# SED Fitting

Marie-Joëlle Gingras

1. INTRODUCTION

We are looking into various methods to infer galaxy properties, especially stellar mass and redshift, from photometric observations. The photometric observations are a combination of Sloan Digital Sky Survey (SDSS) (Blanton et al. 2005; Adelman-McCarthy et al. 2008; Padmanabhan et al. 2008) and Wide-field Infrared Survey Explorer (WISE) (Wright et al. 2010) photometry compiled by Chang et al. (2015). The galaxies' flux densities are known for (up to) 9 photometric bands, SDSS ugriz and WISE w1, w2, w3 and w4. Throughout our analysis, the "known" galaxy properties are taken to be the values listed in Chang et al. (2015).

2. SED FITTING

We first attempt to infer galaxy properties by fitting different galaxy models to the observed photometric data. Combining the different photometric bands, we obtain a spectral energy distribution (SED) which shows the energy emitted as a function of wavelength. For a chosen model, the resulting SED is then fitted to obtain the best model parameters that can describe the observed SED. We tested different analysis techniques and different models.

Fig.1 compares the observed photometric flux densities (black) with the photometric values (red) obtained from the bestfit model (shown as the red spectrum) for Galaxy #33 using emcee+Prospector. emcee is a Markov Chain Monte Carlo Ensemble Sampler and Prospector is a handy package to fit photometric and/or
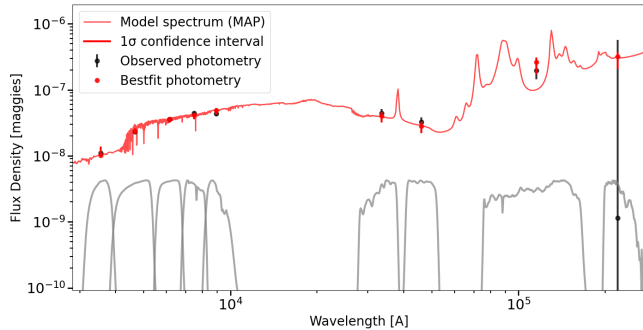
Figure 1. SED of Galaxy # 33. The observed photometry is shown in black, while the red points show the photometry obtained from the bestfit model, shown as the red spectrum, and the red error bars show the 1σ confidence interval. The grey bumps show the photometric bands corresponding to each photometric point.

spectroscopic data using various galaxy/stellar population models (Johnson et al. 2021; Leja et al. 2017). In Fig. 1, the red error bars show the 1σ confidence interval. For the second fitted data point, the bestfit does not lie within the 1σ confidence interval. This is not problematic as all the photometry points are being fitted together such that the bestfit can be obtained even if one point is not in its 1σ confidence interval.

We mostly use two variations of Prospector's parametric SFH (star formation history) model "parametric_sfh" for a delay-tau SFH, where SFR$\sim t_{age}e^{-t_{age}/\tau}$ (SFR: star formation rate). The basic model has 5 free parameters: Stellar mass (M_*), Metallicity, in solar units ($Z/Z_\odot$), Dust optical depth (κ), Star formation rate decay timescale (τ) and Star formation rate time (t_{age}). When exploring parameter space, we take the Log_{10} of M_*, $Z/Z_\odot$ and τ. We also use a variation of this model where the galaxy's redshift, z, is left as a free parameter. Otherwise, the redshift is fixed to its known spectroscopic value from SDSS.

For each model (free or fixed redshift), we tested three fitting methods/packages:

- Prospector's implementation of emcee (Foreman-Mackey et al. 2013): emcee is an affine-invariant Markov Chain Monte Carlo (MCMC) ensemble sampler which samples probability distributions.

- Dynesty (Speagle 2020; Koposov et al. 2022): Dynesty is a nested sampling package which can be used to estimate Bayesian posteriors and evidences.

- Nautilus (Lange 2023): Nautilus is a package which combines importance nested sampling and neural network-based algorithms to accurately and efficiently Bayesian posteriors and evidences.

All three methods aim to solve Bayesian inference problems and characterise the target probability (i.e the posterior distribution of our parameters given the data). The posterior distribution, $P(x|y)$, is:

$$P(x|y) = \frac{p(y|x)p(x)}{p(y)} \tag{1}$$

where $p(y|x)$ is the likelihood of x given y (or vice-versa), p(x) is prior probability (i.e the probability of observing x with no other conditions) and p(y) is the probability of y with no other conditions, also called the evidence.

The evidence can also be expressed as an integral over all possible x values such that: $\mathcal{Z} = p(y) = \int p(y|x)p(x)dx$. Instead of the likelihood, the log of the likelihood (log-likelihood) can be used, which is:

$$ln\mathcal{L} = -0.5\sum_i \left(\left(\frac{y_i - \text{model}(p, x_i)}{\sigma_i} \right)^2 + ln\left(\sigma_i^2\right) \right) \tag{2}$$

where in our case, for each photometric band "i", y_i is the observed flux density, σ_i is its uncertainty, and model(p, x_i) is the flux density predicted by our model with parameter values p at a wavelength x_i.

All three methods require us to choose priors. In this work, we always use flat-priors on our parameters or on the log of our parameters. This means that we set bounds on the possible values of our parameters, where the prior probability is 1 within some chosen range and 0 outside of it. The choices of priors are quite important since: 1) Very wide priors increase computation time 2) Narrow priors can exclude the bestfit parameters from the "accepted" parameter values.

All three methods aim to find highest likelihood regions in the parameter space and then sample the posterior probability distribution. Dynesty and Nautilus not only aim to estimate the posterior distribution but also the evidence.

emcee samples the posterior distribution using MCMC, where numerous walkers explore the parameter space in parallel, moving randomly but favourising positions with higher likelihood, by assigning them greater acceptance probabilities. These acceptance probabilities only depend on the prior probability and the likelihood.

Dynesty uses nested sampling to estimate the Bayesian evidence and to estimate the posterior distribution. Nested sampling estimates the evidence ($\mathcal{Z}$) by integrating the prior, p(x), using shells of constant likelihood:

1. Start with N points sampled from the prior such that all points lie in the prior volume.

2. Remove the point with the smallest likelihood, and replace it with another point sampled from the prior which MUST have a higher likelihood.

3. This reduces the prior volume to where the N points lie. The shell volume associated with the discarded point (ΔV_i) is the change in prior volume.

