

## Spectroscopic ages for 4 million main-sequence dwarf stars from LAMOST DR10 estimated with data-driven approach

JIA-HUI WANG,<sup>1,2</sup> MAOSHENG XIANG,<sup>1,3</sup> MENG ZHANG,<sup>1</sup> JI-WEI XIE,<sup>4,5</sup> JIAN GE,<sup>6</sup> JINGHUA ZHANG,<sup>7,8</sup> LANYA MOU,<sup>1,2</sup> AND JI-FENG LIU<sup>1,2,3</sup>

<sup>1</sup>*National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100012, China*

<sup>2</sup>*School of Astronomy and Space Science, University of Chinese Academy of Sciences, Beijing 100049, China*

<sup>3</sup>*Institute for Frontiers in Astronomy and Astrophysics, Beijing Normal University, Beijing, 102206, China*

<sup>4</sup>*School of Astronomy and Space Science, Nanjing University, Nanjing 210023, China*

<sup>5</sup>*Key Laboratory of Modern Astronomy and Astrophysics, Ministry of Education, Nanjing 210023, China*

<sup>6</sup>*Shanghai Astronomical Observatory, Chinese Academy of Sciences, Shanghai 200030, China*

<sup>7</sup>*South-Western Institute for Astronomy Research, Yunnan University, Chenggong District, Kunming 650500, China*

<sup>8</sup>*CAS Key Laboratory of Optical Astronomy, National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100101, China*

### ABSTRACT

Stellar age determination for large samples of stars opens new avenues for a broad range of astronomical sciences. While precise stellar ages for evolved stars have been derived from large ground- and space-based stellar surveys, reliable age determination for cool main-sequence dwarf stars remains a challenge. In this work, we set out to estimate the age of dwarf stars from the LAMOST spectra with a data-driven approach. We build a training set by using wide binaries that the primary component has reliable isochrone age estimate thus gives the age of the secondary. This training set is further supplemented with field stars and cluster stars whose ages are known. We then train a data-driven model for inferring age from their spectra with the XGBoost algorithm. Given a spectral signal-to-noise ratio greater than 50, the age estimation precise to 10% to 25% for K-type stars, as younger stars have larger relative errors. Validations suggest that the underlying information used for our age estimation is largely attributed to the LAMOST spectral features of chemical abundances. It means our result is a manifestation of stellar chemical clock effectively acted on LAMOST spectra ( $R \simeq 1800$ ). Applying our model to the LAMOST DR10 yields a massive age catalog for  $\sim 4$  million dwarf stars. Statistical properties, such as the age distribution, age-abundance and age-stellar activity relations of the sample stars are discussed. The catalog is publicly accessible and can be helpful for extensive sciences from detection and characterization of Earth-like planets to Galactic archaeology.

*Keywords:* Stellar physics — Stellar spectroscopy — Machine learning

### 1. INTRODUCTION

Stellar ages play a crucial role in understanding the evolution of stars, the formation history of galaxies, and in the search for exoplanets and understanding their evolution (Soderblom 2010; Xiang & Rix 2022; Ge et al. 2022; Chen et al. 2023). Accurate and precise determination of stellar ages is thus of great importance but

usually a challenging task that relies on indirect estimation under the deep involvement of stellar evolution models (Soderblom 2010).

During the past decade, two age dating approaches, isochrone fit and asteroseismology, both based on stellar evolutionary models, have achieved remarkable progress in determining the ages of large samples of field stars. The first approach fits the measured stellar atmospheric parameters with predictions of stellar evolutionary models. This approach is especially effective for stars in the main-sequence turn-off (MSTO) and subgiant phases, during which their positions on the  $T_{\text{eff}}\text{-log } g$  (Kiel) or

wangjh@bao.ac.cn

msxiang@nao.cas.cn

temperature-luminosity (Hertzsprung-Russell; HR) diagram are most sensitive to their ages. Owing to the availability of precise stellar atmospheric parameters for millions of stars from large spectroscopic surveys, robust ages for millions of MSTO and subgiant stars have been derived (Xiang et al. 2015, 2017). Combing the spectroscopic parameters with precise luminosity from astrometric parallax of the Gaia mission (Gaia Collaboration et al. 2016; Brown et al. 2018, 2022), extensive works have been carried out to estimate the ages of stars across a broad range in the parameter space with isochrone fit (e.g. Sanders & Das 2018; Feuillet et al. 2019; Mints et al. 2019; Buder et al. 2021; Vickers et al. 2021; Xiang & Rix 2022; Anders et al. 2022; Sun et al. 2023; Wang et al. 2023). In particular, it has been suggested that, with such a method, stellar age determination can reach a high precision of only a few percent relative errors for subgiant stars (Xiang & Rix 2022; Nataf et al. 2024a).

Asteroseismic analysis of high-precision light curves can precisely yield the size and mass of stars, thus providing crucial constraints of age for evolved stars. A combination of oscillation frequency delivered from the Kepler mission (Borucki et al. 2010) with spectroscopic stellar atmospheric parameters has enabled robust asteroseismic age determination for thousands of evolved stars (e.g. Wu et al. 2017, 2018, 2023; Serenelli et al. 2017; Silva Aguirre et al. 2018; Miglio et al. 2021; Li et al. 2022).

On the other hand, the age determination for dwarf stars, which dominate the stellar populations, remains a great challenge. Asteroseismic parameters has been attempted to constrain the ages for main-sequence dwarf stars (Silva Aguirre et al. 2017), but it works mainly for stars with relatively warm temperatures ( $T_{\text{eff}} \gtrsim 5400$  K), whereas it is much harder for cooler stars as isochrones of different ages are almost indiscernible based on their mass and atmospheric parameters. Gyrochronology, which estimates the stellar age based on a color-dependent relation between age and rotation period, has been proposed as an empirical way of age dating for dwarf stars (Barnes 2003, 2007; Mamajek & Hillenbrand 2008; Angus et al. 2015; Lu et al. 2022; Ye et al. 2024). Similarly, several other empirical approaches, such as age-stellar activity relation and age-lithium abundance relation have been also proposed as a practical age clock for main-sequence dwarf stars (e.g., Skumanich 1972; Baliunas et al. 1995; Sestito & Randich 2005). For these approaches, a good calibration of the empirical age relations is essential but still far from perfect due to either observational challenges or intrinsic complications. As a consequence, the availability of these approaches is mostly restricted to stars younger

than the Sun (Baliunas et al. 1995; Sestito & Randich 2005; Meibom et al. 2015; van Saders et al. 2015; Tripathi et al. 2021).

Stellar chemical abundances have also been suggested to be good age indicators. First, observations show clear correlation among age, [Fe/H], and  $[\alpha/\text{Fe}]$  (e.g. Haywood et al. 2013; Xiang et al. 2017), as the metal-poor, high-alpha stars are generally old, while the metal-rich young stars are relatively young. Second, the abundance ratios that involve a neutron-capture element such as [Y/Mg] has been widely used as chemical clock (e.g. Skúladóttir et al. 2019; Berger et al. 2022; Hayden et al. 2022). Similar to other empirical age relations as mentioned above, a good calibration is essential for age estimation with the chemical clock. This especially matters considering that age-abundance relation may vary with location across the Galaxy (Ratcliffe et al. 2024). Currently, the application of chemical clocks are mainly restricted to stellar abundances from high-resolution spectroscopy, and the sample stars are thus mostly located at solar neighbourhood.

Alternatively, data-driven approach for stellar age estimation has been applied to large survey data sets, and delivered ages for millions of red giant stars (e.g. Martig et al. 2015; Ness et al. 2016; Ho et al. 2017; Wu et al. 2019; Huang et al. 2020; Zhang et al. 2021). This is mainly attributed to the availability of a good set of age reference stars, in particularly the asteroseismic age sample as mentioned above.

In this work, we set out to derive the age of dwarf stars from the LAMOST spectra with a data-driven approach. The LAMOST spectra cover the full range of optical wavelength window (Cui et al. 2012; Yan et al. 2022), such that in principle they contain multiple pieces of information that can indicate the age of a star, for instance, the stellar atmospheric parameters and the chemical abundances. Compared to that for MSTO and subgiant stars, the stellar atmospheric parameters for main-sequence dwarf stars show much less sensitivity on their ages. Therefore, age estimation method suitable for the formers such as isochrone fitting may fail to yield realistic results for the latter, resulting age estimates with both poor precision and large deviation. A data-driven approach is thus an optimized choice to derive ages from the LAMOST spectra for dwarf stars.

The foundation for a data-driven approach is a benchmark (training) sample of stars with precise age determinations. Wide binaries (WB) provide a good solution for this purpose. Millions of wide binaries have been identified from the Gaia astrometric survey (El-Badry & Rix 2018; El-Badry et al. 2021). The component stars in a wide binary system have identical ages but evolve

independently, so that for a binary system composed of a dwarf and a subgiant star, the dwarf's age can be known from the latter. Unlike other tracers that may be biased in parameter space (e.g., age, metallicity, HR diagram), such as open and globular clusters, wide binaries extensively distribute in a broad range of parameter space, making them a good training set for data-driven age estimation for large surveys like LAMOST.

We adopt wide binaries, accomplished with some cluster member stars and field stars that have LAMOST spectra and known ages as the training set, and use the XGBOOST algorithm to build the data-driven model for the age estimation. The paper is organized as follows. Section 2 describes our training samples. Section 3 introduces our data-driven model. Section 4 presents the resultant stellar age catalog and validations. Section 5 presents a discussion of the results. Section 6 concludes.

## 2. DATA

### 2.1. The LAMOST DR10 spectra

The tenth data release (DR10) of the LAMOST Galactic surveys (Zhao et al. 2012; Deng et al. 2012; Liu et al. 2014b; Yan et al. 2022) released 11,817,430 low-resolution ( $R \simeq 1800$ ) spectra, covering the full optical range of 3700–9000Å. The spectroscopic targets are uniformly drawn from the photometric color-magnitude diagram (CMD) with simple yet non-trivial selection function (Yuan et al. 2015; Liu et al. 2017; Chen et al. 2018), leading to a massive spectroscopic sample covering the full stellar parameter space. While the majority of the spectroscopic stars are main-sequence dwarfs, as they are the dominant population of our Galaxy, a large number of stars in brief evolution phases, such as the main-sequence turn-off and subgiant phases, are also effectively sampled (e.g. Xiang et al. 2017; Xiang & Rix 2022).

The allocation of 4000 LAMOST fibers is implemented with 4000 robots, each having an allocation area of 2 arc minutes. As a consequence, only one star can be allocated within a 2-arcmin area for a single visit to avoid fiber cross-talk. However, because the survey has multiple (usually 2 to 3) visits and brightness categories of plates (very bright, bright, medium bright, faint) for a given field defined by a central bright star, as well as overlaps of the adjacent fields, stars can be sampled with much higher spatial density (Liu et al. 2014a; Yuan et al. 2015). This is important for our purpose as we define training sample using wide binary systems that both of the components are observed by LAMOST.

In this work, we use normalized spectra for the age estimation. Our normalization scheme is the same as Xiang et al. (2019). Specifically, the LAMOST spectra

$f(\lambda)$  are divided by a smoothed version of itself  $\bar{f}(\lambda)$  by convolving with a Gaussian kernel,

$$\bar{f}(\lambda_i) = \frac{\sum_j f(\lambda_j) \omega_j(\lambda_i)}{\sum_j \omega_j(\lambda_i)}, \quad (1)$$

$$\omega_j(\lambda_i) = \frac{1}{\sqrt{\pi}L} \exp\left(-\frac{(\lambda_j - \lambda_i)^2}{L^2}\right), \quad (2)$$

where the Gaussian kernel width  $L$  is adopted as a constant value of 50Å. The normalized spectrum  $f_n(\lambda)$  is then obtained via

$$f_n(\lambda) = \frac{f(\lambda)}{\bar{f}(\lambda)}. \quad (3)$$

### 2.2. The age training sample

To obtain a good training sample is the key for data-driven age estimation. Our training sample is constructed combining four data sets,

- First, we build a sample of about 1200 wide binaries for which both components have a LAMOST spectrum, and the primary stars have an effective temperature higher than 5400 K to allow reliable isochrone age estimation.
- Second, to increase the number of old stars in the training sample, we supplement the wide binary sample with about 900 stars that have  $T_{\text{eff}} > 5400$  K and an isochrone age older than 6 Gyr, with age uncertainty less than 10%.
- Third, we further supplement the training sample with 400 high- $\alpha$ , old stars that have  $T_{\text{eff}} < 5400$  K and age older than 9 Gyr. Their ages are inferred from the tight age-metallicity relation of high- $\alpha$  stars.
- Finally, we add the member stars from a number of open clusters to the training set.

