MNRAS **000**, 1–5 (2023) Preprint 29 June 2023 Compiled using MNRAS LATEX style file v3.0

Analyzing the nebular metallicity-stellar mass relation in galaxies

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

29 June 2023

ABSTRACT

We present an investigation into the correlation between stellar mass (M_*) and the metallicity of the intra-galactic gas, commonly referred to as *nebular metallicity* $Z_{\rm neb}$. The objective of this study is to develop a simplified yet informative model elucidating the behavior of metal mass in galaxies during the process of star formation, as well as the underlying physical mechanisms responsible for metal production and removal within galaxies. Subsequently, we will assess the performance of this model by comparing it with observational data obtained from the Sloan Digital Sky Survey (SDSS) and compare our results with the existing 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 forthcoming model aims to address a specific question regarding the relationship between stellar mass and metallicity in galaxies, focusing on both high and low stellar mass galaxies. The mass-metallicity relation (MZR) in galaxies poses a complex problem in contemporary research, which has been extensively explored by researchers such as Tremonti et al. (2004), Davé et al. (2012), and Lilly et al. (2013) Drawing upon their work, this study aims to provide a more intuitive understanding of the association between nebular metallicity and stellar mass by introducing two models: the *closed box* model and the *open box* model.

The *closed box* model assumes no gas inflow or outflow from the galaxy, considering stellar evolution as the sole process for metal enrichment in galaxies. In contrast, the *open box* model allows galaxies to interact with their surroundings, incorporating the concept of a gas reservoir system as proposed by Lilly et al. (2013). To facilitate the discussion, we will introduce certain simplifying assumptions that make the models more accessible.

Our objective is to illustrate the limitations of the *closed box* model and then transition to the more comprehensive *open box* approach, which accommodates various feedback processes and inflow mechanisms. It is important to note that this work serves as an introductory exploration, motivating further studies in this area.

★ E-mail: s.mucedola@campus.unimib.it

2 METHODS AND MATERIALS

In this section, we present the theoretical models utilized for the analysis and provide an overview of the physical variables along with the assumptions made during their formulation.

2.1 Datasets

For this investigation, we utilized the **Sloan Digital Sky Survey (SDSS)** catalogs available online. Specifically, we focused our analysis on the *sdss_mpajhu_catalogue.fits*, which contains information on 53,400 galaxies with a redshift lower than 0.3, allowing us to disregard their redshift distribution. The catalog encompasses the following variables:

- 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 abundance derived using Charlot & Longhetti models, see Charlot & Longhetti (2001). The values are reported as 12 + $\log_{10}$ O/H. For further details, refer to Tremonti et al. (2004).
- Redshift flag: A flag indicating the reliability of the redshift measurement. A value of 0 denotes a reliable redshift, while a value of 1 indicates an unreliable redshift.
- Parameter flag: A flag indicating the reliability of the physical parameters. A value of 1 denotes reliable parameters, while a value of 0 denotes unreliable parameters.

We performed data analysis exclusively on reliable data, filtering the dataset based on both the redshift flag and the parameter flag. Furthermore, we identified and excluded data entries with missing values (NaN) to ensure data integrity.

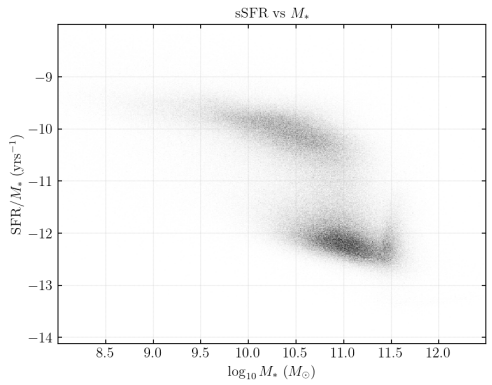

Figure 1. Specific star formation rate (sSFR) vs Stellar mass (M_*), the trend is constant for both populations