4. The evidence is estimated by summing the weight (i.e $w_i = \mathcal{L}_i \Delta V_i$) of every discarded point "i".

5. Redo steps 1-4, until the convergence criterion is met.
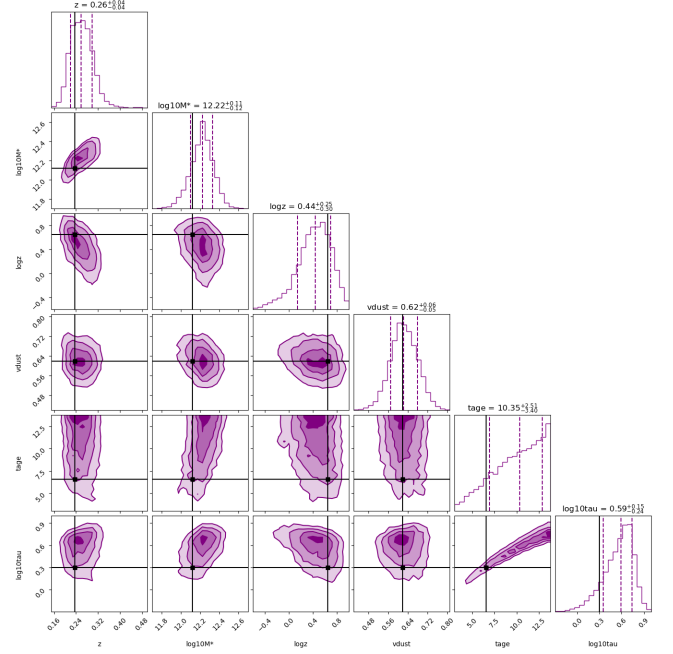
Figure 2. Corner plot obtained from fitting photometric observations from A1835 using the emcee method for a parametric delay-tau SFH model with a free redshift. The black lines show the bestfit values of the parameters and the dotted lines show the 0.16, 0.5 and 0.84 quantiles.

6. The posterior is estimated by taking the samples and weighting them by their corresponding weight w_i.

Since the weight of samples depends on their likelihood, nested sampling, contrary to MCMC, does not require a burn-in since unlikely initial points have extremely low importance. Nautilus is similar to Dynesty during the exploratory phase. However, Nautilus uses neural network-based algorithms to find the most efficient updated volume boundaries. Another difference between the two methods is that Nautilus uses Importance Nested Sampling, instead of Nested Sampling, which can increase the accuracy of the evidence estimate by an order of magnitude.

Initially, we worked with Galaxy #33 but we later changed to Galaxy # 83613 which is known as Abell 1835 or A1835. From now on, in order to compare the different methods, A1835 (Galaxy # 83613) will be used.

Figs. 2, 3, and 4 show the 1D and 2D projections of the posterior probability distribution (corner plot) for the parameters of the parametric delay-tau SFH model, with free redshift, fitted using emcee, Dynesty and Nautilus, respectively. All three analysis methods give very similar results, as shown in the overlaid corner plots in Fig. 5. The solid lines in Fig. 5 show the z and M_* from Chang et al. (2015). While our redshift values agree
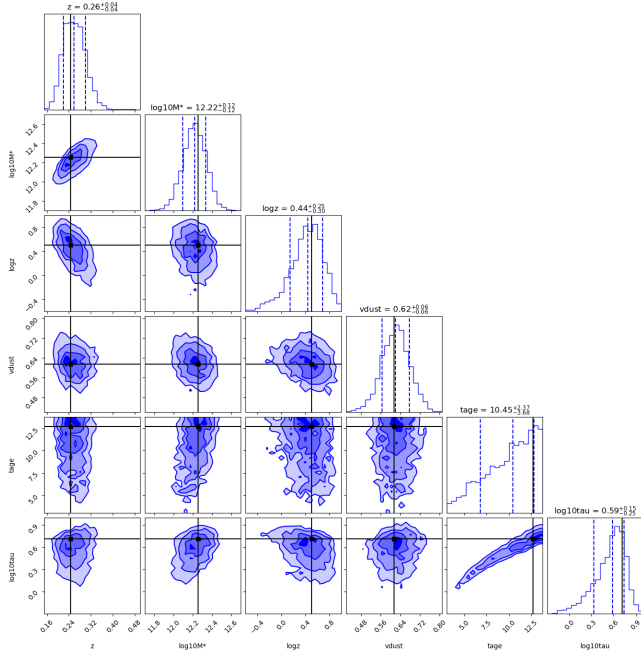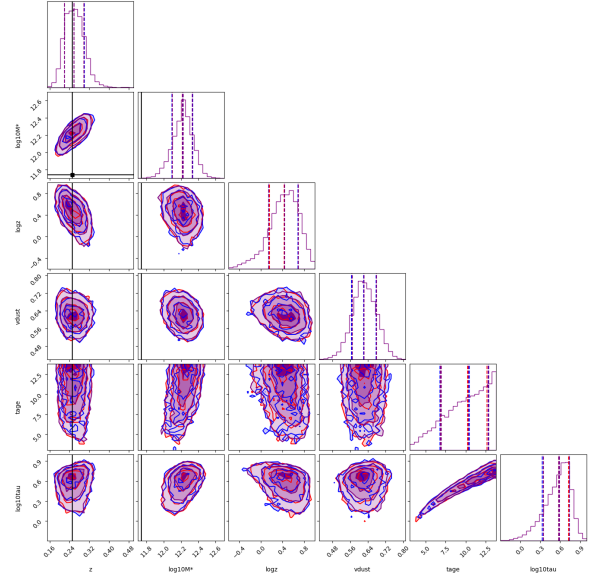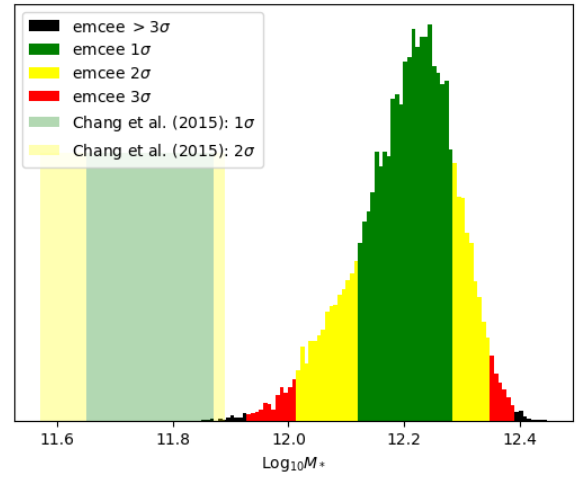
Figure 3. Same as Fig. 2 but using the Dynesty package
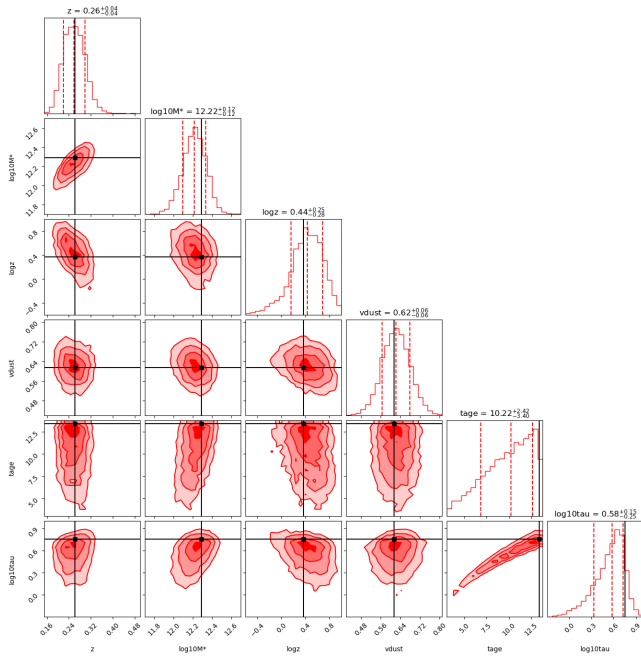
**Figure 4.** Same as Fig. 2 but using the Nautilus package

