NLUAstro 1, 59-74 (2019) Notices of Lancaster University Astrophysics PHYS369: Astrophysics Group project

An analysis of the internal conditions of galaxies at z ≈ 0.1 and the identifying features of galaxy sub-types*

Charlotte Alexander[1], Pascale Desmet[1], Jonathan Dixon[1], Ciara Lithgow[1], Tom Measey[1], John Pollard[1], Phoebe Stainton[1] & David Sobral[1]†
[1] *Department of Physics, Lancaster University, Lancaster, LA1 4YB, UK*

Accepted 31 May 2019. Received 30 May 2019; in original form 22 March 2019

ABSTRACT
We have utilised the eighth data release of the Sloan Digital Sky Survey (SDSS), in combination with CLOUDY photo-ionisation models, to investigate the physical conditions of local galaxies at $z \approx 0.1$. This was done by separating our selected data into active and passive galaxies, and identifying star-forming, starburst and active galactic nucleus (AGN) sources. We found that stellar mass is a key predictor of galaxy type: most (80%) low mass galaxies are starburst galaxies, while the most massive galaxies are passive, with these being ≈ 100% of the population at $z \sim 0.1$ for stellar masses in excess of $10^{11.8}\,M_{\odot}$. We find that AGN are overall rare, but they become fractionally more important at higher stellar masses. CLOUDY photo-ionisation models reveal that star-forming galaxies have a strong linear relationship between temperature, metallicity and stellar mass, recovering the well studied mass-metallicity relation in the local Universe. AGN in SDSS are relatively metal rich at all stellar masses and we find that they are found at density peaks within the cosmic web. Overall, our results reveal a picture in which the most metal poor, lower mass galaxies are still actively assembling their stellar mass, while the most massive sources have assembled the bulk of their mass at higher redshift.

Key words: SDSS, Quenching, AGN, galaxy evolution, photo-ionisation.

1 INTRODUCTION

There are various types of galaxies and identifying which galaxy belongs to each type is a challenge which has been facing astronomers for decades. Galaxies in the Universe can be classified in three major categories: star-forming, passive and active galactic nucleus (AGN). AGN contain a supermassive black hole that is accreting matter, and the emitted light comes from particles losing angular momentum as they fall onto the black hole, which is then converted into friction, and then light (e.g. Cattaneo et al. 2009). The host galaxies can have quite low star formation rates, but the characteristics of individual AGNs depend on properties such as mass, the accretion rate, the orientation of the galactic disk, and the presence of jets (Ananna et al. 2019). As these galaxies are rarely observed in the local Universe, it is possible that every galaxy goes through this phase, but it is very short lived, or that fewer galaxies undergo this stage of evolution, but they remain in this phase for longer.

Passive galaxies are no longer forming stars and this is usually due to one of two effects. Either preventative feedback, where the gas and dust is heated to such an extent that it cannot collapse and form stars. Otherwise it can be due to ejective feedback, where the galaxy expels the dust in the form of things like jets (Lacey et al. 2006). Passive galaxies can still contain AGNs, as many used to be active, but have since evolved and lost the ability to form stars. Star forming galaxies are those that still have enough material to form stars. They have reasonably constant star formation rates until they run out of gas, at which point their star formation rates will begin to decrease (Guo et al. 2019). Starburst galaxies are characterised by a period with an incredibly high star formation rate given the stellar mass of the galaxy. During this time, large numbers of stars are produced, and the galaxy uses up most of its gas reservoir. After this, these galaxies will slowly evolve to become passive (see e.g. Lehnert & Heckman 1996).

Classification of galaxies by type is an efficient way to determine a galaxies properties. Baldwin et al. (1981) defined galaxies based on conditions including metallicity, which led to the creation of BPT diagrams (Baldwin, Phillips, Terlevic), as seen in Figure 6. Use of the BPT di-

* Based on observations obtained with the SDSS, DR8.
† PHYS369 supervisor.

agram allowed the quick and easy classification of galaxies by type during this investigation. The physical conditions of a galaxy such as mass, metallicity and star-formation rate play a vital role in its evolution.

Galaxies in the local Universe are the most often studied, and have brought about some of the most fundamental advances in the field of cosmology and in our understanding of the universe, one notable example being Edwin Hubble's revelation that the velocity at which a galaxy is receding is proportional to its distance from us, leading to what we now know as the Hubble-Lemaitre law (Hubble 1929). As such, we chose to investigate some of the closer galaxies, as opposed to the more distant ones.

Throughout this paper we have assumed Ω_Λ =0.7, Ω_m=0.3 and the Hubble parameter H=0.7, in line with the Λ-CDM model of cosmology (e.g. Sobral et al. 2012). The results of the investigation are presented in Section 4, and they are discussed and analysed in Section 5.

1.1 Motivation

The aim of this project was to use data collected by the SDSS to investigate the physical properties of active galaxies, in order to give some insight into galactic evolution. Galaxies would be categorised into active and passive. Active galaxies would then be classified further into AGN, star-forming, and Starburst. The categorisation process would use the SDSS data and techniques including BPT diagrams (mentioned in 1) and spectroscopic analysis. CLOUDY could then be used to model physical properties of the galaxies that have been classified as active.

Analysis of these calculated properties, spectral line ratios and photometric analysis would also be carried out. This allows conclusions to be drawn on any trends in galactic properties. Using multiple techniques means that the findings encompass a broad range of galactic types and properties. This will increase understanding of our local group and the Milky Way.

1.2 Structure Formation

on large scales the universe is largely isotropic and homogeneous, with small density fluctuations allowing the formation of celestial objects such as galaxies and clusters (Zeldovich 1972). These small fluctuations were present at the beginning of the Universe and were then amplified by cosmic inflation. The Universe was initially radiation dominated and incredibly hot, which ensured that hydrogen and helium were fully ionised. This meant that photons had an incredibly small mean free path before undergoing Thomson scattering. However, as the Universe expanded and cooled, protons and electrons began to form neutral hydrogen in a process called recombination (Grin & Hirata 2010). This made the Universe transparent and allowed photons to propagate through the Universe and be detected as CMB, providing evidence for some initial density fluctuations.

The creation of hydrogen and helium paired with areas of higher densities allowed structures to begin forming. The atoms in these denser regions would be gravitationally attracted to each other, and they would start to clump together, which would then increase the gravitational force they had on other atoms nearby. The force would eventually become large enough that more and more atoms would become bound together, initially forming stars, before forming galaxies, clusters, and super-clusters.

If the distribution of galaxies in our universe is plotted on a sphere representing the sky, it becomes clear that matter has clumped more in certain areas in comparison to others (Springel et al. 2005). The areas lacking any structure are known as voids, and those with a high density of galaxies are known as filaments. combination of these reveals the cosmic web, which is shown for our data in Figure 1. As a result of the universes isotropy, there should be complete uniformity in galaxy types in every direction (Saadeh et al. 2016). However, this assumption does not hold true on smaller scales. In areas of the Universe with large amounts of dust, it is expected that star forming galaxies would be the most common. These galaxies would use the dust to form stars before becoming quenched. In areas with less dust, there should be fewer galaxies overall as there was less material to initially form galaxies, and those that are present are likely to be quenched. In terms of galaxy size, smaller galaxies are statistically more likely to form as a smaller gas cloud is required to form them, so many galaxies in the nearby Universe are classified as small elliptical dwarf galaxies.

1.3 Galaxy types and Classifications

Different galaxy types can be visually shown on galaxy colour-magnitude diagram, which shows the relationship between mass and absolute luminosity. Most galaxies fall into either the blue cloud or the red sequence. Those in the blue cloud are active star forming galaxies and are usually spirals (late type morphology), whereas those in the red sequence are usually either quenched or have low star forming rates Feltzing, S. & Johnson, R. A. (2002). They tend to be elliptical galaxies (early type morphology). The transition phase on the diagram is known as the green valley, and it is very sparsely populated. However, those galaxies that are present are lacking in the necessary components to form stars. This could be due to galaxies having lacked gas for billions of years, or it could be due to gas reservoirs being rapidly destroyed due to galactic mergers or a net outflow of gas due to an AGN. Galaxies like the Milky Way are in the green valley due to their decreasing star formation rate.

BPT diagrams, such as those found in Baldwin et al. (1981), provide a visual method for categorizing different galaxy types based on their ratios on ionised metals to hydrogen (Kewley et al. 2013). They are nebular emission lines that can enable the ionisation method for a given region of nebular gas to be distinguished. The graphs produced, such as those in Baldwin et al. (1981), and Kewley et al. (2006), show clear lines separating different galaxy types. This work is of great importance as it shows a clear link between the levels of metals within a galaxy and the classification of the galaxy, which could lead to a better understanding of how galaxies can evolve, and what factors may influence this.

The most common version consists of [NII]6584/Hα versus [OIII]5007/Hβ (Baldwin et al. 1981). The next two more commonly-used BPT diagnostics are [SII]6717,6731/Hα versus [OIII]5007/Hβ (BPT-SII) and [OI]6300/Hα versus [OIII]5007/Hβ (BPT-OI). These diagrams were studied in numerous work and the dividing lines have been developed

and adapted as a function of the ionisation models and/or observations available (e.g. Kewley et al. 2001, 2013)

1.4 Past SDSS Investigations

Weinmann et al. (2006) study of SDSS data set out to look at the dependence of galaxy type on its properties. This included star formation rate and mass of the galaxy, two properties this report will be considering. However Weinmann et al. (2006) defined their galaxies differently. The use of early and late type galaxies was used. These classifications were less definite than those used in this report, this is due to the classification of each galaxy type used. Early type were typically red, elliptical galaxies with a low SFR (Weinmann et al. 2006). Late type galaxies were blue, spiral galaxies with a high SFR (Weinmann et al. 2006). Galaxies within this sample could have properties of both galaxy types, this led to a subjective classification which means that galaxies can be classified differently depending on those classifying the galaxies. This motivated the use of a unique classification system which could be easily reproduced. The methods used in Weinmann et al. (2006) allowed comparison for methods used in the data classification. However this report used different techniques to define subclasses and therefore the conclusions drawn by Weinmann et al. (2006) will not be directly compared.