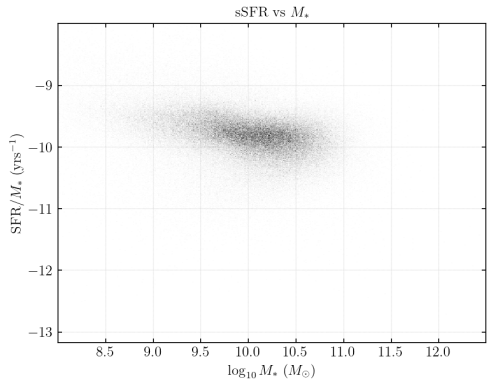

Figure 2. Specific star formation rate (sSFR) vs Stellar mass (M_*) for data with defined nebular metallicity. This is the population displayed in 3

In Figure 1, we observe the presence of two distinct populations, with one exhibiting higher mass compared to the other. To distinguish between these populations, we established a threshold in mass at approximately $10^{10.5} M_\odot$.
Figure 2 further illustrates that the SDSS catalog provides well-defined metallicity values exclusively for the first population, which is characterized by higher star formation rates. This population is commonly referred to as the "star-forming population" or the "main population."
Examining Figure 3, we observe a linear trend in metallicity that levels off for high-mass galaxies. This suggests the need for distinct models for high and low mass galaxies.

2.2 Physical models

In this section, we provide a description of the two aforementioned models, including the mathematical equations and the underlying assumptions used in their formulation. Both models assume that the galaxy resides within a Dark Matter Halo of virialized mass $M_{\rm vir}$. The evolution of the galaxy's gas mass is governed by the following differential equation:

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

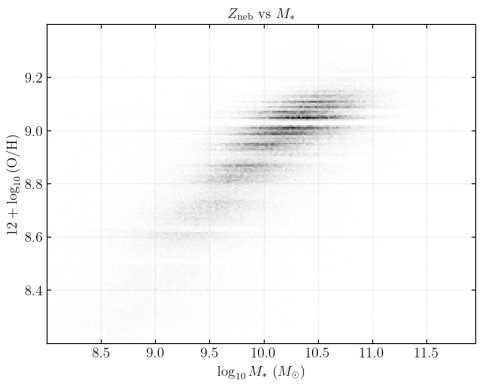

Figure 3. Nebular metallicity vs Stellar Mass plot for the population displayed in Figure 2

where $\dot{M}^{\rm in}_{\rm gas}(t)$ represents the gas inflow rate, $\dot{M}^{\rm out}_{\rm gas}(t)$ denotes the gas outflow rate, and $\dot{M}_*$ represents the star formation rate.

2.2.1 *Closed box definition*

The closed box model represents the simplest scenario for describing the evolution of a galaxy's gas mass. In this model, we assume that the galaxy is unable to interact with its external environment. Consequently, the gas inflow rate and gas outflow rate are both assumed to be zero:

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

Thus, the modified evolution equation for the gas mass becomes:

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

Here, SFR represents the star formation rate. It is reasonable to assume that stars form as a result of the gravitational collapse of gas within galaxies. This collapse occurs when the kinetic pressure of the gas cannot counteract the gravitational force, as determined by the **Jeans criterion** based on the hydrostatic equilibrium condition. For collapse to occur, we require:

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

Here, M_J represents the Jeans mass and is determined by various properties of the gas, such as its metallicity, temperature, density, and so on. If the gas within the galaxy is capable of collapsing, we expect a fraction of it to form stars. We denote this fraction as ϵ (where $0 < \epsilon < 1$), and we assume that the star formation rate is proportional to a power law of the gas mass:

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

We have hypothesized a linear relation with $\alpha = 1$. To test this hypothesis, we can examine the surface density of the galaxy disk and the star formation rate for multiple galaxies. This test has been previously conducted by Schmidt (1959), who found an almost linear trend given by:

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

Our assumption is therefore consistent with previous observations, so we will adopt the following relation