Approximately 30% of the spectra in LAMOST DR10 are repeat observations. In our training sample, for stars with repeat observations, we only use the spectrum with the highest signal-to-noise ratio ( $S/N$ ).

#### 2.2.1. The wide binary sample

We cross-match LAMOST and the Gaia wide binary catalog of El-Badry et al. (2021) using a 3-arcsec criterion, approximately the aperture size of LAMOST fibers. For the Gaia wide binary catalog, we only adopt those with higher than 90% confidence of being true binary, while discarding the others to avoid chance alignment. We restrict our training sample to binaries that both components have LAMOST spectra with  $S/N$

higher than 20 in the SDSS  $g$  band. We further require all the stars to have stellar parameters and abundances estimates from the LAMOST DR9 abundance catalog (Zhang et al. in prep.), which is an updated version of the LAMOST DR5 abundance catalog derived with the DD-PAYNE (Xiang et al. 2019). These criteria lead to 2893 pairs of binaries.

El-Badry et al. (2021) selected binary candidates based on their positions, parallaxes, and proper motions from Gaia eDR3. To obtain a cleaner wide binary sample, we further examine their radial velocity (RV) and metallicity ([Fe/H]) from the LAMOST spectra, which were not considered in El-Badry's work. Stars in a wide binary system are expected to share similar radial velocities, with an RV difference no larger than the maximal relative motion of the components w.r.t. the mass center, which is a few km/s in typical given the wide orbit. Also, the component stars are expected to have almost identical metallicity as they are siblings born from the same gas cloud.

Figure 1 presents a comparison of the RV and [Fe/H] between the binary components. To identify outliers from the comparison, we first define the RV and [Fe/H] criteria as in Eq.4 and Eq.5:

$$\Delta_{\text{RV}} = \frac{(\text{RV}_1 - \text{RV}_2)}{\sqrt{\sigma_{\text{RV}_1}^2 + \sigma_{\text{RV}_2}^2 + \sigma_{\text{intrinsic}}^2}}, \quad (4)$$

$$\Delta_{[\text{Fe}/\text{H}]} = \frac{([\text{Fe}/\text{H}]_1 - [\text{Fe}/\text{H}]_2)}{\sqrt{\sigma_{[\text{Fe}/\text{H}]_1}^2 + \sigma_{[\text{Fe}/\text{H}]_2}^2 + \sigma_{\text{intrinsic}}^2}}, \quad (5)$$

where  $\text{RV}_1$  ( $[\text{Fe}/\text{H}]_1$ ) refers to the radial velocity (metallicity) of the primary star that has bright Gaia G magnitude, and  $\sigma$  refers to the measurement uncertainty.  $\sigma_{\text{intrinsic}}$  refers to the intrinsic differences in binaries. In the processing, we assume  $\sigma_{\text{intrinsic}}$  is 0, which will lead us to a more stringent judgment of true binaries, but give us a purer sample. Binaries that either the RV or the [Fe/H] between the components deviate by larger than  $3\sigma$  are treated as chance alignment and discarded from the sample, except for cool dwarfs with  $T_{\text{eff}} < 4800$  K and  $\log g > 4$ . For these cool dwarfs, only the RV criterion is applied, as their [Fe/H] estimate may suffer large systematics (e.g. Niu et al. 2023).

Furthermore, only the binary systems in which the primary has  $T_{\text{eff}} > 5400$  K, which allow robust isochrone age estimation from their atmospheric parameters, are adopted as our training sample. For subgiant stars specifically, we adopt  $T_{\text{eff}} > 5100$  K instead of  $T_{\text{eff}} > 5400$  K (Figure 2).

Ultimately, these criteria leads to 1213 pairs of wide binaries. The right panel of Figure 2 illustrates that the

training sample covers a broad range of the HR diagram, and has especially a good sampling for the cool dwarf stars.

We determine the age of the primary star components by fitting their atmospheric parameters to the MESA Isochrones & Stellar Tracks (MIST; Choi et al. 2016) with a Bayesian approach, utilizing the DD-Payne spectroscopic parameters ( $T_{\text{eff}}$ ,  $M_{K,\text{spec}}$ ,  $[\text{M}/\text{H}]$ ) of LAMOST DR9, multi-band photometric magnitudes of the Gaia Riello et al. (2021); Gaia Collaboration et al. (2021) and 2MASS (Skrutskie et al. 2006), as well as the Gaia astrometric eDR3 parallax (Gaia Collaboration et al. 2021) as observables. Here the overall metallicity  $[\text{M}/\text{H}]$  is converted from the DD-Payne [Fe/H] and  $[\alpha/\text{Fe}]$ , following the method of Salaris et al. (1993). More details about the age determination are referred to Xiang & Rix (2022).

### 2.2.2. The old, warm star sample

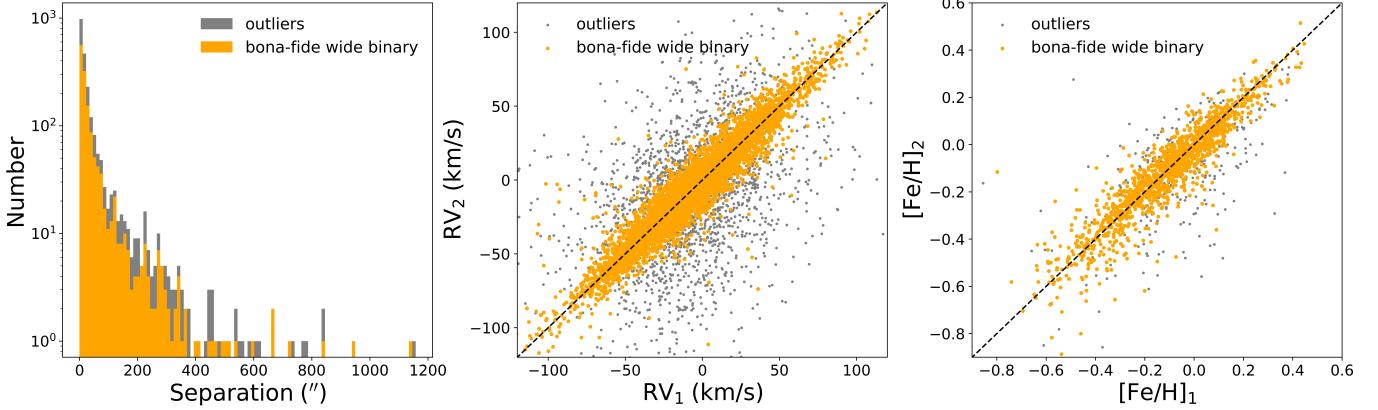
We found that the number of old stars in the binary sample is too small to allow for building a robust data-driven model for age estimation. We therefore supplement the wide binary sample by some field stars that have precise isochrone ages to increase the number of old stars in the training sample.

In doing so, we first determine the lower temperature border below which the isochrone age estimate becomes unreliable, utilizing wide binaries as the test beds. We adopt binary systems for which the primary have an effective temperature higher than 5400 K, and investigate the differences in the isochrone age estimates between the secondary and primary as a function of effective temperature of the secondary. As shown in Figure 3, at the high temperature side the ages of the secondary and primary are estimated consistently, with only a small difference. However, the scatter becomes larger as the temperature decreases. Especially, for stars below about 5400 K, a large proportion of them shows large discrepancy, reflecting the fact that isochrone age estimates in regime is hard.

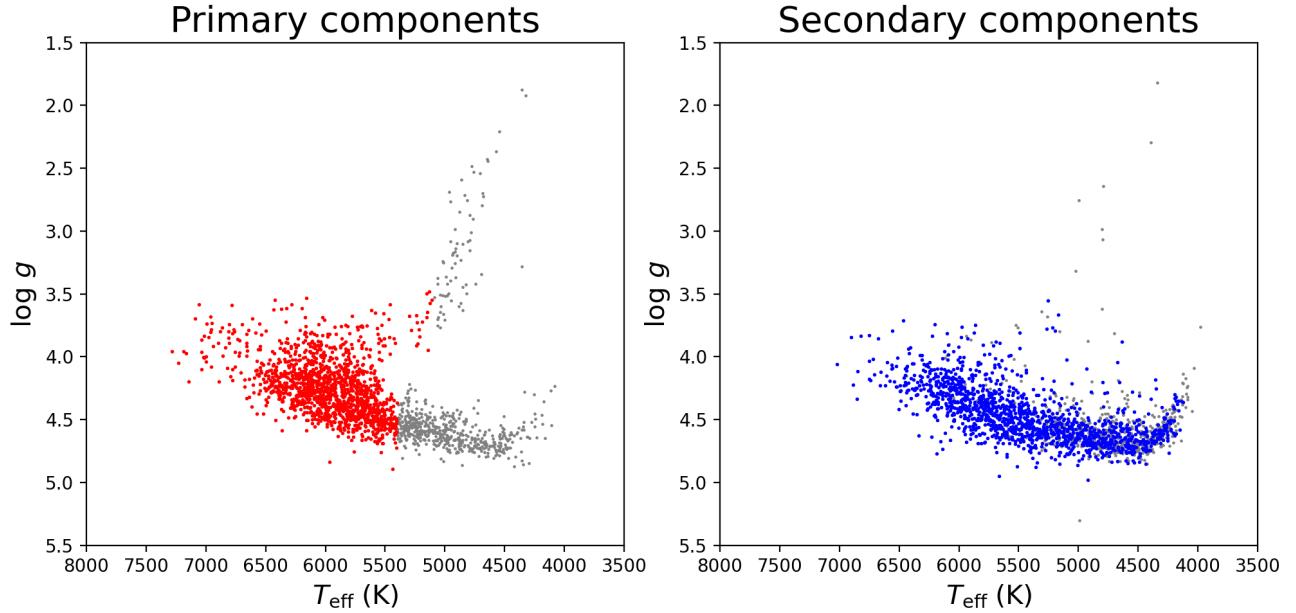
Therefore, in order to make use of the isochrone ages, we only add main-sequence stars with effective temperature higher than 5400 K as a supplementary to wide binary sample. We uniformly select stars based on their effective temperature and ages, requiring an age estimate older than 6 Gyr, and relative age error smaller than 10%. Ultimately, the selection leads a sample of 927 old field stars.

### 2.2.3. The old, cool star sample

For main-sequence stars cooler than 5400 K, the isochrone ages become problematic, and alternative method is required to obtain accurate ages. Xiang &



**Figure 1.** The wide binary sample. The left panel shows the distribution of separation angle of the wide binaries, with a lower separation angle cut of  $3''$ . The middle and right panel show the comparisons of radial velocity and metallicity, respectively, between the individual component stars of the wide binary systems. The orange color delineates bona-fide binaries selected from the radial velocity and metallicity criteria (see text). The gray color delineates outliers, which are probably recognized as wide binaries because of chance alignment.



**Figure 2.** Distribution of wide binaries in the  $T_{\text{eff}} - \log g$  diagram. The left and right panels show respectively the primary and secondary components, defined in (El-Badry et al. 2021). The colored dots are reference stars for age estimation in the current work. For these binaries, the primaries are in either subgiant or main-sequence (turn-off) evolutionary phase and their ages can be reliably determined with isochrone fitting, while the secondaries have a broad distribution in temperature, and are adopted as the training set of our data-driven age estimation.