Brinchmann et al. (2004) investigated the physical properties of local galaxies, focusing on star formation. The positive correlation between SFR and stellar mass within the galaxies was found, with strong agreement across most of their sample (Brinchmann et al. 2004). It was concluded that most stars are produced in galaxies with log(O/H) ratios (thus metallicity) similar to solar values (Brinchmann et al. 2004). The link between star formation and mass, is further tied due to the dust attenuation which affects the calculation of both of these galaxy properties (Brinchmann et al. 2004). Therefore an introduction of a correction factor for dust is important within this report. Brinchmann et al. (2004) state that the Kennicutt (1998) conversion factor is an effective correction to be used in the SFR calculations for the local universe. This correction was utilised in this report.

There are numerous previous studies on the SDSS data. Many of which have a similar objective to this report, however one of the largest differences between them is the classification of galaxies within the reports. The use of spectra is considered within (Folkes et al. 1999), quantifying each galaxy by the spectral emissions it produced. The spectral lines were grouped for classification as; molecular features, Balmer lines and nebula features. This led to unique classifications which did not correspond directly to real galaxy classifications, but the classifications were reproducible. Balmer lines were also used in this report to classify galaxy types. The use of photometry in Bell et al. (2003) to classify galaxies was done to separate early and late type galaxies. This was done by creating and reproducing simple cutoffs in order to separate the data and avoid the misclassification of galaxies (Bell et al. 2003). This report followed a similar simple classification system in order to gain clean, reproducible galaxy classifications. In Tempel, E. et al. (2011) the colour of the galaxies was used to classify them. The use of photometric methods was used in this report for analysis but not to classify the galaxies. It is stated in Folkes et al. (1999) that visual inspection is the most accurate way of classifying galaxies. This technique is used in Nakamura et al. (2004), for morphological classification. In this report this technique was not used due to the type of galaxy classification which was undertaken, and the time consuming nature of this method. However visual inspection was used to check the classification of sources as galaxies by the SDSS in cases with unusual signatures, due to it being the most accurate way to determine some galaxy properties.

2 CATALOGUE

2.1 The Sloan Digital Sky Survey: Data Release 8

The catalogue that was used for the investigation was taken from the 8^{th} Data Release from the Sloan Digital Sky Survey, or SDSS (SDSS a). Located at Apache Point Observatory in New Mexico, USA, the telescope used is a 2.5 m optical telescope which observes the sky with a wide angle in the visible, part of the electromagnetic spectrum (Brinchmann et al. 2004). Over the years, SDSS has made a large variety of data releases each expanding on the previous one to present more precise measurements with more sources. SDSS-III, the third phase, began in August 2008, with surveys focused on galactic structure and chemical evolution. Released in 2011, SDSS Data Release 8 was the first to include extensive photometric information along with spectroscopic data (Aihara et al. 2011). This made DR8 the most ideal data release for the analysis presented in this paper. An advantage to using SDSS data for the investigation is that it covers a large area, over one third of the celestial sphere (Aihara et al. 2011). This large sample area reduces limitations and bias in the data. In addition to this, the survey is also equipped with a deep field of view which allows us to carefully select redshift range for local galaxies. SDSS is equipped with the means to measure and calculate a large range of quantities including mass, star formation rate, and galaxy type. (Brinchmann et al. 2004)

2.2 The CLOUDY Simulation

One of the fundamental elements of this investigation involved the comparison of the data from SDSS to a simulation in order to obtain more information regarding the sources. The simulated data was obtained using a variety of CLOUDY models. CLOUDY is a photo ionisation code and models the ionisation state, of a material that is exposed to an external radiation field (or other heat source) and predicts the values of observables such as the flux of emission and absorption lines in the galactic spectra when given a set of input conditions (Ferland et al. 2013). In our case, this will simulate photons from an undefined source entering a gas cloud with properties that are specified in §2.4. CLOUDY subsequently displays the result of the photons interacting with the cloud in the form of spectroscopic data. The achieved simulated spectra are what was then used as the primary object of comparison to match the recorded data with the simulation. Three different models were created with CLOUDY using different conditions for the code to work from. The first of the three to be used was the BB

(black body) model. This models a single black body and is therefore ideal for simulating galaxies that are star forming as the measured radiation will be produced by stars which we can approximate to have black body behaviour. The second model that was used was the PLAW (power law) model. This model is ideal for simulating AGN since their spectra are the result of a superposition of many black bodies. The combination of black body radiation at a variety of different temperatures approximates to a power law. Finally, age related quantities like mass and temperature are simulated by the BPASS (binary population and spectral synthesis) model (Eldridge et al. 2017). The nature of this model is such that it is only a reliable source of comparison for star forming galaxies since BPASS code doesn't simulate AGN at all.

It is important to note that CLOUDY does not predict physical conditions, only show what the spectrum of a source with the given set of physical conditions might look like. (Sobral et al. 2018b). Our CLOUDY data is the same data used in Sobral et al. (2018b) and Sobral et al. (2018a).

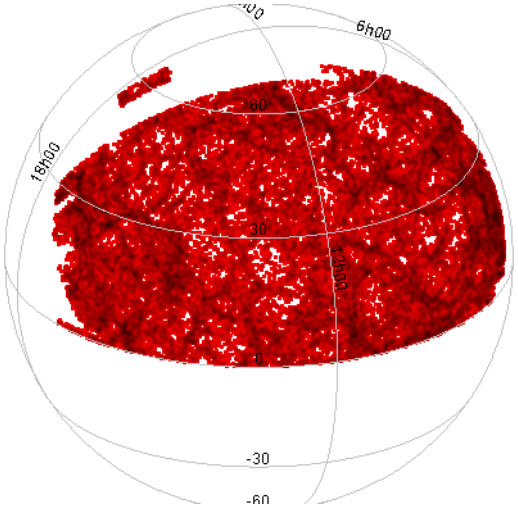

Figure 1. The remaining data plotted according to their R.A. and Dec positions which are in the previously mentioned red-shift range and which will be used for analysis.

2.3 Imposed Conditions on our SDSS Data

The catalogue of data from SDSS DR8 initially contained approximately 1.8 million spectra. This was far more than what was required or manageable as trying to process that much data repeatedly would have taken too long to be practical. We started with two data tables for DR8, one with the spectral information, and the other with some more information about the ID number of the spectra directly from the image capturing process. The table containing the spectral information contained only 850,000 items, so by combining these two tables (by matching items by their RA and DEC positions using the match tool in the astronomical programming software TOPCAT Taylor (2005)) a new smaller table was created. This table contained only 850,000 items, but each data point had all of the information we would be needing.

After this was done there were a few conditions that could be imposed immediately to reduce the number of objects further; first anything that was not a galaxy was removed. The catalogue contained the spectra of stars, quasars and galaxies which had been pre-identified upon the release of the catalogue. By creating a subset and setting the spectral type to be "Galaxy", all other celestial objects were removed as shown in Table 1. However this only decreased the number of objects by about 100. It was then noticed that the RA positions for some of the galaxies was 9999, and their DEC positions were 0. These were obviously false data-points, given that an RA position of 9999 is impossible, and were promptly removed. This brought the total down to 760,000. Now as only the usable data remained, it could then be decided which data was going to be used in the analysis.

The first property considered was the red-shift of the galaxies. A small range was required to ensure that any trends observed further through the analysis was purely due to the conditions within the galaxies. If a large range was used, then the results would be skewed by the fact that only the brightest galaxies will be observable from high red-shifts, which may tend to have certain internal conditions to create this brightness. The galaxies also needed to be comparatively close, as the aim of the project was to study local galaxies. Bearing these considerations in mind, a red-shift range of $0.09 < z < 0.11$ was decided upon. Removing galaxies that were outside of this region produced a sample of 84,000 galaxies, an ideal amount for this investigation.

The remaining data was then plotted on a graph of the sky in TOPCAT and it was noticed that while the large majority of the data was found in one large region, there were a couple of strips of data which came from the other side of the sky.

These strips had been targeted specifically in special data releases, and so could have contained their own trends or conditions, hence the region for them being specifically targeted. For simplicity and to further ensure that our results were not skewed, these strips of data were removed using a condition on the galaxies RA position. After this there were 79,000 galaxies contained within one region of the sky, chosen only for the usability of the data and their red-shift. Since these were the only conditions on the data, they are a representative set from which valid conclusions could be drawn. The remaining data is shown in Figure 1.

In order to carry out analysis of the active galaxies with the CLOUDY data, a way of distinguishing the different types of active galaxies (AGN, star-forming and Starburst) from each other in a way that is easy to see and manipulate was required. A graphical method was the obvious choice and for this a 'BPT diagram' was decided upon. This type of diagram was first used by Baldwin, Phillips and Terlevich, and its purpose was to split astronomical objects in to several distinct groups; being HII regions, planetary nebulae, and objects photoionised by a power law continuum and Shock-wave heating, which are difficult to distinguish (Baldwin et al. 1981). The method we have chosen to use as our primary method of investigation is the one which separates HII regions from regions of AGN (Kewley et al. 2013). HII Regions will correspond to regions of star formation; as the OII regions are produced most commonly by O class stars, which due to their short lifespan will only exist in regions of star formation. After these stars are gone, there would be no source of ionisation left, and we move towards the AGN re-

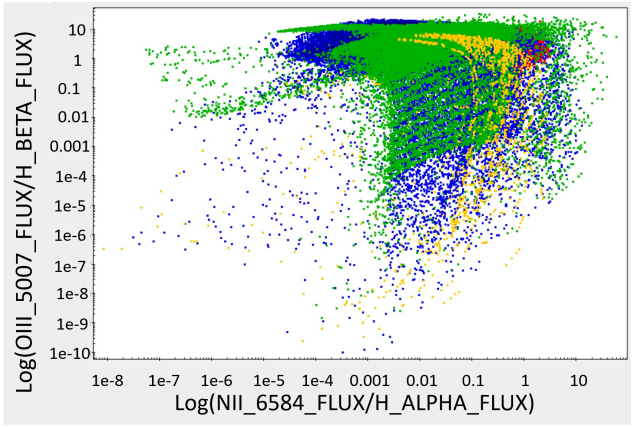

Figure 2. The Red region shows the reduced SDSS data remaining after those with high noise were removed. The other colours show different CLOUDY models (BB - Blue, PLAW - green, BPASS - yellow).

gion. One such graph is shown in Figure 6, where the SDSS was also used.

In order to remove data points which have a high level of noise, and are therefore useless for our spectroscopic analysis, a restriction was placed on the ratio of the emission line flux to the error in the flux. Any spectra that had a ratio of flux divided by the error in flux of greater than three was removed. This removed 51% of the remaining spectra leaving 38,700 galaxies. In doing so, the remaining data is now completely covered by the CLOUDY models

This remaining data was then plotted on a BPT diagram against the data from the CLOUDY simulations (the conditions required to recreate these models are given in §2.4). Figure 2 shows this diagram.

