












# Spectroscopic age estimates for APOGEE red-giant stars: Precise spatial and kinematic trends with age in the Galactic disc<sup>★</sup>

F. Anders<sup>1,2,3</sup>, P. Gispert<sup>1</sup>, B. Ratcliffe<sup>4</sup>, C. Chiappini<sup>4</sup>, I. Minchev<sup>4</sup>, S. Nepal<sup>4,5</sup>, A. B. A. Queiroz<sup>4,5</sup>, J. A. S. Amarante<sup>1,2,3,6</sup>, T. Antoja<sup>1,2,3</sup>, G. Casali<sup>7,8</sup>, L. Casamiquela<sup>1,2,3,9</sup>, A. Khalatyan<sup>4</sup>, A. Miglio<sup>5,6</sup>, H. Perotoni<sup>1,10</sup>, and M. Schultheis<sup>11</sup>

- <sup>1</sup> Departament de Física Quàntica i Astrofísica (FQA), Universitat de Barcelona (UB), Martí i Franquès 1, 08028 Barcelona, Spain  
e-mail: [fanders@fqa.ub.edu](mailto:fanders@fqa.ub.edu)
- <sup>2</sup> Institut de Ciències del Cosmos (ICCUB), Universitat de Barcelona (UB), Martí i Franquès 1, 08028 Barcelona, Spain
- <sup>3</sup> Institut d'Estudis Espacials de Catalunya (IEEC), Gran Capità 2-4, 08034 Barcelona, Spain
- <sup>4</sup> Leibniz-Institut für Astrophysik Potsdam (AIP), An der Sternwarte 16, 14482 Potsdam, Germany
- <sup>5</sup> Institut für Physik und Astronomie, Universität Potsdam, Haus 28 Karl-Liebknecht-Str. 24/25, 14476 Golm, Germany
- <sup>6</sup> Visiting Fellow, Jeremiah Horrocks Institute, University of Central Lancashire, Preston PR1 2HE, UK
- <sup>7</sup> Department of Physics & Astronomy "Augusto Righi", University of Bologna, Via Gobetti 93/2, 40129 Bologna, Italy
- <sup>8</sup> INAF – Astrophysics and Space Science Observatory of Bologna, Via Gobetti 93/3, 40129 Bologna, Italy
- <sup>9</sup> GEPI, Observatoire de Paris, PSL Research University, CNRS, Sorbonne Paris Cité, 5 place Jules Janssen, 92190 Meudon, France
- <sup>10</sup> Universidade de São Paulo, Instituto de Astronomia, Geofísica e Ciências Atmosféricas, Departamento de Astronomia, SP 05508-090 São Paulo, Brazil
- <sup>11</sup> Université Côte d'Azur, Observatoire de la Côte d'Azur, Laboratoire Lagrange, CNRS, Blvd de l'Observatoire, 06304 Nice, France

Received 16 April 2023 / Accepted 8 August 2023

## ABSTRACT

Over the last few years, many studies have found an empirical relationship between the abundance of a star and its age. Here we estimate spectroscopic stellar ages for 178 825 red-giant stars observed by the APOGEE survey with a median statistical uncertainty of 17%. To this end, we use the supervised machine learning technique XGBoost, trained on a high-quality dataset of 3060 red-giant and red-clump stars with asteroseismic ages observed by both APOGEE and *Kepler*. After verifying the obtained age estimates with independent catalogues, we investigate some of the classical chemical, positional, and kinematic relationships of the stars as a function of their age. We find a very clear imprint of the outer-disc flare in the age maps and confirm the recently found split in the local age-metallicity relation. We present new and precise measurements of the Galactic radial metallicity gradient in small age bins between 0.5 and 12 Gyr, confirming a steeper metallicity gradient for ~2–5 Gyr old populations and a subsequent flattening for older populations mostly produced by radial migration. In addition, we analyse the dispersion about the abundance gradient as a function of age. We find a clear power-law trend (with an exponent  $\beta \approx 0.15$ ) for this relation, indicating a relatively smooth radial migration history in the Galactic disc over the past 7–9 Gyr. Departures from this power law may possibly be related to the Gaia Enceladus merger and passages of the Sagittarius dSph galaxy. Finally, we confirm previous measurements showing a steepening in the age-velocity dispersion relation at around ~9 Gyr, but now extending it over a large extent of the Galactic disc ( $5 \text{ kpc} < R_{\text{Gal}} < 13 \text{ kpc}$ ). To establish whether this steepening is the imprint of a Galactic merger event, however, detailed forward modelling work of our data is necessary. Our catalogue of precise stellar ages and the source code to create it are publicly available.

**Key words.** Galaxy: evolution – Galaxy: stellar content – methods: data analysis – methods: statistical – stars: abundances – stars: late-type

## 1. Introduction

Isochrone fitting is frequently used to determine stellar ages of (primarily FGK) field stars, especially in the context of large spectroscopic stellar surveys (e.g. [Santiago et al. 2016](#); [Mints & Hekker 2017](#); [McMillan et al. 2018](#); [Sanders & Das 2018](#); [Lebreton & Reese 2020](#); [Queiroz et al. 2023](#)). While this technique has a long tradition (e.g. [Pont & Eyer 2004](#); [Jørgensen & Lindegren 2005](#)) and works reasonably well for stars close to the main-sequence turn-off and sub-giant branch (but see [Valle et al. 2013](#); [Lebreton et al. 2014](#)), isochrone-fitting for red-giant stars is much more challenging and prone

to sizeable statistical and systematic errors ([Soderblom 2010](#); [Noels & Bragaglia 2015](#)).

Another well-tested (but also model-dependent) method to estimate ages for field stars, including red giants, is offered by asteroseismology ([Chaplin & Miglio 2013](#)). Red-giant stars are particularly interesting for Galactic archaeology studies, since they are numerous, bright, and cover a wide range of stellar ages ([Miglio et al. 2012](#)). Until the upcoming PLATO mission ([Rauer et al. 2014](#), [Miglio et al. 2017](#)), however, red-giant ages derived from joint asteroseismic and spectroscopic constraints are only available for select samples ( $\lesssim 10\,000$  stars) in certain fields, such as the *Kepler* (e.g. [Pinsonneault et al. 2014, 2018](#); [Wu et al. 2018](#); [Miglio et al. 2021](#); [Matsuno et al. 2021](#)), CoRoT ([Valentini et al. 2016](#); [Anders et al. 2017b](#)), K2 (e.g. [Rendle et al. 2019](#); [Valentini et al. 2019](#); [Zinn et al. 2020, 2022](#);

<sup>★</sup> Full Table A.1 is available at the CDS via anonymous ftp to [cdsarc.cds.unistra.fr](https://cdsarc.cds.unistra.fr) (130.79.128.5) or via <https://cdsarc.cds.unistra.fr/viz-bin/cat/J/A+A/678/A158>

Schonhut-Stasik et al. 2023), or TESS Continuous Viewing Zone (Sharma et al. 2018; Silva Aguirre et al. 2020; Mackereth et al. 2021; Wu et al. 2023) fields. Therefore, large-scale spectroscopic surveys such as APOGEE (Majewski et al. 2017), GALAH (De Silva et al. 2015), LAMOST (Cui et al. 2012), or *Gaia* RVS (Gaia Collaboration 2023a; Recio-Blanco et al. 2023), have been aiming at providing empirical spectroscopy-based age estimates for Galactic archaeology studies, typically using asteroseismology as benchmark data (e.g. Martig et al. 2016a; Leung & Bovy 2019; Ciucă et al. 2023; He et al. 2022).

It has long been suggested that, in the absence of reliable stellar ages, precise chemical abundances could be used to determine the approximate age of a star. The works of Nissen (2015, 2016) and Tucci Maia et al. (2016) showed that, at least for solar twins observed at high spectral resolution and high signal-to-noise ratio, this is a viable assumption. The tight relation between  $[Y/Mg]$  and stellar age found in these works, later confirmed by Spina et al. (2018), demonstrated that a combination of elemental abundances probing different nucleosynthetic channels (in the case of  $[Y/Mg]$  an *s*-process and an  $\alpha$  element) may provide a meaningful empirical measure for stellar age. Follow-up works also explored other elemental abundance ratios (e.g. Jofré et al. 2020; Casamiuela et al. 2021) and showed that when leaving the regime of stellar twins, this empirical “chemical clock” needs to include other terms (such as  $T_{\text{eff}}$  or  $[Fe/H]$ ) to provide meaningful results (e.g. Feltzing et al. 2017; Delgado Mena et al. 2019; Casali et al. 2020; Viscasillas Vázquez et al. 2022).

The APOGEE survey, which has taken high-resolution ( $R \simeq 22\,500$ ) high signal-to-noise ( $\gtrsim 70$ ) spectra of mainly red giant stars in the near-infrared *H*-band (Wilson et al. 2019), covering several carbon and nitrogen molecular bands (Allende Prieto et al. 2008), triggered the discovery that in red-giant stars the  $[C/N]$  ratio correlates with stellar mass (and therefore age; Masseron & Gilmore 2015). Subsequent theoretical work by Salaris et al. (2015) showed that  $[C/N]$  measurements alone cannot provide precise and accurate ages for individual stars, but may well provide population ages (see also Lagarde et al. 2017 for a detailed investigation of the underlying mixing effects that produce the  $[C/N]$  dependence on age). The  $[C/N]$  ratio was since used with success to estimate red-giant ages (with  $\simeq 30$ – $40\%$  precision) through a calibration of the  $[C/N]$ -age relation with asteroseismic data from the *Kepler* mission (Martig et al. 2016a; Ness et al. 2016; Wu et al. 2019; Huang et al. 2020; Zhang et al. 2021) or with open clusters (Casali et al. 2019; Spoo et al. 2022).

Using data from the GALAH survey’s third data release (Buder et al. 2021), Hayden et al. (2022) have recently demonstrated that it is possible to infer “spectroscopic” (or “chemical”) stellar ages for main-sequence turn-off stars with a precision of 1–2 Gyr, using not just one abundance ratio, but a combination of many abundances (not even including carbon and nitrogen), by using supervised machine-learning regression, in particular the popular method extreme gradient-boosted trees (XGBoost; Chen & Guestrin 2016). A similar exercise using the same technique has been conducted by He et al. (2022), using red-clump stars observed by the LAMOST survey (which has the advantage of providing also C and N abundances) trained on asteroseismically inferred ages from *Kepler*. Their resulting statistical uncertainties, similar to the earlier results for APOGEE (Martig et al. 2016a; Ness et al. 2016), amount to 31%, which, considering the low resolution of the LAMOST spectra ( $R \sim 1800$ ), is remarkable. Similar precisions were obtained earlier

by Mackereth et al. (2019) using a Bayesian convolutional neural network and APOGEE DR14 data trained on the APOGEE-*Kepler* data (Pinsonneault et al. 2018). Good spectroscopic age estimates have also recently been obtained by Moya et al. (2022) through hierarchical Bayesian modelling for the high-resolution HARPS-GTO sample, by Bu et al. (2020) through Gaussian Process regression on LAMOST data (again using the APOGEE-*Kepler* data as a training set), or by Hasselquist et al. (2020) using The Cannon (Ness et al. 2015) for high-luminosity red giants in the Galactic bulge (using the local giant sample of Feuillet et al. 2018 as a training set).

In this work, we build upon the success of the works of Hayden et al. (2022) and He et al. (2022) and use XGBoost to estimate ages for 178 825 red-giant stars contained in the APOGEE DR17 catalogue (Abdurro’uf et al. 2022). As a training set, we use the carefully produced APOGEE-*Kepler* catalogue of Miglio et al. (2021).

The paper is structured as follows: In Sect. 2 we present the APOGEE data used in this work (including the APOGEE-*Kepler* training set). In Sect. 3 we explain the XGBoost algorithm. The obtained ages are validated in Sect. 4. In Sect. 5, we show that our ages reproduce known chemical, spatial, and kinematic trends with age, suggesting that our inferred ages are reliable even far from the *Kepler* field. Finally, the conclusions are presented in Sect. 6. Our analysis is reproducible online<sup>1</sup>.

## 2. Data

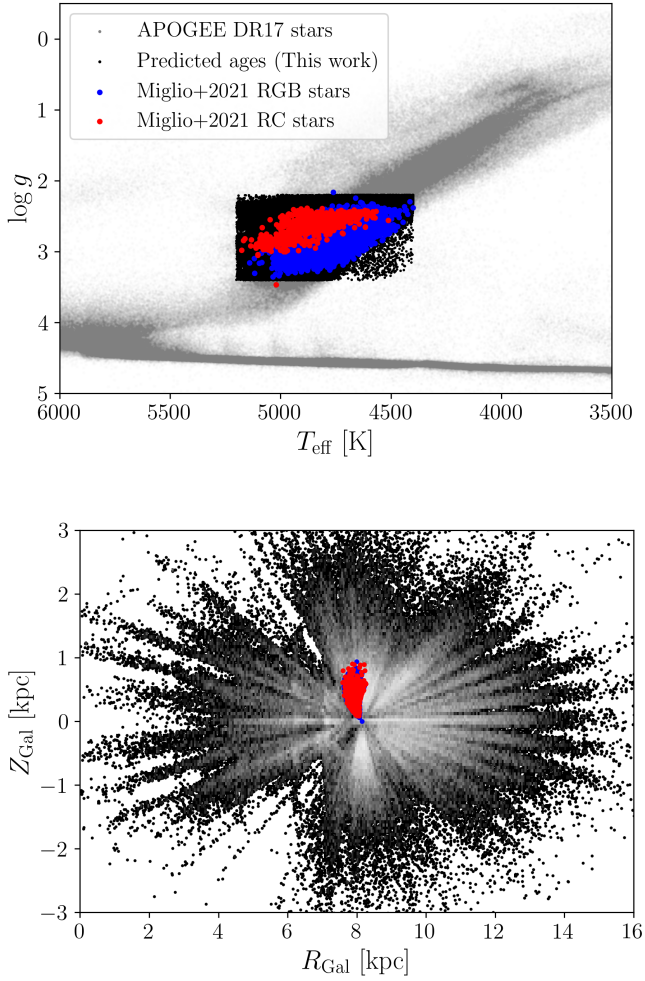
The outcome of a machine-learning regressor can only be as reliable as its training set. Therefore, it is often preferable to use a smaller, but statistically significant and highest-quality, dataset for the training phase. For red-giant stars, the longest asteroseismic time series have been obtained by the *Kepler* satellite (Borucki et al. 2010) during its first phase of observations in which it was continuously observing one large sky area in the sky, towards the constellation of Cygnus.

Using APOGEE DR14 (Abolfathi et al. 2018) spectroscopic follow-up observations of *Kepler* targets (Pinsonneault et al. 2018; Yu et al. 2018), Miglio et al. (2021) inferred masses and ages for more than 5000 giants with available *Kepler* light curves and APOGEE spectra using the Bayesian isochrone-fitting code PARAM (da Silva et al. 2006; Rodrigues et al. 2017). This is arguably an asteroseismic+spectroscopic dataset that is of such a high quality that it can be used for training a machine-learning regressor. The median nominal uncertainties of the asteroseismic ages reported in the Miglio et al. (2021) catalogue are 23% for RGB stars and 10% for RC stars. For this study we cross-matched the results of Miglio et al. (2021) with the APOGEE DR17 allStarLite table (Abdurro’uf et al. 2022), resulting in a preliminary training dataset of 3315 stars. We clean the APOGEE-*Kepler* sample to contain only stars with APOGEE signal-to-noise ratio (SNREV) above 70 and nominal age uncertainties smaller than both 30% and 3 Gyr. We also exclude the few apparently young  $\alpha$ -rich stars (Chiappini et al. 2015; Martig et al. 2015; here defined as  $[\alpha/Fe] > 0.14 + 0.0001 \cdot (\text{age} - 4.1)^5$ ), since their ages (inferred from single-star evolutionary models) are most probably flawed by binary interactions (e.g. Fuhrmann & Chini 2017; Jofré et al. 2023).

As we are interested in using the APOGEE abundance measurements as features for the age estimator, we require that the abundance flags<sup>2</sup> for each of the used elements are equal

<sup>1</sup> [https://github.com/fjaellet/xgboost\\_chem\\_ages](https://github.com/fjaellet/xgboost_chem_ages)

<sup>2</sup> <https://www.sdss4.org/dr17/irspec/abundances/>



**Fig. 1.** Parameter space covered by our APOGEE sample. Top panel: Kiel diagram ( $\log g$  vs.  $T_{\text{eff}}$ ) for the APOGEE DR17 catalogue (grey density). The black dots highlight the stars range for which our method provides chemical stellar age estimates. Bottom panel: Galactic distribution of the 178 825 APOGEE stars with spectroscopic ages in cylindrical coordinates ( $Z_{\text{Gal}}$  vs.  $R_{\text{Gal}}$ ), based on the latest StarHorse distances (Queiroz et al. 2023). In both panels, the training sample (taken from the APOGEE-Kepler catalogue of Miglio et al. 2021) is overlotted in colours.

to zero (meaning that their abundance estimates are unproblematic). The chosen chemical features are:  $T_{\text{eff}}$ ,  $\log g$ , [C/Fe], [Cl/Fe], [N/Fe], [O/Fe], [Na/Fe], [Mg/Fe], [Al/Fe], [Si/Fe], [K/Fe], [Ca/Fe], [Ti/Fe], [V/Fe], [Mn/Fe], [Co/Fe], [Ni/Fe], and [Ce/Fe]. The following abundances were discarded: [P/Fe], [S/Fe], [Cr/Fe], [Fe/H], and [Cu/Fe] – either because there were too few stars with unproblematic ASPCAP flags ([P/Fe] and [Cu/Fe]), because their importance on the final result was negligible ([S/Fe], [Cr/Fe]; see Sect. 3.2), or because we deliberately aim at keeping the age estimate independent from the abundance ratio in question ([Fe/H]). This leaves us with 3060 stars covering a metallicity range between  $-1$  and  $+0.5$ . As can be appreciated in Fig. 1 (top panel), the training data are also limited to a small range in the spectroscopic Hertzsprung–Russell diagram:  $4400 \leq T_{\text{eff}} [\text{K}] \leq 5200$  and  $2.2 \leq \log g [\text{cgs}] \leq 3.4$ , compared to the full APOGEE DR17 data.

When applying our age estimator to the full APOGEE DR17 dataset, we restrict our predictions to the stellar-parameter space covered by the training data (black dots in Fig. 1, top panel),

and to  $[\text{Fe}/\text{H}] > -1$ . We also require clean abundance flags for all included elements, but allow for lower signal-to-noise ratios ( $\text{SNREV} > 50$ ). The bottom panel of Fig. 1 shows the distribution of the training and full sample in Galactocentric cylindrical coordinates, demonstrating the huge gain in area covered with red-giant age estimates – once our method is applied.

