically have lower masses and bluer colors.

Unified Astronomy Thesaurus concepts: Galaxies (573); Late-type galaxies (907); Galaxy evolution (594); Sky surveys (1464); Astronomy databases (83)

Supporting material: machine-readable tables

1. Introduction

The Arecibo Legacy Fast Arecibo L -band Feed Array (ALFALFA) survey provides HI 21 cm line measurements for $\sim 31,500$ galaxies over nearly 7000 deg^2 on the sky, out to a redshift of about 0.06 (Giovanelli et al. 2005; Haynes et al. 2018). As an “untargeted” radio survey, ALFALFA gives a HI-selected view of the low-redshift galaxy population. Primary goals of the survey included determining the HI mass function (e.g., Martin et al. 2012; Moorman et al. 2014; Jones et al. 2018), the HI width function (Papastergis et al. 2011; Moorman et al. 2014), and the HI-selected galaxy correlation function (Martin et al. 2012; Papastergis et al. 2013). The HI-selected galaxy population proved to be surprisingly diverse, including many galaxies that still have massive HI disks at low redshift.

Beyond characterizing the galaxy population in terms of HI alone, ALFALFA is a major resource for comparison with surveys at other wavelengths. Using the 40% complete ALFALFA survey, “ $\alpha.40$,” for example, Huang et al. (2012) determined scaling relations between the HI gas fraction and optical and UV colors for a sample of 9417 galaxies with Sloan Digital Sky Survey (SDSS) and Galaxy Evolution Explorer (GALEX) measurements. They found that the specific star formation rate (sSFR) in galaxies with $M_* < 10^{9.5}$, but not in higher-mass galaxies, is correlated with gas fraction, suggesting that star formation in low-mass galaxies is strongly regulated by HI. Comparing ALFALFA-selected populations with SDSS-selected populations, they found that ALFALFA galaxies have bluer colors, higher SFRs and sSFRs, and lower

star formation efficiencies (possibly caused by higher spin parameters).

The $\alpha.40$ catalog has been the starting point for several follow-up observing programs. For example, the H α 3 survey (Fossati et al. 2013; Gavazzi et al. 2013) is a narrowband optical imaging follow-up survey of ~ 800 galaxies from ALFALFA in the Local Supercluster and the Coma Supercluster. Among other results, it shows a significant decrease in both HI content and sSFR for galaxies closer to cluster centers, with the outside-in quenching expected from ram pressure stripping.

Also building from the $\alpha.40$ catalog, the xGASS survey (Catinella et al. 2018) combines new and previous HI measurements with GALEX and SDSS measurements to create a gas-fraction-limited, stellar-mass-selected sample of 1179 galaxies down to a mass of $M_* = 10^9 M_\odot$. The related xCOLD GASS survey (Saintonge et al. 2017) examines molecular gas in 532 galaxies in the same mass range using CO (1–0) observations with the IRAM 30 m telescope complemented by CO (2–1) observations with the IRAM 30 m and APEX telescopes, HI Arecibo observations, and photometry from SDSS, Wide-field Infrared Survey Explorer (WISE), and GALEX. They confirmed that the molecular gas fractions depend more strongly on the sSFR than on stellar mass.

One particularly promising use of ALFALFA data is in the context of the baryonic Tully–Fisher relation (BTFR). McGaugh et al. (2000) show that for small galaxies, the combination of gas mass and stellar mass has a surprisingly tight correlation with rotation speed. The BTFR is useful because it both places strong constraints on theories of galaxy

formation and evolution and enables the measurement of secondary distances and peculiar velocities. Papastergis et al. (2016) use α .40 to calibrate the BTFR for gas-dominated galaxies.

The primary purpose of the current paper is to facilitate the public use of the recently completed 100% ALFALFA survey in conjunction with observations at other wavelengths. First, we provide SDSS identifications for nearly all (29,638 out of 31,501) ALFALFA galaxies, including \sim 10,000 that do not have SDSS spectroscopy. The SDSS identifications can provide a basis for further crossmatching with other surveys using SDSS photometric IDs as a reference point. We derive absolute magnitudes and stellar masses for each galaxy using optical colors combined with an internal reddening correction designed for small- and intermediate-mass galaxies with active star formation. We also provide measures of stellar masses and SFRs based on infrared and/or ultraviolet photometry for galaxies that are detected by WISE and/or GALEX. In this way, the final catalog puts the HI information from ALFALFA in the context of each galaxy’s stellar content.

This paper is organized as follows. Section 2 describes the identification of HI sources, the assignment of SDSS optical counterparts, optical extinction corrections, and stellar mass calculations based on optical colors. Section 3 describes the stellar mass and SFR calculations using infrared and ultraviolet photometry. Section 4 presents the ALFALFA-SDSS catalog. Section 5 places the ALFALFA-SDSS catalog in the context of three other galaxy catalogs that include stellar mass and/or SFR, showing the overall differences between the ALFALFA-SDSS galaxy population and populations with other selection criteria. Section 6 summarizes the content and scope of the catalog.

Throughout this paper, we adopt the cosmological parameters $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$, and $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$.

2. SDSS Counterparts and Optical Properties

Exploiting the large collecting area of the Arecibo 305 m antenna and the seven-beam ALFA radio camera, ALFALFA mapped \sim 6600 deg² of high galactic latitude sky in spectral line mode, covering a 100 MHz bandwidth corresponding to $2000 < cz < 18,000 \text{ km s}^{-1}$ sampled as 4096 spectral channels, yielding a resolution of 5.5 km s^{-1} at $z \simeq 0$ before smoothing (Giovanelli et al. 2005; Haynes et al. 2011). As described in Appendix A of Haynes et al. (2011), multiple fixed-position drift scan crossings were combined to construct three-dimensional (position–position–velocity) grids. Source identification was performed by applying a Fourier domain-based matched filter algorithm to the processed spectral grids (Saintonge 2007). Source extraction and parameter measurement were performed by interactive analysis of each candidate detection, allowing for localized baseline fitting and parameter extraction.

The size of a single ALFALFA beam is $3\prime.8 \times 3\prime.3$. The grids are constructed on a spatial grid sampled at $1' \times 1'$ to which a Gaussian weight function of $2'$ is applied. This reduces the spatial resolution of the grid to $\sim 4\prime.3 \times 3\prime.8$. As discussed in Section 5 of Giovanelli et al. (2007), the pointing accuracy of the extracted HI sources is limited by the resolution of each ALFA beam, the signal-to-noise ratio (S/N) of the HI emission, and the quality of the telescope’s pointing. The latter has been accounted for by fitting offsets between peaks in the continuum sources associated with each spectral grid and the positions of radio sources cataloged by the Northern VLA

Sky Survey (NVSS; Condon et al. 1998). Examination of both the scatter in the continuum offsets and the difference in positions between the HI sources and their likely optical counterparts (see next section) show that the positional accuracy of high-S/N (>10) sources is $15''\text{--}18''$, but can be much larger, even exceeding $1'$ at lower S/N (Giovanelli et al. 2007; Kent et al. 2008).

2.1. Identification of Optical Counterparts

In addition to the parameters of the HI emission, the ALFALFA catalog includes the position of the “most probable” optical counterpart. The process by which optical counterparts are assigned to the HI sources is discussed in Appendix A and especially Figures 8 and 9 of Haynes et al. (2011). Essentially, the individual interactively performing the source parameter extraction examined, at the same time, several digital extragalactic source catalogs and imaging databases to search for likely optical counterparts to the HI emission. The search box for coincidence was adjusted to take into account the S/N as discussed above, with larger regions searched for sources of lower S/N. Within the positional error box of the vast majority of HI sources lies a star-forming galaxy which was deemed the most probable optical counterpart. Where the optical identification was less obvious, the best candidate was identified based on proximity to the HI centroid, morphological appearance as a star-forming disk galaxy, and, where available, known and coincident radial velocity. The latter information in particular was used as an additional criterion for the identification of a separate category of HI sources (HI code = 2, the “priors”) with S/N below the threshold of 6.5 applied to the high-quality detections (HI code = 1).

The ALFALFA source catalog presented in Haynes et al. (2018) contains 31,501 sources, of which 25,433 (80%) are high quality and 6068 are lower S/N “priors.” As discussed by Haynes et al. (2018) and Leisman et al. (2017), nearly all of the high-quality ALFALFA sources can be identified with a likely optical counterpart. In fact, of the high-S/N sources, only 344 (1.4%) are not assigned a probable optical counterpart, and most of those are probably tidal debris. For the 25,089 high-quality HI sources with identified optical counterparts, the median separation of the position of the identified optical counterpart from the HI centroid is $19\prime.9$; the mean is $23\prime.6$ with a standard deviation of $16\prime.8$. Some of this separation value is due to uncertainties in the optical positions recorded in the ALFALFA database, generally estimated to be $<3\prime.0$.

In Table 3 of Haynes et al. (2011), we presented galaxies in the 40% ALFALFA HI catalog, along with a crossmatch to SDSS DR7. The cross-identification was performed as part of the data reduction process, whereby the user marked the SDSS counterpart interactively by visual inspection of the SDSS image served by the NAVIGATOR tool, as part of the standard data reduction process. Later SDSS data releases presented new cross-identifications with new photometric solutions, so the process of cross-identification was reinitiated for the full ALFALFA catalog in 2017. In the latter case, a more automated approach was adopted, but it made use of the positional matches made during the earlier visual inspection.

For the final ALFALFA-SDSS crossmatch reported here (ALFALFA-SDSS), a more automated approach was performed using the SDSS Cross ID tool.⁷ Matches were sought within a search radius of $0\prime.1$ around the center of the optical

⁷ <http://skyserver.sdss3.org/public/en/tools/crossid/crossid.aspx>

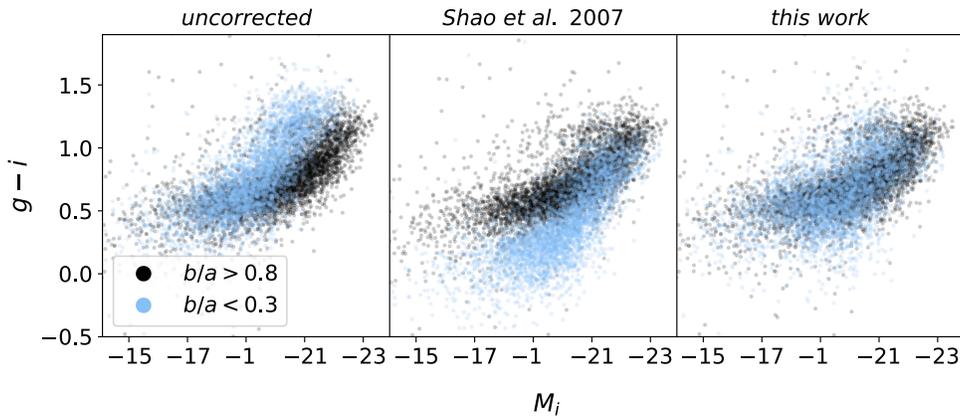


Figure 1. The effect of the internal extinction correction on the color–magnitude diagram for the cases of (left) no correction, (center) the correction from Shao et al. (2007), and (right) the correction adopted in this work (see the text for details). Each panel compares galaxies that have highly inclined disks ($b/a < 0.3$, light blue) with galaxies that appear nearly face on ($b/a > 0.8$, black). We expect these two groups to have similar colors and magnitudes with an appropriate extinction correction.

counterpart identified in the ALFALFA database. Galaxies with anomalous magnitudes were inspected individually. Many of these were galaxies with dust lanes, bright H II regions, or superposed stars, where there were multiple photometric sources. In some cases, it was possible to reassign an appropriate SDSS source for the galaxy as a whole. In other cases, this was not possible.

