The Astronomical Journal, 167:73 (18pp), 2024 February https://doi.org/10.3847/1538-3881/ad12a6

Spectroscopic Distance, Mass, and Age Estimations for APOGEE DR17

Alexander Stone-Martinez[1], Jon A. Holtzman[1], Julie Imig[1,2], Christian Nitschelm[3], Keivan G. Stassun[4], and
Joel R. Brownstein[5]
[1] Department of Astronomy, New Mexico State University, P.O. Box 30001, MSC 4500, Las Cruces, NM 88033, USA
[2] Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA
[3] Centro de Astronomía (CITEVA), Universidad de Antofagasta, Avenida Angamos 601, Antofagasta 1270300, Chile
[4] Department of Physics and Astronomy, Vanderbilt University, VU Station 1807, Nashville, TN 37235, USA
[5] Department of Physics and Astronomy, University of Utah, 115 S. 1400 E., Salt Lake City, UT 84112, USA
Received 2022 September 14; revised 2023 November 14; accepted 2023 November 15; published 2024 January 24

Abstract

We derive distances and masses of stars from the Sloan Digital Sky Survey (SDSS) Apache Point Observatory Galactic Evolution Experiment Data Release 17 using simple neural networks. Training data for distances comes from Gaia EDR3, supplemented by literature distances for star clusters. For masses, the network is trained using asteroseismic masses for evolved stars and isochrone masses for main-sequence stars. The models are trained on effective temperature, surface gravity, metallicity, and carbon and nitrogen abundances. We found that our distance predictions have median fractional errors that range from ≈20% at low log g and ≈10% at higher log g with a standard deviation of ≈11%. The mass predictions have a standard deviation of ±12%. Using the masses, we derive ages for evolved stars based on the correspondence between mass and age for giant stars given by isochrones. The results are compiled into a Value Added Catalog called DistMass that contains distances and masses for 733,901 independent spectra, plus ages for 396,548 evolved stars.

Unified Astronomy Thesaurus concepts: Neural networks (1933); Stellar distance (1595); Stellar masses (1614); Stellar properties (1624); Stellar ages (1581); Abundance ratios (11); Stellar populations (1622)

1. Introduction

The Milky Way offers unique opportunities for studying galaxy formation and evolution. We can measure many of the properties of the nearby stars, such as the three-dimensional position and velocity (through astrometry and spectroscopy), stellar parameters and stellar elemental abundances (through high-resolution spectroscopy), and stellar masses through asteroseimology and binary stars. However, these properties become more challenging to measure at larger distances from the Sun and for dimmer stars, so methods of extending the range of distances and luminosity over which we can reliably measure them is an important step to furthering our understanding of Galactic formation and evolution.

Recently, large-scale surveys such as Apache Point Observatory Galactic Evolution Experiment (APOGEE; Blanton et al. 2017; Majewski et al. 2017), GALAH (Buder et al. 2021), LAMOST (Wang et al. 2021), and Gaia-ESO (Binks et al. 2022) have increased the quantity of high-quality stellar parameters, Gaia (Gaia Collaboration et al. 2016) has provided precise astrometric data in the form of parallaxes and proper motions for large swaths of the Galaxy, and light-curve data from the Kepler space telescope have allowed for asteroseismic measurements of thousands of stars (e.g., Pinsonneault et al. 2018). With asteroseismology, it is possible to derive the evolutionary state, mean density, mass, and radius of stars. However, accurate parallaxes are more difficult to derive for stars at large distance: for example, at 5 kpc the median error in Gaia data is 15%, and at 8 kpc it grows to 34%. Asteroseismic measurements are limited to the few distinct fields that were observed over long periods of time, and out of those fields the asteroseismic measurements are limited to stars that pulsate. Deriving the distances and masses for a larger number of stars than current Gaia observations and asteroseismic measurements can provide is clearly valuable for expanding research into Galactic formation and evolution.

Previous work has used spectroscopic parallax to determine distances; see, e.g., Hawley et al. (2002), Burnett & Binney (2010), Santiago et al. (2015), Das & Sanders (2018), Queiroz et al. (2018), and Leung & Bovy (2019). Queiroz et al. (2018) developed a Bayesian tool, StarHorse, for calculating spectroscopic parallax distances using APOGEE data along with the Gaia parallaxes. Leung & Bovy (2019) developed a different tool that uses machine learning to accomplish the same goal. Their tool, AstroNN, utilizes the spectral data directly to empirically train a model. The large volume of stars with precise Gaia distances makes machine learning an attractive method for calculating spectroscopic parallax distances because it has the advantage of being an empirically driven approach that is not reliant on stellar modeling. Empirically driven methods can capture and utilize small details in the data that stellar models do not model (Ness et al. 2015, 2016; Imig et al. 2022).

An alternative to asteroseismology for the determination of stellar masses is the use of carbon and nitrogen abundance ratios (Martig et al. 2016). As stars ascend the red-giant branch, the observed C and N abundances change at the surface of the star. This is due to the convective regions of stars reaching down deeper into the star and dredging up material processed during the main-sequence phase. In the CNO cycle, the nitrogen-to-oxygen reaction is the slowest, so in regions with CNO reactions, the nitrogen abundance will rise and the carbon abundance will fall. As the star goes through dredge-up, this processed material with the shifted carbon/nitrogen ratio is mixed into the outer layers of the star. In more massive stars, there is more CNO processing and the dredge-up moves more material, which results in the [C/N] ratio changing more in higher-mass stars (Lagarde et al. 2012; Masseron & Gilmore 2015; Salaris et al. 2015; Shetrone et al. 2019; Vincenzo et al. 2021). However, using stellar models to predict the mass is problematic due to uncertainty in the exact mixing process that affects the surface abundances, making the use of model predictions of carbon and nitrogen abundances to estimate stellar mass difficult (Martig et al. 2016). As a result, empirical calibration based on, e.g., asteroseismic masses is an attractive option. A potential issue with empirical calibration lies in the possible primordial variations in the [C/N] ratio (Masseron & Gilmore 2015). This issue is addressed in Section 5.

In this paper, two different neural networks are used to estimate stellar masses and distances using the APOGEE Stellar Parameters and Abundances Pipeline (ASPCAP; Nidever et al. 2015; García Pérez et al. 2016; Majewski et al. 2017) stellar parameters as inputs. This is in contrast to the methodology of Leung & Bovy (2019) who use the full spectrum as the input. Using a smaller number of input parameters means that the model architecture becomes much simpler, which reduces the computational load and reduces the risk of overfitting. An advantage of our model is that the noise in the full spectrum does not influence the model training. ASPCAP utilizes the whole spectra to obtain the stellar parameters and abundances, so the input parameters of our model do not contain the signal noise found in the spectrum. Another advantage of using the parameters instead of the full spectrum is that the model is not intrinsically tied to data from APOGEE, so our model could use stellar parameters derived from other data sources. Finally, training on parameters allows us to present empirical relations between parameters and masses that may be of use in constraining stellar mixing models.

Our models are empirically trained using accurate distance and mass data. For the distance model, the training data come from Gaia targets with low parallax uncertainty, while for the mass model, mass estimates come from asteroseismic measurements from the APOKASC 3 catalog (M. Pinsonneault et al. 2024, in preparation). The mass model is to some extent an expansion of the results presented in Martig et al. (2016) with an expanded data set and using a neural network instead of polynomial relations between masses, stellar parameters, and abundances. Our work is also similar to Das & Sanders (2018), with the main distinctions being the training on a much larger sample of stars from new catalogs (Gaia DR3, APOKASC 3) and the application of the models to all stars for distance predictions, and to all evolved stars for age predictions.

We also present age estimations for giant stars based on the mass predictions from our model. This is possible for evolved stars because stellar mass and main-sequence lifespan are tightly correlated. Evolved stars exist for only the final 1%–10% of the stars lifespan, so we can infer the age of an evolved star using the predicted stellar mass and isochrones. However, ages can only be inferred in this way for evolved stars. We also provide some extended discussion of different training data for the masses.

Section 2 describes all the various data and data sources used for this project. Section 3 covers the machine-learning model and how training and testing the model was performed. Section 4 presents the distance model, including the definition of the training set and validation tests. Section 5 presents the mass model. Section 6 covers determining stellar age from the mass predictions, and Section 7 summarizes the results, including limitations.

Our catalog described in this work can be found at https://www.sdss4.org/dr17/data_access/value-added-catalogs/?vac_id=distmass:-distances,-masses,-and-ages-for-apogee-dr17.

2. Data

Our models use the stellar parameters and abundances from the APOGEE ASPCAP (Nidever et al. 2015; García Pérez et al. 2016; Majewski et al. 2017; Zasowski et al. 2017) pipeline. For this project, the calibrated effective temperature, surface gravity, metallicity ([Fe/H]), carbon([C/Fe]), and nitrogen ([N/Fe]) abundance parameters were used. These data come from APOGEE Data Release 17 (DR17; Abdurro'uf et al. 2022). Spectra from APOGEE are obtained via a multi-object spectrograph (Wilson et al. 2019) attached to the Sloan 2.5 m telescope (Gunn et al. 2006) at the Apache Point Observatory for the northern hemisphere and the Irenee du Pont Telescope (Bowen & Vaughan 1973).

For estimating distances, the model estimates an absolute magnitude, which is then used with the apparent magnitude and an extinction estimation to derive a distance. For the apparent magnitude, we used the 2MASS K-band magnitude. Extinction estimations of stars at lower galactic latitudes ($b < 16^\circ$) are based on the RJCE method (Majewski et al. 2011) using IRAC photometry (Fazio et al. 2004) when available, and WISE (Wright et al. 2010) for all other stars. Data from IRAC are preferred because of the higher spatial resolution, especially in the crowded low galactic latitude fields. At high galactic latitudes, stars are more likely to be behind dust, so the extinctions from Schlegel et al. (1998) are used. It should be noted that any systematic errors in the extinctions will lead to systematic errors in the distance predictions.

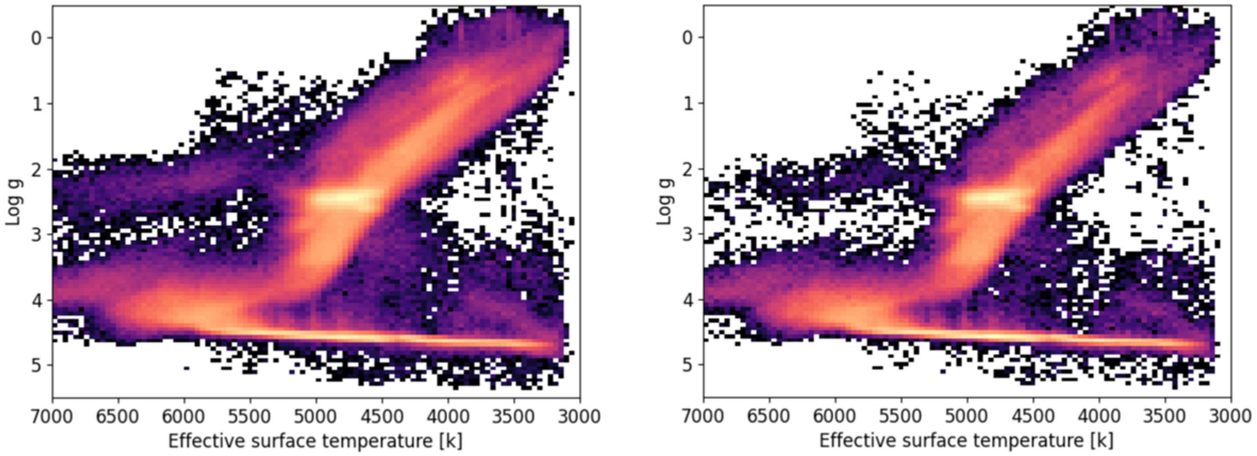
The training labels used for the distance model came from absolute magnitudes derived from Gaia parallax data in DR3 (Gaia Collaboration et al. 2016). Specifically, we used distance estimates made using a weak distance prior, which improves the distance estimate over that obtained by simply inverting the parallax (Bailer-Jones et al. 2018).

The data for mass estimation comes from APOKASC (M. Pinsonneault et al. 2024, in preparation), which is an asteroseismic and spectroscopic survey combining data from APOGEE and the Kepler Asteroseismology Science Consortium (KASC). This is supplemented with additional mass data for subgiants (using asteroseismology) and main-sequence stars (using isochrones) from Serenelli et al. (2013) and Berger et al. (2020), respectively.

