Astronomy & Astrophysics manuscript no. output
November 21, 2022

Deep Mixing In Red Giants

Investigating The Effects Of Mass And Metallicity On The Rate Of Deep Mixing

Roy Lim z5310629
Supervisor: Associate Professor Sarah Martell

UNSW Sydney

November 21, 2022

ABSTRACT

Context. Deep Mixing is a process that occurs in red giants characterized by a decrease in surface $\left[\frac{C}{Fe}\right]$, increase in $\left[\frac{N}{Fe}\right]$ and $^{13}C/^{12}C$ ratio.
While the process has been observed as early as the 1970s, the mixing rates' dependence on mass and metallicity, as well as how the process occurs remains a topic in research.
By understanding Deep Mixing, astronomers may have a more complete understanding of the chemical evolution of stars and galaxies over time.

Aims. Investigate the mixing rates dependence of mass and metallicity of stars
Methods. Using APOGEE stars, create multiple bins based on mass and metallicity. Measure mixing rates for each bin separately and collate/compare results.
Results. Mixing rate is negatively correlated with metallicity while mass did not have a clear effect on mixing rate. The presence of AGB stars and globular clusters in the data may have led to higher apparent rates of mixing. The drop in $\left[\frac{C}{N}\right]$ was shown for stars with mass $> 2M_{\odot}$, which is unexpected.
Conclusions. Further research needs to be conducted with more data sets to assess the impact of mass and metallicity on the mixing rate. In addition, more research is required for low $\left[\frac{C}{N}\right]$ stars and their potential causes. Overall, stars are pretty cool[*Citation Needed*]

Key words. Red Giants, Deep Mixing

1. Introduction

To accurately model the stellar evolution of stars, it is essential to have a deep understanding of how the internal structure and chemical compositions of stars change as they age based on stellar parameters that can be observed such as metallicity and mass. Deep mixing is a currently poorly understood phenomenon, whereby a low mass red giant ($<2.2M_{\odot}$) will experiences a notable change in surface abundance of $\left[\frac{C}{Fe}\right]$ and $\left[\frac{N}{Fe}\right]$, accompanied by an increase between ^{13}C relative to ^{12}C. Previous research on this topic focused on globular clusters or stars that did not have explicit measurements of mass. As a result, deep mixing in stars outside of globular clusters as well as the dependence of the mixing rate on mass is poorly understood.

This paper project aims to help explore more of how mass and metallicity may affect the mixing rate within stars. To do this, the APOKASC catalog is used to train a machine-learning model to predict the mass of stars with the APOGEE catalog. The stars are sorted into bins by mass and metallicity and the mixing rate of each bin is then calculated by measuring how $\left[\frac{C}{N}\right]$ changes as the star ages (as $\log(g)$ decreases).

The paper will be structured like so: section 2 will provide a general overview of useful knowledge /on the topic. Section 3 will explain how a machine learning model was created using the APOKASC catalog as a training set while section 4 will go over the main methodology on how the mixing rate is found, as well as how the significance of mass and metallicity on the mixing rate was evaluated. Section 5 will discuss the results and raise any issues within the data while the conclusion 6 discusses potential future research as well as provide a summary of the key results found.

2. General Background

2.1. Stellar Abundance and Metallicity

In astronomy, any element that is not hydrogen or helium is considered to be a metal. The metallicity of a star is given by the ratio between metals and other elements as a fraction that is given by:

$$Z = 1 - X - Y$$

Where:

- Z is the metallicity.
- X is the fraction of the star (by mass) that is hydrogen.
- Y is the fraction of the star (by mass) that is helium.

In stars, the most common method of determining the relative abundance of elements is by comparing them with the Sun using a logarithmic scale. For convenience, the relative abundance or

ratio between an element to the hydrogen of star ⋆ is commonly denoted as:

$$\left[\frac{Fe}{H}\right]_{\star} = \log\left(\frac{N_{Fe}}{N_H}\right)_{\star} - \log\left(\frac{N_{Fe}}{N_H}\right)_{\odot}$$

Where:

- N_X denotes the number of atoms of X per unit volume.
- $\odot$ indicates that the parameters belong to the Sun.

For simplicity, the ⋆ will be omitted.

One unusual notation that is seen is alpha abundance $\left[\frac{\alpha}{Fe}\right]$. Alpha abundance measures the number of elements formed via the alpha process (whereby a nucleus absorbs an alpha particle). Alpha elements include O, Ne, Mg, Si, S, Ca, and Ti. The α abundance in a population of stars correlates with the star formation rate at the time they formed, and $\left[\frac{\alpha}{Fe}\right]$ in an individual star affects its internal structure and evolution.

To determine the relative abundance between elements A and B, we will need to subtract the relative abundance of A to C and the relative abundance of B to C like so:

$$\begin{aligned}\left[\frac{A}{C}\right] - \left[\frac{B}{C}\right] &= \log\left(\frac{N_A}{N_H}\right)_{\star} - \log\left(\frac{N_A}{N_H}\right)_{\odot} - \log\left(\frac{N_B}{N_H}\right)_{\star} - \log\left(\frac{N_B}{N_H}\right)_{\odot} \\ &= \log(\frac{N_A}{N_B}) - \log(\frac{N_A}{N_B})_{\odot} \\ &= \left[\frac{A}{B}\right] \\ \therefore \left[\frac{A}{B}\right] &= \left[\frac{A}{C}\right] - \left[\frac{B}{C}\right]\end{aligned}$$

This allows for the computation of $\left[\frac{C}{N}\right]$ given $\left[\frac{C}{Fe}\right]$ and $\left[\frac{N}{Fe}\right]$ (which are given in the data set used).

2.2. Movement And Position Of Celestial Objects

On Earth, to determine the location of any place on the surface, longitude and latitude are used. This idea is extended to provide a coordinate system for celestial bodies in the form of declination and right ascension. Declination equates to latitude while right ascension equates to longitude for celestial bodies.
To measure the motion of stars, the proper motion and radial Velocity of stars are measured. Proper motion for declination and right ascension measures the relative change in position in the sky while radial velocity measures the motion of the celestial body towards/away from Earth.

2.3. Other Common Stellar Parameters

This section will briefly go over the stellar parameters that can be observed and include other useful parameters that are not already mentioned.

The observable set of stellar parameters is element abundances (such as $\left[\frac{Fe}{H}\right]$ and $\left[\frac{\alpha}{H}\right]$), effective temperature (T_{eff}), and surface gravity ($\log(g)$).

