



## SPECTROSCOPIC DETERMINATION OF MASSES (AND IMPLIED AGES) FOR RED GIANTS

M. NESS<sup>1</sup>, DAVID W. HOGG<sup>1,2,3</sup>, H.-W. RIX<sup>1</sup>, M. MARTIG<sup>1</sup>, MARC H. PINSONNEAULT<sup>4,5</sup>, AND A. Y. Q. HO<sup>1,6</sup><sup>1</sup>Max-Planck-Institut für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany; [ness@mpia.de](mailto:ness@mpia.de)<sup>2</sup>Center for Cosmology and Particle Physics, Department of Physics, New York University, 4 Washington Pl., Room 424, New York, NY 10003, USA<sup>3</sup>Center for Data Science, New York University, 726 Broadway, 7th Floor, New York, NY 10003, USA<sup>4</sup>Department of Astronomy, The Ohio State University, Columbus, OH 43210, USA<sup>5</sup>Center for Cosmology and Astroparticle Physics, The Ohio State University, Columbus, OH 43210, USA<sup>6</sup>Department of Astronomy, California Institute of Technology, 1200 East California Blvd, Pasadena, CA 91125, USA

Received 2015 October 13; accepted 2016 February 3; published 2016 May 27

## ABSTRACT

The mass of a star is arguably its most fundamental parameter. For red giant stars, tracers luminous enough to be observed across the Galaxy, mass implies a stellar evolution age. It has proven to be extremely difficult to infer ages and masses directly from red giant spectra using existing methods. From the *Kepler* and APOGEE surveys, samples of several thousand stars exist with high-quality spectra and asteroseismic masses. Here we show that from these data we can build a data-driven spectral model using *The Cannon*, which can determine stellar masses to  $\sim 0.07$  dex from APOGEE DR12 spectra of red giants; these imply age estimates accurate to  $\sim 0.2$  dex (40%). We show that *The Cannon* constrains these ages foremost from spectral regions with CN absorption lines, elements whose surface abundances reflect mass-dependent dredge-up. We deliver an unprecedented catalog of 70,000 giants (including 20,000 red clump stars) with mass and age estimates, spanning the entire disk (from the Galactic center to  $R \sim 20$  kpc). We show that the age information in the spectra is not simply a corollary of the birth-material abundances  $[\text{Fe}/\text{H}]$  and  $[\alpha/\text{Fe}]$ , and that, even within a monoabundance population of stars, there are age variations that vary sensibly with Galactic position. Such stellar age constraints across the Milky Way open up new avenues in Galactic archeology.

**Key words:** Galaxy: stellar content – methods: data analysis – methods: statistical – stars: evolution – stars: fundamental parameters – techniques: spectroscopic

**Supporting material:** machine-readable table

## 1. INTRODUCTION

Age dating of stars is fundamental to understanding and reconstructing the formation and evolution history of the Milky Way. Independent measurements for both the elemental abundances and the ages of an extensive set of stars across the Milky Way would be a powerful constraint on galaxy and also on chemical evolution (presuming the chemical information is derived from material from which the stars have formed). Yet, as almost all stars are in equilibrium<sup>7</sup>, age is not a quantity that can be directly measured. Instead, one must rely on measuring instantaneous stellar properties (or “labels”) that correlate with age in a physically understood way, or one that can be calibrated (see Soderblom 2010, for an excellent review). Inevitably, stellar age estimates involve some form of stellar-evolution models, both for stars in clusters and for single field stars.

For the most part, age estimates from spectroscopic surveys have been determined for stars before or just after their main-sequence turnoff. In that regime, stellar evolutionary isochrones are well separated (at a given metallicity), and for well-measured  $T_{\text{eff}}$ ,  $\log g$ , and  $[\text{Fe}/\text{H}]$ , ages follow from isochrone matching. Such stellar parameters are typically derived from high-resolution spectroscopy, which delivers low associated errors on the parameters (e.g., Casagrande et al. 2011; Bensby et al. 2013; Haywood et al. 2013; Bergemann et al. 2014). To date, the largest homogeneous data set of stellar ages in the Galactic disk has been derived in this fashion from the Geneva

Copenhagen Survey (GCS). Yet, all 16,682 main-sequence stars from GCS are located in the immediate solar neighborhood of  $< 0.1$  kpc (Nordström et al. 2004). Recent analogous analyses (e.g., Haywood et al. 2013; Bergemann et al. 2014) have pushed to greater distances but still remain limited to essentially the solar radius.

To map stellar ages throughout the Milky Way, one needs more luminous stars in evolutionary phases that are prevalent across most ages and metallicities. Giant stars satisfy these criteria. They have the advantage that their luminosities and colors vary relatively little with age, which makes age biases in flux-limited samples weaker. Yet, this also means that giant-star isochrones of different ages nearly overlap, making it all but impossible to get precise ages from  $T_{\text{eff}}$ ,  $\log g$ , and  $[\text{Fe}/\text{H}]$  measurements, unless we have tiny errors in these measurements and enormous confidence in the accuracy of stellar isochrones. For reference, consider a typical solar-abundance red giant at  $\log g = 2$  with an age of 5 Gyr. For the PARSEC isochrones (Bressan et al. 2012), age differences of  $\pm 2$  Gyr correspond to changes in  $T_{\text{eff}}$  at fixed  $\log g$  of only  $\approx 10$  K, compared to shifts of  $\approx 50$  K for a 0.10 dex difference in  $[\text{Fe}/\text{H}]$ . Furthermore, core helium burning stars that have experienced significant prior mass loss, red clump stars, are located close in the H-R diagram to less-evolved, first-ascent red giant branch stars. Even if the observational data are exact, absolute comparisons to stellar isochrones are uncertain; the absolute  $T_{\text{eff}}$  of theoretical models, for example, is highly sensitive to the assumed efficiency of convection, typically parameterized with a mixing length. However, for basically all post-main-sequence stars, in particular stars on the red giant branch or in the red clump, the stellar mass should be a

<sup>7</sup> Most stars that are not in equilibrium undergo periodic variations, such as pulsations, but without discernible secular evolution.

powerful constraint on the stars’ age (see, e.g., Martig et al. 2014). In that case, the challenge is reduced to estimating stellar masses for extensive samples of giant stars throughout the Galaxy; these masses then imply ages.

In recent years, asteroseismology surveys such as MOST (Guenther et al. 2005), *Corot* (De Ridder et al. 2009), and *Kepler* (Bedding et al. 2010) have been extremely successful in producing information about stellar interiors and hence masses, in particular for giant stars (e.g., Chaplin & Miglio 2013; Lagarde et al. 2015; Silva Aguirre et al. 2015; Casagrande et al. 2016). These missions operate by taking high-cadence, high-precision stellar photometry over long, uninterrupted time intervals, in which stellar oscillation modes are visible in the Fourier domain. These modes are related to the density and mass of the stars. At present, all these asteroseismological surveys cover only a few directions in the sky and hence a small portion of the Galaxy.

At the same time, there are a number of large spectroscopic surveys, such as APOGEE (Majewski 2015), *Gaia*-ESO (Gilmore et al. 2012), and GALAH (Freeman 2012; De Silva et al. 2015). These surveys are producing high signal-to-noise, high-resolution spectra of hundreds of thousands of stars across the entire sky. These allow measurements of properties including  $T_{\text{eff}}$ ,  $\log g$ ,  $[\text{Fe}/\text{H}]$ , and  $[\text{X}/\text{Fe}]$  for many elements. A star’s surface abundances, in particular  $[\text{Fe}/\text{H}]$  and  $[\alpha/\text{Fe}]$ , hold clues to its age because the plausibility of stars forming from material of a given abundance dramatically varies with time and radius throughout the galaxy: for example, stars were far more likely to have formed from metal-poor but  $\alpha$ -enhanced ISM than they are now. But such age constraints arise from the properties of the birth material, not from the current properties of the star itself, and hence age estimates and chemical evolution of the interstellar medium are inevitably degenerate (see, e.g., Chiappini 2002; Schönrich & Binney 2009).

The question then naturally arises how one can combine the information from these two types of surveys: information about the stellar interior and masses from seismology, and stellar parameters and element abundances from spectroscopy. For stars that have been observed by both kinds of surveys, this can be done at the catalog level (Martig et al. 2014).

As stars evolve to the red giant branch, they develop deep surface-convection zones and dredge up nuclear-processed material in their interiors in a mass-dependent fashion (Iben 1967). Elements whose surface abundance is particularly sensitive to this phenomenon include the light elements Li, Be, B, C (in particular the ratio of C12/C13), and N; C and N are measured by surveys such as APOGEE. There is also evidence, especially in metal-poor stars, for changes in CNO abundances along the giant branch (e.g., Kraft 1994; Yong et al. 2015), which requires a mixing process not usually included in standard models whose origin is still uncertain (Angelou et al. 2012). However, *empirical* correlations between stars of known mass and surface abundances can yield powerful insights even without detailed knowledge of the underlying physics.

Following this approach, Martig et al. (2015) have used red giants in the APOKASC sample of stars, for which seismic parameters are known from *Kepler* (Pinsonneault et al. 2014), and  $T_{\text{eff}}$ ,  $\log g$ ,  $[\text{Fe}/\text{H}]$ ,  $[\alpha/\text{Fe}]$ , and  $[\text{X}/\text{Fe}]$  are measured from APOGEE spectra (Ahn et al. 2014). Martig et al. (2015) have found a tight correlation between the masses determined from the standard seismic scaling relations and the  $[\text{C}/\text{N}]$

measurement from APOGEE. They determine a model for stellar mass and age as a function of C and N abundance measurements.

In this paper, we set out to develop a data-driven and far-reaching connection between the asteroseismic and the spectroscopic results for giant stars, with the ultimate goal of determining stellar masses of giants, and hence ages, directly from spectra. The goal is to derive age estimates that do not simply reflect the abundances of the star’s birth material (such as joint  $[\text{Fe}/\text{H}]$  and  $[\alpha/\text{Fe}]$  estimates); we aim for age estimates that give meaningful results, even at a given  $[\text{Fe}/\text{H}]$  and  $[\alpha/\text{Fe}]$ . We are, however, with spectroscopic data determining only the surface property of stars. The physical properties of mass and derived age can only be inferred, given theoretical expectations from stellar models between these physical quantities and stellar spectra.

For this purpose, we use a set of 1639 APOKASC stars from the APOGEE DR12 spectral sample with stellar mass and  $\log g$  measurements from asteroseismology (Pinsonneault et al. 2014), along with DR12’s  $T_{\text{eff}}$ ,  $[\text{Fe}/\text{H}]$ , and  $[\alpha/\text{Fe}]$  (see Figure 13 in the Appendix). Using *The Cannon* (Ness et al. 2015a), we then generate a data-driven generative model for the five stellar labels  $T_{\text{eff}}$ ,  $\log g$ ,  $[\text{Fe}/\text{H}]$ ,  $[\alpha/\text{Fe}]$ , and mass. This model quantifies the information content at each pixel. Therefore, we can examine the origin of the information on these labels directly in the spectra. We have shown previously, in Ness et al. (2015) and A. Y. Q. Ho et al. (2016, in preparation), that *The Cannon* is successful in delivering the labels of  $T_{\text{eff}}$ ,  $\log g$ ,  $[\text{Fe}/\text{H}]$ , and  $[\alpha/\text{Fe}]$  for APOGEE stars. With a training set of stars with known masses, we can expand the same approach to include a fifth stellar label, mass. With these data, *The Cannon* is a direct framework to characterize the relationship between the surface spectroscopy and interior asteroseismology, in order to jointly infer stellar properties, learning about stellar interiors from surface spectroscopy.

In Section 2 we describe the implementation and application of *The Cannon* for the case at hand. We also lay out the verification of the mass label estimates, and we illustrate where in the spectrum the information on the various labels originates.

We deliver our catalog of mass and inferred ages for the  $\approx 20,000$  red clump stars in the APOGEE survey, which have well-known distances, as well as for the 50,000 red giant stars in the DR12 APOGEE data release that are within the label (stellar parameter) range of our training set.

## 2. METHODS AND DATA

### 2.1. Implementation of *The Cannon* to Include Mass Labels

We make use of *The Cannon* (Ness et al. 2015a), which is a data-driven method for determining stellar parameters and abundances. *The Cannon* is a probabilistic model of stellar spectra—meaning that it produces a likelihood function or a probability density in spectral space—that is itself a function of stellar parameters and chemical abundances (which we collectively call “labels”). The model is not based on physical models, but is instead learned from a training set of stars with (assumed) known labels. This learning is called the “training step.” The model is used to label a new star *not* in the training set by maximizing the likelihood of the label values given the new star’s spectrum. This labeling is called the “test step.” *The Cannon* differs from standard machine-learning methods (such as random forest or deep neural networks) in that it

contains an explicit likelihood function, at both the training step and the test step, so it is able to account for heteroskedastic noise and missing data in the spectra of both the training and test stars.

Generally, we take a spectral model to be characterized by a coefficient vector  $\theta_\lambda$  that predicts the flux at every pixel  $f_{n\lambda}$  for a given label vector  $\ell_n$ :

$$f_{n\lambda} = g(\ell_n|\theta_\lambda) + \text{noise}. \quad (1)$$

In detail, the likelihood function we use for *The Cannon* has a Gaussian form at each measured spectral wavelength, with a mean that is a quadratic function of the labels and a variance that consists of an intrinsic variance added to an observational noise variance (from photon noise and other sources). For our model we use the quadratic-in-labels form of Ness et al. (2015a). This model presumes that the continuum-normalized flux is a polynomial of the stellar labels, written as  $f_{n\lambda} = \theta_\lambda^T \cdot \ell_n + \text{noise}$ , but where  $\theta_\lambda$  now contains 21 elements at every pixel. For the case of the five labels ( $T_{\text{eff}}$ ,  $\log g$ ,  $[\text{Fe}/\text{H}]$ ,  $[\alpha/\text{Fe}]$ ,  $\mathcal{M}$ ), the label vector  $\ell_n$  becomes

$$\begin{aligned} \ell_n \equiv & [1, T_{\text{eff}}, \log g, [\text{Fe}/\text{H}], [\alpha/\text{Fe}], \mathcal{M}, T_{\text{eff}}^2, \\ & T_{\text{eff}} \cdot \log g, T_{\text{eff}} \cdot [\text{Fe}/\text{H}], T_{\text{eff}} \cdot [\alpha/\text{Fe}], T_{\text{eff}} \cdot \mathcal{M}, \log g^2, \\ & \log g \cdot [\text{Fe}/\text{H}], \log g \cdot [\alpha/\text{Fe}], \log g \cdot \mathcal{M}, [\text{Fe}/\text{H}]^2, \\ & [\text{Fe}/\text{H}] \cdot [\alpha/\text{Fe}], [\text{Fe}/\text{H}] \cdot \mathcal{M}, \\ & [\alpha/\text{Fe}]^2, [\alpha/\text{Fe}] \cdot \mathcal{M}, \mathcal{M}^2]. \end{aligned} \quad (2)$$

## 2.2. Data

We have shown in previous work (Ness et al. 2015a) that *The Cannon* does a good job of modeling stellar spectra and delivering stellar parameters and chemical abundances for stars with spectra taken by the APOGEE project (Majewski 2015). APOGEE is a Sloan Digital Sky Survey (SDSS)<sup>8</sup> (Eisenstein et al. 2011) infrared survey of the Milky Way disk, bulge, and halo and has provided H-band spectra (1500–1700 nm) of about 150,000 stars in the public data release DR12. The three labels of  $T_{\text{eff}}$ ,  $\log g$ , and  $[\text{Fe}/\text{H}]$  delivered with *The Cannon* were demonstrated in Ness et al. (2015a). In this work we train on and then determine two additional labels:  $[\alpha/\text{Fe}]$  and mass. We train on  $\log$  mass and infer the subsequent age using stellar-evolution models as described in Section 2.3. Our five labels are provided to *The Cannon* in the training step and delivered by *The Cannon* in the test step.