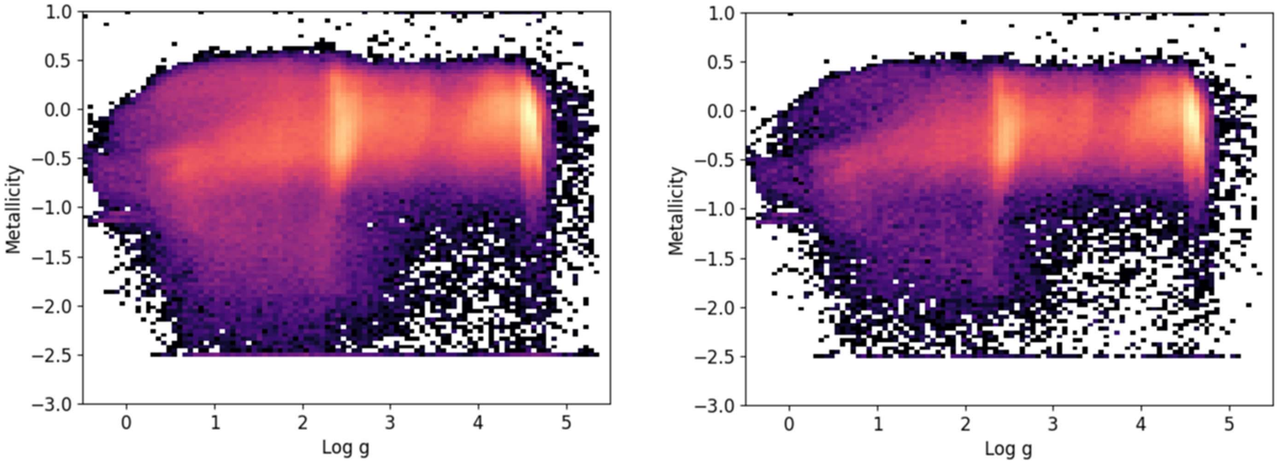
The data sets were trimmed based on several criteria to eliminate bad data from model training. For both models, stars with missing data values for any of the derived parameters were excluded, along with stars flagged with the STAR_BAD bit. Additional cuts and additions to the input data are discussed in following sections for the distance and mass neural nets, because those cuts are specific to their respective neural nets.

3. Machine-learning Background

For this work, two small supervised regression models were trained to predict the absolute magnitude and masses of stars, respectively, using the parameter $\log g$, the effective



(a) Full data set kiel diagram (b) Post uncertainty cut kiel diagram



(c) Full data set log g vs metallicity (d) Post uncertainty cut log g vs metallicity

Figure 1. Density plots showing the Kiel and metallicity diagrams and for the full data set ((a) and (c)) and for the training set after the 15% distance uncertainty cutoff ((b) and (d)). Color scale is logarithmic.

temperature, the metallicity, and the carbon and nitrogen abundances. A supervised machine-learning model is one in which independent data are used to train and grade the predictions of the model. Model training iterates over the training data while changing the internal model parameters with the goal of reducing the value of the loss function, which quantifies the performance of the model.

The data were split into a training, a validation, and a test set. The training set comprised of 64% of the training stars. The rest of the stars are part of the validation set and the test set, comprising 16% and 20% of the data, respectively. The test set is withheld from training and used to determine the model's performance after training is complete. The validation set is also withheld during model training, like the test set, but it is used during training as an unbiased estimate of the model's performance while tuning the model parameters. The sizes of the training and validation set were chosen based on the 80/20 rule, where 20% of the data is used for testing while the remaining 80% is used in training/validation (Tennenholtz et al. 2018). If the model is overfitting the training data, then the loss function of the validation set will change little as the training set loss function decreases. Because the validation set has an indirect influence on model training, it also cannot be fully trusted for testing. After training, the final performance of the trained model is checked using the test set. It is used to verify the model is making reliable predictions and to find issues, such as over/underfitting, not addressed in training.

3.1. Machine-learning Architecture

To train and apply the models, we used the Tensorflow Python package. Tensorflow is an open source library that makes the creation and training of neural networks with many layers relatively easy (Abadi et al. 2015). The model architectures used in both models was a simple single dense layer with eight neurons that feed into an single neuron output layer. This architecture was chosen after iterating over several model architectures and examining performance with the test set. We found that larger models tended to have overfitting and smaller ones did not capture the information needed from the input data. For the loss function, we used the mean squared error normalized by the uncertainties.

The input parameters were normalized and scaled before training by subtracting the mean of each parameter and dividing by the standard deviation. This process reduces the overall model training time. An important note in doing this is

Table 1
List of Star Clusters Used to Supplement the Distances

Cluster	Adopted Distance (Kpc)
M92	8.2
M15	10
N5466	15.9
N4147	19
M2	12
M13	6.8
M3	10
M5	8
M12	5
M107	6
N2243	4.5
N2158	4
N2420	2.5
N188	1.7
M67	0.9
N7789	2.2
N6791	4.078
N5053	17.4
M68	10
N6397	2
M55	5
M22	3
M79	13
N3201	5
M10	4
N6752	4
Omegacen	5
M54	27
Pal5	23
N6544	2.7
N288	8.8
N362	9
N1851	12
M4	2
N2808	9.6
47TUC	5
N6388	10
N6553	6.01

that any data used as input for the models after training will have to be normalized in the same way. The models were then trained until the loss function of the validation set stopped decreasing, which occurred at around 80 iterations for our models (each epoch meaning a complete run-through of the entire data set).

4. Distances

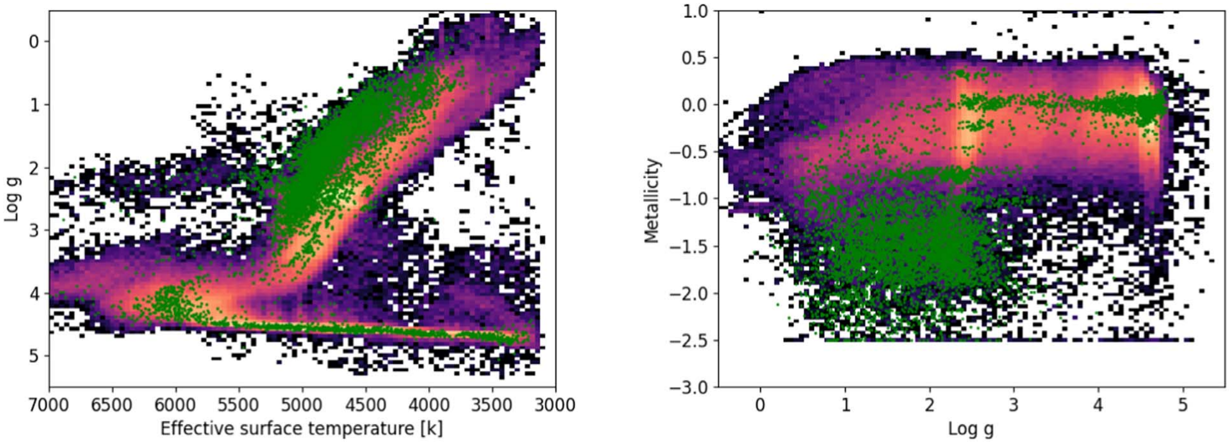
4.1. Training Set

For training, we adopted all stars with Gaia parallax uncertainties less than 15%. These stars tend to be within 2–5 kpc, depending on luminosity. We adopted an upper cutoff of 15% because it reduced the influence of stars with poor distance precision while retaining a large number of stars covering the input parameter space. A smaller cutoff would have resulted in a smaller input parameter space covered by the training set. We also adopted an extinction cutoff of $A_K < 0.5$. This prevents stars with large extinctions from influencing training. The large-extinction stars have higher extinction uncertainties, resulting in larger absolute magnitude uncertainties. This is a problem for our model, which uses the absolute magnitudes as the training labels. The machine-learning model can only be applied over the region of parameter space where the stars in the training set are located. Unfortunately, cutting stars with higher Gaia uncertainty results in parts of the training parameter space having insufficient numbers of stars for training. Figure 1 show the parameter space coverage of the full data set ((a) and (c)) and the training data set ((b) and (d)) after removing high-uncertainty stars and randomly sampling the full data set to create the training set as discussed in Section 3.1. Figure 1(b) shows there is sparser coverage along the upper giant branch, especially at $\log g < 1$. Figure 1(d) shows the same lack of stars at low surface gravity and also that, for stars with $0.5 < \log g < 2.5$, there is a significant drop in the number of stars with $[\mathrm{Fe/H}] < -1$. These cutoffs result in 487,543 stars remaining (0.66% of DR17) in the training set.

To increase the parameter space covered by the training data set, especially in the low-metallicity regime, stellar distances were augmented with distance data for star clusters, for which independent distances are available from main-sequence fitting or other methods. A list of all the clusters used is given in Table 1. This resulted in the addition of 7039 stars with cluster distances that were originally removed due to high parallax uncertainty. Figure 2 shows the increase in parameter space coverage that comes from including cluster stars. The typical uncertainties for these cluster distances are between 7%–10% (van de Ven et al. 2005; Friel et al. 2010; Paust et al. 2010; Shao & Li 2019). The addition of these stars is useful for training our model because they are mostly metal-poor stars, and thus they help fill in a region of parameter space that previously had insufficient numbers of stars.

4.1.1. Binary Removal

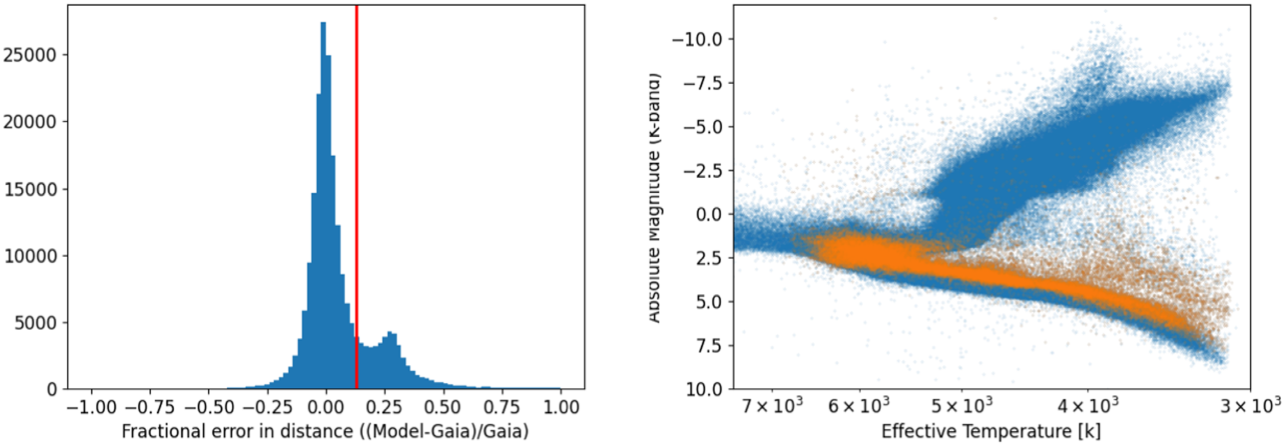
Because the distances are calculated from the apparent magnitude and the model-predicted absolute magnitude, there will be systematic errors for unresolved binaries that will be brighter than their derived parameters would suggest. This situation is further complicated for double-lined spectroscopic binaries, where the derived parameters may be less accurate if the light from multiple stars contributes to the spectrum. This in turn might cause predictions for all stars to have systematic errors because the errors from binary stars will affect the model training. To remove binaries from the training set, we trained and ran a distance model and compared the distances to Gaia distances. We attempted to just use the binary flag and radial velocity scatter from DR17, but these did not remove enough binaries from the training set. Stars in binaries will have a higher residual due to the model predicting a dimmer luminosity for them. This results in a bimodal distribution in fractional error. Stars with a residual fractional distance above 0.13 are then removed from the training set. The threshold was determined by picking the residual value that splits the bimodal distribution so that the high-residual samples are removed. Figure 3 shows the residual plot with the cutoff line (a) and where the cut stars are located on an HR diagram compared to the rest of the data set. This process improved the performance of the model but does also cut out some nonbinary stars. Other methods for binary identification such as radial velocity scatter or eclipsing binaries do not have as high a false-positive rate, but they do not remove enough of the binary stars from the training set to improve the model. It was important to remove as many binaries from the training set as possible, so the high false-positive rate of our method was deemed acceptable. This



(a) kiel Diagram of training space after addition of cluster stars

(b) Log g vs Metallicity of training space after addition of cluster stars

Figure 2. HR diagram showing stars with added cluster distances shown in green (a) Coverage of log *g* and metallicity parameter space showing stars with added cluster distances shown in green (b).



(a) Fractional error of dwarf stars including binaries

(b) kiel diagram showing where the removed stars exist

Figure 3. (a) Residual plot showing dwarf stars over Gaia distance. Red line indicates the 0.13 cutoff threshold (b) Kiel diagram, showing where binary stars live.

method removed 134,092 stars from the training and test sets, leaving a final training set size of 382,673 stars (52% of DR17) and a test set size of 95,668 stars (13% of DR17). We note that, in the full data sample, the model is applied to the binary stars and those will still have systematic errors. However, the systematic errors mostly affect dwarfs, for which Gaia distances are likely to be more reliable than our model. For most star systems with a giant star, the difference in luminosity is much greater than between dwarf companions, so the contamination of the companion star around a giant is less significant.