It must now be mentioned that this final cut in our data is also a condition that is key in identifying passive galaxies; the signal to noise ratio of the Hα line. By using this condition to remove data, most of the passive galaxies are also removed therefore a subset with these spectra removed was only used when investigating the active galaxies.

A summary of these steps is detailed in table 1.

2.4 Cloudy Initial Conditions

The CLOUDY models use the parameters described in Table 2, and varies each of these parameters over a large range of data, details of which are seen in Table 2. Most of the parameters involved have clearly defined physical meanings common outside of astrophysics, such as temperature. The other important parameters are which are more unique to this field of study are: logU (in and out) is a parameter detailing the ionisation strength of the gas in each simulation (Sobral et al. 2018b). The other log parameters; C/O and O/H are both measures of the relative abundances of these elements. This is converted into Solar Metallicity; a comparison to the Sun's metallicity where a Solar Metallicity of 1 indicates a metal content identical to that of the Sun's. In terms of a galaxy, this value would indicate that the majority of stars have the given metal content.

A significant amount of data analysis will be done via plotting our data in the form of a BPT diagram with

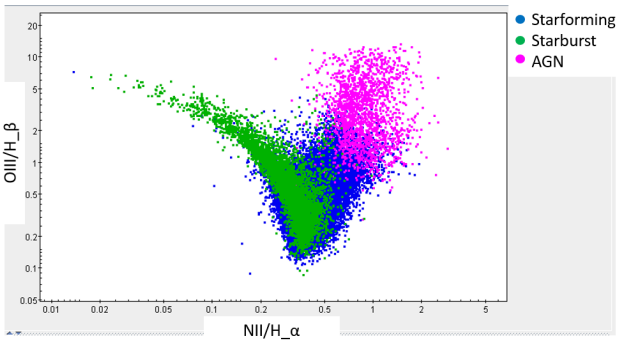

Figure 3. Ratio of NII(6584)/Hα Fluxes against OIII(5007)/Hβ Fluxes, displaying the separation of each galaxy type; starforming(Blue), Starburst(Green) and AGN (Magenta)

x-axis: NII(6584)/Hα Flux ratio
y-axis: OIII(5007)/Hβ Flux Ratio

Figure 3 displays our data, and not only are we able to clearly see the distinction between each type of galaxy, but the shape of the data exactly matches that which is expected to be seen as in Figure 6.

3 METHODOLOGY AND ANALYSIS

The steps and conditions imposed on the data set in order to classify the galaxy types are given in this section. This includes the split between active and passive galaxies. As well further conditions which were imposed onto active galaxies to produce subsets of galaxy types.

The calculations of selected galaxy properties are also outlined.

3.1 Defining Galaxy Types

3.1.1 Passive Galaxies

Passive galaxies will be classified by imposing conditions on the data set. Passive galaxies are not producing any stars, therefore the Hydrogen Hα Balmer line can be used to identify them, as this should have a negligible value. The Hydrogen Hα Balmer line is only produced by the largest but shortest lived O and B stars, which ionise the hydrogen gas surrounding them. These stars have lifetimes on the order of tens of millions of years; once these stars have died, there is no longer a source of production of the Hydrogen Hα Balmer line. Galaxies with a negligible Hydrogen Hα Balmer line emission have not undergone star formation in the last several million years, as no O or B stars are present. Therefore they are passive. It is also possible that galaxies with high amount of error for this line are also passive galaxies, and any recorded flux is simply an error.

The initial condition used to identify passive galaxies is if $3\times$ Hα_{ERR} >Hα, the galaxy is defined as passive.

Initially, conditions were imposed to remove all data points with error on the Hydrogen Hα Balmer line three times greater than the flux itself. This condition was removed so the data could be included once again, allowing this classification to be done. This can be seen in Figure 4.

The classification in Figure 4 left a large amount of the

Table 1. Table showing the process taken to remove unwanted data from the catalogues. At each step in the process, the exact statement used in addition to the one from the previous step to create a a new subset of the original data is quoted (using the subset tool in TOPCAT), as well as the approximate number of spectral sources remaining after each cut. *this condition was only used for the analysis using CLOUDY and all data was included during the identification of passive galaxies.

Condition	Statement used	Sources remaining
Initial Calalogues galSpecInfo-dr8 SDSS (a) & galSpecLine-dr8 SDSS (a)		1,843,000 & 856,400
Combine the two catalogues, matching items by their R.A. & Dec positions		856,400 (100%)
Remove any objects which are not galaxies	equals(SPECTROTYPE_1, "GALAXY")	856,300 (100%)
Remove false data points with RA = 9999 and Dec = 0	$RA\_1 < 9999$ && $DEC\_1 > 0$	760,800 (89%)
Select only galaxies which are within the redshift range desired: $0.09 < z < 0.11$	$0.09 < Z\_1$ && $Z\_1 < 0.11$	84,400 (10%)
Remove Special Release data which appeared in strips on the opposite side of the northern hemisphere to the main bulk of the data	$RA\_1 < 270$ && $RA\_1 > 90$	79,400 (9%)
Remove data with high noise in the emission lines initially used to plot the data ([OIII],[NII], $H\alpha$ & $H\beta$)*	$(H\_BETA\_FLUX/H\_BETA\_FLUX\_ERR) > 3$ && $(NII\_6584\_FLUX/NII\_6584\_FLUX\_ERR) > 3$ && $(OIII\_5007\_FLUX/OIII\_5007\_FLUX\_ERR) > 3$ && $(H\_ALPHA\_FLUX/H\_ALPHA\_FLUX\_ERR) > 3$	39,100 (5%)

Table 2. Starforming region subgroups and number of data points within each subgroup. Temperature is in units of K, Density in units of (n_H,cm^{-3}), Age in units of (log Age, yr). (Sobral et al. 2018b) (Sobral et al. 2018a). Those whose steps have been labelled as '—' had defined steps which were unclear between each subsequent value.

Simulation	Parameter	Minimum value	Maximum value	Step
BB	Temperature (K)	2000	159,000	1000
	logU(in)	-5	1.95	0.05
	logU(out)	-5.38	1.95	0.02
	Density (n_H, cm^{-3})	3	1000	0.5 (log space)
	Solar Metallicity	0.01	3.0	0.05 (log space)
	log(CO)	-0.85	0.5	—
	12 + log(O/H)/	6.7	9.2	0.02
PLAW	PLAW Slope (Slope)	-2	-1	0.05
	logU(in)	-5	2	0.05
	logU(out)	-5	2	0.05 (with slight perturbations)
	Density (n_H, cm^{-3})	3	1000	0.5 (log space)
	Solar Metallicity	0.01	3.0	0.05 (log space)
	log(CO)	-0.85	0.5	—
	12 + log(O/H)/	6.7	9.2	0.01
BPASS	Age (log Age, yr)	6	8.9	0.1
	Metal Stars	0.001	0.04	—
	qH(in)	46	52	—
	logU(out)	-7.99	0.752	—
	Density (n_H, cm^{-3})	3	3000	0.1 (log space)
	Solar Metallicity	0.01	3.1	0.5 (log space)
	log(CO)	-0.35	0.5	—
	12 + log(O/H)/	7.54	9.1	—

data unclassified. In an attempt to increase the number of galaxies classified, the Hα equivalent width line (EQW) was used. A positive EQW value meant that the galaxy was absorbing $H\alpha$ therefore it must be passive. A value for a cut off with low Hα_{EQW} emission values needed to be selected, and this was done through Figure 5. The contribution of different spectral types can be seen clearly and then a cutoff value which does not include large contributions from already classified galaxies (e.g star-forming) was selected for the passive galaxy classification. The cut off value of H$\alpha_{EQW} > -3$, produced a subset of 46% of the data. This overlapped with already classified galaxies making up 1% of the sample, approximately 1000 data points. Due to the large sample size this was deemed acceptable. This led to the classification of passive galaxies with the conditions; H$\alpha_{EQW} > -3$ and $3\times$H$\alpha_{ERR} > H\alpha_{flux}$, which encompassed 46% of the sam-

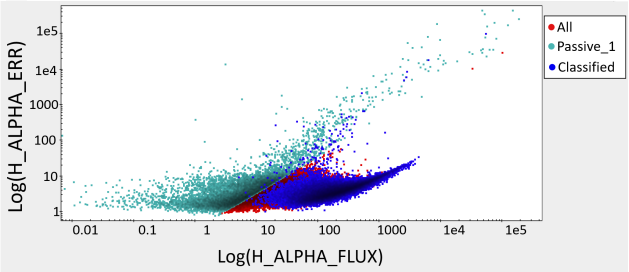

Figure 4. Intial classification of passive galaxies, defined as those with error on Hydrogen Hα Balmer line greater than Hydrogen Hα Balmer line itself. This results in all galaxies marked red as possible passive galaxies under this condition, which constitute approximately 17% of the whole sample.

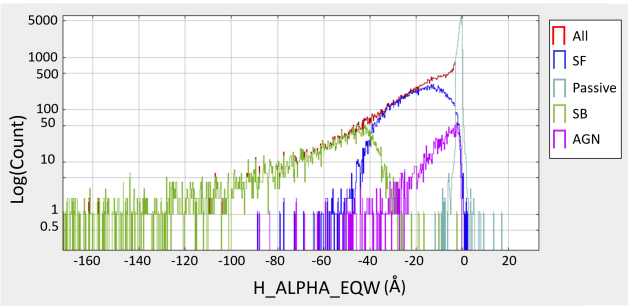

Figure 5. Contribution of each galaxy type to the number of galaxies producing a certain Hα EQW. This allowed a selection of the passive galaxies to be made by selecting the sample at a value which does not include large numbers of other classified galaxies. The passive_3 line is the new sample, with a cut off EQW value of -3.

ple, with only 1% crossover with the other classified galaxy types as in Table 3.

3.1.2 Subgroups of Active Galaxies

The SDSS catalogue included classification of galaxies into sub-types: Starburst, star-forming and AGN. These were set according to the conditions in Table 3, alongside the passive galaxy classifications.

These parameters for the distinction between AGN and star-forming galaxies originate from Kewley et al. (2001). This combination of emission lines is restricted by the range of metallicities and ionisation values which are realistic for Starburst galaxies. This allows a separation of these two galaxy types to be produced, using these conditions, as seen in Figure 6. Starburst galaxies were used in Kewley et al. (2001) due to their strong emission lines. However this separation technique can be used for all star-forming galaxies, as in Sobral et al. (2016), which can be combined with a further classification to define Starburst galaxies.