### 2.2.1. Training Data

The training set is composed of 1639 stars taken from the *Kepler* field, the so-called APOKASC sample (Pinsonneault et al. 2014) of stars observed by APOGEE. This sample of stars has high-quality infrared spectra from APOGEE and also asteroseismological measurements from the *Kepler* mission. The *Kepler* mission (Borucki et al. 2010) took continuous, 30 minute cadence (or higher cadence) photometric observations of more than  $10^5$  stars, providing (at least for giant stars) measurements of the asteroseismological frequencies and frequency splittings that indicate stellar interior density structure. The two global asteroseismic parameters are the  $\nu_{\text{max}}$  and  $\Delta\nu$  quantities. These are the measurements from

*Kepler* that indicate the interior structure of the star (see Pinsonneault et al. 2014, and references therein). The asteroseismic measurements are used—with stellar models—to infer stellar masses and thus provide labels.

Our training set of 1639 APOKASC stars is described in Martig et al. (2014) and selected from the full APOKASC sample in Pinsonneault et al. (2014) based on additional quality cuts. This sample includes only stars with no warning or error in the ASPCAP FLAG parameter provided by APOGEE (Ahn et al. 2014), with no rotation flag set and with errors on the  $\Delta\nu$  and  $\nu_{\text{max}}$  less than 10%. The APOKASC stars comprise a high signal-to-noise ratio (S/N) sample, with an S/N > 80.

We work with the continuum-normalized DR12 spectra, and the method of continuum estimation turns out to be important for performance. We use the aspcapStar files provided by APOGEE, but apply our own signal-to-noise invariant continuum normalization by fitting a low-order polynomial to “true” continuum pixels, as described in Ness et al. (2015a). Five labels are used for training *The Cannon*:  $T_{\text{eff}}$ ,  $\log g$ ,  $[\text{Fe}/\text{H}]$ ,  $[\alpha/\text{Fe}]$ , and  $\log$  mass. The label range of the training data set is shown in the Appendix in Figure 13. The five training labels adopted are from the ASPCAP-corrected values (Mészáros et al. 2013) for the  $T_{\text{eff}}$ ,  $[\text{Fe}/\text{H}]$ , and  $[\alpha/\text{Fe}]$  and the asteroseismic value for  $\log g$ , as determined from the measured  $\nu_{\text{max}}$ . The mass label was determined from the  $\Delta\nu$  and  $\nu_{\text{max}}$  measurements using the standard seismic scaling relation (e.g., Kjeldsen & Bedding 1995), as in Equation (3):

$$M = \left( \frac{\nu_{\text{max}}}{\nu_{\text{max},\odot}} \right)^3 \left( \frac{\Delta\nu}{\Delta\nu_{\odot}} \right)^{-4} \left( \frac{T_{\text{eff}}}{T_{\text{eff},\odot}} \right)^{1.5}. \quad (3)$$

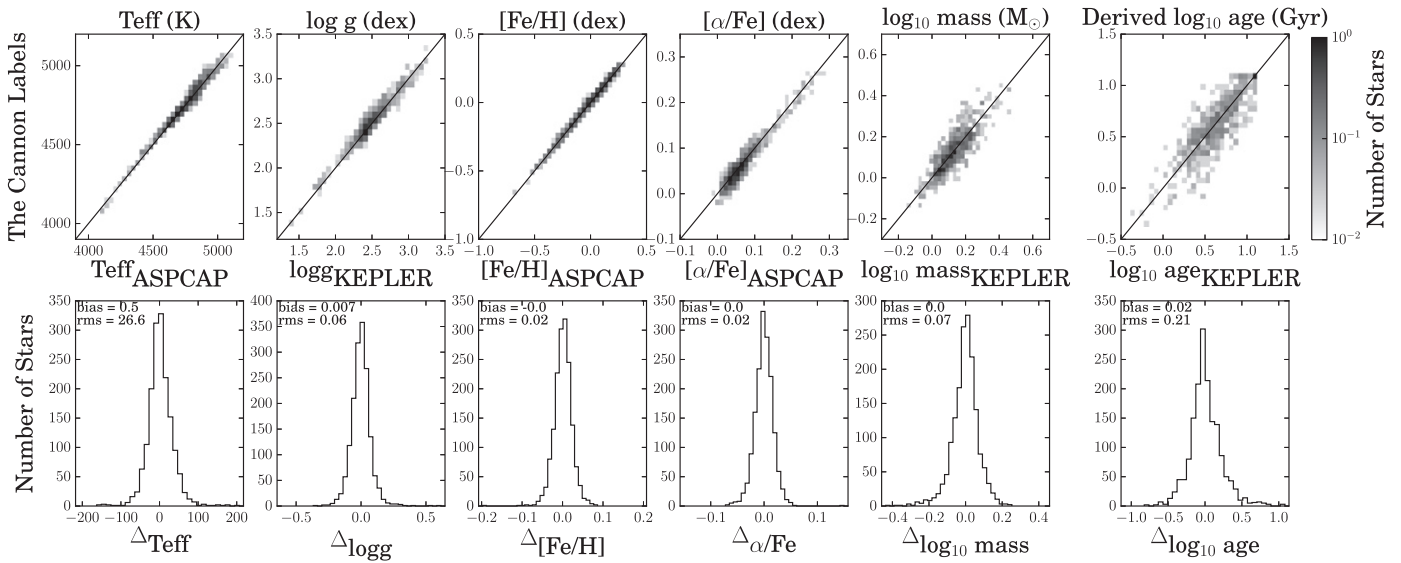
We adopt  $T_{\text{eff},\odot} = 5777$  K,  $\nu_{\text{max},\odot} = 3140$   $\mu\text{Hz}$ ,  $\Delta\nu_{\odot} = 135.03$   $\mu\text{Hz}$ , as per Martig et al. (2014). The solar values  $\Delta\nu_{\odot}$  and  $\nu_{\text{max},\odot}$  are those used for the APOKASC catalog and were obtained by Hekker et al. (2013).

Note that modified scaling relations can be adopted in order to determine mass from the asteroseismic parameters. *The Cannon* is a generalized method, and in all cases, the results at the test step will be directly tied to the assumptions in the training step. *The Cannon* is implemented here as described in Ness et al. (2015a) but using the model in Equation (2), with the mass label coming from the equation described in Equation (3).

The scaling relationships rely on a combination of theoretically motivated and empirical arguments. As such, their absolute values need to be calibrated by comparison with fundamental masses. Radii are in reasonable agreement with parallax (Silva Aguirre et al. 2012) and interferometry (Huber et al. 2012) measurements. However, there appear to be modest but real offsets between the expected and asteroseismic masses of open-cluster red giants (Brogaard et al. 2012) and somewhat larger ones for halo giants (Epstein et al. 2014). These differences may depend on evolutionary state (Miglio et al. 2012) but are otherwise systematic rather than random in nature. We therefore proceed with the masses as indicated by the unmodified relations, cautioning that there could be zero-point differences, metallicity-dependent stretches in the mass scale, and evolutionary-state-dependent changes. *The Cannon* determines the relationship between the stellar mass and metallicity, and the mass range is calibrated across the label space of the training data, which includes stars within the metallicity range  $-0.85 < [\text{Fe}/\text{H}] < 0.30$ . Despite these important caveats, we demonstrate that the relative masses

<sup>8</sup> [www.sdss.org](http://www.sdss.org)





**Figure 1.** Cross validation of the training data set of 1639 stars for the  $T_{\text{eff}}$ ,  $\log g$ ,  $[\text{Fe}/\text{H}]$ ,  $[\alpha/\text{Fe}]$ , and mass labels: the results for *The Cannon*'s labels for training performed on 90% of the APOKASC stars, showing the performance at test time on the 10% of the stars not included in training, run 10 times. The panel on the far right is the derived age label from the mass determined with *The Cannon*, using interpolation with PARSEC isochrones. The 31 stars with a  $\chi^2_{\text{reduced}}$  statistic of  $>2$  (2% of the training data) have been removed.

inferred from these scalings produce sensible inferences about galactic properties. It is straightforward to adopt corrected masses as new calibrations of the scaling relations arise.

### 2.2.2. Test Data

We train our model using the APOKASC sample and then determine our five stellar labels of  $T_{\text{eff}}$ ,  $\log g$ ,  $[\text{Fe}/\text{H}]$ ,  $[\alpha/\text{Fe}]$ , and  $\log$  mass for APOGEE's DR12 red clump catalog (Bovy et al. 2014) and all red giants in APOGEE's DR12 data release that are within the label range of our training data set. The test data are treated in the exact same way as the training data, as described in Ness et al. (2015a), where we work with continuum-normalized APOGEE aspcapStar files and apply our own additional continuum-normalization procedure.

### 2.3. From Masses to Ages

Our asteroseismic calibration set measures only *current* masses; a model is required to map these masses to their initial values. In addition, the mapping of stellar mass to age depends on the adopted input physics, for example, the treatment of convective core overshooting for massive stars as well as the detailed mixture of heavy elements and the assumed initial helium abundance (see Soderblom 2010 for a detailed discussion). For red giants, the importance of mass loss depends sensitively on the luminosity and whether or not the star is a first-ascent red giant or a red clump star. Using the general formulation of Reimers (1975), one would expect on dimensional grounds to have mass loss occur primarily on the upper red giant branch when the surface gravity is low; there could also be mass loss associated with the ignition of helium in a degenerate medium. Mass loss is therefore only likely to be important for red clump stars and for very luminous first-ascent giants; the latter are rare in our sample.

Globular cluster data require modest (of order 0.2 solar masses), integrated mass loss on the giant branch with a stochastic dispersion on the order of 0.03 solar masses (e.g., Lee et al. 1990). The mass loss for higher-metallicity red giants

is less well established, with some suggestion from *Kepler* data for a relatively weak mass loss (Miglio et al. 2012); note, however, that recent summaries of globular cluster data imply a larger scaling constant of order 0.48 (McDonald & Zijlstra 2015). We therefore adopt a modest mass-loss prescription ( $\eta_{\text{Reimers}} = 0.2$ ) to map current onto initial mass, with the caution that this may underestimate the effect for red clump stars. For a recent discussion of the age uncertainties for red giants with asteroseismic masses, see Casagrande et al. (2015).

For our purposes, we are interested primarily in differential ages and in checking whether or not the usage of asteroseismic masses results in plausible age properties, not in rigorous absolute age measurements. In the sections that follow, we explicitly distinguish between the ages of red giant and red clump stars to separate out the red clump sample in which ages depend on the assumptions concerning mass loss from the red giant sample that is relatively insensitive. (Also note that assuming a red clump evolutionary state, for example, instead of a red giant evolutionary state, for interpolating mass to age does not change the age distribution of the sample. Individual stellar age differences are on the order of 5% between these two evolutionary states.) Finally, we note that a real astrophysical sample will include both mergers of low-mass stars and stars that have had their envelopes stripped by a companion; care must be taken in population modeling to distinguish such astrophysical backgrounds from very young or old populations, respectively.

## 3. RESULTS

### 3.1. Validation of the Mass Determination

To determine the uncertainties from *The Cannon* on our individual  $T_{\text{eff}}$ ,  $\log g$ ,  $[\text{Fe}/\text{H}]$ ,  $[\alpha/\text{Fe}]$ , and mass measurements, we perform a take-stars-out test on the set of reference objects. For the take-stars-out test, we train the spectral model iteratively on 90% of the reference spectra and then run the test step on the remaining 10% of the spectra, and we do this 10 times, stepping through each next 10% of the data. Our results

are shown in Figure 1 for the five labels. The top panels show the cross-validation results comparing the input and output labels, and the bottom panels show the histograms of the  $\Delta$  (input–output) for each label. The training labels ( $x$  axis in the top panel of the figure) are from ASPCAP and asteroseismology, as described in Section 2, and the output labels ( $y$  axis in the top panel) are from *The Cannon*. The sixth panel in this figure shows the masses transformed to ages using interpolation between the PARSEC isochrones, where the red clump evolutionary state has been adopted on the isochrones at each age and [Fe/H]. We have removed spectra that could not be well fit by *The Cannon*'s model, where the reduced  $\chi^2_{\text{reduced}}$  of the model from model fit to the data is  $\chi^2_{\text{reduced}} > 2$ , which corresponds to 31 stars removed from the sample of 1639 stars.

This figure shows that *The Cannon*'s purely mathematical approach of label transfer estimates the stellar labels with accuracies of 31 K in  $T_{\text{eff}}$ , 0.07 dex in  $\log g$ , 0.02 in [Fe/H], 0.02 in  $[\alpha/\text{Fe}]$ , and 0.07 dex in  $\log$  mass, or 0.21 dex in the inferred  $\log$  age (Gyr) over the label range of the reference stars. Notably, the uncertainty on the mass (20%) is only slightly larger than the APOKASC catalog uncertainty of 12%. It is important to remember that the objects plotted are the left-out objects, and the spectra of these objects are completely detached from the training step, except that they have the same experimental setup and are drawn from a part of label space that is represented by the remaining reference objects.

### 3.2. The Cannon's Generative Model at the Test Step: The Red Clump Stars

In Figure 2, the spectrum of one of the red clump stars (not in the training set), which is representative of a typical red clump spectra in the APOGEE red clump catalog (discussed in Section 4.1), is shown along with the generative model from *The Cannon* at its stellar labels and the best-fitting model from ASPCAP. The data are shown in black, the synthesized model from *The Cannon* is shown in red, and the best-fit model from ASPCAP is shown in the gray dashed line. The wavelength regions shown in this figure are those for which the highest amplitude of the coefficients are located (Figures 3–6). This example red clump star has parameters of  $T_{\text{eff}} = 4843$  K,  $\log g = 2.5$  dex, [Fe/H] =  $-0.06$  dex,  $[\alpha/\text{Fe}] = 0.04$ , and  $\text{mass} = 1.0$  determined by *The Cannon*.