### 3. Method

Supervised machine-learning regression models can be trained to fit arbitrarily complex relationships between the introduced data. In our case of tabular data, we want to predict an output column (the so-called “label”) based on a set of introduced input parameter columns (so-called “features”). The features (in our case stellar parameters and abundances) are the variables of the data that the model uses to train itself to predict the label (stellar age). Once the model is trained, it can be used to make predictions (age estimates) based on data for which reliable labels do not yet exist.

Borisov et al. (2022) recently benchmark-tested several regression algorithms for tabular data. They found that the best-performing algorithm, in terms of accuracy and speed, was the well-known tree-based algorithm XGBoost. We therefore use this algorithm as the basis for our chemical age estimator in this paper.

#### 3.1. XGBoost

The XGBoost algorithm (Chen & Guestrin 2016) is the culmination of previous development in tree-based machine learning. The basis of tree-based algorithms are decision trees, which are graphical representations of possible solutions to a decision based on certain conditions prompted by the values of the input columns (for an introduction in the astronomical context see e.g. Ivezić et al. 2020, Chapt. 9.7).

A next level of complexity was achieved by random-forest algorithms (Breiman 2001), which randomly select features (columns) to construct a set of decision trees. On top of random-forest algorithms, one can add so-called gradient-boosting methods, which minimise the errors and increase the performance of the random-forest models. These models use a gradient-descent algorithm to minimise the errors of the sequential models. Finally, in contrast to other gradient-boosted methods, XGBoost also performs parallel processing and tree-pruning, handles missing values, and regularises the data to avoid over-fitting.

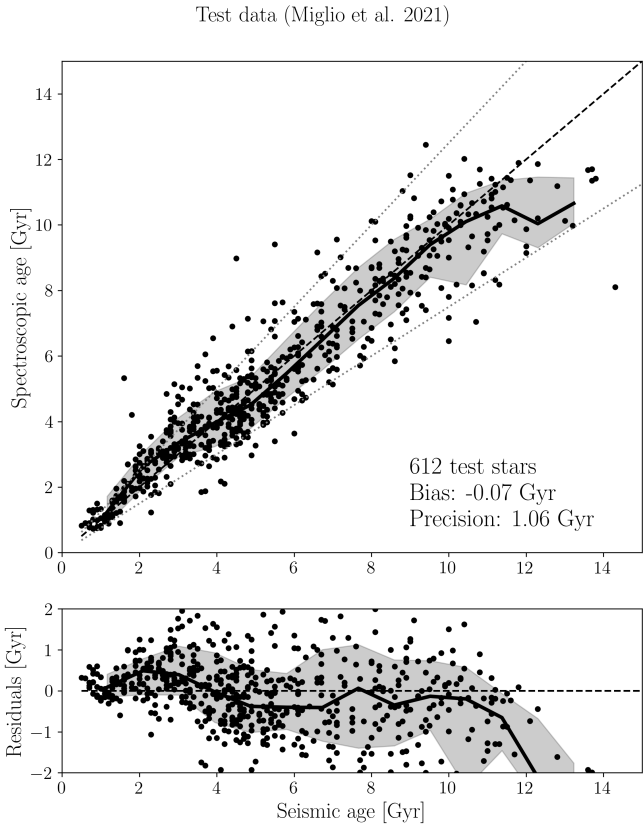
In consequence, the training time of XGBoost is often much smaller compared to other methods with similar predictive power (see Borisov et al. 2022). Since the algorithm is very well explained in the original paper (Chen & Guestrin 2016) as well as in several online resources, we abstain from a deeper explanation here and only mention that XGBoost has a number of hyperparameters that can be tuned to optimise the performance for a given dataset or problem.

For our use case, we use scikit-learn’s (Virtanen et al. 2020) GridSearch<sup>3</sup> to find the optimal hyperparameters for our machine-learning model. The optimal hyperparameters used in our fiducial XGBoost<sup>4</sup> regressor are: `learning_rate = 0.005`, `max_depth = 7`, `min_child_weight = 10`, `n_estimators = 1500`, `subsample = 0.6`. The performance of the default model is shown in Fig. 2 and discussed in detail in Sect. 3.3. In addition,

<sup>3</sup> [https://scikit-learn.org/stable/modules/grid\\_search.html](https://scikit-learn.org/stable/modules/grid_search.html)

<sup>4</sup> <https://xgboost.readthedocs.io/en/stable/>





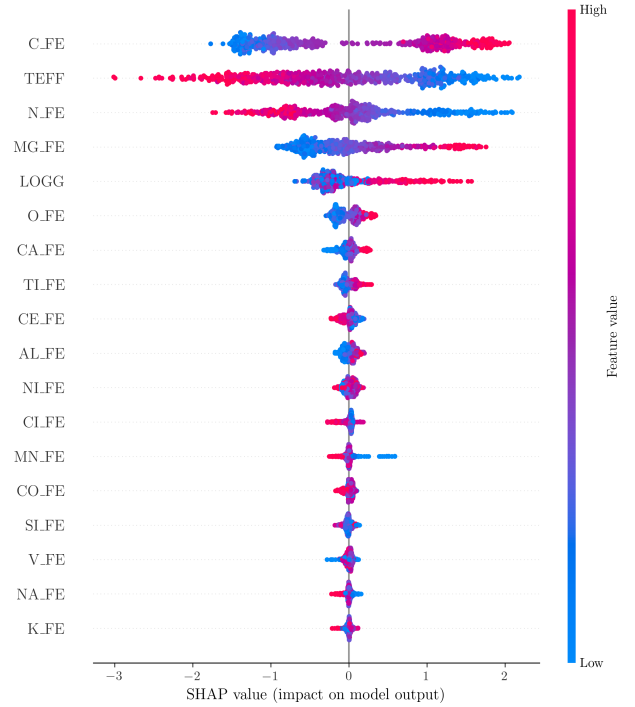
**Fig. 2.** Performance of the XGBoost chemical age estimator for the unseen test data. Top panel: spectroscopic age vs. asteroseismic age (from Miglio et al. 2021), using our default model. The solid black line and the shaded areas delineate the median trend and  $1\sigma$  quantiles. The dashed line delineates the identity line, while the grey dotted lines correspond to  $\pm 25\%$  deviation. Bottom panel: residuals (spectroscopic – seismic age), showing that our model typically provides reliable ages (with systematic residuals within  $\sim 0.5$  Gyr for stars with ages  $\leq 11$  Gyr).

we also provide alternative estimates with individual uncertainties, by using quantile regression (available in the latest version of XGBoost) to include the prediction of  $1\sigma$  confidence intervals in the inference (see Appendix A). In this paper we mainly use the results of our fiducial XGBoost run (column `spec_age_xgb` in Table A.1).

### 3.2. SHAP values

The concept of SHAP (SHapley Additive exPlanations) values provides an elegant way to understand the output of a machine learning model (Lundberg & Lee 2017). In the case of XGBoost, they can be used to understand how each feature has an impact on the predictions of the model. The sum of SHAP values is equal to the difference between the expected output and the baseline output. In other words, a positive value means that the feature increases the output’s model, while a negative value decreases it. For our case, the SHAP values have been used not only to recognize the features that have the greatest impact on the model, but also to discard those that do not contribute to the prediction (the APOGEE DR17 elemental abundances discarded in Sect. 2).

Figure 3 demonstrates that the most influential features in the age prediction are the effective temperature,  $T_{\text{eff}}$ , the chem-



**Fig. 3.** SHAP “bee-swarm” plot ([https://shap.readthedocs.io/en/latest/example\\_notebooks/api\\_examples/plots/bee-swarm.html](https://shap.readthedocs.io/en/latest/example_notebooks/api_examples/plots/bee-swarm.html)). The rows correspond to the input features, ordered by importance. In each row, one dot corresponds to a star in the test dataset, coloured by its (normalised) feature value. The position of each dot indicates how much and in which way each feature contributes to its output label (age).

ical abundances  $[C/Fe]$ ,  $[Mg/Fe]$ ,  $[N/Fe]$ , and the surface gravity,  $\log g$ . In addition, Fig. 3 shows the exact effect each of the features has on the estimated age: for example, for the case of effective temperature (redder dots in the first line), we observe that an increase in its value implies a negative SHAP value, i.e. a decrease in the inferred age. The usefulness of carbon and nitrogen abundances for the prediction of red-giant ages in the absence of asteroseismology is discussed in detail by Martig et al. (2016a). These authors performed a polynomial feature regression on the APOGEE DR12 labels  $\{T_{\text{eff}}, \log g, [M/H], [C/M], [N/M], [C+N/M]\}$  to the APOKASC-1 ages (Pinsonneault et al. 2014) and obtained reasonably precise age estimates for 52 000 APOGEE stars ( $\sim 40\%$ ). Similar precision was achieved by Ness et al. (2016), using the same training data as Martig et al. (2016a), but a different technique (The Cannon, a data-driven regression technique directly trained on the spectra; Ness et al. 2015).

The fact that  $[C/Fe]$ ,  $T_{\text{eff}}$ ,  $[N/Fe]$ ,  $[Mg/Fe]$ , and  $\log g$  are the most important features for age determination is very much in line with the findings of Bu et al. (2020) who, using a variety of machine-learning regression techniques, found that these features contained most information for predicting ages of LAMOST stars (in the absence of reliable mass estimates). Also the work of Ciucă et al. (2023) used these five features ( $+ [Fe/H]$ ) to determine ages for APOGEE stars. We also see from Fig. 3 that most other individual abundance ratios (first  $[O/Fe]$  and other  $\alpha$  elements, then  $[Ce/Fe]$ , and then iron-peak elements) have a smaller impact on the estimated age. Although all abundance measurements are used by XGBoost to some extent to estimate an even more precise age, the precision



improvement of the full model using all abundances over a model that only uses the feature space  $\{[C/Fe], T_{\text{eff}}, [N/Fe], [Mg/Fe], \log g\}$  is only  $\sim 10\%$ . This suggests that while each element is unique, the key information is stored in only a few vectors (in agreement with previous works, e.g. Price-Jones & Bovy 2018; Weinberg et al. 2019; Ratcliffe et al. 2020). Interestingly, Ce is not among the most influential features despite being an *s*-process element for which the correlation with age is strong (e.g. Maiorca et al. 2011; Spina et al. 2018; Magrini et al. 2018; Delgado Mena et al. 2019; Casamiquela et al. 2021; Sales-Silva et al. 2022; Casali et al. 2023). This might be due to the still sizeable uncertainties of this element in the APOGEE DR17 catalogue (see also Hayes et al. 2022).

### 3.3. Biases and uncertainties

The performance of our default model (trained on 80% of the Miglio et al. 2021 data) is shown for the test dataset (the 20% of the data not used in the training phase) in Fig. 2. We observe a very clear linear trend with little wiggles ( $\leq 0.5$  Gyr for almost the full age range), and a determination coefficient of  $R^2 = 0.91$ . Our spectroscopic ages tend to slightly overestimate (by  $< 0.5$  Gyr) ages for stars around 2 Gyr, while for stars between 5 and 7 Gyr our age scale tends to underestimate the seismic age by similar amounts. Beyond seismic ages of  $\sim 11$  Gyr, our method provides more seriously underestimated ages compared to the seismic scale.

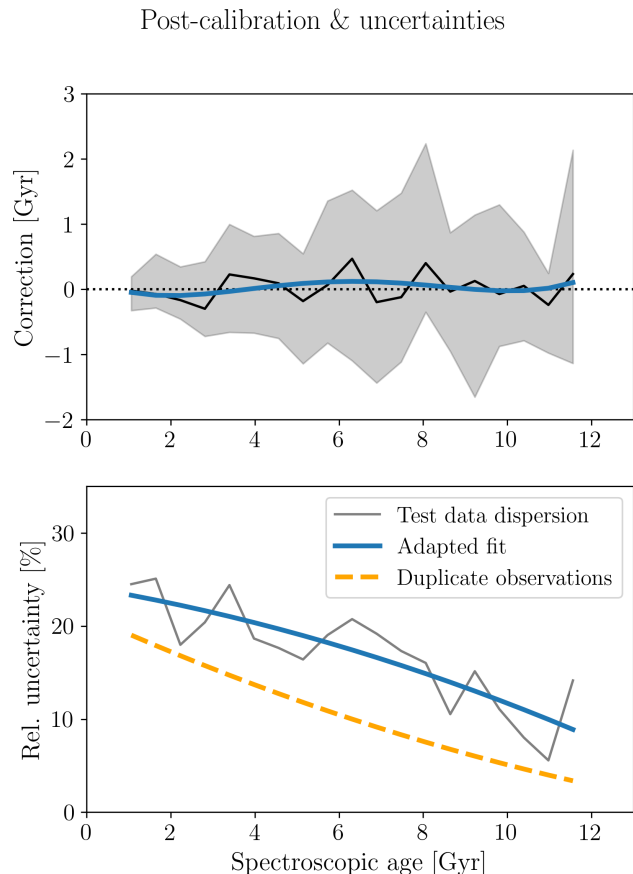
Considering that the age uncertainties associated with the training data are of order 10% for RC and 25% for first-ascent red-giant branch (RGB) stars<sup>5</sup>, the amount of scatter around the one-to-one relation in Fig. 2 is indeed surprisingly small (almost exactly the same values), indicating that the regressor is well trained and that most of the variance stems from the training set itself. In addition, and contrary to the training set, XGBoost is not extrapolating the age scale beyond the age of the Universe (13.7 Gyr), which again is surprising.

We also tested training separate XGBoost regressors for the RC and RGB stars, respectively. However, a global increase in performance could not be appreciated, most probably because the default algorithm also learns the separation between RC and RGB stars from the data (in particular,  $T_{\text{eff}}$ ,  $\log g$ ,  $[C/Fe]$ , and  $[N/Fe]$ ). Our default XGBoost model is therefore capable of estimating meaningful ages for both RC and RGB stars. We appreciate that the performance for the RC stars is significantly better (lower dispersion around the identity line), but the model still works well on average for RGB stars. The statistical age uncertainty of our method is lower than 25% and only weakly dependent on the estimated stellar age (see Fig. 4, bottom panel).

## 4. Validation

Age estimates for field stars are, as explained in the Introduction, highly dependent on the accuracy of stellar evolutionary models, even in the case of the highest quality data. On top of the systematic age biases that are unavoidable and very hard to quantify especially in the case of red-giant stars (related to e.g. assumptions about mass loss, mixing, rotation, etc.; Noels & Bragaglia 2015), there are also sizeable statistical uncertainties. Nevertheless, comparisons to other model-dependent age estimates

<sup>5</sup> We note that these dispersions do not reflect the systematic uncertainties which are considerably higher for red-clump stars (e.g. Casagrande et al. 2016; Anders et al. 2017b) due to the poorly constrained amount of mass loss after the first dredge-up phase.

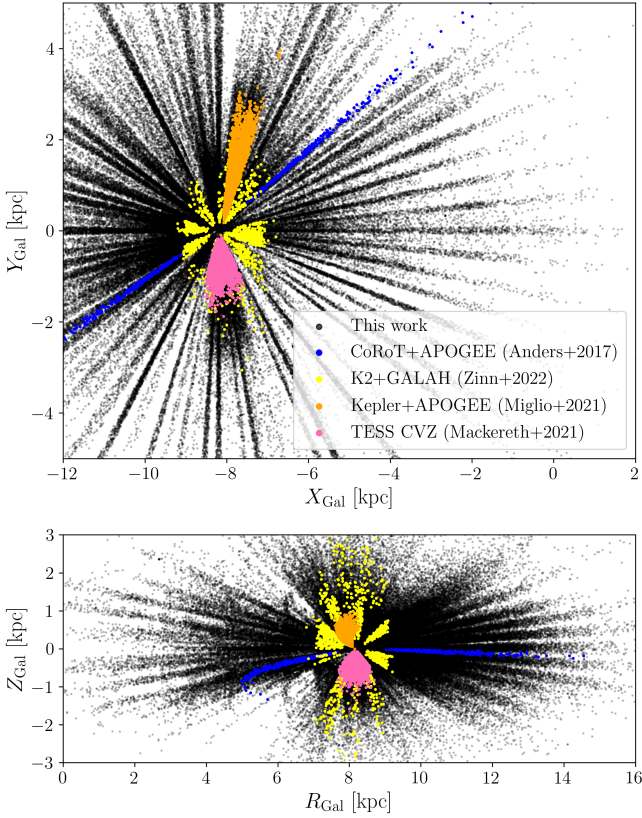


**Fig. 4.** Possible a-posteriori calibration of the age estimates (top panel) and estimation of the internal statistical uncertainties (bottom panel) using the same data as in Fig. 2. The top panel shows the median age bias (test seismic age – spectroscopic age) as a function of the spectroscopic age. The bottom panel shows the relative age error (width of the grey distribution in the top panel normalised by the spectroscopic age). As a sanity check, the orange dotted line in the bottom panel shows the median relative age dispersion among duplicate APOGEE observations. In both panels, we also show the simple polynomial fits used to characterise the bias and  $1\sigma$  age uncertainty in our catalogue.

are of prime importance to estimate the amount of systematic and statistical uncertainties. In this section, we infer internal uncertainties of our method and compare the resulting ages to asteroseismic ages obtained from observing campaigns other than *Kepler*, open cluster ages, and APOGEE DR17 age estimates from other machine-learning methods. We provide additional verification of our ages against isochrone fitting and an empirical  $[C/N]$  calibration in Appendix C.

### 4.1. Duplicate APOGEE observations

More than 11 000 stars in our selected region of the *Kiel* diagram (Fig. 1) are contained more than once in APOGEE DR17 catalogue (for example, stars that have been observed in different fields or even by both the SDSS Telescope and the du Pont Telescope). These duplicate observations can serve to estimate realistic internal uncertainties for stellar parameters and elemental abundances (e.g. Jönsson et al. 2020, Tables 10 and 11). In our case we can use these multiple observations also to assess the uncertainties of our age estimates: the scatter in the stellar parameters and abundances naturally leads to a scatter in the inferred ages. The results of this experiment are



**Fig. 5.** Galactic distribution of asteroseismic+spectroscopic red-giant samples, superimposed on our sample of APOGEE DR17 red-giants with spectroscopic ages determined in this work.

shown in the bottom panel of Fig. 4 (orange dashed curve). Reassuringly, the dispersion among multiple observations (with often significantly different signal-to-noise ratios and, therefore, elemental-abundance precisions) is always lower than the uncertainty estimated from the test dataset.

#### 4.2. Asteroseismic ages in other fields (CoRoT, K2, TESS)

Asteroseismology coupled with spectroscopy can significantly reduce the uncertainties of red-giant ages (e.g. Miglio et al. 2013; Pinsonneault et al. 2014). This is exactly the reason why we have chosen the APOGEE-*Kepler* dataset of Miglio et al. (2021) as a training set for our method. The concordance with this age scale is inherent to the machine-learning method used in this paper, and was demonstrated in Fig. 2 for the unseen test data.