4.2. NMSU Isochrone-based Distances

To compare different techniques and also to document some previous work, we present a separate set of distances determined using observed stellar parameters and isochrones, hereafter referred to as NMSU distances. Distances derived using this technique have been distributed in a previous SDSS Value Added Catalog (VAC)[6] and were used in Weinberg et al. (2019). For these distances, we use PARSEC isochrones (Bressan et al. 2012) with a Kroupa (2001) IMF to identify the expected relative number of stars as a function of effective temperature, surface gravity, metallicity, and absolute magnitude, under the assumption of a constant star formation history. We multiply this by the likelihood of the observed stellar parameters for each star, assuming normally distributed uncertainties given by the DR17 parameter uncertainties, and marginalizing to get the probability distribution function of absolute magnitude. This is then combined with our extinction estimates and apparent magnitudes to provide a probability distribution function in distance. In the subsequent discussion, we use the median of the distance probability distribution function to compare with our neural network distances.

4.3. Verification

Figure 4(a) shows the median absolute deviation (MAD; solid blue line) converted to standard deviation and the percentiles of the absolute deviation range (dashed blue lines) of our distances compared to the Gaia distances, as a function of surface gravity. This difference is largest at lower surface gravities, which we attribute partly to the poorer coverage in

---

[6] https://www.sdss.org/dr14/data_access/value-added-catalogs/?vac_id=apogee-dr14-based-distance-estimations

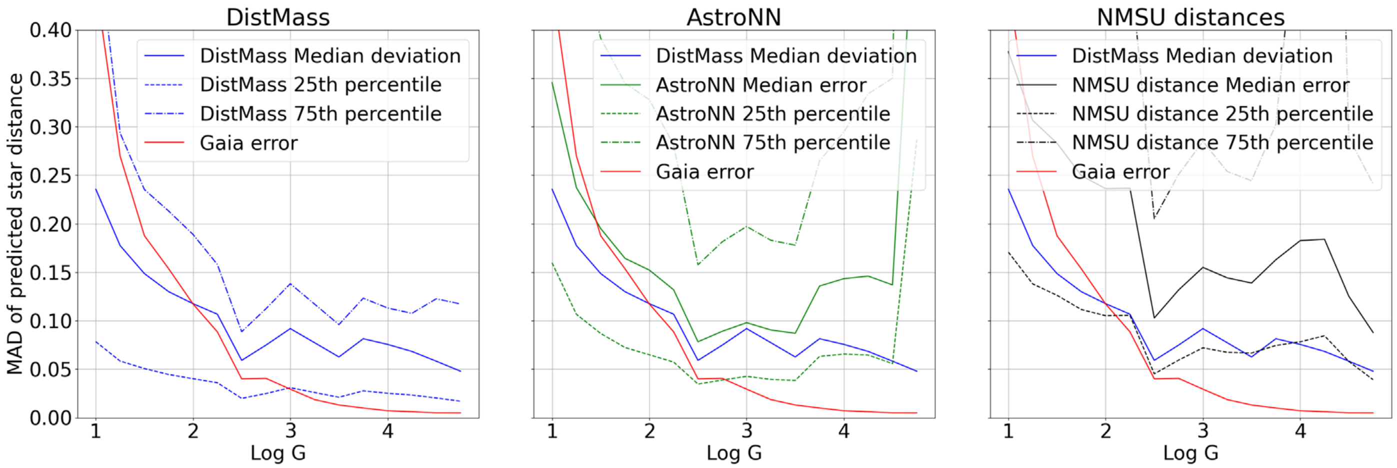


Figure 4. Blue line in all plots shows the MAD of test stars vs. surface gravity for our model. Red shows the Gaia parallax uncertainty. Figure (a) shows just our model compared to the Gaia error, with the dashed lines showing the quartile of the absolute deviation. Figure (b) shows AstroNN's (Leung & Bovy 2019) MAD and quartile ranges in green compared to Gaia and our model. Figure (c) shows the Bayesian NMSU distance MAD in black. The deviations reported here are the listed stellar distance from each method compared against the Gaia distances.

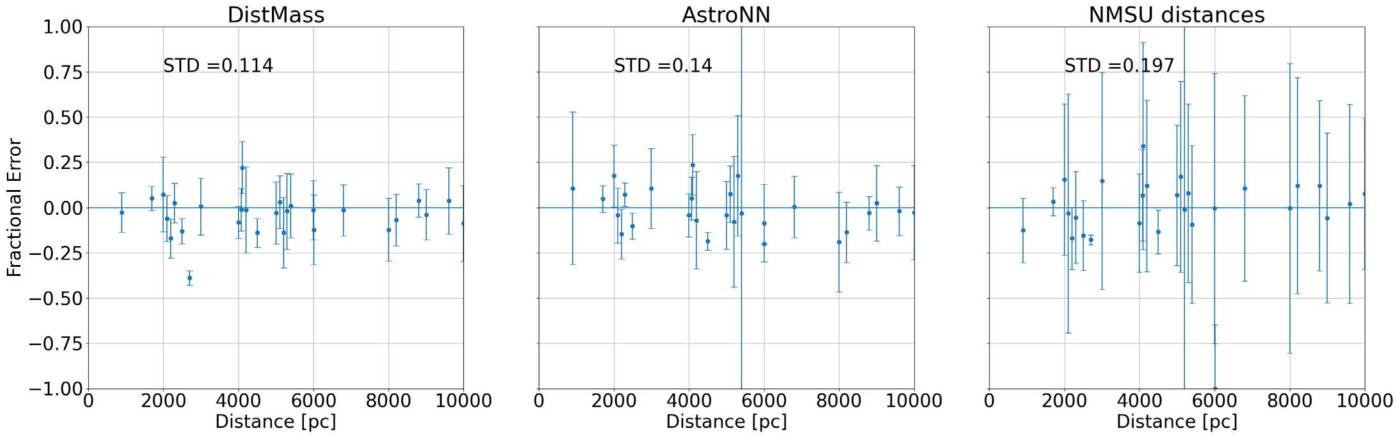


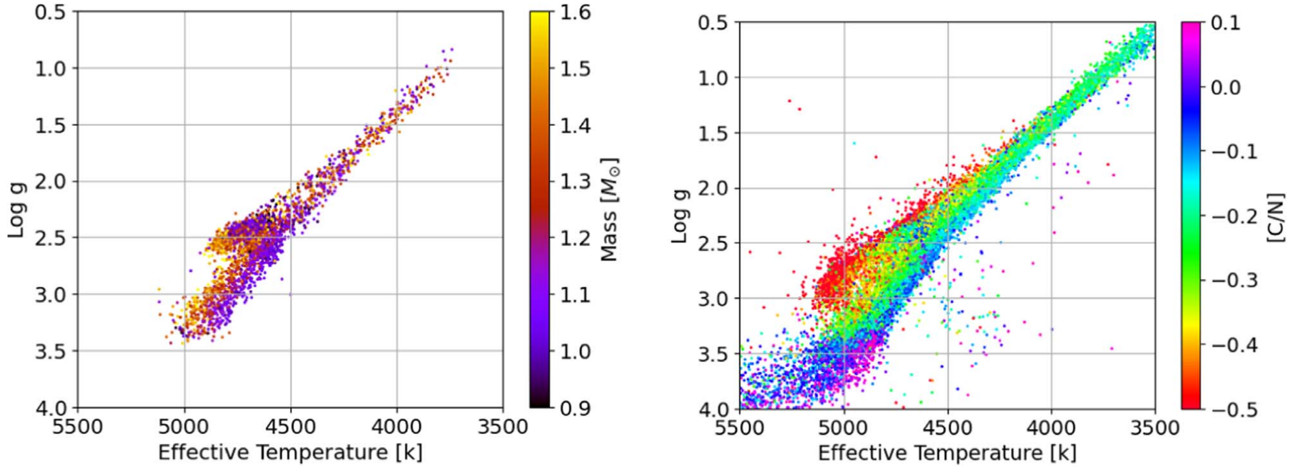
Figure 5. Model fractional error vs. cluster distance for cluster stars. Blue points show the average predicted distance to each cluster. The error bars show the standard deviation of the scatter in distance predictions for stars in each cluster around the mean distance. Scatter of the cluster distance averages is printed for each distance method. Panel (a) shows our model results, panel (b) shows AstroNN (Leung & Bovy 2019), and panel (c) shows NMSU distances. STD of the cluster distance average (orange points) is printed for each plot.

this region of our training set, but also to the larger uncertainties in the Gaia distances to these stars, which tend to be at larger distances. At higher surface gravities, the model achieved a median deviation of less than 10%; of course, for these stars, which are generally close, the Gaia distances are significantly more accurate.

Figures 4(b) and (c) show the same comparisons for AstroNN (Leung & Bovy 2019) and NMSU distances. AstroNN recommends using their distances that are weighted with Gaia distances, but here we compare their unweighted distance predictions because otherwise we would be comparing to Gaia for nearby stars. Figures 4(b) and (c) shows that our neural network (in blue) tends to have more accurate distance predictions than AstroNN and NMSU distances across all log *g* values. This may be due to AstroNN being influenced by noise in the full spectra, thus reducing its accuracy, while the NMSU distances are subject to uncertainty in the isochrones and the chosen IMF.

Additional validation comes from distances to cluster stars, both from the scatter in distances to cluster members and from comparison to literature distances. Figure 5(a) shows the performance of the neural network on stars in the clusters given in Table 1. The predicted distances for each star in each cluster were averaged to give estimated cluster distances shown with the blue points in Figure 5(a). The error bars show the standard deviation of the scatter in distance predictions for stars in each cluster around the mean distance. The median fractional error across the cluster mean distances is −0.06 and the standard deviation of the fractional error in cluster mean distances is 0.114. Model performance was also judged by examining the median scatter of predicted star distances around the true cluster distance for each cluster; we find a median standard deviation of 0.14 in fractional distance.

Figures 5(b) and (c) show the same comparison for AstroNN and NMSU distances. The median fractional errors of the mean cluster distances are closer to zero than in our model, with 0.038 for AstroNN and −0.042 for NMSU, but the standard deviations of the mean cluster distances are higher at 0.14 and 0.197, respectively. The median standard deviations within clusters for the comparison models are 0.37 and 0.4 for AstroNN and NMSU, respectively.



(a)Asteroseismic masses, $[Fe/H] = 0 \pm 0.025$ (b) [C/N] ratio, $[Fe/H] = 0 \pm 0.025$

Figure 6. (a) Kiel diagram of APOKASC stars <1.6 $M_\odot$ with $-0.05 < [\mathrm{Fe/H}] < 0.05$ and color coded by asteroseismic mass. (b) Kiel diagram showing the subgiant and red-giant branch for stars with $-0.025 < [\mathrm{Fe/H}] < 0.025$ color coded by [C/N].

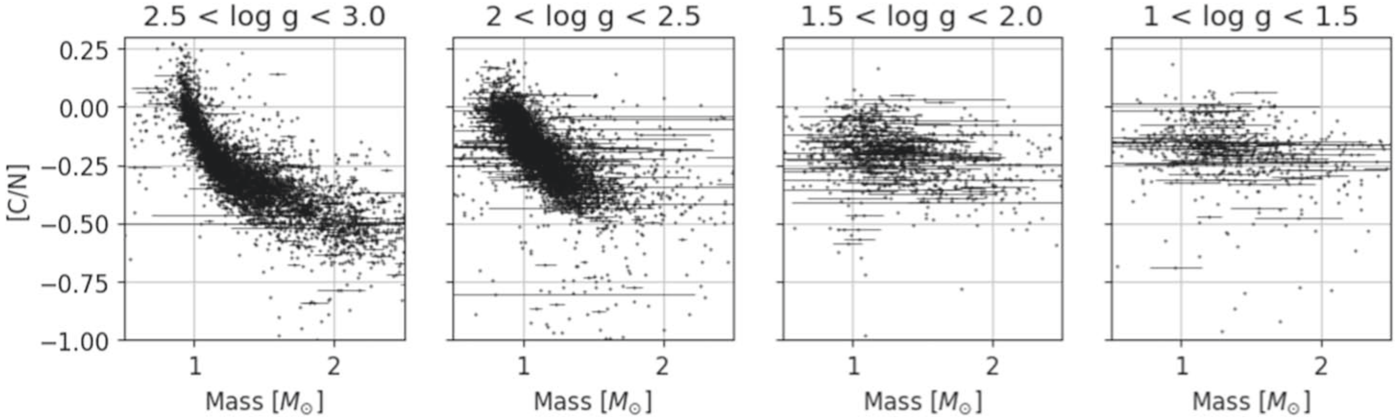


Figure 7. [C/N]–mass relation binned by log g. The masses shown here are the corrected Mosser masses from APOKASC. Each group of 10 points has an error bar for mass.

Uncertainties for our distance predictions come from a quadratic fit of the median fractional error versus log g. We do not have any inclusion of the input uncertainties. We recommend using the distances from our model for stars with larger Gaia parallax uncertainty, but for stars with small Gaia parallax uncertainty, the Gaia distances are preferred. Our catalog contains a column with our model distances weighted with the Gaia distances by the quoted uncertainties in each measurement.