The effective temperature is computed by using the Stefan-Boltzmann equation, which states that for a spherical star, the Luminosity (L) can be defined as:

$$L = 4\pi R^2 \sigma T_{\text{eff}}^4$$

Where:

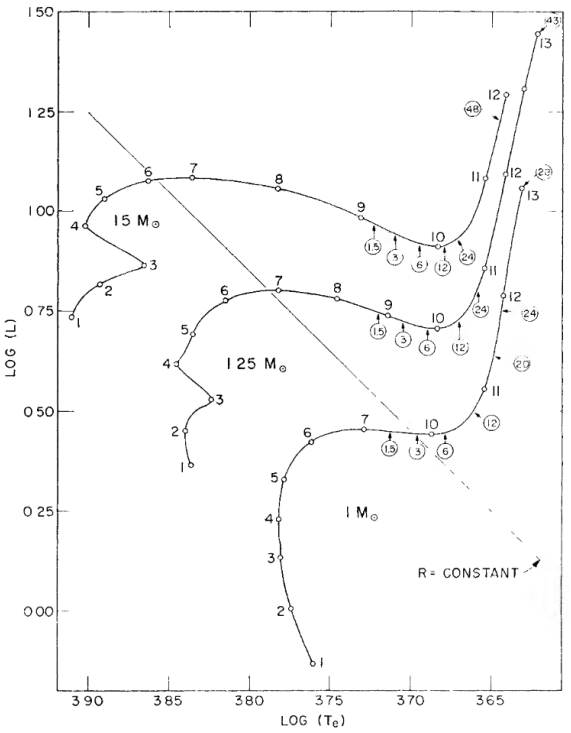

Fig. 1. Evolutionary paths for Population I stars of mass $M/M_{\odot}$ = 1, 1.25, and 1 5. Luminosity L is in solar units and surface temperature T_e is in degrees K. Circled numbers represent the factors by which surface ^{7}Li abundance has been depleted relative to its main sequence value. The straight line is a locus of constant radius R. (Iben (1967))

- R is the radius of the star;
- σ is the Stefan-Boltzmann constant (equal to $5.670400 \times 10^{-8} Wm^{-2}K^{-4}$
- T_{eff} is effective temperature.

This equation also approximates the star to a black body. The equation can then be rearranged to get T_{eff}.

Surface gravity is measured logarithmically as $\log(g)$ is calculated with gravity(g) being in $cm\ s^{-2}$.
The mass of the Sun is used as a unit of measurement and is usually denoted with the subscript $\odot$ like so:

$$1M_{\odot} \approx 1.99 \times 10^{30} kg$$

2.4. Deep Mixing and First Dredge Up

During a low/intermediate-mass star's evolution to become a red giant from being a main sequence star, it will undergo two main events (first dredge up and deep mixing) that cause the surface abundance of carbon and nitrogen to change. As the stars evolve from the main sequence, the core will no longer be undergoing hydrogen fusion, which causes the core to collapse inward due to gravitational pressure. This causes the convective zone near the near-surface of the star to begin to expand towards the core of the star.

The convective zone will transport the C-poor, N-rich material from the interior to the surface, lowering the $\left[\frac{C}{N}\right]$ at the surface ((Iben 1967)).

This process is commonly known as first dredge up (FDU) and occurs around point 6 as shown in figure 1. At the end of

FDU, the convective zone returns to its original depth.

Although the core is no longer undergoing hydrogen fusion, there is a region outside the core that is hot enough for H burning. This H-burning shell moves outwards towards the surface of the star over time.

Within an RGB star, the region between the H-burning shell and the surface is separated by a mean molecular weight barrier (also known as $\mu-$barrier), which is left behind at the lowest point reached by the convection zone during FDU.

When the H-burning shell has expanded outwards far enough to cross the $\mu-$ barrier, the star experiences a brief slow-down in its evolution and deep mixing should occur. The luminosity of the star at this point depends on the depth of FDU, which is controlled by the metallicity of the star. In a single-metallicity population of stars, the pause in stellar evolution causes a pile-up at a particular luminosity, which is referred to as the RGB bump (RGBb). In section 4, we use isochrones from the Dartmouth Stellar Evolution Database (Dotter et al. (2008)) to identify what the surface gravity $\log(g)$ will be where the RGBb is for different bins created based on mass and metallicity. Due to the higher speed of stellar evolution in high metallicity stars, stars with mass greater than $2.2M_{\odot}$ should not reach the RGBb before the end of their RGB lifetime, meaning deep mixing should not occur for those stars(Nataf et al. (2013)).

Thermohaline mixing (one of the suspected mechanisms of deep mixing) should only occur after the RGBb. Once the hydrogen-burning shell has crossed the $\mu-$barrier, mixing should be possible between the H-burning shell and the radiative zone above (Angelou et al. (2011)).

Since we are investigating deep mixing as a phenomenon in RGB stars, it is expected that mass has a negative correlation with the mixing rate. This is due to higher mass stars evolving faster and spending less time as an RGB star relative to stars with lower mass, leading to less time for mixing to occur (Iben (1967)). Similarly, lower metallicity also leads to lower mixing rates as lower metallicity stars evolve faster relative to high metallicity stars and can convert carbon to nitrogen more efficiently.

In addition, the fusion of carbon to nitrogen takes less time than the fusion of nitrogen into other elements (oxygen via the CNO cycle), this leads to a depletion of carbon and an increase in nitrogen over time. As a result, we expect stars $\left[\frac{C}{N}\right]$ to decrease steadily.

Deep mixing and the mixing rate can be quantified by looking at the $\frac{d[C/N]}{d\log(g)}$ since surface gravity tends to decrease with age (as stars age, their radius increases and thus decreasing surface gravity).

2.5. APOGEE and APOKASC

APOGEE stands for Apache Point Observatory Galactic Evolution Experiment (Majewski et al. (2017)). APOGEE provides abundance for many elements needed in the project such as Carbon, Nitrogen, and Oxygen, as well as the effective temperature and surface gravity. The stars observed consist largely of RGB and main sequence stars, making it ideal for observing the change in surface abundance. The data were obtained at `https://www.sdss4.org/dr17/irspec/spectro_data/` and data release 17 was used for this project.

The stars use 2MASS ID, which can then be used to cross-identify stars from other surveys.

The APOKASC catalog combines asteroseismic data from the Kepler satellite to provide mass estimates for stars in APOGEE, which will be used to train a machine-learning model later on in the project. The data used in this project comes from the second release (APOKASC-2, Pinsonneault et al. (2018)).