Here we compare our age estimates for the full APOGEE DR17 data with other age estimates that used similar techniques as Miglio et al. (2021), but on other asteroseismic and spectroscopic data. The main reason is that we need to assess whether the stellar populations contained in the training set (i.e. the *Kepler* field) are representative enough – so that our method can also successfully be used to derive ages for stars in other parts of the Galaxy. To verify this, we study the CoRoT-APOGEE data presented in Anders et al. (2017b), the K2-GALAH data (Zinn et al. 2022), and the first Galactic Archaeology data from the Southern continuous viewing zone of the TESS satellite (TESS CVZ; Mackereth et al. 2021). The Galactic distribution of all these targets, along with our APOGEE DR17 stars with spectroscopic age estimates, is shown in Fig. 5. The distances for

the Zinn et al. 2022 and for the Mackereth et al. (2021) stars in this plot were taken from Queiroz et al. (2023) and Anders et al. (2022), respectively.

Figure 6 (top left panel) presents the direct comparison with the asteroseismic ages obtained by Anders et al. (2017b), Zinn et al. (2022), and Mackereth et al. (2021). The plot shares some similarities with the comparison plot for test sample shown in Fig. 2: in the youngish regime, the XGBoost ages appear overestimated with respect to the seismic scale for all three external datasets, albeit much more pronounced than in the *Kepler* sample. In the intermediate regime (between 4 and 10 Gyr), the concordance is better on average, and there is no consistent trend among the three external seismic samples (in this regime, the age scale of Zinn et al. 2022 is lower, while the one of Mackereth et al. 2021 is higher than ours). For the oldest ages (>10 Gyr), the XGBoost ages appear underestimated with respect to the seismic scale, at least for the case of the TESS-CVZ sample.

Finally, it is important to recall that also the asteroseismic+spectroscopic catalogues used for comparison here are prone to sizeable statistical and systematic uncertainties. Due to the shorter time series of especially CoRoT and K2, these uncertainties are typically much larger than the ones associated to the *Kepler* training data. To illustrate this, the typical statistical age uncertainties for each of the considered catalogues are shown as error bars in Fig. 6.

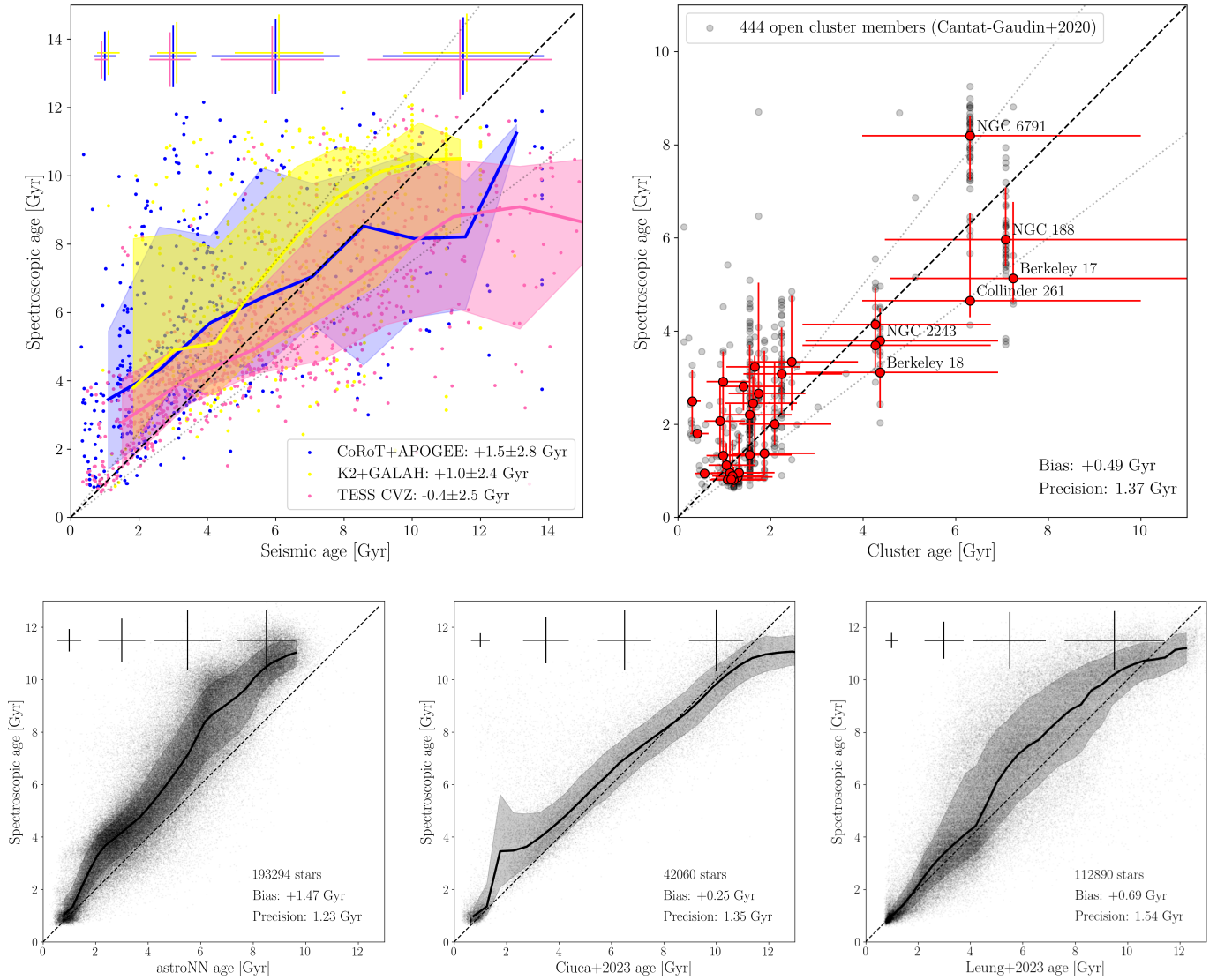
#### 4.3. Open clusters (Cantat-Gaudin et al. 2020)

Figure 6 (top right panel) shows the comparison of our spectroscopic age estimates with open-cluster (OC) ages obtained by Cantat-Gaudin et al. (2020) by *Gaia* DR2 colour-magnitude diagram fitting with an artificial neural network trained on a mixture of simulations and real data (mostly stemming from Bossini et al. 2019). As indicated by the error bars in the figure, the OC age estimates also come with sizeable uncertainties. Keeping this in mind, the concordance between our age estimates and the OC ages is not too bad: on average, our ages are greater than the OC ages by 0.48 Gyr, and the dispersion is 1.37 Gyr (~40% larger than the uncertainties estimated from the test dataset in Fig. 2).

In spite of the statistical concordance, however, we observe a tendency to overestimate age for a group of clusters younger than 2.5 Gyr. Queiroz et al. (2023) observed a similar behaviour in their comparison of sub-giant star age estimates with the OC age scale of Cantat-Gaudin et al. (2020). This suggests that either the OC age scale of Cantat-Gaudin et al. (2020) has to be revised (perhaps by enlarging the OC training set with reliable ages, or by better taking into account metallicity effects) or that both the isochrone ages of Queiroz et al. (2023) and our XGBoost ages suffer from similar (but independent) systematics. Part of these systematics are caused by the coldest members of young clusters (i.e. the largest giants), which is in part expected, since the spectroscopic abundances of young stars (<500 Myr) can be significantly biased (see e.g. Spina et al. 2022, Sect. 5.6).

#### 4.4. Machine-learning age estimates for APOGEE DR17

The lower panels of Fig. 6 show comparisons of our age estimates with three independent attempts at producing field-star age estimates for APOGEE DR17 red-giant stars, also using machine-learning techniques and APOGEE-*Kepler* data as their training set. In all panels of Fig. 6 we show comparisons to our uncalibrated age scale (the differences between calibrated and uncalibrated ages are marginal).



**Fig. 6.** Comparisons of our uncalibrated age estimates with ages from the literature. The lines and shaded bands in each panel delineate the running median and  $1\sigma$  quantiles, respectively. Top left panel: age estimates from recent asteroseismic catalogues (obtained from shorter time series than the *Kepler* data), as indicated in the legend. Top right: open cluster ages as listed in the catalogue of [Cantat-Gaudin et al. \(2020\)](#). Grey dots correspond to individual cluster members (with membership probabilities  $>0.95$ ), while the red errorbars show the median statistics in each cluster with more than two members (the ones older than 4 Gyr are annotated). Bottom row: machine-learning field star estimates for APOGEE DR17 using the same training dataset as used in this paper ([Miglio et al. 2021](#)). Bottom left: DR17 *astroNN* ages ([Leung & Bovy 2019](#)). Bottom middle: age estimates from [Ciucă et al. \(2023\)](#). Bottom right: age estimates obtained by [Leung et al. \(2023\)](#) using a random-forest regressor trained on the APOGEE DR17 spectral latent space.

[Leung & Bovy \(2019\)](#) presented a neural-network algorithm, *astroNN*, that is capable of determining stellar atmospheric parameters, elemental abundances, and additional desirables such as distance and age from APOGEE spectra. The DR17 *astroNN* catalogue is available at the SDSS DR17 website<sup>6</sup>.

The age estimation of *astroNN* follows the procedure of [Mackereth et al. \(2019\)](#) and uses a similar training set as ours. Instead of the [Miglio et al. \(2021\)](#) ages, however, the *astroNN* ages are based on the APOKASC-2 catalogue ([Pinsonneault et al. 2018](#)), complemented with the 96 low-metallicity stars of [Montalbán et al. \(2021\)](#). In view of the similarity of the training sets (and the overall philosophy of obtaining

age estimates from spectroscopy alone), it is not surprising that the comparison shown in Fig. 6 (lower left panel) shows much less scatter, albeit with a significant systematic trend (which is most likely inherited from the different age scale in their training set): except for the youngest stars, our ages are typically larger than the *astroNN* estimates by a factor of  $\approx 1.2$ . This, however, is reassuring, since the age plateau present in the *astroNN* catalogue at  $\lesssim 9$  Gyr is inconsistent with the age of the Galactic disc (which is  $>12$  Gyr; e.g. [Fuhrmann et al. 2017](#); [Rendle et al. 2019](#)).

In a different attempt at determining age estimates for APOGEE DR17 red giants (but also using the data of [Miglio et al. 2021](#) as a training set), [Ciucă et al. \(2023\)](#) used a Bayesian neural-network architecture, BINGO ([Ciucă et al. 2021](#)), to infer precise stellar age estimates for 68 360 stars with exquisite APOGEE signal-to-noise ratios (SNREV  $> 100$ ).

<sup>6</sup> [https://www.sdss4.org/dr17/data\\_access/value-added-catalogs](https://www.sdss4.org/dr17/data_access/value-added-catalogs)



As input, Ciucă et al. (2023) use the stellar parameters  $T_{\text{eff}}$  and  $\log g$  as well as the abundances [Fe/H], [Mg/Fe], [C/Fe] and [N/Fe], each with their associated uncertainties. Following the method of Das & Sanders (2019), the hidden parameters (a.k.a. weights and biases) of a neural network with two fully-connected layers were optimised during the training phase. Then, by marginalising over the posterior distribution of the neural network parameters, a posterior age distribution for each individual star was obtained.

The lower middle panel of Fig. 6 shows the comparison of our spectroscopic age estimates with the ages obtained by Ciucă et al. (2023). Except for the small jump at around  $\approx 1.8$  Gyr (which is partly due to low statistics in that age regime), we find a remarkable concordance between the two age scales, with little systematic differences (+0.24 Gyr) and scatter (1.34 Gyr). The formal uncertainties of the BINGO ages are also similar to ours.

Very recently, Leung et al. (2023) presented a new catalogue of APOGEE DR17 age estimates that addresses some of the issues present in the *astroNN* catalogue. In particular, they present the implementation of a novel technique that they dub “variational encoder-decoder”, which reduces the dimensionality of the APOGEE spectra to a four-dimensional latent space that does not contain strong elemental-abundance information. In a second step, Leung et al. (2023) trained a random-forest regressor to learn the relationship between the latent-space parameters and stellar ages, using the APOGEE-*Kepler* data of Miglio et al. (2021). They found that this method delivers more precise and accurate results than their earlier *astroNN* attempt, delivering age uncertainties of  $\sim 22\%$ . Perhaps the main advantage of their method over other attempts (including ours) is that their ages do not depend on [Mg/Fe] (see Sect. 4.5).

The lower right panel of Fig. 6 compares our spectroscopic age estimates with the estimates of Leung et al. (2023). In the young regime ( $< 4$  Gyr), we find very good agreement and little dispersion between the two age scales. For the intermediate and old regime, both systematic differences and dispersion increase, so that the overall concordance between the two methods is worse than in the case of Ciucă et al. (2023), and we also see more dispersion than in the comparison to the *astroNN* ages. This is to some degree expected, because the method of Leung et al. (2023) explicitly excludes the use of part of the information contained in the stellar spectrum (also indicated by the larger uncertainties; see error bars in Fig. 6). Nevertheless, the comparison is reassuring, in particular with regards to the younger stars where we see very little systematic trends.

#### 4.5. Caveats

Some caveats have to be taken into account when using our age estimates. Most importantly, the age estimates critically depend on the quality of the APOGEE DR17 atmospheric parameters and abundances. Since our fiducial method does not provide individual uncertainty estimates, the user may decide whether to apply further quality cuts to the ones already imposed.

We are extrapolating the relationships found in the *Kepler* data to the whole Galactic volume covered by the APOGEE DR17 red-giant sample. This could potentially be dangerous as well, and it is to some degree surprising that it works so well. The reason for that is most certainly stellar migration: the stellar population of the *Kepler* field has such a great variety in ages and abundances because its stars came from vastly different birth positions in the Galactic disc (see e.g. Minchev et al. 2013; Lagarde et al. 2021; Miglio et al. 2021; Ratcliffe et al. 2023), so that this caveat is milder than it could have been.

By construction, our age estimates are not independent of chemical abundances (most importantly, [C/Fe], [N/Fe], and [Mg/Fe]). This tendentially makes inferences of the age-abundance relations slightly circular, at least for the elemental abundance ratios that are most important for the XGBoost model.

Despite the large sample size and coverage of the Galactic disc, our sample is affected by some important selection effects. Most importantly, it is restricted to stars with [Fe/H]  $> -1$ , and therefore to disc stars. A detailed comparison of our sample with Milky Way models will require careful forward modelling to take into account these selection effects.

While traditional chemical clocks such as  $[\alpha/\text{Fe}]$  or  $[\text{Y}/\text{Mg}]$  should be quite robust also for close binaries, post-common-envelope phase binaries, and binary merger products, our age estimates also rely on stellar parameters and chemical abundances that are heavily affected by non-binary evolution. This is the reason why, despite the fact that we eliminated the conspicuous “young  $\alpha$ -rich” stars (a.k.a.  $[\alpha/\text{Fe}]$ -enhanced over-massive stars) from the training set, we still find some of them in the full APOGEE sample. In particular, we find 706 stars (0.4% of the total sample) with [Mg/Fe]  $> 0.2$  and spectroscopic ages  $\leq 6$  Gyr. These stars are very likely products of binary evolution; their ages should not be used. Another small part of our dataset (561 stars) that is most likely affected by effects of binary evolution are rapidly rotating ( $v \sin i > 10 \text{ km s}^{-1}$ ) red giants – they were recently analysed using APOGEE DR16 by Patton et al. (2023). We include flags for both these categories in our results, along with flags for stars that are slightly bluer (4137 stars) or redder (280 stars) than the bulk of the training set (and that thus may have less reliable ages). There are 5383 flagged stars in total.

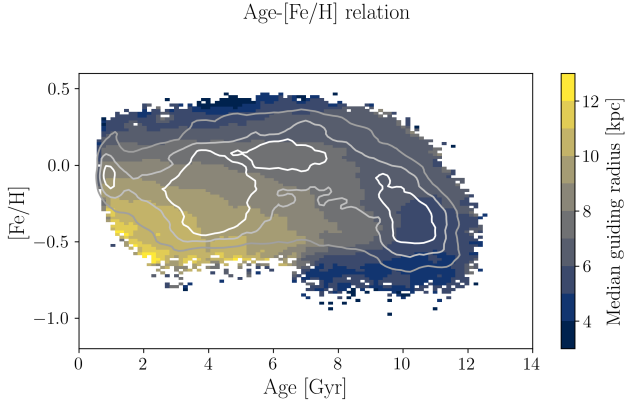
## 5. Results

We applied our XGBoost model to 193 478 APOGEE DR17 stars with SNREV  $> 50$ , [Fe/H]  $> -1$ , and clean abundance flags (see Sect. 2) located in the small  $T_{\text{eff}} - \log g$  box highlighted in Fig. 1, thus avoiding extrapolation of our results into a regime of stellar parameters not covered by the training set. After cleaning for duplicate APOGEE\_IDs (keeping the results with the highest SNREV), we provide spectroscopic age estimates for 178 825 unique APOGEE DR17 stars. To further ensure that these results are meaningful, we show a number of validation plots that reproduce well the expected chemical, positional, and kinematic trends with age. When interpreting these plots, we recall that our absolute age scale is likely subject to some systematics, as illustrated in Sect. 4. For a detailed analysis of the chemical evolution of the Galactic disc with this sample, including a proper treatment of stellar radial migration, we refer the reader to Ratcliffe et al. (2023).

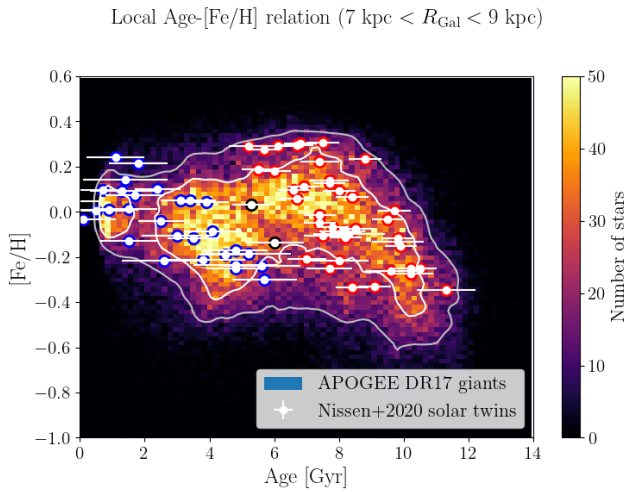
### 5.1. Age-metallicity relation

An important abundance ratio that was not used on purpose in our default XGBoost model is [Fe/H]. We can therefore analyse the age-metallicity relation of our sample without too much fear of circularity (although one could plausibly argue that [Fe/H] is also directly related with the other abundance ratios; see e.g. Ness et al. 2019).