Starburst galaxies are star-forming galaxies which are undergoing the highest rates of star formation. Therefore their emission of Hα emission line will be larger than the other star-forming galaxies. This produces a large equivalent width of Hα for Starburst galaxies which can be used to separate them out from other star-forming galaxies (SDSS b).

3.2 Calculated Properties

Star Formation Rate (SFR) was calculated using the Hα Flux line for the star-forming and Starburst galaxies in our data set. However a value was calculated for all data points (thus galaxy types), but it had no real meaning due to the non stellar sources which it was produced by in some galaxy types.

The Hydrogen-α spectral line was used due to its production by large, short lived O and B stars ionising the hydrogen gas which surround them. Their presence and production of Hydrogen Hα Balmer lines in a galaxy implies that the galaxy in undergoing or have recently undergone star formation (Maungkorn & Kriwattanawong 2017). This is due to the short lifetime which the massive O and B stars which produce this spectral line have. Therefore if the flux is present the galaxies have recently undergone star formation (Sobral et al. 2011).

The Hydrogen Hα Balmer line is also used due to its advantageous qualities that it holds over other spectral lines, such as limited extinction, allowing a more accurate estimate of SFR (Sobral et al. 2013).

3.2.1 Luminosity Distance

First the luminosity distance needed to be calculated using the redshift of the objects. Due to the small range of redshift in the sample, the calculated luminosity distances have a small range, 400-494 Mpc. The distance was calculated using a TOPCAT function.

This gave the distance of the objects, which allows the luminosity to be extrapolated by using the flux of the Hydrogen Hα Balmer line from the survey, and the fact that flux decreases with a $\frac{1}{r^2}$ relationship.

$$L_{H\alpha} = 4\pi f_{H\alpha} d_L^2 \tag{1}$$

Where flux_{H_α} has the units with a correction factor $10^{-17}\,\text{erg}\,\text{s}^{-1}\,\text{cm}^{-2}$ d_L is the luminosity distance in [cm].

3.2.2 Star Formation Rate

The $H\alpha$ flux line is affected by dust extinction. This leads to lower values of the Hydrogen Hα Balmer line than what should be measured from these galaxies due to the dust absorbing a proportion of the radiation. This needs to be accounted for in order to make the star formation rate as accurate as possible. As in (Best et al. 2012) The ratio of Hα Balmer line to Hβ Balmer line is used as an estimate of the size of the dust extinction in the Hα Balmer line. The correction factor used is given by (Best et al. 2012) as

$$A_{H\alpha} = 6.531 \log 10\left(\frac{H\alpha}{H\beta}\right) - 2.981 \tag{2}$$

The star formation rate is finally produced by converting the $H\alpha$ flux and the dust extinction correction term into a SFR. We used (Kennicutt 1998) to give the correction factor, to be able to use $H\alpha$ flux. This was calculated using the Salpeter IMF, and is assumed that star formation is constant and continuous across all the galaxies.

$$SFR_{H\alpha} = 7.9 \times 10^{-42} \times 10(\log_{10}(L_{H\alpha}) + A_{H\alpha}) \tag{3}$$

Table 3. All of the conditions used to classify the galaxy types. The steps are taken from all data points to the smallest subset (2%) These conditions are imposed after the steps in Table1 were undertaken to produce the reduced data set (SDSS b).

Galaxy Property Used	Conditions	Galaxy Type	% Data Set
H_α Balmer Line Emission	$H_{\alpha\_EQW} > -3$	Passive	17%
Noise of H_α Line	$3\times H_{\alpha\_ERR} > H_{alpha\_flux}$	Passive	46% (with above)
Metallicities and Ionisation	$\log_{10}(O_{III}/H_\beta) > 0.7$ - $1.2(\log_{10}(N_{II}/H_\alpha) + 0.4)$	AGN	2%
Metallicities and Ionisation	$\log_{10}(O_{III}/H_\beta) < 0.7$ - $1.2(\log_{10}(N_{II}/H_\alpha) + 0.4)$	star-forming	32%
High H_α Balmer Emission	$H_\alpha > 50$Å	Starburst	6%
Galaxies classified as passive, as well as either AGN, SB or SF		Crossover	1%
-	-	Unclassified	15%

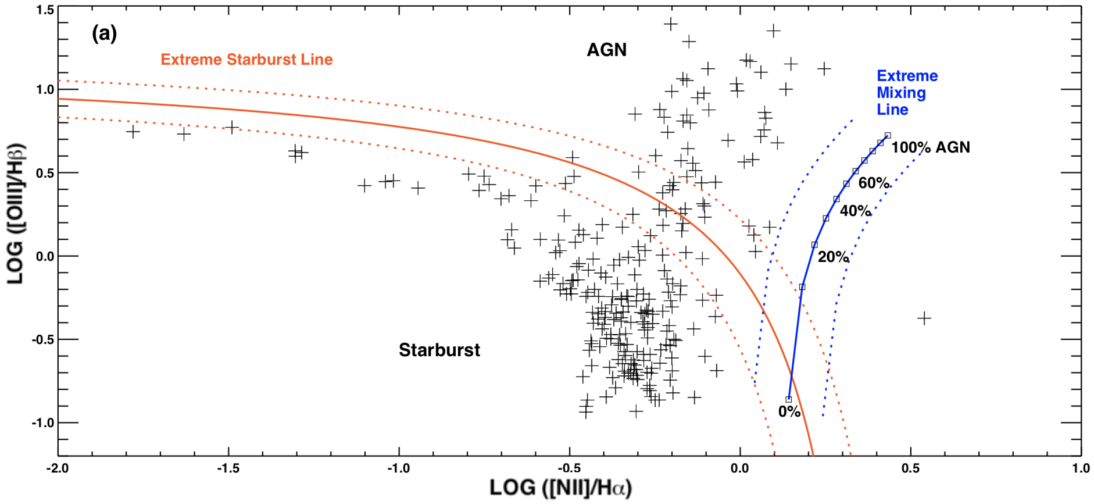

Figure 6. Separation of AGN and Starburst galaxies using limit of ionisation and metallicity parameters defined through the spectral line ratios used. (Kewley et al. 2001)

This produces a star formation rate in $M_\odot$ yr^{-1}. This rate was calculated for all of the galaxies with a recorded $H\alpha$ flux. This included AGN galaxies, as they had a $H\alpha$ emission line. However the flux of this line was not produced by star formation in these galaxies, but from the black hole in the centre of the galaxy instead. This is due to the high temperatures which these black holes can reach. This leads to the emission of $H\alpha$ radiation. This has led to AGN galaxies now having a sizable calculated star formation rate despite not being star-forming galaxies. This meant that the use of SFR had to be considered carefully when producing plots, to ensure that the information being plotted was physically possible and not just a fabricated value of SFR.

4 RESULTS

4.1 CLOUDY Trends

Using the simulated values of flux from CLOUDY, and plotting this over our SDSS BPT diagrams produced the following graphs in the following subsection. The data has been split into our active galaxy group types; star-forming, Starburst and AGN. Each analysis will be done independently, and will be seeking to characterize the overall trends of our parameters through the system.

4.1.1 Starforming Galaxies

For our star formation regions, only the PLAW CLOUDY simulation is not appropriate for consideration. The BB and BPASS simulations were used alongside the SDSS data in order to produce graphs which show the simulated variance of CLOUDY's initial parameters across the region. The CLOUDY data here has been selected by eye around the outline of the star-forming data. this was done in order to focus the data onto the more densely populated areas of the star-forming region. There are 631 data points in the star-forming region for the BB simulation, and 4658 data points from the BPASS simulation in the star-forming region.

Firstly let us consider the BB model and the variance of temperature across the star-forming region. Temperature tends to increase with an increasing ratio of NII/Hα flux as we move towards the AGN region and away from the Starburst region. While the highest temperature values (as

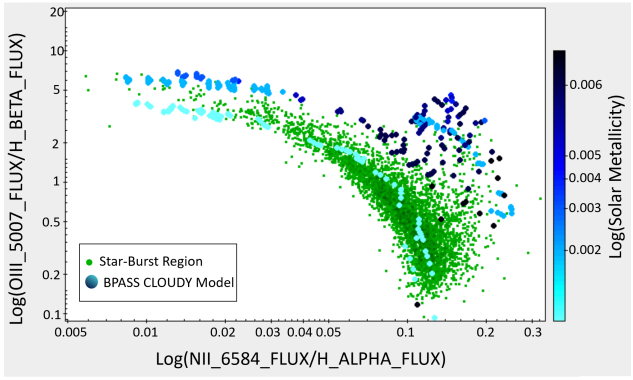

Figure 7. Scale of simulated metallicity in the BPASS simulation superimposed onto the Starburst galaxy region.

seen in black) are dotted around the region, it must be noted that these are most likely an artifact from the simulation; the galaxies that lie beneath are unlikely to have the stated temperature exactly. Rather, we should look for the overall trend of data within the region; and we are able to see a trend of increasing temperature towards the AGN region from the Starburst region. Consider §4.2.1 for a more detailed analysis of temperature variation.

The BB simulation was also used to consider the variance of Solar Metallicity across the star-forming region. Most of the galaxies have a metallicity at or below 1 which is to be expected; in star-forming regions we expect to see sun-like stars as a majority, so this result is expected. As before, these "odd" values are likely artefacts of the simulation. The BPASS simulation is also valid within the star-forming regime, and thus we are able to construct similar BPT diagrams for the BPASS simulation.

4.1.2 Starburst Galaxies

Starburst galaxies are not a standalone type of galaxy, rather they are a phase which star-forming galaxies undergo at some stage of their lifetime. A Starburst galaxy is simply defined as a galaxy undergoing a period of extremely high star formation. This phase is most often triggered by a close encounter, or collision, with another galaxy. The exceptionally fast rate of star formation means that the available fuel in the galaxy is quickly used up, so this phase is relatively short when compared to the galactic lifetime (Lehnert & Heckman 1996). Initial graphs showing OIII/Hβ and NII/Hα absorption lines produced a curve of approximate equal metallicity. To investigate this further, the simulated data from the CLOUDY BPASS data set was superimposed on the SDSS data. Figure 7 shows this data with the CLOUDY data colour coded to show metallicity. Sure enough, the data shows that the bulk of the Starburst curve has a low overall metallicity. This implies that galaxies undergoing their Starburst phase are younger on average then their star-forming counterparts who have produced more metals through supernovae in their longer lifetimes.