2.6. Data Processing

To ensure that the results are accurate, the data is first filtered by removing stars that are flagged as bad in parameters relevant to the project.
The flags used are shown below:

- ASPCAPFLAG
- FE_H_FLAG
- C_FE_FLAG
- N_FE_FLAG
- O_FE_FLAG
- RV_FLAG

ASPCAPFLAG is a flag used to show analysis on a variety of stellar parameters are uncertain (such as bad measurements for effective temperature or surface gravity). The elemental flags such as FE_H_FLAG are for stars with uncertain abundances (in this case, it means that $\left[\frac{Fe}{H}\right]$ is unreliable). RV_FLAG is related to a star's radial velocity measurements.

3. Prediction of Mass

The APOKASC catalog with mass derived using asteroseismic data will be used to train a machine learning model that will predict the mass of all the stars used in the APOGEE data set. Two machine learning methods will be used and compared to determine which will more accurately determine the mass.

To do this, the data set will be split into 80% being used to train the data. 20% of the data is then used to validate the predictions by calculating the correlation coefficient (r) and mean square error ($RMSE$). These two can be calculated using the formulas:

$$r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum\sum(x_i - \bar{x})^2(y_i - \bar{y})^2}}$$

$$RMSE = \frac{\sqrt{\sum_{i=1}^{N}(\hat{y}_i - y_i)^2}}{N}$$

The correlation coefficient varies between 1 and -1 where it measures the difference between the mean values (represented by $\bar{x}$ and $\bar{y}$) versus the actual values (represented as x_i and y_i).

For root mean square error, $\hat{y}_i$ is the predicted mass, y_i represents actual mass and N represents the total number of observations. For an ideal model, the RMSE should be as close to 0 (as this means the prediction match up with the actual masses) while the correlation should be as close to 1 as possible (meaning the mass can be predicted perfectly using the parameters given).

For each method, we provide seven variables or features to the regressor to analyze:

- Effective Temperature $T_{\rm eff}$
- Surface Gravity $log(g)$
- Metallicity $\left[\frac{Fe}{H}\right]$
- Alpha Abundance $\left[\frac{\alpha}{Fe}\right]$
- Carbon Abundance $\left[\frac{C}{Fe}\right]$

- Nitrogen Abundance $\left[\frac{N}{Fe}\right]$
- Oxygen Abundance $\left[\frac{O}{Fe}\right]$

T_{eff} and $\log(g)$ were chosen as they vary with time in a way that is dependent on $\left[\frac{\alpha}{Fe}\right]$ and $\left[\frac{Fe}{H}\right]$. Carbon and nitrogen abundance are chosen as they are only changed during FDU and deep mixing, which are dependent on metallicity and mass. Oxygen abundance should decrease slightly during FDU and deep mixing, and it is measured using similar spectral features as carbon and nitrogen.

The two methods are provided by the **Scikit-Learn** Python library, specifically **RandomForestRegressors** and **LinearRegression** were used.

3.1. Linear Regression

Linear Regression is by far the simplest way of predicting mass by assuming that each variable can be interpreted to add a linear factor as shown below.

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n$$

where $x_1, x_2, \ldots, x_n$ are the features used to predict y (mass) while the $\beta_1, \beta_2, \ldots, \beta_n$ are the associated weight placed on each feature. The model will attempt to minimize the distance of each data point from the least square regression line.

Linear regression runs into the problem of working most optimally when parameters are independent of each other, which is not true in this case. An example of this in stars is that surface gravity and effective temperature are usually linked as both are dependent on the size of the star.

3.2. Linear Regression Results

The following two graphs show the results when predicting mass using linear regression.

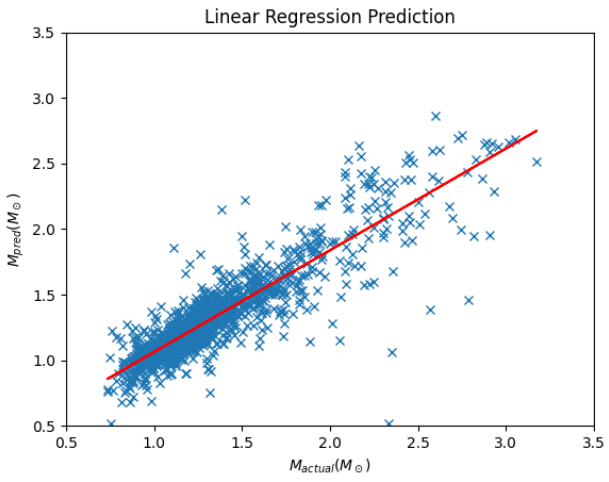

Fig. 2. Linear Regression Results

Linear regression tends to be overestimate near the 1.0 to 1.5 $M_\odot$ range while underestimating for the <1 $M_\odot$ and >1.75 $M_\odot$ range as shown in figure 3. This is shown by the predicted mass deviating more from the actual mass for stars with a mass between 2 and 3$M_\odot$.

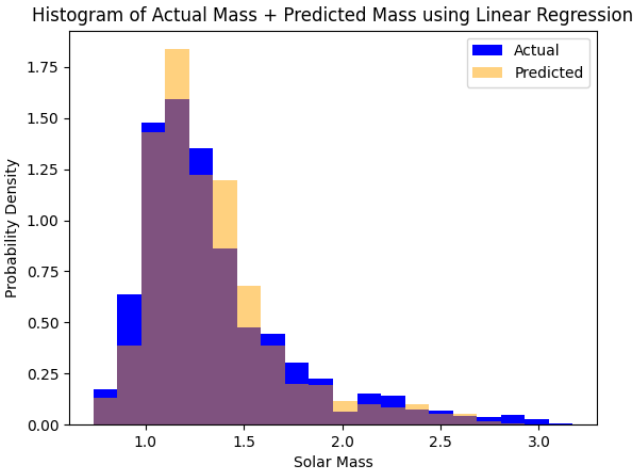

Fig. 3. Linear Regression Fit Histogram

3.3. Random Forest Regression (RFR)

Random forest involves using a collection of decision trees to create predictions. Decision trees provide estimates by using the data to create binary rules (for example, if alpha abundance is greater than 0, increase the mass by 0.5$M_\odot$), which determines which path is traversed down a tree. The traversal of a path provides an estimate for mass. As a result, we can tweak the model by varying how many paths there are by varying how many splits there are and the depth of the tree. The accuracy of the tree is improved by using a Random Forest, which consists of multiple decision trees whose results are averaged to create the final results.