Most galaxies with a clear SDSS counterpart are assigned an optical photometry flag of “1” (28,267 objects). Galaxies with a clear SDSS counterpart but with large photometric uncertainties (as described in the section on optical photometry below) are assigned the flag “2” (1371 objects). Galaxies with no clear SDSS counterparts are assigned the flag “0” if they were outside the SDSS footprint (1296 objects) and “3” otherwise (567 objects).

Distances to each galaxy are estimated using the process described in Haynes et al. (2018).

2.2. Optical Photometry and Extinction

Following the recommendations on the SDSS website, we use SDSS *cmodel* mags to calculate galaxy absolute magnitudes and model mags for colors. Galaxies with *g*- or *i*-band errors greater than 0.05 were assigned a photometry flag of “2” and excluded from color–magnitude diagrams and stellar mass calculations based on optical magnitudes. Excluding these galaxies cuts out the majority of objects with anomalous colors that suggest magnitude uncertainties beyond the formal errors calculated by the SDSS pipeline. Inspection of individual galaxies indicates that several different factors contribute to the anomalous magnitudes, including contamination by nearby stars and shredding of large galaxies. We also note that the SDSS provides a “clean” parameter, but using this as a flag preferentially excludes the small, blue galaxies that dominate our sample; a full 43% of our galaxies would be left out, most of which look normal otherwise. Our flag based on magnitude uncertainties leaves out only 4% of the sample.

We correct optical photometry for foreground extinction by the Milky Way using the $E(B - V)$ map of Schlegel et al. (1998) with the $R_V = 3.1$ reddening curve of Schlafly & Finkbeiner (2011). We do not apply the additional 14% recalibration to lower values for extinction suggested by Schlafly & Finkbeiner (2011),

because not all authors agree the calibration should be lower. For example, the Planck Collaboration et al. (2014) find the values in Schlegel et al. (1998) to be too low (by about 8%) rather than too high. We note that for the typical values of $E(B - V)$ in our sample, around 0.06, a shift of about 10% in $E(B - V)$ corresponds to a change in magnitude of only 0.02 in the *g* band and 0.01 in the *i* band. Indeed, the scatter in the results from different methods is larger than the systematic shift. For the purposes of error propagation, we adopt an uncertainty of 20% in the values of *g*- and *i*-band galactic extinction corrections (in magnitudes), and an uncertainty of 0.02 for the $g - i$ color (see Green et al. 2015, 2019).

Galactic extinction values provided by the SDSS for Data Release 15 are the same as our chosen values; they also use the conversions from $E(B - V)$, but not the 14% recalibration, of Schlafly & Finkbeiner (2011). (The SDSS pipeline does not apply galactic extinction corrections automatically; this is left to the user.)

We also correct optical photometry for extinction internal to each galaxy. A simple and standard way of doing this is to estimate the extinction in magnitudes using $A_\lambda = \gamma_\lambda \log_{10}(a/b)$, where a/b is the axial ratio obtained from $\text{exp}AB_r$ in SDSS, and γ_λ is a constant for each filter (see Giovanelli et al. 1994; Shao et al. 2007). However, if we use the same γ for all the galaxies in our sample, this method clearly overcorrects the photometry for fainter galaxies. As shown in Figure 1, the “corrected” $g - i$ colors for edge-on faint galaxies, for example, are bluer than those for face-on galaxies by about half a magnitude. Indeed, there has long been evidence that there is less internal extinction in less luminous galaxies (e.g., Tully et al. 1998). There is also increasing evidence for still more complicated extinction effects, including a nonlinear dependence of A_λ on $\log_{10}(a/b)$, and differences in extinction based on additional parameters, including bulge-to-disk ratios, colors, gas content, and surface brightness (Masters et al. 2010; Devour & Bell 2016; Kourkchi et al. 2019).

Our approach here is to provide a simple correction that captures the observed overall dependence of extinction on absolute magnitude for ALFALFA galaxies, which are nearly all members of the star-forming blue cloud (Haynes et al. 2011). We use a value of γ that is zero for absolute magnitudes fainter than -17 and that changes linearly for brighter

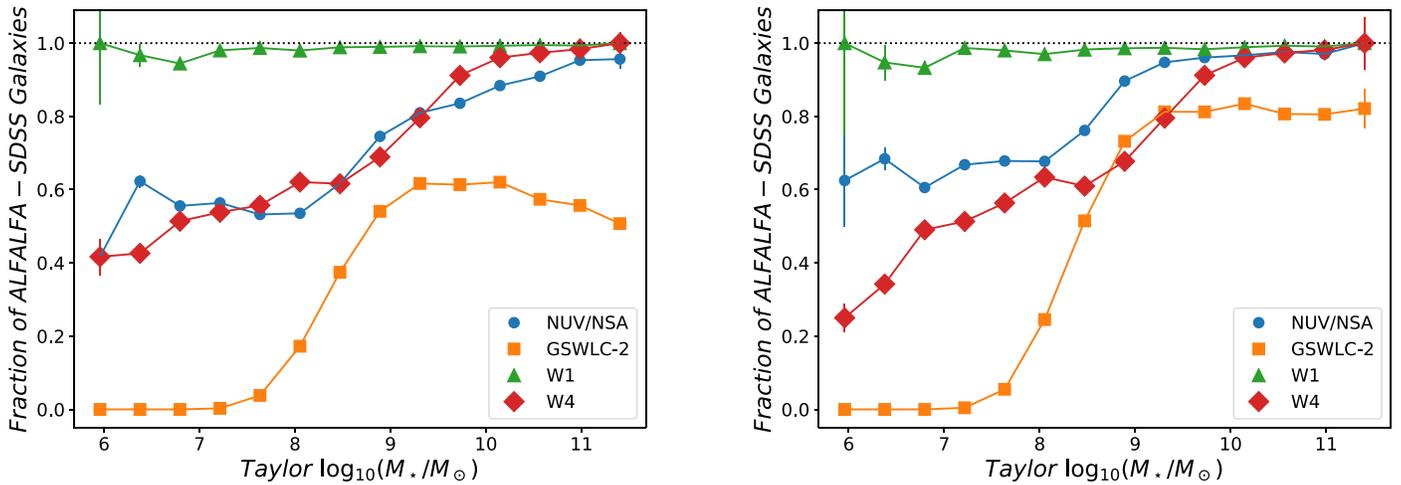


Figure 2. (Left) Fraction of ALFALFA-SDSS galaxies included in the NSA (blue circles), GSWLC-2 (orange squares), unWISE W1 (green triangles), and W4 (red diamonds) surveys as a function of stellar mass. We use the full survey areas for this comparison. (Right) Same as the left panel, but for the survey overlap regions only.

magnitudes according to

$$\begin{aligned} \gamma_g &= -0.35M_g - 5.95, M_g < -17 \\ \gamma_i &= -0.15M_i - 2.55, M_i < -17 \end{aligned} \quad (1)$$

with an uncertainty of ± 0.3 in γ , and where the magnitudes M_g and M_i are corrected for foreground Galactic extinction only. This correction is designed to be consistent with the results of Masters et al. (2010) and Devour & Bell (2016) for the case of star-forming galaxies with small and intermediate magnitudes. We do not recommend using it for passive (red sequence) galaxies or for very massive galaxies, both of which show evidence for less extinction than that given by these equations. Figure 1 illustrates the effect of the adopted internal extinction correction on the galaxy color–magnitude diagram, compared with using either no correction at all or the correction from Shao et al. (2007). Each panel compares galaxies that have highly inclined disks ($b/a < 0.3$) with galaxies that appear nearly face on ($b/a > 0.8$). We expect these two groups to have similar colors and magnitudes with an appropriate extinction correction.

2.3. Derivation of Stellar Masses Using SDSS Photometry

The stellar mass of a galaxy is a fundamental property that, when combined with information about whether the galaxy is star-forming or quenched, can closely predict many other galaxy properties, including color, shape, and metallicity (e.g., Balogh et al. 2004; Brinchmann et al. 2004; Blanton et al. 2005; Mouhcine et al. 2007; Brough et al. 2013). Characterizing galaxies in terms of their stellar mass is powerful and intuitive, and provides a way to compare results across different types of galaxy surveys. Stellar mass can also be used in conjunction with gas mass to determine distances according to the Tully–Fisher relation (Tully & Fisher 1977). In the context of the ALFALFA-SDSS catalog, we are particularly interested in the BTFR (McGaugh et al. 2000), which improves the simple Tully–Fisher relation for small galaxies by including gas mass as well as stellar mass. This gives us an additional motivation for characterizing the ALFALFA galaxies in terms of their stellar mass.

Stellar masses can be estimated by comparing spectra and/or broadband photometry to stellar population synthesis models (Kauffmann et al. 2003; Chen et al. 2012; Maraston et al. 2013). The GALEX-SDSS-WISE Legacy Catalog 2 (GSWLC-2; Salim et al. 2016, 2018), for example, uses UV/optical/IR SED fitting to calculate stellar masses and SFRs.

For the purposes of providing stellar masses for the ALFALFA galaxy population as a whole, however, we cannot take advantage of the values provided by these catalogs because of the lack of sufficient overlap between them and our survey volume. This is especially true for the part of ALFALFA in the “fall sky” (galactic southern hemisphere) not covered by the SDSS spectroscopic surveys. Furthermore, many existing catalogs that provide stellar masses for large samples of galaxies (e.g., Blanton et al. 2011; Salim et al. 2016) are incomplete at the low stellar masses typical of ALFALFA galaxies.

To illustrate these coverage and completeness issues, we show the fraction of ALFALFA-SDSS galaxies detected by several surveys in Figure 2. In the left panel, we show the fraction of ALFALFA galaxies included in the GSWLC-2 (orange squares) when we compare the full regions covered by each survey. To estimate completeness, we limit the comparison to a region where the surveys overlap: $140^\circ < \text{R.A.} < 230^\circ$, $0^\circ < \text{decl.} < 35^\circ$, and $z < 0.05$. In the right panel of Figure 2, we see that even in the overlap region, the GSWLC-2 misses a large fraction of ALFALFA galaxies with $\log_{10}(M_*/M_\odot) < 9$. With an eye toward calculating SFRs as well as stellar masses (Sections 3.2, 3.1), we also show the fraction of ALFALFA galaxies that have GALEX NUV fluxes (blue circles) reported in the NASA-Sloan Atlas (NSA; Blanton et al. 2011), and the fraction detected at W4 ($22 \mu\text{m}$) in the unWISE catalog (red diamonds; Lang 2014; Lang et al. 2016). While these surveys include a higher fraction of ALFALFA galaxies than the GSWLC-2, they still miss the majority of low-mass galaxies. Almost all ALFALFA-SDSS galaxies seem to be detected in unWISE W1 ($3.4 \mu\text{m}$; green triangles), and we will utilize this to calculate an IR-based stellar mass in Section 3.1.