Figure 7 shows the age-metallicity relation (AMR) of our whole APOGEE sample, colour-coded by median guiding-centre radius (estimated as  $R_{\text{guide}} = R_{\text{Gal}} \cdot v_{\Phi}/v_{\text{circ}}$ ), assuming a flat rotation curve ( $v_{\text{circ}} = 229.76 \text{ km s}^{-1}$ ; Schönrich et al. 2010; Bovy et al. 2012) and using proper motion measurements



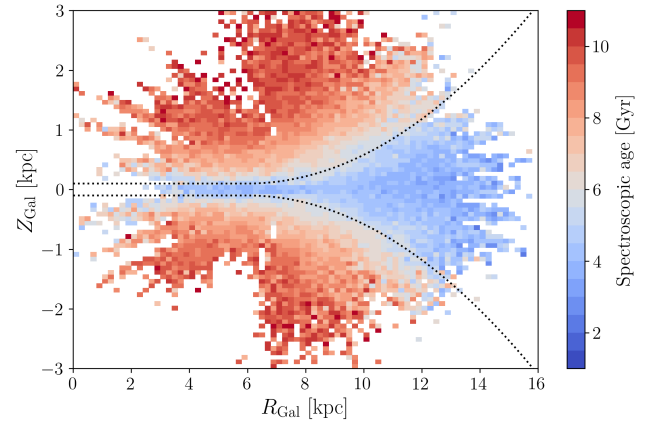
**Fig. 7.** Age–[Fe/H] relation, colour-coded by median guiding radius and smoothed with a Gaussian kernel. Overplotted are iso-density contours corresponding to 20, 40, and 60 stars per bin. Only bins containing at least three stars are shown.



**Fig. 8.** Age–[Fe/H] relation for the solar vicinity ( $7 \text{ kpc} < R_{\text{Gal}} < 9 \text{ kpc}$ ,  $|Z_{\text{Gal}}| < 1 \text{ kpc}$ ), as a 2D histogram and as smoothed iso-density contours corresponding to 10 and 30 stars per bin. Overplotted are the 72 solar twins studied by Nissen et al. (2020), with the edge colour of the symbols corresponding to the groups defined by their Fig. 3.

from *Gaia* DR3 (Gaia Collaboration 2023b). Density contours are overplotted. The overall picture is consistent with both predictions from classical multi-zone GCE models (e.g. Chiappini 2009), GCE models with radial mixing (e.g. Minchev et al. 2014; Kubryk et al. 2015; Johnson et al. 2021), and cosmological zoom-in simulations of Milky-Way-like galaxies (e.g. Renaud et al. 2021; Lu et al. 2022b): the colour-code clearly indicates a gradual inside-out formation of the disc within the last  $\sim 7$  Gyr, with a rapid enrichment in the inner disc (see also Joyce et al. 2023), and a rather recent onset of star formation in the outermost parts of the disc.

Restricting ourselves to the extended solar neighbourhood ( $7 \text{ kpc} < R_{\text{Gal}} < 9 \text{ kpc}$ ,  $|Z_{\text{Gal}}| < 1 \text{ kpc}$ ), we obtain the local AMR shown in Fig. 8. It resembles the one reconstructed from GALAH data by Sahlholdt et al. (2022, see also Queiroz et al. 2023, Fig. 18) and confirms the bimodality in the AMR hinted by Nissen et al. (2020) and later corroborated by the works of Jofré (2021) and Xiang & Rix (2022). Following the terminology of the latter authors, the younger of the two branches in the AMR corresponds to the late, dynamically quiescent phase of disc evolution,



**Fig. 9.** Median age as a function of position in Galactocentric coordinates. Only bins containing more than three stars are shown, no smoothing is applied. The dotted lines trace approximately the region in which the median age equals to 5 Gyr as a quadratic function of  $R_{\text{Gal}}$ , clearly showing the effect of the young disc’s flare.

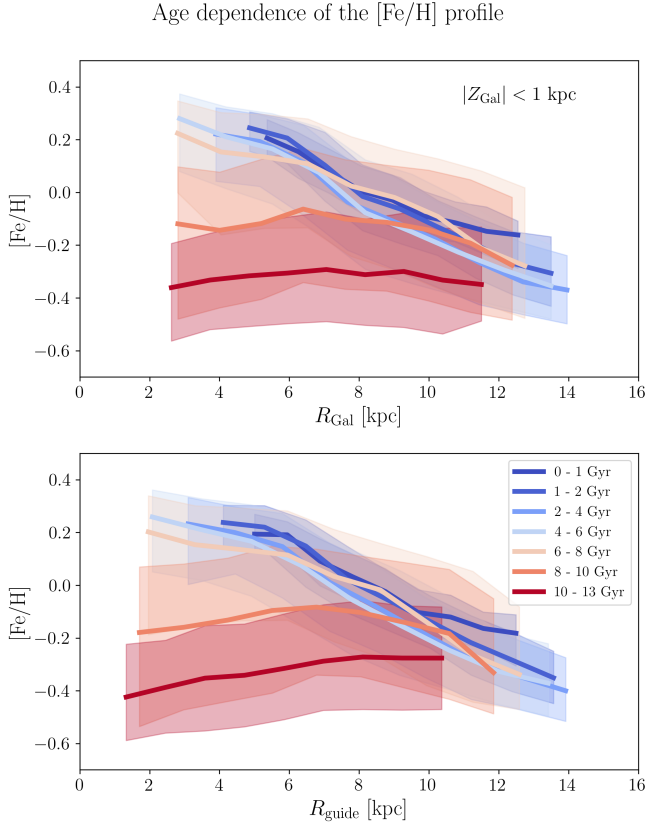
while the older one comprises the in-situ halo (not present in our sample due to the  $[\text{Fe}/\text{H}] > -1$  criterion) and the old,  $[\alpha/\text{Fe}]$ -enhanced disc. This is in line with traditional scenarios of dual disc formation (e.g. Chiappini et al. 1997; Fuhrmann et al. 2017). The fact that we clearly see the split in the AMR further validates the high precision of our age estimates.

However, we remind ourselves that the AMRs shown in Figs. 7 and 8 are a convolution of (i) the underlying evolution of the ISM in the Milky Way disc over time, (ii) stellar radial migration, (iii) observational uncertainties (most importantly in terms of age), and (iv) the selection function of our sample. Therefore, deciphering the past chemical-enrichment history is an endeavour that requires taking into account radial migration (e.g. Minchev et al. 2018; Frankel et al. 2018, 2019, 2020; Beraldo e Silva et al. 2021; Lu et al. 2023; Ratcliffe et al. 2023).

## 5.2. Spatial age trends

Another interesting plot is the median age map in cylindrical ( $R_{\text{Gal}}$  vs.  $Z_{\text{Gal}}$ ) coordinates shown in Fig. 9. When interpreting this plot, we recall that in this work we have been extrapolating the Milky Way disc’s intrinsic age-abundance relations from the *Kepler* field to the whole range of the Galaxy covered by our sample ( $0 \lesssim R_{\text{Gal}} \lesssim 16 \text{ kpc}$ ). We also remember that selection effects have not been taken into account in the creation of this map (nor in any other of the shown plots).

Even with these caveats, Fig. 9 clearly shows the vertical age gradient we expect to see, and also the clear signature of the flared outer disc (as illustrated by the dotted lines that follow the 5 Gyr isochrone in this plot with a quadratic function starting at  $R_{\text{Gal}} \sim 6 \text{ kpc}$ ). In fact, not only the younger populations display flaring, but mono-age populations are expected to show nested flares, as argued by Minchev et al. (2015). The prediction of their model was a strong age gradient in the morphological thick disc ( $|Z_{\text{Gal}}| > 0.5 \text{ kpc}$ ) that was indeed found in earlier APOGEE data (Martig et al. 2016b, see also Imig et al. 2023). We can now confirm this result with much larger statistics and extent in  $R_{\text{Gal}}$ . This is in contrast to associating age with high- $[\alpha/\text{Fe}]$  mono-abundance populations (MAPs), which show no flaring (Bovy et al. 2016, but see Lian et al. 2022) due to significant negative age gradients in a given MAP (Minchev et al. 2017).

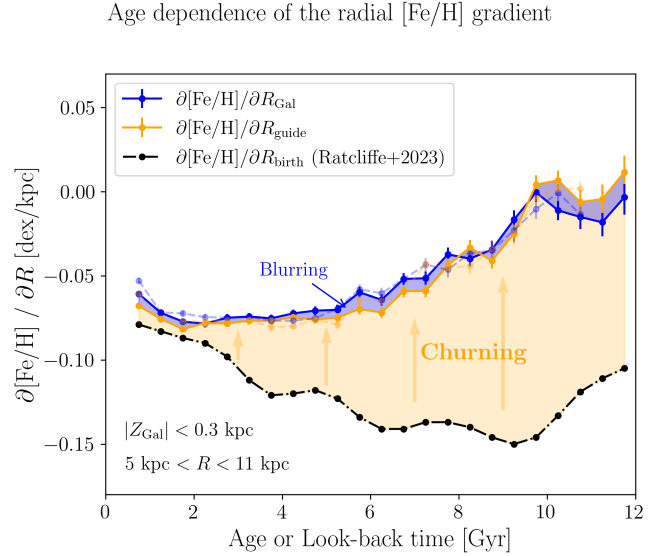


**Fig. 10.** Radial [Fe/H] profile in bins of age. Top panel: with respect to  $R_{\text{Gal}}$ . Bottom panel: with respect to the guiding-centre radius,  $R_{\text{guide}}$ . The youngest bin contains 2807 stars, all other bins more than 9000 stars.

Mackereth et al. (2017) studied the age-metallicity structure of the Milky Way disc using 31 000 APOGEE DR12 stars with [C/N] age estimates inferred by Martig et al. (2016a). The authors fitted parametric density profiles to the (volume-corrected) APOGEE star counts and found radially broken exponential radial profiles as well as flared exponential vertical profiles. The same behaviour could be reproduced by Beraldo e Silva et al. (2020) with a simulation including clumpy star formation in the early epochs of the disc. This big picture is still consistent with the APOGEE DR17 data (but see Sysoliatina & Just 2022).

### 5.3. Galactic radial [Fe/H] profile as a function of age

The Galactic radial abundance profile is an important observable that helps to constrain the chemical evolution of the Milky Way (e.g. Trimble 1975; Tinsley 1980). First discovered in nebular tracers (Hawley 1978; Peimbert & Serrano 1980; Afflerbach et al. 1997), the presence of a negative radial abundance gradient in the Milky Way was confirmed in field stars soon thereafter (Janes 1979; Luck & Bond 1980). Since the 1980s, studies of the Galactic radial abundance profile as a function of age thrived, as larger samples of tracers with distance, age, and abundance were compiled. Most importantly, open clusters (e.g. Panagia & Tosi 1981; Lyngå 1982; Friel 1995; Friel et al. 2002; Magrini et al. 2009) and planetary nebulae (e.g. Maciel et al. 2005; Stanghellini & Haywood 2010, 2018) were used to infer the age dependence of the Milky Way’s abundance profile, sometimes with contradictory results. With the advent



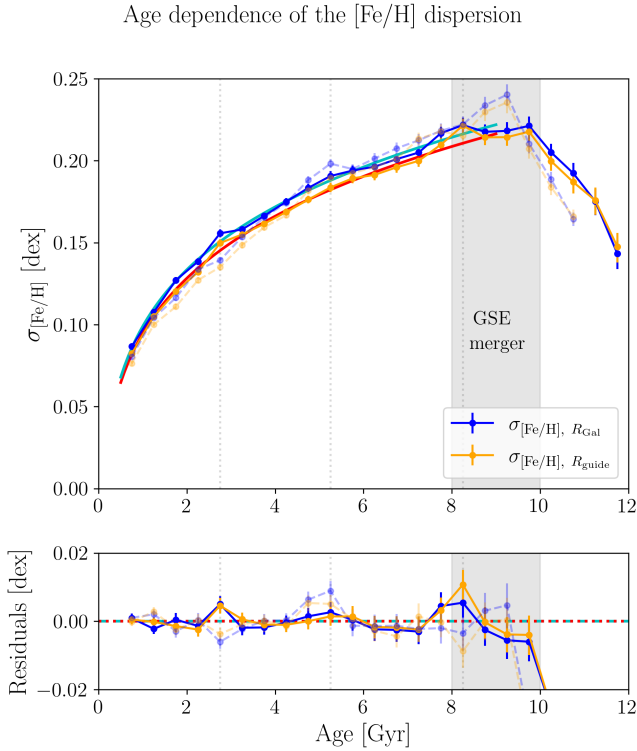
**Fig. 11.** Age dependence of the radial [Fe/H] abundance gradient in terms of present-day Galactocentric distance (blue) and guiding-centre radii (orange). Each point was obtained by 3-parameter (slope, intercept, and dispersion) Bayesian fits to the [Fe/H]– $R$  distribution, using only data in the respective age bin, restricted to  $|Z_{\text{Gal}}| < 0.3$  kpc and  $5 \text{ kpc} < R_{\text{Gal/guide}} < 11$  kpc (39 920 and 40 629 stars, respectively). The faint dashed lines and errorbars correspond to the gradients when using the alternative set of ages obtained with XGBoost quantile regression (see Appendix A). In black we overplot the radial [Fe/H] gradients in the interstellar medium across look-back time obtained by Ratcliffe et al. (2023), calculated for the same age bins. The difference between the blue and the orange line (blue-shaded region) can be attributed to radial heating (blurring), while the difference between the orange and the black line (orange-shaded region) highlights the influence of radial migration (churning). Beyond  $\sim 10$  Gyr, the gradients are consistent with zero.

of large spectroscopic surveys, it also became possible to use larger samples of field stars with isochrone ages to infer the age dependence of the abundance gradient (e.g. Nordström et al. 2004; Casagrande et al. 2011; Xiang et al. 2015).

In parallel, also Galactic chemical-evolution modelling became more important and helped to explain the metallicity-gradient observations, typically by assuming an inside-out formation of the Galactic disc (e.g. Ferrini et al. 1994; Koeppen 1994; Hou et al. 2000; Chiappini et al. 2001; Naab & Ostriker 2006). Nevertheless, the model comparison to different tracers can lead to opposite conclusions, since each tracer population is affected by different biases (in terms of age, distance, abundance, and selection effects). A more comprehensive review on the history of determinations of the age dependence of the radial metallicity gradient can be found in Sects. 1 and 6 of Anders et al. (2017a).

Anders et al. (2017a) studied the age dependence of the radial metallicity profile close to the Galactic plane ( $|Z_{\text{Gal}}| < 0.3$  kpc) using a sample of 418 red-giant stars observed by both APOGEE (DR12; Alam et al. 2015) and the CoRoT satellite (Baglin et al. 2006). Thanks to our machine-learning approach we have now dramatically increased the sample size of field red giants with statistically meaningful ages to  $> 100\,000$ , obtaining a much less geometrically biased sample. We therefore reanalyse the age dependence of the Galactic radial metallicity profile in Figs. 10 and 11.





**Fig. 12.** Age dependence of the [Fe/H] dispersion around the radial [Fe/H] abundance gradient shown in Fig. 11. Simple power-law fits to the data (for ages  $<7$  Gyr) are overplotted, with power-law coefficients 0.12 and 0.15 for the dispersions about the present-day abundance gradient (red) and the abundance gradient with respect to the guiding radius (cyan), respectively. The lower panel shows the residuals with respect to these fits. As in Fig. 11, the faint dashed lines and error bars correspond to the gradients when using the alternative set of ages obtained with XGBoost quantile regression. The vertical lines highlight regions of departure from the power-law trend, possibly connected to episodes of enhanced star formation (or enhanced radial migration). The grey-shaded area marks the age interval in which we expect to see signatures from the Gaia Sausage Enceladus (GSE) merger event.

Figure 10 shows the radial [Fe/H] profile of the APOGEE red-giant sample in seven broad age bins. In the top panel we display the [Fe/H] profiles (running median and  $1\sigma$  quantiles) as a function of current Galactocentric distance, while in the bottom panel we show the same as a function of guiding-centre radius. The differences between both panels, at least for the younger populations, are very subtle.

The radial metallicity distribution measured for the youngest populations (0–1 Gyr, 1–2 Gyr, 2–4 Gyr bins) are dominated by a strong linear ( $\partial[\text{Fe}/\text{H}]/\partial R \approx -0.06 \text{ dex kpc}^{-1}$ ) gradient, with a typically symmetric dispersion that gradually increases with age. The gradient of the 1–5 Gyr populations is also slightly steeper (by  $0.015 \text{ dex kpc}^{-1}$ ) than the gradient of the youngest bin ( $<1$  Gyr), in agreement with previous results from field stars (e.g. Nordström et al. 2004; Casagrande et al. 2011; Anders et al. 2017a; Wang et al. 2019) and recent open-cluster studies (Spina et al. 2021; Netopil et al. 2022; Myers et al. 2022; Gaia Collaboration 2023a), while in slight tension with other studies that found a rather monotonic flattening of the metallicity gradient with increasing age (Chen & Zhao 2020; Vickers et al. 2021).

For older ages, it is expected that the radial metallicity gradient (and also the age-velocity dispersion relation, see Sect. 5.4)

for open clusters behaves differently from field stars, due to a survival bias of clusters that migrate outwards (Lyngå & Palouš 1987; Anders et al. 2017a; Spina et al. 2021). Nevertheless, the importance of radial migration for the gradient evolution is still debated in the open-cluster community (Spina et al. 2022; Magrini et al. 2023). In the sample of field stars used in this work, we find that the [Fe/H] gradient significantly flattens (and even inverts when considering  $R_{\text{guide}}$ ) for the oldest age bins (Fig. 10). This weakening in the metallicity gradient, along with increased scatter about the running median, is an expected consequence of radial migration (Kubryk et al. 2013; Minchev et al. 2013), and is in agreement with other works (e.g. Casagrande et al. 2011; Xiang et al. 2015; Anders et al. 2017a).

Unfortunately, the age dependence of the radial abundance gradient does not allow us to directly infer the evolution of the abundance gradient in the interstellar medium. In fact, just as in the case of the AMR discussed in Sect. 5.1, the abundance profiles shown in Fig. 10 are a convolution of Galactic chemical evolution, stellar mixing (heating and migration, or “blurring” and “churning” in the terminology of Schönrich & Binney 2009), observational uncertainties, and selection effects.

Figure 11 shows the result of Bayesian linear fits (see Appendix D for details) to the radial [Fe/H] profiles in smaller age bins and restricted to a thinner slice of the Galactic disc ( $|Z_{\text{Gal}}| < 0.3 \text{ kpc}$ ) and an  $R$  range between 5 kpc and 11 kpc. As in Anders et al. (2017a), these fits account for both the observational uncertainties in [Fe/H] and an intrinsic abundance scatter about the linear trend. The present-day abundance gradients of mono-age populations in terms of Galactocentric distance ( $\partial[\text{Fe}/\text{H}]/\partial R_{\text{Gal}}$ ) are shown in blue, while the gradients in terms of guiding-centre radius ( $\partial[\text{Fe}/\text{H}]/\partial R_{\text{guide}}$ ) are shown in orange. For illustration, we also show the [Fe/H] gradient with respect to the birth radius inferred by Ratcliffe et al. (2023), using the same age bins. These values can be thought of as a reconstruction of the Galactic radial abundance gradient in the interstellar medium (ISM) at a certain look-back time<sup>7</sup>. As indicated by the blue-shaded region in Fig. 11, the difference between the blue and the orange curves is produced by radial heating (blurring) of stellar orbits, flattening the observed radial abundance gradient only slightly in each age bin. On the other hand, the much larger difference between the black and the orange curve is produced by radial migration (churning), as highlighted by the orange-shaded area.