4.1.3 AGN Galaxies

For the investigation regarding AGN the only appropriate simulation code for CLOUDY is the PLAW code. This model is the envelope of several different black body radiation curves and approximated to a power law. The upper highlighted region of the graph in Figure 3 shows which of the data points in the survey have been identified as AGN. To make the comparison with CLOUDY, a region was manually selected within the simulation to match the area from the data once the PLAW model had been superimposed. The selection was made by eye using the manual subset function on the software. One of the more important quantities to consider throughout this analysis in the PLAW model is density.

4.2 CLOUDY - Parameter Comparison

Defining subgroups in each of the galaxy types enables a more qualitative analysis of the data, allowing for direct comparison of parameters which were not practical to do in §4.1.

Distinct regions were defined in each data set in order to calculate average values for a number of parameters in each region, both from CLOUDY (temperature, density, solar metallicity) and using the SDSS data (mass to the 50^{th} percentile), calculated SFR). This enabled us to track the evolution of these CLOUDY parameters alongside the physical parameters; the main comparative being mass.

In this section, the error bars for each quantity are given by

$$\mathrm{Err} = \frac{\sigma}{\sqrt{n_i}} \tag{4}$$

where σ is the standard deviation of a given parameter in the subgroup, and n_i is the number of data points of either the SDSS (for Mass, SFR) or the relevant CLOUDY model (for the other parameters in said subgroup).

4.2.1 Starforming Galaxies

The subgroups shown in Figure 8 were chosen to group the data into approximately equal regions based on area, as opposed to the number of stars within the region, which varies greatly.

It is acknowledged that by utilising a higher number of subgroups would increase accuracy in the data obtained: the most accurate means of comparison would be to have each subgroup contain roughly the same number of data points from each model. This was judged to be unnecessary, as the focus here is not the values contained in the quantitative data, but to use this to identify any trends that appear. For the star-forming region, seen in Figure 8, the subgroups were defined aesthetically in order to determine if there was a particular trend when moving from the Starburst side towards AGN, defined as in Figure 3.

Figure 9 shows the evolution of BB simulated temperature with mass: the more massive galaxies have a higher predicted temperature, with quite a strong positive correlation. The line of best fit has a gradient of $\approx$ 50500 K per $\log(M/M_\odot)$, which is a very large temperature gradient. Unsurprisingly, the hottest (and hence most massive) of these regions are those towards the AGN region, which is in agreement to that discussed in Section 4.1.1. We expect

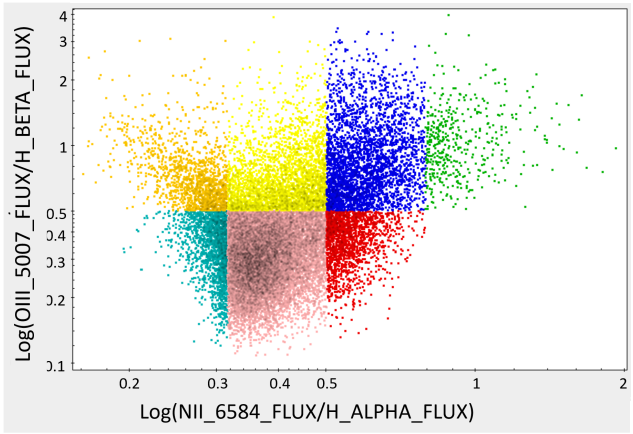

Figure 8. Illustration of the star-forming regions, Low1(L1 - Cyan), Low2(L2 - Pink), Low3(L3 - Red), High1(H1 - Orange), High2(H2 - Yellow), High3(H3 - Blue) and High4(H4 - Green)

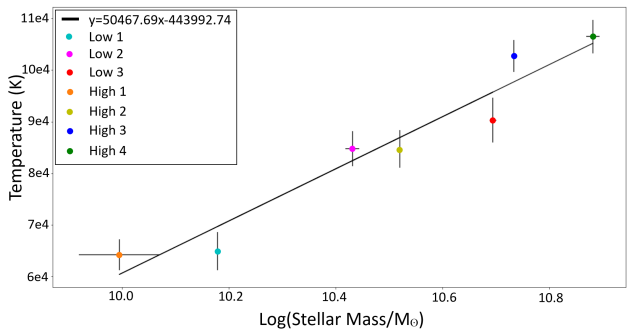

Figure 9. Variance of Mass (x, $\log(M/M_{\odot})$) with Temperature (y, K) across the star-forming region, as predicted by the BB simulation. The line of best fit has the equation $y = 50500x + 444000$. Mass is seen to increase with temperature.

this physically; one such suggestion for the cause of quenching in AGN is virial shock heating Gabor et al. (2010), which would increase the temperature of AGN-leaning galaxies.

A more useful parameter to consider is solar metallicity across the star-forming region. BB predicts a gradient of 0.5 $Z/Z_{\odot}$ per $\log(M/M_{\odot})$, with an intercept of -4.67 $Z/Z_{\odot}$, whereas BPASS predicts a gradient of 0.35 $Z/Z_{\odot}$ per $\log(M/M_{\odot})$, with an intercept of -3.12 $Z/Z_{\odot}$. The similarity of the two models should be noted here as in Section 4.1.1, the BB and BPASS metallicity variation through the regions did not exhibit any clear structure.

Kewley & Ellison (2008) also finds a positive correlation between mass and metallicity over a range of calibration values using 12+log(O/H), a parameter that was also utilized by CLOUDY models used here. It is suggested in Tremonti et al. (2004) that dwarf (low mass) galaxies are more likely to be metal depleted, due to both the ubiquity of galactic winds (galactic winds occur throughout the galaxy), and its effectiveness in removing metals from galaxy potential wells (winds are strong enough to eject metals from the galaxy at speeds quicker than the escape velocity).

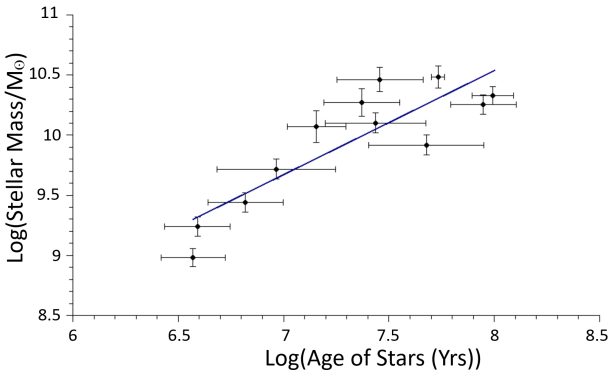

Figure 10. Variance of Mass (y, $\log(M/M_{\odot})$) with Age (x, log(age, yrs)) across the star-forming region, as predicted by the BPASS simulation. The line of best fit has the equation $y = 0.93x + 8.7$. Age is seen to increase with Mass.

4.2.2 Starburst Galaxies

The Starburst region is a subgroup of the star-forming region, and thus many of the same conclusions are expected to also hold true here. In order to produce a quantitative analysis of the data, the Starburst region was split into twelve groups, primarily divided evenly based on a the OIII/Hβ axis, with several more groups encompassing the remaining data. Again, the average and standard deviation of variables were calculated within each group using the method detailed at the start of Section 4.2. By comparing average mass to average age of a given group with the BPASS simulation, Figure 10 was created. This shows a clear linear relationship between the log-age of a Starburst galaxy and its log-mass.

As a comparison to the star-forming region, the SFR and mass averages in each subgroup have been plotted for the star-forming and Starburst regions as displayed in Figure 11.

By considering Figure 11 we can see that star-forming galaxies and Starburst galaxies are separated by mass; with the Starburst galaxies occupying the lower stellar mass region. By considering analysis of Figure 10, it was determined that log-age is proportional to log-mass. It can be hypothesized that the star-forming galaxies are older than Starburst galaxies on average - having lower masses, and some may have already had a Starburst phase when they were younger.

4.2.3 AGN Galaxies

To observe AGN on a quantitative scale, the region from the data set representing AGN was split into ten different regions. As before, each region was analysed by taking mean values and standard deviations, and calculating errors as detailed at the start of Section 4.2.

Due to the sparse nature of the PLAW model throughout the AGN region of the BPT diagram, the data from the groups would have been very biased towards values of low NII/Hβ flux. As such, the primary focus of investigation will be to compare the average SFR to the average mass within the AGN region. This also serves to follow Section 4.2.2.

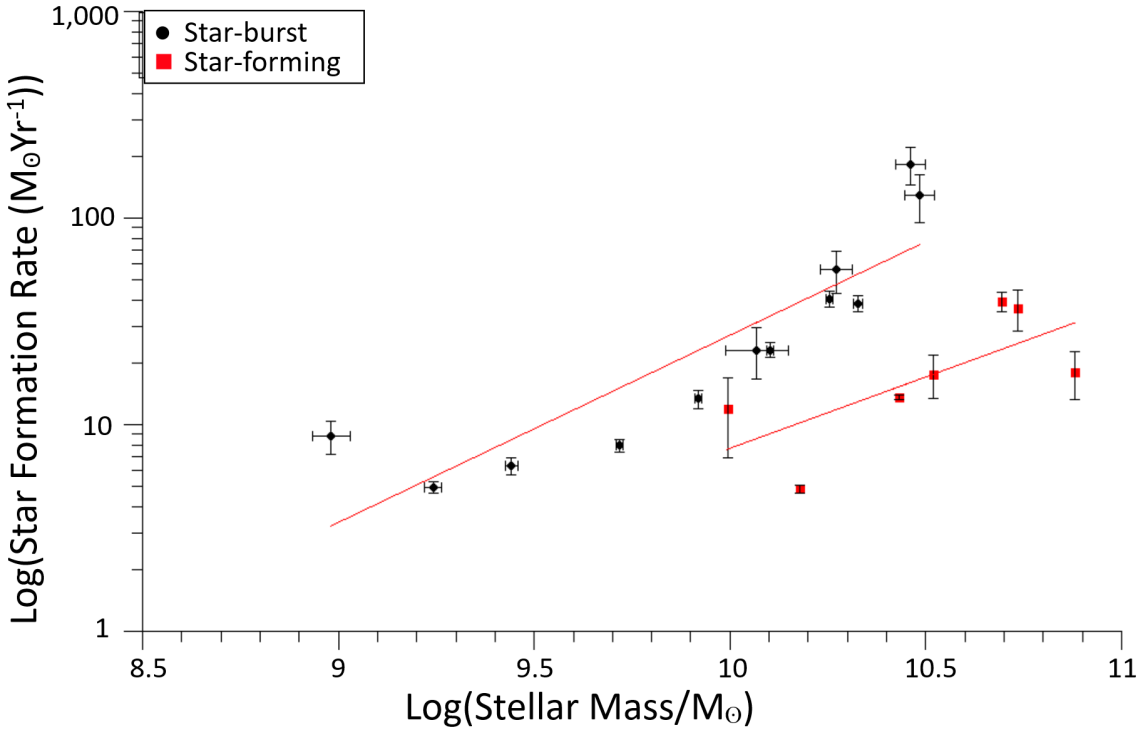

Figure 11. Comparison between star-forming and Starburst regions from the consideration of Mass (y, log($M/M_\odot$)) and SFR (x, log(Stellar Mass/$M_\odot$). Both have a positive correlation, though occupy different regions of the diagram.