Figure 2 illustrates that the generated spectral model from *The Cannon* provides a very good fit to the survey spectra. In fact, the generative model from *The Cannon* is a better fit to the data than the best-fit synthetic model from ASPCAP. That the model from *The Cannon* is a good fit to the data demonstrates that the five labels that we use to train, as well as our polynomial model, are sufficient to very well describe the behavior of the flux of a typical red clump star, given the training set of reference stars from the APOKASC catalog. Note in this figure that one of the regions where ASPCAP performs most poorly is at the Brackett line (see Section 3.3.1), which is highly  $\log g$  sensitive (Figure 3). This may indicate a problem with the model stellar atmospheres or its associated oscillator strength for this feature (or the lines it is blended with).

### 3.3. Which Parts of the H-band Spectra Constrain the Labels?

*The Cannon* is a generative model that determines a coefficient at every pixel or wavelength. These coefficients describe how the flux depends on the stellar labels, given the

model (in this work, in Equation (2)). A near-zero coefficient for a given pixel indicates that the flux at that pixel is independent of the labels. Conversely, the largest values of the coefficients are where the spectra change most significantly with the label or labels. Here we examine the origin of the highest coefficients for the first-order, linear coefficients. We use the first-order coefficients to identify some key regions of the spectra that contain the most information with respect to the labels and to determine which elements and molecules in particular these coefficients correlate with.

Figures 3–6 show the first-order coefficients of our model described in Equation (2) over a narrow wavelength range ( $\approx 30$  Å), centered on where the first-order coefficients reach their largest amplitude. The zeroth-order coefficient is shown in the top panel, and the first-order linear terms  $\theta_l$ , where  $l = T_{\text{eff}}$ ,  $\log g$ , [Fe/H],  $[\alpha/\text{Fe}]$ , and mass, are shown in the middle panels. For a given set of labels, we can use *The Cannon*'s model to generate the spectra (using all coefficients). The generated spectra are shown in the bottom panel of each figure for a representative set of stellar labels. These spectra are made at three steps across each stellar label, for each respective first-order coefficient. This directly illustrates how the flux changes with each label in regions where the coefficient associated with that label is highest.

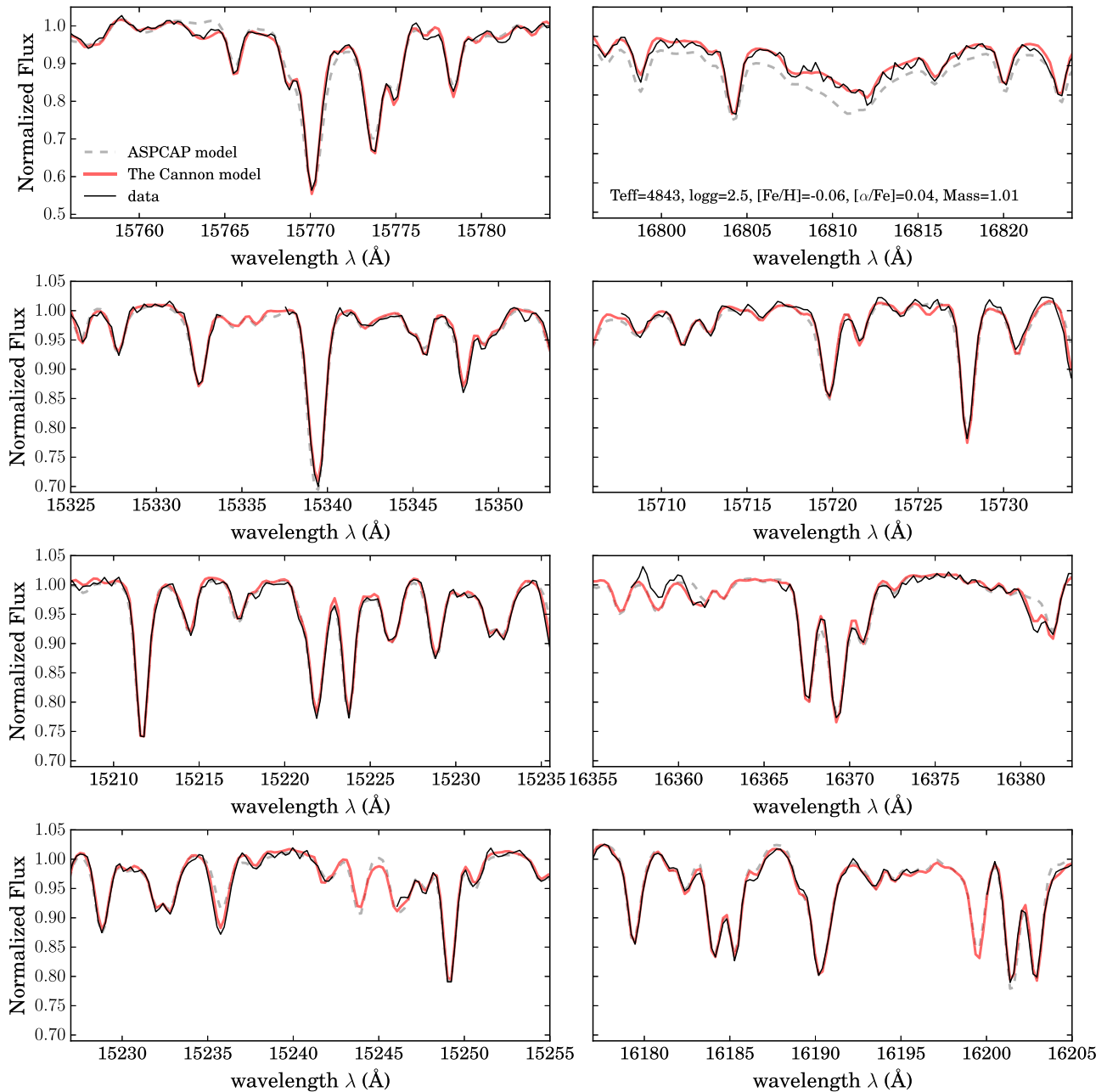
We use the DR12 APOGEE line list (Shetrone et al. 2015), Kurucz model atmospheres (Castelli & Kurucz 2004), and the stellar synthesis code MOOG (Snedden et al. 1979) to determine which elements correspond to the absorption features in the spectra where the highest first-order coefficients are located. The elements and molecules are marked on the zeroth-order coefficient spectra in the top panel of each of the figures. The zeroth-order coefficient vector  $\theta_0$ , the baseline spectrum of the model, is, essentially, the intersect spectrum of the training set of stars.

The absorption features in the H band are heavily blended with OH, CN, CO, and  $^{12}\text{C}$  molecules, and the figures indicate which absorption features are composed of blends of molecules and elements at the stellar-parameter space of APOGEE stars. The elements that show the most significant changes with the labels show a gratifying accord with the expectations from stellar physics, and these are discussed below for each of the five labels.

#### 3.3.1. Spectral Dependencies on $\log g$

Figure 3 shows two  $30$  Å regions of the spectra centered on the two highest first-order  $\log g$  coefficient amplitudes,  $\theta_l = \theta_{\log g}$ . The three panels at left show the highest  $\log g$  coefficient, and the three panels at right show the second-highest coefficient. Relevant elements and molecules that correspond to the absorption features are marked at the top on the baseline spectrum of the model. The middle panel of Figure 3 shows the first-order coefficients  $\theta_l$  that are linear in  $T_{\text{eff}}$ ,  $\log g$ , [Fe/H],  $[\alpha/\text{Fe}]$ , and mass. The coefficients have all been normalized to their largest absolute value, such that an amplitude of  $\theta_l = 1$  for any coefficient is at the highest value.

The bottom panel of Figure 3 shows a generated spectrum from *The Cannon*'s model for a reference set of stellar parameters, which are the mean of the training set labels or the fiducial spectra. These reference labels are set at  $T_{\text{eff}} = 4761$  K, [Fe/H] = 0.0,  $[\alpha/\text{Fe}] = 0.06$ , and  $\text{mass} = 0.3 M_{\text{star}}$  for three different  $\log g$  values of  $\log g = 1.5$ ,  $\log g = 2.1$ , and  $\log g = 3.3$ . From the center-left panel of Figure 3 it is



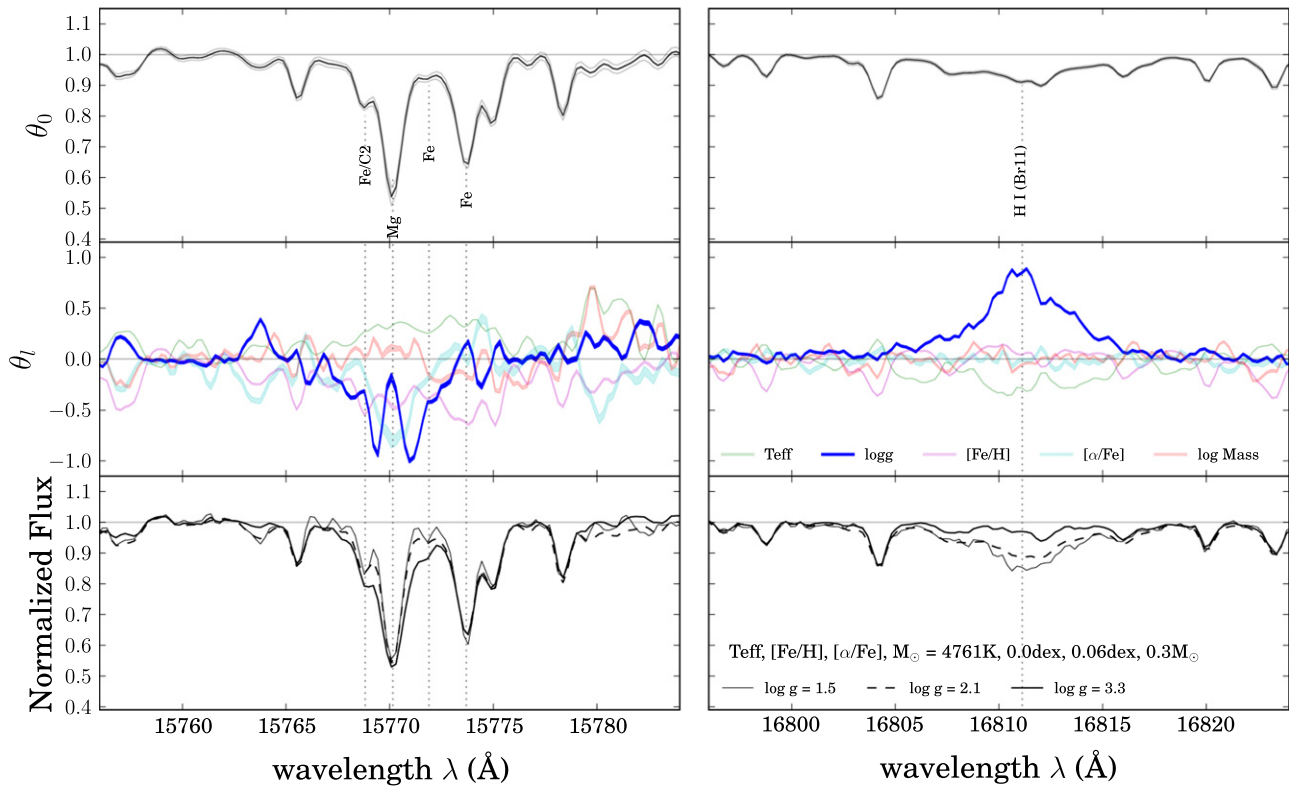
**Figure 2.** A spectrum of one of the red clump stars in the catalog of Bovy et al. (2014) shown in black, with the best-fit model from *The Cannon* in red. The eight wavelength regions shown correspond to the highest first-order coefficient amplitudes shown in Figures 2–5 (from top to bottom). The best-fit model from ASPCAP is shown in the gray dashed line.

clear that the flux at any given pixel can correlate with multiple labels. Typically, some coefficients, like  $T_{\text{eff}}$  and  $\log g$ , show an inverse relationship between label amplitude and flux.

The location of the highest-amplitude  $\log g$  linear coefficient, which is shown in the top left-hand panel of Figure 3, corresponds to a strong Mg feature in the APOGEE spectra. Importantly, the highest amplitude of this coefficient corresponds not to the core of the Mg feature but to the wings, and more strongly so for the upper-wavelength side of the feature. The core of the Mg feature in fact corresponds to a significantly lesser amplitude of the coefficient; clearly in the case of  $\log g$ , there is a dramatic reduction in the information content of these pixels in the core of the feature. Note that where the  $\log g$  coefficient decreases from the wings to the core (in blue), the

$[\alpha/\text{Fe}]$  label (in cyan) increases, so the largest amount of information in this region for the  $[\alpha/\text{Fe}]$  label is, conversely, from the core of this feature. This Mg feature at 15770.15 Å is one of the two strongest Mg features (along with the Mg feature at 15753.29 Å) across the APOGEE H-band spectral region.

That the strongest coefficient in  $\log g$  comes from the wings of a strong Mg line in the H-band APOGEE spectral region is well aligned with empirical analyses in other, more comprehensively studied wavelength regions. The wings of strong lines are known  $\log g$  indicators (Gray 2008). Specifically, the wings of Mg lines in the optical wavelength region, which are sensitive to pressure broadening, are used by Fuhrmann et al. (1997) to derive  $\log g$  for F and G main-sequence stars.



**Figure 3.** The zeroth- and first-order coefficients ( $\theta_0$ ,  $\theta_l$ ) of the model trained on the ASPCAP stars showing the two 30 Å wavelength regions where the  $\log g$  coefficient (in dark blue) reaches its highest absolute amplitudes. The zeroth coefficient (at top) describes the intersect spectrum, and a number of spectral absorption features are marked. The middle panel shows all coefficients, each normalized to their highest amplitude, for all  $l = T_{\text{eff}}$ ,  $\log g$ ,  $[\text{Fe}/\text{H}]$ ,  $[\alpha/\text{Fe}]$ , and mass coefficients, where all coefficients are shown in transparent lines, except for the  $\log g$ , which is of interest here. The bottom panels show the generated spectra from *The Cannon*'s model for three increasing values of  $\log g$ , which span the full range of  $\log g$  values in the training set. The selected parameters for this star are set around the fiducial point of the stars in the training set and represent a typical APOKASC spectra. The panel at left where the  $\log g$  coefficient reaches its highest amplitude shows the  $\log g$  information is concentrated in the wings of the strong Mg feature in the spectra, unlike the  $[\alpha/\text{Fe}]$  information, which is in the core, as seen from the cyan coefficient. The panels at right show the broadening of the flux as a consequence of the H Brackett-11 line which, like Mg, is similarly gravity sensitive due to pressure broadening effects (see the text). The labels of the fiducial spectra are indicated in the bottom right-hand panel.

Similarly, Brackett lines (as well as Balmer and Paschen lines) are sensitive to pressure (Stark) broadening and are therefore excellent tracers of  $\log g$  in stars. The second-highest amplitude coefficient for the first-order linear  $\log g$  coefficient in the APOGEE spectral region is at the Brackett feature at  $\approx 16810$  Å, as shown in the right-hand panels of Figure 3. The bottom panel of this figure (at right) shows how significantly the flux varies as a function of  $\log g$  for this feature. In addition to being second-highest in amplitude, the sign of this coefficient for this feature is positive and opposite that of the wings of the Mg line. As seen in the bottom panel at left, the wings of the Mg feature deepen with increasing  $\log g$ , whereas for the Brackett feature at right, the spectral profile flattens with increasing  $\log g$  for any given set of stellar  $T_{\text{eff}}$ ,  $[\text{Fe}/\text{H}]$ ,  $[\alpha/\text{Fe}]$ , and mass parameters, directly demonstrating the inverse relative relationship between the two features.