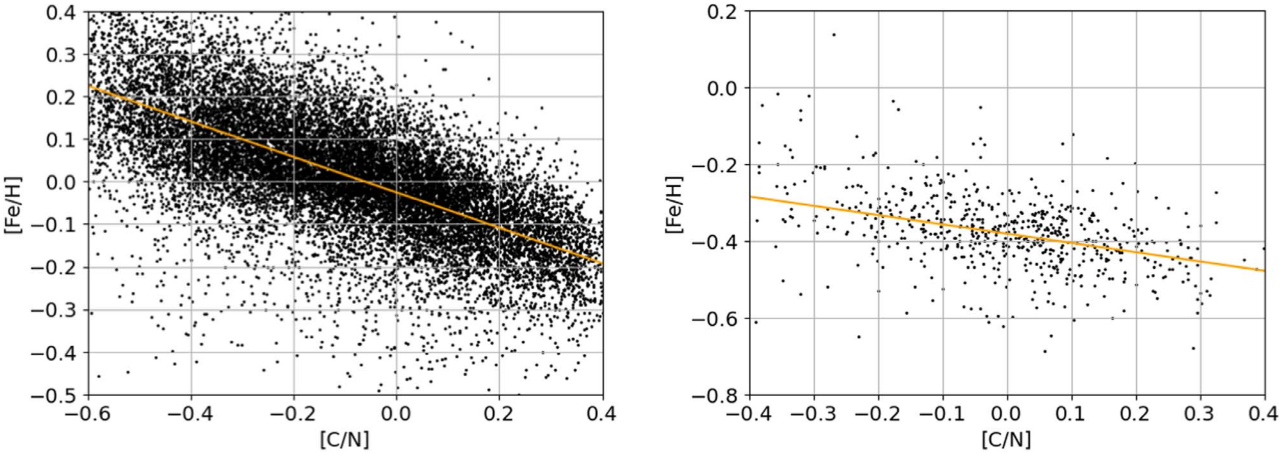
5. Masses

A second goal of this paper is to develop a model that estimates the mass (and age) of stars using the stellar parameters from APOGEE DR17.

Figure 6(a) shows a Kiel diagram for stars from the APOKASC catalog (M. Pinsonneault et al. 2024, in preparation) with solar [Fe/H] and color coded by mass, demonstrating that, especially along the upper giant branch, stars of different masses cannot be distinguished by Teff and log g alone. Fortunately, the [C/N] ratio can be used to separate stars of different masses (Martig et al. 2016). As a star evolves onto the red-giant branch, the first dredge-up results in a drop of the [C/N] ratio, where higher-mass stars have a larger change. Figure 6(b) shows a Kiel diagram of solar [Fe/H] stars from APOGEE DR17 color coded by [C/N], where the dredge-up mixing process is apparent near log $g = 3.5$.

Mass/age determination for evolved stars using [C/N] requires a correlation between the measured [C/N] and stellar mass across parameter space. Figure 7 shows the relation between [C/N] and APOKASC mass along with some representative error bars binned by log g. At higher log g, the [C/N]–mass relation is tight and defined; however, at low log g, the relation is less definitive. There is little physical reason for the [C/N]—mass relation to change at low log g for most stars, although extra mixing may need to be considered for metal-poor stars ($[\mathrm{Fe/H}] < -0.5$) (Salaris & Cassisi 2008; Shetrone et al. 2019). The error bars shown in Figure 7 grow larger at lower log g values, with an average mass uncertainty reaching up to $0.3 M_\odot$. Therefore, the explanation for the less clearly defined [C/N]—mass relation at lower log g is the increased uncertainty in the APOKASC masses.

Figure 6(b) shows a range in the pre-first dredge-up [C/N] along the subgiant branch stars below log g 3.3. The range in pre-first dredge-up [C/N] is a result of variations in the birth [C/N] of these stars. The first dredge-up process does not completely erase the original birth [C/N], therefore these variations in the birth [C/N] will influence the observed [C/N]



(a) Sub-giant stars from APOGEE DR17 $4 < log\,g < 3.5$ and $5500K < T_{eff} < 4750K$

(b) RGB stars from APOKASC $3 < log\,g < 2.5$ and $1.5 < M_{\odot} < 1.75$

Figure 8. (a) [C/N] as a function of [Fe/H] for subgiant stars. Slope shows the [Fe/H]–birth [C/N] relation. (b) [C/N] as a function of [Fe/H] for RGB stars. Star selection was limited to a narrow mass range in order to highlight the [C/N]–[Fe/H] relation.

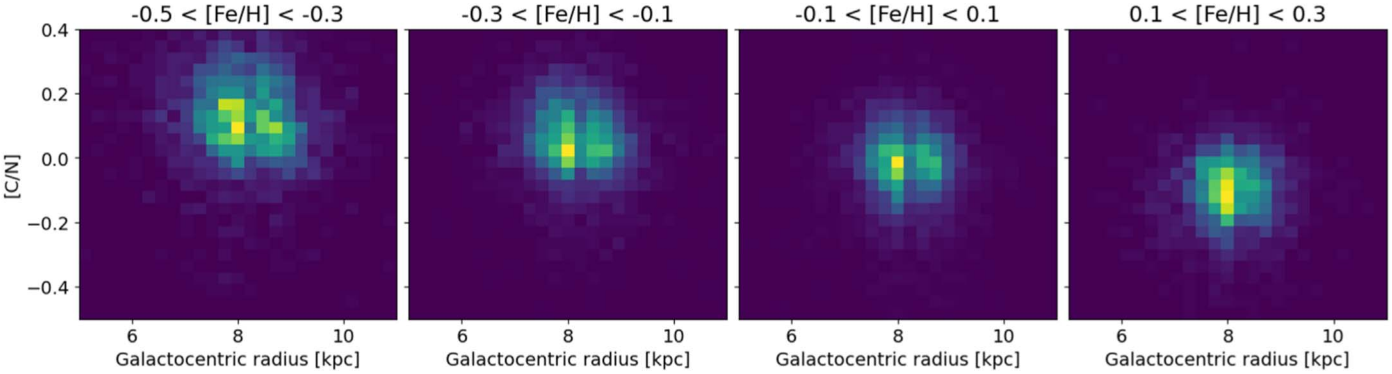


Figure 9. Histograms of [C/N] vs. Galactocentric radius for pre-dredge-up subgiants binned by [Fe/H]. The location of the distribution is different for each [Fe/H] bin, due to the [C/N]–[Fe/H] relation. There is no apparent trend in the [C/N]–[Fe/H] relation over the examined range. However, the range we can examine the [C/N] variation is limited to 2 kpc by the lower luminosity of subgiants.

in red-giant branch stars and consequently can affect stellar mass/age predictions based on [C/N] abundances if the birth abundances are not taken into account (Martig et al. 2016; Vincenzo et al. 2021; J. Roberts et al. 2024, in preparation). Vincenzo et al. (2021) and J. Roberts et al. (2024, in preparation) observed that birth [C/N] correlates with [Fe/H] and [Mg/Fe]. Figure 8(a) shows the [C/N]–[Fe/H] relation for subgiant stars, which means the measurement is of birth [C/N]. Figure 8(b) also shows [C/N] and [Fe/H] but for APOKASC RGB stars in a narrow mass bin, which indicates that the [Fe/H]–[C/N] relation is important after first dredge-up. This clearly shows that including [Fe/H] in chemical mass/age is important. While J. Roberts et al. (2024, in preparation) also suggest [Mg/Fe] correlates with birth [C/N], the additional use of [Mg/Fe] would be problematic for our mass/age estimations and future use cases of our estimations. [Mg/Fe] has been shown to have a correlation with age (da Silva et al. 2012; Haywood et al. 2013; Bensby et al. 2014), but the [Mg/Fe]–age correlation has been found to vary across the Galactic disk (Feuillet et al. 2018; Silva Aguirre et al. 2018; Katz et al. 2021; Viscasillas Vázquez et al. 2022; Imig et al. 2023). The spatial variation in the [Mg/Fe]–age correlation could cause any chemical stellar mass/age prediction model to present spatial systematics if [Mg/Fe] was included in the model. We also intend our mass/age predictions to be useful for mapping the [Mg/Fe]–age correlation in order to better constrain Galactic chemical evolution models. Including [Mg/Fe] in the prediction model would tie the [Mg/Fe] abundances to our mass/age estimations, and therefore our estimations could not be used as an independent test of [Mg/Fe] chemical evolution. As a result of this, we exclude [Mg/Fe] and include [Fe/H] for our mass/age estimation models.

In order to leverage the [C/N]–mass relationship and estimate the masses of evolved stars, we have built a second neural network model that uses the effective temperature, surface gravity, metallicity, carbon, and nitrogen abundances as input parameters. The training data is constructed from asterosiesmic masses in the APOKASC catalog, which uses data from Kepler observations as described in Section 1.

5.1. Spatial Variations in the [C/N]–Mass Relation

Because our goal is to reliably predict stellar masses across a large volume of the Galaxy, we require the training stars to be representative of the entire sample we predict ages for. Our training data are limited to stars in the original Kepler field, which does not sample a wide range of Galactocentric distance, although we apply the age model to all of the stars in DR17,

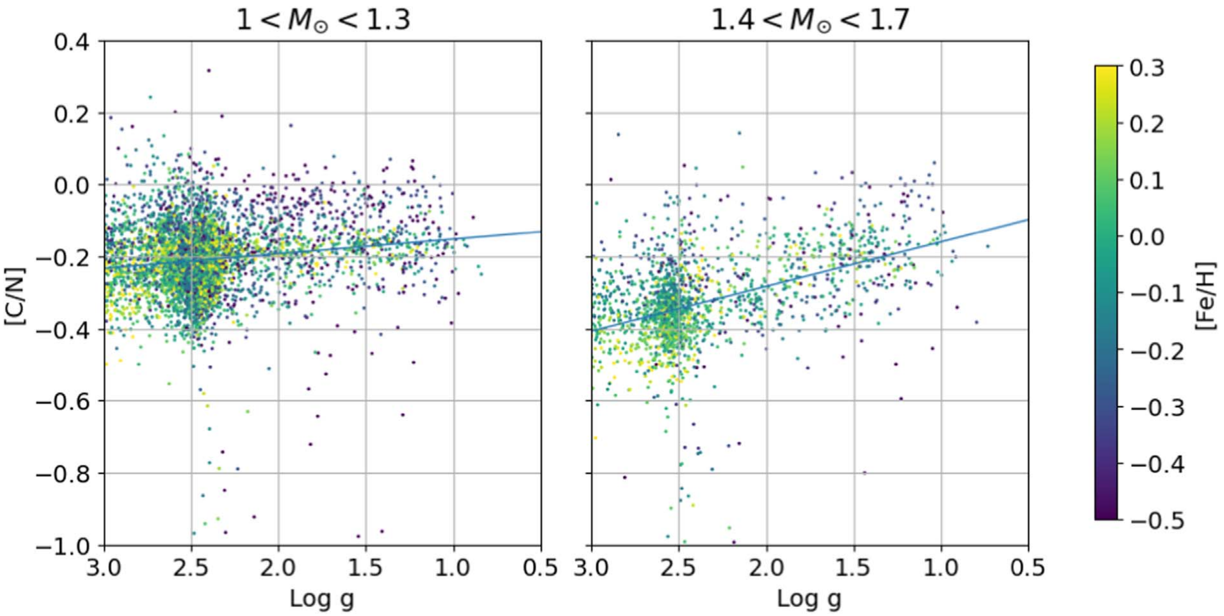


Figure 10. [C/N] vs. log g binned by asteroseismic mass. Best-fit line shows the positive trend in [C/N].

which does span a large region of the Galaxy. This requirement is part of the reason we excluded [Mg/Fe] from the mass/age estimation model, because the [Mg/Fe]–age correlation varies across different regions of the Galaxy as described in Section 5. We make the assumption that the training stars are representative of the [C/N]–[Fe/H]–mass relation throughout the Galaxy. However, if the [C/N]–[Fe/H]–mass relation also varies as a function of Galactocentric radii, then our mass/age estimations will have systematic errors that depend on Galactic location.

We test our assumption by searching for spatial variations of the [C/N]–[Fe/H] relation in subgiant stars. If there is little variation, then we can be more confident about our assumption that the Kepler field is representative of the Galaxy. This test can be applied beyond the Kepler field, because it does not rely on the asteroseismology from Kepler. Martig et al. (2016) did this test with DR12 data, and we have repeated this test using the more extensive DR17 data. Figure 9 shows a histogram of [C/N] versus Galactocentric radius for all DR17 pre-dredge-up stars. We define pre-dredge-up stars to have $4 < log\ g < 3.5$ and $5500\ \mathrm{K} < T_{\mathrm{eff}} < 4750\ \mathrm{K}$. Figure 9 does not indicate any spatial trends in [C/N], which is consistent with the results of Martig et al. (2016). Unfortunately, even with the more extensive DR17 data, subgiants are only included within 2 kpc of the Sun, so our results at larger distances are still subject to potential systematics if the [Fe/H]-[C/N] relation is different at those distances.

5.2. [C/N] Systematic

At low log g, the reported [C/N] from DR17 rises for the stars of the same mass, as shown in Figure 10. We expect stars to either keep a constant [C/N] after first dredge-up or a drop in [C/N] from extra mixing. However, Figure 10 appears to show stars undergoing "unmixing" with a higher [C/N] as they evolve up the RBG, which is nonphysical as discussed by Fraser et al. (2022). This appears to be a systematic error in the reported [C/N] that varies with log g. However, in principle, our model trained with this systematic and therefore is unaffected by it. However, this highlights that our learned [C/N]–mass relation is tied to the DR17 data, and comparisons with models that use non-DR17 data must take this into account.