4.3 Fractional Distribution Graphs

4.3.1 Star Formation Rate

Using the value of SFR calculated in §3.2, the difference between galaxy types can be investigated. The star formation rate for each data entry was calculated, including galaxy types for which this is an unrealistic calculation, such as AGN. For AGN, Hα is likely coming from the accretion disk. It is instead indicative of another very ionising source within these galaxies; their black hole. The star formation rate calculated for these galaxies is not a real parameter and should be ignored.

Star formation rate is appropriate for use with galaxies which have been classified by their Hα emissions in §3. Therefore only star-forming and Starburst classifications remain to be compared by star formation rate.

There is a larger contribution from star-forming galaxies across all star formation rates. This is likely due to there being more star-forming galaxies in the sample than Starburst. It was expected that Starburst galaxies would have a greater contribution at high star formation rates, due to their properties (and subsequent classification) §3, which required a 50Å Hα EQW. The absence of high SFR Starburst galaxies could be due to a small number of Starburst data points, which causes any errors/anomalous values to be propagated from star-forming galaxy data points, onto any possible trends.

In order for star formation rate to be a more useful parameter in this investigation into galactic properties, the data set needs to be examined more closely. The extreme values of Hα flux which led to large SFR need to examined to remove any corrupted values which may reside. The classifications of galaxies in §3 could also be reviewed further to ensure galaxy types were accurately assigned.

4.3.2 Mass

The mass of each galaxy can be used as an indicator of trends in each galaxy type. As in Brinchmann et al. (2004), the expectation was held of star-forming galaxies to dominate at small masses, due to the large proportion of star formation which occurs in smaller mass galaxies. The opposite is true of AGN galaxies, which are predicted to dominate at larger masses. This is likely due to greater likelihood of smaller galaxies to still be star-forming.This is in contrast with the lower probability of lower mass galaxies having an AGN. This is a consequence of the formation of a super-massive black hole being linked to galaxy mergers, and therefore it is unlikely that a low mass galaxy would have undergone such an event, reducing the likelihood of a low mass AGN galaxy Wang & Kauffmann (2008).

Including passive galaxies in the data set drastically alters the fraction of galaxies represented at each mass. The dominance of passive galaxies at high mass is due to their mass being mostly incorporated into stars, which are detected by the SDSS. There is a peak from AGN, representing the trend seen in our results. However, the number of AGN galaxies is far less than the number of passive galaxies in the data set, leading to the smaller contribution by AGN at high mass in Figure 12

The dominance of star-forming galaxies at low masses can still be seen. This is due to the evolution that the other

galaxy types have undergone (e.g. star formation or galaxy merger), which reduces the likelihood of them being recorded as low mass by the SDSS survey.

Figure 12 shows the fractional contribution to mass across all classified data. Starburst galaxies dominate at low mass due to the reduced likelihood of Starburst events occurring in massive galaxies (Bergvall, Nils et al. 2016). This may also be due to the measurement of mass in the SDSS survey which accounts for the stellar mass, excluding contributions from dust and non luminous material. Starburst (and star-forming) galaxies which have a lower proportion of their material as stars are recorded as low mass by the SDSS survey. The star-forming galaxies peak in their contribution in the mid mass range, but still account for large proportions of the population at all masses. This is due to the large range of galaxies which are star forming, and the period which star formation can occur over. The proportion of their mass which is stellar material widely varies, producing a wide range of mass values for star-forming galaxies.

Through looking at mass using Figure 12, key trends of the galaxy types can be inferred. The use of the mass calculated by the SDSS survey, which infers the galaxy mass through stellar source emission, creates a strong mass dependence for all of the galaxy classifications. The lower mass values are dominated by galaxies still producing stars (Star-forming and Starburst). This is due to these galaxies still having large amounts of mass in non stellar material (in order to produce stars) but this means that there is a smaller amount of stellar mass in the galaxy. This gives the low mass domination by these galaxy types. As galaxy mass increases the dominance is held by passive galaxies. These galaxies have finished producing stars, typically due to a lack of star-forming material, therefore their mass is accounted for in stellar mass, and is therefore recorded as high mass by the SDSS survey. AGN galaxies have a small peak at higher masses due to the mass contribution from their black holes. AGN galaxies do have a low contribution to the high mass values, but this is due to their low proportion of the data set (1%) which is translated into Figure 12. The trends produced were as expected due to the properties of each galaxy type.

4.3.3 Metallicity/Line Ratios

When comparing the ratio of NeIII flux and Hγ flux, it is noticeably higher for AGN galaxies at almost all ratios. The only exception to this occurs when the ratio is 3, where there seem to be more star forming galaxies with a higher ratio. As AGN galaxies are typically more metal rich, the sudden change in the dominant galaxy type was unexpected, and it was decided that this point could be anomalous, or that the ratio of NeIII to Hγ was a poor metric to use when attempting to classify galaxies. Due to the large error bars and the lack of a general trend to the graph plotting this data, it was concluded that this point was likely to be anomalous and that NeIII was an inappropriate metal to use when classifying galaxies.

AGN galaxies also dominate when the ratio of OIII 4959 to Hβ is compared. This galaxy type contributes at least 80% of the measured galaxies at every ratio, suggesting that AGN galaxies have higher metallicities than their star forming counterparts. This is to be expected as AGN galaxies contain a dense nucleus that has the energy required to ionise atoms several times over whilst many star forming galaxies will not have the required energy. This means that AGN galaxies have larger quantities of ionised metals, which is reflected in the ratios they produce.

When looking at the ratio between NII flux and Hα flux, AGN galaxies only dominate above ratios of 0.8. Before this point, star forming galaxies accounted for between 95% and 100% of all galaxies at ratios between 0 and 0.4, and approximately 80% of the galaxies at a ratio of 0.63. At a ratio of 0.8, AGN galaxies then begin to dominate, contributing roughly 67% of galaxies at this point. The domination of AGN galaxies at higher NII flux/Hα flux ratios still suggests a higher metallicity than star forming galaxies, but it seems that star forming galaxies are more capable of producing singly ionised nitrogen in comparison to doubly ionised oxygen, which is likely due to the smaller amount of energy required for this process. Again, this suggests that AGN galaxies have higher quantities of NII compared to star forming galaxies, which gives them higher ratios. This in turn suggests that AGN galaxies are more metal rich and have the capacity to ionise more atoms, which is likely due to higher temperatures that can be found near the active galactic nucleus or in denser dust clouds. In comparison, the thermal energy required for ionisation in star forming galaxies will be contained within the stars themselves, meaning very little is available to ionise the atoms outside of the stars. Furthermore, metals have higher ionisation energies than elements such as hydrogen, so this further reduces the chances of metal atoms becoming ionised, which makes it highly unlikely that these galaxies could produced doubly ionised metals. To compound this further, stars that contain the energy required to doubly ionise oxygen are likely to be O and B type stars, which are produced in significantly smaller numbers and have drastically shorter lifespans than cooler stars. This means that double ionised metals are even less likely to be produced, and that the rate at which they are produced will fall off significantly as the star forming galaxies begin to run out of the materials required to form O and B type stars.

4.4 Photometry

As seen in Figure 13, passive galaxies have higher $\log_{10}$(U-G) values, ranging from 0.2-0.4, which indicates that they are redder than star forming galaxies which have $\log_{10}$(U-G) values ranging from approximately 0.06-0.31. As passive galaxies are no longer forming new stars, any stars that are still left would be older redder stars so this graph is supported by the theory. There is a small amount of overlap between the two, but that is possibly due to star forming galaxies that are starting to become passive or star forming galaxies with smaller star formation rates. The range of $\log_{10}$(R-Z) values is much larger for star forming galaxies which suggests that they emit light from a wider range of the electromagnetic spectrum compared to passive galaxies. This could be due to the types of ionised atoms present in the star forming galaxies, each of which can produce light of different colours.

For star forming regions, most of the data points fall between -0.54 to -0.27. The $\log_{10}$(U-R) values for passive galaxies are between 0.4 and 0.525, and the values for star forming

