Analyzing the nebular metallicity-stellar mass relation in galaxies

Samuel Mucedola,[1]*
[1] University of Milano-Bicocca, Milano, IT

27 June 2023

ABSTRACT
We present an investigation on the relation between stellar mass (M_*) and the metallicity of the intra-galactic gas, also known as *nebular metallicity* $Z_{\rm neb}$. This paper sets the aim to develop a simplified but yet informative model about how the metal mass in galaxies behaves during star formation and which physical processes partake in producing or removing metals from galaxies. After developing such model, we will show how it performs comparing to observations obtained from the Sloan Digital Sky Survey (SDSS) and we will compare the results with the literature.

Key words: Metallicity – Stellar Mass – Galaxy evolution

1 INTRODUCTION

The model we are about to propose tries to answer the following specific question:

What is the Stellar mass vs Metallicity relation for high and low stellar mass galaxies?

The mass-metallicity relation (MZR) in galaxies is a complex problem in modern day research. It has been tackled already by many such as Tremonti et al. (2004), Davé et al. (2012) and Lilly et al. (2013). The following discussion will rely on their work and it will try to provide a more intuitive understanding of the relationship between nebular metallicity and stellar mass by introducing two models: the *closed box* model and the *open box model*. The latter borrows the idea of a gas reservoir system Lilly et al. (2013), which adopts a much more sophisticated system. We will make this model more accessible by introducing some constraining assumptions to simplify the discussion.
The idea is to show the limits of the *closed box* model which assumes no gas inflow or outflow from the galaxy, effectively leaving stellar evolution as the only possibility for metal enrichment in galaxies, and then move into an *open box* approach, which allows galaxies to communicate with their surroundings. The open box is a very general model which can be extended for a series of feedback processes and inflows of various types, which is why the work done here is not sufficient but serves as an introduction to motivate further studies.

2 METHODS AND MATERIALS

In this section we present the theoretical models used for the analysis, describing the physical variables and the assumption that went into their formulation.

* E-mail: s.mucedola@campus.unimib.it

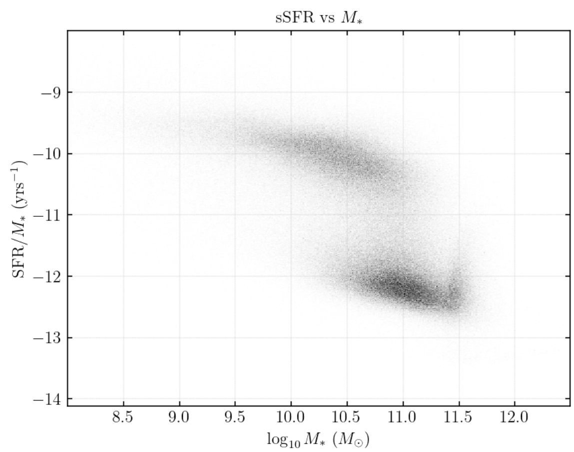

Figure 1. Specific star formation rate (sSFR) vs Stellar mass (M_*), the trend is constant for both populations

2.1 Datasets

For this investigation we used some catalogs **"Sloan Digital Sky Survey" (SDSS)** available online, in particular we narrowed our investigation to the *sdss_mpajhu_catalogue.fits* which contains 53400 galaxies at a redshift lower than 0.3, which allows us to neglect their redshift distribution. The catalogue includes variables like

- Right ascension and declination
- Redshift best estimate
- Stellar mass: The median estimate of the log10 total stellar mass PDF using model photometry ($M\odot$)
- Star formation rate: The median estimate of the log10 total SFR PDF, per unit year. This is derived by combining nebular emission line measurements and model fits to the integrated photometry when emission line measurements are not possible.
- Nebular metallicity : The median estimate of the Oxygen abun-