The incomplete sampling of ALFALFA galaxies in existing catalogs drives us to calculate independent stellar mass estimates from SDSS photometry. As an optically based

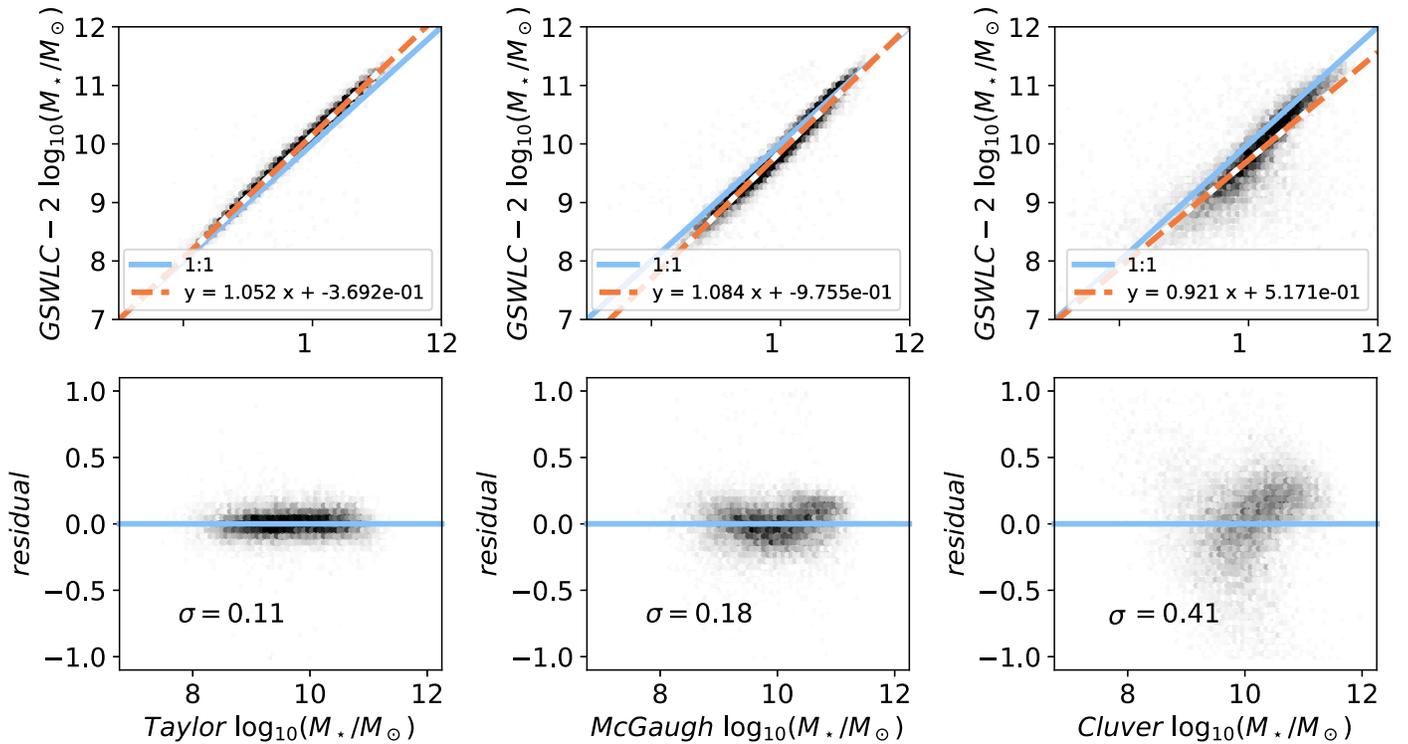


Figure 3. Top panels: GSWLC-2 stellar masses vs. (left) stellar mass estimates from the SDSS $g - i$ color (Taylor et al. 2011), (center) unWISE W1 (McGaugh & Schombert 2015), and unWISE W1 – W2 color (Cluver et al. 2014). The light blue line shows a one-to-one relation, and the orange dashed line shows a linear fit. Bottom panels: residuals from the fitted relation for each mass estimate. For both the Taylor et al. (2011) and McGaugh & Schombert (2015) mass estimates, we are able to remove systematic offsets with respect to stellar mass.

measure of stellar mass, we adopt the method of Taylor et al. (2011), which can be used for all the galaxies in the matched ALFALFA-SDSS catalog, including those without optical spectra. This simple prescription is based on optical color and magnitude according to

$$\log M_*/L_i = -0.68 + 0.70(g - i). \quad (2)$$

Taylor et al. (2011) show that their method agrees with more complicated ones. We denote the stellar mass determined using the Taylor method as $M_{*,\text{Taylor}}$.

About 40% of the galaxies in the ALFALFA-SDSS catalog have stellar masses from the GSWLC-2, and we compare our stellar mass estimates with theirs in the top-left panel of Figure 3. The two mass estimates are tightly correlated, but the $M_{*,\text{Taylor}}$ estimates tend to be slightly below the GSWLC-2 masses, and the offset increases with stellar mass. We provide a translation between the two mass estimates by fitting a linear relationship:

$$\begin{aligned} \log_{10}(M_*/M_\odot)_{\text{GSWLC-2}} \\ = 1.052 \log_{10}(M_*/M_\odot)_{\text{Taylor}} - 0.369. \end{aligned} \quad (3)$$

When we use this fit to correct the $M_{*,\text{Taylor}}$ estimates, we are able to remove the dominant systematic offsets. This is borne out in the bottom-left panel, which shows the residuals between the GSWLC-2 and corrected Taylor stellar masses. The dispersion of the residuals is only 0.11 dex.

In the ALFALFA-SDSS catalog (Section 4), we include the (uncorrected) values for $M_{*,\text{Taylor}}$ as well as the GSWLC-2 values for the $\sim 40\%$ of galaxies for which they are available.

3. Infrared and Ultraviolet Properties

3.1. Derivation of Stellar Masses Using unWISE Photometry

Infrared photometry provides an additional method for estimating stellar masses and is available for a large fraction of the galaxies in the ALFALFA-SDSS catalog thanks to the all-sky coverage of NASA’s Wide-field Infrared Survey Explorer (WISE⁸; Wright et al. 2010). WISE mapped the sky in four infrared bands: $3.4 \mu\text{m}$ (W1), $4.6 \mu\text{m}$ (W2), $12 \mu\text{m}$ (W3), and $22 \mu\text{m}$ (W4) with an angular resolution of $6''1$, $6''4$, $6''5$, and $12''0$. The unWISE catalog is derived from a reprocessing of WISE imaging (Lang 2014) and has two big advantages over the existing WISE catalog (AllWISE): deeper imaging and improved modeling of crowded regions. It detects sources at a 5σ level to about ~ 0.7 mag fainter than the AllWISE catalog, doubling the number of galaxies detected between redshifts $0 < z < 1$. In the end, unWISE offers more accurate photometry for extended sources than the AllWISE catalog, and we therefore choose to use it in our study for the identification of infrared counterparts.

There are two databases associated with unWISE. One is the unWISE Catalog, a point-source catalog with sources identified by a unique unWISE ID (Schlafly et al. 2019). The other is the unWISE/SDSS Forced Photometry catalog, which takes SDSS sources and shapes, and fits for unWISE fluxes that best match the unWISE images (Lang et al. 2016). ALFALFA sources were matched to unWISE sources by SDSS objID number in this catalog (D. Lang 2020, private communication). This yields 29,088 ALFALFA sources with an unWISE match.

⁸ The Wide-field Infrared Survey Explorer (WISE) All-Sky Data Release is available at <http://wise2.ipac.caltech.edu/docs/release/allsky/>.

Table 1
Basic Optical Properties of Cross-listed Objects in the ALFALFA-SDSS Catalog

AGC	Flag	SDSS objID	R.A.	Decl.	V_{helio}	D	σ_D	Ext _g	Ext _i	expAB _r	σ_{expAB_r}	cmodel _i	σ_{cmodel_i}
(1)	(2)	(3)	J2000	J2000	km s ⁻¹	Mpc	Mpc	mag	mag	(11)	(12)	mag	mag
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
1	1	1237679455462228052	0.656670	16.65222	5839	82.8	2.2	0.11	0.06	0.77	0.01	15.07	0.01
3	1	1237679502171701451	0.692920	18.88583	7883	107.6	2.3	0.11	0.06	0.47	0.01	13.08	0.01
4	1	1237678660887445652	0.737080	4.20889	8621	118.0	2.3	0.09	0.05	0.64	0.01	13.88	0.01
6	1	1237679476933722181	0.790420	21.95972	6561	88.8	2.2	0.15	0.08	0.73	0.01	13.52	0.01
7	1	1237656496724639763	0.796670	15.96500	11223	155.2	2.2	0.15	0.08	0.37	0.01	13.27	0.01
8	1	1237652944786292742	0.811670	16.14556	1050	13.2	1.3	0.15	0.08	0.79	0.01	10.46	0.01
10	1	1237669682261983347	0.835420	8.61861	11941	165.4	2.1	0.24	0.13	0.89	0.01	13.41	0.01
11	1	1237680297818390670	0.839580	22.10250	4445	62.8	2.4	0.17	0.09	0.72	0.01	14.80	0.01
12	1	1237663234988769589	0.835000	29.79722	6980	95.0	2.4	0.19	0.10	0.59	0.01	14.26	0.01
13	1	1237663234452947057	0.871670	27.35139	7749	105.9	2.2	0.18	0.09	0.78	0.01	13.52	0.01

(This table is available in its entirety in machine-readable form.)

We consider two methods of estimating stellar masses using infrared photometry from unWISE. McGaugh & Schombert (2015) calculate stellar mass using the WISE W1 band, which is dominated by light from old stars (Jarrett et al. 2013):

$$M_{\star}/L_{W1} = 0.45M_{\odot}/L_{\odot}. \quad (4)$$

We denote stellar masses determined using this method as $M_{\star, \text{McGaugh}}$. The center column of Figure 3 compares the GSWLC-2 masses with the $M_{\star, \text{McGaugh}}$ estimates. We are able to remove the dominant systematic offsets between the two mass estimates using a linear fit. However, the scatter between the McGaugh and GSWLC-2 masses is systematically larger than for the Taylor-GSWLC-2 masses. As discussed in Salim et al. (2016), these differences may be related to uncertainties in the contribution from post-main-sequence stars to W1 flux. The best-fit conversion between the McGaugh & Schombert (2015) and GSWLC-2 masses is

$$\begin{aligned} \log_{10}(M_{\star}/M_{\odot})_{\text{GSWLC-2}} \\ = 1.084 \log_{10}(M_{\star}/M_{\odot})_{\text{McGaugh}} - 0.9755 \end{aligned} \quad (5)$$

for

$$8 < \log_{10}(M_{\star}/M_{\odot})_{\text{McGaugh}} < 11.$$

Another measure of stellar mass, derived by Cluver et al. (2014), combines the WISE W1 luminosity and the WISE W1 – W2 color:

$$\log_{10}(M_{\star}/L_{W1}) = -2.54(W_{3.4 \mu\text{m}} - W_{4.6 \mu\text{m}}) - 0.17. \quad (6)$$

Here, $W_{3.4 \mu\text{m}} - W_{4.6 \mu\text{m}}$ is the rest-frame color of the source and $L_{W1}(L_{\odot}) = 10^{-0.4(M - M_{\odot})}$, where $M_{\odot} = 3.24$ and M is the absolute magnitude of the source in the W1 band. We denote stellar masses determined using this alternative method as $M_{\star, \text{Cluver}}$. We compare this stellar mass estimate with the GSWLC-2 mass in the right column of Figure 3. The Cluver et al. (2014) stellar masses are reasonably consistent for $\log_{10}(M_{\star}) > 10$, but the scatter and offset are large for lower-mass galaxies. Therefore, we do not report a best-fit relation, and we do not include the Cluver et al. (2014) stellar masses in the catalog (Section 4).

3.2. Derivation of Star Formation Rates Using IR and UV Photometry

Several well-characterized indicators are used to trace recent star formation, including the direct detection of UV radiation from massive stars, and the infrared emission from dust grains that absorb some of the UV light and reradiate it in the infrared. When both UV and infrared fluxes are available, the combination provides an SFR tracer that is robust against extinction. The combination of WISE and GALEX makes these measurements possible for large samples of nearby galaxies (e.g., Salim et al. 2018; Leroy et al. 2019).