A drawback of this regression method is that it tends to not extrapolate well outside the training set and overestimates the most common elements in the training set.

3.3.1. RFR Results

The following two graphs show the results when predicting mass using a Random Forest Regression:

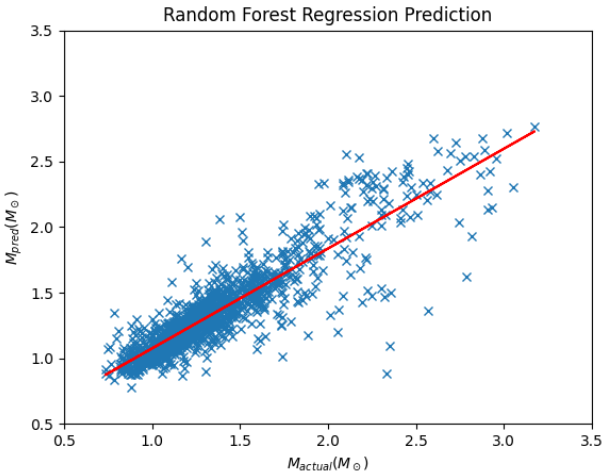

Fig. 4. Random Forest Regression Predictions

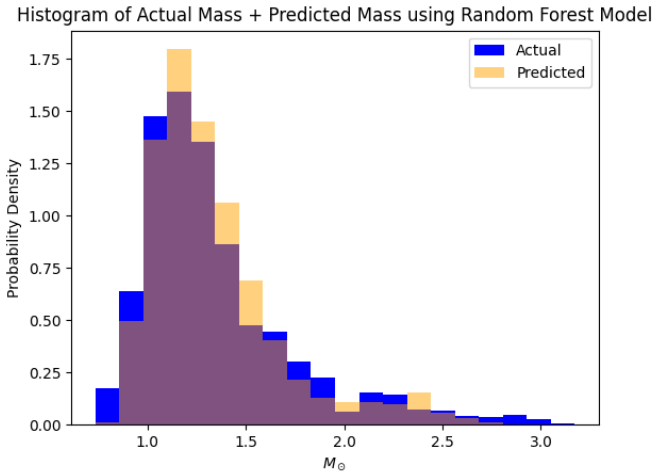

Fig. 5. Random Forest Regression Fit Histogram

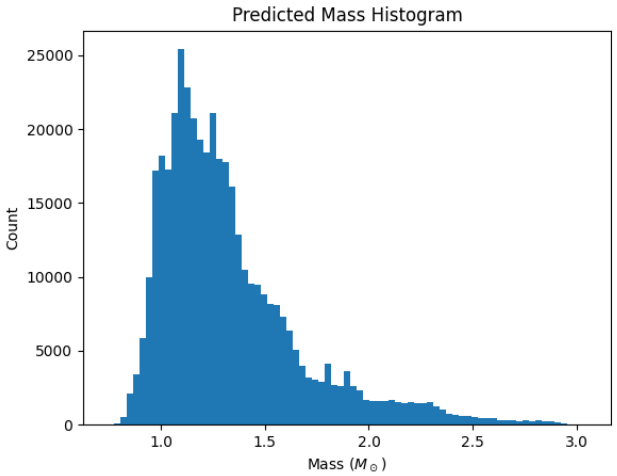

Fig. 6. Histogram of APOGEE Catalogues predicted mass using Random Forest Regressor and APOKASC as a training set.

As shown in Figure 5, one of this model's weaknesses in not being able to extrapolate for low ($< 1M_\odot$) and high ($> 2M_\odot$) stars is apparent. The lack of high mass and low metallicity stars may have affected the distribution of mass produced by the model. As a result, the number of low ($< 1M_\odot$) and high ($> 2M_\odot$) stars are underestimated as shown in 5.

3.4. Results and Verdict

Model	RMSE ($M_\odot$)	Correlation Coefficient
Random Forest Regression	0.342	0.883
Linear Regression	0.341	0.873

Given the result, both methods are similar in error and correlation. Random Forest was chosen over Linear Regression due to the higher correlation coefficient.

3.5. Prediction Of Mass

Using the Random Forest Regressor, the predicted masses for APOGEE stars are shown below:

The results show that most stars in APOGEE are between 1-1.5$M_\odot$, which illustrates the model's tendency to overestimate common data points in the training set as shown in figure 7.

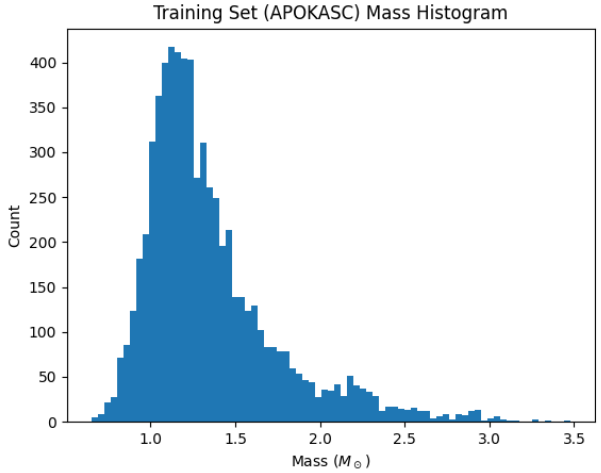

Fig. 7. Histogram of mass in the training set (APOKASC). There is a large number of stars between 1 and 1.5$M_\odot$, which the regressor model seems to have accounted for as shown in figure 6 where similar peaks exist.

4. Deep Mixing

To investigate the rate of mixing as a function of mass and metallicity, there are 3 distinct steps:

1. Select RGB stars;
2. Determine where deep mixing occurs;
3. Analyse the rate of mixing by separating the data into bins based on mass and metallicity and comparing the change in $\left[\frac{C}{N}\right]$ over surface gravity.

4.1. Selecting Red Giants

The first step to investigating deep mixing is to determine which stars are red giants.

After a main sequence star core runs out of hydrogen to fuse, the star will begin its evolution into a red giant. When the star becomes a red giant, it is characterized by a large radius, which causes the surface gravity and surface temperature to be lower relative to typical main-sequence stars. This difference can be visually discerned as shown in figure 8.

Stars with $log(g) < 3$ are chosen to exclude stars that are on the main sequence or early in the evolutionary stage between the main sequence and red giant from the analysis.

4.2. Overview of Red Giant Characteristics

Before we look at deep mixing, let's take a look at the data set.