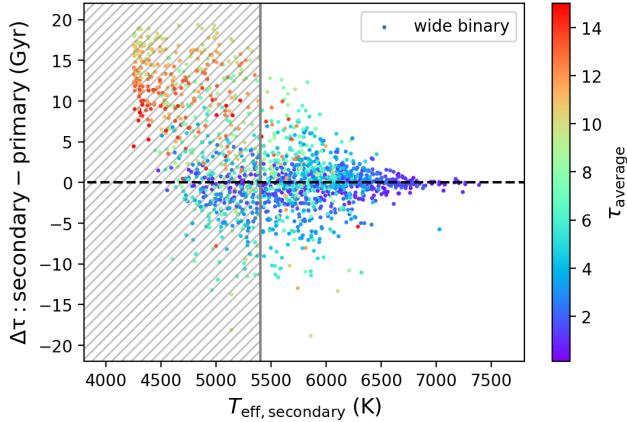
Rix (2022) found that for high- $\alpha$  stars, there exists a very clear and tight age-metallicity relation in the range of  $-1.0 \lesssim [\text{Fe}/\text{H}] \lesssim 0.5$ , and  $8 \lesssim \tau \lesssim 13$  Gyr, and they provided a linear fit to the relation as  $\bar{\tau}([\text{Fe}/\text{H}]) = 9.38 + (-3.79) \times [\text{Fe}/\text{H}]$ . Intrinsic dispersion to this relation is estimated to be only  $\sim 0.8$  Gyr.

In order to add cool old stars into the training set, here we randomly select 400 cool ( $T_{\text{eff}} < 5400$  K) stars that belong to the high- $\alpha$  sequence in the  $[\text{Fe}/\text{H}]-[\alpha/\text{Fe}]$

plane, and estimate their ages using the above age- $[\text{Fe}/\text{H}]$  relation. We assign a 10% error in the resultant age estimate.

#### 2.2.4. Age training sample from clusters

In addition to all the samples listed above, we also include member stars of 4 open clusters (OCs) observed by LAMOST in our training sample. OCs are good complementary training samples at the younger end



**Figure 3.** The difference in isochrone age estimates between the secondary and primary stars of wide binaries as a function of effective temperature of the secondary. All the primary stars have an effective temperature higher than 5400 K. The color represents the average age of the binaries. The shaded area delineates regime where the isochrone ages are unreliable.

( $\tau \lesssim 1$  Gyr). All member stars in a single cluster are assigned to be a constant age, adopted from literature value. Table 1 presents a detailed list of the information for the clusters adopted. For cluster stars, we assign 10% error in their ages.

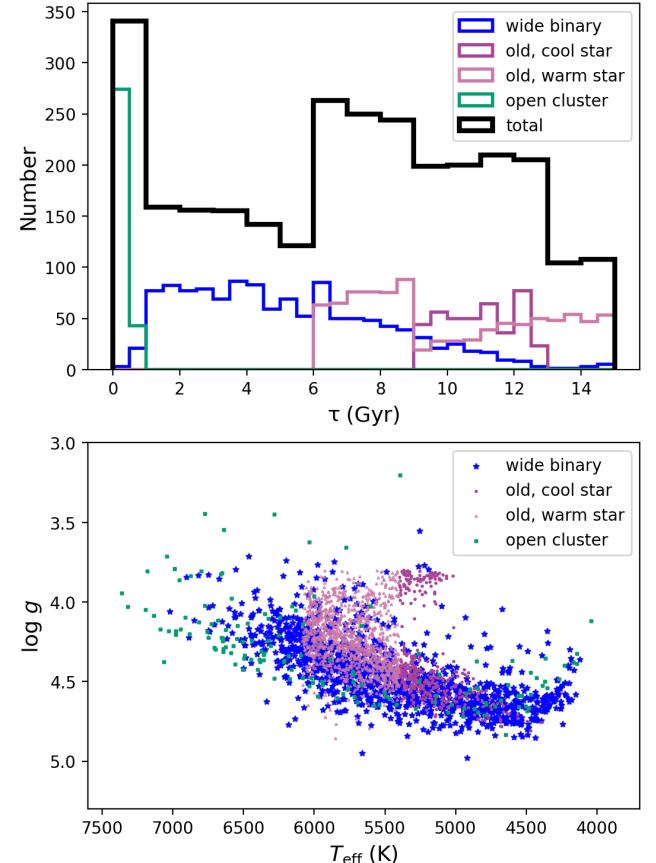
The age distribution of all training samples is shown in Figure 4, with different components marked in different colors.

### 3. METHOD

In this section, we introduce our data-driven model for age determination based on XGBOOST algorithm, as well as data pre-processing with principal component analysis (PCA), followed by an examination of the algorithms with a validation set.

#### 3.1. *EXTreme Gradient Boosting*

Extreme Gradient Boosting (XGBOOST) is a powerful machine learning algorithm used for solving classification and regression problems (Chen & Guestrin 2016). It is an ensemble learning method that boosts the performance of predictions by combining multiple weak learners, typically decision trees, into a strong learner. The core idea behind XGBOOST is to iteratively train a sequence of decision tree models and adjust the weights of each model based on their prediction errors. XGBOOST uses an objective function to guide model training and optimizes model performance by minimizing the objective function, which contains the loss function  $\mathcal{L}$  as well as the regularizations.



**Figure 4.** Distribution of the training samples. The upper panel shows their age distribution. Different colors represent different types of samples as marked in the figure. Black thick line represent the overall age distribution. The lower panel shows their distribution in the  $T_{\text{eff}}\text{-log } g$  diagram.

In our model, the spectra after principal component analysis (PCA) are used as inputs to the XGBOOST model, which is a  $2857 \times 20$  dimensional array. We construct the regression relationship between spectra and stellar age by XGBOOST. We customized the loss function  $\mathcal{L}$  and considered the age error obtained from the MIST isochrones. Prior to this we mapped age and age error to logarithmic space as shown in Eq.6.

$$\begin{aligned} \log \tau &= \log_{10} \tau \\ \sigma_{\log \tau} &= \sigma_\tau / \tau \log_{10}(e) \\ \mathcal{L} &= \sum_i \left( \frac{\log \tau_{\text{pred},i} - \log \tau_{\text{ref},i}}{\sigma_{\log \tau,i}} \right)^2 \end{aligned} \quad (6)$$

where  $\tau$  is the reference age of the training sample from the MIST stellar isochrones,  $\sigma_\tau$  is the age error of the training sample.  $\log \tau_{\text{pred}}$  is the age predicted by the XGBOOST model.

In order to avoid overfitting and ensure the convergence of the training process, we fine-tuning the learn-

**Table 1.** Age of clusters adopted in our training sample

Cluster name	Number of stars	Age(Gyr)	Reference
Hyades	61	0.625	Perryman et al. (1998)
NGC2168	154	0.175	Bouy et al. (2015)
NGC2099	78	0.346-0.55	Wu et al. (2009); Hartman et al. (2009)
Pleiades	68	0.115	Basri et al. (1996)

NOTE— 1. "Number of stars" depends on the LAMOST observations.

2. Only half of the member stars are used for training, while the other half is used for validation.

ing rates, number of iterations, and hyperparameters for the model.

### 3.2. Spectral Principal Component Analysis

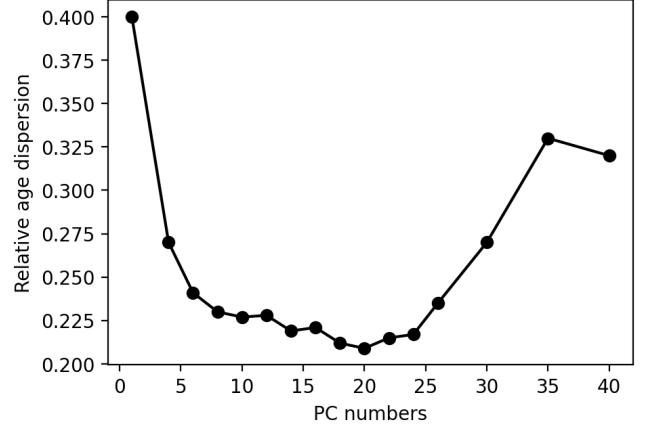
Robustness of a machine learning method is usually very sensitive to data noise. For the LAMOST spectra, most of the features that are informative for stellar physical parameters are in the blue wavelength regime( $\lesssim 5000 \text{ \AA}$ ), which only has low  $S/N$  for cool stars. Therefore, a de-noising pre-processing of the spectra is helpful to improve the robustness of the age estimates. Here we perform a principal component analysis (PCA) on the spectra for such a purpose.

PCA converts, in a variance-conservative manner, most of the information of a high-dimensional data set into a small number of principal components via linear combinations of orthogonal basis functions, while leaving the noise information to other components with small variance. Reconstruction of a noise-reduced spectrum with PCA has been applied to the LAMOST spectral analysis in previous work (e.g. Xiang et al. 2021). Here we only briefly introduce our model choice, but we omit the math aspect.

We use the full sample of 2857 spectra to compute the principal components (PCs) and choose the first 20 PCs for the spectra reconstruction. Here the number of PCs adopted is determined by a brute-force search approach. Specifically, we examine a series of PC numbers from 1 to 40, and implement the age regression on a validation set and calculated the relative age dispersion to determine the optimal number of PCs. As shown in Figure 5, using 20 PCs yields the best performance, and the precision of the age estimates no longer improves with increasing PC numbers. We then use the PCA spectra for age regression. Figure 6 shows an example of the PCA reconstructed spectra as well as the noise spectra (residuals).

### 3.3. Validating the age estimation with validation set

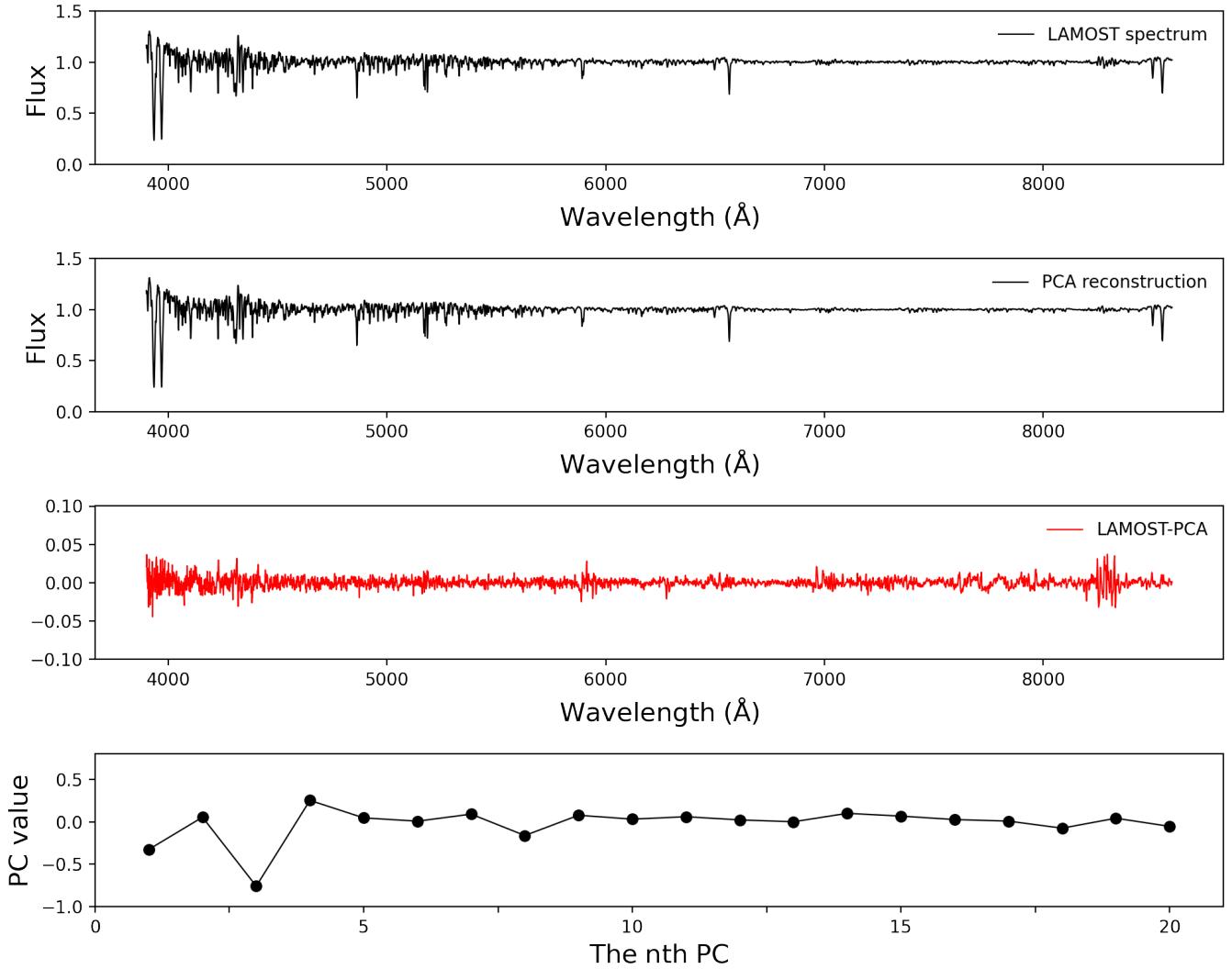
We divide our age sample into training and validation sets. Eighty percent of the wide binary stars and field