Figure 5. Comparing corner plots from all three methods by combining Fig. 2, 3 and 4. The black lines show the redshift and stellar mass values from Chang et al. (2015).


Figure 6. Stellar mass distribution of the samples obtained using emcee to fit the galaxy model with free redshift. The confidence intervals for 1σ, 2σ, 3σ and more than 3σ are shown in green, yellow, red and black, respectively. The 1σ and 2σ confidence intervals for the fit by Chang et al. (2015) are shown by the shaded green and yellow regions, respectively.

with Chang et al. (2015) within less than 1σ, the stellar mass values obtained from our fits are almost 4σ away from the Chang et al. (2015) value. This discrepancy is clearly seen in Fig. 6, where we observe very little overlap between the stellar mass distribution obtained from the emcee chains and the 2σ confidence interval from Chang et al. (2015).

The corner plots in Fig. 2- 4, as well as the correlation matrix in Fig. 7, show some degeneracies between the fitted parameters for the redshift and the stellar mass and for t_{age} and τ. Stellar mass is correlated with redshift because objects further away need to be more massive to be detected (assuming a mass or flux limited survey) than closer objects. The degeneracy between t_{age} and τ
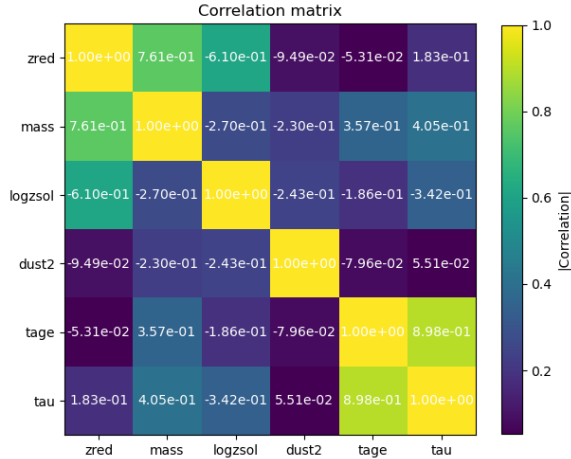
Figure 7. Correlation matrix for the free parameters from the emcee fit corresponding to the corner plot shown in Fig. 2.
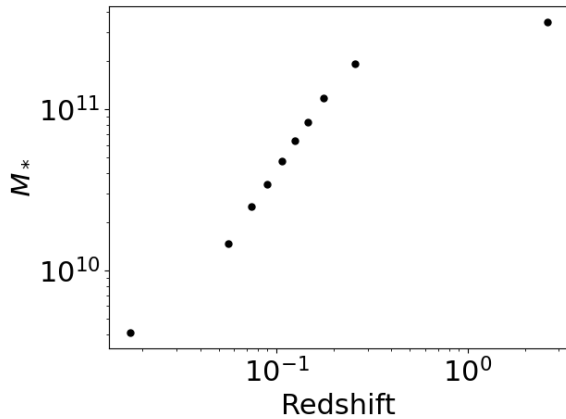
Figure 8. Comparison between the redshift and stellar masses of galaxies using the sample from Chang et al. (2015). The galaxies are separated in 10 redshift bins. Jackknife is used to approximate the uncertainties on the mean stellar masses. The errors are very small and therefore not seen in the plot.

means that older galaxies have a slower decay in their SFR.

Figure 8 shows the relation between redshift and stellar mass for the galaxy sample from Chang et al. (2015). The galaxies are separated in 10 redshift bins (which are not independant). We used Jackknife to approximate the uncertainties on the mean stellar masses. We used 5000 Jackknife subsets such that our errors are close to the "analytic" errors (i.e $\sqrt{N}$, where N is the number of galaxies in each bin) although Jackknife errors are generally larger than analytic errors. If too few subsets are used, the Jackknife errors get very big and overestimate the "true" errors. On the other hand, if too many Jackknife subsets are used, the sample for each iteration of the Jackknife will basically be the same such that the errors will be artificially small.

Table 1 lists the bestfit parameters for each method as well as their reduced χ^2 and Bayesian evidence. As seen in Fig. 2- 4, the bestfit parameters are not always close to the median of their marginalized 1D distribution, sometimes they do not even lie within the 1σ confidence interval. This can happen when a very specific combination of parameters has a sudden increase in likelihood compared to its surroundings (in parameter space). As the bestfit is the one combination of parameters with largest likelihood, it is more prone to sudden changes than the median of the marginalized distributions.