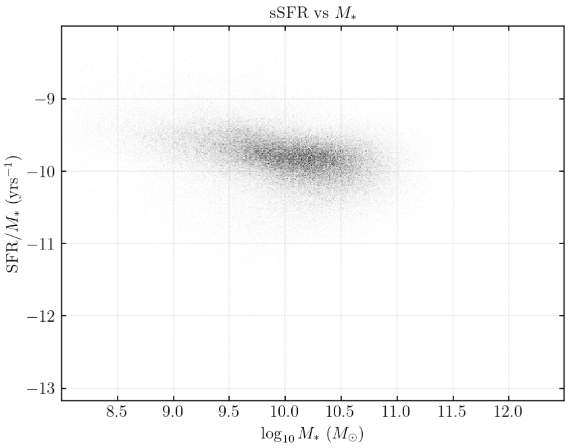

Figure 2. Specific star formation rate (sSFR) vs Stellar mass (M_*) for data with defined nebular metallicity. This is the population displayed in Figure 3

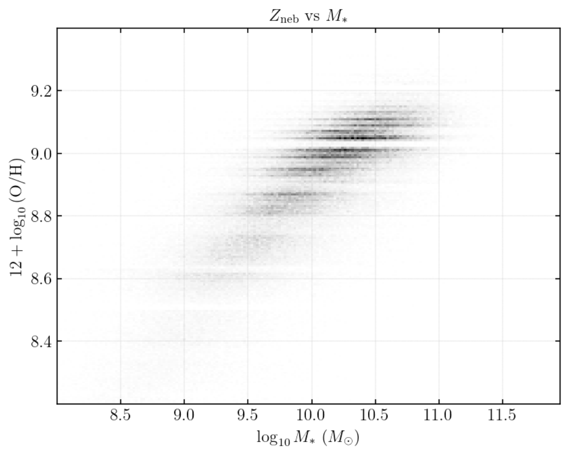

Figure 3. Nebular metallicity vs Stellar Mass plot for the population displayed in Figure 2

dance derived using Charlot & Longhetti models, see Charlot & Longhetti (2001). The values are reported as $12 + \log_{10}$ O/H. See Tremonti et al. (2004) for more.

- Redshift flag: Redshift is reliable if this flag is 0, unreliable redshift if this flag is non-zero
- Parameter flag: Physical parameters are reliable if this flag is 1, unreliable if this flag is 0

We analyzed data that is considered reliable so we filtered the data both for the redshift flag and the parameter flag, then we found that many data had a nan value for many entries and we made sure to filter these out too.
In fig.1 we notice the presence of two well defined populations, one more massive than the other, we will define the dividing line between the two with a **treshold** in mass $\sim 10^{10.5} M_\odot$. Fig.2 also shows that the SDSS catalogue provides defined metallicity values only for the first population, the one with higher SFR, which is generally called **star forming population** or also main population. In fig.3 we see a linear trend in metallicity that flattens out for high mass galaxies, this seems to suggest two different models for high and low mass galaxies.

2.2 Physical models

The two models previously mentioned are described in this section, with the related maths and the assumptions that went into their formulation. Both models assume that the galaxy is located inside a Dark Matter Halo of virialized mass $M_{\rm vir}$ and use the following differential equation for the evolution of the galaxy's gas mass

$$\frac{dM_{\rm gas}}{dt} = \dot{M}^{\rm in}_{\rm gas}(t) - \dot{M}^{\rm out}_{\rm gas}(t) - \dot{M}_* \tag{1}$$

2.2.1 Closed box definition

The closed box model is the simplest setup one can think about when describing the galaxy's gas mass evolution. We assume the galaxy can't communicate with the outside, this means that

$$\dot{M}^{\rm in}_{\rm gas}(t) = \dot{M}^{\rm out}_{\rm gas}(t) = 0$$

So the new evolution equation for the gas mass becomes

$$\frac{dM_{\rm gas}}{dt} = -\dot{M}_* = -\mathrm{SFR} \tag{2}$$

Where SFR stands for star formation rate. Now, it's fair to assume that stars form from the gravitational collapse of gas in galaxies, which can only happen if the fluid's kinetic pressure can't counteract the gravitational force, this is the **Jeans criterion**, which relies on the hydrostatic equilibrium condition. For the collapse we will require that

$$M_{\rm gas} > M_J(T, Z, ...) \tag{3}$$

Where M_J is the Jeans' mass and it's a combination of the properties of the gas, such as metallicity, temperature, density and so on. If the gas of the galaxy is able to collapse, then we expect a part of it to form stars, we will call this fraction $\epsilon < 1$ and we will require that

$$\mathrm{SFR} = \epsilon\ (M_{\rm gas})^\alpha \tag{4}$$

We hypothesized a linear relation, so $\alpha = 1$. To test this hypothesis, you could plot the surface density of the galaxy disk and the star formation rate for many galaxies. This test has already been done by Schmidt (1959), what was found is an almost linear trend

$$\Sigma_{\rm SFR} \propto (\Sigma_{\rm gas})^\alpha \qquad \alpha \in (1, 1.5) \tag{5}$$