**Figure 5.** The relationship between the number of PCs and the precision of age estimates.

stars are adopted as the training set, while the remaining 20% are taken as the validation set. For each of the OCs, half of the member stars are used for training and the other half for validation. As the number of training samples is limited, the XGBOOST model training may suffer non-negligible uncertainty due to too sparse sampling. We therefore perform 20 training realizations with a bootstrap approach. In each realization, we randomly select the training and validation sets. Age predictions from the 20 training realizations are combined by taking the average value to minimize possible method error due to the training process. Specifically, the average age for the  $i$ th star in training (validation) set is defined as follows:

$$\bar{\tau}_{\text{pred},i} = \frac{1}{N_i} \sum_{j=1}^{N_i} \tau_{\text{pred},i,j} \quad (7)$$

$N_i$  is the number of realizations that the  $i$ th star is used for training (validation), which ranges from 2 to 20 as we require at least two realizations to compute average age.  $\tau_{\text{pred},i,j}$  denotes the predicted age of the  $i$ th star using the  $j$ th training realization. It is worth noting that for the final LAMOST DR10 age catalog (see



**Figure 6.** An example for PCA reconstruction of the LAMOST spectra for de-noising. From top to bottom, the first panel is the normalized LAMOST spectrum. The second panel is the PCA-reconstructed spectrum. The third panel is the residual by subtracting the PCA-reconstructed spectrum from the LAMOST spectrum. The fourth panel is the value for the first 20 principal components (PCs).

Section 4.1), the ages are averaged over 20 realizations for all stars.

Figure 7 shows the comparison of the  $\bar{\tau}_{\text{pred},i}$  with the ground-truth values for both the training and validation sets. The figure shows a good agreement between the predicted age and the ground-truth age, except for the youngest and oldest end. For the validation set, the youngest stars ( $< 0.5$  Gyr) may be overestimated significantly, while quite a number of the oldest stars ( $\gtrsim 13$  Gyr) are significantly underestimated (a similar effect is also found for the training set). However, the agreement on the whole is remarkable: the relative age dispersion for the overall sample is 15% for the training sample, and 18% for the validation sample. Given these

two numbers are close to each other, we believe there is no significant overfitting in our model training process.

As our training set is a mix of several sources, we found that in the validation set the relative age dispersion is  $\sim 23.7\%$  for wide binary, and  $\sim 12.5\%$  for field stars. The latter is smaller because we have set much more strict quality cut, in terms of both spectral  $S/N$  and age precision, for selecting the training and validation set, as described in Sect. 2.2. When dividing the validation set in hot ( $T_{\text{eff}} > 5400\text{K}$ ) and cool samples, we found that the relative age dispersion is  $\sim 19\%$  for the hot sample, and  $\sim 15\%$  for the cool sample. The latter is lower because of two main reasons: first, the hot sample contains the majority of open cluster member stars, whose ages are young and thus have large relative errors

(Sect. 5.1). Second, the cool sample contains the vast majority of old, cool high-alpha stars, which have good age estimates through the tight abundance-age relationship. After removing these old, cool high-alpha stars, the age dispersion for the cool sample becomes 23%.

### 3.4. Age error estimation

To assign an error to each age estimate, we perform an Monte Carlo (MC) analysis by adding random Gaussian errors to the original spectrum and deriving the age repeatedly. Thirty iterations are implemented for each spectrum, and the standard deviation of the individual age estimates is adopted as the error estimate. Such an analysis takes the measurement errors of the spectral flux into account, thus giving a statistical uncertainty of the age estimate. A complete treatment of the age uncertainty should also consider other error sources, such as intrinsic error of the method, systematic age errors in the training set, etc. We will examine possible systematic errors in the age estimates in Section 4.

## 4. RESULT

We applied our data-driven model to all low-resolution spectra of dwarf and subgiant stars in LAMOST DR10. In this section, we present statics of this large stellar age sample, and explore the correlation between age and other stellar physical quantities, including elemental abundances and activity.

### 4.1. The LAMOST DR10 stellar age sample

Owing to limitation of the training set, here we confine the analysis only to dwarf and subgiant stars. Therefore, we exclude red giants from the age sample, and the selection of red giants adopts the following criteria:

$$\log g \lesssim 3.8 \quad (8)$$

$$\log g \lesssim -0.0045 \times T_{\text{eff}} + 27.1 \quad (9)$$

as shown in the shaded window in Figure 8. The  $T_{\text{eff}}$  and  $\log g$  are from the updated LAMOST DR10 catalog (Mou et al. in prep.). In addition, we restrict the  $T_{\text{eff}}$  to the range of 4000–6800 K, in accordance with the coverage of the training sample. These criteria lead to a catalog of 3,868,513 stars in our resultant age catalog <sup>1</sup>

Figure 9 shows the spatial distribution of the number density and the age trend of the remaining sample. The upper panels of the Figure shows that our sample covers a large portion of the Galactic halo and anti-center disk, with a particularly dense sampling of the nearby

disk ( $8 \lesssim R \lesssim 10$  kpc). The mean stellar age shows a clear increasing trend with Galactic latitudes (bottom-left panel), and the age distribution in the  $R$ - $Z$  plane (bottom-right panel) of the disk exhibits a negative radial age trend, along with a strong flaring phenomenon. These features are consistent well with previous work (Ness et al. 2016; Martig et al. 2016; Xiang et al. 2017; Sanders & Das 2018; Wu et al. 2019; Huang et al. 2020; Anders et al. 2022).

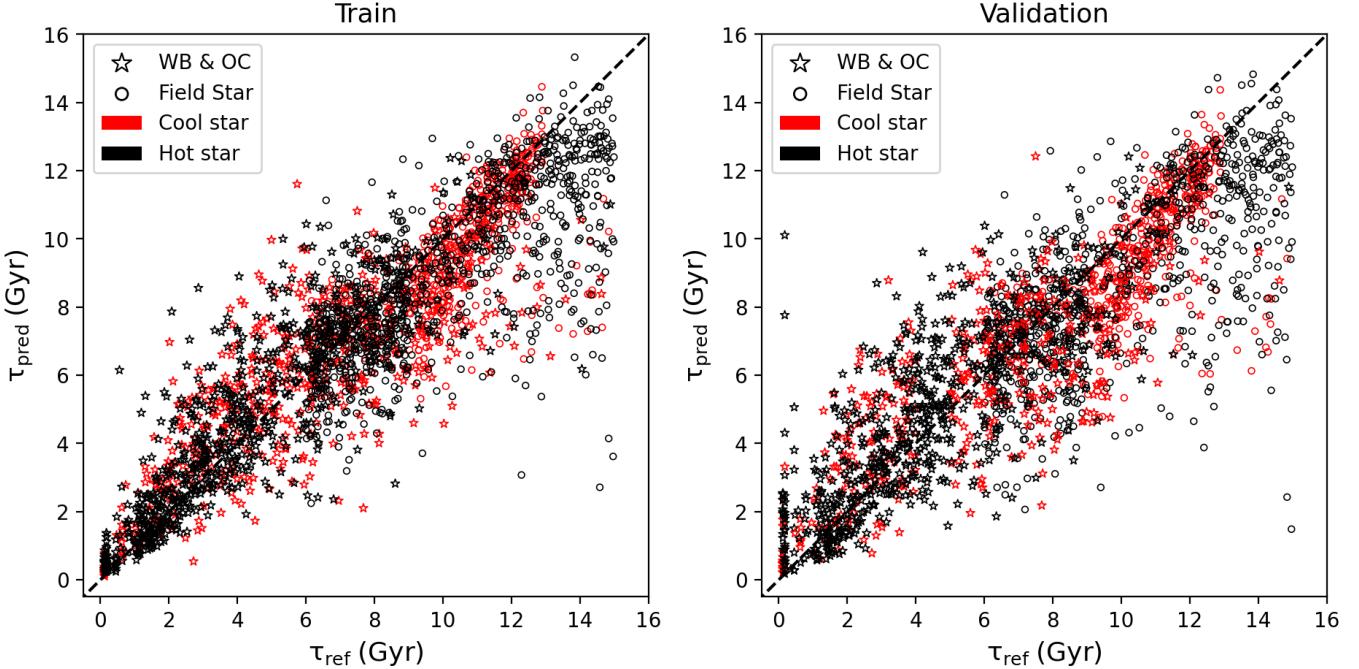
The bottom-right panel of Figure 9 also shows an unexpected trend in the inner disk of  $R < 8$  kpc that younger ages appear to occur at higher places (larger  $Z$ ) as  $R$  decreases. This is in contrast with the expectation that the inner and higher disk is older, reflecting an earlier, turbulent assembly history of the Galactic formation. However, as has been discussed in Xiang et al. (2017), this is an artifact due to the selection effect. Young stellar populations are intrinsically more luminous so that they probe to a larger distance along each sightline of the survey. Such an effect is especially severe for the inner disk because in those sky areas usually only bright stars (VB plates;  $r < 14$  mag) are targeted by the LAMOST survey. An accurate mapping of the age distribution requires a detailed correction of the survey selection effect, which is beyond the scope of this paper.

A large fraction ( $\sim 30\%$ ) of spectra in our sample are repeat observations of common stars. As a validation of the age estimates, Figure 10 shows the dispersion of  $\log(\text{age})$  between repeat observations as a function of spectral  $S/N$ . Here for repeat observation samples, we restrict to those that have comparable  $S/N$  ( $0.8 < \frac{S/N_1}{S/N_2} < 1.2$ ). The dispersion is then computed with a Gaussian fit to the  $\log(\text{age})$  difference between the repeat observations, and divided by  $\sqrt{2}$ . The figure shows a dispersion decreasing from  $\sim 0.18$  (41% in linear space) at  $S/N = 5$  to  $\sim 0.08$  (18% in linear space) at  $S/N > 100$ , slightly depending on the spectral type. F/G-type stars with higher temperatures have lower dispersion than the cooler, K dwarfs, which is expected as the ages of K dwarfs are harder to estimate due to their more crowded distribution in the  $H$ – $R$  diagram, which means that their locations in  $H$ – $R$  diagram have less sensitivity to their ages.

### 4.2. Ages of star clusters

The ages of star clusters provide an independent validation for our method. Here we present the results for three open clusters targeted by the LAMOST survey, namely, Pleiades (M45), and NGC 2420, M67 (NGC 2682). We use Gaia astrometric data to select cluster member stars based on their spatial positions (RA, Dec) and proper motions (PMRA, PMDec). In

<sup>1</sup> The catalog are publicly available via <https://zenodo.org/records/15824429>



**Figure 7.** Comparison of the XGBoost age estimates and reference (ground-truth) ages for our samples. The left panel shows results for the training sample, while the right panel shows results for the validation sample. Note the vertical axis refer to the averaged ages from different traing realizations (see text). Different tracers, namely wide binaries and open clusters (WB & OC), and field stars, are shown different samples as marked in the Figure. Different colors represent different temperatures, with red for  $T_{\text{eff}} < 5400\text{K}$  and black for  $T_{\text{eff}} > 5400\text{K}$ . The dispersion in the relative age difference is found to be 15% for the training sample and 18% for the validation sample.