We provide multiple measures of the total SFR for galaxies that are detected by WISE and/or GALEX. First, we calculate SFRs using the 22 μm flux and the conversions from Kennicutt & Evans (2012) based on calibrations from Rieke et al. (2009). A total of 23,895 galaxies have a detection in W4 (we require $W4 > 0$). We convert the unWISE W4 magnitudes from Vega to AB by adding 6.620 (Jarrett et al. 2011) and then to Janskys using a flux zero point of 3631 Jy. We compute νL_{ν} by multiplying the flux in Jansky by the frequency at 22 μm and by $4\pi D^2$, where D is the flow-corrected distance from Haynes et al. (2018; see Table 1). According to Kennicutt & Evans (2012), the SFR is then

$$\log_{10}(\text{SFR}_{22}) = \log_{10}(\nu L_{\nu_{22}}) - 42.69. \quad (7)$$

We denote SFRs determined this way as $\log_{10}(\text{SFR}_{22})$.

For galaxies that have GALEX near-ultraviolet fluxes in the NASA-Sloan Atlas (Blanton et al. 2011), we calculate the NUV SFRs (Hao et al. 2011; Murphy et al. 2011; Kennicutt & Evans 2012) and corrected NUV SFRs (Hao et al. 2011; Kennicutt & Evans 2012), respectively, as

$$\log_{10}(\text{SFR}_{\text{NUV}}) = \log_{10}(\nu L_{\nu}) - 43.17 \quad (8)$$

and

$$\log_{10}(\text{SFR}_{\text{NUV,corr}}) = \log_{10}(\nu L_{\nu_{\text{corrected}}}) - 43.17. \quad (9)$$

where the corrected NUV spectral energy density $\nu L_{\nu_{\text{corrected}}}$ is the sum $\nu L_{\nu}(\text{NUV}) + 2.26\nu L_{\nu}(22 \mu\text{m})$. A total of 22,848 ALFALFA sources have NUV detections. (We note that a substantial part of the sky is missing in GALEX FUV because the FUV detector failed early in the mission; for this reason, we use NUV estimates.)

In Figure 4, we compare the three measures of the SFR with the values from the GSWLC-2 (Salim et al. 2016, 2018). The

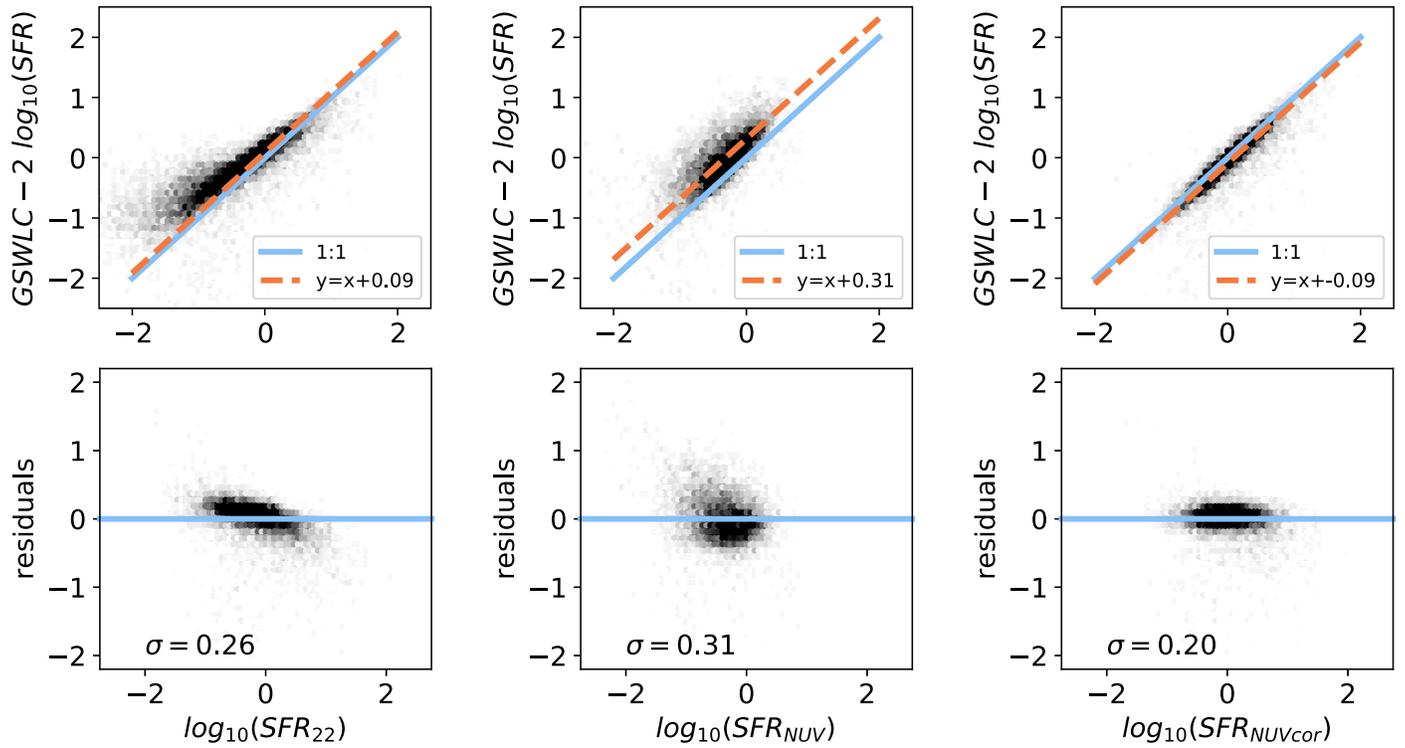


Figure 4. Top panels: GSWLC-2 SFR vs. estimates from (left) WISE W4, (center) GALEX near-UV, and (right) corrected near-UV. The light blue line shows the one-to-one relation, and the orange dashed line has a slope of 1 but an intercept that is the best-fit zero-point offset between the two indicators. Bottom panels: residuals between the two SFR indicators after correcting for the zero-point offset. The corrected NUV SFR provides the most reliable estimate of SFR and should be used when available.

GSWLC-2 SFRs are based on SED fitting from UV/optical photometry jointly with the mid-IR flux from $22\ \mu\text{m}$, or $12\ \mu\text{m}$ if $22\ \mu\text{m}$ is not detected (Salim et al. 2018). In the left column, we show the GSWLC-2 SFR versus the WISE $22\ \mu\text{m}$ SFR. The two measures of SFR are in good agreement for $\log_{10}(\text{SFR}) > 0$ but the $22\ \mu\text{m}$ SFRs fall below the GSWLC-2 values for lower SFRs and lower-mass galaxies. This is expected due to the lower metallicity and dust in lower-mass galaxies, which results in less extinction. We find an average offset of 0.09, but this is due mostly to the offsets observed in lower-SFR galaxies. We do not fit a linear relationship because this would make the $\text{SFR}_{22} > 0$ values inaccurate. Instead, we prefer to use the $22\ \mu\text{m}$ SFR with the caveat that they will underestimate the true SFR for low-mass, low-SFR galaxies. The NUV SFR (middle column) underestimates the total SFR by an average of 0.31 due to dust absorption, which again is a strong function of stellar mass. When the NUV flux is corrected for emission that is absorbed and reradiated in the IR, the inferred SFR is much closer to GSWLC-2 values (right column). Therefore, when both NUV and IR are available, the corrected NUV SFR should be used. Otherwise, the $22\ \mu\text{m}$ SFR is the next best option. NUV alone is the least reliable, and we do not include it in our catalog.

4. The ALFALFA-SDSS Catalog

Table 1 presents the ALFALFA 100% catalog including basic SDSS DR15 properties of crossmatched galaxies.

The table is organized as follows:

1. Column 1—AGC number (entry number in the AGC catalog).

2. Column 2—photometry flag. 0: outside the SDSS footprint, 1: SDSS photometry with uncertainties less than 0.05 in g and i (*good photometry*), 2: SDSS photometry with uncertainties greater than 0.05 in g and/or i (*bad photometry*), 3: no SDSS counterpart identified, despite being within the SDSS footprint.
3. Column 3—SDSS DR15 object ID of the optical counterpart.
4. Column 4—R.A. (J2000) of the optical counterpart or HI centroid, if no optical counterpart has been identified.
5. Column 5—decl. (J2000) of the optical counterpart or HI centroid, if no optical counterpart has been identified.
6. Column 6—heliocentric velocity (cz) of the HI profile midpoint in km s^{-1} .
7. Column 7—distance in megaparsec estimated as described in Haynes et al. (2018).
8. Column 8—uncertainty in distance from Haynes et al. (2018).
9. Column 9— g -band Galactic extinction in mag, as described in Section 2.2.
10. Column 10— i -band Galactic extinction in mag, as described in Section 2.2.
11. Column 11— expAB_r axial ratio b/a in the r band from SDSS.
12. Column 12—uncertainty in expAB_r axial ratio b/a in r band from SDSS.
13. Column 13—SDSS i -band c model magnitude.
14. Column 14—uncertainty in the i -band c model magnitude from SDSS.

We include additional information for each galaxy, including all the photometric measurements from SDSS needed to

calculate the derived optical properties in a table at <http://egg.astro.cornell.edu/alfalfa/data/index.php>.

Table 2 presents the derived properties of the cross-listed objects in the ALFALFA-SDSS catalog, including absolute magnitude, color, stellar mass, HI mass, and SFR. Stellar masses and SFR have been calculated using the methods described in Sections 2.3, 3.1, and 3.2. We also provide stellar mass and SFR from GSWLC-2, when available. Table 2 is organized as follows:

1. Column 1—AGC number (entry number in the AGC catalog).
2. Column 2— g -band internal extinction factor γ_g in mag, as described in Section 2.2.
3. Column 3— i -band internal extinction factor γ_i in mag, as described in Section 2.2.
4. Column 4—corrected absolute i -band magnitude in mag obtained using SDSS i -band *cmodel* magnitude corrected for galactic and internal extinction, as described in Section 2.2.
5. Column 5—uncertainty in corrected absolute i -band magnitude.
6. Column 6—corrected $g - i$ color obtained in mag using the SDSS g - and i -band model magnitude corrected for galactic and internal extinction, as described in Section 2.2.
7. Column 7—uncertainty in the corrected $g - i$ color.
8. Column 8—stellar mass from SDSS optical photometry in logarithmic solar units using the Taylor method $M_{\star, \text{Taylor}}$, as described in Section 2.3.
9. Column 9—uncertainty in $\log M_{\star, \text{Taylor}}$.
10. Column 10—stellar mass from infrared unWISE photometry in logarithmic solar units using the McGaugh method $\log M_{\star, \text{McGaugh}}$, as described in Section 3.1.
11. Column 11—uncertainty in $\log M_{\star, \text{McGaugh}}$.
12. Column 12—stellar mass from GSWLC-2 in logarithmic solar units.
13. Column 13—uncertainty in $\log M_{\star, \text{GSWLC-2}}$.
14. Column 14—SFR from unWISE infrared photometry using the flux at the frequency at $22 \mu\text{m}$ SFR_{22} in logarithmic $M_{\odot} \text{yr}^{-1}$, as described in Section 3.2.
15. Column 15—uncertainty in SFR_{22} .
16. Column 16—corrected near-ultraviolet SFR from GALEX NUV photometry $\text{SFR}_{\text{NUV, cor}}$ in logarithmic $M_{\odot} \text{yr}^{-1}$ for galaxies with NUV fluxes available in NASA-Sloan Atlas, as described in Section 3.2.
17. Column 17—uncertainty in $\text{SFR}_{\text{NUV, cor}}$.
18. Column 18—SFR from GSWLC-2 in logarithmic $M_{\odot} \text{yr}^{-1}$.
19. Column 19—uncertainty in $\text{SFR}_{\text{GSWLC-2}}$.
20. Column 20—HI mass in logarithmic solar units $\log M_{\text{HI}}$ from Haynes et al. (2018).
21. Column 21—uncertainty in $\log M_{\text{HI}}$ from Haynes et al. (2018).

The full table is available online at the following website: <http://egg.astro.cornell.edu/alfalfa/data/index.php>.