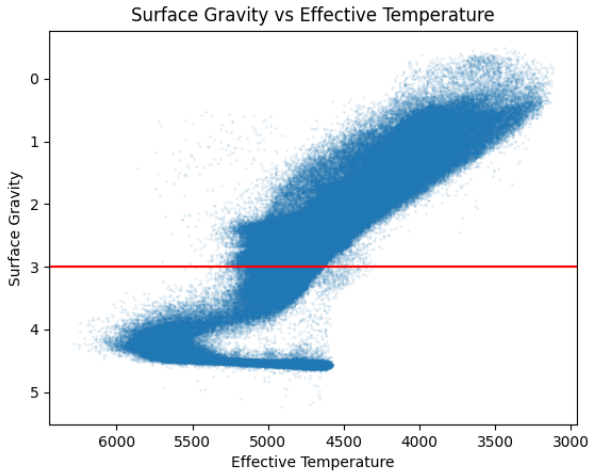

Fig. 8. Determining Red Giants: Stars that have a surface gravity lower than 3 were selected as indicated by the horizontal line.

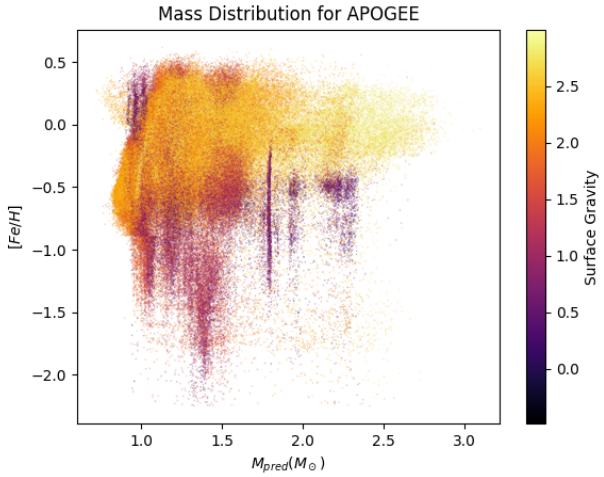

Fig. 9. Distribution of Mass vs Surface Gravity and Metallicity

From Figure 9, we can notice that there is a lack of intermediate/high mass ($M > 2-2.5M_\odot$), low metallicity ($[Fe/H] < -1.0$) stars.

4.3. Locating the Red Giant Branch Bump (RGBb)

The RGB bump is where theoretically deep mixing should begin for low-mass (<$2.2M_\odot$) stars. This can be seen in table 1 where for each bin with different metallicity and mass, the location/surface gravity of the RGBb was determined using synthetic isochrones from the Dartmouth Stellar Evolution Database (Dotter et al. (2008)).

4.4. Measuring Deep Mixing Rates

Deep mixing is a process that lowers carbon and increases nitrogen abundance. To assess the strength of how metallicity and mass affect this process, the change of $\left[\frac{C}{N}\right]$ over $\log(g)$ will be measured. To do this, stars are divided into bins based on their metallicity and mass.
To determine the deep mixing rate for each bin, the average metallicity and mass were computed before and after the theoretical RGB bump (surface gravity of the RGBb ±1).

The average metallicity and surface gravity are used to compute the deep mixing rate as shown below in figure 10:

According to the literature, the drop in $\left[\frac{C}{N}\right]$ is expected when the star reaches the RGBb (as shown by the red, vertical lines in figure 10) as this is where deep mixing occurs. All the graphs show a downward trend in $\left[\frac{C}{N}\right]$, with the decrease in $\left[\frac{C}{N}\right]$ most evident as the metallicity of the stars decreases. There is a lack of stars with surface gravity between $\log(g) = 1$ and $\log(g) = 2$ for the higher mass bins (more than $2.2M_\odot$) as well as a general lack of stars that have low mass and low metallicity.

For some plots (for example, the plots with metallicity between -0.25 and -0.75), the $\left[\frac{C}{N}\right]$ seems to recover at $\log(g) \approx 0-1$ as it has not declined relative to stars that have not reached the RGBb, which goes against the literature as the $\left[\frac{C}{N}\right]$ should be steadily declining and no known mechanism allows $\left[\frac{C}{N}\right]$ to increase.

The table below lists the measured gradient for each bin. Note that to compute the gradient, surface gravity is decreasing (this represents the progression of time as surface gravity decreases as a star ages to become a red giant), hence leading to a negative gradient.

Gradient	Average Mass	Average $\left[\frac{Fe}{H}\right]$
-0.098488	0.962165	-0.145190
-0.033388	1.122561	-0.028432
-0.044620	1.368209	-0.048708
-0.060148	1.726248	-0.016210
-0.031951	2.423191	-0.030428
-0.112865	0.944973	-0.474528
-0.163348	1.108145	-0.414558
-0.152119	1.352239	-0.395188
-0.139684	1.728852	-0.423012
-0.204246	2.291261	-0.356759
-0.139754	0.942655	-0.844139
-0.301095	1.101735	-0.957578
-0.293591	1.362183	-1.002540
-0.302892	1.738166	-0.938415
-0.460493	2.322289	-0.977833
-0.192639	0.941702	-1.435758
-0.728427	1.139367	-1.431329
-0.462092	1.352376	-1.456436
-0.663044	1.767123	-1.509758
-0.539132	2.348773	-1.519398
-0.234738	0.951835	-1.897233
-1.049823	1.188714	-1.897159
-0.536546	1.338187	-1.878491
-0.403549	1.803693	-1.901571
-0.545997	2.361422	-1.948989

By plotting the data we have above, we get the following graph in figure 11:

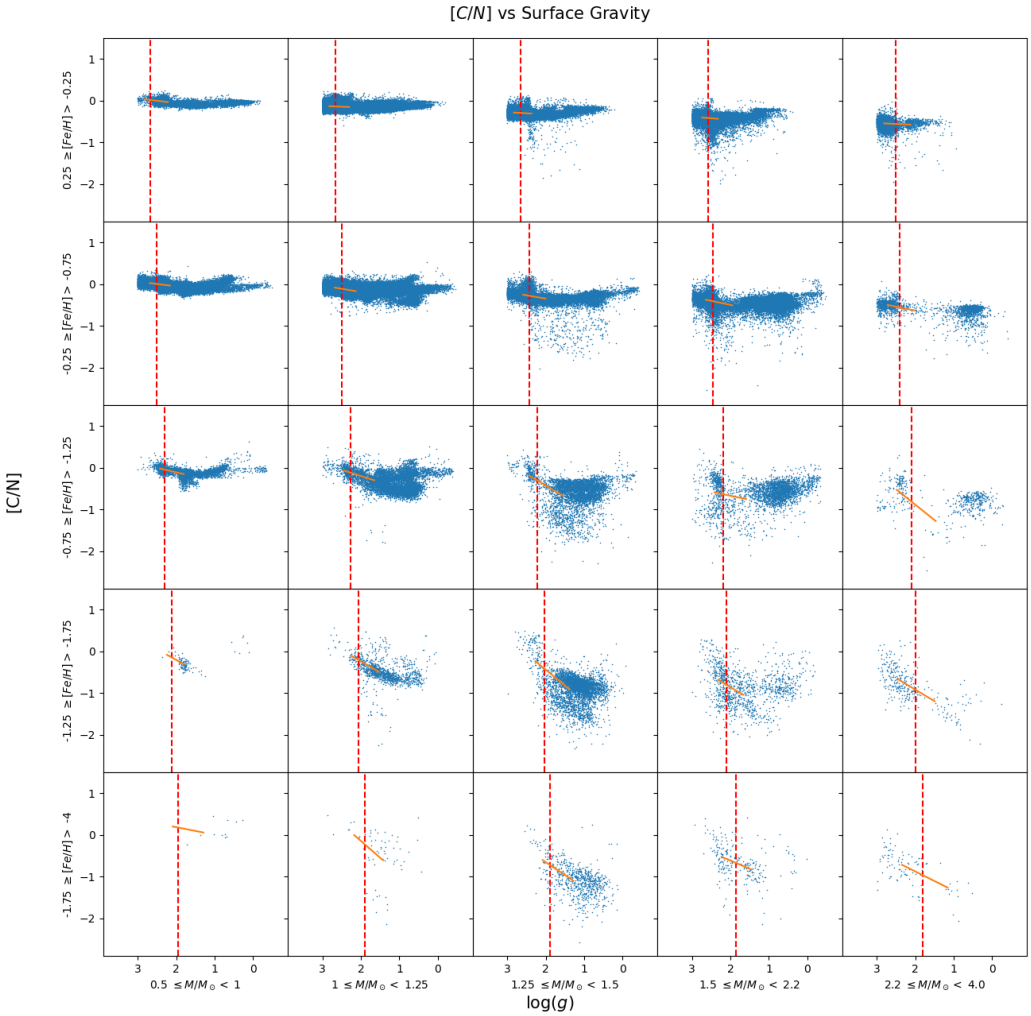

Fig. 10. The $\left[\frac{C}{N}\right]$ abundance ratio for stars, separated by mass and metallicity. The vertical red lines indicate where deep mixing should theoretically occur. The orange lines show the change between the average $\left[\frac{C}{N}\right]$ over surface gravity after deep mixing has begun.

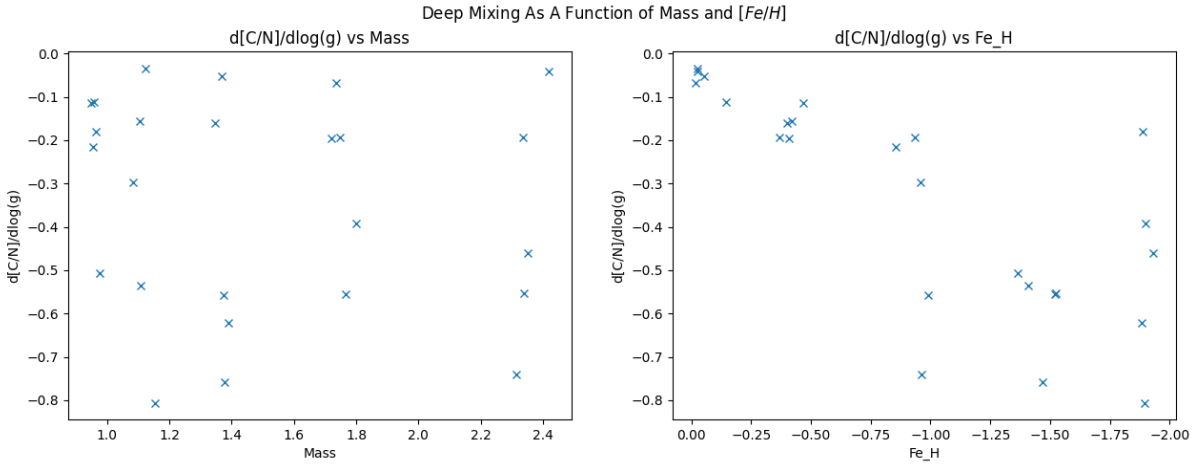

Fig. 11. Deep Mixing As A Function Of Mass and Metallicity. The correlation with metallicity is very negative while the relationship with mass is unclear.

Metallicity and Mass Ranges	0.25 ≤ [Fe/H] < -0.25	0.75 ≤ [Fe/H] < -0.75	-0.75 ≤ [Fe/H] < -1.25	-1.25 ≤ [Fe/H] < -1.75	-1.75 ≤ [Fe/H] < -2.25
$0.50 \leq M/M_{\odot} < 1.00$	2.502	2.330	2.158	1.985	1.838
$1.00 \leq M/M_{\odot} < 1.25$	2.549	2.312	2.090	1.915	1.175
$1.25 \leq M/M_{\odot} < 1.50$	2.562	2.324	2.087	1.882	1.765
$1.50 \leq M/M_{\odot} < 2.20$	2.328	2.211	2.090	1.947	1.779

Table 1. Location of RGBb in surface gravity for different mass and metallicity bins

5. Discussion

5.1. Deep Mixing As A Function Of Mass And Metallicity

As shown in figure 11, the rate of deep mixing has a clear relationship with metallicity, with lower metallicity leading to a much higher level of deep mixing. The relationship between mass and deep mixing is less apparent.
The literature predicts a negative correlation with both metallicity and mass. We expect a sharp cutoff in mixing for stars with mass greater than $2.2M_{\odot}$ as stars with high mass will evolve past the RGB phase without the H-burning shell reaching the $\mu-$barrier, where deep mixing theoretically occurs. This is not shown within the data as the high mass ($2.2M_{\odot} < M < 4M_{\odot}$) stars still showed a drop in $\left[\frac{C}{N}\right]$.

5.2. Anomalous Dip in $\left[\frac{C}{N}\right]$

Figure 10 shows that for all the stars after the deep mixing has occurred, which is shown by $\left[\frac{C}{N}\right]$ decreasing near the red line that indicates where the RGBb is, the $\left[\frac{C}{N}\right]$ should keep decreasing at a steady rate as the star ages (as shown by $\log(g)$ decreases).