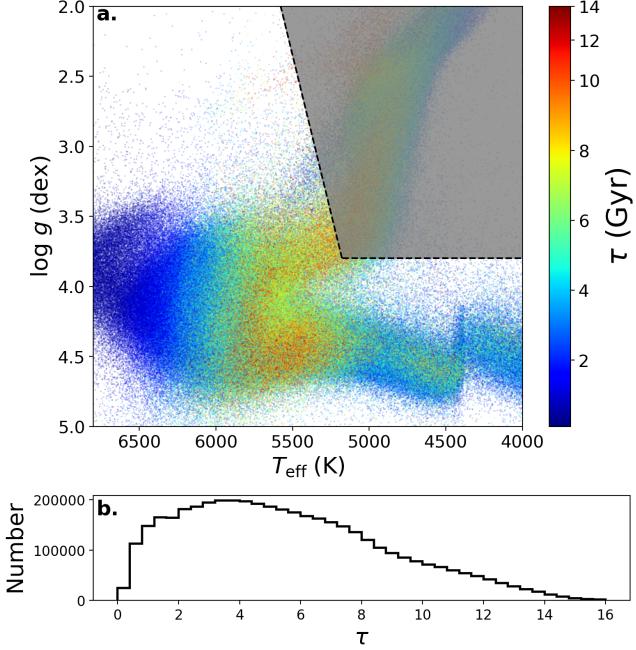
doing so, we first select LAMOST stars within a given radius around the cluster’s center, and characterize the distribution of proper motion for the remaining stars with Gaussian fit. Stars deviating from the mean values of the Gaussian fit by  $3\sigma$  are discarded as outliers. We restrict the analysis to stars that have a spectral  $S/N$  higher than 30. The resultant age distribution is shown in Figure 11.

For Pleiades, we adopt a cluster center of (RA, Dec)=(56.59359°, +24.12223°) and a radius of 220.2 arc minutes. The mean proper motion is (PMRA, PMDec) = (19.88, -45.41)mas/yr, and the dispersion ( $\sigma_{\text{PMRA}}$ ,  $\sigma_{\text{PMDec}}$ ) = (1.24, 1.43) mas/yr. This leads to 35 member stars in our sample, all of which were only used in the validation set while training the XGBoost model. The left panel of Figure 11 shows that their age distribution exhibits a peak at 0.39 Gyr, with a standard deviation of 0.23 Gyr. This result is systematically older than literature values (0.075–0.15 Gyr; Basri et al. 1996) by 0.3 Gyr. This is likely due to an overestimate of our results for stars at the youngest border of the training set. For stars with ages close to or beyond the boundaries of the training sample, the XGBoost method tends to generate a lower age at the old end and a higher age

at the young end. This is assumed to be a common bias effect for many machine learning methods (Ting 2024).

For NGC2420, we adopt a cluster center of (RA, Dec)=(114.6020°, +21.5750°) and a radius of 19.8 arc minutes. The mean proper motion is (PMRA, PMDec) = (-1.22, -2.05)mas/yr, and the dispersion  $\sigma$  is (0.10, 0.09)mas/yr. This leads to 20 member stars in our sample. The middle panel of Figure 11 shows that the member stars spread an age between 1 and 3 Gyr, with a significant peak at 1.81 Gyr, and the standard deviation is 0.79 Gyr, corresponding to a relative age uncertainty of  $\sim 45\%$ . The result is consistent with the literature, which is  $2.0 \pm 0.2$  Gyr (von Hippel & Gilmore 2000).

For M67, we adopt a cluster center of (RA, Dec)=(132.8460°, +11.8140°) and a radius of 49.8 arc minutes. The mean proper motion is (PMRA, PMDec) = (-10.96, -2.91)mas/yr, and the dispersion  $\sigma$  is (0.22, 0.21)mas/yr. This leads to 80 member stars in our sample. The right panel of Figure 11 shows that the member stars have a median age of 3.75 Gyr, and a standard deviation is 1.2 Gyr, corresponding to a relative age uncertainty of  $\sim 30\%$ . This result is also consistent well with the literature, which suggests 4 Gyr (e.g Richer et al. 1998).



**Figure 8.** Distribution of the LAMOST star sample. The upper panel (a) shows the distribution in the  $T_{\text{eff}}$ - $\log g$  diagram, color-coded by the stellar age estimates. To ensure precise age estimation, here only stars with  $S/N > 50$  are presented. The shaded area masks red giant stars that are not the focus of the current work. The bottom panel (b) shows the age distribution of the full stellar age catalog, with a total number of 4,233,948 stars.

#### 4.3. Age comparison with literature

Silva Aguirre et al. (2017) determined the ages of 66 main-sequence (turn-off) stars utilizing asteroseismic and spectroscopic parameters as the constraints to stellar evolution models. Most of these stars have effective temperatures greater than 5400 K, and are close to the main-sequence turn-off in the HR diagram, while only 3 stars have effective temperatures below 5400 K. We cross-matched this asteroseismic age sample with our LAMOST DR10 age catalog, and obtained 31 common sources. Silva Aguirre et al. (2017) employed seven different pipelines to calculate the ages, and the results among these pipelines showed a good consistency. Here we only compare our results with ages from their V&A pipeline. As shown in Figure 12, our age estimates show a good consistency with theirs, with a difference of approximately 20%.

Xiang et al. (2017) provided the ages and masses of approximately 930,000 MSTO stars and subgiants in the Galactic disk from the LAMOST survey<sup>2</sup>. They used a

Bayesian algorithm to infer these stellar parameters by matching the effective temperature  $T_{\text{eff}}$ , absolute magnitude  $M_V$ , metallicity [Fe/H], and  $\alpha$  abundance [ $\alpha/\text{Fe}$ ] with stellar isochrones. Here we compare our age estimates with Xiang et al. (2017) for common stars with  $S/N$  greater than 30. As shown in Figure 13, the results exhibit a clear and tight consistency. The error varies with age, from approximately 0.3 Gyr at 1 Gyr to about 1.5 Gyr at 11 Gyr. We notice a slight systematic difference in the figure at the younger part ( $\tau \lesssim 4$  Gyr). Upon inspection, we suspect this is likely caused by the different stellar evolution models used, as age of the training sample in the current work is determined using the MIST isochrone, while Xiang et al. (2017) determined ages using the Yonsei-Yale (YY) stellar isochrones (Kim et al. 2001; Demarque et al. 2004; Casamiquela et al. 2024; Nataf et al. 2024b). These two isochrones adopt different schemes of core overshooting, which has a significant impact on the age determinations of young stars with intermediate mass.

#### 4.4. Age-activity relation

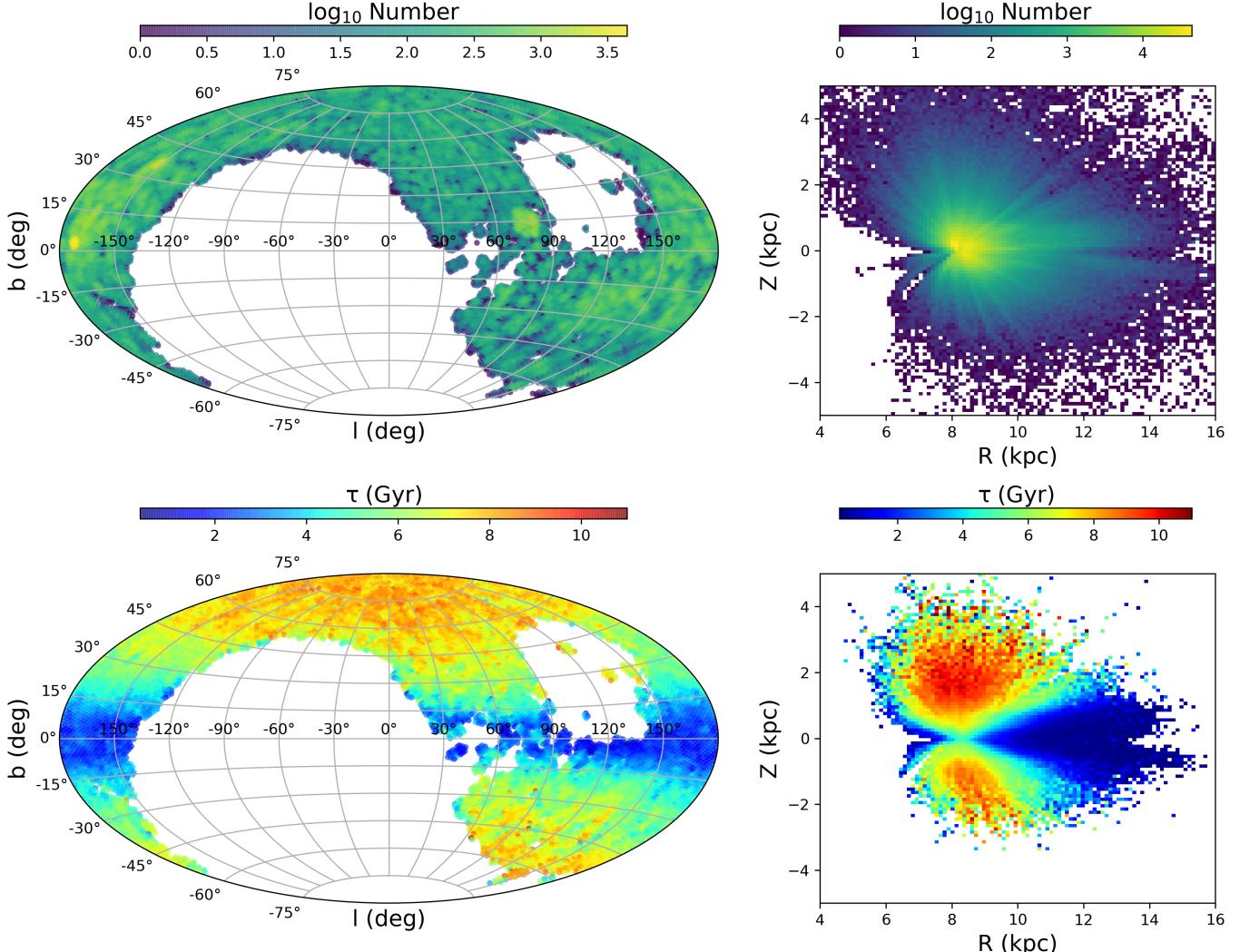
It has been extensively discussed that the chromospheric activity of a star correlates with its age, and this relationship has been effectively utilized to estimate the ages of young dwarf stars. (e.g. Skumanich 1972; Barnes 2010; Soderblom 2010; Chen et al. 2021; Ye et al. 2024; Han et al. 2024). Here we investigate the relation between age and stellar activity with our sample. We adopt the stellar S-index from Zhang et al. (2024) as the activity indicator. Figure 14 presents the stellar S-index as a function of age for different types of stars in our sample. All of the F-(6000-7000K), G-(5300-6000K), and K-type(4500-5300K) dwarf stars exhibit a clear age-activity relation for relatively young stars, as younger stars show higher S-index. For old stars, such a relation however disappears for all of the F-, G-, and K-type dwarfs, and the S-index exhibits a flat distribution with age. These trends are qualitatively consistent well with literature (Pace & Pasquini 2004; Zhao et al. 2011), validating the effectiveness of our age estimates.

Figure 14 also shows that the terminating age when the age-activity relation flattens varies with stellar types, from about 6 Gyr for K dwarfs to about 2 Gyr for F dwarfs. This indicates that as the temperature increases, the terminating age of dwarf stars moves toward the younger end. This feature is consistent with the prediction trend of theoretical evolution models.(Meibom et al. 2015; van Saders et al. 2016)

#### 4.5. Relation between age and chemical abundances

In this section, we investigate the relation between the age and [X/Fe] as well as [X/Mg]. For demonstrating

<sup>2</sup> <http://dr4.lamost.org/v2/doc/vac>

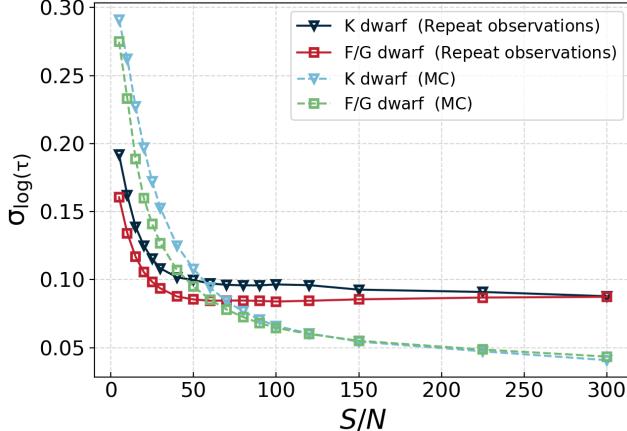


**Figure 9.** Spatial density and age distribution of our dwarf star sample. The left panels show the stellar distribution in the Galactic coordinate system ( $l, b$ ), color-coded by the stellar density (upper-left panel) and median age (lower-left panel). The right panels show the stellar distribution in the  $R$ - $Z$  plane of the Galactic cylindrical coordinate system, color-coded by the stellar density (upper-right panel) and median age (lower-right panel). Only stars with a spectral S/N higher than 30 are used to generate this figure.

our age estimation method, here we focus on the K dwarf sample with  $4800 < T_{\text{eff}} < 5400$  K, whose ages are hard to be determined precisely with isochrone fitting (Section 5.1). We note that, as one expects our data-driven age estimates are largely constrained by features of stellar abundances in the spectra, i.e., the age estimates are correlated with the abundances, the age-abundance trends shown below are thus not independent measurements. Nevertheless, it is still valuable to look into these trends as an internal examination of possible intrinsic age-abundance trends, and is particularly informative on which chemical species contribute the majority of the age information that our method has implicitly utilized.