So our assumption is consistent with previous observations, so we will use

$$\mathrm{SFR} = \epsilon M_{\rm gas} \tag{6}$$

Substituting into the differential equation and solving for $M_{\rm gas}$ yields

$$M_{\rm gas} = M_0 e^{-\epsilon(t-t_0)} \tag{7}$$

Where t_0 is the time at which the galaxy formed and $M_0 = M(t_0)$ is its initial mass, due to the fact that at primordial stages the galaxy is just a collection of gas particles and has not formed any stars yet so $M_{\rm gas}(t_0) = M_{\rm tot}$ Now we can solve for the stellar mass

$$M_* = \int_{t_0}^{t} dt\ \mathrm{SFR} = -\int dt\ \epsilon M_0\ e^{-\epsilon(t-t_0)} \tag{8}$$

$$M_* = M_0 \left(1 - e^{-\epsilon(t-t_0)}\right) \tag{9}$$

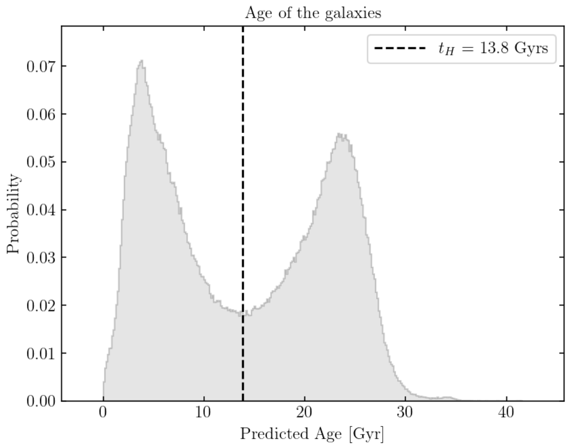

Figure 4. Distribution of the age of galaxies according to the closed box model, t_H is the Hubble time

Now we have all the tools to try to explain the graph in fig.1, that is a constant trend in the sSFR vs M_* relation, we calculate the SFR as the derivative of the stellar mass and find

$$\mathrm{SFR}(M_*) = M_* \frac{\epsilon}{e^{\epsilon(t-t_0)} - 1} \tag{10}$$

So it seems that the constant trend at the time of observation $t = 0$ is justified if

$$\frac{\epsilon}{e^{-\epsilon t_0} - 1} = \mathrm{const} \longrightarrow t_0 = \mathrm{const} \tag{11}$$

If we assume that our efficiency is constant with mass, notice that t_0 = const actually means **constant with the stellar mass**, this means that if the efficiency is independent with time, the age of the galaxy is independent of the stellar mass, this allows us to assume a very simple stellar profile like the **single burst model**.
The value for ϵ that satisfies the constant with stellar mass condition is given by Kennicutt (1998), which gives two results for ϵ, the one that satisfies the condition in eq.10 requires the introduction of the dynamical time for galaxies. Since this overcomplicates our model, we will use the result of $\epsilon = 0.25\ \mathrm{Gyr}^{-1}$ which assumes SFR = $M_*^{1.4}$.

We can validate this model by simply plotting the predicted age for our two populations (fig.4). The plot seems to estimate well the age of one of two populations. This population is the one with defined metallicity so we can use this as a predictive model.

2.2.2 *Metallicity in the closed box*

We start by defining the metallicity as $Z = M(\mathrm{O})/M(\mathrm{H})$, the oxygen content in galaxies will be referred as M_Z so the metal mass of the galaxy, while $M(H) \sim M_{\mathrm{gas}}$ due to the large imbalance of hydrogen content in gas.
In a closed box model the only processes that change the metal content of the galaxy are:

- Star formation: which deprives the galaxy of its heavier elements to produce the core of stars, this will enter our equation as $-Z$ SFR
- Supernovae explosions: which gives back the metals to the galaxy, this is regulated by the so called *metal yield fraction* y, so y SNR, with SNR standing for supernovae rate.