5. Comparison with Other Catalogs

As an untargeted radio survey, the ALFALFA population is selected for HI gas content, with an additional bias toward galaxies with narrow HI line widths (Giovanelli et al. 2005). In this section, we place the ALFALFA-SDSS catalog in the

context of three other galaxy catalogs that include stellar masses: the NSA, S4G, and GSWLC-2 catalogs. Our goal is to emphasize the overall differences between the ALFALFA-SDSS galaxy population and those from catalogs with selection effects that are related to different physical properties.

For the purpose of comparing the ALFALFA-SDSS galaxy population with each of the other three catalogs, we limit the galaxies to a volume where the surveys overlap in order to account for differences in the local environment (e.g., clusters versus voids) and differences in the limiting distance. Figure 5 outlines the overlap volumes between ALFALFA and the comparison catalogs using orange hatched rectangles. The exact R.A. and decl. ranges of each region are provided in the following subsections. The galaxy populations of each catalog are displayed with light blue symbols in each panel. The dark blue dashed lines trace the area in the sky covered by ALFALFA.

Two of the comparison catalogs, the NSA and the S4G, do not include a corresponding SDSS object ID. For these, we find cross-identifications by searching for matches where the position difference on the sky is less than $15''$ and the radial velocity difference is less than 300 km s^{-1} . Note that this matching process is different from the more involved process described in Section 2 used to create the ALFALFA-SDSS catalog itself. For the third comparison catalog, GSWLC-2, we find cross-identifications by matching with the SDSS object ID. Table 3 summarizes the population statistics by catalog, first for each catalog separately and then for the overlap volumes.

5.1. Comparison with the NSA

The NASA-Sloan Atlas version *v1_0_1* (NSA; Blanton et al. 2011) includes images, photometric parameters, and spectroscopic parameters based on SDSS and GALEX data for 641,409 nearby galaxies to a redshift of $z = 0.15$. Calculations are optimized for galaxies out to this redshift and are designed to be an improvement over those in the SDSS pipeline. Future versions of the NSA are expected to incorporate information from the Two-Micron All-Sky Survey (2MASS) and WISE. Absolute magnitudes in the NSA include k -corrections based on photometry using the *kcorrect* package (Blanton & Roweis 2007). The *kcorrect* package estimates metallicities, SFRs, stellar masses, and mass-to-light ratios as it determines the best k correction; these estimates are included in the NSA as well.

Based on a match of the entire catalog, we find that 22,860 NSA galaxies are also in ALFALFA. (The NSA *v1_0_1* has a parameter for membership in the ALFALFA catalog, but this was based on an early, incomplete version of the catalog; there are many more matches now.)

To compare the galaxy populations in ALFALFA and the NSA, we first select galaxies that lie within a volume that is common to both surveys. Specifically, we use the following selection criteria: $z < 0.05$, and

$$140^\circ < \text{R.A.} < 230^\circ,$$

$$0^\circ < \text{decl.} < 35^\circ.$$

We compare the color versus stellar mass in Figure 6. In the left panels, we show the distribution of galaxies that are common to both surveys (blue symbols) and galaxies that are in the NSA but not in ALFALFA (orange contours). Note that in the left panels we show the $M_g - M_i$ color (corrected for galactic

Table 2
Derived Properties of Cross-listed Objects in the ALFALFA-SDSS Catalog

AGC	γ_g	γ_i	$M_{i\text{corr}}$	$\sigma_{M_{i\text{corr}}}$	$(g-i)_{\text{corr}}$	$\sigma_{(g-i)_{\text{corr}}}$	$\log M_*$ Taylor	$\sigma_{\log M_*}$ Taylor	$\log M_*$ McGaugh	$\sigma_{\log M_*}$ McGaugh	$\log M_*$ GSWLC	$\sigma_{\log M_*}$ GSWLC	$\log \text{SFR}_{22}$	$\sigma_{\log \text{SFR}_{22}}$	$\log \text{SFR}_{\text{NUVcor}}$	$\sigma_{\log \text{SFR}_{\text{NUVcor}}}$	$\log \text{SFR}$ GSWLC	$\sigma_{\log \text{SFR}}$ GSWLC	M_{HI}	$\sigma_{M_{\text{HI}}}$	
(1)	mag	mag	mag	mag	mag	mag	$\log(M_\odot)$	(9)	$\log(M_\odot)$	(11)	$\log(M_\odot)$	(13)	$\log(M_\odot \text{ yr}^{-1})$	(15)	$\log(M_\odot \text{ yr}^{-1})$	(17)	$\log(M_\odot \text{ yr}^{-1})$	(18)	(19)	$\log(M_\odot)$	(21)
1	0.73	0.39	-19.62	0.07	0.48	0.06	9.34	0.05	9.78	0.02	9.67	0.05	
3	1.40	0.77	-22.39	0.11	1.06	0.14	10.85	0.11	10.89	0.02	-0.40	0.07	-0.22	0.03	10.28	0.05	
4	1.25	0.68	-21.66	0.07	0.83	0.09	10.39	0.07	10.62	0.02	0.12	0.03	0.25	0.02	9.74	0.06	
6	1.21	0.64	-21.38	0.07	0.77	0.07	10.24	0.05	11.23	0.02	1.50	0.02	1.40	0.03	9.29	0.07	
7	1.59	0.86	-23.14	0.13	1.00	0.19	11.10	0.14	11.21	0.01	11.39	0.01	-0.06	0.05	-0.05	0.04	0.01	0.12	10.36	0.05	
8	0.69	0.48	-20.26	0.22	1.64	0.05	10.41	0.09	9.32	0.09	-1.97	0.09	-1.56	0.09	8.96	0.10	
10	1.70	0.87	-22.86	0.04	1.11	0.04	11.07	0.03	11.22	0.01	0.44	0.03	0.59	0.03	10.64	0.05	
11	0.55	0.34	-19.33	0.10	0.65	0.07	9.33	0.06	9.58	0.03	-2.18	0.66	-0.90	0.04	9.25	0.06	
12	1.01	0.56	-20.85	0.09	0.77	0.10	10.03	0.08	10.29	0.02	-0.15	0.03	9.79	0.05	
13	1.27	0.70	-21.77	0.06	1.05	0.06	10.59	0.05	10.79	0.02	0.12	0.02	0.21	0.03	9.24	0.08	

(This table is available in its entirety in machine-readable form.)

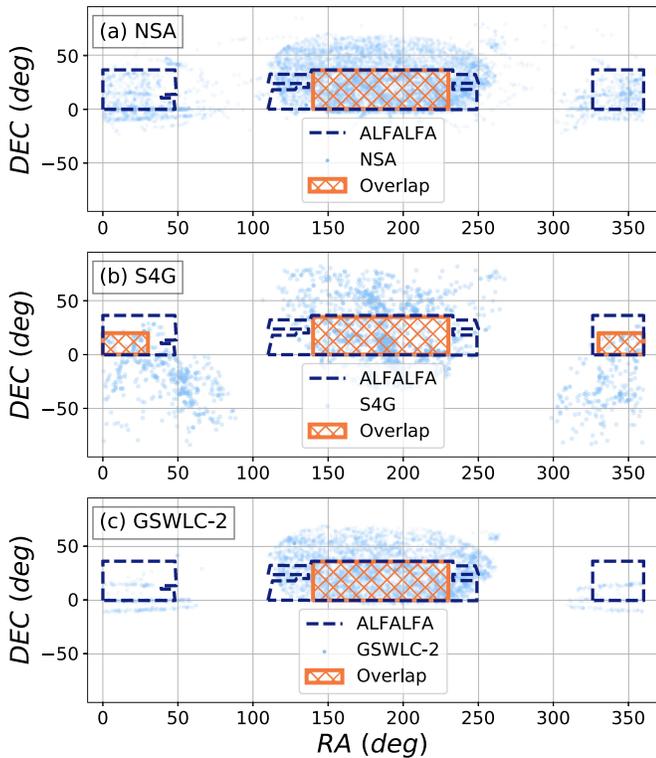


Figure 5. Decl. vs. R.A. of galaxies in (top) NSA, (center) GSWLC-2, and (bottom) S4G with $c_z < 15,000 \text{ km s}^{-1}$ (NSA and GSWLC-2) and $c_z < 3000 \text{ km s}^{-1}$ (S4G) shown in light blue symbols. The dark blue dashed line traces the area covered by ALFALFA. The orange hatched areas show the overlap regions used to compare ALFALFA with each survey. See the text for the exact R.A. and decl. coordinates of the regions used for comparison.

Table 3
Population Statistics by Catalog

Catalog	Total Number	Matched with ALFALFA
Full catalogs		
ALFALFA	31501	31501
NSA	641409	22860
S4G	2352	736
GSWLC-2	640659	15009
Overlap comparison volumes		
NSA : ALFALFA	42999 : 15467	13547
S4G : ALFALFA	607 : 16811	541
GSWLC-2 : ALFALFA	34628 : 15467	10425

Note. For each overlap comparison volume, we list the total number of galaxies in each catalog within this volume and the number of ALFALFA galaxies within this volume, separated by a colon (column two) as well as the number of matched galaxies (column three). The limits of these volumes in terms of R.A., decl., and redshift are described in the text.

extinction but not internal extinction) and stellar mass that come from the NSA catalog (*SERSIC_ABSMAG*, *EXTINCTION*, and *MASS*) because not all of the galaxies are in ALFALFA. The left plot and histograms of the $g - i$ color show that the NSA galaxy population is dramatically different from the ALFALFA galaxy population in that it includes the red sequence as well as the blue cloud. In the right panels, we show the complementary comparison, namely galaxies that are in ALFALFA but not in the NSA (light blue symbols), and we again compare to the population that is in common (blue

contours). Note that in the right-hand panels, we show our values of $g - i$ (Section 2.2) and M_* (Section 2.3), because not all of the galaxies are in the NSA. The ALFALFA galaxies that are not in the NSA (lighter blue symbols) are lower mass and bluer than those that are in the NSA (blue symbols).

5.2. Comparison with the S4G

The Spitzer Survey of Stellar Structure in Galaxies (S4G) is designed to provide a large baseline sample for the distribution of stellar mass within galaxies out to about 40 Mpc. It contains infrared images and photometric parameters for 2352 galaxies extending down to stellar masses $\sim 10^7 M_\odot$ that have been mapped using the IRAC 3.6 and $4.5 \mu\text{m}$ channels (Sheth et al. 2010; Muñoz-Mateos et al. 2013; Querejeta et al. 2015). It covers a much larger area of the sky than ALFALFA, but out to a smaller redshift of $z = 0.01$. The overlap volume with ALFALFA used to compare galaxy populations is shown in the middle panel of Figure 5. In the north galactic hemisphere, the overlap region is defined as $z < 0.01$ and

$$138^\circ < \text{R.A.} < 232^\circ,$$

$$0^\circ < \text{decl.} < 35^\circ.$$

We add an additional region in the southern galactic hemisphere of $z < 0.01$ and

$$0^\circ < \text{R.A.} < 30^\circ \text{ or } 330 < \text{R.A.} < 360^\circ,$$

$$0^\circ < \text{decl.} < 20^\circ.$$

The matching statistics are reported in Table 3, and the stellar masses and colors are shown in Figure 7. While most of the S4G galaxies are small and blue relative to large optical surveys, ALFALFA galaxies are still bluer than the S4G population. The samples have similar masses, although the resulting mass distributions are somewhat sensitive to the exact choice of overlap region. We investigated using slightly different criteria for the overlap region, including the same cut in redshift but different boundaries for the position on the sky. In each case, the ALFALFA galaxies were bluer. However, in some cases, the ALFALFA galaxies had, on average, slightly lower masses, and in other cases, they had slightly higher masses.