The stellar abundances we utilized are derived from the LAMOST spectra with the DD-PAYNE. While the early version of LAMOST DD-PAYNE abundance catalog (Xiang et al. 2019) only contains elements with atomic numbers smaller than Zn (except for Ba), an updated version of DD-PAYNE abundance catalog on LAMOST DR9 has included the abundances for a number of heavier elements, including Sr, Zr, Y, Ce, Ba, Nd, La, Sm, and Eu (Zhang et al. in prep.). We thus use this updated LAMOST DR9 abundance catalog for our analysis.

Figure 15 shows the stellar number density distribution in age-abundance planes. In particular, the figure shows a clear separated double-sequence feature in



**Figure 10.** Measurement error of the logarithmic age as a function of  $S/N$ . Results for F/G (5300-7000 K) and K dwarf (4500-5300 K) stars are shown in different symbols, are labeled in the Figure. Both measurement error estimated with a Monte-Carlo approach (see text) and dispersion among repeat observations are presented, as labeled in the Figure.

the age-[Mg/Fe] plot. Stars in the sequence with higher [Mg/Fe] have older ages, with ages older than 10 Gyr, while stars in the low-[Mg/Fe] sequence are systematically younger<sup>3</sup>. Such a double-sequence trend is in good agreement with previous work (e.g. Wu et al. 2018). Beyond this double-sequence feature, Figure 15 shows clear age-abundance correlations for most of the elements. For stars older than 4 Gyr, the mean values of [C/Fe], [Mg/Fe], [Al/Fe], [Si/Fe], [Ca/Fe], and [Ti/Fe] show a monotonous decrease with decreasing age, the [Cr/Fe] and [Ni/Fe] show a nearly flat trend with age, while the mean values of [N/Fe], [Mn/Fe], and [Fe/H] show an monotonous increase with decreasing age. However, for stars younger than 4 Gyr, these trends are either flattened or reversed for some of the elements, such as [N/Fe], [Ca/Fe], [Ti/Fe], [Mn/Fe], and [Fe/H].

Similar age-abundance trends are also clearly presented for heavier, s- and r-process elements (Figure 16). There is a strong positive age trend for [Sr/Fe], [La/Fe], and [Nd/Fe] for stars older than 4 Gyr, but the trends are reversed for stars younger than 4 Gyr. Albeit less strong, similar trends for [Zr/Fe], [Sm/Fe], and [Eu/Fe] are also visible. The [Y/Fe] and [Ba/Fe] show almost a flat or slightly negative trend for stars older than 4 Gyr, but exhibit a clear negative trend for stars younger than 4 Gyr. The [Ce/Fe], however, exhibit a negative age trend for all ages.

<sup>3</sup> see Section A in the Appendix for more details about the old stars in low-[Mg/Fe] sequence.

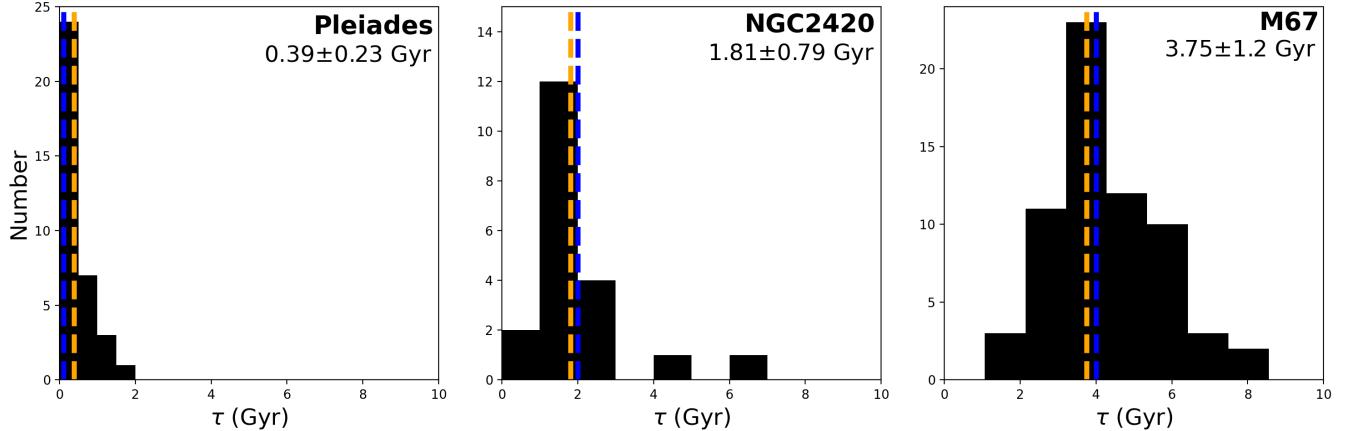
The age-abundance trends are qualitatively consistent with literature. Using data from the Gaia-ESO survey, Smiljanic et al. (2016) found that [Al/Fe] was positively correlated with age, which is consistent with our results. Sun et al. (2023) obtained the same trend of [Al/Fe] using MSTO from the GALAH survey. The slightly negative trend with age for [Ba/Fe] is consistent with the results of high-resolution spectroscopy (Delgado Mena et al. 2019), and with the GALAH survey (Hayden et al. 2022; Sharma et al. 2022). Sharma et al. (2022) also found that the age-[Eu/Fe] relationship showed a positive correlation trend, which is consistent with our result for stars old than  $\tau > 4$  Gyr. Ratcliffe et al. (2024) suggested that the age-abundance trend changes with the birth radius of the star. This is probably the reason for the broken age-abundance relations seen in our sample, as we expect the old stars in our sample to be mostly born at the inner disk, while the young stars ( $\tau \lesssim 4$  Gyr) are mostly born at the outer disk (e.g. Wu et al. 2021; Lu et al. 2022).

These well-presented age-abundance relations imply that the LAMOST spectra indeed contain a wealth of features that can be used to estimate the stellar age, at least from a chemical clock. The results also suggest that the DD-PAYNE abundance estimates are robust, and can provide constraints on their ages.

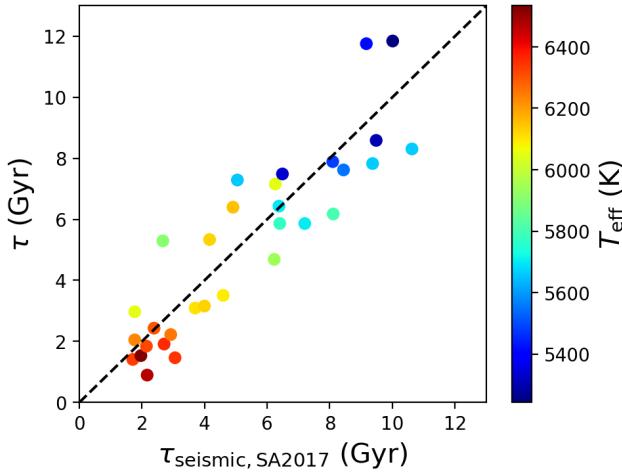
## 5. DISCUSSION

While we have shown that the XGBOOST regression approach can pragmatically estimate the age of cool main-sequence dwarf stars from the LAMOST spectra to an internal precision of 20-40%, it is important to understand the underlying clocks that our method has utilized implicitly for estimating the age. This is especially essential for better understanding limitations in the age estimates. For this purpose, we have made the following experiments:

- First, we examine the precision of age estimation with isochrone fitting method utilizing the stellar atmospheric parameters, i.e.,  $T_{\text{eff}}$ ,  $\log g$ , and [Fe/H], for a typical cool main-sequence star. A comparison of the precision between the isochrone age and the data-driven age provides insights into which extent the latter relies on clocks of stellar evolution.
- Second, we build chemical clocks to infer stellar age with LAMOST abundances. A comparison of precision between the chemical age and the data-driven age tells to which extent the latter relies on clocks of Galactic chemical evolution.



**Figure 11.** Age distribution of cluster member stars. From left to the right panels are the results for Pleiades (M45), NGC2420, and M67 (NGC2682). Median age and standard deviation of the individual member stars are marked in the Figure. The vertical orange line represents the median age, and the blue line represents the literature age.



**Figure 12.** A comparison of our age estimates with the asteroseismic ages of Silva Aguirre et al. (2017). Colors represent the effective temperatures from LAMOST spectra.

- Third, we inspect how the data-driven age estimation relies on stellar activity by masking activity-sensitive lines, including Ca II H&K and H $\alpha$  lines.

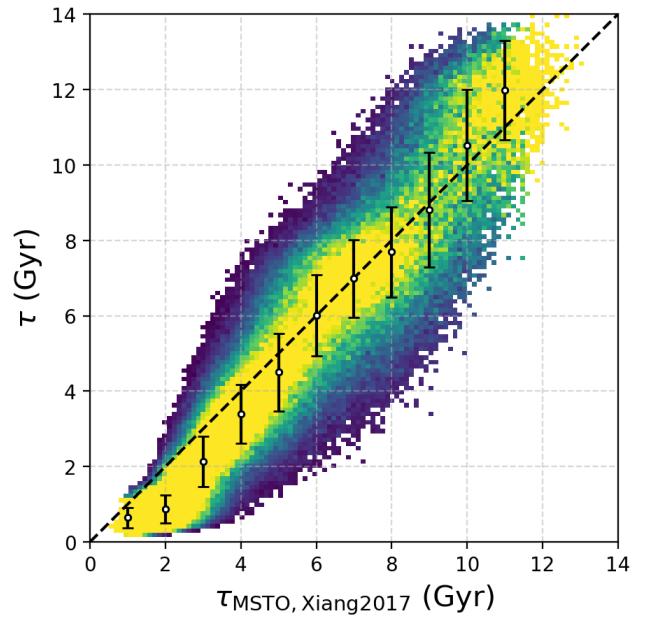
### 5.1. Precision of isochrone ages for K dwarfs

Stellar age can be estimated from their atmospheric parameters via fitting to stellar isochrones. For convenience, let's denote this isochrone age estimate as,

$$\tau_{\text{iso}} = f(T_{\text{eff}}, \log g, [\text{Fe}/\text{H}]). \quad (10)$$

Then one may have an approximate estimate of the uncertainty in the age estimate via,