The picture drawn in Fig. 11 depends slightly on the exact age scale used, as exemplified by the fits using the XGBoost quantile regression ages (see Appendix A) shown as faint dashed lines in that figure. It is also affected by the systematics discussed in Sect. 4, and in particular less trustworthy for older ages (both in terms of the age estimates and the estimated birth radii and abundance gradients; see e.g. Lu et al. 2022a). However, we see that radial migration is indeed a main driver in shaping the observed Galactic abundance gradient for mono-age populations, even for relatively small ages ( $<1$  Gyr), and becoming more and more important for older ages (e.g. for ages older than 10 Gyr half of the solar neighbourhood’s thin-disc stars should have migrated outward; Beraldo e Silva et al. 2021). This picture is also in excellent agreement with the main result of

<sup>7</sup> The reconstruction of the radial abundance gradients as a function of look-back time is based on the method of Lu et al. (2023). These authors found, using different suites of cosmological hydrodynamical simulations of Milky-Way-like galaxies, a tight anticorrelation between the ISM gradient at a certain look-back time and the 5%–95% inter-percentile range of stellar metallicities found in the corresponding age interval.

Frankel et al. (2020) who concluded that the secular orbit evolution in the Galactic disc is dominated by diffusion in angular momentum, with radial heating being less important by one order of magnitude. A deeper analysis of the actual evolution of the Galactic abundance gradients (not only for [Fe/H] but also for other elements) using our age estimates is presented in Ratcliffe et al. (2023), revealing substructure in the evolution of the abundance gradients that possibly correspond to enhanced star formation episodes triggered by passages of the Sgr dSph satellite and the GSE (see the wiggles in the black curve in Fig. 11).

Our Bayesian fits of the radial metallicity gradient between 5 and 11 kpc in age bins of 0.5 Gyr (shown in Fig. 11) also deliver precise estimates of the [Fe/H] dispersion about the gradient. Figure 12 shows these dispersion measurements as a function of age, again for both cases, with respect to  $R_{\text{Gal}}$  and  $R_{\text{guide}}$ . As expected (and first demonstrated by Spina et al. 2021 for open clusters), the dispersions about the radial abundance gradient with respect to the guiding radius are consistently smaller by a small amount (typically 0.01–0.02 dex).

In addition, thanks to the dramatically increased sample size and therefore much smaller statistical uncertainties, we can now quantify the behaviour of the [Fe/H] dispersion over time much better than in Anders et al. (2017a). We find that the increase of the abundance dispersion between 1 and 7 Gyr can be well fit by a power law with an exponent around  $\sim 0.15$  (solid coloured lines in Fig. 12). For even younger ages ( $< 500$  Myr), this relation is expected to flatten, since the intrinsic [Fe/H] dispersion of younger tracers is finite and not much smaller than 0.05 dex (Genovali et al. 2014; Wenger et al. 2019; Spina et al. 2021). However, our red-giant sample does not cover this age range. The increase in abundance scatter with age is obviously created by radial migration (e.g. Haywood 2008; Schönrich & Binney 2009; Minchev et al. 2014; Kubryk et al. 2015). The smooth functional form of this increase with age suggests that radial migration (churning) can be described by a sub-diffusive process with a constant efficiency over a large phase of the Galactic disc's lifetime (see also Anders et al. 2020).

Hints of departures from the power-law trends in Fig. 12 are seen in the age bins 2.5–3 Gyr, 5–5.5 Gyr, and possibly 8–8.5 Gyr. For example, an enhanced [Fe/H] dispersion with respect to the overall trend can be appreciated around 2.5–3 Gyr (when using our fiducial age scale). Interestingly, this age corresponds to an epoch for which an enhanced star formation rate has been reported by several studies using *Gaia* (Bernard 2018; Mor et al. 2019; Isern 2019; Alzate et al. 2021) and spectroscopic survey data (Johnson et al. 2021; Sahlholdt et al. 2022; Spitoni et al. 2023). Another small enhancement in  $\sigma_{[\text{Fe}/\text{H}], R_{\text{Gal}}}$  and  $\sigma_{[\text{Fe}/\text{H}], R_{\text{guide}}}$  seems to occur in the 5–5.5 Gyr bin (in our age scale), coinciding with an epoch of enhanced star formation proposed in the literature (Ruiz-Lara et al. 2020; Alzate et al. 2021; Sahlholdt et al. 2022), possibly related to the first passage of the Sgr dSph galaxy (Ruiz-Lara et al. 2020). Recent simulations indeed suggest that Sgr passages can temporally enhance migration (Carr et al. 2022). As also demonstrated in Fig. 12, however, these small departures from the main power-law trend are not always robust to a change in the used age scale.

The grey-shaded area in Fig. 12 marks the approximate time frame in which the last significant merger event, the *Gaia* Sausage Enceladus (GSE) merger (Helmi et al. 2018; Belokurov et al. 2018), probably affected the Milky Way disc. For example, Chaplin et al. (2020) found an upper limit of  $> 11.0 \pm 1.5$  Gyr by age-dating the in-situ halo star  $\nu$  Ind. Montalbán et al. (2021) tightened this constraint further using

detailed asteroseismic modelling of 96 old stars observed with *Kepler* and found that the merger happened  $\sim 10$  Gyr ago.

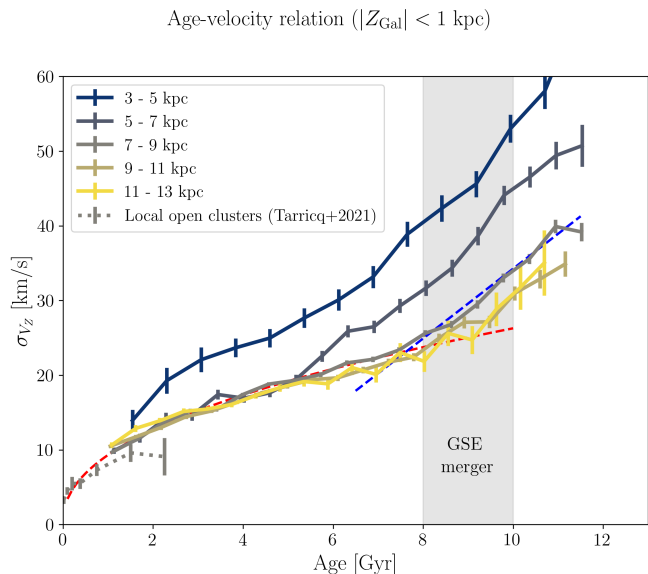
In the oldest regime ( $\geq 9$ –10 Gyr in our age scale), we see a sharp decrease of  $\sigma_{[\text{Fe}/\text{H}]}$  with age (i.e. an increase with cosmic time). This could potentially be related to the turbulent formation of the early Milky Way disc, either through numerous mergers (e.g. Brook et al. 2004, 2005) or by internal interactions in turbulent and clumpy discs (e.g. Bournaud et al. 2009; Forbes et al. 2012; Amarante 2020; Bird et al. 2021), or both (Belokurov & Kravtsov 2022). However, the expectation from recent Milky-Way-like simulations such as FIRE-2 (Wetzel et al. 2023), NIHAO-UCD (Buck et al. 2020), or Auriga (Grand et al. 2017) is rather that rapid variations in star formation, outflow rate, and depletion time during early stages of evolution result in large variations in the elemental abundances of old stars, especially for  $[\text{Fe}/\text{H}] < -1$  (Belokurov & Kravtsov 2022, Fig. 19 – although their analysis is not strictly focused on the disc and concerns the total metallicity spread, not subtracting the gradient). Another option is that the GSE merger triggered a burst in star formation in the Galactic disc (An et al. 2023; Ciucă et al. 2023), as a result of which we see an enhanced [Fe/H] dispersion about the radial abundance gradient of the  $\sim 8$ –9 Gyr red-giant population that then drops gradually. We note, however, that the enhanced [Fe/H] dispersion can also be produced by a temporally enhanced radial migration efficiency (again related to the merger). We also recall that in the old age regime our disc sample can potentially be biased by the metallicity range for which our age estimation is valid ( $[\text{Fe}/\text{H}] > -1$ ). Again, a full forward modelling (and/or a training set including enough metal-poor stars) is necessary to reliably extend this analysis into the oldest regime.

#### 5.4. Age-velocity dispersion relation

While our age estimates critically depend on the quality of the APOGEE stellar parameters and abundances (except for [Fe/H]), they are completely independent from kinematics and position in the Galaxy. As an example for a kinematic relationship with age that is often used in Galactic archaeology, we show in Fig. 13 the dispersion in vertical velocity  $v_z$  as a function of age (here restricted to  $|Z_{\text{Gal}}| < 1$  kpc) – often called the age-velocity relation (AVR).

Since Gliese (1956) and von Hoerner (1960) found significant differences in the motions of nearby stars as a function of their spectral type, the AVR in the solar neighbourhood is an important observable that constrains the vertical heating history of the Milky Way's disc. It has been studied by many authors using a variety of tracers in the past 50 years (e.g. Wielen 1974; Mayor 1974; Byl 1974; Meusinger et al. 1991; Nordström et al. 2004; Casagrande et al. 2011; Aumer et al. 2016; Ting & Rix 2019; Raddi et al. 2022).

Figure 13 shows the AVR of our APOGEE DR17 sample in 2 kpc wide bins of Galactocentric distance between 3 kpc and 13 kpc, with each bin containing more than 14 000 stars (except for the 3–5 kpc bin with 6000 stars and a rather distinct AVR behaviour, probably due to the influence of the Galactic bar; Saha et al. 2010; Grand et al. 2016). Since the age range of our sample does not cover the youngest stars, we also plot the AVR recently obtained by Tarricq et al. (2021) using a sample of 418 *Gaia*-confirmed OCs with radial velocities in the solar neighbourhood. The authors found that the AVR in the young regime can be described by a power law with  $\beta \approx 0.2$ . Although the OC sample has thus much lower statistics and is possibly affected by the cluster destruction bias in the 2–3 Gyr bin (Anders et al. 2017a; Spina et al. 2022), we see that the trend



**Fig. 13.** Age – velocity dispersion relationship (AVR), in bins of Galactocentric distance, for stars close to the Galactic plane ( $|Z_{\text{Gal}}| < 1$  kpc). Also plotted are the results of Tarricq et al. (2021) for open clusters in the solar vicinity. The red dashed line corresponds to a simple power-law fit ( $\beta = 0.46$ ) for ages  $< 7$  Gyr in the Galactocentric distance bin 7–9 kpc, while the blue dashed line corresponds to a linear fit for ages between 7 and 11 Gyr in the same  $R_{\text{Gal}}$  bin. The shaded region highlights the age range in which we see a steepening in the AVR, potentially related to the GSE merger event.

of growing velocity dispersion continues in our field star sample, regardless of Galactocentric distance. Indeed, we find that the trend in our age range between 0.5 Gyr and 7 Gyr can be well described with a power law for  $R_{\text{Gal}} > 5$  kpc (in accordance with e.g. Aumer et al. 2016; Ting & Rix 2019 or the simulation of Agertz et al. 2021), but with a notably different power-law index from the OC sample ( $\beta \approx 0.46$  for the 7–9 kpc bin; red dashed line in Fig. 13).

In part, the higher velocity dispersion for older stars in the inner disc seen in Fig. 13 is expected from the exponential disc profile in galactic discs (for the case of a flat rotation curve and hydrostatic equilibrium:  $\sigma_{v_z}(R) \propto \sqrt{\Sigma(R) \cdot h_z}$ ; e.g. Eq. (7) in van der Kruit & Freeman 2011). We also see hints (in the smallest age bin) of the expected trend that stars in the outer disc are born with a higher velocity dispersion (due to the smaller vertical force acting on the gas in the outer disc). However, the velocity dispersions in the intermediate age regime ( $\sim 2$ –6 Gyr) are surprisingly congruent from 5 kpc out to 13 kpc (different from what is seen in modern simulations; e.g. Fig. 14 in Agertz et al. 2021), which possibly indicates that vertical heating happened with similar efficiency over a large part of the Galactic disc within the past  $\sim 6$ –8 Gyr. We caution, however, that age errors will always tend to smooth out any abrupt changes in the AVR (e.g. Martig et al. 2014).

For  $R_{\text{Gal}} > 7$  kpc populations older than  $\sim 8$  Gyr, we see evidence for a departure from the power-law behaviour observed between 0.5 and 7 Gyr, in the sense that the AVR steepens (in the 5–7 kpc bin, a first steepening happens already around 5.5 Gyr in our age scale). This change of trend is exemplified by the blue dashed line in Fig. 13, which corresponds to a linear fit of the AVR in the 7–9 kpc bin restricted to the age range between 7 and 11 Gyr. Similar broken trends in the AVR can be seen in all  $R_{\text{Gal}}$  bins. Except for the innermost  $R_{\text{Gal}}$  bin, the exact onset of

the steepening appears to depend slightly on  $R_{\text{Gal}}$ . Similar trends for the solar vicinity have already been observed previously (e.g. Meusinger et al. 1991; Quillen & Garnett 2000; Yu & Liu 2018). Indeed, Quillen & Garnett (2000) already found, using the data of Edvardsson et al. (1993) together with the revised age estimates from Ng & Bertelli (1998) an abrupt increase by a factor of two in the stellar velocity dispersions at an age of  $9 \pm 1$  Gyr and proposed that the Milky Way suffered a minor merger 9 Gyr ago, which created the thick disc. This means that we have not discovered a new feature of the AVR, but confirmed a previously observed trend with much better statistics and individual age precision. In addition, we can now study the AVR over a much larger portion of the Galactic disc, finding tendencies with  $R_{\text{Gal}}$  that can potentially be used for a more detailed comparison to chemo-dynamical models of the Milky Way.

On the simulation side, the AVR has been shown to be affected by both internal and external kinematic processes. For example, Grand et al. (2016) found, using cosmological zoom-in simulations of Milky Way-sized galaxies, that in most cases the Galactic bar is the dominant heating agent, while spiral arms, radial migration, and adiabatic heating play a secondary role. In some of the Grand et al. (2016) simulations, however, the strongest source of vertical heating were external perturbations from massive satellites. Using cosmological simulations, Martig et al. (2014) showed that the AVR depends on the merger history at low redshift, and that jumps in the solar vicinity’s AVR often correspond to (minor) merger events. However, Martig et al. (2014) also showed that in the presence of age errors these jumps are considerably blurred, so that only the largest merger events are detectable (and, depending on the age error, not as jumps but as changes in slope of the AVR). Yurin & Springel (2015) also found that substructures are important for heating the outer parts of stellar discs but do not significantly affect their inner parts. Saha et al. (2010) related the vertical heating exponent in the inner parts of galaxies to the growth rate of the bar potential, while in the outer regions they found that vertical heating was often dominated by transient spiral waves and mild bending waves.

Regarding the break/steepening in the AVR that we see in Fig. 13 around  $\sim 8$  Gyr, some work relying on simulations suggests that this is an effect of a GSE-like merger event (Quillen & Garnett 2000; Martig et al. 2014), while others associate it with the gradual settling of the early, turbulent Galactic disc (Bird et al. 2013; Navarro et al. 2018; Bird et al. 2021; Yu et al. 2021). We suggest that this controversy could potentially be settled with the data presented in this work (that cover an extensive range of the Galactic disc), by comparing the AVR in radial bins to simulations, taking into account selection effects.

The kinematics of mono-age, mono-metallicity APOGEE populations were studied in more detail by Mackereth et al. (2019), albeit with less data (65 719 stars) and not showing a prominent steepening in the AVR. Their results depend in part on the accuracy of the inferred ages (obtained using a Bayesian convolutional neural network and APOGEE DR14 spectra). We note, however, that the AVR, as other relationships with age, is not only blurred by statistical age errors, but also affected by systematic age errors and radial migration. Its full meaning is more clearly revealed when separating stars by their birth radii (Minchev et al. 2018; Bird et al. 2021). We also note that a more robust analytical description of the vertical heating history of the Galactic disc can be obtained by analysing the stellar kinematics in action space (Ting & Rix 2019).



## 6. Conclusions

In this paper we successfully use the XGBoost algorithm to accurately estimate spectroscopic ages for 178 825 APOGEE DR17 red-giant stars located close to the red clump ( $2.2 < \log g < 3.4$ ,  $4400 \text{ K} < T_{\text{eff}} < 5200 \text{ K}$ ) with a median statistical uncertainty of 17%. These estimates are formally independent of  $[\text{Fe}/\text{H}]$ . In accordance with previous studies, we find that our red-giant age estimates are mostly driven by  $T_{\text{eff}}$ ,  $\log g$ ,  $[\text{C}/\text{Fe}]$ , and  $[\text{N}/\text{Fe}]$ . To verify the obtained ages (which are tied to the asteroseismic age scale of Miglio et al. 2021), we compare our results with the age scale of open clusters, asteroseismic+spectroscopic red-giant samples observed with CoRoT, K2, and TESS, as well as other machine-learning age estimates from the literature, finding acceptable agreement, modulo some (expected) systematic trends, with most of the test samples (none of which can be regarded as absolute benchmarks). Our analysis is reproducible via github<sup>8</sup> and the age catalogue (see Appendix A) can be accessed in that repository<sup>9</sup> or via CDS.

We then investigate some chemo-kinematic relationships with stellar age that are often used in Galactic archaeology. We find that the ages are precise enough to confirm the split in the local age-metallicity relation found by Nissen et al. (2020) using high-resolution spectroscopy of solar analogues (Fig. 8). Analysing the Milky Way disc's age map (Fig. 9), we find a clear imprint of the flaring in the outer disc, impressively confirming that our age-estimation method (which is completely independent of sky position and distance) works well for the full range probed by the APOGEE DR17 giant population.

In Sect. 5.3 (Figs. 10–12), we analyse the Milky Way disc's radial metallicity profile. We present new and precise measurements of the Galactic radial metallicity gradient close to the Galactic plane ( $|Z_{\text{Gal}}| < 0.3 \text{ kpc}$ ,  $5 \text{ kpc} < R_{\text{Gal}} < 11 \text{ kpc}$ ) in 0.5 Gyr bins between 0.5 and 12 Gyr, confirming a steeper metallicity gradient for 2–5 Gyr old populations and a subsequent flattening for older populations mostly produced by radial migration (churning). In addition, our fits to the radial  $[\text{Fe}/\text{H}]$  profile allow us to analyse, with unprecedented precision, the dispersion about the radial abundance gradient as a function of age (Fig. 12). We see a clear power-law trend (with a power-law index  $\beta \approx 0.15$ ), indicating a smooth radial migration history in the Galactic disc over the past  $\sim 7$  Gyr. Departures from this power law are detected at ages of  $\approx 9$  Gyr (possibly related to the Gaia Sausage Enceladus merger) and  $\approx 5$  Gyr (with lower significance, possibly related to a passage of the Sgr dSph galaxy; Carr et al. 2022).

