Leibniz-Institut für Astrophysik Potsdam

Contact Email: `fanders@aip.de`

Automatic Classification of APOGEE giant stars with ASPCAP?

Friedrich Anders, Cristina Chiappini, Thaíse S. Rodrigues (INAF Padova)

Motivation

Information on the evolutionary stage of a star is crucial for the inference of stellar distances, extinctions, masses and ages, especially when similar stellar parameter regions are occupied by different populations (Fig. 1). Asteroseismically determined evolutionary states from *Kepler* [1] are also valuable for calibration of the APOGEE Stellar Parameters and Chemical Abundances Pipeline (ASPCAP, [2,3]).

Recently, Bovy et al. [4] have used stellar evolution models to perform cuts in colour–stellar parameter space in order to obtain a low-contamination APOGEE red clump (RC) sample with distance uncertainties of $\lesssim$ 5%. Their approach was tailored to minimize the sample contamination by red giant branch (RGB) stars; the trade-off is a non-trivial selection function (see Fig. 2). Instead, we would like to use the whole APOGEE red giant sample and get best possible distances for all stars.

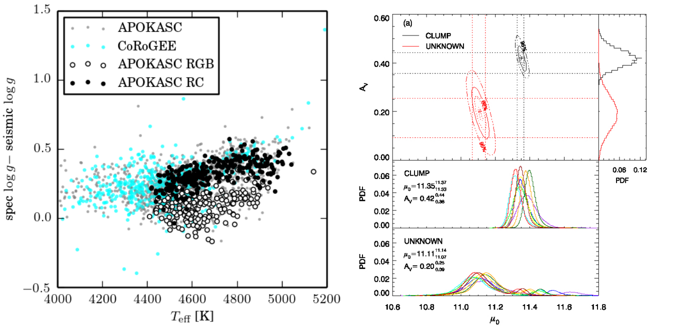

Figure 1: Two examples for the usefulness of discrimination between red giant branch and red clump stars. Left: The calibration relations for ASPCAP surface gravity are different for the two populations. Right: Knowledge of the evolutionary state of a star yields much more precise distance, age and extinction estimates. Figure from Rodrigues et al. [5]

Analysis

Using 505 stars with known evolutionary state from the joint APOGEE-Kepler (APOKASC) dataset [6], we test the possibility of using only the output parameters of ASPCAP to determine the evolutionary stage (1st ascent Red Giant Branch or He-burning Red Clump star) of giant stars in the APOGEE survey (see Fig. 2).

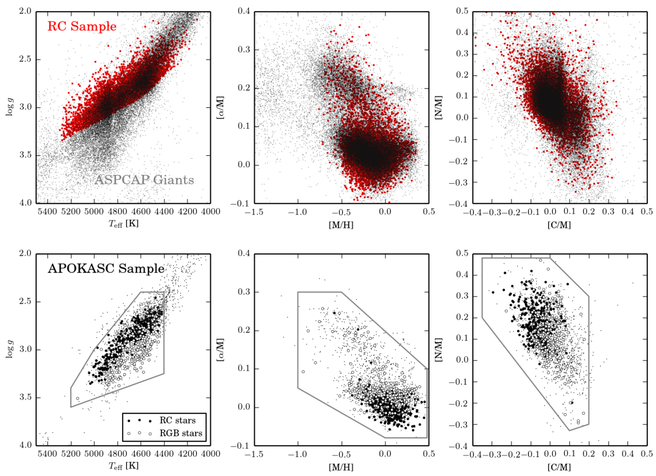

Figure 2: Three slices through the ASPCAP parameter space. From left to right: The ASPCAP $T_{\rm eff}$ — $\log g$ diagram, the "chemical plane" diagram [M/H] vs. [α/M], and the [N/M] vs. [C/M] diagram for the ASPCAP allStar sample and the APOGEE Red Clump Catalog (upper panels) and for the APOKASC sample (lower panels). The polygons in the lower panels mark the parameter space covered by the APOKASC training dataset.

Results

Using different simple classification methods available (e.g., Ivezić et al. 2013 [7]), we show that – at least within the parameter space covered by the APOKASC test dataset – it may be possible to assign meaningful evolutionary stage information by only using spectroscopic information.