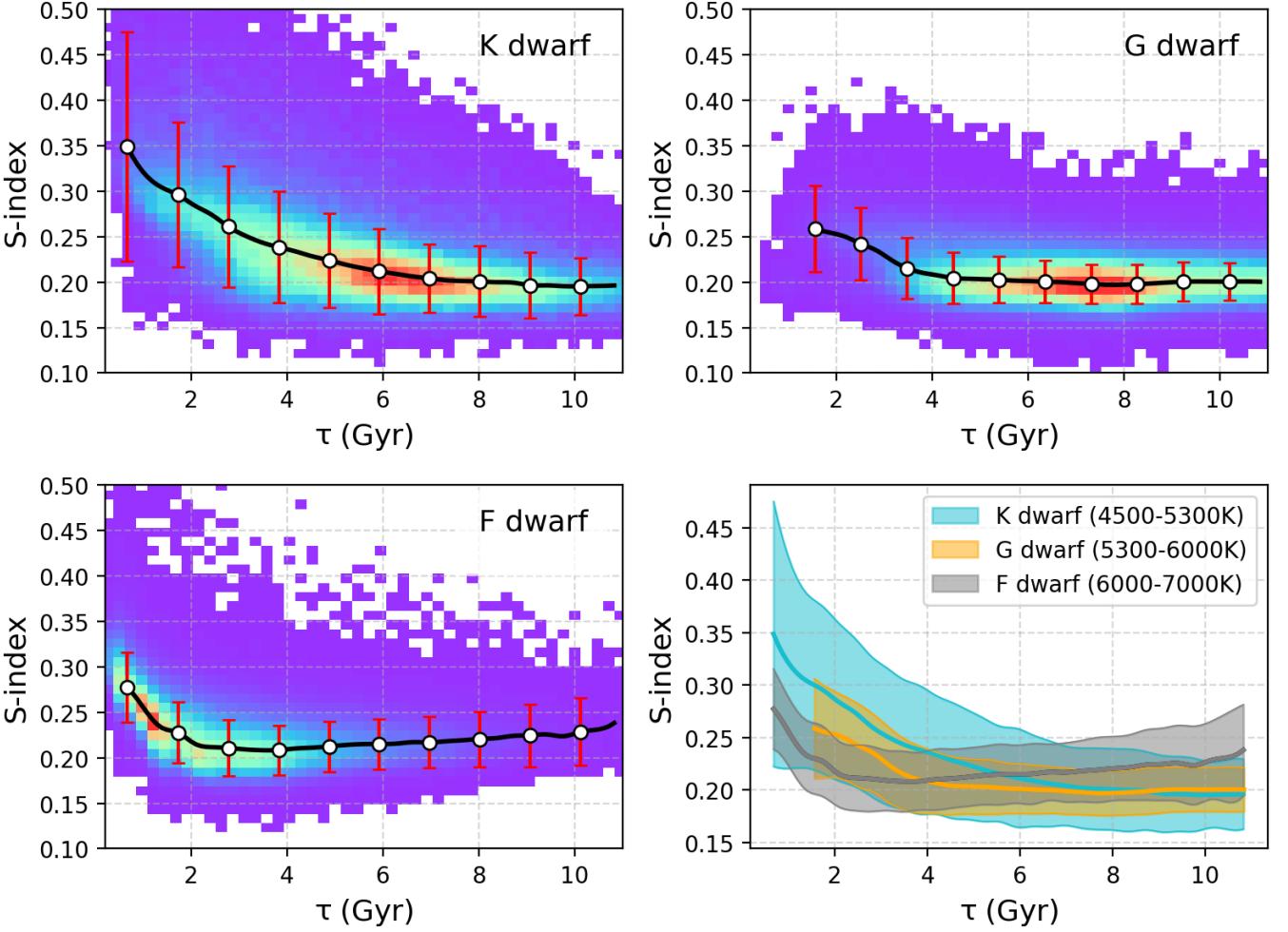
$$\delta_{\tau} = \sqrt{\sum_i \left( \frac{\partial \tau}{\partial X_i} \right)^2 \delta_{X_i}^2} \quad (11)$$



**Figure 13.** A comparison of our age estimates with the isochrone ages of Xiang et al. (2017). Color represents stellar density normalized to the peak value row-by-row. The white points correspond to the median ages from 1 Gyr to 11 Gyr, with the error bars representing the dispersion.

where  $X_i$  refers to  $T_{\text{eff}}$ ,  $\log g$ , and  $[\text{Fe}/\text{H}]$ ,  $\delta_{X_i}$  refers to uncertainties in these atmospheric parameters.

For  $K$  dwarf stars with  $T_{\text{eff}} \simeq 5000$  K, typical internal uncertainties in the LAMOST stellar parameters are  $\delta T_{\text{eff}} = 20$  K,  $\delta \log g = 0.05$ ,  $\delta [\text{Fe}/\text{H}] = 0.07$ , given a spectral  $S/N$  higher than 50. To have an estimate of  $\delta_{\tau}$ , we deduce a local derivative  $\frac{\partial \tau}{\partial X}$  in Eq.11 for each fiducial star  $X_i$  of concern using the MIST stellar isochrones. Figure 17 shows the resultant relative age error estimates for several fiducial  $K$  dwarfs of dif-



**Figure 14.** Stellar activity S-index as a function of age for stars. The upper-left, upper-right, and bottom-left panels show the results for K dwarfs, G dwarfs, and F dwarfs, respectively. The dots and error bars are the mean and standard deviations of S-index in different age bins. The bottom right panel compares the results for these F, G, and K dwarf stars.

ferent ages, selected from MIST isochrone grid points, at both  $[Fe/H] = 0$  and  $[Fe/H] = -0.5$ . The relative age error increases linearly with decreasing age in logarithmic space, from 50% for stars of 10 Gyr to  $\sim 1000\%$  for stars of 0.5 Gyr. This corresponds to a constant age error of 5 Gyr in absolute scale, owing to the nearly even distribution of stellar isochrones of different ages in the parameter space of concern, i.e.,  $\frac{\partial \tau}{\partial X}$  is nearly constant.

Such an isochrone age error is apparently much larger than the data-driven age error we obtained in the current work, as the latter has a typical value of  $\simeq 15\%$  for stars of 10 Gyr, and 30-50% for young stars of  $\tau \lesssim 2$  Gyr (Figure 17). On one hand, this clearly illustrates that isochrone ages for main-sequence *K* dwarfs suffer huge uncertainties. On the other hand, it implies that stellar atmospheric parameters only have contributed a small portion of the information that our data-driven approach used for age inference, while most of the informa-

tion for age inference must be from other stellar clocks, presumably chemical abundances (Sect. 5.2).

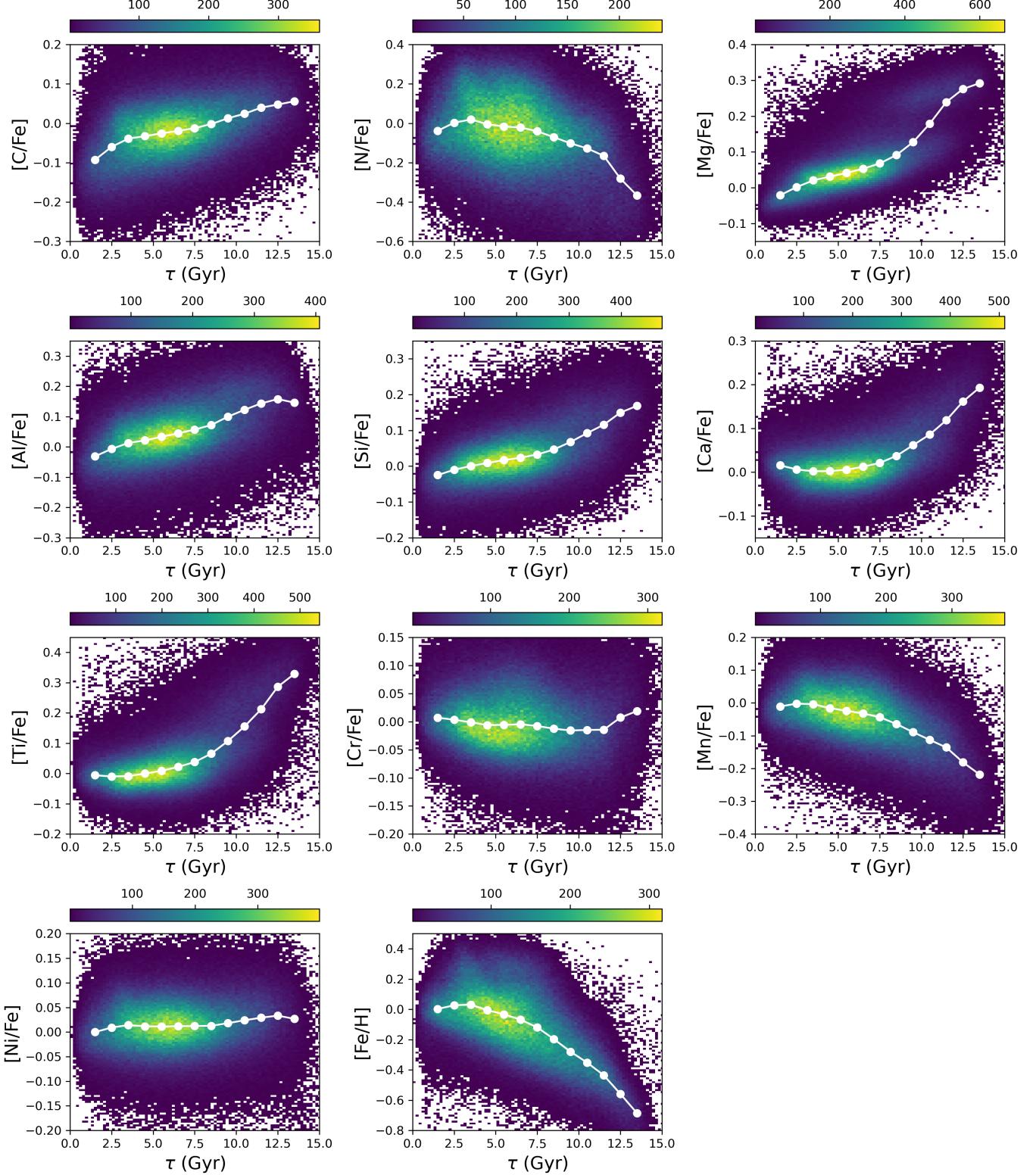
## 5.2. Age estimates from chemical clocks

Section 4.5 has shown that there are tight correlations between our age estimates and the chemical abundances. These correlations may have served as chemical clocks that provide major information for the data-driven age estimation.

To have a quantitative examination, here we derive age estimates directly from the chemical abundances by building an empirical chemical clock,

$$\tau_{[X/H]} = f(T_{\text{eff}}, \log g, [X/H]), \quad (12)$$

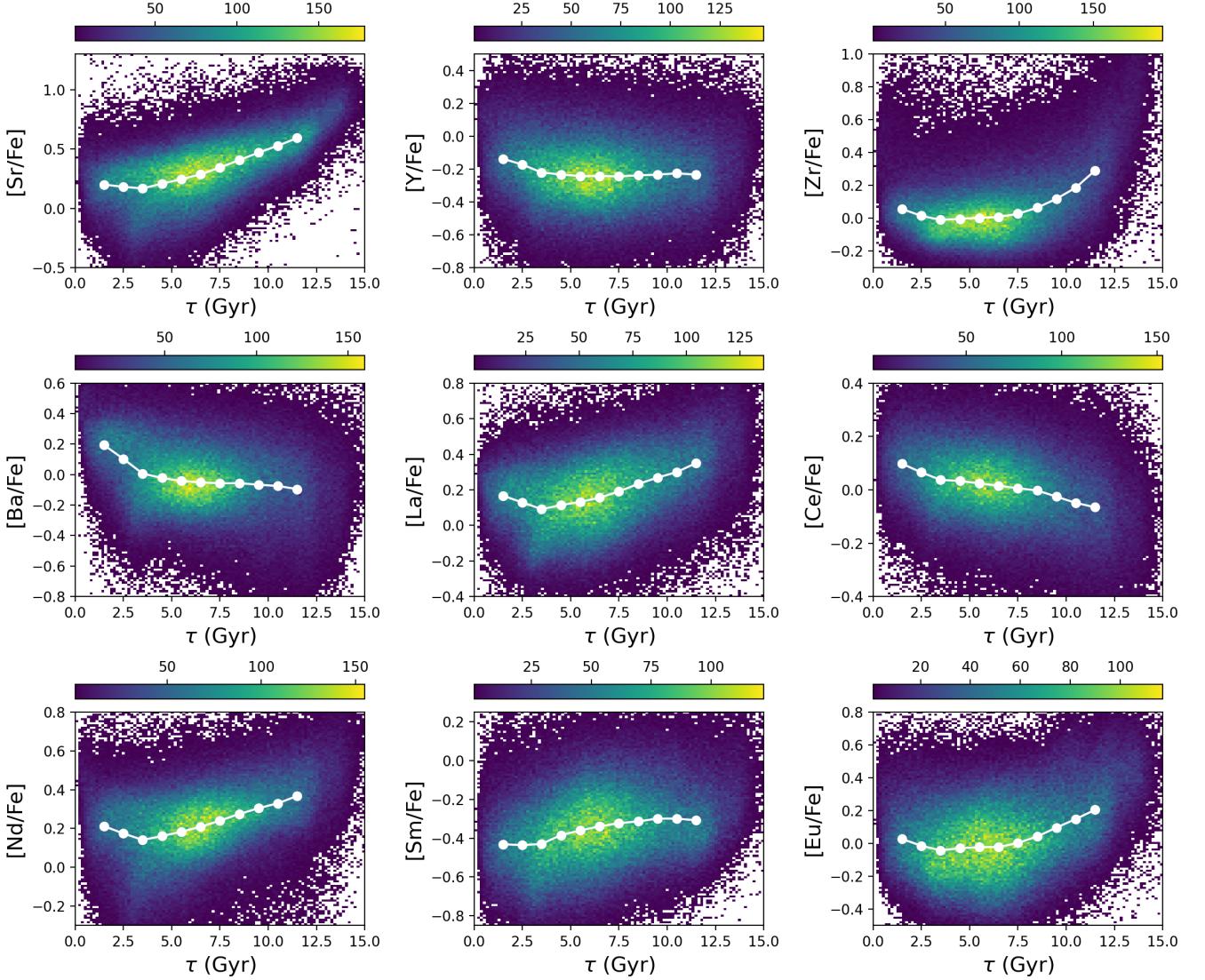
where  $[X/H]$  refers to 12 element abundances, namely,  $[Fe/H]$ ,  $[C/Fe]$ ,  $[N/Fe]$ ,  $[Mg/Fe]$ ,  $[Ca/Fe]$ ,  $[Si/Fe]$ ,  $[Ti/Fe]$ ,  $[Ba/Fe]$ ,  $[Cr/Fe]$ ,  $[Mn/Fe]$ ,  $[Ni/Fe]$  and  $[Al/Fe]$ , derived with DD-PAYNE from the LAMOST spectra.