However, looking in certain bins, some stars are noticeably lower in $\left[\frac{C}{N}\right]$ than the majority of stars. A good example is bins in the middle of figure 10 (specifically with $-0.75 >= \left[\frac{Fe}{H}\right] > -1.25$ and $1.25 \leq M/M_{\odot} < 1.5$).
This phenomenon cannot be explained with just deep mixing and is likely due to other factors.
There may be two reasons for this, (1) being Globular Clusters and (2) being AGB stars. To investigate this, stars with $\left[\frac{C}{N}\right]$ less than -1.0 will be extracted and analyzed below to determine whether they could be explained by either being stars that belong to a globular cluster or being an AGB star. Figure 12 shows which stars are extracted.

5.2.1. Globular Clusters

Globular clusters are collections of gravitationally bound stars. There are around 150 clusters in the Milky Way, they tend to orbit in the halo and tend to be old and metal-poor. While their formation is still a topic under debate, stars within clusters are hypothesized to form at the same time and place as each other, leading to a similar composition. The dip we see in our data may be partially explained if the data includes stars in globular clusters since around half of the stars in globular clusters are expected to have be $\left[\frac{C}{N}\right]$-weak. There is not currently a model for globular cluster formation that fully explains why this is the case.

To investigate whether Globular Clusters are the likely cause, stars with a low $\left[\frac{C}{N}\right]$ are selected:

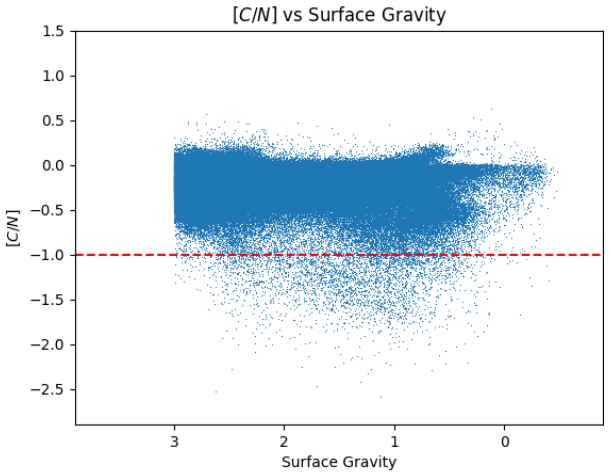

Fig. 12. $\left[\frac{C}{N}\right]$ vs Surface Gravity. Everything below the red horizontal line is considered a low $\left[\frac{C}{N}\right]$ star and will be extracted for investigation. The graph uses all stars with a surface gravity lower than 3.

To verify that the stars in question are from Globular Clusters, they must fit two criteria:

1. Similar movement (Proper motion in declination, proper motion in right ascension, and radial velocity).
2. Similar position (right ascension and declination).

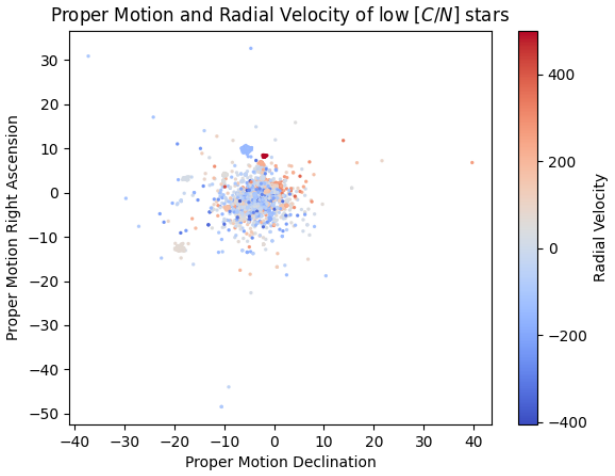

Fig. 13. Comparing the movement of low $\left[\frac{C}{N}\right]$ stars. The coloring is indicative of radial velocity.

Using the figures 13 and 14, we can see that some of the low $\left[\frac{C}{N}\right]$ stars belong to certain globular clusters as shown by the overlap between the orange (low $\left[\frac{C}{N}\right]$ stars) and blue (all stars) in

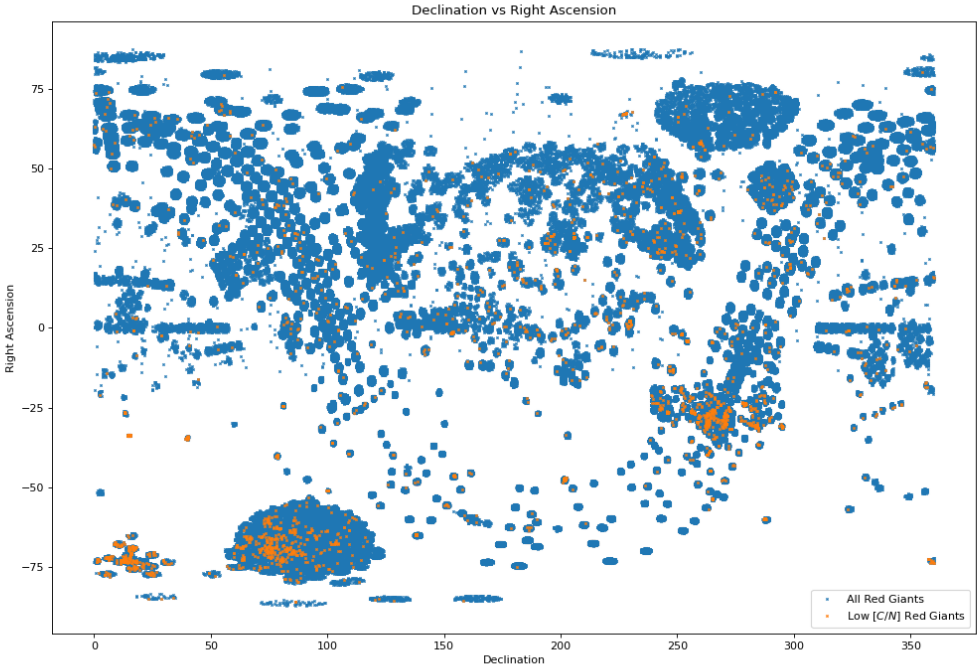

Fig. 14. Comparing the position of low $\left[\frac{C}{N}\right]$ stars. The orange data points are for low $\left[\frac{C}{N}\right]$ red giants while the blue data points consist of all the stars in the APOGEE data set.

figure 14. However, as shown in figure 13, few well-defined stars may be in clusters (exhibiting similar position, proper motion, and radial velocity). This suggests that most of the $\left[\frac{C}{N}\right]$ -weak stars do not appear to be in any globular clusters.