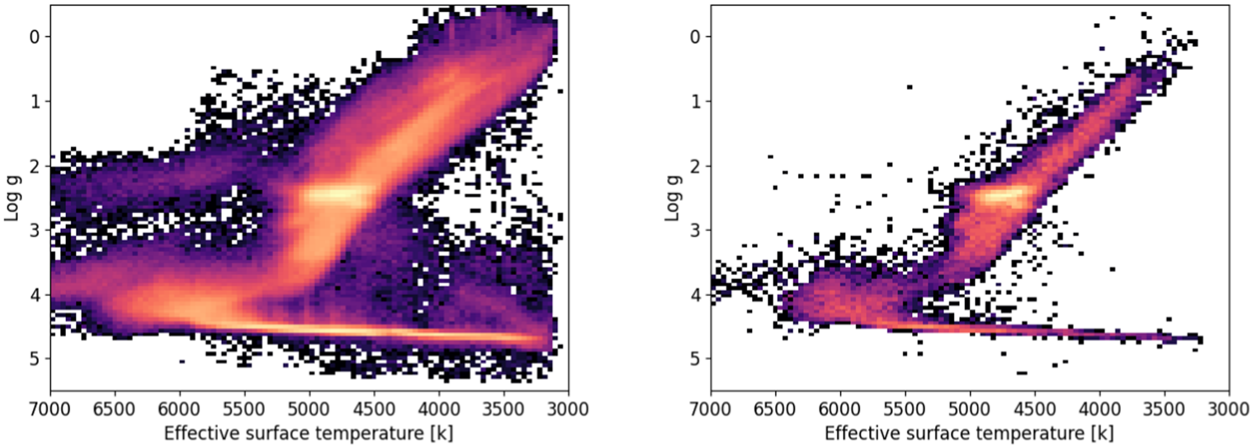
5.3. Training Space Constraints

The model should not extrapolate too far beyond the space covered by training, otherwise the mass predictions will be impacted by poor model extrapolations. We also want to remove stars with high mass uncertainties from the training set, otherwise they could degrade the model's performance. To achieve this, we chose to cut stars with an APOKASC uncertainty of >10%, resulting in a training set of 13,458 stars. These cuts result in significant gaps in parameter space, which are shown in Figure 11. In particular, training data are limited at log $g < 1.5$ and $[\mathrm{Fe/H}] < -0.5$, implying that the model will begin to rely on extrapolation for stars at lower log g and more metal-poor stars.

5.4. APOKASC Radii Corrections

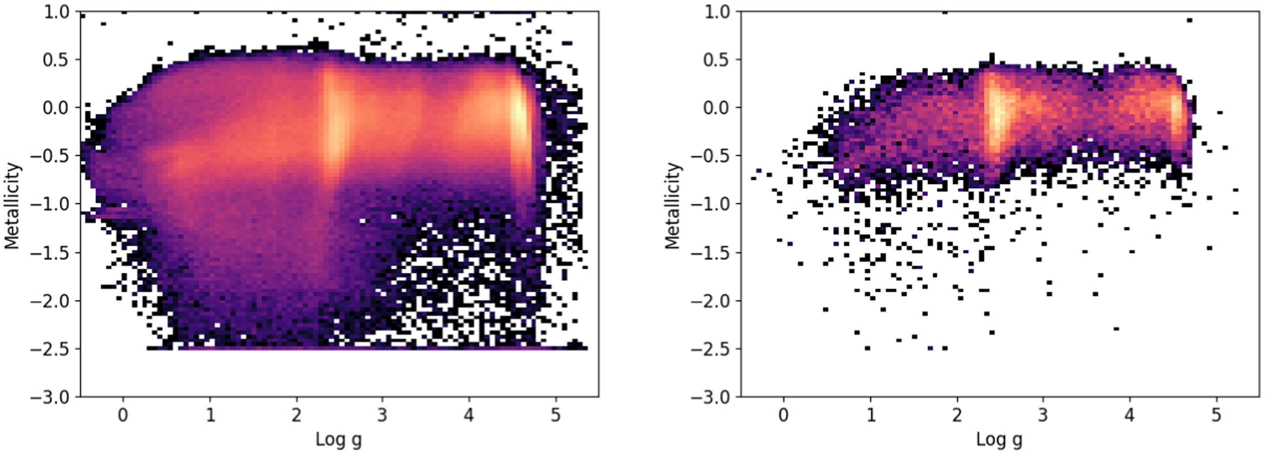
The training set for the mass model uses asteroseismic masses from APOKASC 3 (M. Pinsonneault et al. 2024, in preparation). APOKASC 3 contains contains stellar mass estimations using three different seismology results from Mosser et al. (2013), Sharma et al. (2016), and White et al. (2011). Each of these is included in APOKASC 3, with both the raw seismology masses as well as masses corrected based on a comparison with radii determined by Gaia (Zinn et al. 2019).

To attempt to assess whether there is a preference for one of these data set, we consider mass distributions for stars of different log. We would expect that, in a given region of the Galaxy, the distributions of masses for low and high log g stars to be equivalent. In Figure 12, we show normalized histograms of the masses present in each of the six seismology results at two different distance bins. Different rows present each of the three seismology results, and the columns show the Gaia uncorrected masses (left) and corrected masses (right). The



(a) Full data set kiel diagram

(b) Post mass uncertainty cut kiel diagram



(c) Full data set metallicity

(d) Post mass uncertainty cut metallicity

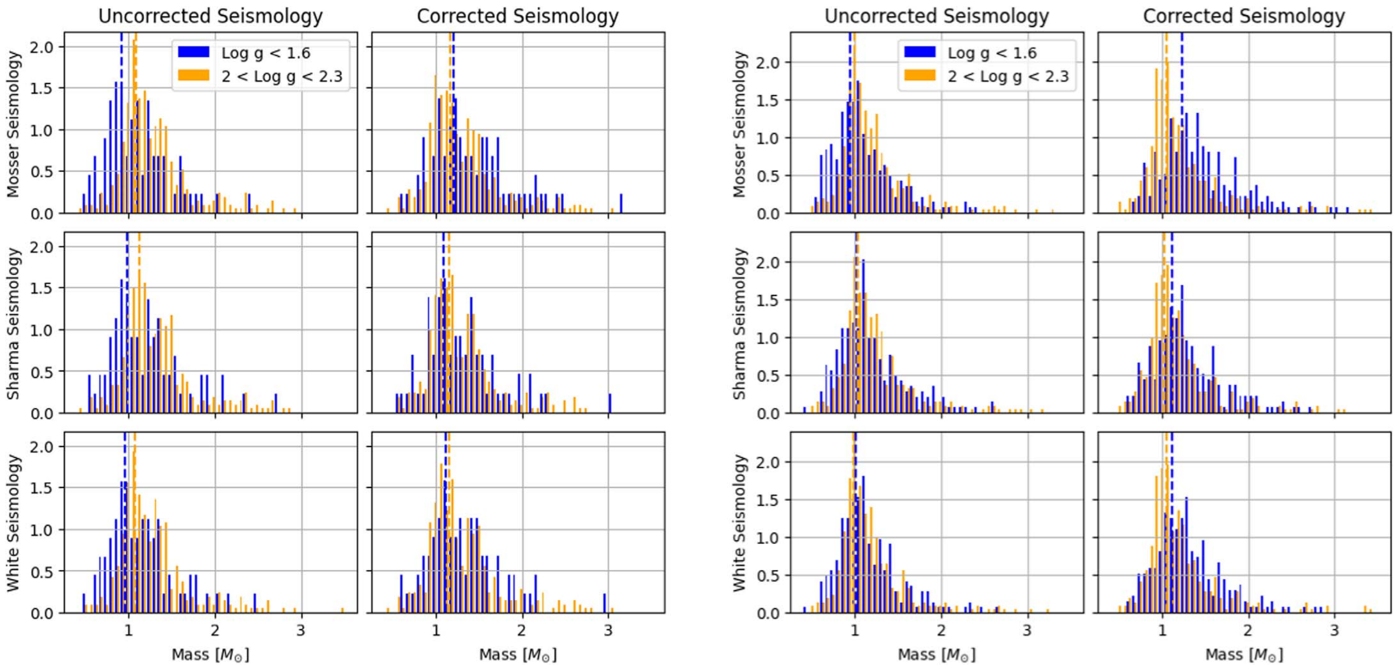
Figure 11. Density plots showing the kiel and metallicity diagrams and for the full data set ((a) and (c)) and for the training set after the 10% mass uncertainty cutoff ((b) and (d)). Brighter and colored regions contain a higher density of stars.

histograms are split into a low log *g* bin (blue) and high log *g* bin (orange), and a vertical line shows the mean mass for the log *g* bins. In every seismology result, the Gaia corrections shift the low log *g* stars to higher estimated masses. High log *g* stars have very small corrections to their masses, as expected, because the higher log *g* stars have very small radii (Zinn et al. 2019). In most of the results, the shifts are fairly mild. In the closer distance bin, the corrections bring the mass distributions into alignment; however, in the greater distance bin, the corrections appear to push the distributions out of alignment. In the greater distance bin, the corrected Mosser et al. (2013) results have the most unequal distributions. So it is not clear if using the corrected seismology results is superior. However, as discussed below, models trained with the corrected results lead to a paucity of low-mass stars across the Galaxy.

5.5. Implementation

In order to better understand the effects of the small differences between the seismologies and corrections, we used each as the training set for six independent mass prediction models. Figure 13 shows the derived [C/N]–mass relations for each model, color coded by [Fe/H] and binned by log *g*. Each column is a different model trained on the six different seismologies/corrections, and each row is a different log *g* bin. At high log *g*, there are few differences between the model predictions. However, at low log *g*, the corrections for each seismology appear to shift the [C/N]–mass relation to higher masses. The shift is large enough that there is an absence of low-mass stars in the model predictions for the models trained on the Gaia corrected masses. The absence of low-mass stars translates to an absence of old stars in the resulting age predictions, which is not an expected feature of the Galaxy nor is it apparent in the training data from Figure 12, which does include low-mass stars. The likely cause of this is the Gaia corrections shifting low log *g* stars to higher masses and the models learning a shifted [C/N]–mass relation.

As discussed in Section 5.4, we expect that, in a given region of the Galaxy, the distributions of masses for low and high log *g* stars will be equivalent. Figure 14 shows histograms of the predicted mass distributions from our six models split into a low log *g* bin (blue) and a high log *g* bin (orange). The mass distributions from the corrected and uncorrected Mosser models have medians that are offset from each other. The



(a) $1 < Distance[kpc] < 2$

(b) $2 < Distance[kpc] < 3$

Figure 12. Mass distribution histograms for the six mass data sets in two distance bins. Left columns show uncorrected masses, and the right columns show the corrected masses. The histograms are split into a low log g bin (blue) and high log g bin (orange).

distributions from the Sharma and White models are equivalent. The shift to higher masses is also apparent for the corrected mass models. In general, the Sharma and White corrected and uncorrected models treat the low log g stars similarly to high log g stars.

Given that it is challenging to know which seismology result is the best one for model training, we elected to produce a model using each set of seismology results and corrections. This resulted in six sets of mass and age predictions in our catalog. However, we note that the mass distributions are more unequal for the Mosser models. We also note that the corrected mass models lack low-mass (old-age) stars in their predictions. For these reasons, Imig et al. (2023) used the masses and ages from our model trained on uncorrected Sharma masses.

For our reported uncertainties, we chose a flat uncertainty of 10% for all of our predicted masses. This value comes from the median absolute deviation between our predictions and the APOKASC masses for stars at log $g \approx 2.5$. At low log g, the median absolute deviation becomes much larger, but this likely results from the fact that the APOKASC uncertainties grow larger at low log g as well. We make the assumption that the uncertainties are similar between high and low log g. This is assumption is based on the similar mass distribution medians shown in Figure 14. It should be noted, however, that this assumption is weaker for the Mosser models, because their distributions are not well matched.

For the discussion of ages in the following section, we use the ages from the model using the corrected Mosser seismology, because that is what M. Pinsonneault et al. (2024, in preparation) chose to use for their analysis and this is also the seismology used for training the Mackereth et al. (2019) stellar ages. However, our final catalog contains masses and derived ages for all six sets of seismology.