Table 1 shows that the best reduced χ^2 is obtained when using emcee. Reduced $\chi^2 = \frac{1}{\nu} \left(\frac{y_i - \mathrm{model}(p_{\mathrm{bf}}, x_i)}{\sigma_i} \right)^2$, where ν is the number of degrees of freedom and p_{bf} are the bestfit parameters. While emcee has the best reduced χ^2, the Dynesty and Nautilus methods have significant benefits over emcee. While plotting the chains obtained from emcee and looking at the evolution of their scatter can help figure out if the fit has converged, the evolution (with computational time) of the evidence and/or the importance weight PDF obtained by Dynesty and Nautilus is a clear indicator of fit convergence. Fig. 9 shows a figure which can be easily made using Dynesty. Also, since Dynesty and Nautilus estimate the evidence, we can use the Bayes factor to compare different models. Since Nautilus uses Importance Nested Sampling, it obtains a more accurate evidence than Dynesty. It is also significantly faster than either Dynesty or emcee. Overall, Nautilus seems to be the best method for SED fitting.

Fig. 10- 12 show the fits obtained when A1835's redshift is fixed to its known SDSS value when fitting our galaxy model using emcee, Dynesty and Nautilus, respectively. Once again, Fig. 13 shows that the results for the different methods agree together. Using the Bayesian evidence (listed in Table 1 and 2) to compare our model with and without fixing the redshift, we obtain that the fixed-z model is ~ 5 times more likely for Dynesty and ~ 3 times more likely for Nautilus than for the free-z model. Comparing Fig. 5 and Fig. 13, we do not see significant differences in the parameter distributions.

3. MACHINE LEARNING (ML)

Instead of using a galaxy/SFH model to infer galaxy properties, we look into using machine learning. As for

Table 1. SED fitting bestfit parameters for the model with redshift as a free parameter

Method	z	$\mathrm{Log_{10}(M_*/M_\odot)}$	$\mathrm{Log_{10}(Z/Z_\odot)}$	κ	t_{age} (Gyr)	τ (Gyr^{-1})	Reduced χ^2	Bayesian evidence ($\times 10^{60}$)
emcee	0.24	12.12	0.65	0.62	6.58	0.30	1.44	
Dynesty	0.25	12.26	0.50	0.62	12.75	0.71	6.33	1.45
Nautilus	0.26	12.29	0.36	0.61	13.45	0.76	6.24	1.48

Table 2. SED fitting bestfit parameters for the model with redshift fixed to the value from SDSS

Method	$\mathrm{Log_{10}(M_*/M_\odot)}$	$\mathrm{Log_{10}(Z/Z_\odot)}$	κ	t_{age} (Gyr)	τ (Gyr^{-1})	Reduced χ^2	Bayesian evidence ($\times 10^{60}$)
emcee	12.15	0.55	0.62	6.38	0.29	1.09	
Dynesty	12.28	0.45	0.59	13.62	0.75	4.05	7.35
Nautilus	12.29	0.48	0.60	13.26	0.73	3.82	5.66
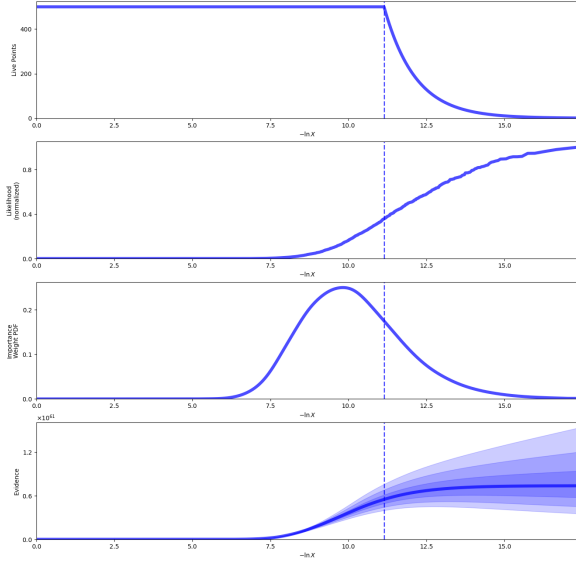
Figure 9. Output figure from Dynesty showing, from top to bottom, the number of live points, the likelihood, the importance weight probability distribution function (PDF) and the evidence. The clear peak (goes up and later down) in the importance weight PDF and the flattening of the evidence shows that the fit has converged.

the SED fitting, we use the same flux densities as our input parameters. In this case, we do not use their uncertainties. Contrary to the SED fitting method, we use observations from thousands, even tens of thousands of galaxies from the Chang et al. (2015) sample. We require that our sample of galaxies have observations for all 9 photometric bands as well as having physical output
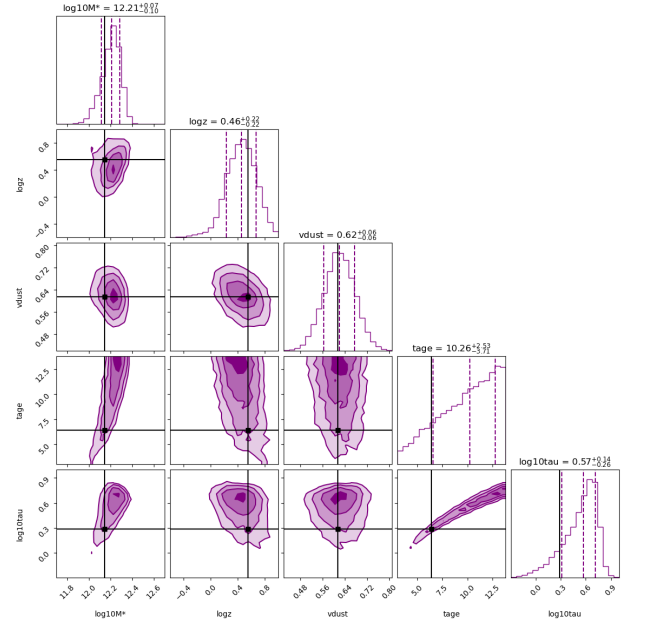
Figure 10. Corner plot obtained from fitting the SED of A1835 using the emcee method for a parametric delay-tau SFH model while fixing the redshift to the spectroscopically measured SDSS redshift (z=0.252). The black lines show the bestfit values of the parameters and the dotted lines show the 0.16, 0.5 and 0.84 quantiles.
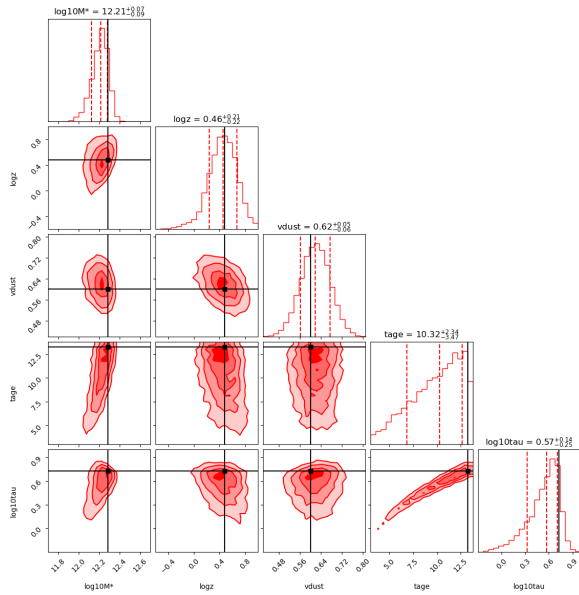
parameters when required. Throughout our analysis, we normalize our data to avoid biases in our results caused by differences in parameters' scales/orders of magnitude.