### 3.3.2. Spectral Dependencies on $T_{\text{eff}}$

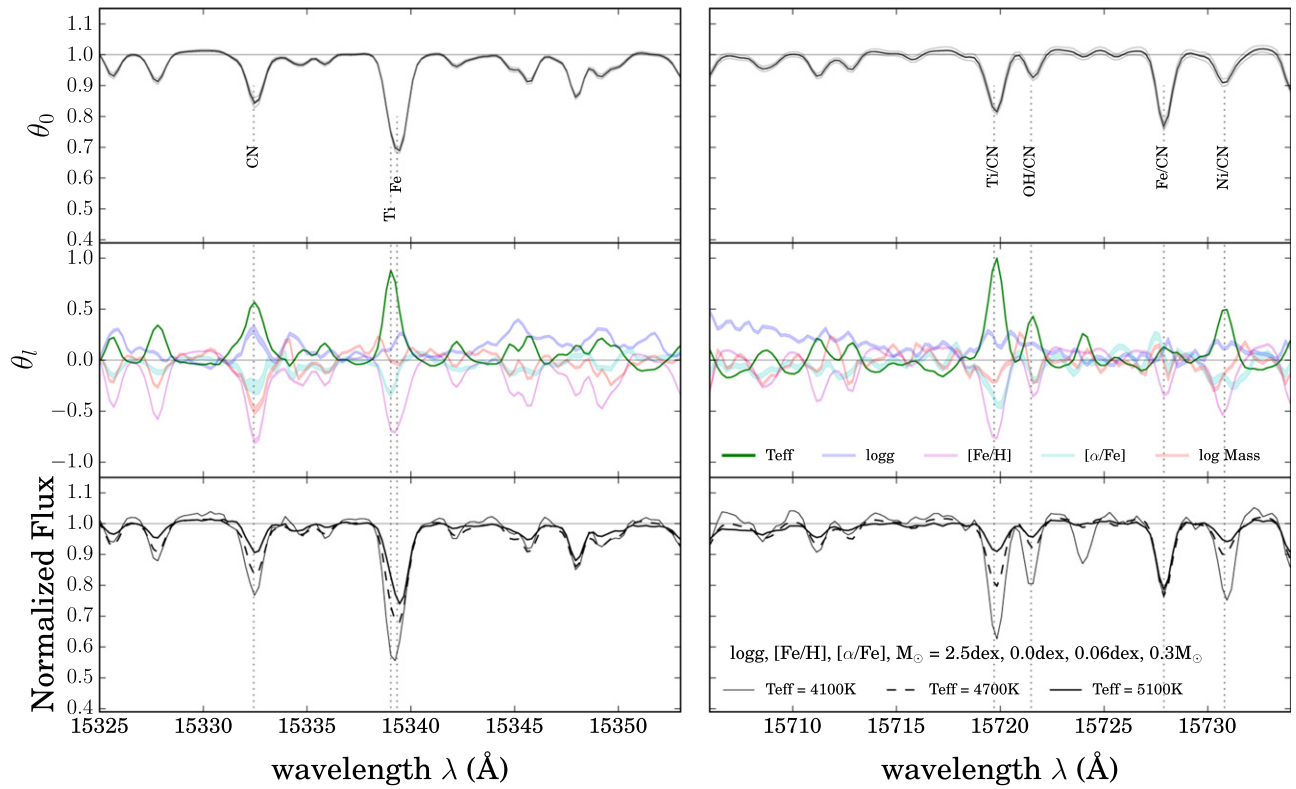
Figure 4 shows the same information as for Figure 3 but for the two highest  $T_{\text{eff}}$  coefficients,  $\theta_l = \theta_{T_{\text{eff}}}$ , centered on  $\approx 15338$  and  $15720$  Å. The highest  $T_{\text{eff}}$  coefficients correspond to the cores of two Ti lines in the H-band spectra (one of which is blended also with Fe and the other with CN). The temperature coefficient is typically positive in the APOGEE spectral region,

with exceptions, for example at the Brackett feature shown in Figure 3 (where it is inversely correlated with the  $\log g$  coefficient). As seen in Figure 4, is typically strongly anticorrelated with  $[\text{Fe}/\text{H}]$  and  $[\alpha/\text{Fe}]$ . This anticorrelation reflects that, in a spectrum at a given  $[\text{Fe}/\text{H}]$ , as the temperature increases, the lines weaken and so the flux decreases, whereas at a given  $T_{\text{eff}}$ , as the metallicity increases, the lines strengthen and the flux increases.

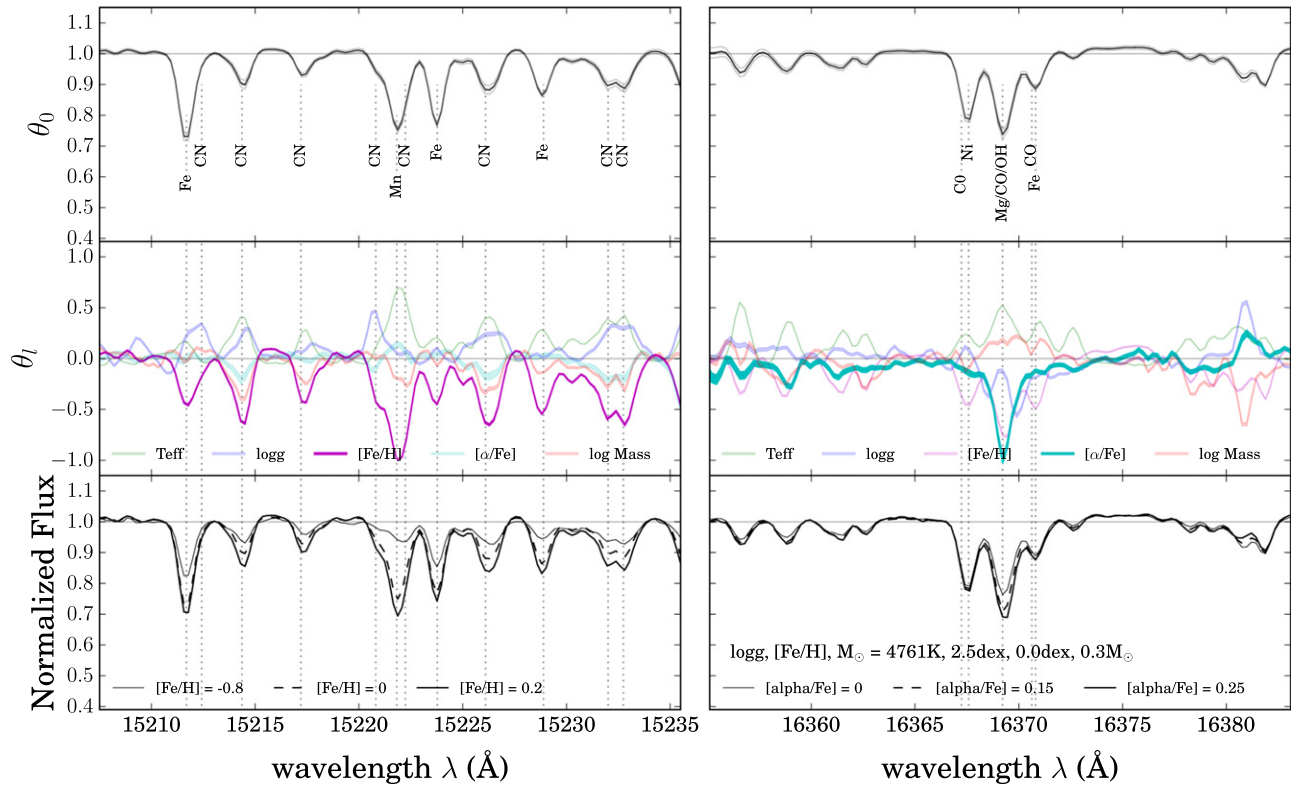
As we have a coefficient at *every* wavelength, which we can map to the chemical elements and molecules in the spectra using the APOGEE line list, we can interpret the spectral relationship between labels and flux in more detail than for an integrated absorption feature itself. For example, there is an asymmetry in the variation of the  $T_{\text{eff}}$  label in the left-hand bottom panel of Figure 4. This asymmetry likely reflects the changing ratio of the blends within this absorption feature (in this case, the feature is a blend of Ti and Fe, which are offset within this feature in their central wavelength).

The coupling of the data-driven model to stellar physics and mapping to the elements or molecules that determine the flux has important applications for stellar astrophysics. Here our aim is simply to verify that the information in the spectra or regions of highest spectral dependence on the labels originates from genuinely sensible and plausible chemical features in spectral space.



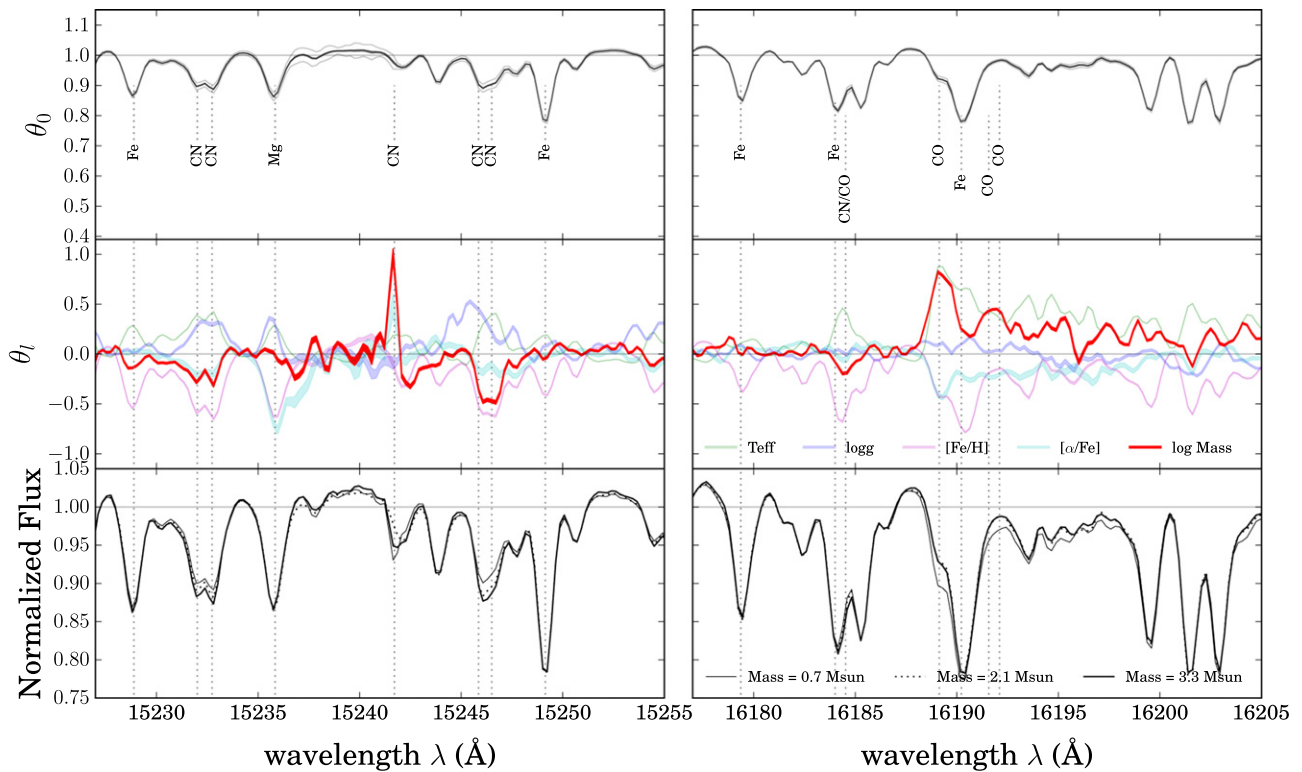


**Figure 4.** Same as for Figure 3 but showing the two 30 Å regions centered at the highest  $T_{\text{eff}}$  coefficient (in green). The  $T_{\text{eff}}$  coefficient is typically anticorrelated with the  $[\text{Fe}/\text{H}]$  and  $[\alpha/\text{Fe}]$  coefficients and traces the absorption profiles, reaching the highest amplitudes at the core of metal lines across the entire spectrum.



**Figure 5.** Same as for Figure 3 but showing the two 30 Å regions centered at the highest  $[\text{Fe}/\text{H}]$  coefficient (in magenta), at left, and the highest  $[\alpha/\text{Fe}]$  coefficient (in cyan), at right. The largest coefficient in  $[\text{Fe}/\text{H}]$  is from a Mn (iron peak) line that correlates with  $[\text{Fe}/\text{H}]$ . Note that for this line in the middle panel at left the  $[\alpha/\text{Fe}]$  coefficient is near zero where the  $[\text{Fe}/\text{H}]$  information is highest. The highest amplitude of the  $[\alpha/\text{Fe}]$  coefficient is seen in the core of one of the strong Mg lines in the APOGEE spectra, shown in the panel at right.





**Figure 6.** Same as for Figure 3 but showing the two 30 Å regions centered at the highest mass coefficients (in red). The bottom panels show that the spectra change line shape and depth with varying mass, and the mass coefficients are highest in the region of the CN and CO molecular blends. In particular, these blends corresponding to the highest mass coefficient in this figure comprise both  $^{12}\text{C}$  and  $^{13}\text{C}$ . Where the mass coefficients are negative, the flux is shallower at lower mass. Where the mass coefficient is positive, the flux is deeper at lower mass.

### 3.3.3. Spectral Dependencies on $[\text{Fe}/\text{H}]$ and $[\alpha/\text{Fe}]$

Figure 5 is demonstrative of the highest  $[\text{Fe}/\text{H}]$  and  $[\alpha/\text{Fe}]$  coefficient in the spectra,  $\theta_{[\text{Fe}/\text{H}]}$  and  $\theta_{[\alpha/\text{Fe}]}$ . The  $[\text{Fe}/\text{H}]$  and  $[\alpha/\text{Fe}]$  labels are typically correlated with the cores of all of the absorption features in the spectra, particularly for the  $[\text{Fe}/\text{H}]$  label, as seen at left. This is unsurprising as the overall metallicity,  $[\text{M}/\text{H}]$ , of a star simply correlates with the  $[\text{Fe}/\text{H}]$ , and the  $[\alpha/\text{Fe}]$  is known to increase with  $[\text{Fe}/\text{H}]$  and flattens to a plateau at high  $[\alpha/\text{Fe}]$  and low  $[\text{Fe}/\text{H}]$ , subject to the star-formation rate and initial mass function. For many (but not all) absorption features, the  $T_{\text{eff}}$  shows an inverse correlation with temperature, as seen in the left-hand panel of Figure 4.