Finally, as an example of a kinematic relationship with age, we study the age-velocity dispersion relation (AVR) in 2 kpc wide radial bins around the Galactic centre (Fig. 13). We find that the AVR for ages up to  $\sim 7$  Gyr can be well described by a power law with an exponent around 0.5, and confirm earlier measurements reporting a pronounced steepening of the AVR at around  $\sim 9$  Gyr. Indeed, we find that this behaviour extends over a large extent of the Galactic disc ( $5 \text{ kpc} < R_{\text{Gal}} < 13 \text{ kpc}$ ), while in the inner disc ( $R_{\text{Gal}} < 5 \text{ kpc}$ ) the AVR shows a more complex behaviour that we attribute to efficient heating by the Galactic bar.

Our results reproduce the expected chemical, positional, and kinematic trends with age. This suggests that chemical clocks, and more broadly speaking, weak chemical tagging, is a viable method in Galactic Archaeology. The use of precision asteroseismology coupled with efficient machine-learning algorithms such as XGBoost may allow for greater efficiency and accuracy in estimating ages for millions of stars in the era of large spectroscopic surveys. We recall that some of our conclusions (for

example, the exact time dependence of the radial abundance gradient) are sensitive to our absolute age scale and the systematics it may be suffering from (see Sect. 4). The knowledge of precise and accurate ages, kinematics, and detailed chemical abundances of stars still holds the key to uncovering the formation and evolution of our Galaxy.

**Acknowledgements.** This work was partially funded by the Spanish MICIN/AEI/10.13039/501100011033 and by the “ERDF A way of making Europe” funds by the European Union through grant RTI2018-095076-B-C21 and PID2021-122842OB-C21, and the Institute of Cosmos Sciences University of Barcelona (ICCUB, Unidad de Excelencia “María de Maeztu”) through grant CEX2019-000918-M. F.A. acknowledges financial support from MCIN/AEI/10.13039/501100011033 through grants IJC2019-04862-I and RYC2021-031638-I (the latter co-funded by European Union NextGenerationEU/PRTR). B.R. and I.M. acknowledge support by the Deutsche Forschungsgemeinschaft under the grant MI 2009/2-1. C.C. acknowledges support from the Jesús Serra foundation. J.A.S.A. acknowledges funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 852839). H.P. acknowledges support from FAPESP proc. 2018/21250-9 and 2022/04079-0. This project was developed in part at the Lorentz Center workshop Mapping the Milky Way, held 6–10 February, 2023 in Leiden, the Netherlands. We thank Ioana Ciucă (Canberra) for sharing her APOGEE DR17 BINGO results, and Letizia Stanghellini (Tucson) and Tiancheng Sun (Beijing) for sending comments on the first version of the manuscript. Finally, we thank the anonymous referee for a very useful report that helped improve the quality of the paper. The preparation of this work has made use of TOPCAT (Taylor 2005), NASA's Astrophysics Data System Bibliographic Services, as well as the open-source Python packages xgboost (Chen & Guestrin 2016), astropy (Astropy Collaboration 2018), and NumPy (van der Walt et al. 2011). The figures in this paper were produced with matplotlib (Hunter 2007). This work has made use of data from the European Space Agency (ESA) mission Gaia (www.cosmos.esa.int/gaia), processed by the Gaia Data Processing and Analysis Consortium (DPAC, www.cosmos.esa.int/web/gaia/dpac/consortium). Funding for the DPAC has been provided by national institutions, in particular the institutions participating in the Gaia Multilateral Agreement. 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 web site 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, the Chilean Participation Group, the French Participation Group, Harvard-Smithsonian Center for Astrophysics, Instituto de Astrofísica de Canarias, The Johns Hopkins University, Kavli Institute for the Physics and Mathematics of the Universe (IPMU)/University of Tokyo, 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 Observatory of China, New Mexico State University, New York University, University of Notre Dame, Observatório 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. T.A. acknowledges funding by MCIN/AEI/10.13039/501100011033 as well as the “ESF Investing in your future” and “ERDF A way of making Europe” funds of the European Union and the European Union's “Next Generation EU”/PRTR, through grants RYC2018-025968-I, PID2021-125451NA-I00, and CNS2022-135232.

## References

- Abdurro'uf, N., Accetta, K., Aerts, C., et al. 2022, *ApJS*, 259, 35  
 Abolfathi, B., Aguado, D. S., Aguilar, G., et al. 2018, *ApJS*, 235, 42  
 Adibekyan, V. Z., Santos, N. C., Sousa, S. G., & Israelian, G. 2011, *A&A*, 535, L11  
 Afflerbach, A., Churchwell, E., & Werner, M. W. 1997, *ApJ*, 478, 190  
 Agertz, O., Renaud, F., Feltzing, S., et al. 2021, *MNRAS*, 503, 5826  
 Alam, S., Albareti, F. D., Allende Prieto, C., et al. 2015, *ApJS*, 219, 12  
 Allende Prieto, C., Majewski, S. R., Schiavon, R., et al. 2008, *Astron. Nachr.*, 329, 1018  
 Alzate, J. A., Bruzual, G., & Díaz-González, D. J. 2021, *MNRAS*, 501, 302

<sup>8</sup> [https://github.com/fjaellet/xgboost\\_chem\\_ages](https://github.com/fjaellet/xgboost_chem_ages)

<sup>9</sup> [https://github.com/fjaellet/xgboost\\_chem\\_ages/blob/main/data/spec\\_ages\\_published.fits](https://github.com/fjaellet/xgboost_chem_ages/blob/main/data/spec_ages_published.fits)

- Amarante, J. A. S., Beraldo e Silva, L., Debattista, V. P., & Smith, M. C., 2020, *ApJ*, **891**, L30
- An, D., Beers, T. C., Lee, Y. S., & Masseron, T. 2023, *ApJ*, **952**, 66
- Anders, F., Chiappini, C., Minchev, I., et al. 2017a, *A&A*, **600**, A70
- Anders, F., Chiappini, C., Rodrigues, T. S., et al. 2017b, *A&A*, **597**, A30
- Anders, F., Chiappini, C., Santiago, B. X., et al. 2018, *A&A*, **619**, A125
- Anders, F., Buck, T., Frankel, N., & Minchev, I. 2020, in *XIV.0 Scientific Meeting (virtual) of the Spanish Astronomical Society*, 115
- Anders, F., Khalatyan, A., Queiroz, A. B. A., et al. 2022, *A&A*, **658**, A91
- Astropy Collaboration (Price-Whelan, A. M., et al.) 2018, *AJ*, **156**, 123
- Aumer, M., Binney, J., & Schönrich, R. 2016, *MNRAS*, **462**, 1697
- Baglin, A., Auvergne, M., Boisnard, L., et al. 2006, in *36th COSPAR Scientific Assembly*, **36**, 3749
- Belokurov, V., & Kravtsov, A. 2022, *MNRAS*, **514**, 689
- Belokurov, V., Erkal, D., Evans, N. W., Koposov, S. E., & Deason, A. J. 2018, *MNRAS*, **478**, 611
- Beraldo e Silva, L., Debattista, V. P., Khachatryan, T., & Nidever, D. 2020, *MNRAS*, **492**, 4716
- Beraldo e Silva, L., Debattista, V. P., Nidever, D., Amarante, J. A. S., & Garver, B. 2021, *MNRAS*, **502**, 260
- Bernard, E. J. 2018, *IAU Symp.*, **334**, 158
- Bird, J. C., Kazantzidis, S., Weinberg, D. H., et al. 2013, *ApJ*, **773**, 43
- Bird, J. C., Loebman, S. R., Weinberg, D. H., et al. 2021, *MNRAS*, **503**, 1815
- Borisov, V., Leemann, T., Seßler, K., et al. 2022, *IEEE Transactions on Neural Networks and Learning Systems*, 1
- Borucki, W. J., Koch, D., Basri, G., et al. 2010, *Science*, **327**, 977
- Bossini, D., Vallenari, A., Bragaglia, A., et al. 2019, *A&A*, **623**, A108
- Bournaud, F., Elmegreen, B. G., & Martig, M. 2009, *ApJ*, **707**, L1
- Bovy, J., Allende Prieto, C., Beers, T. C., et al. 2012, *ApJ*, **759**, 131
- Bovy, J., Rix, H.-W., Schlafly, E. F., et al. 2016, *ApJ*, **823**, 30
- Breiman, L. 2001, *Mach. Learn.*, **45**, 5
- Brook, C. B., Kawata, D., Gibson, B. K., & Freeman, K. C. 2004, *ApJ*, **612**, 894
- Brook, C. B., Gibson, B. K., Martel, H., & Kawata, D. 2005, *ApJ*, **630**, 298
- Bu, Y., Kumar, Y. B., Xie, J., et al. 2020, *ApJS*, **249**, 7
- Buck, T., Obreja, A., Macciò, A. V., et al. 2020, *MNRAS*, **491**, 3461
- Buder, S., Sharma, S., Kos, J., et al. 2021, *MNRAS*, **506**, 150
- Byl, J. 1974, *MNRAS*, **169**, 157
- Cantat-Gaudin, T., Anders, F., Castro-Ginard, A., et al. 2020, *A&A*, **640**, A1
- Carr, C., Johnston, K. V., Laporte, C. F. P., & Ness, M. K. 2022, *MNRAS*, **516**, 5067
- Casagrande, L., Schönrich, R., Asplund, M., et al. 2011, *A&A*, **530**, A138
- Casagrande, L., Silva Aguirre, V., Schlesinger, K. J., et al. 2016, *MNRAS*, **455**, 987
- Casali, G., Magrini, L., Tognelli, E., et al. 2019, *A&A*, **629**, A62
- Casali, G., Spina, L., Magrini, L., et al. 2020, *A&A*, **639**, A127
- Casali, G., Grisoni, V., Miglio, A., et al. 2023, *A&A*, **677**, A60
- Casamiquela, L., Soubiran, C., Jofré, P., et al. 2021, *A&A*, **652**, A25
- Chaplin, W. J., & Miglio, A. 2013, *ARA&A*, **51**, 353
- Chaplin, W. J., Serenelli, A. M., Miglio, A., et al. 2020, *Nat. Astron.*, **4**, 382
- Chen, T., & Guestrin, C. 2016, KDD '16: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785
- Chen, Y. Q., & Zhao, G. 2020, *MNRAS*, **495**, 2673
- Chiappini, C. 2009, in *The Galaxy Disk in Cosmological Context*, eds. J. Andersen, B. M. Nordström, & J. Bland-Hawthorn, 254, 191
- Chiappini, C., Matteucci, F., & Gratton, R. 1997, *ApJ*, **477**, 765
- Chiappini, C., Matteucci, F., & Romano, D. 2001, *ApJ*, **554**, 1044
- Chiappini, C., Anders, F., Rodrigues, T. S., et al. 2015, *A&A*, **576**, L12
- Ciucă, I., Kawata, D., Miglio, A., Davies, G. R., & Grand, R. J. J. 2021, *MNRAS*, **503**, 2814
- Ciucă, I., Kawata, D., Ting, Y. S., et al. 2023, *MNRAS*, in press, <https://doi.org/10.1093/mnras/1/slad033>
- Cui, X.-Q., Zhao, Y.-H., Chu, Y.-Q., et al. 2012, *Res. Astron. Astrophys.*, **12**, 1197
- Das, P., & Sanders, J. L. 2019, *MNRAS*, **484**, 294
- da Silva, L., Girardi, L., Pasquini, L., et al. 2006, *A&A*, **458**, 609
- Delgado Mena, E., Tsantaki, M., Adibekyan, V. Z., et al. 2017, *A&A*, **606**, A94
- Delgado Mena, E., Moya, A., Adibekyan, V., et al. 2019, *A&A*, **624**, A78
- De Silva, G. M., Freeman, K. C., Bland-Hawthorn, J., et al. 2015, *MNRAS*, **449**, 2604
- Edvardsson, B., Andersen, J., Gustafsson, B., et al. 1993, *A&A*, **275**, 101
- Feltzing, S., Howes, L. M., McMillan, P. J., & Stokutė, E. 2017, *MNRAS*, **465**, L109
- Ferrini, F., Molla, M., Pardi, M. C., & Diaz, A. I. 1994, *ApJ*, **427**, 745
- Feillet, D. K., Bovy, J., Holtzman, J., et al. 2018, *MNRAS*, **477**, 2326
- Forbes, J., Krumholz, M., & Burkert, A. 2012, *ApJ*, **754**, 48
- Frankel, N., Rix, H.-W., Ting, Y.-S., Ness, M., & Hogg, D. W. 2018, *ApJ*, **865**, 96
- Frankel, N., Sanders, J., Rix, H.-W., Ting, Y.-S., & Ness, M. 2019, *ApJ*, **884**, 99
- Frankel, N., Sanders, J., Ting, Y.-S., & Rix, H.-W. 2020, *ApJ*, **896**, 15
- Friel, E. D. 1995, *ARA&A*, **33**, 381
- Friel, E. D., Janes, K. A., Tavares, M., et al. 2002, *AJ*, **124**, 2693
- Frinchaboy, P. M., Thompson, B., Jackson, K. M., et al. 2013, *ApJ*, **777**, L1
- Fuhrmann, K., & Chini, R. 2017, *MNRAS*, **471**, 1888
- Fuhrmann, K., & Chini, R. 2018, *ApJ*, **858**, 103
- Fuhrmann, K., Chini, R., Kaderhandt, L., & Chen, Z. 2017, *MNRAS*, **464**, 2610
- Gaia Collaboration (Brown, A. G. A., et al.) 2021, *A&A*, **649**, A1
- Gaia Collaboration (Recio-Blanco, A., et al.) 2023a, *A&A*, **674**, A38
- Gaia Collaboration (Vallenari, A., et al.) 2023b, *A&A*, **674**, A1
- Genovali, K., Lemasle, B., Bono, G., et al. 2014, *A&A*, **566**, A37
- Gliese, W. 1956, *ZAp*, **39**, 1
- Grand, R. J. J., Springel, V., Gómez, F. A., et al. 2016, *MNRAS*, **459**, 199
- Grand, R. J. J., Gómez, F. A., Marinacci, F., et al. 2017, *MNRAS*, **467**, 179
- Hasselquist, S., Zasowski, G., Feuillet, D. K., et al. 2020, *ApJ*, **901**, 109
- Hawley, S. A. 1978, *ApJ*, **224**, 417
- Hayden, M. R., Sharma, S., Bland-Hawthorn, J., et al. 2022, *MNRAS*, **517**, 5325
- Hayes, C. R., Masseron, T., Sobeck, J., et al. 2022, *ApJS*, **262**, 34
- Haywood, M. 2008, *MNRAS*, **388**, 1175
- He, X.-J., Luo, A. L., & Chen, Y.-Q. 2022, *MNRAS*, **512**, 1710
- Hekker, S., & Johnson, J. A. 2019, *MNRAS*, **487**, 4343
- Helmi, A., Babusiaux, C., Koppelman, H. H., et al. 2018, *Nature*, **563**, 85
- Hou, J. L., Prantzos, N., & Boissier, S. 2000, *A&A*, **362**, 921
- Huang, Y., Schönrich, R., Zhang, H., et al. 2020, *ApJS*, **249**, 29
- Hunter, J. D. 2007, *Comput. Sci. Eng.*, **9**, 90
- Imig, J., Price, C., Holtzman, J. A., et al. 2023, *ApJ*, **954**, 124
- Isern, J. 2019, *ApJ*, **878**, L11
- Ivezić, Ž., Connolly, A. J., VanderPlas, J. T., & Gray, A. 2020, *Statistics, Data Mining, and Machine Learning in Astronomy: A Practical Python Guide for the Analysis of Survey Data* (Princeton, NJ: Princeton University Press)
- Janes, K. A. 1979, *ApJS*, **39**, 135
- Jofré, P. 2021, *ApJ*, **920**, 23
- Jofré, P., Jorissen, A., Van Eck, S., et al. 2016, *A&A*, **595**, A60
- Jofré, P., Jackson, H., & Tucci Maia, M. 2020, *A&A*, **633**, L9
- Jofré, P., Jorissen, A., Aguilera-Gómez, C., et al. 2023, *A&A*, **671**, A21
- Johnson, J. W., Weinberg, D. H., Vincenzo, F., et al. 2021, *MNRAS*, **508**, 4484
- Jönsson, H., Holtzman, J. A., Allende Prieto, C., et al. 2020, *AJ*, **160**, 120
- Jørgensen, B. R., & Lindegren, L. 2005, *A&A*, **436**, 127
- Joyce, M., Johnson, C. I., Marchetti, T., et al. 2023, *ApJ*, **946**, 28
- Koepfen, J. 1994, *A&A*, **281**, 26
- Kubryk, M., Prantzos, N., & Athanassoula, E. 2013, *MNRAS*, **436**, 1479
- Kubryk, M., Prantzos, N., & Athanassoula, E. 2015, *A&A*, **580**, A127
- Lagarde, N., Robin, A. C., Reylé, C., & Nasello, G. 2017, *A&A*, **601**, A27
- Lagarde, N., Reylé, C., Chiappini, C., et al. 2021, *A&A*, **654**, A13
- Lebreton, Y., & Reese, D. R. 2020, *A&A*, **642**, A88
- Lebreton, Y., Goupil, M. J., & Montalbán, J. 2014, in *EAS Publ. Ser.*, eds. Y. Lebreton, D. Valls-Gabaud, & C. Charbonnel, 65, 99
- Leung, H. W., & Bovy, J. 2019, *MNRAS*, **489**, 2079
- Leung, H. W., Bovy, J., Mackereth, J. T., & Miglio, A. 2023, *MNRAS*, **522**, 4577
- Lian, J., Zasowski, G., Mackereth, T., et al. 2022, *MNRAS*, **513**, 4130
- Lu, Y., Buck, T., Minchev, I., & Ness, M. K. 2022a, *MNRAS*, **515**, L34
- Lu, Y. L., Ness, M. K., Buck, T., & Carr, C. 2022b, *MNRAS*, **512**, 4697
- Lu, Y., Minchev, I., Buck, T., et al. 2023, *Nat. Lett.*, submitted [arXiv:2212.04515]
- Luck, R. E., & Bond, H. E. 1980, *ApJ*, **241**, 218
- Lundberg, S., & Lee, S.-I. 2017, *ArXiv e-prints* [arXiv:1705.07874]
- Lyngå, G. 1982, *A&A*, **109**, 213
- Lyngå, G., & Palouš, J. 1987, *A&A*, **188**, 35
- Maciel, W. J., Lago, L. G., & Costa, R. D. D. 2005, *A&A*, **433**, 127
- Mackereth, J. T., Bovy, J., Schiavon, R. P., et al. 2017, *MNRAS*, **471**, 3057
- Mackereth, J. T., Bovy, J., Leung, H. W., et al. 2019, *MNRAS*, **489**, 176
- Mackereth, J. T., Miglio, A., Elsworth, Y., et al. 2021, *MNRAS*, **502**, 1947
- Magrini, L., Sestito, P., Randich, S., & Galli, D. 2009, *A&A*, **494**, 95
- Magrini, L., Spina, L., Randich, S., et al. 2018, *A&A*, **617**, A106
- Magrini, L., Viscasillas Vázquez, C., Spina, L., et al. 2023, *A&A*, **669**, A119
- Maiorca, E., Randich, S., Busso, M., Magrini, L., & Palmerini, S. 2011, *ApJ*, **736**, 120
- Majewski, S. R., Schiavon, R. P., Frinchaboy, P. M., et al. 2017, *AJ*, **154**, 94
- Mannucci, F., Della Valle, M., & Panagia, N. 2006, *MNRAS*, **370**, 773
- Martig, M., Minchev, I., & Flynn, C. 2014, *MNRAS*, **443**, 2452
- Martig, M., Rix, H.-W., Silva Aguirre, V., et al. 2015, *MNRAS*, **451**, 2230
- Martig, M., Founesneau, M., Rix, H.-W., et al. 2016a, *MNRAS*, **456**, 3655
- Martig, M., Minchev, I., Ness, M., Founesneau, M., & Rix, H.-W. 2016b, *ApJ*, **831**, 139
- Masseron, T., & Gilmore, G. 2015, *MNRAS*, **453**, 1855
- Matsuno, T., Aoki, W., Casagrande, L., et al. 2021, *ApJ*, **912**, 72
- Mayor, M. 1974, *A&A*, **32**, 321