To visualize our data (as well making our dataset less complex and reducing computation times), we reduced
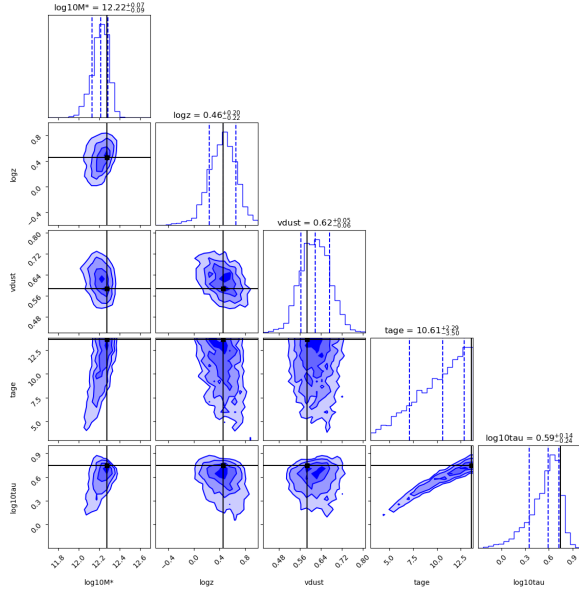
**Figure 11.** Same as Fig. 10 but using the Dynesty package.

Figure 12. Same as Fig. 10 but using the Nautilus package.

the dimensionality of our parameter space from 9 parameters to 2. We considered two methods.

1. Principal Component Analysis (PCA): PCA is a method that reduces the dimensionality of a dataset by finding new, more meaningful, dimensions. PCA looks at the data space and finds the axes with the most variance. All the principal components are orthogonal.
   Pros: Simple to implement. Fast computation when obtaining few dimensions. Identifies the most important variables/components. Useful for
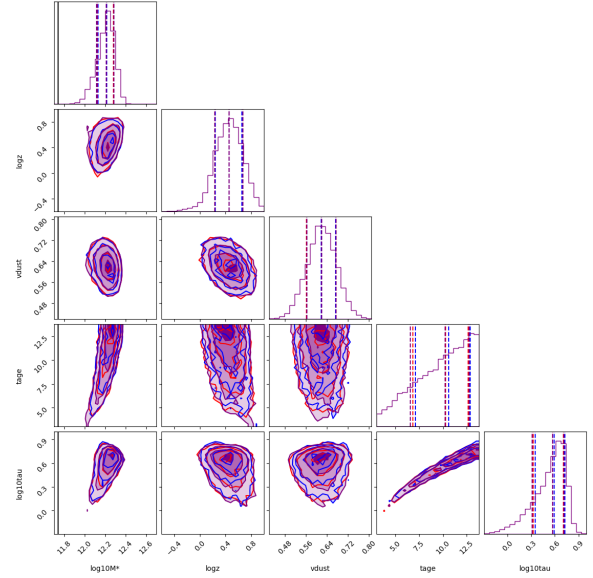
Figure 13. Comparing corner plots from all three methods by combining Fig. 10, 11 and 12. The black lines show the stellar mass value from Chang et al. (2015).

noise reduction.
Cons: Requires linear relations such that it will miss any non-linear structures.

2. Uniform Manifold Approximation and Projection (UMAP): It is a non-linear dimensionality reduction technique. "UMAP is constructed from a theoretical framework based in Riemannian geometry and algebraic topology." - McInnes et al. (2018)
   Pros: It can capture both local and global structures. Can be used to find non-linear relationships in your dataset. Great method to use for clustering and data visualization.
   Cons: Computationally expensive for large datasets. Sensitive to the choice hyperparameters.

For our dimensionality reduction, since we do not expect our data to have linear relationships and that we are specifically looking into clustering, we choose to use UMAP over PCA.

One significant downside with using UMAP is that it is very dependent on the choice of hyperparameters such as the number of neighbours, the random seed, the effective minimum distance between embedded points and the method used to initialize the low dimensional embedding. We opted to use PCA as our method to initialize the low dimensional embedding, applying the first two PCA components to our input data. Since PCA is another dimensionality reduction technique, we chose to use it for initializing the input data. We use 10 neighbours, although taking 5 or 20 neighbours does not significantly change our results. Figure 14 shows the

distribution of the 10 000 galaxies from our sample with the x and y axis showing the two new dimensions. The data is not well separated into clusters. There is a main cluster and one or two tiny clusters in the upper left. The tiny clusters correspond to galaxies where the fit from Chang et al. (2015) did not work properly.

While UMAP was not successful in making distinct clusters, Fig. 14 shows patterns in the UMAP associated with stellar mass, redshift and dust attenuation. Overall, stellar masses increase as the value of the y-axis increases, although some lower stellar mass galaxies lie at the top of the UMAP. The southern border of the UMAP has lower redshift galaxies and the galaxies become more redshifted further from it. The top left clumps are mostly made of higher redshift galaxies. In general, more redshifted galaxies are towards the top left of the map. Dust attenuation roughly increases along the y-axis.

