

The dimensionality of stellar abundance space for red clump stars from APOGEE data

Natalie Price-Jones, supervised by Professor Jo Bovy

ABSTRACT

Recent spectroscopic surveys have provided abundant data on stars in the Milky Way. We will analyze a subset of data from the APO Galactic Evolution Experiment (APOGEE) survey, in the form of red clump stars. These metal enriched stars offer a complex abundance space to explore, which we will do by studying the spectra themselves, rather than derived quantities. This direct use of the spectra facilitates the characterization of errors and systematic effects on absorption features. An understanding of these uncertainties allows accurate investigation of the multi-dimensional abundance space, which may reveal trends in abundances with stellar age and position throughout the Milky Way. It also provides a comparative method to assess uncertainties associated with current abundance calculations.

1. Introduction

A star’s history, from its formation environment to its migratory motions, can be traced through its composition. Calculating elemental abundances from stellar spectra establishes this composition. However, current methods for calculating these abundances require robust theoretical models and complete understanding of measurement uncertainties. At present, these methods still exhibit systematic trends with other stellar properties, such as effective temperature T_{eff} or surface gravity $\log(g)$ (Holtzman et al. 2015). Removing these contaminating effects from abundance calculations entirely may reveal statistically significant correlations between abundances and stellar age or location within a galaxy.

Recent large-volume spectroscopic survey data from the APO Galactic Evolution Experiment (APOGEE) (Majewski et al. 2015), and upcoming data from the Gaia mission (Jordan 2008) will offer rich spectral information about hundreds of thousands of stars. This data offers a prime opportunity for abundance analysis. However, even with data reduction pipelines in place, interpreting that volume of data to find meaningful results is challenging.

To this end, we will employ an efficient method for investigating abundance space (Bovy 2015). The method, described in §3, uses spectra directly, rather than quantities derived by fitting multi-parameter model spectra. This reduces complexity in calculations of uncertainty, and incorporates an understanding of the systematic effects of other stellar properties. This will allow for a more accurate investigation the dimensionality of abundance space for a sample of stars, including probing the relative significance of particular elements and searching out abundance-position trends.

2. Data Set

This work will focus on analyzing the spectra of a sample of red clump stars from APOGEE’s Data Release 12. These stars are selected as belonging to the red clump according to cuts in gravity, effective temperature, and metallicity, and leave us with an initial sample of 19936 stars (Bovy et al. 2015). These stars trace, to some extent, the stellar population of the Milky Way through the volume of the APOGEE survey. A star’s time in the red clump is short compared to its lifetime, so the sample is biased towards locating younger stars. Data on these stars comes in the form of infrared (H-band) spectra covering a wavelength range from $1.514\ \mu\text{m}$ to $1.696\ \mu\text{m}$ (with some small nm gaps between detectors)(Majewski et al. 2015). APOGEE’s DR12 pipeline provides data on abundances of 15 elements (C, N, O, Na, Mg, Al, Si, S, K, Ca, Ti, V, Mn, Fe, Ni) as well as effective temperature (T_{eff}), surface gravity ($\log(g)$), metallicity (Z), and alpha-element enhancements ($[\alpha/\text{Fe}]$)(Holtzman et al. 2015). We will perform our analysis on each element for which an abundance is given.

3. Methods

Starting with our sample of APOGEE spectra, we will perform fits to remove non-abundance stellar properties. Flux values from each spectrum at a given pixel p ($F(p)$) are fit with a low order polynomial in T_{eff} , $\log(g)$ and Z , so some function for flux is found: $f(p, T_{\text{eff}}, \log(g), Z)$ (see Figure 1 for an example fit). At pixels (p_X) known to host absorption features for an element X, residuals are calculated from the fit $\delta = F(p_X) - f(p_X, T_{\text{eff}}, \log(g), Z)$. Residuals for element X pixels are weighted by normalized window functions which trace each X absorption feature and are then summed for each spectrum. This gives some scatter in element X for each star, since other systematic effects have ostensibly removed by the fit. It remains to characterize what portion of this scatter is due to measurement noise or intrinsic effects particular to the element. The measurement noise can be determined from uncertainties in $F(p)$, and intrinsic element scatter can be estimated from a sample of open cluster stars (Mészáros et al. 2015), which should all have similar abundance properties. The remaining scatter for each star after these effects have been accounted for will represent some variation in the 15 dimensional abundance space (see Figure 2 for a histogram showing the spread in element residuals compared with measurement uncertainty). The importance of these variations will be investigated through principal component analysis (PCA), finding the linear combinations of element residuals with the most significant variance.

4. Timeline

- Nov 2015, Q: Can we see scatter in residuals of individual elements above our level of intrinsic and measurement uncertainty?