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

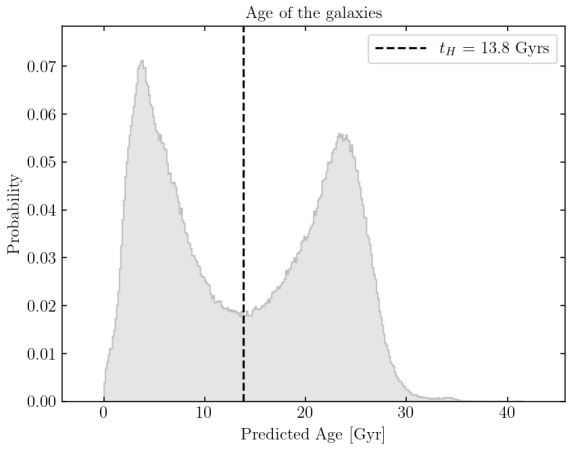

Figure 4. Distribution of the age of galaxies according to the closed box model, t_H is the Hubble time

By substituting this into the differential equation and solving for $M_{\rm gas}$, we obtain:

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

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

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

This yields:

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

With these equations, we can now attempt to explain the graph in Figure 1, which shows a constant trend in the specific star formation rate (sSFR) versus stellar mass (M_*) relation. We calculate the SFR as the derivative of the stellar mass and find:

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

Hence, the constant trend at the time of observation ($t = 0$) can be justified if:

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

Indeed, if we assume that our efficiency (ϵ) is constant with mass, the condition $t_0 = $ const implies a constant value with respect to the stellar mass. This means that if the efficiency remains independent of time, the age of the galaxy is independent of the stellar mass. Consequently, we can adopt a simple stellar profile, such as the **single burst model**.
The value for ϵ that satisfies the constant-with-stellar-mass can be obtained from Kennicutt (1998). They provide two possible results for ϵ, but using the value that satisfies the condition in equation 10 would introduce the dynamical time for galaxies, which would overly complicate our model. Therefore, we will utilize the result $\epsilon = 0.25\ {\rm Gyr}^{-1}$, which assumes a star formation rate (SFR) proportional to $M_*^{1.4}$.
We can verify the validity of this model by creating a plot that showcases the predicted ages for both populations (see Figure 4). The plot indicates a successful estimation of the age for one of the populations, specifically the population with known metallicity. Hence, we can consider this model as a potentially effective predictive tool for estimating ages based on the metallicity criterion.

2.2.2 *Metallicity in the closed box*

In the closed box model, the metallicity (Z) is defined as the ratio of oxygen mass (M(O)) to hydrogen mass (M(H)). The metal mass of the galaxy is denoted as M_Z, while the hydrogen mass is approximated by the gas mass ($M_{\rm gas}$) due to the significant predominance of hydrogen in the gas.
In this model, there are two primary processes that influence the metal content of the galaxy:

- Star formation: This process reduces the metal content of the galaxy as heavier elements are consumed in the formation of stellar cores. Mathematically, this is represented as $-Z$ SFR.
- These events contribute to the replenishment of metals in the galaxy. The rate at which this occurs is regulated by the metal yield fraction (y) and the supernova rate (SNR). Thus, the metal contribution from supernovae can be expressed as y · SNR.

Under the assumption that the principal yield component of supernova explosions is due to Type II supernovae, we can approximate the supernova rate (SNR) to be equal to the star formation rate (SFR) since most of the massive stars explode as supernovae. Substituting the expression for metallicity (Z), we arrive at the following differential equation:

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

This differential equation has an analytical solution:

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

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

Assuming ϵ to be constant with time, we find

$$M_Z(t) = {\rm y}\ M_{\rm gas}(t_0)\ \epsilon^{-1}\ (t - t_0)\ e^{-\epsilon\ (t-t_0)} \tag{14}$$

And finally

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

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

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

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

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

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 relax certain assumptions and allow the galaxy to interact with its environment, including the possibility of gas inflow and outflow. However, we will make a stringent assumption

that the total mass of the galaxy remains unchanged. This assumption leads us to the following condition:

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

This allows us to maintain the same gas profile of the closed box model. In the open box model, we consider two additional terms for the metal mass evolution:

- Inflow of metal mass: $\dot{M}^{\rm in}_Z = Z_{\rm igm}\dot{M}^{\rm in}_{\rm gas}$. This term accounts for the inflow of metal mass due to the accretion of intergalactic gas (IGM). However, since observations indicate that the IGM is highly pristine and has negligible metallicity ($Z_{\rm IGM} = 0$), we can disregard this term in our analysis.
- Outflow of metal mass: $\dot{M}^{\rm out}_Z = -Z\eta$ SNR. This term represents the feedback from supernovae explosions, where metals are expelled from the galaxy. The outflow rate is assumed to be proportional to the metallicity Z, and the parameter η represents the efficiency of metal ejection. The supernovae rate (SNR) is considered to be proportional to the star formation rate (SFR). This outflow of metal mass is particularly relevant for low and medium mass galaxies, while it becomes less significant for high mass galaxies (above a threshold of approximately $10^{11} M\odot$). We will only consider this source of feedback in our model, although it is important to note that other factors can contribute to metal outflows in reality.

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

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

we can adjust the terms, this gives

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

and simplify Z with SFR

$$\frac{dM_Z}{dt} + M_Z(1+\eta)\epsilon = \mathrm{ySFR} \tag{20}$$

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

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

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 $M_{\rm gas}$ and solving the integral steps we end up with

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

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

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

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

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

which means that

$$Z^{\rm open} = \frac{\mathrm{y}}{\eta}. \tag{25}$$

The metallicity in our open box model doesn't depend on the age of the galaxy.

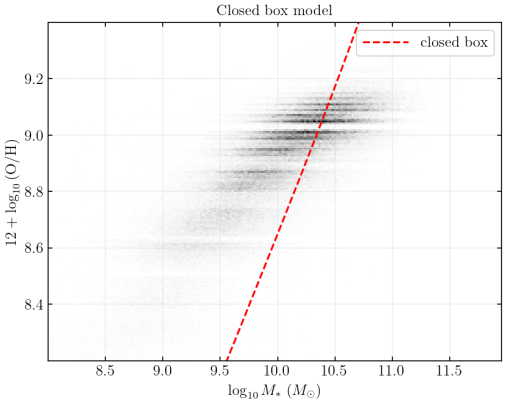

Figure 5. Closed box model over the data population (not a fit)

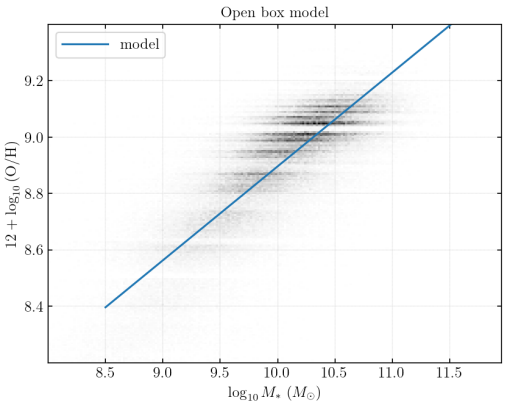

Figure 6. Open box model with one free parameter a fitted to the data

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

$$Z^{\rm open} = \frac{\mathrm{y}}{a\ (M_*/M_0)^{-1/3}} = \frac{\mathrm{y}}{a}\ (M_*/M_0)^{1/3} \tag{26}$$

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

2.3 Results

The closed box model shows poor agreement with the data, except for a few cases at high mass. This might be caused by the small number of high mass isolated galaxies, since these are the only ones that could be accurately represented by such a model (fig.5).