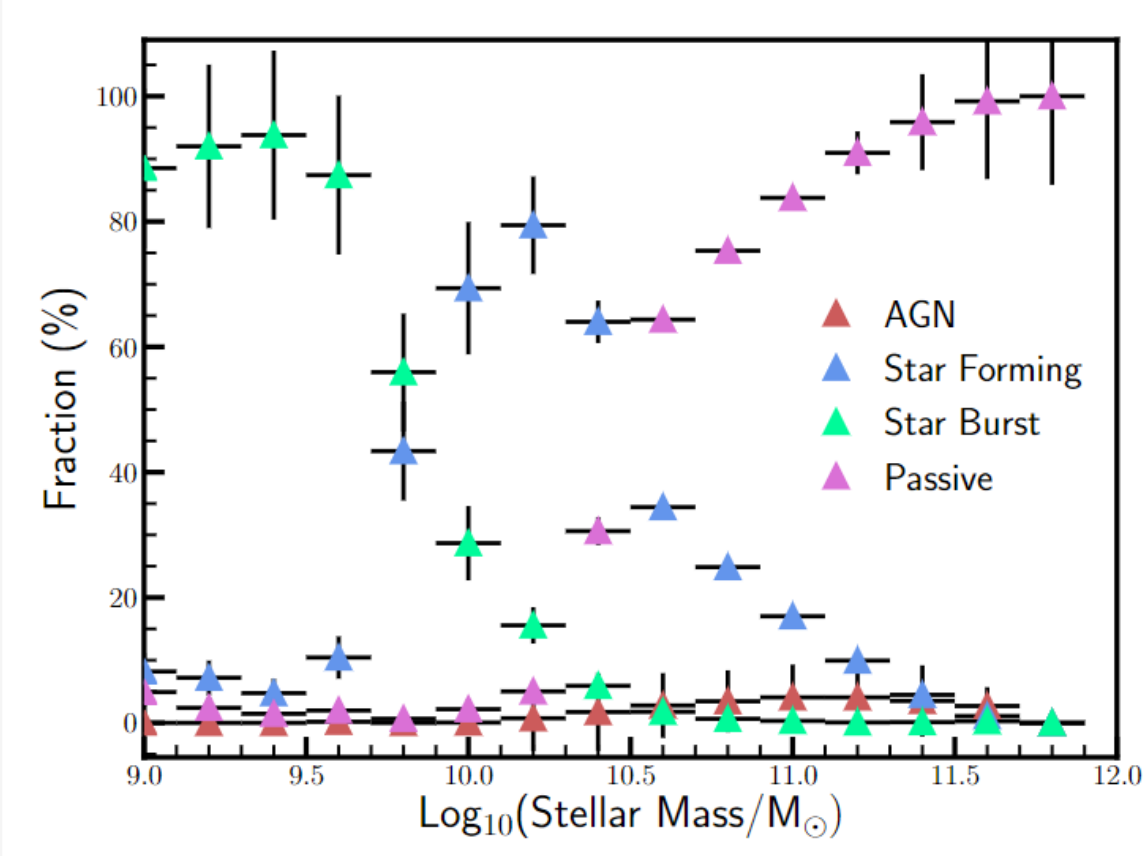

Figure 12. The fractional contributions of each galaxy type to the overall measured mass, using the final conditions for passive galaxies found in section 3.1.

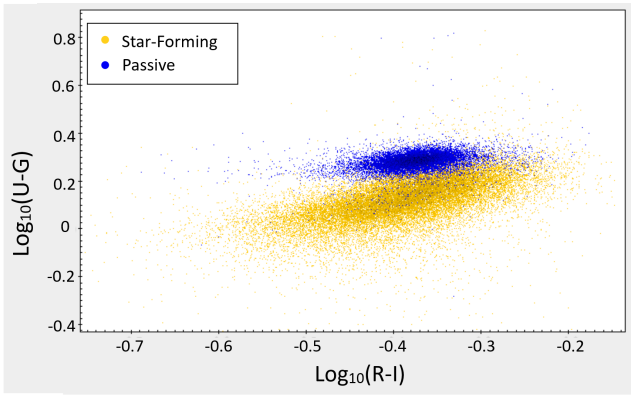

Figure 13. A photometry graph showing the effective reddening of passive and star-forming galaxies when the filters used are $\log_{10}$(R-I) vs $\log_{10}$0(U-G)

values are between 0.175 and 0.425. The small ranges in both values for the passive galaxies suggest that many of them share similar traits, and therefore the variation in reddening will be smaller. The star forming galaxies have larger ranges in their values for both $\log_{10}$(U-R) and $\log_{10}$(R-I), however the range in $\log_{10}$(U-R) is smaller than in the previous two graphs. As with the passive galaxies, this reduced range suggests less variation in the reddening of the star forming galaxies, and gives the graph a smaller gradient.

All the photometry graphs resemble colour magnitude graphs, in that there is a clear "red sequence", "blue cloud", and "green valley". The red sequence is shown by the passive galaxies, the blue cloud is shown by the star-forming galaxies, and the green valley is the overlap between these two groups. The main difference between the photometry graphs produced by the research data and the standard graph is the shape of the groups. The red sequence is usually longer and more oval shaped, whereas the blue cloud is usually closer to being circular. Our photometry graphs show both groups as resembling ovals with the blue cloud being longer than the red sequence which is somewhat unexpected. However, the blue cloud does have a wider range of y-axis values, which is what is shown on more typical colour-magnitude diagrams. This resemblance suggests that the data we have found agrees well with data that has already been collected whilst the disparity in the sizes and shapes of the star-forming group suggests that a study with more data may provide a better fit to the typical graphs.

It should also be noted that the shape of these photometry graphs will change over time with blue cloud galaxies migrating to the green valley and then to the red sequence. One possible path for this progression was suggested by Faber et al. (2007), in which it was suggested that galaxies in the blue cloud undergo gas-rich mergers and use up their remaining gas by undergoing a Starburst phase. These galaxies can then undergo gas-free mergers and move up fully into the red sequence.

5 DISCUSSION AND ANALYSIS

5.1 Fraction Graphs

The results from the photometry analysis concluded that it is obvious that passive galaxies are redder, which is as expected due to them no longer forming stars. The star-forming galaxies are seen as less red as they are still forming O and B type stars, which appear blue and will easily outshine multiple redder stars within the same galaxy.

Overall, we find AGN are relatively metal rich. This is demonstrated by every fraction plots having AGN galaxies dominating at higher ratios. This matches the assumption that the main process for production of these lines is photoionisation by an active galactic nuclei (Storchi-Bergmann et al. 1998).

AGN are found to be more prominent at higher masses, which is likely due to their high metallicities and galactic nuclei, both of which will add a significant amount of mass. Conversely, star-forming galaxies are likely to be found at lower masses due to the lack of mass contribution in these galaxies from stellar sources, which the SDSS survey contributes as galaxy mass. This is due to star-forming galaxies having a large mass contribution from gas and non luminous sources. Passive galaxies, as seen in Figure 12, dominate at higher masses. This is due to the majority of the material within these galaxies being accounted for as stars.

The star formation rates calculated by both the research team and the SDSS are limited due to the use of Hα emission line. Whilst Hα emission line is an incredibly sensitive source, such that it is reasonably unaffected by dust and small changes are discernible , it is usually not possible to pinpoint the exact source of the emission. This means that although a SFR can be calculated for all data points in the set, it does not ensure it is a meaningful value. This can be due to the nature of the source of Hα emission line. For example AGN galaxies have a Hα emission from their supermassive black hole, and were ignored when drawing conclusions from SFR values.

There were also a large number of unclassified galaxies that had large star formation rates and large error bars. Upon further investigation, the majority of these galaxies could be classified as passive, which we would expect to have low star formation rates. This supports the conclusion that the star formation rates calculated by the SDSS and the research group were unreliable.

5.2 Mass vs SFR

In §4.2.3, it was noticed that each galaxy type was separated into distinct regions based on the average masses of galaxies within each type's subgroups. On average, the Starburst region had the largest spread of masses; including a large proportion of low mass galaxies.

5.3 Summarising the CLOUDY Data

The CLOUDY simulation enabled us to make predictions on how several quantities were likely to vary throughout star-forming, Starburst and AGN regions. The important physical parameters to consider from cloudy are Density, Solar Metallicity, Temperature and Age, and how they vary across each region, as well has how they vary in terms of our measured parameters such as Mass and SFR.

The star-forming regions and Starburst regions have a lot of overlap. They both represent galaxies that produce a great number of stars, with their main difference being the luminosity and rate of star production (Calzetti et al. 2000). In terms of what CLOUDY predicts, the main focus has been on the star-forming regions. Let's first consider the variance of BB Temperature. Figure 9 details the variance throughout the star-forming region. We see Temperature increase as we move towards the AGN region, as well as Temperature increase with the mass of a galaxy. In combination, this leads to the implication that AGN galaxies are more massive than star-forming galaxies. This is quite surprising; given that AGN galaxies eject a large amount of material (Crenshaw et al. 2003) we would expect to see the opposite. This could however explained if AGN galaxies are more frequent after mergers; which naturally have larger masses in general. More research would be required to investigate this.

We can also again consider the mass to metallicity graphs considered in §4.2.1. The two graphs have been combined into Figure 14, and we can see with greater ease the difference between the two. At lower masses, the BPASS model predicts a larger prediction whereas at higher masses the BB model predicts a larger solar metallicity for each subgroup. There is a cross-over point at which the 2 models are in agreement at a Solar Metallicity value of 0.49 and a $\log(M/M_{\odot})$ value of 10.26.

With regards to the AGN region, particular focus needs to be brought to BPT diagrams. From plotting the data following the ratios it was immediately possible to identify the AGN region as being separate from the others with higher values for OIII/Hβ and NII/Hα. Further confirmation of this is visible in the log(C/O) plots. Specifically, it is possible to see how this quantity on CLOUDY defines a very neat curve on the BPT to separate star forming galaxies from AGN. When comparing the CLOUDY model to the data it is clear that whilst star forming galaxies stay on the lower C/O ratio portion of the plot, AGN are definitely concentrated in the section of the plot with much higher C/O ratio.

5.4 Further Work

Due to the nature of this project the amount of further work which could be undertaken with the data is large. There are endless different ways which the data could be analysed or defined. However, to be realistic, certain tasks should be focused upon. The task of defining galaxies using different parameters is an area which could be given particular importance by using different methods/parameters, as discussed in Section 1.4, or by defining different galaxy subgroups. The different classifications could then be compared for overlap of sources or for different trends produced by each classification. This could be done with an aim to produce the best definitions of galaxy types possible. Another further research direction is to attempt to calculate SFR in multiple ways. Both by using different emission lines as well as using different corrections and assumptions to try and find the most effective and accurate method of calculation. A missing spectral line which could provide further research is the 4000 Å break line. The inclusion of this line in the data set could allow improved passive galaxy definition, as well

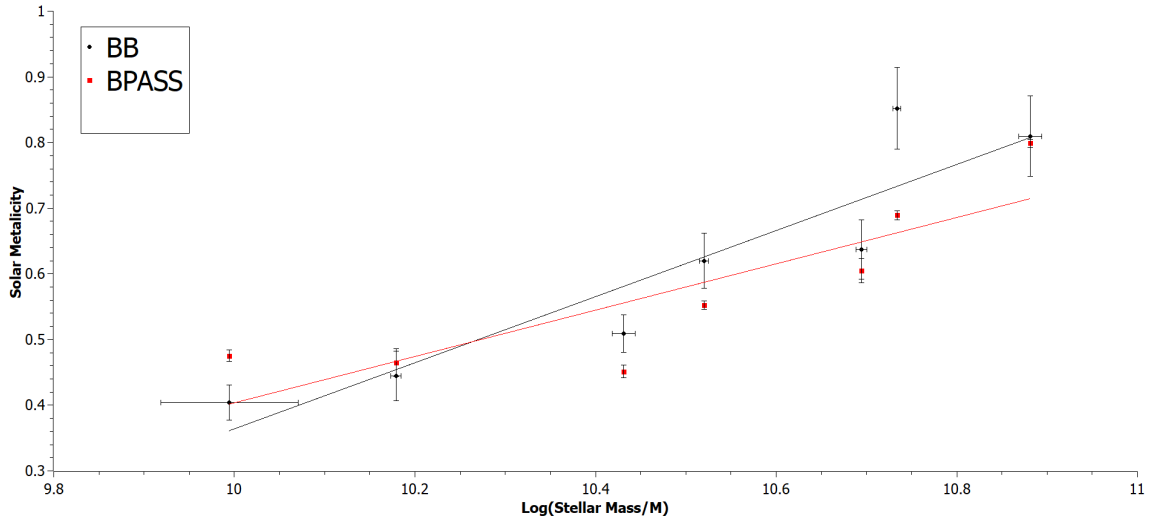

Figure 14. Combination of the BB and BPASS simulations of the Mass (x, $\log(M/M_{\odot})$) to Solar Metallicity (y, $Z/Z_{\odot}$) ratio in the star-forming region. The lines of best fit are $y = 0.35x - 3.12$ for BPASS and $y = 0.5x - 4.67$ for BB.

as another method to calculate SFR. In order to expand the results collected we could expand the section of the SDSS survey selected to include a greater range of redshift values. This would then allow comparison of trends across different redshift ranges. Consequently a comparison of galaxy trends beyond the local universe could be developed which would allow us to see how galaxies differ as they move further away.