If you assume the principal yield component of supernovae explosion to be due to Type II supernovae, you can assume SNR = SFR since almost all of the massive stars explode, so their explosion rate is almost exactly their birth rate. Substituting Z we find

$$\frac{dM_Z}{dt} + M_Z\ \epsilon = \mathrm{y\ SFR} \tag{12}$$

which has the following analytical solution

$$M_Z(t) = e^{-A(t)} \cdot \int_{t_0}^{t} \mathrm{y\ SFR}(t')\ e^{A(t')}\ dt' \tag{13}$$

with $A(t) = \int_{t_0}^{t} \epsilon\ dt = \epsilon\ (t - t_0)$

Assuming ϵ to be constant with time, we find

$$M_Z(t) = \mathrm{y}\ M_{\mathrm{gas}}(t_0)\ \epsilon\ (t - t_0)\ e^{-\epsilon\ (t-t_0)} \tag{14}$$

And finally

$$Z(t) = \frac{M_Z(t)}{M_{\mathrm{gas}}(t)} = \mathrm{y} \cdot \epsilon\ (t - t_0) \tag{15}$$

By substituting the expression for the age used for fig.4 we can recast this as a function of the stellar mass

$$Z(M_*) = \mathrm{y}\ \log\left(\frac{1}{1 - M_*/M_0}\right) \tag{16}$$

In the result section we plot this curve over the data, using the definition of metallicity provided by the dataset

$$\tilde{Z} = 12 + \log_{10}(Z(M_*)) \tag{17}$$

For the values of the metal yield and M_0 we used the y vs M_0 relation in Tremonti et al. (2004).

2.2.3 *Open box definition*

In the open box model we review certain assumptions and allow the galaxy to interact with the environment. This now means that you can have inflow or outflow and the total mass of your galaxy can change. For our purposes we will make a very stringent assumption that we didn't justify which is that **the total mass of the galaxy remains unchanged** which puts us in the situation of having

$$\dot{M}^{\mathrm{in}}_{\mathrm{gas}}(t) = \dot{M}^{\mathrm{out}}_{\mathrm{gas}}(t) \neq 0$$

This allows us to maintain the same gas profile of the closed box model, which we know is unphysical for one population but consistent with the one with defined metallicity. The metal mass now has an addition:

- $\dot{M}^{\mathrm{in}}_Z = Z_{\mathrm{igm}} \dot{M}^{\mathrm{in}}_{\mathrm{gas}}$ the inflow of metal mass is mainly due to accretion of inter galactic gas, IGM for short. Observations show that this gas is always very pure, so we will assume the presteen gas approximation and say that $Z_{\mathrm{IGM}} = 0$ so we don't have to worry about this term
- $\dot{M}^{\mathrm{out}}_Z = -Z\eta$ SNR this is the feedback due to the supernovae explosion rate, this outflow is very relevant for low and medium mass galaxies but not for high mass galaxies. The Z factor is the result of the assumption of *homogeneous distribution of metals in the galaxy* The actual threshold at which it starts to become less relevant is somewhere in the ballpark of $10^{11} M_\odot$ (ref?) so it's a good enough approximation for our entire population. We will consider only this source of feedback but it's not the only factor.

With that said we can now solve for M_Z the following equation, keeping the same SN Type II assumption

$$\frac{dM_Z}{dt} = [y + Z]\ \text{SFR} - Z\eta\ \text{SFR} \tag{18}$$

we can adjust the terms, this gives

$$\frac{dM_Z}{dt} + Z(1+\eta)\ \text{SFR} = y\text{SFR} \tag{19}$$

and simplify Z with SFR

$$\frac{dM_Z}{dt} + M_Z(1+\eta)\epsilon = y\text{SFR} \tag{20}$$

we recognize this differential equation to be the one encountered in the closed box model, the solution is

$$M_Z(t) = e^{-A(t)} \cdot \int_{t_0}^{t} y\ \text{SFR}(t')\ e^{A(t')}\ dt' \tag{21}$$

with $A(t) = \int_{t_0}^{t}(1+\eta)\epsilon\ dt = (1+\eta)\epsilon\ (t - t_0)$

After substituting the SFR with the time dependent equation for M_{gas} and solving the integral steps we end up with

$$M_Z = \frac{y}{\eta} M_0\ e^{-\epsilon(t-t_0)} \tag{22}$$

Now this is similar to the solution we found with the closed box model, in fact

$$M_Z^{\text{open}} = \frac{\epsilon}{\eta} \frac{1}{t - t_0} M_Z^{\text{closed}} \tag{23}$$