- Characterize measurement scatter by drawing from a Gaussian determined by the uncertainty in flux measurements.
- Fit open clusters (from Mészáros et al. (2015)) in T_{eff} according to the outline in §3 and take residuals from this fit as representative of intrinsic abundance scatter.
- Dec 2015, Q: What is the dimensionality of residual space for all elements? Which elements are most significant for distinguishing stars?
 - Use principal component analysis to identify elements of primary importance
- Jan/Feb 2016, Q: Is there any positional dependence of mono-abundance populations (MAPs) identified in the remaining dimensions of residual space?
 - Collect stars from small bins in the multidimensional residual space (i.e. stars in the same MAP) and plot their radial and vertical positions.
 - Determine if stars belonging to different MAPs inhabit different sections of the galaxy, as well as positional variations for stars within a MAP.
- Mar/Apr 2016, Q: Can we understand the reasons certain elements may be more significant after PCA by modeling spectra?
 - Create model spectra by varying abundances of the 15 elements.
 - Perform fitting described in §3 and compare with fitting done on APOGEE data. Ascertain whether model spectra created with derived abundances for the red clump stars can reproduce residual results from the data set.

REFERENCES

- Bovy, J. 2015, ArXiv e-prints, arXiv:1510.06745
- Bovy, J., Rix, H.-W., Schlafly, E. F., et al. 2015, ArXiv e-prints, arXiv:1509.05796
- Holtzman, J. A., Shetrone, M., Johnson, J. A., Zamora, O., & Zasowski, G. 2015, AJ, 150, 148
- Jordan, S. 2008, Astronomische Nachrichten, 329, 875
- Majewski, S. R., Schiavon, R. P., Frinchaboy, P. M., et al. 2015, ArXiv e-prints, arXiv:1509.05420
- Mészáros, S., Martell, S. L., Shetrone, M., et al. 2015, AJ, 149, 153

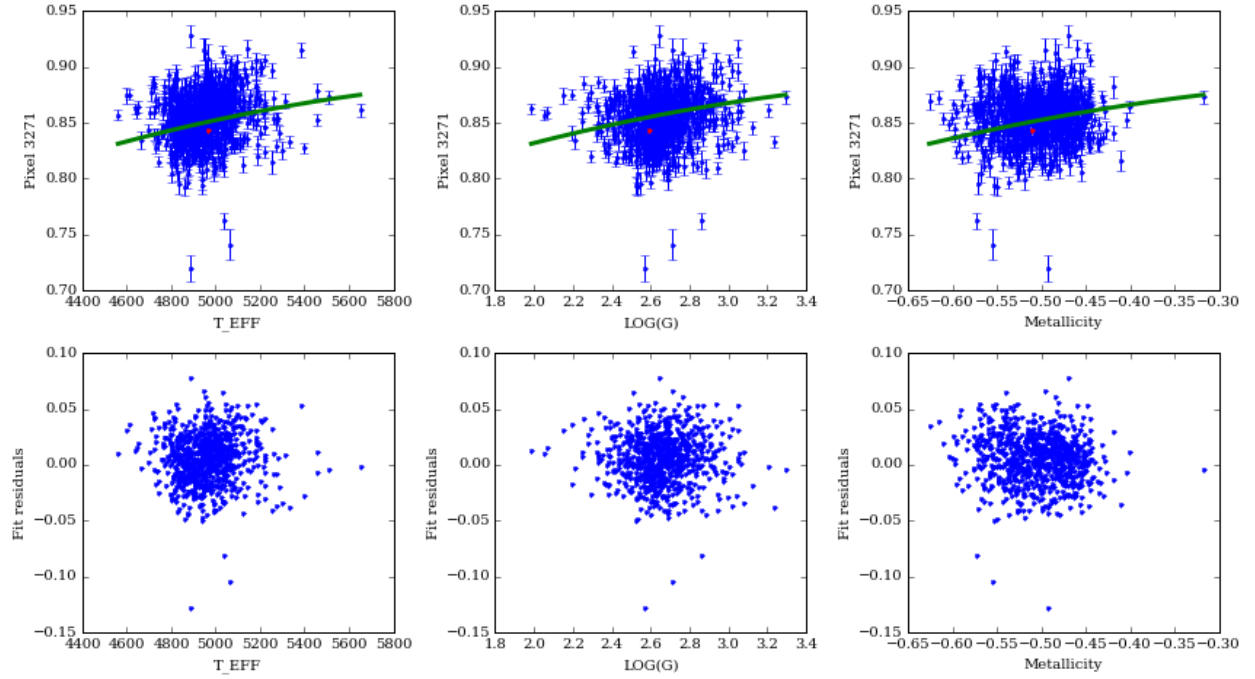


Fig. 1.— Sample second order fit to flux values at pixel 3271 for each star in the sample. Blue dots are the flux values, the green line is the simultaneous fit second order fit in three parameters: T_{eff} , $\log(g)$ and Z . Red dots indicate points with large uncertainty (>0.1 in the normalized flux units). These points were plotted without errorbars but are still included in the fit. Each upper panel is the fit plotted against one of its dimensions, the lower panels are the residuals from these results, assumed to be due to variations in abundances. This pixel corresponds to an absorption feature for Mg.