The strongest  $[\text{Fe}/\text{H}]$  coefficient corresponds to a core of a (blended) Mn feature, the flux of which changes dramatically as a function of  $[\text{Fe}/\text{H}]$  over the range of  $-0.8 < [\text{Fe}/\text{H}] < +0.2$ , as shown in the bottom panel of Figure 4. Mn is one of the Fe-group elements (in addition to V, Ti, Cr, Co, and Ni), and this element is known to correlate directly with  $[\text{Fe}/\text{H}]$  (see Bergemann 2008; Battistini & Bensby 2015).

The largest coefficient in  $[\alpha/\text{Fe}]$  corresponds to the core of the strong (alpha-element) Mg line at  $\approx 16370$  Å (which is also blended with CO and OH), and unlike the  $\log g$  coefficient, it is the core of the line that correlates with  $[\alpha/\text{Fe}]$ . Note that for the  $\log g$  coefficient at this blended Mg feature in the middle panel of Figure 5 (right), the  $\log g$  coefficient is  $\approx 0$  at the very center of the line profile. The  $\log g$  coefficient increases to a much larger amplitude in the wings of the feature.

### 3.3.4. Spectral Dependencies on Mass

Finally, having verified that *The Cannon* delivers physically sensible origins of the  $T_{\text{eff}}$ ,  $\log g$ ,  $[\text{Fe}/\text{H}]$ , and  $[\alpha/\text{Fe}]$  labels, we examine the origin of the mass information from which we can infer the age of APOGEE stars. These first four labels are most straightforward to convincingly derive: indeed they are standard labels that are routinely determined from a stellar spectrum. However, delivering a mass label, mass, directly from stellar spectra marks a significant step forward in the exploitation of stellar spectra.

With the exception of a few specific indices that have been used previously to derive mass and inferred age, this work is the first claim of the success of a generalized approach for the extraction of stellar mass from spectra. Mathematically this works (as per the cross validation), and now we examine the interpretability of the mass label in terms of direct spectral signatures. For example, consider the case of main-sequence stars. In this case, mass is correlated strongly with effective temperature and more weakly with surface gravity in a composition-dependent fashion. A *Cannon*-like approach with mass labels would therefore be likely to have very similar spectral correlations, producing something equivalent to isochrone fitting for photometry. Red giants have a wide range in  $\log g$ , and different mass tracks differ only subtly in effective temperature, with a very strong metallicity dependence. One might therefore fear that any mass estimates would have very large random uncertainties, which is clearly not the case based on our results from Section 3.1.

From previous analyses in the ultraviolet and optical wavelength regions, we might expect spectral mass indicators, if present, to be realized in (1) chromospheric activity (emission), (2) dredge-up effects (and changing line strengths and profiles of particular elements), or (3) some combination of individual elemental abundances that reflect the enrichment history of the Milky Way with time (changing element ratios in the spectra).

Figure 6 shows the two largest coefficients in the log mass label. The information for the mass label is from (the relatively weak) CN and CO molecular features. Although we show only two regions as demonstrative, we have verified that the five highest mass coefficient amplitudes all correspond uniquely to predominantly CN but also CO molecular features. The relationship between mass and CN is consistent with the discovery by Martig et al. (2015), which shows that the [C/N] ratio calculated from APOGEE’s delivered catalog of C and N abundances in data release DR12 correlates with the mass and inferred age of the APOKASC stars. Salaris et al. (2015) also demonstrate the theoretical basis for the [C/N] ratio as an age indicator from after the first dredge-up. Martig et al. (2015) use the C and N abundances to create a model from these abundances and known masses of the APOKASC stars. With *The Cannon* this information is similarly exploited, only at the spectral level: we do not inform *The Cannon*’s generative model about the origin of the information (instead we rely on stellar physics to interpret the regions where the information is highest).

From the synthesized spectra in the bottom left-hand panel, it is apparent that it is not only the line strength that changes with the mass coefficient but also the line profile. Furthermore, the mass coefficient correlates with the [Fe/H] coefficient at the regions of the CN blends, at left, and anticorrelates with the [Fe/H] coefficient at the CO molecular feature, at right. Where the coefficient is positive, the flux of the model becomes larger at lower mass, whereas when the coefficient is negative, the line strength is weaker at smaller mass.

The changes in spectra as a function of mass are in general very subtle compared to the other labels. This is likely responsible for the relatively large scatter in the mass label determined with the take-stars-out test, shown in Figure 1. Note the other stellar labels have been historically well determined from spectra, even without general mathematical methods like *The Cannon* (which can optimally exploit all of the available and potential information). The correlations contained in the traditional stellar labels of  $T_{\text{eff}}$ ,  $\log g$ , [Fe/H], and  $[\alpha/\text{Fe}]$  are more straightforward to extract (e.g., [Fe/H] correlates with the cores of most absorption features in the spectra). This highlights the strength of an approach like *The Cannon* to determine and quantify the information that can be truly extracted from data, particularly as a function of signal to noise.

Examining the CN molecular regions in more detail, the two CN regions shown in Figure 6 with the highest mass coefficients are in fact a blend of CN molecules containing both  $^{12}\text{C}$  and  $^{13}\text{C}$ . Similarly, the CO feature at right is a blend of both  $^{12}\text{C}$  and  $^{13}\text{C}$ . It is this ratio that may drive the changing line profile as a function of mass and may play an important role in delivering the mass information from APOGEE spectra. This is because the  $^{12}\text{C}/^{13}\text{C}$  ratio is known as one of the best diagnostics of deep mixing in stellar interiors and so is known to contain information with respect to stellar mass.

Changes in isotope ratios complement information from the carbon-to-nitrogen ratio, and the combination is more powerful than any one indicator. In addition to their diagnostic power for mass, they also serve as markers of evolutionary state; carbon isotope ratios have already been used in the literature to differentiate first-ascent giants from red clump stars (Tautvaišienė et al. 2013, and references therein). The empirical data therefore naturally account for both the traditional first dredge-up effect, incorporating material processed in the core of the main-sequence precursor (Tautvaišienė et al. 2010), and in situ giant branch mixing (Gilroy & Brown 1991). This is true even in the absence of a predictive theory for the origin of the latter phenomenon.

### 3.4. Mass (and Age) Determination at a Given [Fe/H] and $[\alpha/\text{Fe}]$

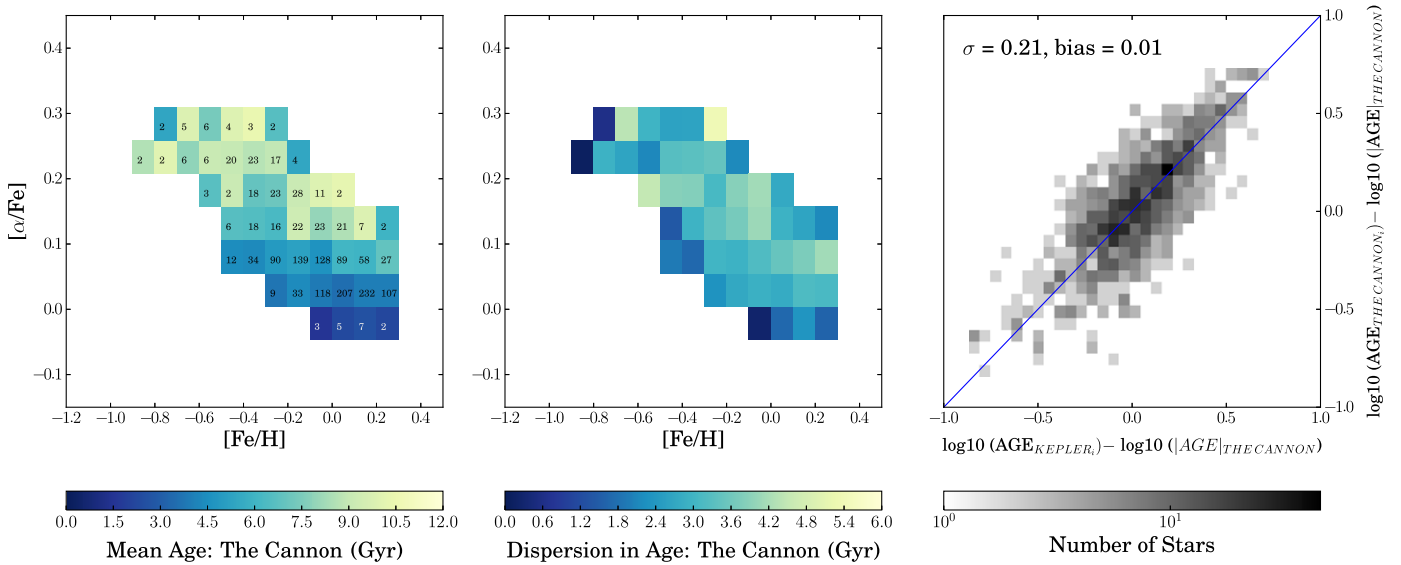
We use small regions of monoabundance space, that is, regions of a small range in [Fe/H] and  $[\alpha/\text{Fe}]$ , to demonstrate that we have a bona fide spectral mass label, from which we infer stellar ages (Bovy et al. 2012; Rix & Bovy 2013). We can thereby show that our mass/age label does not reflect simply some combination of the other four labels. We wish to illustrate, in particular, that the mass label that we use to infer age is not simply another expression of the  $[\alpha/\text{Fe}]$  label. The  $[\alpha/\text{Fe}]$  label itself is often used as an overall age proxy in abundance studies given gross expectations from stellar evolution and chemical yields in stellar populations.

In Figure 7, the APOKASC set of reference stars used to train *The Cannon* is shown in the  $[\alpha/\text{Fe}]$ –[Fe/H] plane. These stars are binned into small monoabundance boxes in this figure, and the panel at the far left indicates how many stars are in each of these bins. The color map represents the mean age and the age dispersion from *The Cannon* that is obtained in cross validation.

The input label for the inferred age is from the seismic scaling relations for these objects (from *Kepler*), and the output label is derived from the inferred age from the mass label output by *The Cannon* in cross validation (the take-star-out test in Section 2). The far right-hand panel of Figure 7 shows the individual age label for each star, from *The Cannon* (on the  $x$  axis) and from *Kepler* (on the  $y$  axis), subtracted from the mean age value in each age monoabundance bin. This is done for each bin and combined in this right-hand panel in the figure. If there were no additional information in each of the monoabundance bins with respect to age, that is, if the age information was simply a reincarnation of the  $[\alpha/\text{Fe}]$  label, then there would be no expected correlation between the difference in *The Cannon* and *Kepler* and the mean age. That there is a 1:1 relation between these two axes reflects that *The Cannon* works mathematically to determine the mass label and that the mass label within a monoabundance bin carries additional information.

## 4. MASSES AND AGES FOR APOGEE RED GIANT STARS IN DR12

For the following sections, the results for the ages of stars are inferred from their output mass label determined by *The Cannon*. We train on log mass, as described in Section 2.2, where the mass label for the APOKASC stars has been determined using the standard seismic scaling relations. We transform the output mass to age, as described in Section 2.3, for mapping the age



**Figure 7.** The  $[\alpha/\text{Fe}]$ - $[\text{Fe}/\text{H}]$  planes at left and center for the training set of APOKASC stars, colored by mean age of the stars in each bin at left, and the age dispersion at center, where the number of stars in each bin is indicated at left. The far right-hand panel shows the stars in the  $[\text{Fe}/\text{H}]$ - $[\alpha/\text{Fe}]$  plane where the individual stars  $i$  have been subtracted from the mean age in their respective bin, taking both the age from *The Cannon* and the age from *Kepler*. The panel on the far right shows the correlation in the input and output age labels delivered by *The Cannon* in cross validation within a narrow bin in abundance space.

distribution of the red giant stars in APOGEE’s DR12 across the disk, as shown in Figures 9 and 10. Our stellar labels determined by *The Cannon* for 70,000 red giant stars from DR12 (including  $\approx 20,000$  red clump stars) are provided in an online table, with a partial extract shown in Table 1.

Our fundamental assumption in our reported masses and inferred ages is that the efficiency of the dredge-up process is determined foremost by the star’s mass and metallicity. Therefore, it is important that our training sample spans much of the relevant mass (age) and  $[\text{Fe}/\text{H}]$  range. In turn, however, this implies (quite robustly) that the CNO abundances do not depend explicitly on the star’s orbit in the Galaxy, beyond the spatially varying probability of finding a star with a given mass and  $[\text{Fe}/\text{H}]$ . Our masses are calibrated across the stellar parameter range of the APOKASC training set described in Section 2.2.1. For our selection of stars for which we report masses and inferred ages, we take only stars from DR12 that are within the stellar parameter range of our training set of APOKASC stars, including for the C and N abundance space. These are stars that have experienced the first dredge-up. We make the following cuts using the ASPCAP DR12 parameters. This cut also excludes the metal-poor stars, for which the accuracy of the standard seismic scaling relations decreases (Epstein et al. 2014):

$$\begin{aligned} 4054 < T_{\text{eff}} < 5120 \text{ K} \\ 1.35 < \log g < 3.3 \text{ dex} \\ -0.85 < [\text{Fe}/\text{H}] < 0.30 \text{ dex} \\ -0.03 < [\alpha/\text{Fe}] < 0.30 \text{ dex}. \end{aligned}$$

For the 70,000 stars that remain after this selection in ASPCAP parameters, *The Cannon*’s labels compare very well to ASPCAP’s:  $T_{\text{eff}} = 16 \text{ K} \pm 50 \text{ K}$ ,  $\log g = 0.08 \pm 0.14 \text{ dex}$ ,  $[\text{Fe}/\text{H}] = 0.01 \pm 0.04 \text{ dex}$ , and  $[\alpha/\text{Fe}] = 0 \pm 0.03 \text{ dex}$ . As demonstrated in Martig et al. (2015) (Figures 1 and 2), our selection of stars within a narrow parameter range includes only those stars that have undergone the dredge-up, and this is not dependent on the spatial position. This is validated in Figure 12 of Martig et al.

(2015) by examining the  $[\text{C}/\text{N}]$  ratio of pre-dredge-up stars, which is constant as a function of galactic position, using distances derived following Ness et al. (2015b). We find that 90% of the stars within the stellar range of the training set also have a measured  $[\text{C}/\text{N}]$  and  $[(\text{C}+\text{N})/\text{M}]$  abundance that is within the abundance space of the training stars. We have examined the age distribution of these outlying stars compared to the 90th percentile that falls within the bounds of the training set and find these age distributions to be the same. Nevertheless, we exclude the 10% of the DR12 stars that have reported ASPCAP C and N abundances as a function of  $[\text{Fe}/\text{H}]$  that fall outside of the space spanned by the training set. We report the  $\chi^2$  statistic in Table 1 for each star in addition to our stellar labels; it is indicative of the fidelity of the stellar labels, including the mass, and we find a few percent of stars have a high  $\chi^2$  statistic, due to *The Cannon*’s model not being able to well fit the data.