For the APOKASC test dataset, both the Bayesian Gaussian Mixture Model (GMM) and the Quadratic Disriminant Analysis (QDA) classifiers attain a completeness (= $\frac{\#\,\text{True Positives}}{\#\,\text{True Positives} + \#\,\text{False Negatives}}$) of $>$ 95% and a contamination (= $\frac{\#\,\text{False Positives}}{\#\,\text{True Positives} + \#\,\text{False Positives}}$) of $\lesssim$ 5%, only using the ASPCAP parameters $T_{\rm eff}$, $\log g$ and [M/H] (Fig. 3). Adding other spectroscopic or photometric parameters does not significantly improve the discrimiation.

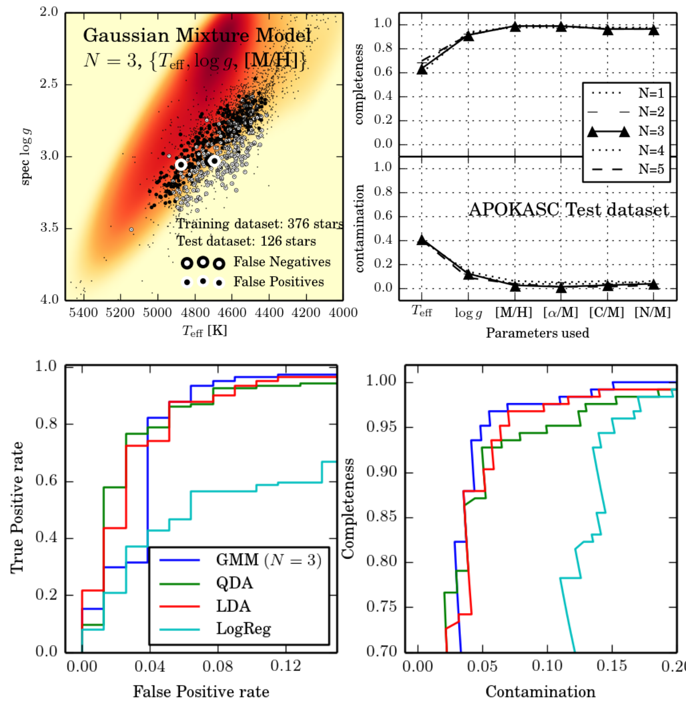

Figure 3: Performance of different classification methods for our dataset. Top left: Illustration of the Bayesian Gaussian Mixture Model (GMM) classifier for $N = 3$ Gaussians in $\{T_{\rm eff}, \log g, [{\rm M/H}]\}$ space. The color code corresponds to the marginalized RC probability density. Top right: Completeness and contamination achieved by the GMM classifier for different numbers of Gaussians and spectroscopic parameters used. Lower panels: ROC curves (left) and completeness-contamination curves (right) for the 6-parameter ASPCAP data using the most efficient classifiers for our dataset [8].

Caveats

As the RC and the RGB populations trace different masses and ages, it is difficult to disentangle pipeline bugs from "real" population effects. It is therefore necessary to test the classification methods also in other fields for which asteroseismically determined evolutionary stages are available (i.e., the CoRoT fields LRa01 and LRc01).

Ideally, we would also like to cover the full parameter space of the horizontal branch, in order to find a robust classifier for the whole APOGEE giant sample.

References

[1] Stello, D., Huber, D., Bedding, T. R., et al. 2013, ApJL 765, L41; [2] Mészáros, S., Holtzman, J., García Pérez, A. E. et al. 2013, AJ 146, 133; [3] García Pérez, A. E., et al. 2014, in prep.; [4] Bovy, J., Nidever, D. L., et al. 2014, ApJ, in press; [5] Rodrigues, T. S., Girardi, L., et al. 2014, MNRAS, submitted; [6] Pinsonneault, M., Epstein, C., et al. 2014, ApJS, submitted; [7] Ivezić, Ž., Connolly, A., VanderPlas, J. and Gray, A., *Statistics, Data Mining, and Machine Learning in Astronomy*, Princeton University Press, 2013; [8] We use the python packages `astroML` and `sklearn`.