5.3. Comparison with the GSWLC

The GALEX-SDSS-WISE Legacy Catalog (GSWLC, Salim et al. 2016, 2018) contains galaxies within the GALEX footprint, whether or not they were detected in the UV. The catalog includes physical properties (e.g., stellar mass, SFR) of about 650,000 galaxies with SDSS redshifts below 0.3. There are two versions of the catalog: the GSWLC-1 and the GSWLC-2. Both versions contain the same sources (exactly the same number in the same order), and they both use identical photometry. GSWLC-2 has more accurate SFRs from joint UV + optical + mid-IR SED fitting while GSWLC-1 contains SFRs from the UV + optical SED fitting. For our comparison with ALFALFA, we use GSWLC-2. The redshifts reported in the GSWLC-2 catalog are from SDSS.

The overlap volume employed to compare galaxy populations in GSWLC-2 and ALFALFA is the same as that for the NSA: $z < 0.05$ and

$$140^\circ < \text{R.A.} < 230^\circ,$$

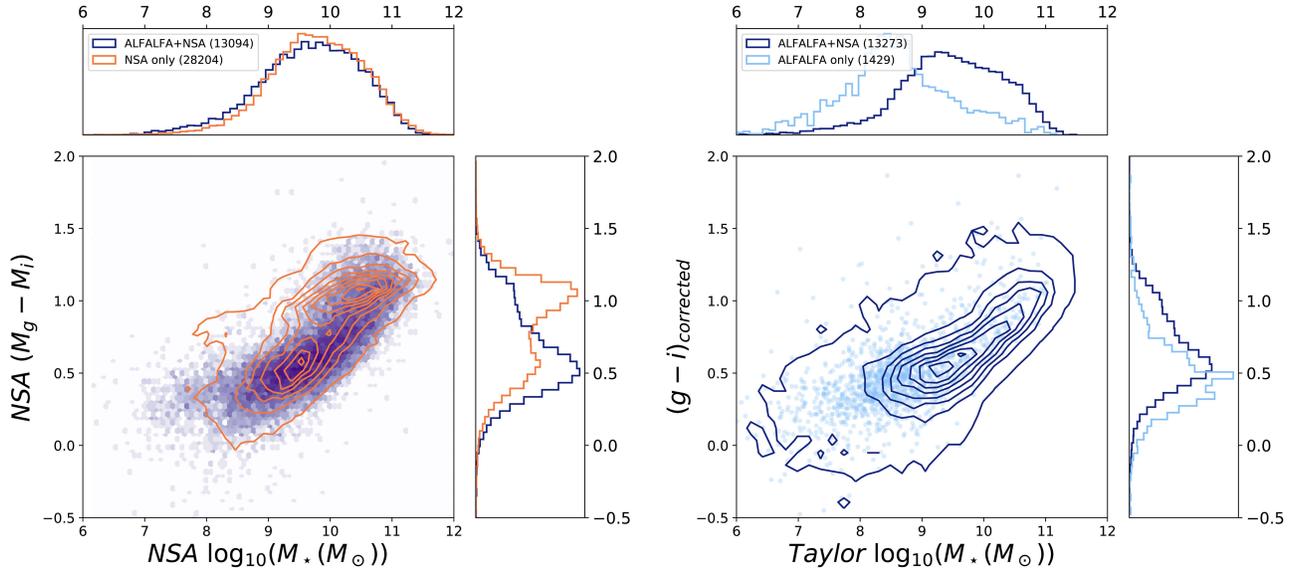


Figure 6. (Left) $M_g - M_i$ color vs. stellar mass (both from the NSA) for galaxies that lie in a volume common to both the ALFALFA and NSA surveys. Galaxies that are detected by both surveys are shown in blue, and the galaxies in the NSA but not in ALFALFA are shown as orange contours. The ALFALFA sample is dominated by blue galaxies whereas the NSA is dominated by red galaxies. (Right) Complementary comparison, showing the $g - i$ corrected color vs. stellar mass calculated following Taylor et al. (2011) for galaxies in common to both surveys (blue contours) and galaxies in ALFALFA but not in the NSA (light blue). When compared to the NSA galaxies, ALFALFA galaxies are bluer and include more lower-mass galaxies.

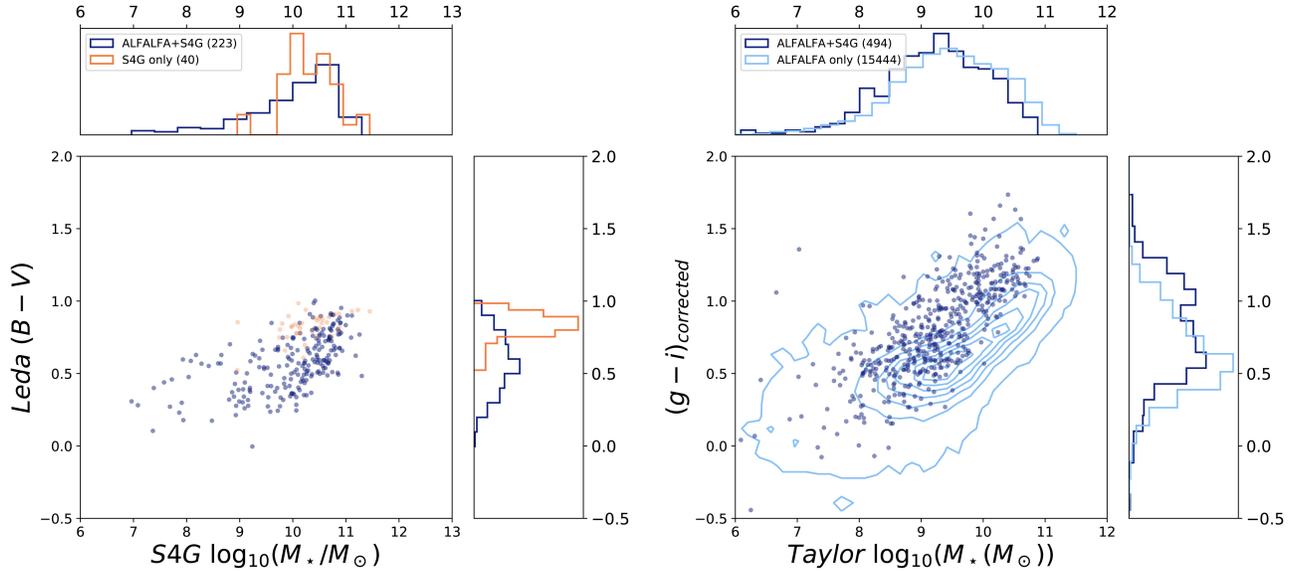


Figure 7. (Left) $B - V$ color vs. stellar mass (both from the S4G catalog) for galaxies that lie in a volume common to both the ALFALFA and S4G (Sheth et al. 2010) surveys. Galaxies that are detected by both surveys are shown in blue, and galaxies in the S4G but not in ALFALFA are shown in orange. (Right) Complementary comparison, showing the $g - i$ corrected color vs. stellar mass calculated following Taylor et al. (2011), for galaxies in common to both surveys (blue) and galaxies in ALFALFA but not in the S4G (light blue contours). When compared to the S4G galaxies, ALFALFA galaxies are bluer and have similar masses.

$$0^\circ < \text{decl.} < 35^\circ$$

This overlap region is outlined with an orange hatched rectangle in the bottom panel of Figure 5.

The matching statistics are reported in Table 3, and the colors and stellar masses are shown in Figure 8. For colors, we use SDSS $g - i$ model magnitudes for both ALFALFA and GSWLC-2 galaxies, corrected for both galactic and internal extinction as described in Section 2.2. The left panel uses stellar masses from the GSWLC-2 catalog while the right panel uses stellar mass values calculated using the Taylor et al. (2011) method, as explained in Section 2.3. The left panel compares galaxies found in both surveys (blue symbols) to

galaxies found in GSWLC-2 but not in ALFALFA (orange contours). The right panel compares galaxies found in both surveys (blue contours) to galaxies found in ALFALFA but not in GSWLC-2 (light blue symbols). Comparing the left and right panels of Figure 8, we see that galaxies in GSWLC-2 tend to be redder and more massive compared to the ALFALFA population.

For the GSWLC-2, we also compare the galaxy populations in terms of SFRs and sSFRs. In Figure 9, we plot SFR versus stellar mass for the GSWLC-2 and ALFALFA galaxies. The left panel of Figure 9 compares galaxies that are found in both ALFALFA and GSWLC-2 (blue symbols) to those found in GSWLC-2 but not in ALFALFA (orange contours). The

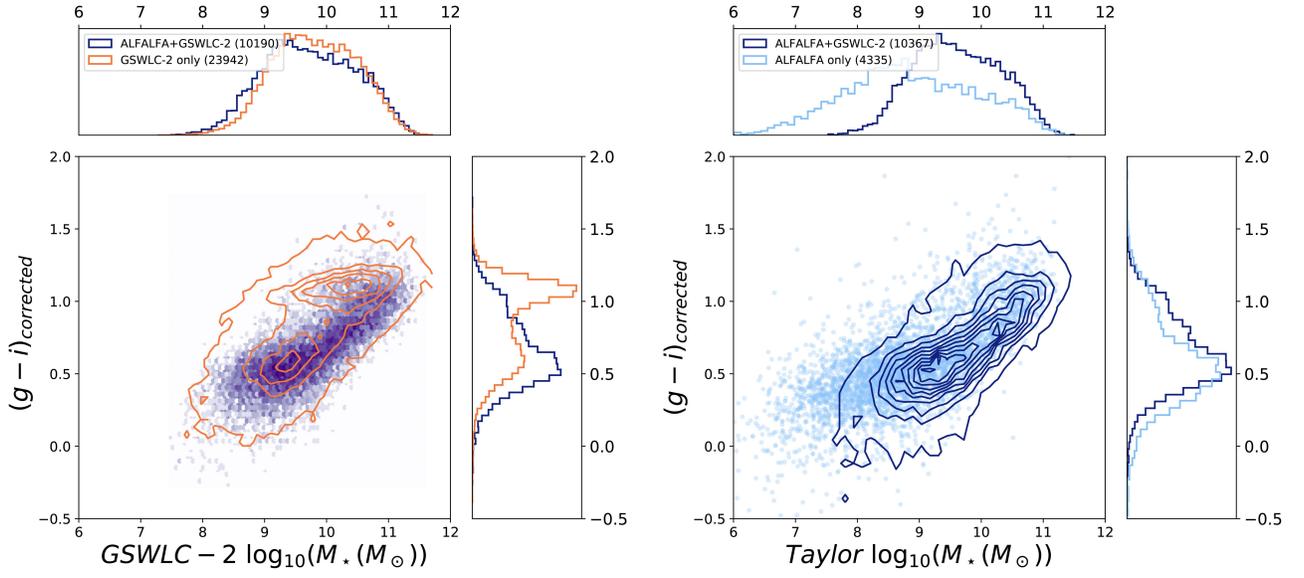


Figure 8. (Left) $g - i$ color vs. stellar mass from the GSWLC-2 catalog for galaxies that lie in a volume common to both the ALFALFA and GSWLC-2 (Salim et al. 2016) surveys. Galaxies that are detected by both surveys are shown in blue, and the galaxies in the GSWLC-2 but not in ALFALFA are shown as orange contours. The ALFALFA sample is dominated by blue galaxies whereas the GSWLC-2 is dominated by red galaxies. (Right) Complementary comparison, showing the $g - i$ corrected color vs. stellar mass calculated following Taylor et al. (2011) for galaxies in common to both surveys (blue contours) and galaxies in ALFALFA but not in the GSWLC-2 (light blue) as a function of stellar mass determined using the Taylor et al. (2011) method. When compared to the GSWLC-2 galaxies, ALFALFA includes more lower-mass (and slightly bluer) galaxies.