And since we are considering the same gas profile for both models this holds

$$Z^{\text{open}} = \frac{\epsilon}{\eta} \frac{1}{t - t_0} Z^{\text{closed}} \tag{24}$$

which means that

$$Z^{\text{open}} = \frac{y}{\eta} \tag{25}$$

the metallicity in our open box model doesn't depend on the age of the galaxy.
What is left is to understand the stellar mass dependence of the SN feedback rate. We will recover the result cited in Davé et al. (2012), which suggests a dependence of the type $\eta \propto M_*^{-1/3}$. We will introduce a proportionality constant and use

$$Z^{\text{open}} = \frac{y}{a\ M_*^{-1/3}} = \frac{y}{a}\ M_*^{1/3} \tag{26}$$

again, we recovered the values of M_0 and y from Tremonti et al. (2004)

2.3 Results

The closed box model performs the worst, it fails to represent the data apart from some instances at high mass. This might be due to the small number of high mass isolated galaxies, since these are the only ones that could be fairly represented by such a model (fig.5).

The open box model is reasonable, we estimated the proportionality constant a by instancing a non linear fit, the `curvefit` routine of `scipy.optimize`. Which gives

$$a = 4.44 \pm 1.03 \times 10^{-5} M_\odot^{1/3} \tag{27}$$

The resulting plot shows a simplified fit but still reasonable as 1st order approximation (fig.6)

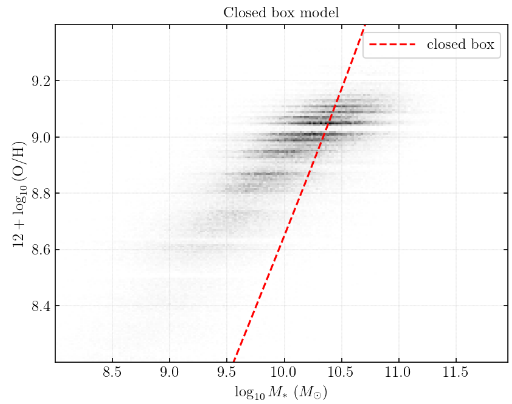

Figure 5. Closed box model over the data population (not a fit)

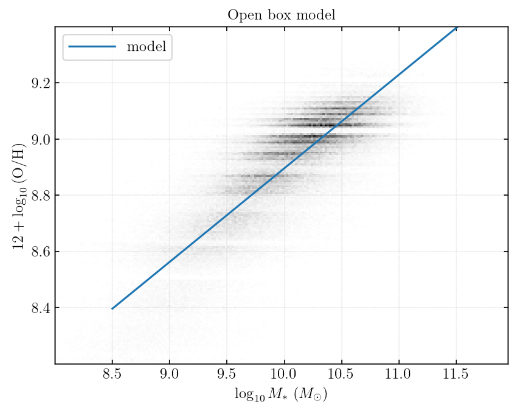

Figure 6. Open box model with one free parameter a fitted to the data

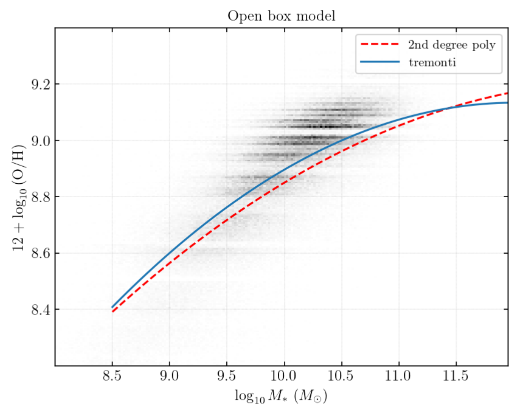

Figure 7. Tremonti vs our 2nd degree polynomial fit

As suggested by Tremonti et al. (2004), a more data driven approach could help us understand the subtleties of our model when you add in all the other feedbacks. We emulated the work by Tremonti and ran a Polynomial Regression to estimate the coefficients and compare them with Tremonti's (fig.7).