6 CONCLUSIONS

Throughout our investigation we have used SDSS data in order to investigate the physical properties of galaxies by categorising them into active and passive, with further sub-classification of active galaxies into star-forming, Starburst and AGN galaxies. Spectroscopic and photometric techniques were utilised in order to produce fractional distribution plots and flux line ratio graphs. This allowed properties including mass, SFR, and metallicity, to be identified for different galaxy types. The comparison between different classifications also made it possible to analyse how properties such as reddening differ between galaxy types. We also made use of three CLOUDY models (BB, BPASS and PLAW) to model other physical conditions such as temperature and solar metallicity throughout the active galaxy regions. Our results show that:

- Of our selected data sources, 46% have been classified as passive. 40% are defined as a type of active galaxy (Star-forming, Starburst or AGN), and 15% of data points were unclassified. §3.1.1
- The Mass-SFR relation reveals that star-forming, Starburst and AGN galaxies are split according to their masses, with Starburst being the least massive on average, moving to AGN galaxies being the most massive.
- The Mass-SFR relation also revealed that Starburst galaxies for a given mass had a larger SFR than star-forming galaxies, which in turn have a larger SFR than AGN galaxies. For star-forming and Starburst galaxies, SFR was seen to increase with mass. The SFR of AGN galaxies cannot be calculated as the H_{α} emission lines seen are from the matter accretion by the AGN itself.
- There is a strong positive correlation between mass and metallicity within the star-forming region, reflected by both the BB and BPASS model. The two lines of best fit are (x - Solar Metallicity, y - $\log(M/M_{\odot})$) y=0.5x-4.67 for BB and y=0.35x-3.12 for BPASS. There is an intercept at the Solar Metallicity value of 0.49 and a $\log(M/M_{\odot})$ value of 10.26. For low masses, galaxies BPASS predicts a higher Solar Metallicity; for higher masses, BB predicts a higher Solar Metallicity.
- There appears to be a link between metallicity and density in AGN galaxies. This is alluded to by the fact that denser AGN in the BPT diagrams coincide with the regions of higher metallicity.
- The ratio of carbon and oxygen in a galaxy seems to be related to the galaxy type. Upon inspection of BPT diagrams on CLOUDY, when adding an auxiliary axis displaying log(C/O), a clear line is formed that neatly separates AGN from star-forming galaxies.
- It can be concluded that passive galaxies are always redder than their active counterparts. This is due to them being dominated by colder stars which appear more red on the spectrum.

It has been shown that clear trends exist between certain physical properties of a galaxy such as mass, SFRand their classification. This is particularly obvious when considering metallicities, where AGN galaxies always dominate; and redness, where passive galaxies always appear redder. star-forming galaxies have shown the strongest trends and benefited most from the CLOUDY analysis; strong positive correlations were found between both mass-temperature, as well as mass-metallicity.

The results which were drawn in this report agree with those expected from each each galaxy classification. The use and calculation of certain parameters, such as SFR, led to SNAGs due to the values not necessarily agreeing with initial expectations. However, many of the conclusions that we have

drawn agree with the properties expected for varying galaxy types in the local universe.

ACKNOWLEDGMENTS

The group would like to thank David Sobral for his supervision during the Astrophysics Group Project, and for providing unconditional support and valuable insight. Thanks to Lancaster University Physics Department for the inception of the PHYS369 module, allowing astrophysics to be pursued practically and this investigation to be made. We have benefited immensely from the public available programming language Python, including NumPy & SciPy, Matplotlib, Astropyand the Topcat analysis program (Taylor 2005). This research has made use of the VizieR catalogue access tool, CDS, Strasbourg, France. We have benefited from publicly available data from the Sloan Digital Sky Survey (SDSS a). Funding for SDSS-III has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, and the U.S. Department of Energy Office of Science. The SDSS-III web site is http://www.sdss3.org/. 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, Carnegie Mellon University, University of Florida, the French Participation Group, the German Participation Group, Harvard University, the Instituto de Astrofisica de Canarias, the Michigan State/Notre Dame/JINA Participation Group, Johns Hopkins University, Lawrence Berkeley National Laboratory, Max Planck Institute for Astrophysics, Max Planck Institute for Extraterrestrial Physics, 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.

REFERENCES

Aihara H., et al., 2011, ApJS, 193, 29
Ananna T. T., et al., 2019, ApJ, 871, 240
Baldwin J. A., Phillips M. M., Terlevich R., 1981, PASP, 93, 5
Bell E. F., McIntosh D. H., Katz N., Weinberg M. D., 2003, ApJS, 149, 289
Bergvall, Nils Marquart, Thomas Way, Michael J. Blomqvist, Anna Holst, Emma Östlin, Göran Zackrisson, Erik 2016, A&A, 587, A72
Best P. N., Sobral D., Cirasuolo M., Smail I., Matsuda Y., Geach J. E., 2012, MNRAS, 420, 1926
Brinchmann J., Charlot S., White S. D. M., Tremonti C., Kauffmann G., Heckman T., Brinkmann J., 2004, MNRAS, 351, 1151
Calzetti D., Armus L., Bohlin R. C., Kinney A. L., Koornneef J., Storchi-Bergmann T., 2000, ApJ, 533, 682
Cattaneo A., et al., 2009, Nature, 460, 213
Crenshaw D. M., Kraemer S. B., George I. M., 2003, ARA&A, 41, 117
Eldridge J. J., Stanway E. R., Xiao L., McClelland L. A. S., Taylor G., Ng M., Greis S. M. L., Bray J. C., 2017, Publications of the Astronomical Society of Australia, 34, e058
Faber S. M., et al., 2007, ApJ, 665, 265
Feltzing, S. Johnson, R. A. 2002, A&A, 385, 67
Ferland G. J., et al., 2013, Revista mexicana de astronomía y astrofísica, 49, 137
Folkes S., et al., 1999, MNRAS, 308, 459
Gabor J. M., DavÃl' R., Finlator K., Oppenheimer B. D., 2010, MNRAS, 407, 749
Grin D., Hirata C. M., 2010, Phys. Rev. D, 81, 083005
Guo H., et al., 2019, ApJ, 871, 147
Hubble E., 1929, Proceedings of the National Academy of Sciences, 15, 168
Kennicutt Jr. R. C., 1998, ARA&A, 36, 189
Kewley L. J., Ellison S. L., 2008, ApJ, 681, 1183
Kewley L. J., Dopita M., Sutherland R. S., Heisler C., Trevena J., 2001, ApJ, 556
Kewley L. J., Groves B., Kauffmann G., Heckman T., 2006, MNRAS, 372, 961
Kewley L. J., Maier C., Yabe K., Ohta K., Akiyama M., Dopita M. A., Yuan T., 2013, ApJ, 774, L10
Lacey C. G., Baugh C. M., Frenk C. S., Helly J. C., Malbon R., Cole S., Bower R. G., Benson A. J., 2006, MNRAS, 370, 645
Lehnert M. D., Heckman T. M., 1996, ApJ, 472, 546
Maungkorn S., Kriwattanawong W., 2017, Journal of Physics: Conference Series, 901, 012004
Nakamura O., Fukugita M., Brinkmann J., Schneider D. P., 2004, AJ, 127, 2511
SDSS, The Eighth SDSS Data Release (DR8), `http://www.sdss3.org/dr8/`
SDSS, The Optical Spectra Catalogs, `https://www.sdss.org/dr12/spectro/catalogs/#Objectinformation`
Saadeh D., Feeney S. M., Pontzen A., Peiris H. V., McEwen J. D., 2016, Physical Review Letters, 117, 131302
Sobral D., N. Best P., Matsuda Y., Smail I., Geach J., Cirasuolo M., 2011, MNRAS, 420
Sobral D., Smail I., Stott J. P., Best P. N., Geach J. E., Cirasuolo M., Matsuda Y., Kurk J., 2012, MNRAS, 428, 1128
Sobral D., Smail I., Best P., Geach J., Matsuda Y., Stott J., Cirasuolo M., Kurk J., 2013, MNRAS, 428, 1128
Sobral D., Stroe A., Koyama Y., Darvish B., Calhau J., Afonso A., Kodama T., Nakata F., 2016, MNRAS, 458, 3443
Sobral D., et al., 2018a, MNRAS, 477, 2817
Sobral D., et al., 2018b, MNRAS, 482, 2422
Springel V., et al., 2005, Nature, 435, 629
Storchi-Bergmann T., Schmitt H. R., Calzetti D., Kinney A. L., 1998, AJ, 115, 909
Taylor M. B., 2005, in Shopbell P., Britton M., Ebert R., eds, Astronomical Society of the Pacific Conference Series Vol. 347, Astronomical Data Analysis Software and Systems XIV. p. 29
Tempel, E. Saar, E. Liivamägi, L. J. Tamm, A. Einasto, J. Einasto, M. Müller, V. 2011, A&A, 529, A53
Tremonti C. A., et al., 2004, ApJ, 613, 898
Wang L., Kauffmann G., 2008, MNRAS, 391, 785
Weinmann S. M., van den Bosch F. C., Mo H. J., Yang X., 2006, MNRAS, 366, 2
Zeldovich Y. B., 1972, MNRAS, 160, 1P

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