#### 4.1. Stellar Ages for the Red Clump Sample

The APOGEE DR12 sample comprises primarily red giant stars plus a valuable subset of  $\approx 20,000$  red clump stars, identified by Bovy et al. (2014). These stars have individual distance uncertainties of 5%. These red clump stars cover a large radial extent of the disk, spanning distances of 4–15 kpc, and are located predominantly at heights  $|z| < 3.0 \text{ kpc}$  from the plane. The red clump sample is a representative and unbiased sample of Milky Way disk stars and has an expected age distribution peaking at about 1.8 Gyr with a tail out to old ages (see Figure 15 of Bovy et al. 2014). We take our model, trained using the reference APOKASC stars, and determine the stellar parameters and masses for these red clump stars. We then infer ages by interpolating in label space onto PARSEC isochrones.

The red clump may seem to be a surprising choice to use for age studies because stars in this evolutionary state are known to experience stochastic and significant mass loss relative to prior epochs. However, we do account for this mass loss, and, given

**Table 1**

Partial Column Excerpt from the Online Table of Six Stellar Labels ( $T_{\text{eff}}$ ,  $\log g$ ,  $[\text{Fe}/\text{H}]$ ,  $[\alpha/\text{Fe}]$ , Mass, and Age) Determined by *The Cannon* for 50,000 Red Giant Stars and 20,000 Red Clump Stars in APOGEE's Data Release DR12

star ID (2MASS)	$T_{\text{eff}}$ (K)	$\log g$ (dex)	$[\text{Fe}/\text{H}]$ (dex)	$[\alpha/\text{Fe}]$ (dex)	$\ln \text{Mass}$ ( $M_{\text{star}}$ )	$\ln \text{Age}$ (Gyr)	$\sigma(T_{\text{eff}})$ (K)	$\sigma(\log g)$ (dex)	$\sigma([\text{Fe}/\text{H}])$ (dex)	$\sigma([\alpha/\text{Fe}])$ (dex)	$\sigma(\log \text{mass})$ ( $M_{\text{star}}$ )	$\sigma(\log \text{age})$ (Gyr)	$\chi^2_{\text{reduced}}$
2M21353892+4229507	4085.31	1.39	-0.002	0.019	0.463	1.17	1.272	0.007	0.002	0.002	0.012	0.038	1.29
2M21354775+4233120	4685.85	2.84	0.07	0.165	0.022	2.158	8.43	0.018	0.006	0.005	0.033	0.095	2.3
2M21360285+4231145	4493.81	1.72	-0.431	0.025	0.236	1.487	4.96	0.018	0.006	0.005	0.034	0.103	2.4
2M21360302+4250260	4687.45	2.55	0.041	0.042	0.779	0.137	4.652	0.013	0.004	0.004	0.027	0.077	1.2

**Note.** The errors quoted are the formal errors from *The Cannon* for the uncertainties on the labels (see Figures 1 and 11). The mass column in this table is for training on mass derived from seismic scaling relations, and the age column in this table is derived from training on age from Martig et al. (2015) for the same set of 1639 reference stars from APOKASC.

(This table is available in its entirety in machine-readable form.)



other uncertainties, the age estimates for red clump stars are not dramatically more unreliable than those for first-ascent giant branch stars (see Casagrande et al. 2015 for a recent discussion). The higher age uncertainties are also compensated for to some degree by having more reliable distances.

#### 4.1.1. The Stellar Age Distribution of the Milky Way's Disk across 4–15 kpc

We have determined the masses and (from PARSEC isochrones) inferred the ages for the  $\approx 20,000$  red clump stars that have distances known to approximately 5%. We use these results to show the age distribution of the Milky Way's disk. The full catalog of the stellar labels determined with *The Cannon* for the red clump sample is included in Table 1. This data set represents the largest homogeneous sample of stars in the Milky Way with mass and associated age labels and extends the age mapping of the Milky Way from the previous local neighborhood only (GCS) to trace the inner to outer disk, from 4 to 15 kpc.

Figure 8 shows the median age of the red clump stars in the  $[\alpha/\text{Fe}]$ – $[\text{Fe}/\text{H}]$  plane (top left) and the density distribution of these stars (top right). We include the 17,065 stars with a  $\chi_{\text{reduced}}^2 < 2$  in this figure, which excludes 15% of the sample (all stars with their corresponding  $\chi_{\text{reduced}}^2$  statistic are given in Table 1). Most of the red clump stars are located in the low-alpha sequence. We select the low-alpha-sequence stars to examine the trends of the metallicity,  $[\text{Fe}/\text{H}]$ , as a function of radius, for the young compared to the intermediate-age stars. This selection is for all stars below the dashed line in the density distribution of the clump stars, at the top right of Figure 8. In the selection of these low-alpha stars we are selecting stars that should represent a single sequence of chemical enrichment. We therefore might expect differences in the distribution of  $[\text{Fe}/\text{H}]$  with radius for this sequence as a function of age. This difference would be not due to different formation histories but instead due to Galaxy evolution processes for this population over time (e.g., Roškar et al. 2008; Schönrich & Binney 2009).

The bottom panels of Figure 8 show density maps of the  $[\text{Fe}/\text{H}]$  of the youngest stars (at left) and the intermediate-age stars (at right) as a function of radius. At bottom left, there are 1669 stars with ages  $< 1$  Gyr, and, at right, there are 6716 stars with ages  $> 5$  Gyr. Note that there is an apparent overdensity at about 8 kpc across all  $[\text{Fe}/\text{H}]$  for the intermediate-age selection. These are the stars in the *Kepler* field in the sample.

Importantly, the median age of the red clump sample is not the median age of the population from which it is drawn. The red clump age distribution, from stellar evolution theory, is peaked at young ages. As discussed in Bovy et al. (2014), the red clump population is a long-lived evolutionary phase (and one for which precise distances can be determined) and is an excellent population tracer. At the same time, the fraction of the mass in the red clump is a function of the overall star-formation history or age distribution of the Milky Way's disk. The red clump, while being an excellent tracer of the Milky Way disk, does not represent the unbiased stellar distribution function of ages in the Milky Way disk.

The  $[\text{Fe}/\text{H}]$  distribution for intermediate-age stars as a function of  $R_{\text{GAL}}$  is less tightly correlated with radius compared to the youngest stars in the red clump sample. That the  $[\text{Fe}/\text{H}]$ –radius correlation weakens with age likely reflects dynamical evolution processes in the Milky Way that redistribute the stars in the disk, such as radial migration. Intermediate-age stars,

being longer lived, would have experienced a more significant dynamical timescale over which these processes take effect and so are scattered more from their original birth radii. The youngest stars have been subject to a shorter dynamical evolution history, and their current origin likely more tightly traces the origin of their birthplace, reflected in the correlation between radius and  $[\text{Fe}/\text{H}]$ , tracing the chemical enrichment of the gas, which increases toward the center of the Galaxy.

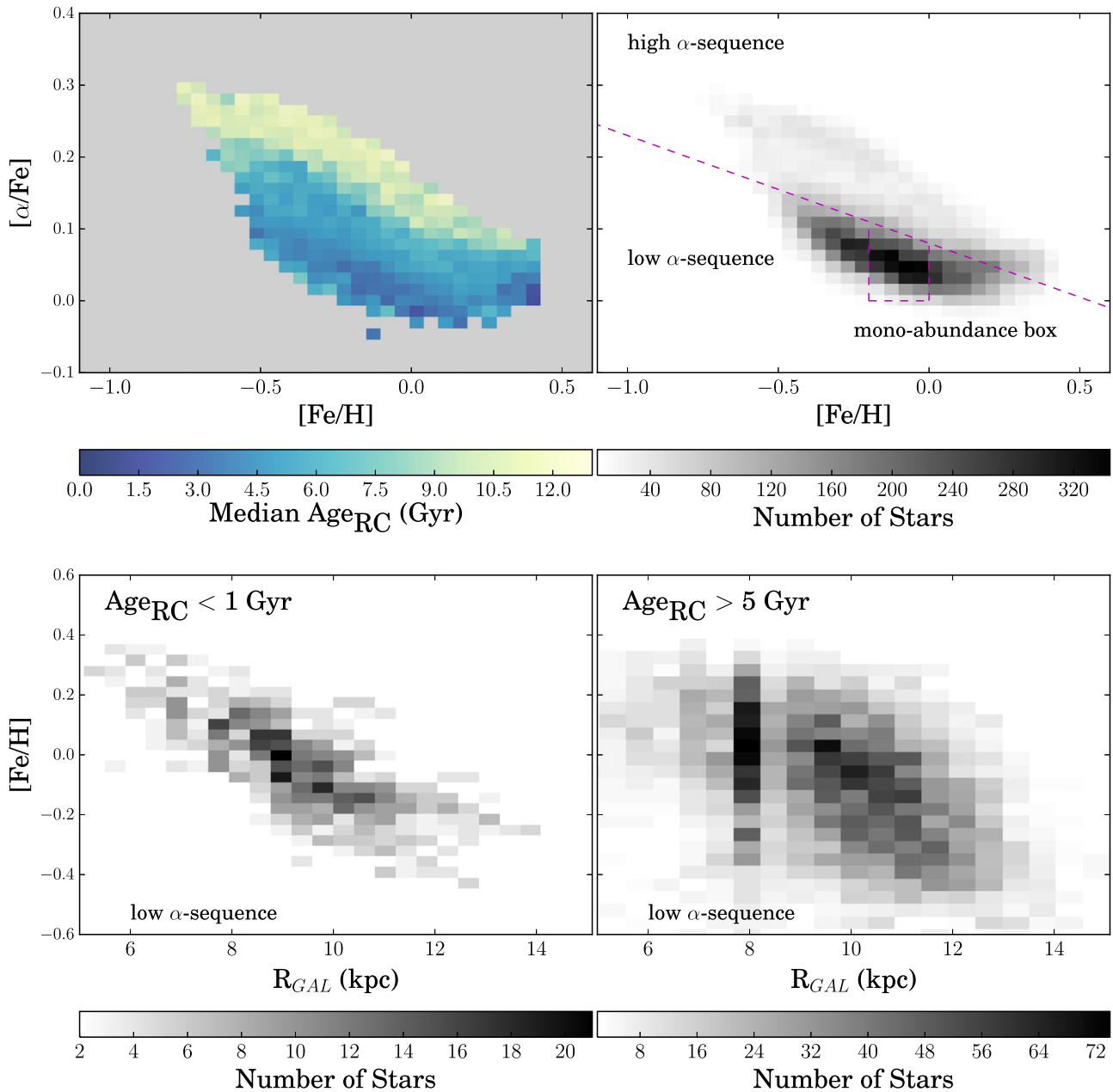
The top right-hand panel of Figure 8 shows a small box in the  $[\text{Fe}/\text{H}]$ – $[\alpha/\text{Fe}]$  plane from which we select stars for conditioning our age analysis on abundances. We use this monoabundance box to investigate and compare the mean trends of age across the disk ( $R_{\text{GAL}}, z$ ), contrasted with that for all stars, in demonstrating the information in the age label, even conditioned on abundances.

Figure 9 shows the  $(R_{\text{GAL}}, z)$  distribution of the red clump sample colored by median age across 4–15 kpc for (1) all stars, at left, (2) the low-alpha sequence, second from left, and (3) the monoabundance sample bin shown in the top right panel of Figure 8, third from left. The distribution of young, intermediate-age, and old stars, for all of the stars (far left panel), is shown in the histogram at far right.

Figure 9 demonstrates that the stars in a narrow  $|z|$  range in the plane are typically young, spanning the radial extent of the sample. There are fewer stars in the low-alpha sequence far from the plane, and the low-alpha sequence is dominant at larger radii (e.g., Hayden et al. 2015), but the same trends are seen in all three panels of age distributions. Older stars are present, preferentially at smaller radii, as seen most clearly in the far left panel, and these are typically located farther from the plane than younger stars. Stars transition to older ages farther from the plane as the radius increases, and there is an apparent vertical flaring in the age distribution with radius, with younger stars also dominating the ages at larger heights from the plane at the largest radii.

The histogram at far right shows the very different distributions of young and old stars. The number of youngest stars is strongly peaked near  $z \sim 0$  as these stars are concentrated toward the plane, suggesting ongoing star formation in the gas-enriched regions of the Galaxy. The older stars show a much broader distribution and extend to larger heights from the plane and are present in a larger relative fraction at smaller radii, preferentially at larger  $z$ . The younger stars extending out to larger radii, including farther from the plane, support an inside-out formation scenario for the Milky Way. These distributions, which show young stars also at large heights from the plane, imply that younger stars are also born at relatively large heights from the plane.

The center and far-right panels show a restricted distribution in  $z$ , as the young-alpha sequence is concentrated to the plane. Nevertheless, even conditioned on abundances, there are the same apparent age trends seen in the left-hand panel. Old stars are preferentially seen at larger heights from the plane, and for the low-alpha sequence, very few old stars are present at the largest radial extents of the sample. For the monoabundance selection in the panel at the far right, there are a handful of old stars present across the radial extent of the sample, preferentially at the largest heights from the plane. At the same time, as also seen for the center panel, there are young stars seen at large  $z$  across all  $R$ . Clearly, the young-alpha sequence does not represent a homogeneously young population, and the age label demonstrates that, conditioned on abundances, older stars are



**Figure 8.** The  $\approx 17,065$  red clump sample of stars. The top left panel shows all stars colored by median age, where the median age<sub>RC</sub> does not represent the median age of the population from which it was drawn. The density distribution of this sample is shown at the top right, and the sequences we use for examining the age of the disk of the low- $\alpha$  sequence and monoabundance population bins are indicated. The bottom panels show the  $[\text{Fe}/\text{H}]-R_{\text{GAL}}$  distribution for the young and intermediate-age stars in the low- $\alpha$  sequence.

distributed differently from younger stars in the disk of the Milky Way.