6. Ages

6.1. Age Determination from Mass

For evolved stars, there is a relation between the stellar mass and age. Using the predictions from the mass model, we provide age estimations for stars with log $g < 3.5$ using MIST isochrones (Choi et al. 2016). The relation between mass and age for evolved stars works because stars only occupy the giant branch for the final 1%–10% of their lives (Pols et al. 1998). The MIST isochrones were used to make a function with initial stellar mass and metallicity as inputs to produce an estimated age. The function assumes all giants of the same mass and [Fe/H] to have the same age. The function covers a mass range of 0.8–8 $M_\odot$ and a metallicity range of −0.5 to 0.5 [Fe/H]. Because the function uses the mass estimates from the mass model, the age estimation is tied to the performance of the mass model, so the uncertainties for our age estimation are determined by propagating the mass uncertainty through the mass-to-age function. The metallicities from ASPCAP and the predicted masses from the neural net were used for the age prediction. Due to the nature of the mass-to-age conversion, the age uncertainties are more sensitive to mass uncertainties at lower masses. Thus, the typical mass uncertainty of ±10% will result in an age uncertainty of [+0.3, −0.21] log(age) for a $1M_\odot$ star, while for a $2M_\odot$ star, the same mass uncertainty gives an age uncertainty of [0.067, −0.063] log(age).

6.2. Mass Loss

As stars ascend into the red-giant phase, they undergo mass loss. According to the Reimers (1977) prescription, mass loss is expected to be important primarily for the upper red-giant branch, due to low surface gravity. Metallicity is also a factor in driving mass loss, with higher-metallicity stars losing more mass via stellar winds (Vink et al. 2001). There also could be

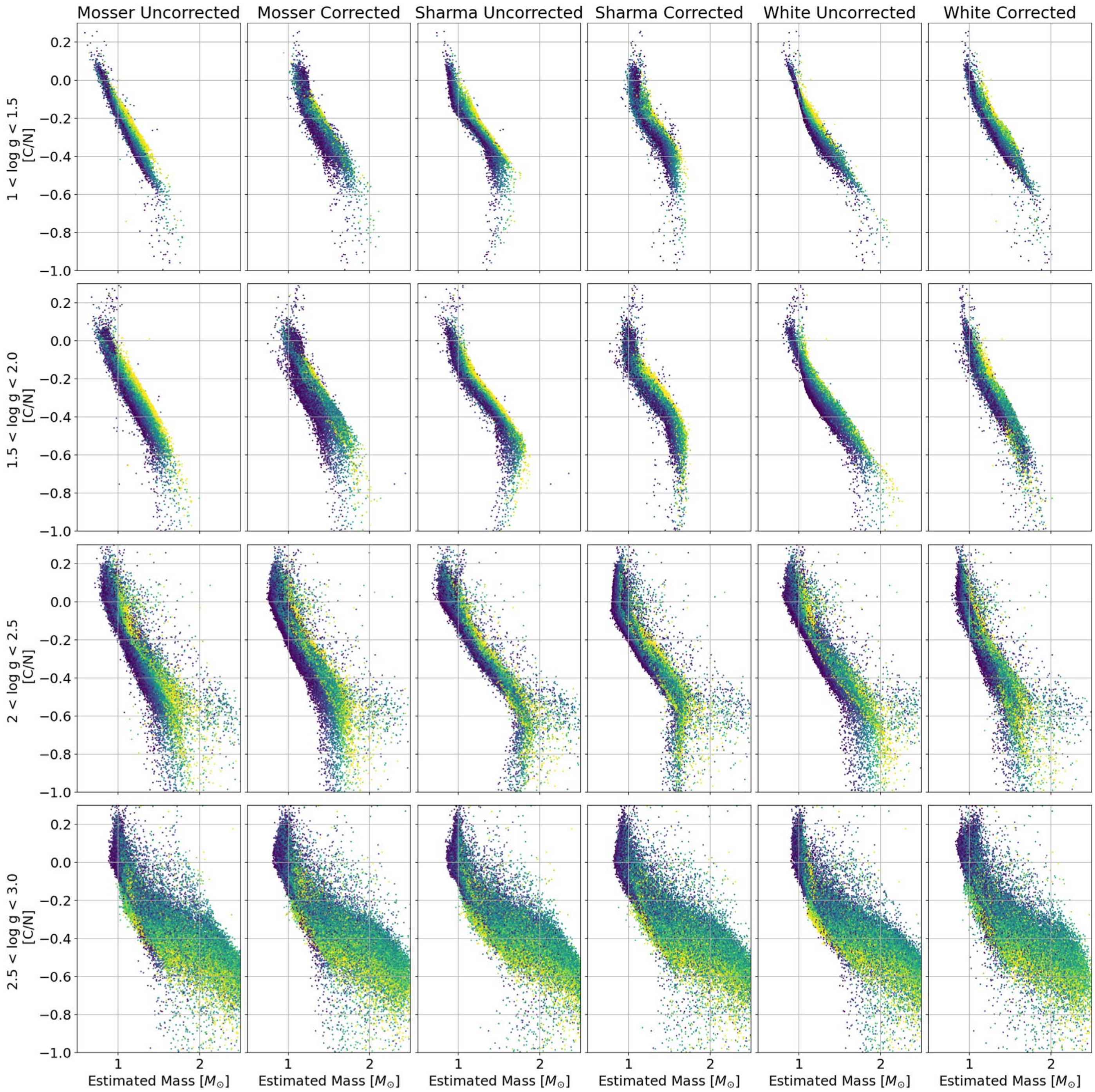


Figure 13. [C/N]–mass relation for the ML model predictions trained on all six seismologies/corrections. Color coded by [Fe/H] and binned in log g intervals of 0.5.

mass loss associated with the helium flash that stars under 2.25 solar masses go through (Catelan et al. 1996; Tailo et al. 2021); for our sample of stars, the most affected region is likely the red clump.

Because the asteroseismic masses used for training are current masses and our neural network is trained only on empirical data, the model does not have any knowledge of stellar mass loss. The mass predictions of the model are estimates of current star mass. If mass loss is significant, then we would expect stars with higher mass loss to have overpredicted ages.

However, Miglio et al. (2012) examined the mass loss for clusters N6819 and N6791 and show that little appreciable mass loss is observed in RGB stars, with the exception of high-luminosity stars at the tip of RGB, which would carry over to RC stars as well. It should be noted that these clusters span from solar to supersolar metallicity. Both Miglio et al. (2012) and Handberg et al. (2017) examined N6819 mass loss with asteroseismic data and constrained the possible mass loss to $<0.03 M_\odot$, which Handberg et al. (2017) noted as almost insensitive to systematics. While the expected mass loss depends on the stellar parameters, these studies suggest that

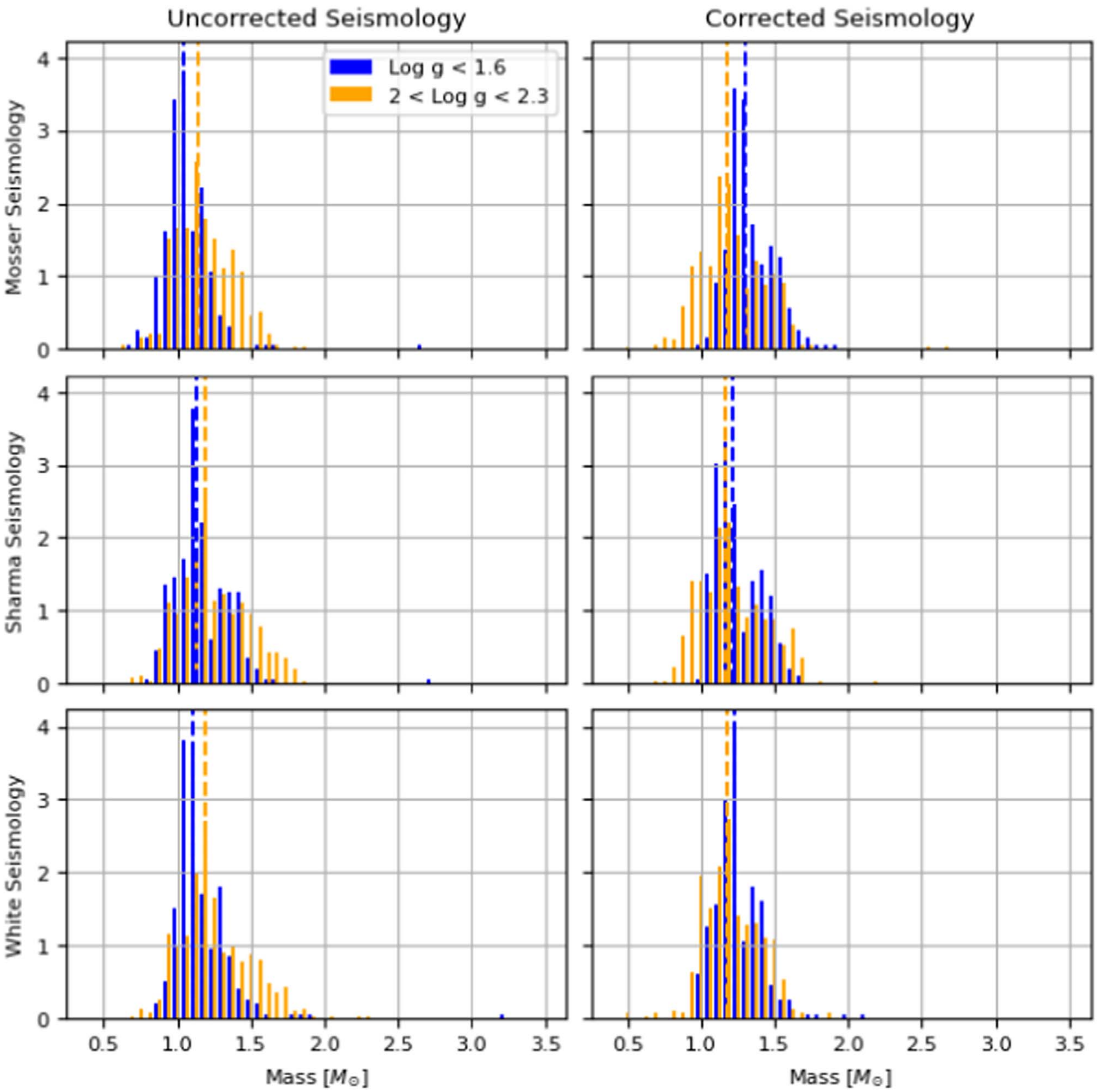


Figure 14. Histograms showing the distribution of masses predicted by the six models. The left column shows uncorrected masses, and the right column shows the corrected masses. The histograms are split into a low log *g* bin (blue) and high log *g* bin (orange).

it is unlikely to be higher than the constraint of $0.03M_\odot$. For a 1 solar mass star that loses $0.03M_\odot$, the age prediction will be overpredicted by 11%; for a 2 solar mass star, the overprediction would be 3%.

6.3. Parameter to Age Relations

Figure 15 shows the derived ages as a function of the [C/N] in three narrow [Fe/H] bins color coded by log *g*. All three show, in dashed red, the [C/N] to age relation from Spoo et al. (2022), which was determined by calibrating a fit with the ages of open clusters versus the average [C/N] ratio of a cluster. However, that calibration does not include any metallicity dependence, which is important to the [C/N]–mass/age relation, as discussed in Section 5, and is apparent across plots in Figure 15. The comparison to Spoo et al. (2022) serves both as another test of our age predictions and as a demonstration of the importance of including metallicity in any age/mass predictions involving [C/N].

6.4. Validation

The AstroNN catalog contains stellar ages calculated by Mackereth et al. (2019), which are derived using a Bayesian convolutional neural network trained on the full APOGEE spectra and asteroseismic ages, as opposed to our method, which uses the derived stellar parameters and abundances. The comparison to Mackereth et al. (2019) is not an independent validation of our ages, because they use the same asteroseismic training set, but they do serve as a quick sanity check for our ages. Figure 16 compares our results with the AstroNN results, using only stars within the parameter space covered by the

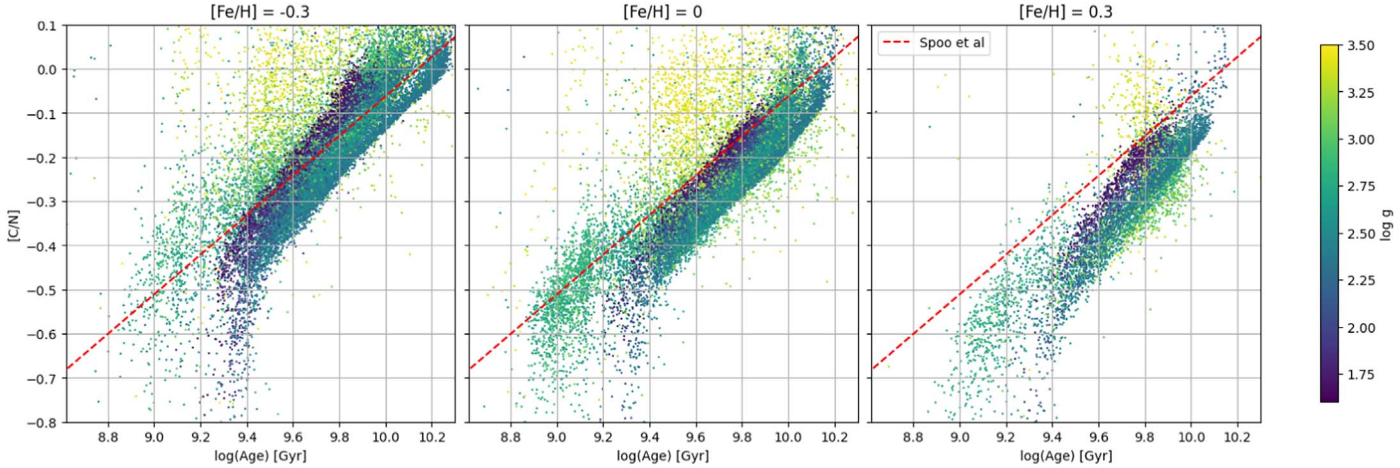


Figure 15. [C/N] vs. estimated age binned by [Fe/H] and color coded by log g. Ages are from the model trained on the corrected Mosser seismology. Red dashed line shows the [C/N] age calibration from Spoo et al. (2022).