Testing unsupervised machine learning on our compressed data, we tried different clustering methods. As the distribution of our points has an unusual shape, we elected to use a clustering method capable of dealing with such data such as spectral clustering, DBSCAN or HDBSCAN. In an attempt to reduce human input, the spectral clustering method was discarded as it requires the user to decide the number of clusters. Finally, we decided to use HDBSCAN as it can find clusters with varying densities.
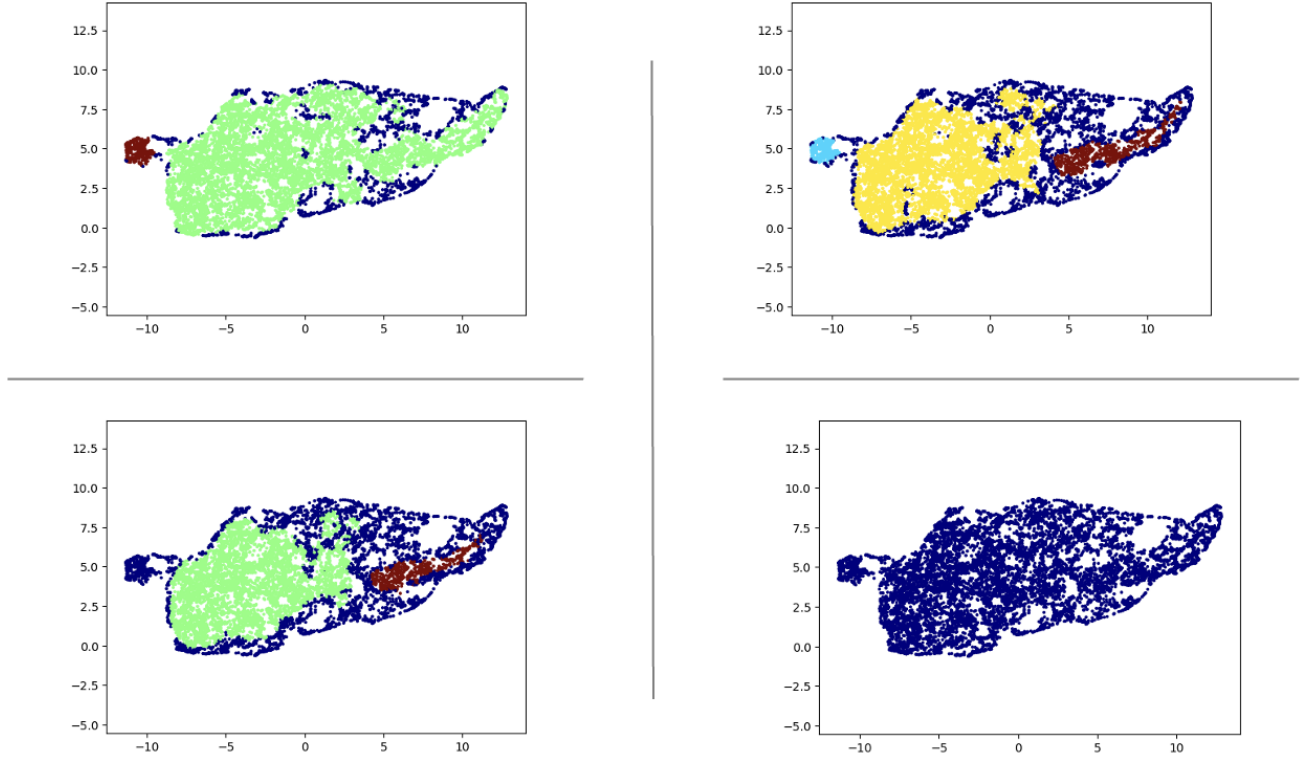
Fig. 15 shows the clusters identified by HDBSCAN for our compressed data. The different panels show how the choice of hyperparameters can signficantly affect our results. The different panels show the results for various minimum cluster size: Top left panel (100), Top right panel (200), Bottom left panel (300) and Bottom right panel (500). We are unable to make distinct clusters based on galaxy properties, regardless of the minimum cluster size chosen. When the minimum cluster size is less or equal to 200, the left clump is correctly identified as different than the rest of the data. Other clusters identified do not show significant differences in their galaxy properties. In an attempt to get better results, we tried using the uncompressed data instead of the lower dimension data. This did not yield better results.

As Chang et al. (2015) contains photometric observations as well as redshift and stellar masses for more than 800 000 galaxies, we attempted to use supervised machine learning to train our algorithm to predict galaxies' stellar masses and redshifts. We use a Multi-Layer Perceptron (MLP) regressor neural network model since it is easy to implement and can learn non-linear functions. As redshift and stellar mass are continuous vari-
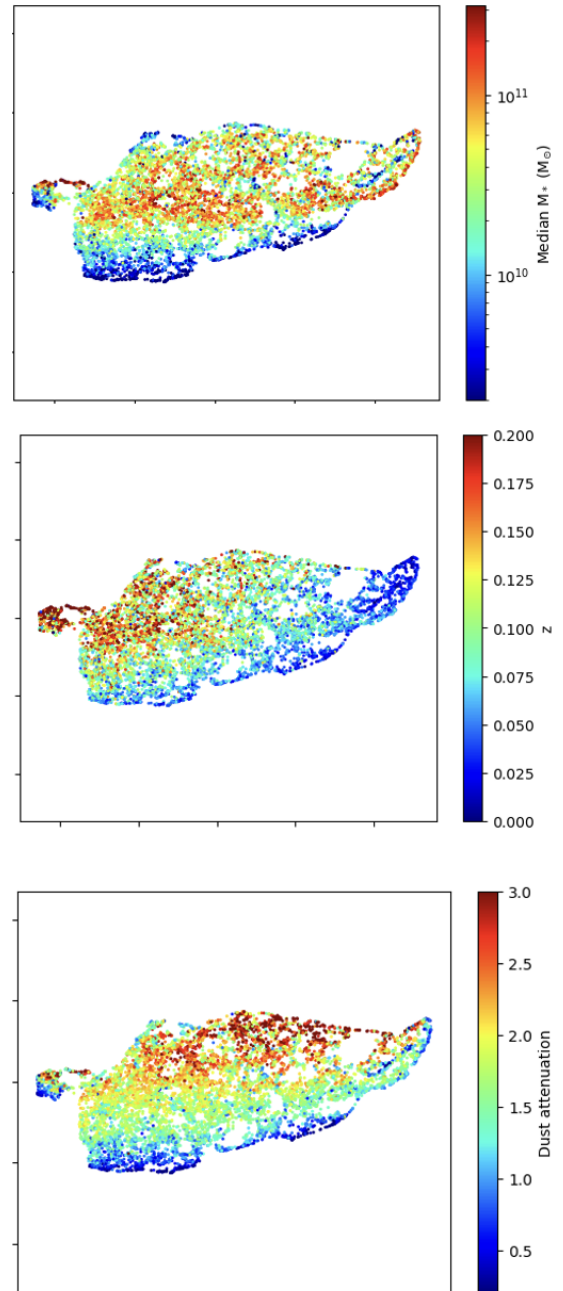

Figure 14. UMAP obtained for a sample of 10^4 galaxies when reducing the dimensionality of our data (9 dimensions for 9 photometric bands) to two dimensions. The two new dimensions are the x and y axis of each panel. From the top, the colormaps represent the stellar mass, redshift and dust attenuation of the different galaxies.

ables, we require the use of a regressor model. Also, our model can predict more than one parameter simultaneously, such that we can predict both redshift and stellar mass at the same time.