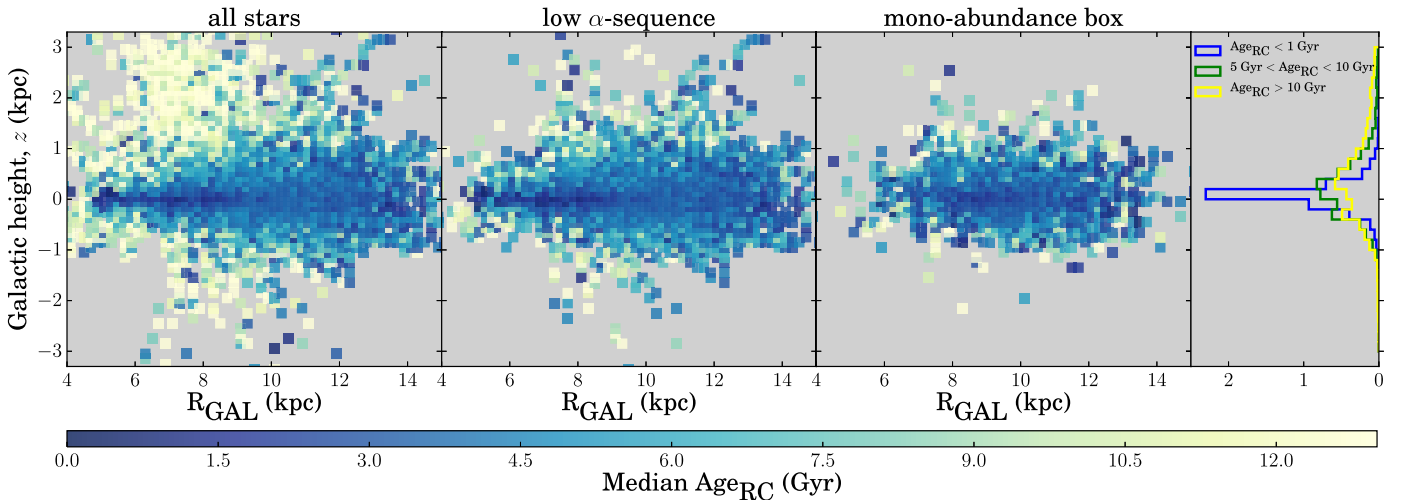
#### 4.2. Stellar Ages for the Red Giant Sample

In addition to the red clump stars, we have determined stellar masses and inferred ages (assuming the red giant evolutionary state) for the 50,000 red giant stars in DR12 that span the label range of our training set in  $T_{\text{eff}}$ ,  $\log g$ ,  $[\text{Fe}/\text{H}]$ , and  $[\alpha/\text{Fe}]$ . We include the labels for these red giant stars in Table 1.

In Figure 10 we show the fractional age distribution for the 50,000 red giant stars in DR12. Note the 5% of stars where the  $\chi^2_{\text{reduced}} > 3$  for *The Cannon* model have been removed. The distances to all of the red giant stars have been determined via

interpolation to PARSEC isochrones, from the stellar parameters and by adopting the RJCE-WISE extinction value for that line of sight, provided in the APOGEE DR12 data (Majewski et al. 2015; Zasowski et al. 2013). The panels are the same as for Figure 9 except now shown in terms of fractional ages (stars with ages  $< 5$  Gyr) and for a larger extent in ( $R_{\text{GAL}}, z$ ), as the red giant stars cover a much larger spatial region than the red clump alone. The distance uncertainty for the red giant stars is much larger than for the red clump, at about 30%. Distances toward the bulge are particularly uncertain and likely underestimated because of the high and differential reddening in this direction (see Ness et al. 2015).

Figure 10 demonstrates that the highest fraction of young red giant stars is in the plane of the disk, and this youngest fraction



**Figure 9.** ( $R_{\text{GAL}}, z$ ) maps of the median age of the APOGEE red clump stars showing  $\approx 17,065$  stars at far left, the low-alpha-sequence only in the second panel from left, and a small abundance bin in the low-alpha sequence, third panel from left. The final panel, at right, is a histogram of the different age distributions as a function of age, showing all stars across  $z$ , for a young, intermediate-age, and old selection.

flares in height with increasing radius (see the far left panel). The stars in the outermost region of the disk are predominantly young, and stars at the solar radius at large heights from the plane are almost all old. At a given height from the plane, the stars are on average younger moving out in radius from the center of the Galaxy. For the low-alpha sequence only (middle panel), the stars toward the center of the Galaxy comprise almost exclusively old stars, and stars in the outer regions are predominantly young. Young stars appear at all heights from the plane, even conditioned on abundances for the low-alpha sequence and the monoabundance population. Overall, the trends of the red giant sample are the same as that of the red clump sample.

## 5. DISCUSSION

We have provided three demonstrations of the validity of the stellar masses and ages determined with *The Cannon*. First, mathematically, *The Cannon* works and can return labels of  $T_{\text{eff}}$ ,  $\log g$ ,  $[\text{Fe}/\text{H}]$ ,  $[\alpha/\text{Fe}]$ , and mass for APOGEE spectra, which we validate with a take-stars-out test (see Figure 1). As shown with this cross validation, we can determine log masses to an accuracy of 0.07 dex and infer log ages from these masses to an accuracy of 0.21 dex.

The generative model well matched to the data (see Figure 2), from *The Cannon*'s best-fit labels, verifies that these five labels and the polynomial model (see Section 2.1) are sufficient to very well model the behavior of the flux with the labels at test time (see Figure 2). In fact, the data-driven model of *The Cannon* trained on these five labels only (no individual abundances) provides a better match to the real data than the synthetic stellar models utilized by ASPCAP.

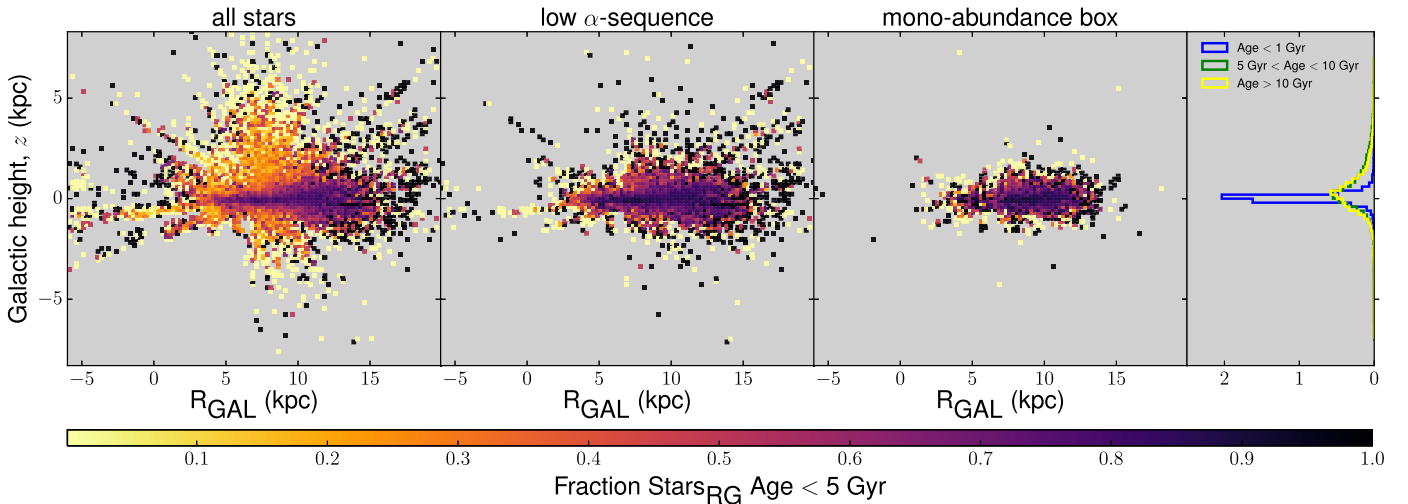
Second, we have shown that the spectral mass (or age) indicators discovered by *The Cannon* are associated in the space of the actual spectra with elements that can be “dredged up” (see Figure 6). Specifically, the mass information comes from the CN and CO molecules in the spectra. Although the mass information in the APOGEE spectral region originates from these features, in other wavelength regions it could derive from different elements or molecules. If mass information is present,

it can be determined using *The Cannon* for other surveys, such as GALAH (Freeman 2012).

Third, we show using the red clump sample of Bovy et al. (2014) that the ages of stellar structures in the Milky Way follow gross expectations, even conditioned on abundances. To demonstrate that we have a real age indicator and not a simple proxy for chemical enhancement that is tightly correlated with age (such as  $[\alpha/\text{Fe}]$ ), we have examined the age information within small monoabundance boxes in  $[\text{Fe}/\text{H}]-[\alpha/\text{Fe}]$  space for the training sample. Figure 7 shows that there is age information within the monoabundance bins. Furthermore, Figure 8 demonstrates the different  $[\text{Fe}/\text{H}]-$ radial profiles for the young and old red clump populations conditioned on abundances. For the low-alpha sequence only, the stars show an  $[\text{Fe}/\text{H}]$  distribution with radius that is consistent with radial mixing processes that are expected to be relevant for the intermediate-age and old populations but not the youngest stars.

The mean age map of the Milky Way disk as traced by the red clump stars shown in Figure 9 confirms the common wisdom that disk thickness depends on age. Moving out in radius, younger stars are present at larger and larger heights from the plane, and at small radii the youngest stars are located in significant fraction only in the plane of the disk. In the median age maps shown in Figure 9, there are old stars present even for the low-alpha sequence. Therefore, the low-alpha sequence is not a homogeneously young population. The oldest stars are located preferentially at larger heights from the plane compared to the younger stars, which truncate in their distribution nearer to the plane. The age distribution trends seen in the red clump sample as a function of  $(R, z)$  shown in Figure 9 are also seen in the red giant sample shown in Figure 10, which spans a larger extent in  $(R, z)$ .

For our analysis of stellar ages presented in Figures 7–10, we transform our mass labels into stellar age, as described in Section 4.1. Mapping the output mass labels from *The Cannon* to a stellar age using stellar models enforces fixed upper and lower age limits. It is also possible, however, to use *The Cannon* to train directly on log age rather than mass. In this case, there is no physical constraint on minimum or maximum ages at the test step.



**Figure 10.**  $(R_{\text{GAL}}, z)$  maps of the median age of the APOGEE red giant stars showing all  $\approx 50,000$  stars at far left, the low-alpha sequence only of these stars in the second panel from left, and a small abundance bin in the low-alpha sequence, third panel from left. The final panel, at right, is a histogram of the different age distributions as a function of age, showing all stars across  $z$ , for a young, intermediate-age, and old selection.

We provide in Table 1 a partial extract showing our stellar labels for 70,000 red giant stars from DR12 (including  $\approx 20,000$  red giant stars). This table is available in full online. We tabulate the  $T_{\text{eff}}$ ,  $\log g$ ,  $[\text{Fe}/\text{H}]$ , and  $[\alpha/\text{Fe}]$  as well as the stellar mass label determined by *The Cannon* for training on log mass and the stellar age label determined by *The Cannon* for the case of training directly on log age. For training on log age directly rather than log mass, the same set of reference stars is used. The label space of these stars is shown in the Appendix. The ages for the reference set of stars for training have been determined by Martig et al. (2015), who used interpolation between PARSEC isochrones with optimized scaling relations, as a function of evolutionary state.

There are several promising avenues for improving our results. Improved absolute calibrations for asteroseismic mass and radius would be highly desirable. Our methodology would also benefit from quantifying mass loss and its stochastic uncertainty, especially for red clump stars. A more complete stellar population study should also include corrections for the products of interacting binary star evolution and include the impact of the IMF and star-formation history on the derived mass and age distributions. There is also the possibility of using the mass trends identified in this paper to quantify first dredge-up and in situ red giant branch mixing as a function of mass and the initial abundance mixture and to test physical theories of stellar structure and evolution.

It is a pleasure to thank Maria Bergemann (MPIA) John Bochanski (Rider), Morgan Fouesneau (MPIA), Ricardo Schiavon (Liverpool John Moores University), Dan Foreman-Mackey (UW), Amelia Stutz (MPIA), and Ben Weiner (Arizona State) for valuable discussions and contributions. This project made use of the NASA Astrophysics Data System and open-source code in the *numpy* and *scipy* packages.

The research has received funding from the European Research Council under the European Union’s Seventh Framework Programme (FP 7) ERC Grant Agreement n. [321035].

We thank the Kavali Institute Theoretical Physics Galactic Archeology Program: this research was supported in part by the

National Science Foundation under Grant No. NSF PHY11-25915.

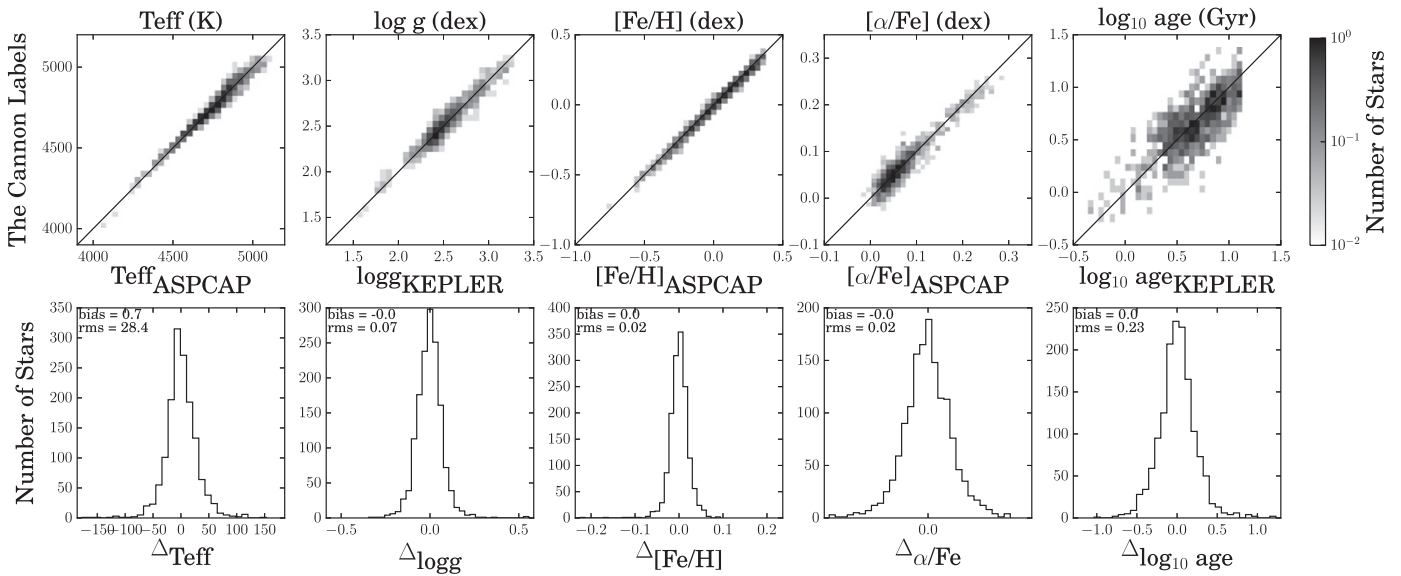
Funding for SDSS-III has been provided by the Alfred P. Sloan Foundation, the participating institutions, the National Science Foundation, and the U.S. Department of Energy Office of Science. The SDSS-III website is <http://www.sdss3.org/>.

SDSS-III is managed by the Astrophysical Research Consortium for the participating institutions of the SDSS-III Collaboration, including the University of Arizona, the Brazilian Participation Group, Brookhaven National Laboratory, Carnegie Mellon University, the University of Florida, the French Participation Group, the German Participation Group, Harvard University, the Instituto de Astrofísica de Canarias, the Michigan State/Notre Dame/JINA Participation Group, Johns Hopkins University, Lawrence Berkeley National Laboratory, Max Planck Institute for Astrophysics, Max Planck Institute for Extraterrestrial Physics, New Mexico State University, New York University, Ohio State University, Pennsylvania State University, the University of Portsmouth, Princeton University, the Spanish Participation Group, the University of Tokyo, the University of Utah, Vanderbilt University, the University of Virginia, the University of Washington, and Yale University.