- McMillan, P. J., Kordopatis, G., Kunder, A., et al. 2018, *MNRAS*, 477, 5279
- Mennekens, N., Vanbeveren, D., De Greve, J. P., & De Donder, E. 2010, *A&A*, 515, A89
- Meusinger, H., Reimann, H. G., & Stecklum, B. 1991, *A&A*, 245, 57
- Miglio, A., Montalbán, J., & Noels, A. 2012, *Astrophys. Space Sci. Proc.*, 26, 165
- Miglio, A., Chiappini, C., Morel, T., et al. 2013, *MNRAS*, 429, 423
- Miglio, A., Chiappini, C., Mosser, B., et al. 2017, *Astron. Nachr.*, 338, 644
- Miglio, A., Chiappini, C., Mackereth, J. T., et al. 2021, *A&A*, 645, A85
- Minchev, I., Chiappini, C., & Martig, M. 2013, *A&A*, 558, A9
- Minchev, I., Chiappini, C., & Martig, M. 2014, *A&A*, 572, A92
- Minchev, I., Martig, M., Streich, D., et al. 2015, *ApJ*, 804, L9
- Minchev, I., Steinmetz, M., Chiappini, C., et al. 2017, *ApJ*, 834, 27
- Minchev, I., Anders, F., Recio-Blanco, A., et al. 2018, *MNRAS*, 481, 1645
- Mints, A., & Hekker, S. 2017, *A&A*, 604, A108
- Montalbán, J., Mackereth, J. T., Miglio, A., et al. 2021, *Nat. Astron.*, 5, 640
- Mor, R., Robin, A. C., Figueras, F., Roca-Fàbrega, S., & Luri, X. 2019, *A&A*, 624, L1
- Moya, A., Sarro, L. M., Delgado-Mena, E., et al. 2022, *A&A*, 660, A15
- Myers, N., Donor, J., Spoo, T., et al. 2022, *AJ*, 164, 85
- Naab, T., & Ostriker, J. P. 2006, *MNRAS*, 366, 899
- Navarro, J. F., Yozin, C., Loewen, N., et al. 2018, *MNRAS*, 476, 3648
- 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
- Ness, M. K., Johnston, K. V., Blancato, K., et al. 2019, *ApJ*, 883, 177
- Netopil, M., Oralhan, İ. A., Çakmak, H., Michel, R., & Karataş, Y. 2022, *MNRAS*, 509, 421
- Ng, Y. K., & Bertelli, G. 1998, *A&A*, 329, 943
- Nissen, P. E. 2015, *A&A*, 579, A52
- Nissen, P. E. 2016, *A&A*, 593, A65
- Nissen, P. E., Christensen-Dalsgaard, J., Mosumgaard, J. R., et al. 2020, *A&A*, 640, A81
- Noels, A., & Bragaglia, A. 2015, in *Asteroseismology of Stellar Populations in the Milky Way* (New York: Springer International Publishing), *Astrophys. Space Sci. Proc.*, 39, 167
- Nordström, B., Mayor, M., Andersen, J., et al. 2004, *A&A*, 418, 989
- Panagia, N., & Tosi, M. 1981, *A&A*, 96, 306
- Patton, R. A., Pinsonneault, M. H., Cao, L., et al. 2023, *MNRAS*, submitted [arXiv:2303.08151]
- Peimbert, M., & Serrano, A. 1980, *Rev. Mex. Astron. Astrofis.*, 5, 9
- Pinsonneault, M. H., Elsworth, Y., Epstein, C., et al. 2014, *ApJS*, 215, 19
- Pinsonneault, M. H., Elsworth, Y. P., Tayar, J., et al. 2018, *ApJS*, 239, 32
- Pont, F., & Eyer, L. 2004, *MNRAS*, 351, 487
- Price-Jones, N., & Bovy, J. 2018, *MNRAS*, 475, 1410
- Queiroz, A. B. A., Anders, F., Santiago, B. X., et al. 2018, *MNRAS*, 476, 2556
- Queiroz, A. B. A., Anders, F., Chiappini, C., et al. 2020, *A&A*, 638, A76
- Queiroz, A. B. A., Chiappini, C., Perez-Villegas, A., et al. 2021, *A&A*, 656, A156
- Queiroz, A. B. A., Anders, F., Chiappini, C., et al. 2023, *A&A*, 673, A155
- Quillen, A. C., & Garnett, D. R. 2000, *ArXiv e-prints* [arXiv:astro-ph/0004210]
- Raddi, R., Torres, S., Rebassa-Mansergas, A., et al. 2022, *A&A*, 658, A22
- Ratcliffe, B. L., Ness, M. K., Johnston, K. V., & Sen, B. 2020, *ApJ*, 900, 165
- Ratcliffe, B., Minchev, I., Anders, F., et al. 2023, *MNRAS*, 525, 2208
- Rauer, H., Catala, C., Aerts, C., et al. 2014, *Exp. Astron.*, 38, 249
- Recio-Blanco, A., de Laverny, P., Palicio, P. A., et al. 2023, *A&A*, 674, A29
- Renaud, F., Agertz, O., Read, J. I., et al. 2021, *MNRAS*, 503, 5846
- Rendle, B. M., Miglio, A., Chiappini, C., et al. 2019, *MNRAS*, 490, 4465
- Rodrigues, T. S., Bossini, D., Miglio, A., et al. 2017, *MNRAS*, 467, 1433
- Ruiz-Lara, T., Gallart, C., Bernard, E. J., & Cassisi, S. 2020, *Nat. Astron.*, 4, 965
- Saha, K., Tseng, Y.-H., & Taam, R. E. 2010, *ApJ*, 721, 1878
- Sahlholdt, C. L., Feltzing, S., & Feuillet, D. K. 2022, *MNRAS*, 510, 4669
- Salaris, M., Pietrinferni, A., Piersimoni, A. M., & Cassisi, S. 2015, *A&A*, 583, A87
- Sales-Silva, J. V., Daflon, S., Cunha, K., et al. 2022, *ApJ*, 926, 154
- Sanders, J. L., & Das, P. 2018, *MNRAS*, 481, 4093
- Santiago, B. X., Brauer, D. E., Anders, F., et al. 2016, *A&A*, 585, A42
- Schonhut-Stasik, J., Zinn, J. C., Stassun, K. G., et al. 2023, *AJ*, submitted [arXiv:2304.10654]
- Schönrich, R., & Binney, J. 2009, *MNRAS*, 396, 203
- Schönrich, R., Binney, J., & Dehnen, W. 2010, *MNRAS*, 403, 1829
- Sharma, S., Stello, D., Buder, S., et al. 2018, *MNRAS*, 473, 2004
- Silva Aguirre, V., Stello, D., Stokholm, A., et al. 2020, *ApJ*, 889, L34
- Soderblom, D. R. 2010, *ARA&A*, 48, 581
- Spina, L., Meléndez, J., Karakas, A. I., et al. 2018, *MNRAS*, 474, 2580
- Spina, L., Ting, Y. S., De Silva, G. M., et al. 2021, *MNRAS*, 503, 3279
- Spina, L., Magrini, L., & Cunha, K. 2022, *Universe*, 8, 87
- Spitoni, E., Recio-Blanco, A., de Laverny, P., et al. 2023, *A&A*, 670, A109
- Spoo, T., Tayar, J., Frinchaboy, P. M., et al. 2022, *AJ*, 163, 229
- Stanghellini, L., & Haywood, M. 2010, *ApJ*, 714, 1096
- Stanghellini, L., & Haywood, M. 2018, *ApJ*, 862, 45
- Sun, W. X., Huang, Y., Wang, H. F., et al. 2020, *ApJ*, 903, 12
- Sysoliatina, K., & Just, A. 2022, *A&A*, 666, A130
- Tarricq, Y., Soubiran, C., Casamiquela, L., et al. 2021, *A&A*, 647, A19
- Taylor, M. B. 2005, in *TOPCAT & STIL: Starlink Table/VOTable Processing Software*, eds. P. Shopbell, M. Britton, & R. Ebert, *ASP Conf. Ser.*, 347, 29
- Ting, Y.-S., & Rix, H.-W. 2019, *ApJ*, 878, 21
- Tinsley, B. M. 1980, *Fund. Cosmic Phys.*, 5, 287
- Trimble, V. 1975, *Rev. Mod. Phys.*, 47, 877
- Tucci Maia, M., Ramírez, I., Meléndez, J., et al. 2016, *A&A*, 590, A32
- Valentini, M., Chiappini, C., Miglio, A., et al. 2016, *Astron. Nachr.*, 337, 970
- Valentini, M., Chiappini, C., Bossini, D., et al. 2019, *A&A*, 627, A173
- Valle, G., Dell'Omodarme, M., Prada Moroni, P. G., & Degl'Innocenti, S. 2013, *A&A*, 549, A50
- van der Kruit, P. C., & Freeman, K. C. 2011, *ARA&A*, 49, 301
- van der Walt, S., Colbert, S. C., & Varoquaux, G. 2011, *Comput. Sci. Eng.*, 13, 22
- Vickers, J. J., Shen, J., & Li, Z.-Y. 2021, *ApJ*, 922, 189
- Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, *Nat. Methods*, 17, 261
- Viscasillas Vázquez, C., Magrini, L., Casali, G., et al. 2022, *A&A*, 660, A135
- von Hoerner, S. 1960, *Fortschr. Phys.*, 8, 191
- Wang, C., Liu, X. W., Xiang, M. S., et al. 2019, *MNRAS*, 482, 2189
- Weinberg, D. H., Holtzman, J. A., Hasselquist, S., et al. 2019, *ApJ*, 874, 102
- Wenger, T. V., Bailer, D. S., Anderson, L. D., & Bania, T. M. 2019, *ApJ*, 887, 114
- Wetzel, A., Hayward, C. C., Sanderson, R. E., et al. 2023, *ApJS*, 265, 44
- Wielen, R. 1974, *Highlights Astron.*, 3, 395
- Wilson, J. C., Hearty, F. R., Skrutskie, M. F., et al. 2019, *PASP*, 131, 055001
- Wu, Y., Xiang, M., Bi, S., et al. 2018, *MNRAS*, 475, 3633
- Wu, Y., Xiang, M., Zhao, G., et al. 2019, *MNRAS*, 484, 5315
- Wu, Y., Xiang, M., Zhao, G., et al. 2023, *MNRAS*, 520, 1913
- Xiang, M., & Rix, H.-W. 2022, *Nature*, 603, 599
- Xiang, M.-S., Liu, X.-W., Yuan, H.-B., et al. 2015, *Res. Astron. Astrophys.*, 15, 1209
- Yong, D., Casagrande, L., Venn, K. A., et al. 2016, *MNRAS*, 459, 487
- Yu, J., & Liu, C. 2018, *MNRAS*, 475, 1093
- Yu, J., Huber, D., Bedding, T. R., et al. 2018, *ApJS*, 236, 42
- Yu, S., Bullock, J. S., Klein, C., et al. 2021, *MNRAS*, 505, 889
- Yurin, D., & Springel, V. 2015, *MNRAS*, 452, 2367
- Zhang, M., Xiang, M., Zhang, H.-W., et al. 2021, *ApJ*, 922, 145
- Zinn, J. C., Stello, D., Elsworth, Y., et al. 2020, *ApJS*, 251, 23
- Zinn, J. C., Stello, D., Elsworth, Y., et al. 2022, *ApJ*, 926, 191



## Appendix A: Age catalogue

Table A.1 shows the datamodel for the catalogue that we release with this work. Columns 2-4 correspond to the fiducial results obtained with the default XGBoost (package version 1.7.6) run described in Sect. 3. Columns 6-9 correspond to results obtained with XGBoost quantile regression<sup>10</sup>, available in package version 2.0.0-dev. The advantage of these latter results is that they come with more meaningful individual uncertainties, the disadvantage is that the median values suffer from a greater regression-towards-the-mean effect, i.e. noticeably smaller ages in the old regime. It also results in a higher number of young  $\alpha$ -rich stars (1 222 vs. 706). For most other stars, the differences

between the two methods are small. In most of this paper, as well as in Ratcliffe et al. (2023), we mainly use the columns `spec_age_xgb` and `spec_age_xgb_uncert`.

Some of the caveats described in Sect. 4.5 are encoded as warning flags in the human-readable column `spec_age_xgb_flag` (`spec_age_xgb_quantilereg_flag` for the quantile regression ages). For the 5383 affected stars (5549 in the quantile regression case) it can take the values `BLUER_THAN_TRAINING_SET`, `REDDER_THAN_TRAINING_SET`, `HIGH_VSINI`, `APPARENTLY_YOUNG_ALPHA_RICH`, and combinations thereof. The preparation steps necessary for the production of the age catalogue can be reproduced and modified following the procedure in the accompanying jupyter notebooks.

**Table A.1.** Data model of our APOGEE DR17 spectroscopic age catalogue.

#	Column name	Format	Description
1	<code>APOGEE_ID</code>	String (19A)	APOGEE ID (DR17)
2	<code>spec_age_xgb</code>	Float	Spectroscopic age from fiducial XGBoost model
3	<code>spec_age_xgb_calib</code>	Float	Calibrated spectroscopic age (see Fig. 4, top panel)
4	<code>spec_age_xgb_uncert</code>	Float	Age uncertainty estimate (see Fig. 4, bottom panel)
5	<code>spec_age_xgb_flag</code>	String (51A)	Human-readable warning flag for potentially problematic stars
6	<code>spec_age_xgb_quantilereg</code>	Float	Spectroscopic age from XGBoost quantile regression
7	<code>spec_age_xgb_quantilereg_calib</code>	Float	Calibrated age from XGBoost quantile regression
8	<code>spec_age_xgb_quantilereg_sig1</code>	Float	Lower $1\sigma$ uncertainty from XGBoost quantile regression
9	<code>spec_age_xgb_quantilereg_sigu</code>	Float	Upper $1\sigma$ uncertainty from XGBoost quantile regression
10	<code>spec_age_xgb_quantilereg_flag</code>	String (52A)	Human-readable warning flag for potentially problematic stars

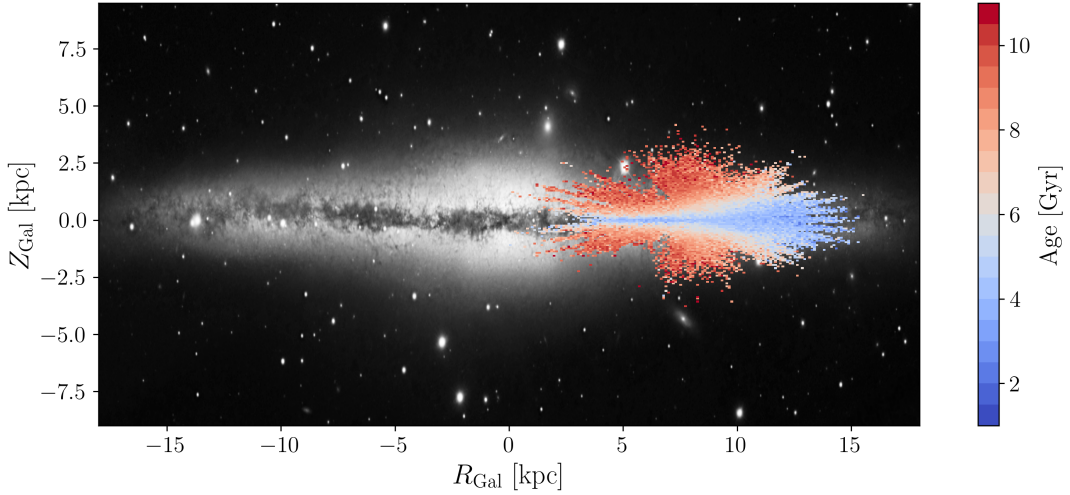
**Notes.** The full table is available at the CDS.

<sup>10</sup> [https://scikit-learn.org/stable/auto\\_examples/ensemble/plot\\_gradient\\_boosting\\_quantile.html](https://scikit-learn.org/stable/auto_examples/ensemble/plot_gradient_boosting_quantile.html)

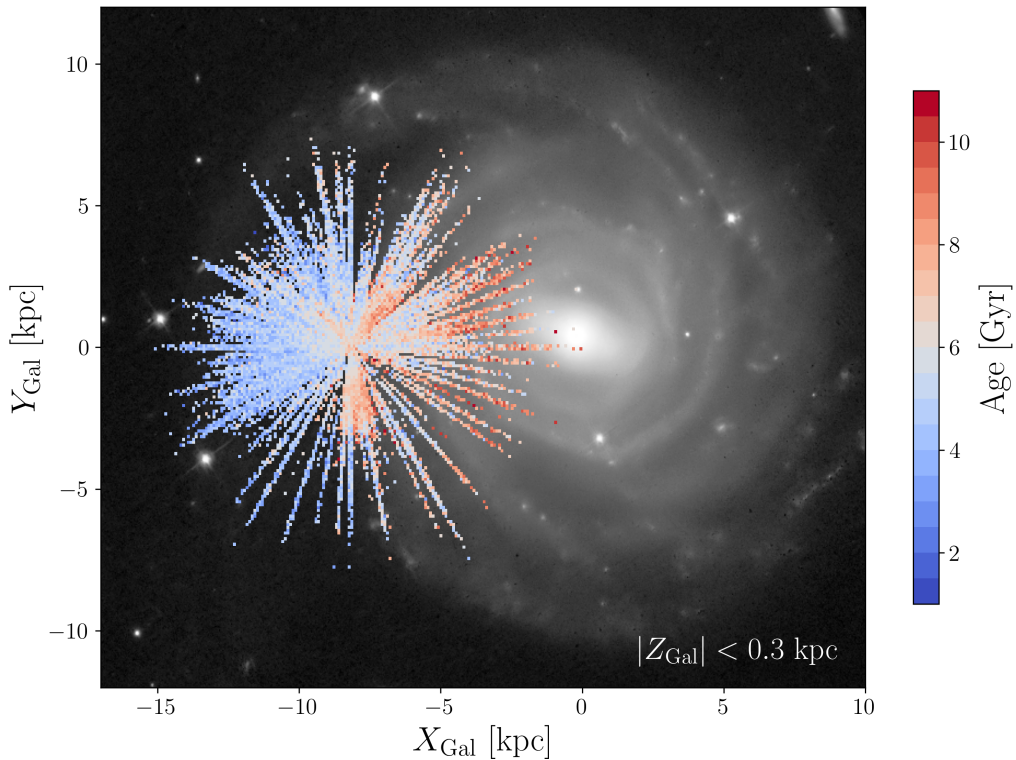
## Appendix B: Additional age maps

For illustration purposes, we include two additional figures highlighting the Galactic coverage of the present APOGEE spectroscopic age data. Figure B.1 shows the median age per pixel in

cylindrical coordinates, similar to Fig. 9, but now using equally-scaled axes, and with a Milky-Way-like edge-on galaxy (NGC 891) as a background image. Figure B.2 shows a top-down view of the Galactic disc colour-coded by median age, with the face-on Milky-Way analogue UGC 12158 in the background.