Figure 15. Clusters identified by HDBSCAN for our compressed data. The different panels show the effects of varying the minimum cluster size: Top left panel (100), Top right panel (200), Bottom left panel (300), Bottom right panel (500).
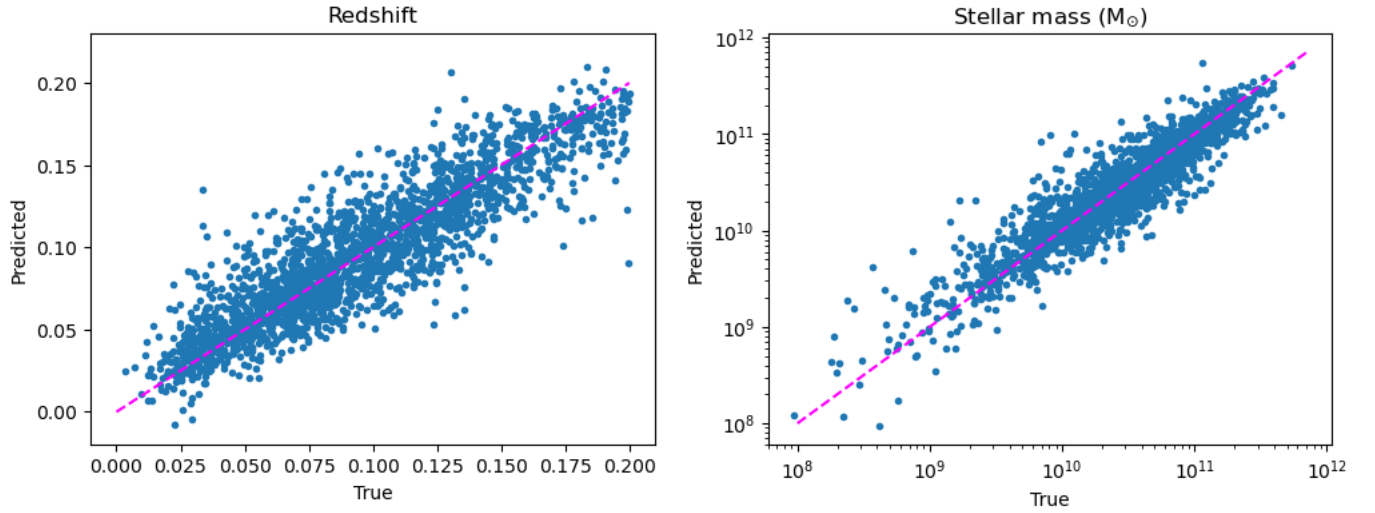
Figure 16. Comparison between predicted galaxy properties and "true" (i.e from Chang et al. (2015)) galaxy properties. The left panel shows the redshift while the right panel shows the stellar mass.

Fig. 16 compares the redshifts (left) and stellar masses (right) predicted by our algorithm against their "true" values (from Chang et al. (2015)). Our algorithm is generally able to predict galaxies' redshifts and stellar masses, with some scatter. It is worth noting that the training set plays a crucial role in the success of supervised ML algorithms. It is important that the properties of the galaxies which we are trying to predict be represented in the training data.

4. COMPARISON

We will compare the methods that had the most success in the machine learning category and the SED fitting category: MLP regressor neural network model vs Nautilus.

Accuracy: For A1835, our ML algorithm predicts a redshift of 0.12 and $Log_{10}\frac{M_*}{M_\odot}$ of 11.18. Nautilus predicts a redshift of $0.26^{+0.04}_{-0.04}$ and $Log_{10}\frac{M_*}{M_\odot}$ of $12.22^{+0.12}_{-0.12}$. Chang et al. (2015) lists values of z=0.252 and $Log_{10}\frac{M_*}{M_\odot}$ = 11.74. This also agrees with known literature values for A1835 of z=0.2514 and $Log_{10}\frac{M_*}{M_\odot}$ = 11.76 (Gingras et al. 2024). Fig. 17 compares the distributions of redshifts (left) and stellar masses (right) from Chang et al. (2015) (grey) with the "true" redshift and stellar mass of A1835 (dotted black line). The red line shows the bestfit result from ML and the green histograms show the marginalized distributions obtained from Nautilus. Clearly, the redshift and stellar mass predicted for A1835 using our ML algorithm are not accurate. However, as shown by the grey histograms and the black dotted lines in Fig. 17, A1835 has large redshift and stellar mass compared to the Chang et al. (2015) sample. This biases our estimates since the model is being trained with galaxies with smaller z and M_*. It would be preferable to train the model with galaxies which are more similar to the object studied. The predicted redshift value from Nautilus is much more accurate, agreeing with the "true" values within less than 1σ. However, both ML and SED fitting do not predict A1835's stellar mass accurately, with a discrepancy of almost 4σ for the prediction from the Nautilus SED fit.

Both methods have limitations and strengths. As previously discussed, for ML, the training set is very important. The model should be trained with objects with similar properties as the ones we are trying to estimate. Otherwise, this will bias the estimates and give bad re-sults. The choice of hyperparameters can also play an important role in ML algorithms. Although other ML algorithms can predict uncertainties, this is not the case for the one I chose. Also, we cannot use priors, or establish ranges of values allowed for the estimated quantities. However, ML does not require a model so it can estimate galaxy properties even if the model is completely unknown. Also, ML is great for predicting many properties for many galaxies simultaneously.

In comparison, SED fitting with Nautilus has to be run for every galaxy, which is much longer to run for many objects than ML. It also relies on the model being used. Results hinge on the choice of model. However, both with Dynesty and Nautilus, we can test different models using the Bayesian evidence to find which model better describes the data. It is also useful to know if our fit has converged. Our SED fitting method also gives us uncertainties on our parameters as well as posteriors and corner plots such that we can learn about covariance between the different parameters.