$$y = -1.492 + 1.847 \log_{10}(M_*) - 0.0803 \left(\log_{10}(M_*)\right)^2 \quad \text{Tremonti}$$

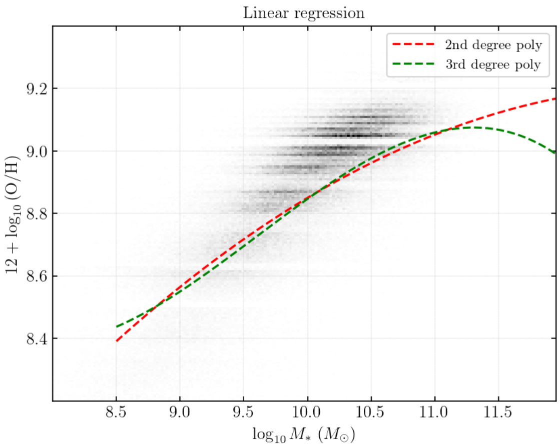

Figure 8. 2nd and 3rd degree polynomial fits

$$y = 0.66 + 1.389 \log_{10}(M_*) - 0.056 \left(\log_{10}(M_*)\right)^2 \quad \text{2nd degree fit}$$

We also suggest a third polynomial variant because for us it's more representative of the plateau at high masses (fig.8).

$$\begin{aligned} y =& 52.91 - 14.74 \log_{10}(M_*) + 1.598 \left(\log_{10}(M_*)\right)^2 + \\ & - 0.0564 \left(\log_{10}(M_*)\right)^3 \qquad \text{3rd degree fit} \end{aligned}$$

3 CONCLUSIONS

We were able to present and validate a simplified model that explains the Metallicity vs Stellar Mass relation. Our assumptions made this possible so we suggest for future studies to relax a few of these assumptions. The main suggestion would be to change the gas mass profile for the open box model, allowing for a reservoir system in equilibrium, also extending the differential equation for the metal's mass allowing for some sort of inflow that could be due to other effects such as mergers. Finally one could add other outflows due to AGN feedback, which becomes more relevant at higher masses. This last one is the easiest of the extension of the model and could resolve the plateau problem at high masses.

3.1 Tremonti y vs M_0 relation

The data used for y and M_0 are retrieved from a table in Tremonti et al. (2004), we computed a simple interpolation to extend the range of the metal yield. The one used in both models assume a M_0 of $10^{11.35} M_\odot$ and so y = $10^{-2.01}$ which was given by the table, we ended up not needing the extended values. The fit is shown in fig.9.

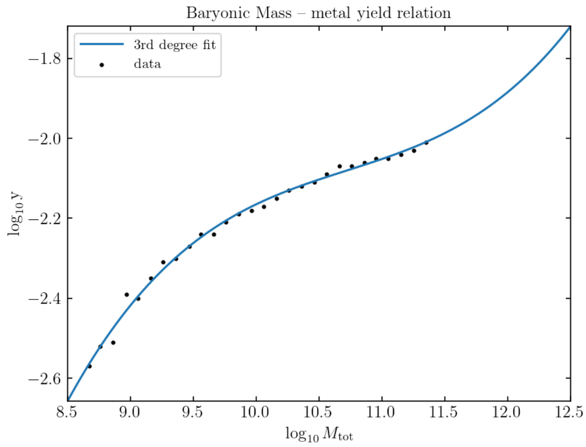

Figure 9. Fit of observation data of baryon mass vs effective yield of supernovae

REFERENCES

Charlot S., Longhetti M., 2001, Monthly Notices of the Royal Astronomical Society, 323, 887
Davé R., Finlator K., Oppenheimer B. D., 2012, MNRAS, 421, 98
Kennicutt R. C., 1998, Annual Review of Astronomy and Astrophysics, 36, 189
Lilly S. J., Carollo C. M., Pipino A., Renzini A., Peng Y., 2013, ApJ, 772, 119
Schmidt M., 1959, ApJ, 129, 243
Tremonti C. A., et al., 2004, ApJ, 613, 898