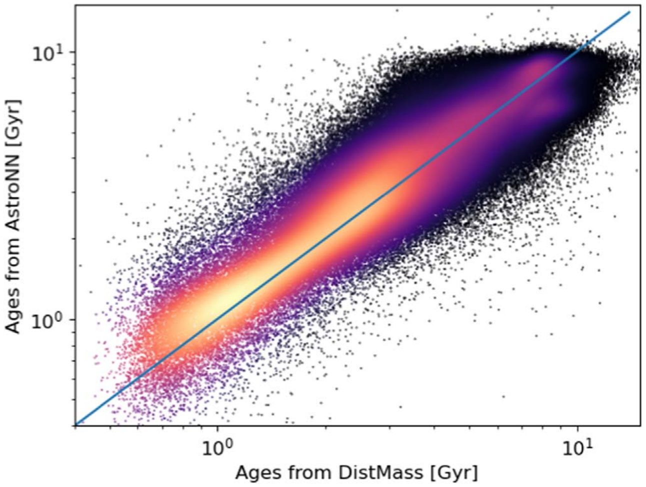


Figure 16. DistMass ages vs. Mackereth et al. (2019) ages for stars within the bounded training region for the mass model.

training set of the mass model shown in Figure 11. The figure shows that our age predictions are generally in agreement with the ages from Mackereth et al. (2019).

Star clusters are composed of stars that all should have largely the same age, so they can act as a tool to indicate prediction reliability. Figure 17 shows age predictions from our model and Mackereth et al. (2019) for a selection of stellar clusters along with the scatter of age predictions for each cluster by the models. The star cluster comparisons in Figure 17 show that our model has good agreement with Mackereth et al. (2019), and it presents fairly tight age distributions within each cluster.

Some of the clusters shown in Figure 17 have inconsistencies with the literature age. NGC 6388 is predicted by both our ages and Mackereth et al. (2019) to be younger than the listed age we used for N6388, with an age of 3.65 Gyr. The literature age for NGC 6388 is from Catelan et al. (2006), who note the age relative to the age of the cluster 47 TUC as $\Delta t \approx -0.5 \pm 1.6$ Gyr (NGC 6388 being younger). We used the age of 47 TUC from Brogaard et al. (2017) along with the information from Catelan et al. (2006) to calculate an age of 11.3 Gyr for NGC 6388. Given that N6388 is a globular cluster, the young age predicted by our model and Mackereth et al. (2019) is likely to be an underestimate, and the literature age is probably closer to reality. A plausible explanation for the discrepancy is there are multiple populations of stars in N6388. Mészáros et al. (2020) note an N-C anticorrelation in many globular clusters, including NGC 6388, that is likely due to pollution from an older population of FG stars. The result of the pollution is a lower birth [C/N] for a large population of stars, which would appear as the mass/age models underpredicting the ages of the stars.

Another cluster we examined, N7789, has a bimodal distribution of predicted ages for both our method and Mackereth et al. (2019). Figure 18 shows a Kiel diagram of stars in NGC 7789 color coded by DistMass age. We see a lower age prediction in the red clump, which means that the mass prediction is higher in that region. N7789 is the only cluster we examined that had this feature. Confirming with APOKASC asteroseismology is not an option, as there are few stars in N7789 with APOKASC measurements, although additional asteroseismology from K2 and TESS may help address this issue. Red clump stars do experience mass loss, as described in Section 6.2; however, mass loss would result in an age overprediction, while the predicted ages for red clump stars in N7789 are younger. Girardi et al. (2010) suggests there are peculiarities in the red clump of N7789, specifically that there may be a faint secondary red clump.

7. Conclusion

Simple neural networks were used to derive distances and masses for stars in SDSS Data Release 17. We also used the predicted masses to derive ages for select giant stars. Our distance model derived distances for 733901 unique spectra with good coverage of parameter space, with the exception of very low log g and metal-poor stars. These distance predictions have an uncertainty of 20% for stars at log $g < 2$ as determined from comparison with the Gaia training set, and uncertainties below 10% for stars with log $g > 2.5$. Comparisons of distance predictions to clusters indicates that the model has a median fractional error of −6% and a scatter of 11.4%.

The mass model was also applied to all the stars in the data set; however, the training parameter space is sparser, with a limited number of stars at log $g < 1.5$ and $[Fe/H] < -0.5$, which limits the set of stars the mass model provides valid

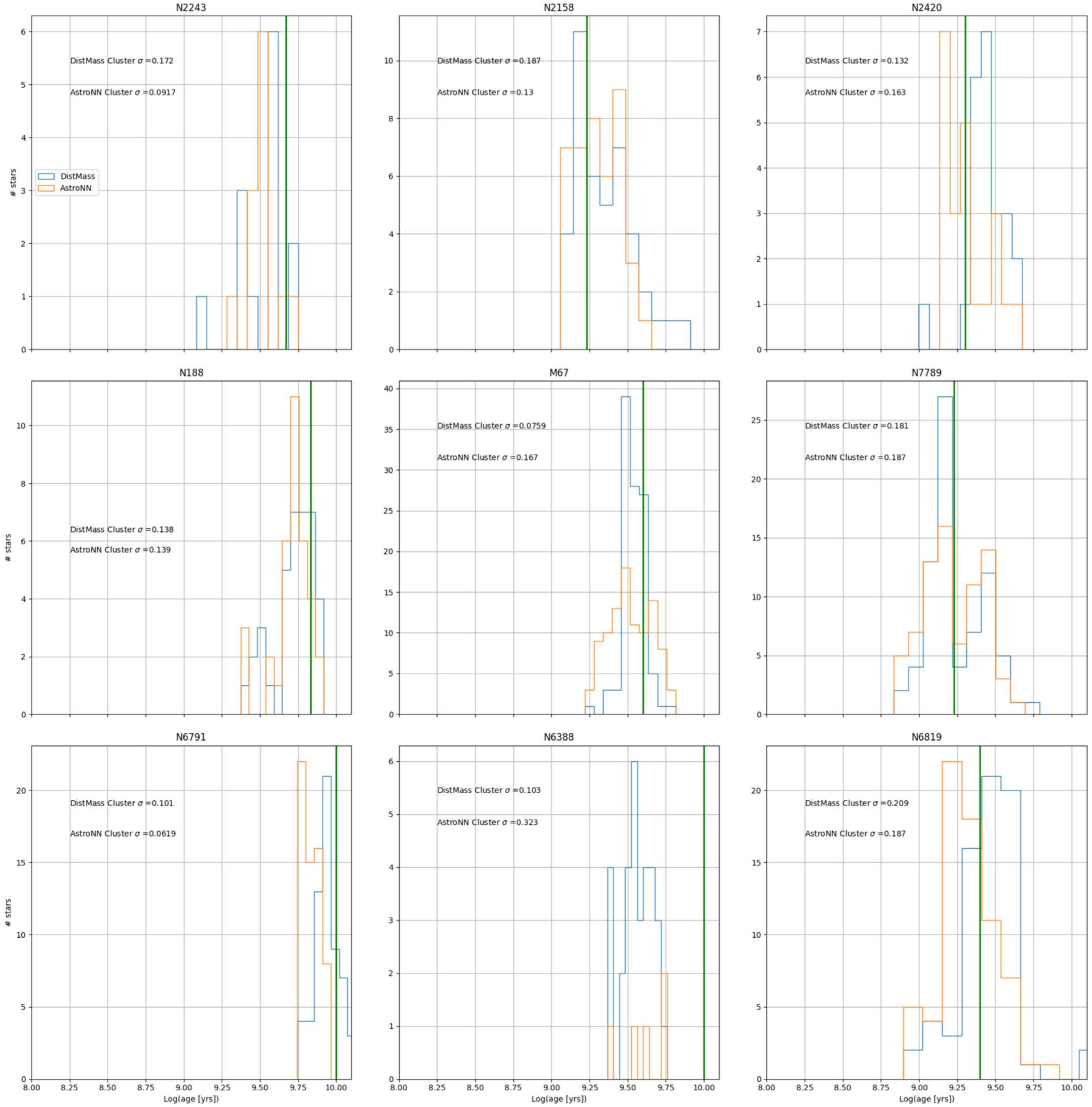


Figure 17. DistMass and Mackereth et al. (2019) age histograms for stars in clusters. The DistMass ages used here are from the corrected Mosser seismology.

predictions for. We derived masses for 506,817 unique spectra that are inside the limited coverage of the training parameter space, and we find that the mass model has median fractional errors of <0.12.

Stellar age prediction was performed using the mass model predictions and stellar isochrones. Our age prediction is only possible for giant stars because it relies on the assumption that a giant star is in the final 1-10% of its lifespan. Because it also relies on the mass model predictions, it is further limited to stars within the mass model training parameter space. In total, we have age predictions for 271,615 unique spectra. To examine the validity of the age predictions, we compared the ages of stars inside clusters and found that the average standard deviation of age in log(age) is <0.1. Some clusters had much higher scatter, or in the case of N7789, an unexplained bimodal distribution of ages. However, most clusters we examined showed self-consistent age predictions or had plausible known reasons for having high scatter. We also presented some initial examinations of the C/N—age relation using our data compared to work from Spoo et al. (2022).

All of the derived distances, masses, and ages are available in a VAC for SDSS Data Release 17[7] (Abdurro'uf et al. 2022). The catalog provides all the input parameters and training

---

[7] www.sdss4.org/dr17/data_access/value-added-catalogs/?vac_id=distmass:-distances,-masses,-and-ages-for-apogee-dr17

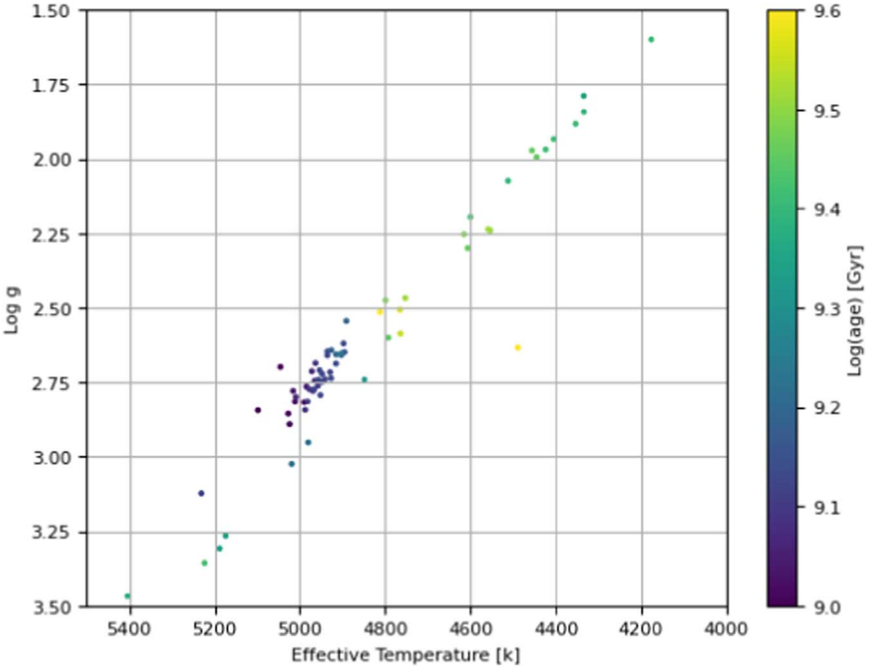


Figure 18. Kiel diagram showing stars in NGC 7789 color coded by DistMass. The DistMass ages used here are from the corrected Mosser seismology.