5.2.2. AGB (Asymptotic Giant Branch Stars)

AGB stands for Asymptotic Giant Branch. AGB stars describe a subset of low/intermediate-mass stars with $M < 8M_{\odot}$ in the later stage of evolution relative to a red giant. AGB stars will have gone through both FDU and deep mixing and thus should have a lower $\left[\frac{C}{N}\right]$ as a result.

These stars are characterized by being asymptotically close to the red giant branch stars (RGB) by having lower temperature vs surface gravity relative to regular red giants. This makes AGB stars difficult to separate from RGB stars based on photometry or stellar parameters due to their asymptotic nature causing overlaps in position.

The likely AGB stars have a surface temperature of around 6000-5000K and a surface gravity of around 0.5 to 2 in 15 as they are warmer and have lower gravity relative to most RGB stars.

There may be AGB stars that are not identified in the main group of stars.

5.2.3. Verdict

While some of the low $\left[\frac{C}{N}\right]$ stars in our data set are globular cluster members and AGB stars, they alone do not explain all of the low $\left[\frac{C}{N}\right]$ stars observed in figure 12.

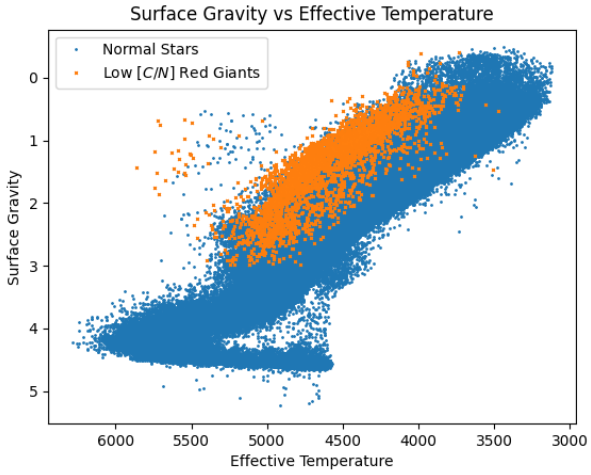

Fig. 15. Verifying whether low $\left[\frac{C}{N}\right]$ stars are AGB stars. Orange data points are for low $\left[\frac{C}{N}\right]$ red giants while the blue data points consist of all the stars in the APOGEE data set. The orange spots that are above the main group of stars are potential AGB stars.

5.3. Recovery in $\left[\frac{C}{N}\right]$

The theory predicts that once the $\left[\frac{C}{N}\right]$ has decreased as a result of deep mixing and first dredge up, the relative abundance between carbon and nitrogen should continue to decline as the star keeps evolving (as indicated by the decrease of $\log(g)$). However, in certain panels in figure 10, for stars with metallicity $\left[\frac{Fe}{H}\right]$ between -0.25 and -0.75, there is a notable increase in $\left[\frac{C}{N}\right]$ where the shape

of the data seems to curve upwards after the RGB bump instead of steadily decreasing as the literature predicts.

This phenomenon has not been noted in previous literature and hence warrants follow-up.

6. Conclusions

To summarise, within this project we were able to:

- Predict the mass of stars in the APOGEE data set by using APOKASC data and an appropriate regressor
- Analyse the mixing rate as a function of mass and metallicity
- Investigate the presence of low $\left[\frac{C}{N}\right]$ stars

The results demonstrate a negative correlation between mixing rate and metallicity while the relationship between mixing rate and mass could not be shown.
In addition, the literature predicted that for stars with mass greater than $2.2M_{\odot}$, deep mixing should not occur and $\left[\frac{C}{N}\right]$ should not drastically decrease. This was not observed as $\left[\frac{C}{N}\right]$ seems to decrease after the RGB bump.

Within the data set, two different phenomena are unexpected that may require further research to accurately determine the cause.

- Presence of low $\left[\frac{C}{N}\right]$ stars, whose causes could not be inferred.
- A recovery in $\left[\frac{C}{N}\right]$ as surface gravity decreases.

6.1. Future Work

For this project, there is a multitude of different paths where future research can take place.

The method used (Random Forest Regression) tends to be reflective of the data set, and it tends to struggle to estimate the mass of stars that do not lie within the range of the training set (this leads to an underestimation of the number of high and low mass stars due to the rarity of these stars in the training data set APOKASC). More advanced machine learning techniques coupled with a more extensive training set will hopefully produce more accurate and reliable results.

AGB stars can be more accurately identified via asteroseismic data rather than simply graphing their surface temperature and surface gravity due to their internal structure differing from other stars. In addition, the prediction of mass may also benefit from additional parameters.

More research needs to be conducted to verify what is causing some stars to have lower $\left[\frac{C}{N}\right]$ relative to the majority. A more rigorous analysis of whether $\left[\frac{C}{N}\right]$ - weak stars are in globular clusters should be done by cross-referencing the relative positions of known globular clusters with the APOGEE data set.

While the APOGEE data set allows for the analysis of $\left[\frac{C}{N}\right]$, it lacks lithium abundance, which is another element that can be used to verify the presence of deep mixing. This may be remedied by using other data sets. During the project, the GALAH catalog was used but due to the lack of stars with reliable C and N abundance measurements, the results are not included. Other data sets should be used that contain lithium to further investigate deep mixing effect on lithium abundance.

Acknowledgements. This research project is largely inspired by Christopher J. Brockett's (z5075621) Honours Thesis.

References

Angelou G. C., Lattanzio J. C., Church R. P., Stancliffe R. J., Dearborn D. S., Smith G. H., Tout C. A., Smith V. V., 2011, in Kerschbaum F., Lebzelter T., Wing R. F., eds, Astronomical Society of the Pacific Conference Series Vol. 445, Why Galaxies Care about AGB Stars II: Shining Examples and Common Inhabitants. p. 151
Dotter A., Chaboyer B., Jevremović D., Kostov V., Baron E., Ferguson J. W., 2008, ApJS, 178, 89
Iben Icko J., 1967, ApJ, 147, 624
Majewski S. R., et al., 2017, AJ, 154, 94
Nataf D. M., Gould A. P., Pinsonneault M. H., Udalski A., 2013, ApJ, 766, 77
Pinsonneault M. H., et al., 2018, ApJS, 239, 32