**Fig. B.1.** Median age as a function of position in Galactocentric coordinates, similar to Fig. 9, but using equal scales on both axes and showing a greater extent of the Galactic disc. We also show an  $r$ -band image of the nearby edge-on galaxy NGC 891 in the background for illustration.



**Fig. B.2.** Median age as a function of position in Galactocentric Cartesian coordinates  $\{X_{\text{Gal}}, Y_{\text{Gal}}\}$  (top-down view of the Galaxy), showing a slice of Galactic disc close to the Galactic plane ( $|Z_{\text{Gal}}| < 0.3$  kpc). An  $r$ -band image of the nearby face-on spiral galaxy UGC 12158 is shown in the background for illustration.

## Appendix C: Additional validation

### C.1. StarHorse ages for APOGEE DR17 (Queiroz et al. 2023)

The top panel of Fig. C.1 shows a direct comparison with the age estimates of Queiroz et al. (2023) for the APOGEE DR17 catalogue. The authors used the Bayesian isochrone-fitting code StarHorse (Queiroz et al. 2018, 2020) to simultaneously obtain stellar parameters for the APOGEE DR17 catalogue, cross-matched with *Gaia* EDR3 astrometry (Gaia Collaboration 2021) as well as several photometric catalogues. Qualitatively, the comparison reveals a similar systematic trend as the comparison to the seismic and open-cluster ages shown in the top panels of Fig. 6: good concordance for the youngest ages followed by prominent positive ( $\sim +1$  Gyr) residuals for ages between 2 and 4 Gyr, after which the trend regresses towards zero residuals again. We note, however, that Queiroz et al. (2023) only published the ages inferred for sub-giant and main-sequence turn-off stars, since isochrone ages for red giants are still prone to sizeable systematic uncertainties – which also explains the larger scatter.

### C.2. [C/N] ages (Spoo et al. 2022)

Since the APOGEE discovery of a systematic trend of [C/N] with the position in the [Mg/Fe] vs. [Fe/H] diagram and, therefore, possibly stellar age (Masseron & Gilmore 2015), several authors have tried to calibrate a relation between the observed [C/N] ratio and the age of red giant stars (e.g. Martig et al. 2016a; Ness et al. 2016; Casali et al. 2019). To obtain a solid calibration for evolved stars, Spoo et al. (2022) used the APOGEE DR17 data from the Open Cluster Chemical Abundances and Mapping survey (OCCAM; Frinchaboy et al. 2013; Myers et al. 2022). Using 94 open clusters tagged as "highly reliable", they obtained the following relationship:

$$\log[\text{Age}(\text{yr})]_{\text{DR17}} = 10.14 (\pm 0.08) + 2.23 (\pm 0.19) \cdot [\text{C}/\text{N}],$$

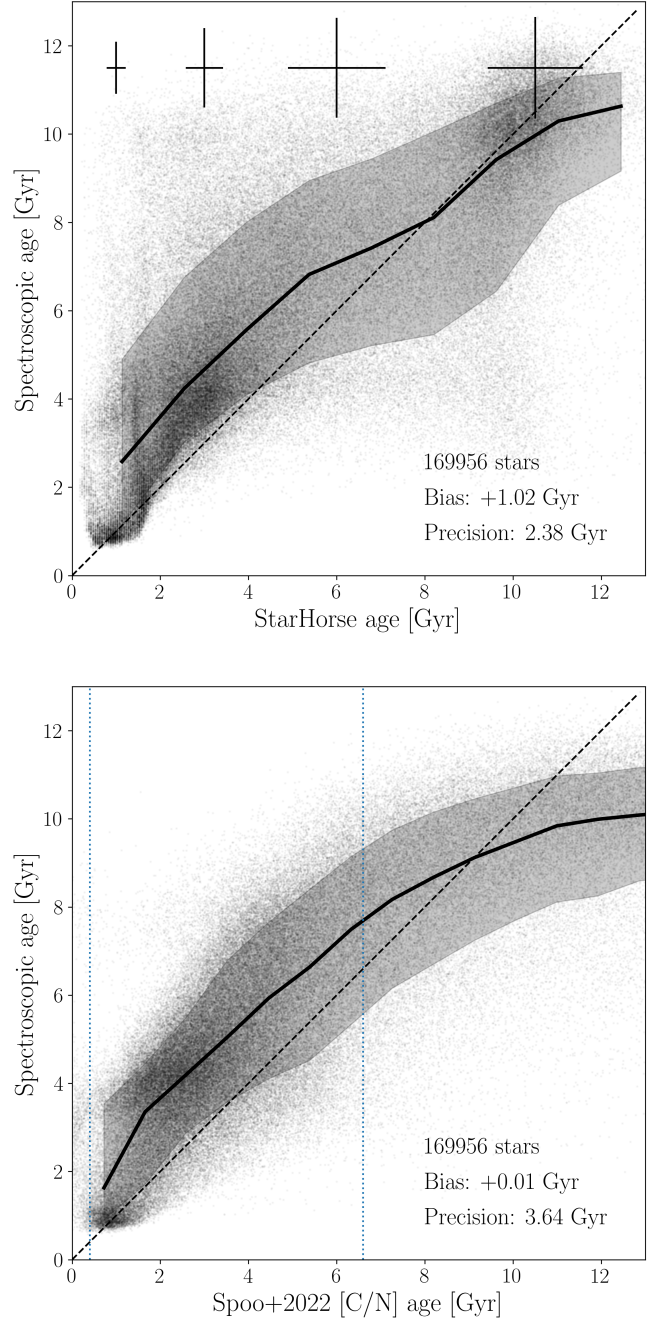
valid for ages between 0.4 Gyr and 6.6 Gyr.

The lower panel of Fig. C.1 shows the comparison of our spectroscopic age estimates with the calibration obtained by Spoo et al. (2022). Again, we find similar trends as in those visible in Fig. 6, albeit with much larger scatter, which is expected, since the empirical relation of Spoo et al. (2022) only uses two abundance measurements to infer an age estimate.

### C.3. Age-abundance relations

From Fig. 3 we learned that the five most influential features for our spectroscopic age estimator are [C/Fe],  $T_{\text{eff}}$ , [N/Fe], [Mg/Fe], and  $\log g$ . Therefore, before we discuss some of the abundance trends with age, a note of caution is due: ideally one wishes to study stellar abundance ratios as a function of an age estimate that is both accurate and independent from the spectroscopically inferred abundances. In our case, however, the inferred ages rely on the measured abundances, and therefore our derived abundance ratio (e.g. [Mg/Fe], Fig. C.3) vs. age trends should only be viewed as sanity checks, not as proofs of the validity of the concept of chemical clocks.

Figure C.2 shows the [Mg/Fe] vs. [Fe/H] diagram colour-coded by age, for the solar-vicinity sample of Delgado Mena et al. (2017, 2019) and for our APOGEE DR17 giant sample. The figure demonstrates that our data reproduce the expected age trends in this diagram, and that

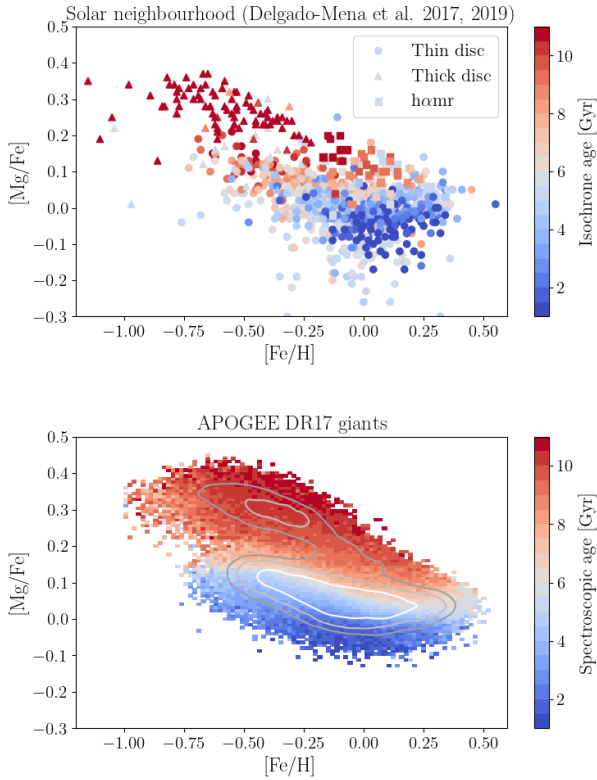


**Fig. C.1.** Additional comparisons of our uncalibrated age estimates with other field-star age estimates for APOGEE DR17. Top: Unpublished red-giant star age estimates from StarHorse (Queiroz et al. 2023). Bottom: Empirically calibrated DR17 [C/N] age estimates from Spoo et al. (2022). The dotted vertical lines enclose the age regime in which their [C/N] age calibration is valid.

these trends persist over the large portion of the Galactic disc covered by our sample. It also confirms that the so-called high- $\alpha$  metal-rich population (h $\alpha$ mr, Adibekyan et al. 2011) is dominantly old (between 9 and 11 Gyr), indicating that this population was most likely formed in the very inner disc and/or bulge and migrated outwards (e.g. Anders et al. 2018; Queiroz et al. 2021; Ratcliffe et al. 2023).

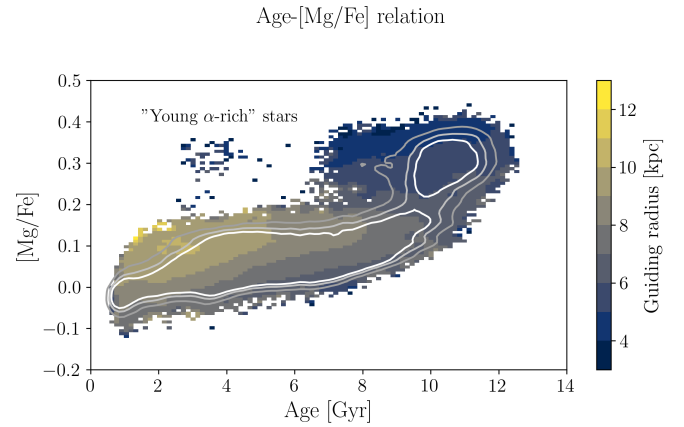
As a typical example of an elemental-abundance ratio that is often referred to as a chemical clock, we show the age-[Mg/Fe] relation for the full APOGEE sample in Fig. C.3. The iso-density





**Fig. C.2.**  $[Mg/Fe]$  vs.  $[Fe/H]$  Tinsley diagram colour-coded by stellar age. Top panel: scatter plot for the high-resolution solar-vicinity sample of Delgado Mena et al. (2019), using the same symbols as in the original publication. Bottom panel: Median age per pixel in the  $[Mg/Fe]$  vs.  $[Fe/H]$  diagram for our 200 000 APOGEE DR17 giants. Overplotted are iso-density contours corresponding to 40, 80, and 160 stars per bin.

contours in this plot show that, in accordance with expectations from classical chemical evolution models (e.g. Chiappini 2009), the relationship is almost flat, except for the oldest ages, where we see a sharp transition from the high- $[Mg/Fe]$  regime to the low- $[Mg/Fe]$  regime, related to the onset of type-Ia supernovae (e.g. Mannucci et al. 2006; Mennekens et al. 2010). The colour code in Fig. C.3 shows how the age- $[Mg/Fe]$  varies with guiding radius, from a shorter and steeper relation in the outer parts of the disc (yellow colours) to a very flat relation (for ages < 8 Gyr) inside the solar circle.

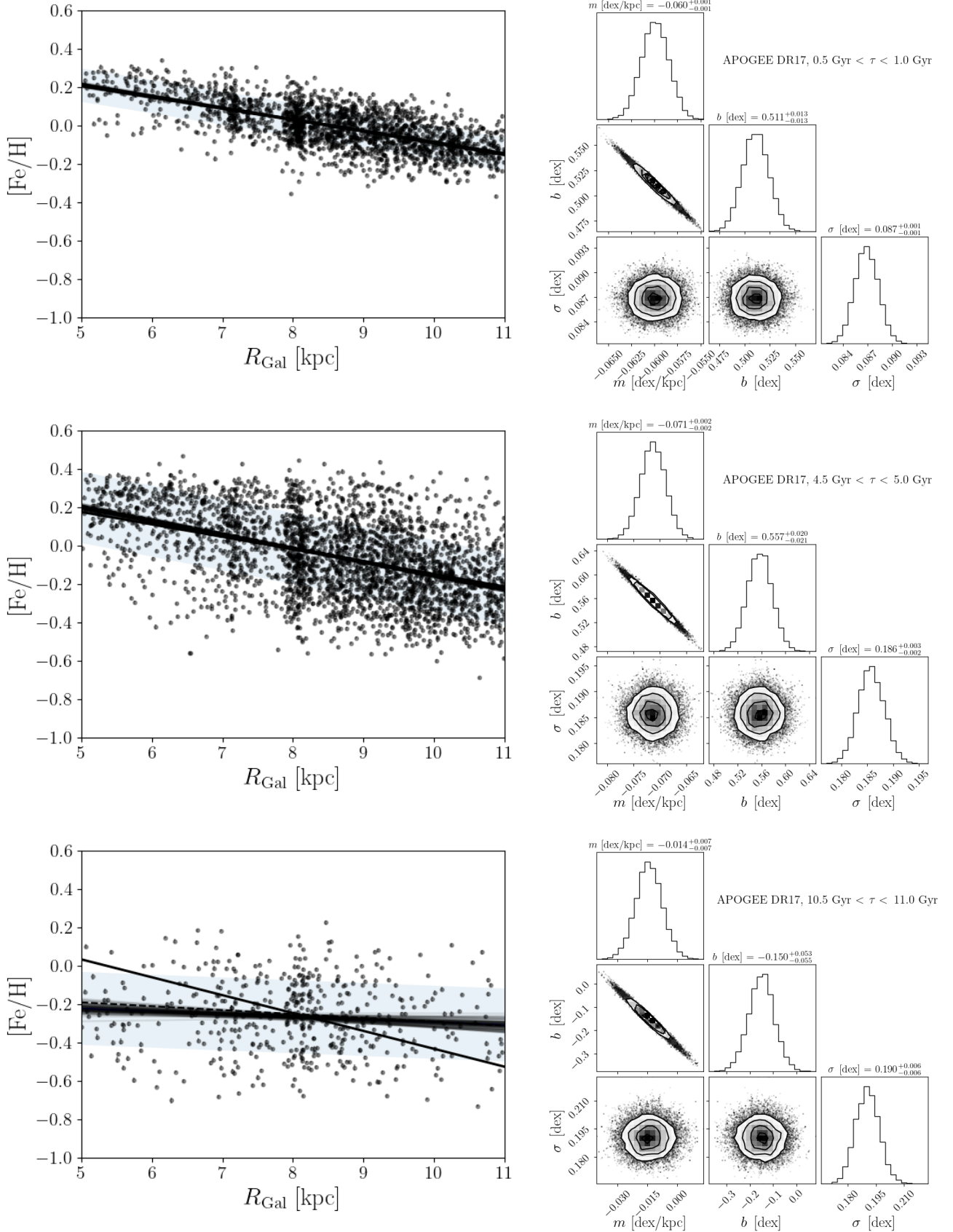


**Fig. C.3.** Age- $[Mg/Fe]$  relation, in the same style as Fig. 7. The small population of young  $\alpha$ -rich stars is highlighted.

Figure C.3 also highlights the presence of a small population of so-called young  $\alpha$ -rich stars. These stars, first discovered in APOGEE data (Chiappini et al. 2015; Martig et al. 2015), have since been demonstrated to be products of binary stellar evolution – they have been rejuvenated by mass transfer from a stellar companion (e.g. Jofré et al. 2016; Yong et al. 2016; Fuhrmann & Chini 2017, 2018; Hekker & Johnson 2019; Sun et al. 2020; Zhang et al. 2021; Jofré et al. 2023). Therefore, their (isochrone-derived) ages are catastrophically underestimated. The presence of the few young  $\alpha$ -rich stars in the full APOGEE sample (not present in the training set) reminds us that our age estimates are not completely based on the stellar photospheric compositions alone, but also depend on effective temperature and gravity, which may also mimic a younger stellar age for a given chemical composition. We therefore flag these stars as problematic in our catalogue (see Sect. A).

## Appendix D: Fitting the radial $[Fe/H]$ distributions

In Figs. 11 and 12 (Sect. 5.3) we presented the results of Bayesian fits to the radial  $[Fe/H]$  abundance distributions in age bins of 0.5 Gyr. The method (in particular, the implementation of the likelihood and posterior) is described in Anders et al. (2017a). In this appendix we show the detailed results of these fits for some age bins.



**Fig. D.1.** Examples of the fits to the  $[Fe/H]$  vs.  $R_{Gal}$  distributions for three age bins (from top to bottom: 0.5-1 Gyr, 4.5-5 Gyr, 10.5-11 Gyr). Left panels: distribution of the red-giant stars in each age bin. The thick black line shows the result of a naive least-squares linear fit. The thin grey lines show 30 realisations drawn from the linear gradient + intrinsic scatter posterior, while the shaded band corresponds the  $1\sigma$  dispersion around the gradient. Right panels: posterior distributions of the fit parameters ( $m = \partial[Fe/H]/\partial R$ ,  $b$  (intercept at  $R = 0$ ), and  $\sigma$  (intrinsic  $[Fe/H]$  dispersion)).