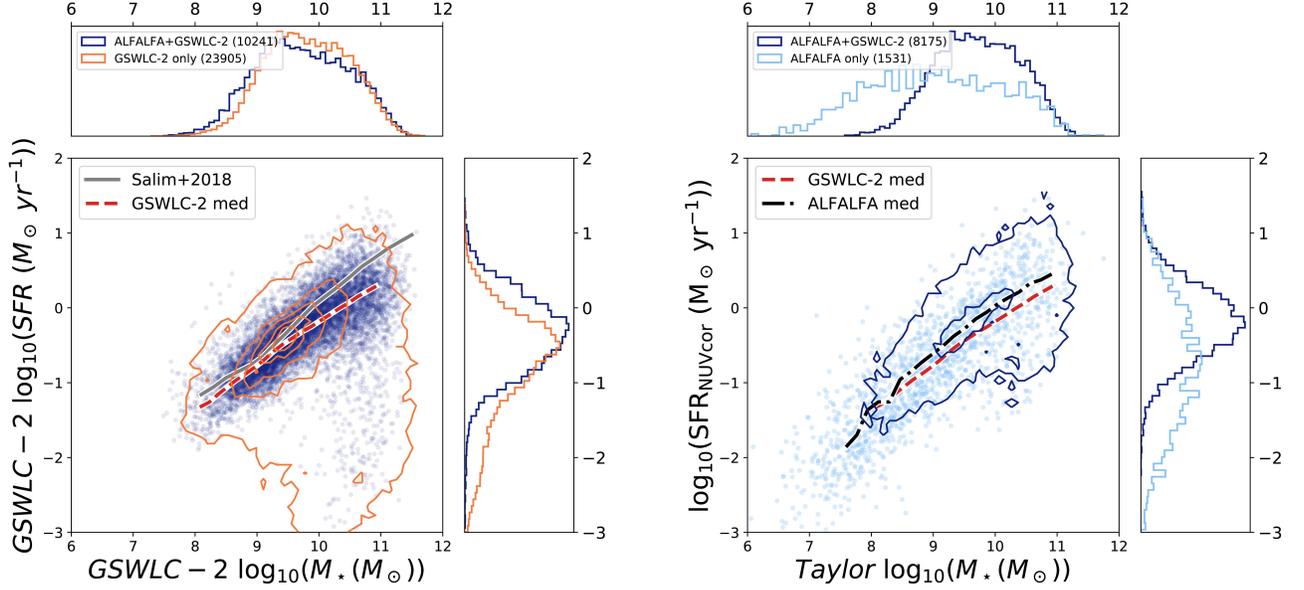


Figure 9. (Left) Star formation rate vs. stellar mass (both from the GSWLC-2 catalog) for galaxies in a common volume to both the ALFALFA and GSWLC-2 surveys. Galaxies that are detected in both surveys are shown in blue, and galaxies found in GSWLC-2 but not in ALFALFA are shown as orange contours. The red dashed line is the median for the star-forming GSWLC-2 galaxies in the overlap region and the gray line represents the fit to the full GSWLC-2 sample from Salim et al. (2018). The slopes are different for $\log_{10}(M_*/M_\odot) > 9.5$. (Right) Complementary comparison, showing the NUV-corrected star formation rate vs. stellar mass calculated following Taylor et al. (2011), for galaxies in common to both surveys (blue contours) and galaxies found in ALFALFA but not in GSWLC-2 (light blue). The red dashed line is the median for the GSWLC-2 galaxies in the overlap region, and the black dashed-dotted line is the median SFR vs. stellar mass for the ALFALFA galaxies. The star-forming main sequences for the GSWLC-2 and ALFALFA samples are similar but slightly shifted toward higher SFRs (see the text for discussion).

majority of the galaxies fall on the star-forming main sequence. The red dashed line in the left panel of Figure 9 shows the median SFR versus stellar mass for the star-forming GSWLC-2 galaxies in the overlap region, where we define star-forming as $\log_{10}(sSFR) > -11$ according to the criteria of Salim et al. (2018). For comparison, the gray line shows the main sequence derived from the full GSWLC-2 (Salim et al. 2018). We find that the slopes are significantly different for

$\log_{10}(M_*/M_\odot) > 9.5$. It is possible that the offset between the two lines is due to evolutionary effects between the ALFALFA galaxies and the higher-redshift galaxies that are more typical of the GSWLC-2 sample.

In the right panel, we again show the galaxies that are in both the ALFALFA and GSWLC-2 samples (blue contours), but now compare to the galaxies that are in ALFALFA but not in GSWLC-2 (light blue symbols). Here we use NUV-corrected

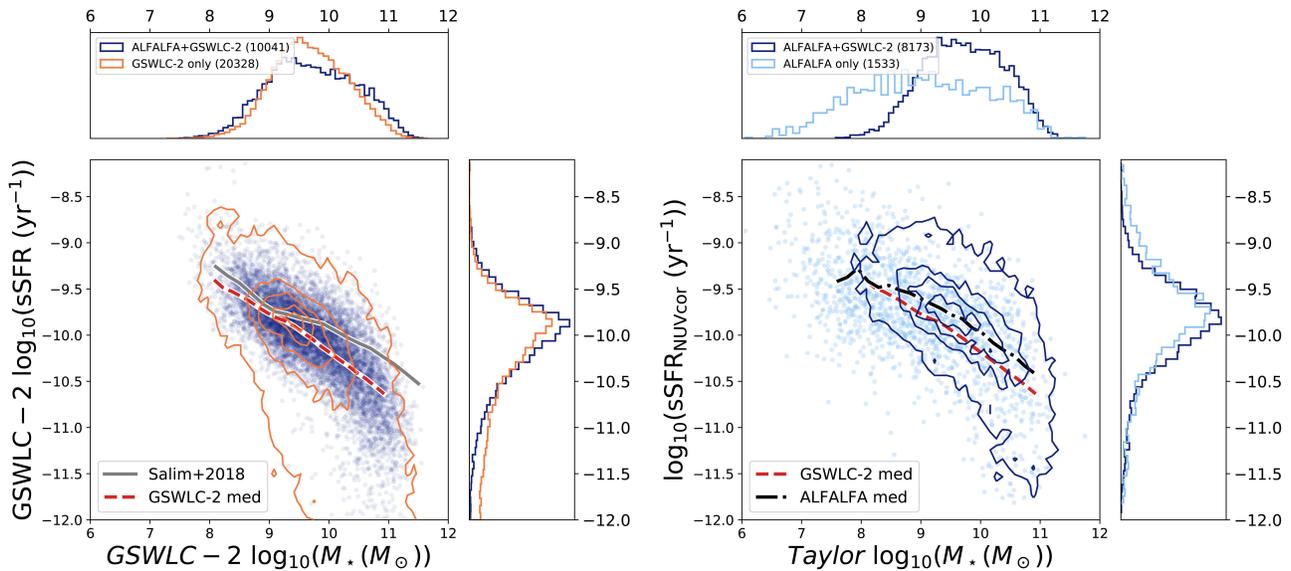


Figure 10. (Left) Specific star formation rate vs. stellar mass (both from the GSWLC-2 catalog) for galaxies in a volume common to both ALFALFA and GSWLC-2 surveys. Galaxies that are detected in both surveys are shown in blue, and galaxies in GSWLC-2 but not in ALFALFA are shown as orange contours. The red dashed line is the median for the GSWLC-2 galaxies in the overlap region and the gray line represents the fit to the full GSWLC-2 sample from Salim et al. (2018). The slopes are significantly different for $\log_{10}(M_*/M_\odot) > 9.5$. (Right) Complementary comparison, showing the NUV-corrected specific star formation rate vs. stellar mass calculated following Taylor et al. (2011) for galaxies in common to both surveys (blue contours) and galaxies in ALFALFA but not in GSWLC-2 (light blue). The red dashed line is the median for the GSWLC-2 galaxies in the overlap region, and the black dashed-dotted line is the median sSFR vs. stellar mass for the ALFALFA galaxies. The median sSFRs of the ALFALFA galaxies are similar to those of the GSWLC-2 galaxies but slightly shifted toward higher sSFRs (see text for discussion).

SFRs and the Taylor et al. (2011) stellar masses. We show the median SFR versus stellar mass for the ALFALFA star-forming galaxies ($\log_{10}(\text{sSFR}) > -11$) with the black dashed-dotted line, and we again show the median SFR for the star-forming GSWLC-2 galaxies with the red dashed line for comparison. We see that the star-forming main sequences for the GSWLC-2 and ALFALFA samples are similar. However, there is an offset toward higher SFRs. Higher SFRs for the ALFALFA galaxies are not surprising given that ALFALFA is a HI-selected sample.

Similarly, in Figure 10, we compare the sSFR versus stellar mass for GSWLC-2 and ALFALFA galaxies. The general trend is that as stellar mass increases, the sSFR decreases. In the left panel, we show the median sSFR versus stellar mass for the GSWLC-2 galaxies in our overlap comparison sample with the red dashed line. The gray line shows the relationship for the full GSWLC-2 (Figure 3 from Salim et al. 2018). The trends differ significantly at $\log_{10}(M_*/M_\odot) > 9.5$, and we again attribute the offset to the lower redshift-cut that we apply to the overlap sample. The right panel of Figure 10 shows the ALFALFA galaxies. The median sSFR of the ALFALFA galaxies (black dashed-dotted line) is close to the GSWLC-2 galaxies, but again the ALFALFA relation is offset toward higher sSFR.

Note that we have made no effort to correct for the incompleteness that we undoubtedly suffer in detecting galaxies with the lowest SFRs. A full analysis of completeness, both in terms of stellar mass and SFRs, is needed to constrain the behavior of the star-forming main sequence for dwarf galaxies. Such an analysis is beyond the scope of this paper.

It is worth noting that the different selection effects between the catalogs do not create very dramatic differences in the relationship between stellar mass and SFR.

6. Summary

We present a HI-optical catalog of matches between the recently completed 100% ALFALFA survey (ALFALFA) and SDSS. The ALFALFA-SDSS catalog contains 31,501 galaxies. We provide SDSS identifications for nearly all ALFALFA galaxies (29,638 galaxies) including 10,233 that do not have SDSS spectroscopy. The SDSS identifications in the ALFALFA-SDSS catalog can be used as a basis for further crossmatching with other surveys at other wavelengths. Most galaxies with SDSS photometry have uncertainties less than 0.05 mag in g and i (good photometry) and are assigned a code of “1” in our catalog (28,267 objects). Galaxies with uncertainties greater than 0.05 in g and/or i (bad photometry) are assigned a code of “2” (1371 objects). Galaxies with no clear SDSS counterpart are assigned a code of “0” if they are outside the SDSS footprint (1296 objects) and “3” otherwise (567 objects). We present the observed (Table 1) and derived properties (Table 2) for the entire ALFALFA sample, including absolute magnitude, color, stellar mass, HI mass, and SFR. In addition, we include magnitude-dependent internal extinction estimates that differ significantly from previous work and are better suited to the low-mass galaxies that dominate the ALFALFA sample.

We explore different methods to calculate the stellar mass based on SDSS optical (Taylor et al. 2011) and unWISE infrared photometry (Cluver et al. 2014; McGaugh & Schombert 2015). We find that the Taylor method using optical SDSS gives the best agreement with SED-derived stellar masses from the GSWLC-2. We also explore different methods to calculate the SFR using unWISE infrared and/or NUV GALEX photometry (Hao et al. 2011; Murphy et al. 2011; Kennicutt & Evans 2012). The corrected NUV SFR agrees most closely with the SFR estimates from the GSWLC-2 and should be used when available.

We place the ALFALFA-SDSS catalog in the context of three other galaxy catalogs that include stellar mass (NSA, GSWLC-2, and S4G) and SFR (GSWLC-2). In this way, we show how the ALFALFA-SDSS sample compares with other optically selected catalogs. We find that ALFALFA-SDSS galaxies are generally less massive and bluer. This is not surprising. Because ALFALFA detects the H I gas that fuels star formation, ALFALFA galaxies tend to be star-forming and blue. The sensitivity of ALFALFA allows for the detection of small galaxies within this population. Meanwhile, the constraints introduced by SDSS in the ALFALFA-SDSS Catalog are based on photometry only—not spectroscopy, which is limited to brighter galaxies on average.