In contrast, the open box model performs better and provides a more realistic representation of galaxy evolution. We estimated the proportionality constant a by instancing a non linear fit, the `curvefit` routine of `scipy.optimize`. Which gives

$$\mathrm{a} = 4.44 \pm 1.03 \times 10^{-5} \tag{27}$$

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

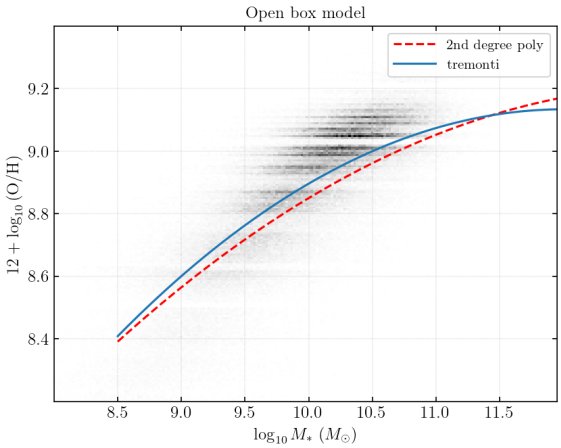

Figure 7. Tremonti vs our 2nd degree polynomial fit

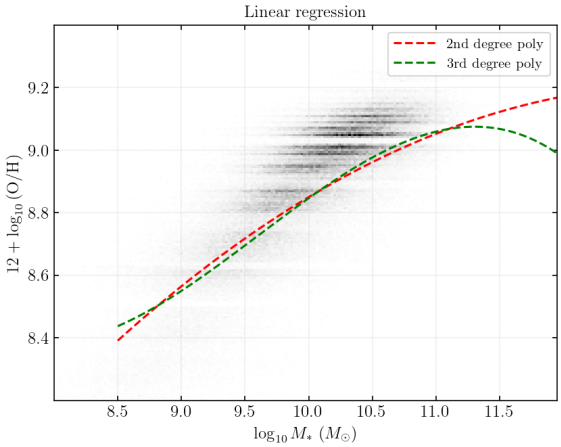

Figure 8. 2nd and 3rd degree polynomial fits

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

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

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

We also looked into a third polynomial variant since it appears to be more representative of the plateau at high masses (fig.8).

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

2.4 Tremonti yield vs baryonic mass relation

Regarding the Tremonti yield vs baryonic mass relation, we used data from a table in Tremonti et al. (2004) to obtain the values for the yield factor (y) and the initial mass (M_0). We performed a simple interpolation to extend the range of the metal yield, although we found that the values provided in the table were sufficient for our analysis. The fitting of the yield factor is illustrated in Fig. 9. In both models we used $M_0 = 10^{11.35} M_\odot$ and $y = 10^{-2.01}$.

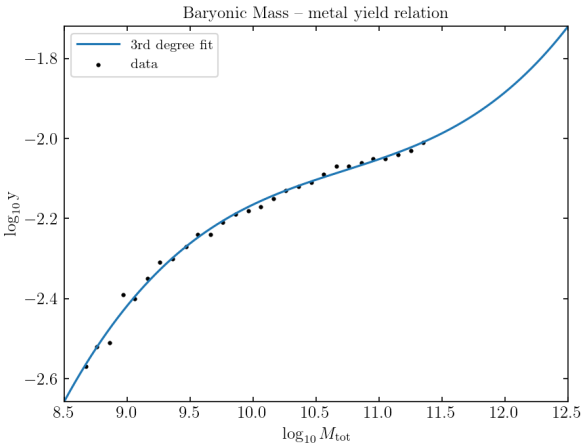

Figure 9. Fit of observation data of baryon mass vs effective yield of supernovae

3 CONCLUSIONS

In conclusion, we have presented and validated a simplified model that successfully explains the Metallicity vs Stellar Mass relation observed in galaxies. However, our model is based on several assumptions, and we suggest that future studies consider relaxing some of these assumptions to improve the model's accuracy.

One possible improvement is to modify the gas mass profile in the open box model to allow for a reservoir system in equilibrium. This would better account for the gas inflow and outflow processes that occur in galaxies.

Additionally, extending the differential equation for the metal mass to include other sources of inflow, such as gas accretion from mergers, could provide a more comprehensive description of the metallicity evolution.

Furthermore, incorporating additional outflows due to AGN (Active Galactic Nuclei) feedback could address the plateau problem observed at high masses. This extension would account for the impact of AGN feedback on the metallicity distribution of galaxies, particularly at higher masses.

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