5. APPLICATIONS TO RESEARCH

The techniques which I have described here, both for the SED fitting and for machine learning can be used in interesting ways for my research. I work on cooling in galaxies and galaxy clusters. To understand what causes cooling in cluster cores, I study the similarities and differences in the properties of the intracluster medium (ICM) for galaxy cores with and without signs of cooling. I can use ML and clustering algorithms to compare the ICM properties of various systems and find similarities/differences in cool cores and non-cool cores. I may be able to find a clear criterion to distinguish between cool core and non-cool core clusters. I also have access to a database of cluster core properties, which I can use to try to predict the properties of similar objects. I can also apply the work which I have done with SED fitting to galaxy/cluster cores. Many cooling clusters have multiwavelength observations. With Nautilus, I can fit galaxy SED faster and infer some of the host galaxy properties such as stellar mass and star formation rate. I can also try to expand the range of the SED from Prospector such that I can fit galaxy SEDs across a larger range of wavelengths. Using Nautilus, Bayesian evidence and other methods to compare the probability of different models, I can compare various models describing ICM properties as a function galactocentric radius to learn about the evolution of the ICM inside galaxies.

REFERENCES

Adelman-McCarthy, J. K., Agüeros, M. A., Allam, S. S., et al. 2008, ApJS, 175, 297, doi: 10.1086/524984

Blanton, M. R., Schlegel, D. J., Strauss, M. A., et al. 2005, AJ, 129, 2562, doi: 10.1086/429803
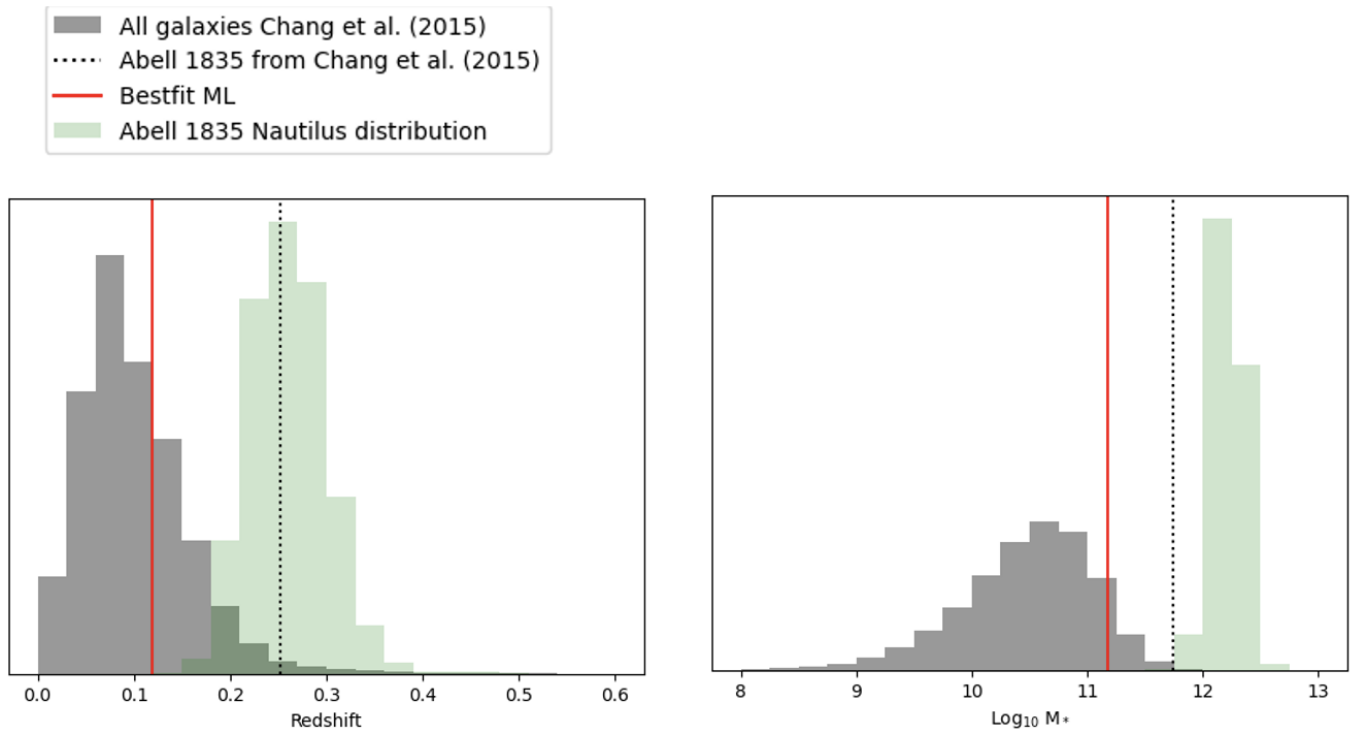
Figure 17. Comparison of the distributions of redshifts (left) and stellar masses (right) from Chang et al. (2015) (grey) with the redshift and stellar mass of A1835 (dotted black line), also from Chang et al. (2015). The red line shows the bestfit results from ML and the green histogram shows the marginalized distributions obtained from Nautilus.

Chang, Y.-Y., van der Wel, A., da Cunha, E., & Rix, H.-W. 2015, ApJS, 219, 8, doi: 10.1088/0067-0049/219/1/8

Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306, doi: 10.1086/670067

Gingras, M.-J., Coil, A. L., McNamara, B. R., et al. 2024, arXiv e-prints, arXiv:2404.02212, doi: 10.48550/arXiv.2404.02212

Johnson, B. D., Leja, J., Conroy, C., & Speagle, J. S. 2021, ApJS, 254, 22, doi: 10.3847/1538-4365/abef67

Koposov, S., Speagle, J., Barbary, K., et al. 2022, joshspeagle/dynesty: v2.0.3, v2.0.3, Zenodo, doi: 10.5281/zenodo.7388523

Lange, J. U. 2023, Monthly Notices of the Royal Astronomical Society, 525, 3181, doi: 10.1093/mnras/stad2441

Leja, J., Johnson, B. D., Conroy, C., van Dokkum, P. G., & Byler, N. 2017, ApJ, 837, 170, doi: 10.3847/1538-4357/aa5ffe

McInnes, L., Healy, J., & Melville, J. 2018, arXiv e-prints, arXiv:1802.03426, doi: 10.48550/arXiv.1802.03426

Padmanabhan, N., Schlegel, D. J., Finkbeiner, D. P., et al. 2008, ApJ, 674, 1217, doi: 10.1086/524677

Speagle, J. S. 2020, MNRAS, 493, 3132, doi: 10.1093/mnras/staa278

Wright, E. L., Eisenhardt, P. R. M., Mainzer, A. K., et al. 2010, AJ, 140, 1868, doi: 10.1088/0004-6256/140/6/1868