We further compare the ALFALFA and GSWLC-2 samples in terms of the $\text{SFR}-M_*$ and $\text{sSFR}-M_*$ relations. The median relationships are similar, but we find evidence for a population of low-mass, low-sSFR galaxies in ALFALFA that are not included in the GSWLC-2. In addition, the slope of the $\text{sSFR}-M_*$ relation appears flatter for the ALFALFA sample. We emphasize that these comparisons are meant to show how the ALFALFA galaxy population differs overall from the populations in other catalogs because of their different selection effects. An analysis of the true scaling relations among galaxy properties requires a more thorough analysis of the completeness of each survey, which we reserve for future work.

We are grateful for the contributions of all members of the Undergraduate ALFALFA Team. The authors thank D. Lang and S. Salim for assistance with utilizing their catalogs. This work has been supported by NSF grants AST-1211005, AST-1637339, and AST-1637271. The ALFALFA team at Cornell has been supported by NSF grants AST-1107390 and AST-1714828 and by the Brinson Foundation. R.A.F. gratefully acknowledges support from NSF grants AST-0847430 and AST-1716657.

From 2011 to 2018, the Arecibo Observatory was operated by SRI International under a cooperative agreement with the National Science Foundation (AST-1100968) and in alliance with Ana G. Méndez-Universidad Metropolitana, and the Universities Space Research Association. Currently, the Arecibo Observatory is a facility of the National Science Foundation operated under cooperative agreement (AST-1744119) by the University of Central Florida in alliance with Universidad Ana G. Méndez (UAGM) and Yang Enterprises (YEI), Inc.

Funding for the SDSS and SDSS-II has been provided by the Alfred P. Sloan Foundation, the participating institutions, the National Science Foundation, the US Department of Energy, the NASA, the Japanese Monbukagakusho, the Max Planck Society, and the Higher Education Funding Council for England. The SDSS website is <http://www.sdss.org/>. The SDSS is managed by the Astrophysical Research Consortium for the participating institutions.

Funding for the NASA-Sloan Atlas has been provided by the NASA Astrophysics Data Analysis Program (08-ADP08-0072) and the NSF (AST-1211644).

Funding for SDSS-III has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, and the US Department of Energy. The SDSS-III website is <http://www.sdss3.org>. SDSS-III is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS-III Collaboration

including the University of Arizona, the Brazilian Participation Group, Brookhaven National Laboratory, University of Cambridge, University of Florida, the French Participation Group, the German Participation Group, the Instituto de Astrofísica de Canarias, the Michigan State/Notre Dame/JINA Participation Group, Johns Hopkins University, Lawrence Berkeley National Laboratory, Max Planck Institute for Astrophysics, New Mexico State University, New York University, Ohio State University, Pennsylvania State University, University of Portsmouth, Princeton University, the Spanish Participation Group, University of Tokyo, University of Utah, Vanderbilt University, University of Virginia, University of Washington, and Yale University.

The Galaxy Evolution Explorer (GALEX) is a NASA Small Explorer. The mission was developed in cooperation with the Centre National d'Etudes Spatiales of France and the Korean Ministry of Science and Technology.

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.

This research has made use of the NASA/IPAC Infrared Science Archive, which is funded by the National Aeronautics and Space Administration and operated by the California Institute of Technology.

The construction of GSWLC was funded through NASA award NNX12AE06G.

This research made use of Astropy,⁹ a community-developed core Python package for Astronomy (Astropy Collaboration et al. 2013, 2018), matplotlib (Hunter 2007), and TOPCAT (Taylor 2005).

Facilities: Arecibo, IRSA, Spitzer, WISE, GALEX.

Software: astropy, matplotlib, TOPCAT.

ORCID iDs

Adriana Durbala  <https://orcid.org/0000-0002-3406-5502>

Rose A. Finn  <https://orcid.org/0000-0001-8518-4862>

Mary Crone Odekon  <https://orcid.org/0000-0003-0162-1012>

Martha P. Haynes  <https://orcid.org/0000-0001-5334-5166>

Rebecca A. Koopmann  <https://orcid.org/0000-0002-3144-2501>

Aileen A. O'Donoghue  <https://orcid.org/0000-0003-1664-6255>

References

- Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, *AJ*, 156, 123
- Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, *A&A*, 558, A33
- Balogh, M. L., Baldry, I. K., Nichol, R., et al. 2004, *ApJL*, 615, L101
- Blanton, M. R., Eisenstein, D., Hogg, D. W., Schlegel, D. J., & Brinkmann, J. 2005, *ApJ*, 629, 143
- Blanton, M. R., Kazin, E., Muna, D., Weaver, B. A., & Price-Whelan, A. 2011, *AJ*, 142, 31
- Blanton, M. R., & Roweis, S. 2007, *AJ*, 133, 734
- Brinchmann, J., Charlot, S., White, S. D. M., et al. 2004, *MNRAS*, 351, 1151
- Brough, S., Croom, S., Sharp, R., et al. 2013, *MNRAS*, 435, 2903
- Catinella, B., Saintonge, A., Janowiecki, S., et al. 2018, *MNRAS*, 476, 875
- Chen, Y.-M., Kauffmann, G., Tremonti, C. A., et al. 2012, *MNRAS*, 421, 314
- Cluver, M. E., Jarrett, T. H., Hopkins, A. M., et al. 2014, *ApJ*, 782, 90

⁹ <http://www.astropy.org>

- Condon, J., Cotton, W., E. W., G., et al. 1998, *AJ*, 115, 1693
- Devour, B. M., & Bell, E. F. 2016, *MNRAS*, 459, 2054
- Fossati, M., Gavazzi, G., Savorgnan, G., et al. 2013, *A&A*, 553, A91
- Gavazzi, G., Fumagalli, M., Fossati, M., et al. 2013, *A&A*, 553, A89
- Giovanelli, R., Haynes, M. P., Kent, B. R., et al. 2005, *AJ*, 130, 2598
- Giovanelli, R., Haynes, M. P., Kent, B. R., et al. 2007, *AJ*, 133, 2569
- Giovanelli, R., Haynes, M. P., Salzer, J. J., et al. 1994, *AJ*, 107, 2036
- Green, G. M., Schlafly, E. F., Finkbeiner, D. P., et al. 2015, *ApJ*, 810, 25
- Green, G. M., Schlafly, E. F., Zucker, C., Speagle, J. S., & Finkbeiner, D. P. 2019, *ApJ*, 887, 93
- Hao, C.-N., Kennicutt, R. C., Johnson, B. D., et al. 2011, *ApJ*, 741, 124
- Haynes, M. P., Giovanelli, R., Kent, B. R., et al. 2018, *ApJ*, 861, 49
- Haynes, M. P., Giovanelli, R., Martin, A. M., et al. 2011, *AJ*, 142, 170
- Huang, S., Haynes, M. P., Giovanelli, R., & Brinchmann, J. 2012, *ApJ*, 756, 113
- Hunter, J. D. 2007, *CSE*, 9, 90
- Jarrett, T. H., Cohen, M., Masci, F., et al. 2011, *ApJ*, 735, 112
- Jarrett, T. H., Masci, F., Tsai, C. W., et al. 2013, *AJ*, 145, 6
- Jones, M. G., Haynes, M. P., Giovanelli, R., & Moorman, C. 2018, *MNRAS*, 477, 2
- Kauffmann, G., Heckman, T. M., White, S. D. M., et al. 2003, *MNRAS*, 341, 33
- Kennicutt, R. C., & Evans, N. J. 2012, *ARA&A*, 50, 531
- Kent, B. R., Giovanelli, R., Haynes, M. P., et al. 2008, *AJ*, 136, 713
- Kourkchi, E., Tully, R. B., Neill, J. D., et al. 2019, *ApJ*, 884, 82
- Lang, D. 2014, *AJ*, 147, 108
- Lang, D., Hogg, D. W., & Schlegel, D. J. 2016, *AJ*, 151, 36
- Leisman, L., Haynes, M. P., Janowiecki, S., et al. 2017, *ApJ*, 842, 133
- Leroy, A. K., Sandstrom, K. M., Lang, D., et al. 2019, *ApJS*, 244, 24
- Maraston, C., Pforr, J., Henriques, B. M., et al. 2013, *MNRAS*, 435, 2764
- Martin, A. M., Giovanelli, R., Haynes, M. P., & Guzzo, L. 2012, *ApJ*, 750, 38
- Masters, K. L., Nichol, R., Bamford, S., et al. 2010, *MNRAS*, 404, 792
- McGaugh, S. S., & Schombert, J. M. 2015, *ApJ*, 802, 18
- McGaugh, S. S., Schombert, J. M., Bothun, G. D., & de Blok, W. J. G. 2000, *ApJL*, 533, L99
- Moorman, C. M., Vogeley, M. S., Hoyle, F., et al. 2014, *MNRAS*, 444, 3559
- Mouhcine, M., Baldry, I. K., & Bamford, S. P. 2007, *MNRAS*, 382, 801
- Muñoz-Mateos, J. C., Sheth, K., Gil de Paz, A., et al. 2013, *ApJ*, 771, 59
- Murphy, E. J., Condon, J. J., Schinnerer, E., et al. 2011, *ApJ*, 737, 67
- Papastergis, E., Adams, E. A. K., & van der Hulst, J. M. 2016, *A&A*, 593, A39
- Papastergis, E., Giovanelli, R., Haynes, M. P., Rodríguez-Puebla, A., & Jones, M. G. 2013, *ApJ*, 776, 43
- Papastergis, E., Martin, A. M., Giovanelli, R., & Haynes, M. P. 2011, *ApJ*, 739, 38
- Planck Collaboration, Abergel, A., Ade, P. A. R., et al. 2014, *A&A*, 571, A11
- Querejeta, M., Meidt, S. E., Schinnerer, E., et al. 2015, *ApJS*, 219, 5
- Rieke, G. H., Alonso-Herrero, A., Weiner, B. J., et al. 2009, *ApJ*, 692, 556
- Saintonge, A. 2007, *AJ*, 133, 2087
- Saintonge, A., Catinella, B., Tacconi, L. J., et al. 2017, *ApJS*, 233, 22
- Salim, S., Boquien, M., & Lee, J. C. 2018, *ApJ*, 859, 11
- Salim, S., Lee, J. C., Janowiecki, S., et al. 2016, *ApJS*, 227, 2
- Schlafly, E. F., & Finkbeiner, D. P. 2011, *ApJ*, 737, 103
- Schlafly, E. F., Meisner, A. M., & Green, G. M. 2019, *ApJS*, 240, 30
- Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, *ApJ*, 500, 525
- Shao, Z., Xiao, Q., Shen, S., et al. 2007, *ApJ*, 659, 1159
- Sheth, K., Regan, M., Hinz, J. L., et al. 2010, *PASP*, 122, 1397
- Taylor, E. N., Hopkins, A. M., Baldry, I. K., et al. 2011, *MNRAS*, 418, 1587
- Taylor, M. B. 2005, in *ASP Conf. Ser. 347, Astronomical Data Analysis Software and Systems XIV*, ed. P. Shopbell, M. Britton, & R. Ebert (San Francisco, CA: ASP), 29
- Tully, R. B., & Fisher, J. R. 1977, *A&A*, 500, 105
- Tully, R. B., Pierce, M. J., Huang, J.-S., et al. 1998, *AJ*, 115, 2264
- Wright, E. L., Eisenhardt, P. R. M., Mainzer, A. K., et al. 2010, *AJ*, 140, 1868