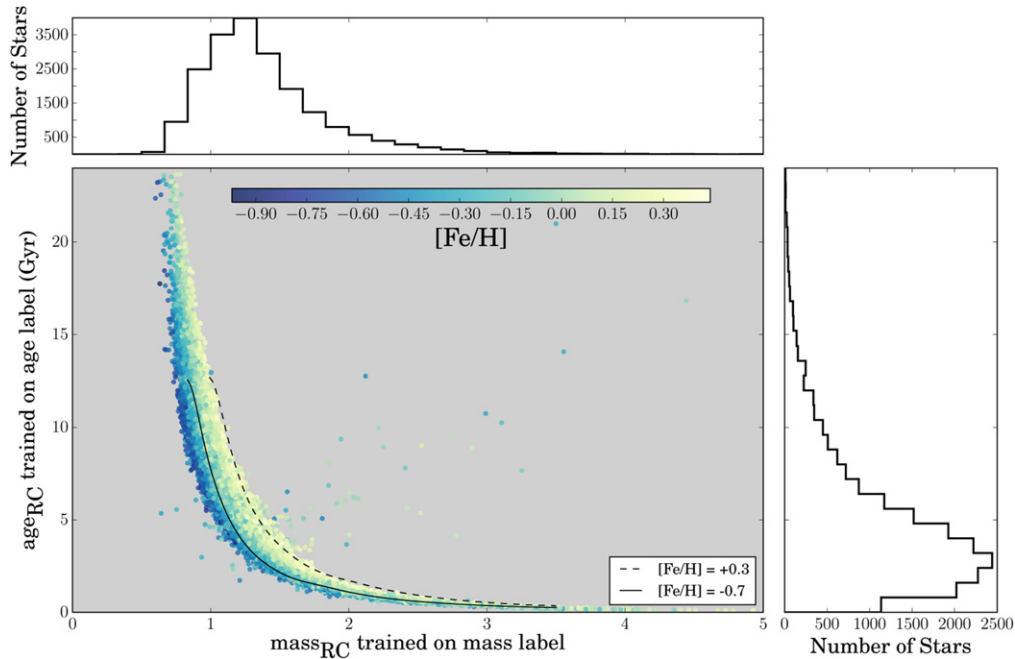
## APPENDIX

In Table 1, we provide the stellar parameters of  $T_{\text{eff}}$ ,  $\log g$ ,  $[\text{Fe}/\text{H}]$ ,  $[\alpha/\text{Fe}]$ , mass, and age for the DR12 red clump stars and red giant stars that are within the label range of our training set. The mass label from *The Cannon* is obtained for training on log mass, and the age label from *The Cannon* is obtained for directly training on log age. The mass label we provide can be used to infer stellar ages using interpolation between any selected stellar isochrones and given a set of assumptions. The age inferred from the mass label from *The Cannon*, as described in Section 4.1, was used to generate the stellar ages presented in Figures 7–10. In training on mass and inferring age, there are stars that are artificially truncated to the maximum age from the isochrones (where masses determined by *The Cannon* are lower than the smallest value from the stellar evolution tracks).





**Figure 11.** Cross validation of the training data set of 1639 stars for the  $T_{\text{eff}}$ ,  $\log g$ ,  $[\text{Fe}/\text{H}]$ ,  $[\alpha/\text{Fe}]$ , and age labels: the results for *The Cannon*'s labels for training performed on 90% of the APOKASC stars, showing the performance at test time on the 10% of the stars not included in training, run 10 times.



**Figure 12.** The mass label for the red clump sample from training on mass, compared to the age label for the red clump sample from training on age. The age distribution peaks at 2.5 Gyr for the red clump sample, which is colored by  $[\text{Fe}/\text{H}]$ . Almost all (>95%) of the stars are within  $0 < \text{age} < 14$  Gyr. Two theoretical mass and red clump age tracks are shown, at  $[\text{Fe}/\text{H}] = +0.3$  and  $-0.7$ , from PARSEC isochrones.

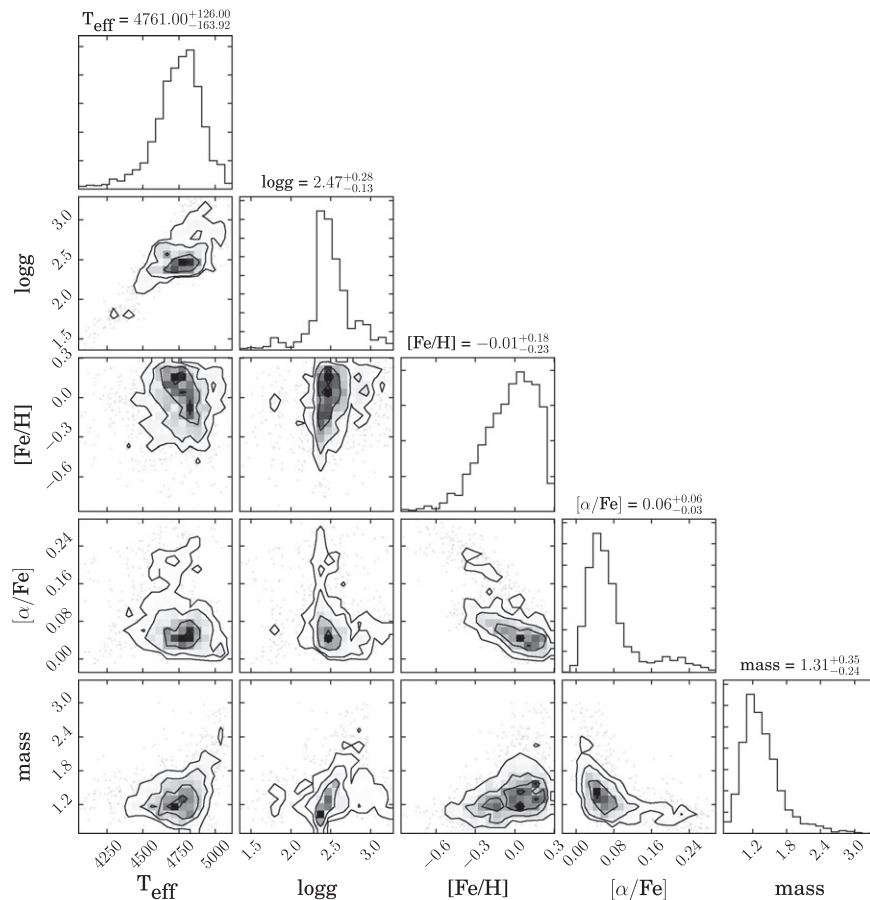
Training on log age directly instead of log mass, *The Cannon* works mathematically in the same way, as described in Section 2.1. The cross-validation result for training on age directly, instead of mass, is shown in Figure 11. The uncertainties on the labels are similar to that of Figure 1, for training on mass. There is no physical limit in the test step of *The Cannon* that prohibits ages (or masses) that exceed or are smaller than that of the training set. *The Cannon* therefore is not constrained to a physically allowed regime. Therefore, training on log age results in a small subset of stars that are older than the age of the universe at the test step, although the vast majority of stars are in physically realistic label space,  $0 < \text{age} < 14$  Gyr:

only 5% of stars are outside of this range and typically have large associated  $\chi^2_{\text{reduced}}$  values. The comparison for the age label determined for the red clump sample of stars, training on age, and the mass label determined for the red clump stars, training on mass, is shown in Figure 12. Two PARSEC tracks for red clump masses and ages are shown, demonstrating that the data follow theoretical expectations.

The label range of our training set of 1639 APOKASC stars is provided in Figure 13.

Our code and documentation are located on Github.<sup>9</sup>

<sup>9</sup> <https://github.com/mkness/TheCannon/>



**Figure 13.** The label space of the training data. This figure was made using the `corner.py` routine in Foreman-Mackey et al. (2014).

## REFERENCES

- Ahn, C. P., Alexandroff, R., Allende Prieto, C., et al. 2014, *ApJS*, **211**, 17
- Angelou, G. C., Stancliffe, R. J., Church, R. P., Lattanzio, J. C., & Smith, G. H. 2012, *ApJ*, **749**, 128
- Battistini, C., & Bensby, T. 2015, *A&A*, **577**, A9
- Bedding, T. R., Huber, D., Stello, D., et al. 2010, *ApJL*, **713**, L176
- Bensby, T., Yee, J. C., Feltzing, S., et al. 2013, *A&A*, **549**, A147
- Bergemann, M. 2008, *PhST*, **133**, 014013
- Bergemann, M., Ruchti, G. R., Serenelli, A., et al. 2014, *A&A*, **565**, A89
- Borucki, W. J., Koch, D., Basri, G., et al. 2010, *Sci*, **327**, 977
- Bovy, J., Nidever, D. L., Rix, H.-W., et al. 2014, *ApJ*, **790**, 127
- Bovy, J., Rix, H.-W., & Hogg, D. W. 2012, *ApJ*, **751**, 131
- Bressan, A., Marigo, P., Girardi, L., et al. 2012, *MNRAS*, **427**, 127
- Brogaard, K., Vandenberg, D. A., Bruntt, H., et al. 2012, *A&A*, **543**, A106
- Casagrande, L., Schönrich, R., Asplund, M., et al. 2011, *A&A*, **530**, A138
- Casagrande, L., Silva Aguirre, V., Schlesinger, K. J., et al. 2015, arXiv:1510.01376
- Casagrande, L., Silva Aguirre, V., Schlesinger, K. J., et al. 2016, *MNRAS*, **455**, 987
- Castelli, F., & Kurucz, R. L. 2004, arXiv:astro-ph/0405087
- Chaplin, W. J., & Miglio, A. 2013, *ARA&A*, **51**, 353
- Chiappini, C. 2002, *Ap&SS*, **281**, 253
- De Ridder, J., Barban, C., Baudin, F., et al. 2009, *Natur*, **459**, 398
- De Silva, G. M., Freeman, K. C., Bland-Hawthorn, J., et al. 2015, *MNRAS*, **449**, 2604
- Eisenstein, D. J., Weinberg, D. H., Agol, E., et al. 2011, *AJ*, **142**, 72
- Epstein, C. R., Elsworth, Y. P., Johnson, J. A., et al. 2014, *ApJL*, **785**, L28
- Foreman-Mackey, D., Price-Whelan, A., Ryan, G., et al. 2014, `triangle.py`, v0.1.1, doi:10.5281/zenodo.11020
- Freeman, K. C. 2012, in ASP Conf. Ser. 458, Galactic Archaeology: Near-Field Cosmology and the Formation of the Milky Way, ed. W. Aoki et al. (San Francisco, CA: ASP), 393
- Fuhrmann, K., Pfeiffer, M., Frank, C., Reetz, J., & Gehren, T. 1997, *A&A*, **323**, 909
- Gilmore, G., Randich, S., Asplund, M., et al. 2012, *Msngr*, **147**, 25
- Gilroy, K. K., & Brown, J. A. 1991, *ApJ*, **371**, 578
- Gray, D. F. 2008, *The Observation and Analysis of Stellar Photospheres* (Cambridge: Cambridge Univ. Press)
- Guenther, D. B., Kallinger, T., Reegen, P., et al. 2005, *ApJ*, **635**, 547
- Hayden, M. R., Bovy, J., Holtzman, J. A., et al. 2015, *ApJ*, **808**, 132
- Haywood, M., Di Matteo, P., Lehnert, M., Katz, D., & Gomez, A. 2013, arXiv:1305.4663
- Hekker, S., Elsworth, Y., Mosser, B., et al. 2013, *A&A*, **556**, A59
- Holtzman, J. A., Shetrone, M., Johnson, J. A., et al. 2015, *AJ*, **150**, 148
- Huber, D., Ireland, M. J., Bedding, T. R., et al. 2012, *ApJ*, **760**, 32
- Iben, I., Jr. 1967, *ApJ*, **147**, 650
- Kjeldsen, H., & Bedding, T. R. 1995, *A&A*, **293**, 87
- Kraft, R. P. 1994, *PASP*, **106**, 553
- Lagarde, N., Miglio, A., Eggenberger, P., et al. 2015, *A&A*, **580**, A141
- Lee, Y.-W., Demarque, P., & Zinn, R. 1990, *ApJ*, **350**, 155
- Majewski, S. R. 2015, arXiv:1509.05420
- Majewski, S. R., Zasowski, G., & Nidever, D. L. 2011, *ApJ*, **739**, 25
- Martig, M., Fouesneau, M., Rix, H.-W., et al. 2015, arXiv:1511.08203
- Martig, M., Rix, H.-W., Silva Aguirre, V., et al. 2014, arXiv:1412.3453
- McDonald, I., & Zijlstra, A. A. 2015, *MNRAS*, **448**, 502
- Mészáros, S., Holtzman, J., García Pérez, A. E., et al. 2013, *AJ*, **146**, 133
- Miglio, A., Brogaard, K., Stello, D., et al. 2012, *MNRAS*, **419**, 2077
- Ness, M., Hogg, D. W., Rix, H.-W., Ho, A. Y. Q., & Zasowski, G. 2015, *ApJ*, **808**, 16
- Ness, M., Zasowski, G., Johnson, J. A., et al. 2015b, arXiv:1512.04948
- Nordström, B., Mayor, M., Andersen, J., et al. 2004, *A&A*, **418**, 989
- Pinsonneault, M. H., Elsworth, Y., Epstein, C., et al. 2014, *ApJS*, **215**, 19
- Reimers, D. 1975, *MSRSL*, **8**, 369
- Rix, H.-W., & Bovy, J. 2013, *A&ARv*, **21**, 61
- Roškar, R., Debattista, V. P., Quinn, T. R., Stinson, G. S., & Wadsley, J. 2008, *ApJL*, **684**, L79
- Salaris, M., Pietrinferni, A., Piersimoni, A. M., & Cassisi, S. 2015, arXiv:1509.06904

- Schönrich, R., & Binney, J. 2009, [MNRAS](#), **396**, 203
- Silva Aguirre, V., Casagrande, L., Basu, S., et al. 2012, [ApJ](#), **757**, 99
- Silva Aguirre, V., Davies, G. R., Basu, S., et al. 2015, [MNRAS](#), **452**, 2127
- Snedden, C., Lambert, D. L., & Whitaker, R. W. 1979, [ApJ](#), **234**, 964
- Shetrone, M., Bizyaev, D., Lawler, J. E., et al. 2015, [ApJS](#), **221**, 24
- Soderblom, D. R. 2010, [ARA&A](#), **48**, 581
- Tautvaišienė, G., Barisevičius, G., Chorniy, Y., Ilyin, I., & Puzeras, E. 2013, [MNRAS](#), **430**, 621
- Tautvaišienė, G., Edvardsson, B., Puzeras, E., Barisevičius, G., & Ilyin, I. 2010, [MNRAS](#), **409**, 1213
- Yong, D., Grundahl, F., & Norris, J. E. 2015, [MNRAS](#), **446**, 3319
- Zasowski, G., Johnson, J. A., Frinchaboy, P. M., et al. 2013, [AJ](#), **146**, 81