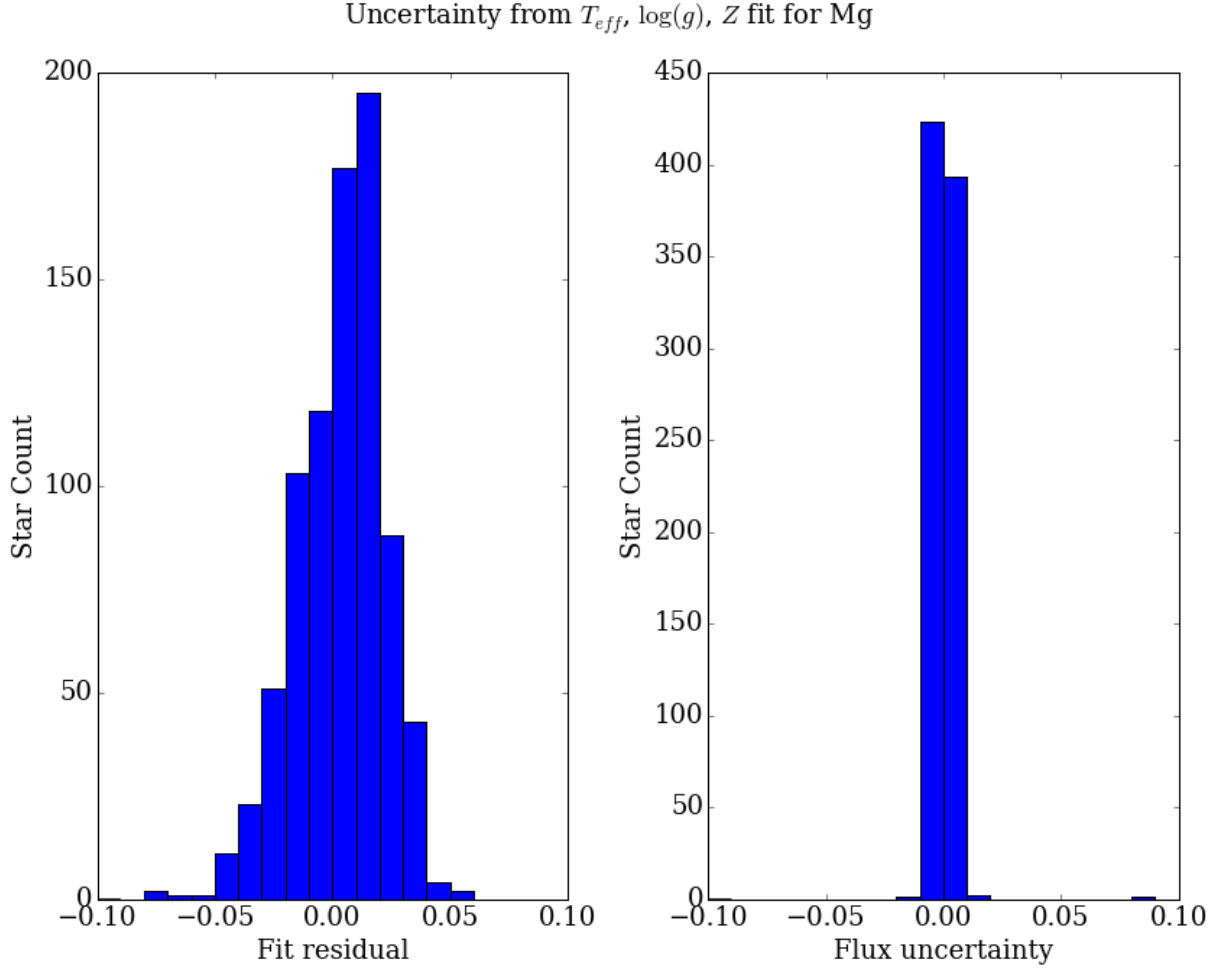


Fig. 2.— The left panel shows a histogram of the weighted residuals for several pixels associated with Mg absorption from each stellar spectrum. The right panel is a histogram of the fractional uncertainty in the flux value. This is calculated by drawing randomly from a normal distribution with a standard deviation of the measurement uncertainty at each relevant pixel and dividing it by the flux value at that pixel. These fractional results are then weighted and summed in the same way as the residual values to produce the histogram in the right panel.