Table 2
DistMass Catalog Data Contents

Data	Header Name	Description	Source
APOGEE ID	APOGEE_ID	ID from APOGEE	SDSS DR17
APSTAR ID	APSTAR_ID	ID from APSTAR	SDSS DR17
Effective temperature	TEFF	Stellar effective temperature	SDSS DR17
Surface gravity	LOGG	Surface gravity in log space	SDSS DR17
Metallicity	M_H	Stellar metallicity	SDSS DR17
Carbon abundance	C_FE	Stellar carbon abundance	SDSS DR17
Nitrogen abundance	N_FE	Stellar nitrogen abundance	SDSS DR17
Gaia distance	GAIAEDR3_DIST	Distance from Gaia used for training	Gaia DR3
Extinction	EXTINCTION	Extinction magnitude used as described in Section 2	SDSS DR17
Apparent mag	MAG	Apparent K magnitude from APOGEE	SDSS DR17
Absolute mag	ABS_MAG	The predicted absolute magnitude from our model	DistMass
Absolute mag error	ABS_MAG_ERR	Estimated error in our magnitude prediction	DistMass
DistMass distance	DISTANCE	Predicted distance from our model	DistMass
DistMass distance error	DISTANCE_ERR	Error in our predicted distance	DistMass
Weighted DistMass distance	DISTANCE_WEIGHTED	DistMass distances weighted with Gaia distances	DistMass
DistMass mass uncorrected Sharma	MASS_UNCOR_SS	Estimated mass from the uncorrected Sharma model	DistMass
DistMass mass corrected Sharma	MASS_COR_SS	Corrected Sharma model mass	DistMass
DistMass mass uncorrected Mosser	MASS_UNCOR_MO	Uncorrected Mosser model mass	DistMass
DistMass mass corrected Mosser	MASS_COR_MO	Corrected Mosser model mass	DistMass
DistMass mass uncorrected White	MASS_UNCOR_TW	Uncorrected White model mass	DistMass
DistMass mass corrected White	MASS_COR_TW	Corrected White model mass	DistMass
DistMass mass error	MASS_ERR	Mass error estimate for all models	DistMass
DistMass age uncorrected Sharma	AGE_UNCOR_SS	Estimated age from our uncorrected Sharma model	DistMass
DistMass age corrected Sharma	AGE_COR_SS	Estimated age from the corrected Sharma model	DistMass
DistMass age uncorrected Mosser	AGE_UNCOR_MO	Estimated age from the uncorrected Mosser model	DistMass
DistMass age corrected Mosser	AGE_COR_MO	Estimated age from the corrected Mosser model	DistMass
DistMass age uncorrected White	AGE_UNCOR_TW	Estimated age from the uncorrected White model	DistMass
DistMass age corrected White	AGE_COR_TW	Estimated age from the corrected White model	DistMass
DistMass age error	AGE_ERR	Estimated error in our age estimates for all models	DistMass
BITMASK	BITMASK	Contains flags to indicate notes about a given star	
NMSU distance	NMSU_DIST	Distances from method described in 4.2	DistMass

labels used by the models, such as effective temperature, surface gravity, and metallicity. Table 2 shows the complete contents of the catalog. A bitmask is provided that contains flags for indicating whether a star was a member of the training set, whether it is outside the training parameter space for the mass model, and the source of the mass label; see Table 3.

Potential uses of the catalog could include testing stellar mixing models, examining various star clusters, probing the

Table 3
Bitmask Description

Bit	Bit Location	Description
1	000001	Marks stars that were members of distance model training set
2	000010	Stars outside the covered parameter space of the mass model
3	001000	Training mass from APOKASC (Pinsonneault et al. 2018)
4	010000	Training mass from Berger et al. (2020)
5	100000	Training mass from Serenelli et al. (2013)

history of the Milky Way, etc. Imig et al. (2023) use the distances and ages from our catalog to produce chemical and age maps of the Milky Way. Future work will include using our methods to produce a catalog utilizing new asteroseismic masses from K2 and TESS for the SDSS-V Milky Way Mapper project.

Acknowledgments

A.S-M., J.A.H., and J.I gratefully acknowledge support from NSF grant AST-1909897.

Funding for the Sloan Digital Sky Survey IV has been provided by the Alfred P. Sloan Foundation, the U.S. Department of Energy Office of Science, and the Participating Institutions.

SDSS-IV acknowledges support and resources from the Center for High Performance Computing at the University of Utah. The SDSS website is www.sdss.org.

SDSS-IV is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS Collaboration including the Brazilian Participation Group, the Carnegie Institution for Science, Carnegie Mellon University, Center for Astrophysics—Harvard & Smithsonian, the Chilean Participation Group, the French Participation Group, Instituto de Astrofísica de Canarias, The Johns Hopkins University, Kavli Institute for the Physics and Mathematics of the Universe (IPMU)/University of Tokyo, the Korean Participation Group, Lawrence Berkeley National Laboratory, Leibniz Institut für Astrophysik Potsdam (AIP), Max-Planck-Institut für Astronomie (MPIA Heidelberg), Max-Planck-Institut für Astrophysik (MPA Garching), Max-Planck-Institut für Extraterrestrische Physik (MPE), National Astronomical Observatories of China, New Mexico State University, New York University, University of Notre Dame, Observatário Nacional/MCTI, The Ohio State University, Pennsylvania State University, Shanghai Astronomical Observatory, United Kingdom Participation Group, Universidad Nacional Autónoma de México, University of Arizona, University of Colorado Boulder, University of Oxford, University of Portsmouth, University of Utah, University of Virginia, University of Washington, University of Wisconsin, Vanderbilt University, and Yale University.

ORCID iDs

Alexander Stone-Martinez https://orcid.org/0000-0003-4761-9305
Jon A. Holtzman https://orcid.org/0000-0002-9771-9622
Julie Imig https://orcid.org/0000-0003-2025-3585
Christian Nitschelm https://orcid.org/0000-0003-4752-4365
Keivan G. Stassun https://orcid.org/0000-0002-3481-9052
Joel R. Brownstein https://orcid.org/0000-0002-8725-1069

References

Abadi, M., Agarwal, A., Barham, P., et al. 2015, TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems, https://www.tensorflow.org/
Abdurro'uf, Accetta, K., Aerts, C., et al. 2022, ApJS, 259, 35
Bailer-Jones, C. A. L., Rybizki, J., Fouesneau, M., Mantelet, G., & Andrae, R. 2018, AJ, 156, 58
Bensby, T., Feltzing, S., & Oey, M. S. 2014, AA, 562, A71
Berger, T. A., Huber, D., Gaidos, E., van Saders, J. L., & Weiss, L. M. 2020, AJ, 160, 108
Binks, A. S., Jeffries, R. D., Sacco, G. G., et al. 2022, MNRAS, 513, 5727
Blanton, M. R., Bershady, M. A., Abolfathi, B., et al. 2017, AJ, 154, 28
Bowen, I. S., & Vaughan, A. H. 1973, ApOpt, 12, 1430
Bressan, A., Marigo, P., Girardi, L., et al. 2012, MNRAS, 427, 127
Brogaard, K., VandenBerg, D. A., Bedin, L. R., et al. 2017, MNRAS, 468, 645
Buder, S., Sharma, S., Kos, J., et al. 2021, MNRAS, 506, 150
Burnett, B., & Binney, J. 2010, MNRAS, 407, 339
Catelan, M., de Freitas Pacheco, J. A., & Horvath, J. E. 1996, ApJ, 461, 231
Catelan, M., Stetson, P. B., Pritzl, B. J., et al. 2006, ApJL, 651, L133
Choi, J., Dotter, A., Conroy, C., et al. 2016, ApJ, 823, 102
da Silva, R., Porto de Mello, G. F., Milone, A. C., et al. 2012, A&A, 542, A84
Das, P., & Sanders, J. L. 2018, MNRAS, 484, 294
Fazio, G. G., Hora, J. L., Allen, L. E., et al. 2004, ApJS, 154, 10
Feuillet, D. K., Bovy, J., Holtzman, J., et al. 2018, MNRAS, 477, 2326
Fraser, A. E., Joyce, M., Anders, E. H., Tayar, J., & Cantiello, M. 2022, ApJ, 941, 164
Friel, E. D., Jacobson, H. R., & Pilachowski, C. A. 2010, AJ, 139, 1942
Gaia Collaboration, Prusti, T., de Bruijne, J. H. J., et al. 2016, A&A, 595, A1
García Pérez, A. E., Allende Prieto, C., Holtzman, J. A., et al. 2016, AJ, 151, 144
Girardi, L., Rubele, S., & Kerber, L. 2010, in IAU Symp. 266, Star Clusters: Basic Galactic Building Blocks Throughout Time and Space, ed. R. de Grijs & J. R. D. Lépine (Cambridge: Cambridge Univ. Press), 320
Gunn, J. E., Siegmund, W. A., Mannery, E. J., et al. 2006, AJ, 131, 2332
Handberg, R., Brogaard, K., Miglio, A., et al. 2017, MNRAS, 472, 979
Hawley, S. L., Covey, K. R., Knapp, G. R., et al. 2002, AJ, 123, 3409
Haywood, M., Di Matteo, P., Lehnert, M. D., Katz, D., & Gómez, A. 2013, AA, 560, A109
Imig, J., Holtzman, J. A., Yan, R., et al. 2022, AJ, 163, 56
Imig, J., Price, C., Holtzman, J. A., et al. 2023, ApJ, 954, 124
Katz, D., Gómez, A., Haywood, M., Snaith, O., & Di Matteo, P. 2021, A&A, 655, A111
Kroupa, P. 2001, in ASP Conf. Ser. 228, Dynamics of Star Clusters and the Milky Way, ed. S. Deiters et al. (San Francisco, CA: ASP), 187
Lagarde, N., Decressin, T., Charbonnel, C., et al. 2012, AA, 543, A108
Leung, H. W., & Bovy, J. 2019, MNRAS, 489, 2079
Mackereth, J. T., Bovy, J., Leung, H. W., et al. 2019, MNRAS, 489, 176
Majewski, S. R., Schiavon, R. P., Frinchaboy, P. M., et al. 2017, AJ, 154, 94
Majewski, S. R., Zasowski, G., & Nidever, D. L. 2011, ApJ, 739, 25
Martig, M., Fouesneau, M., Rix, H.-W., et al. 2016, MNRAS, 456, 3655
Masseron, T., & Gilmore, G. 2015, MNRAS, 453, 1855
Mészáros, S., Masseron, T., García-Hernández, D. A., et al. 2020, MNRAS, 492, 1641
Miglio, A., Brogaard, K., Stello, D., et al. 2012, MNRAS, 419, 2077
Mosser, B., Dziembowski, W. A., Belkacem, K., et al. 2013, AA, 559, A137
Ness, M., Hogg, D. W., Rix, H.-W., Ho, A. Y. Q., & Zasowski, G. 2015, ApJ, 808, 16
Ness, M., Hogg, D. W., Rix, H.-W., et al. 2016, ApJ, 823, 114
Nidever, D. L., Holtzman, J. A., Prieto, C. A., et al. 2015, AJ, 150, 173
Paust, N. E. Q., Reid, I. N., Piotto, G., et al. 2010, AJ, 139, 476
Pinsonneault, M. H., Elsworth, Y. P., Tayar, J., et al. 2018, ApJS, 239, 32

Pols, O. R., Schröder, K.-P., Hurley, J. R., Tout, C. A., & Eggleton, P. P. 1998, MNRAS, 298, 525
Queiroz, A. B. A., Anders, F., Santiago, B. X., et al. 2018, MNRAS, 476, 2556
Reimers, D. 1977, A&A, 61, 217
Salaris, M., & Cassisi, S. 2008, Evolution of Stars and Stellar Populations (Hoboken, NJ: Wiley)
Salaris, P., Adriano, P., Piersimoni, A., Anna, M., & Cassisi, S. 2015, AA, 583, A87
Santiago, B. X., Brauer, D. E., Anders, F., et al. 2015, A&A, 585, A42
Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, ApJ, 500, 525
Serenelli, A. M., Bergemann, M., Ruchti, G., & Casagrande, L. 2013, MNRAS, 429, 3645
Shao, Z., & Li, L. 2019, MNRAS, 489, 3093
Sharma, S., Stello, D., Bland-Hawthorn, J., Huber, D., & Bedding, T. R. 2016, ApJ, 822, 15
Shetrone, M., Tayar, J., Johnson, J. A., et al. 2019, ApJ, 872, 137
Silva Aguirre, V., Bojsen-Hansen, M., Slumstrup, D., et al. 2018, MNRAS, 475, 5487
Spoo, T., Tayar, J., Frinchaboy, P. M., et al. 2022, AJ, 163, 229
Tailo, M., Milone, A. P., Lagioia, E. P., et al. 2021, MNRAS, 503, 694
Tennenholtz, G., Zahavy, T., & Mannor, S. 2018, arXiv:1802.05846
van de Ven, G., van den Bosch, R. C. E., Verolme, E. K., & de Zeeuw, P. T. 2005, AA, 445, 513
Vincenzo, F., Weinberg, D. H., Montalbán, J., et al. 2021, arXiv:2106.03912
Vink, J. S., de Koter, A., & Lamers, H. J. G. L. M. 2001, A&A, 369, 574
Viscasillas Vázquez, C., Magrini, L., Casali, G., et al. 2022, A&A, 660, A135
Wang, S., Zhang, H.-T., Bai, Z.-R., et al. 2021, RAA, 21, 292
Weinberg, D. H., Holtzman, J. A., Hasselquist, S., et al. 2019, ApJ, 874, 102
White, T. R., Bedding, T. R., Stello, D., et al. 2011, ApJ, 743, 161
Wilson, J. C., Hearty, F. R., Skrutskie, M. F., et al. 2019, PASP, 131, 055001
Wright, E. L., Eisenhardt, P. R. M., Mainzer, A. K., et al. 2010, AJ, 140, 1868
Zasowski, G., Cohen, R. E., Chojnowski, S. D., et al. 2017, AJ, 154, 198
Zinn, J. C., Pinsonneault, M. H., Huber, D., et al. 2019, ApJ, 885, 166