**Figure 15.** Relations between age and elemental abundances for K dwarf stars. The abundances are derived from the LAMOST spectra with the DD-Payne (see text). The dots in white color are the median abundances in individual age bins.

This empirical relation is again constructed from our training sample with the XGBoost method. We then

apply this chemical clock to our test sample. As shown in Figure 18, the chemical age is in good consistency



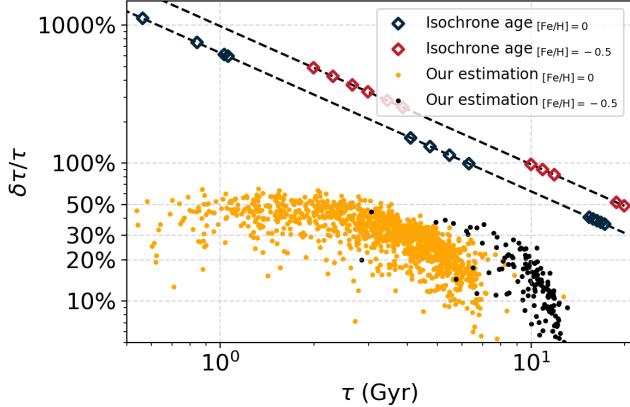
**Figure 16.** Same as Figure 15, but for abundances of heavier, s- and r-process elements.

with the reference age, and the overall dispersion in the relative age difference is 17%, almost identical to the data-driven age (18%, Figure 7). Inspired by this comparison, in an upcoming work we will provide another set of stellar ages for the LAMOST stars, specifically derived with the chemical clock, using the abundances for more than 20 elements, including the heavier, s- and r-process ones. These results suggest that our data-driven approach of age estimation is largely identical to a chemical clock, that is, the information from which the data-driven approach infers the age estimates is mostly from spectral features of chemical abundances. This is particularly true for cool main-sequence dwarf stars, however, for hotter main-sequence or subgiant stars, the isochrone age may have comparable or even better precision to the chemical age, stellar evolution thus serves as the main

clock for the current results. On the other hand, as discussed in Section 4.5, because our age estimates rely (implicitly) on chemical abundances, the age-abundance correlation from our sample is no longer an independent derivation. We emphasize that one needs to exercise great caution in any studies about the age-abundance relation using our age catalog.

### 5.3. Contribution from stellar activity

The LAMOST spectrum contains a few lines that are good indicators of stellar activity, including the Ca II H and K lines, and the  $H_{\alpha}$  line. To examine how our data-driven age estimation depends on these features, we have made a test for the age estimation by using the LAMOST spectra windows after masking these activity-sensitive lines. Indeed, after performing PCA, much of



**Figure 17.** Relative age error as a function of age for K dwarfs ( $T_{\text{eff}} = 5000$  K). The orange and black dots are age estimates for two different metallicities, as labeled in the Figure. As a comparison, the expected errors of isochrone age estimates for K dwarfs are shown in diamonds (see text). The dashed line is a linear fit to the diamonds.

the activity information in the spectra may have been removed. As shown in the right panel of Figure 18, the resultant age estimates are almost identical to those from the full spectra, meaning that these activity-sensitive lines contributed only a negligible portion of the information for the data-driven age estimation. The stellar age-activity relations presented in Sect. 4.4 are therefore reliable. At the same time, this also implies that the chemical information in the spectrum is enough to constrain age measurements, and the effect of chemical tagging is no less than that of activity information, even for young stars.

It has been shown that there are tight correlations among age, stellar activity, and stellar intrinsic photometric noise (Zhang et al. 2024; Ye et al. 2024). Our age catalog thus provides a piece of key information to select quiet stars with low-level intrinsic photometric noise for planet detection with transit observations, such as the coming Earth 2.0 mission (Ge & The ET team 2024).

#### 5.4. Limitation of the age estimates

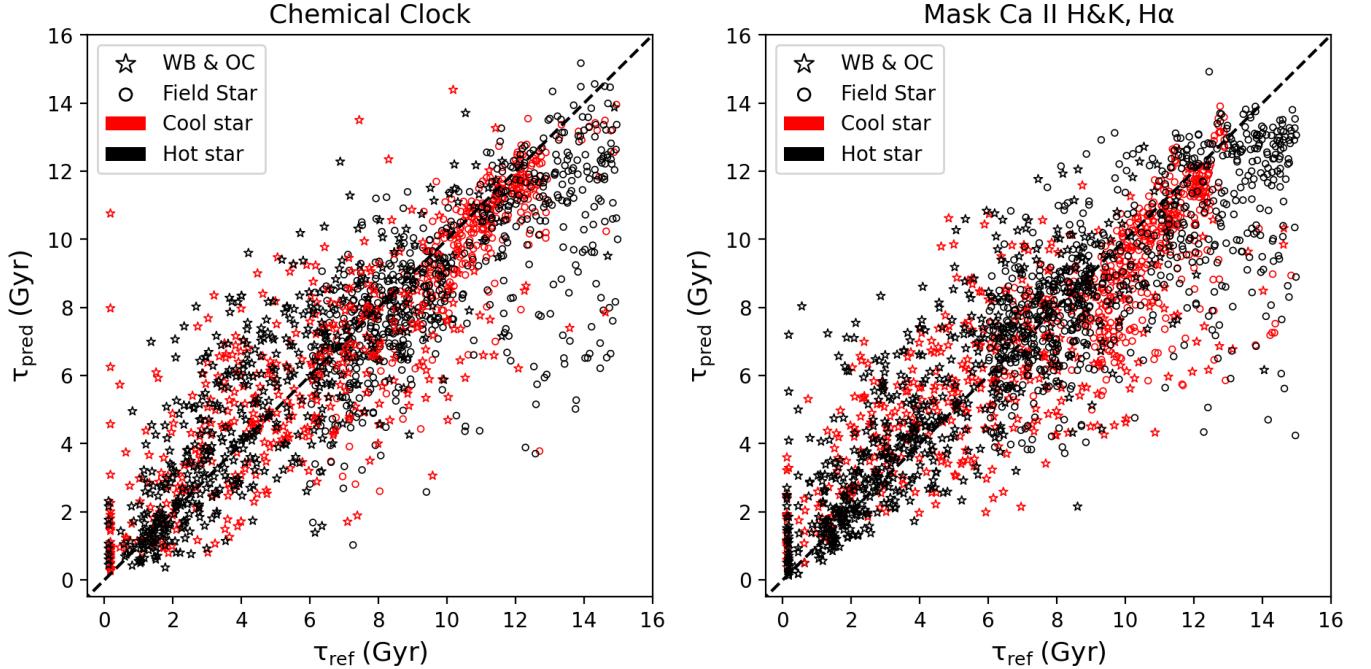
While our method has achieved robust age estimates for given feasible spectral  $S/N$ , one has to be in cautious about some caveats in the age estimates. First of all, as mentioned above (Sect. 4.2), while we can distinguish between young and old stars, there maybe a systematic error in absolute age estimation at the youngest end ( $<1$ Gyr). We believe that the systematic error maybe insignificant only for stars older than about 1 Gyr. In addition, the age needs to be used with caution in the following specific cases: i) the age estimates can be unreliable for binary and multiple star systems; ii) the age estimates can be incorrect for stars accreted from dwarf

galaxies, which have different chemical evolution histories compared to native stars. This is a consequence of the fact that our training sample is not constructed to contain a significant number of stars from dwarf galaxies; iii) our age estimates are assumed to be inapplicable for pre-main sequence stars, as these stars may have a different age-spectra relation compared to main-sequence stars.

## 6. CONCLUSION

It is challenging to determine the age of cool main-sequence dwarf stars accurately. In this work, we have employed a data-driven method to estimate age of dwarf stars from the LAMOST low-resolution spectra. We take 1213 wide binaries of main-sequence dwarf plus sub-giant systems, 927 warm field stars and 400 cool dwarfs in the old disk with known age estimates, and 317 cluster member stars as the training set. For wide binaries, the age of the cool main-sequence dwarf companion is known from its warmer primary, whose age is determined via isochrone fit. We adopt the XGBOOST algorithm to train the data-driven model for age estimation from the LAMOST spectra. Our main results are listed below:

- Given a LAMOST spectral S/N higher than 50, our method achieved a typical precision of 0.1 dex in the log Age estimates for K-type dwarf stars, corresponding to  $\sim 23\%$  relative age uncertainty in linear scale, while the uncertainty increases to 0.15 dex at  $S/N \simeq 20$ . Additionally, the relative age uncertainty also varies with age. For K-dwarfs, it is 50% at 1 Gyr and decreases to 15% at 10 Gyr.
- We found that for cool main-sequence stars ( $T_{\text{eff}} \lesssim 5400K$ ), chemical abundances in the spectra provided most of the information for the data-driven age inference, while other information such as stellar parameters ( $T_{\text{eff}}$ ,  $\log g$ ,  $[{\text{Fe}}/{\text{H}}]$ ) and stellar activity indicators (Ca II H and K lines,  $H_\alpha$  line) contribute little to the age inference. As a result, one needs to be very cautious in using our age estimates for these stars for studying the stellar age-abundance correlations.
- Our results reveal a clear age-stellar activity relation for relatively young stars ( $\tau \lesssim 5$  Gyr), with a detailed trend varying with spectral types, while older stars essentially show a flattened age-activity relation.
- We applied the method to the full dataset of LAMOST DR10, and obtained the age estimates for about 4 million stars. This age catalog is expected to be helpful for i) the study of star formation and



**Figure 18.** Same as Fig. 6, but for stellar age estimates with alternative methods. The left panel shows the age estimated from an empirical chemical clock built in this work. The right panel shows the data-driven spectroscopic age estimates after masking Ca II H, K lines, which may be affected by stellar activity (see discussions in Section 5).

assembly history of our Galaxy; ii) the study of temporal variations of the stellar initial mass function (IMF); iii) the selection of quiet stars with low intrinsic photometric noise for planet detection through transiting observations, such as that on the coming Earth 2.0 mission.

- The current sample of wide binary stars in our training set is limited. We expect that the ongoing LAMOST Phase-III survey will significantly increase the spectroscopic wide binary sample, and our method can therefore achieve a better age estimation in the next data release.

We acknowledge financial support from the NSFC through grant Nos. 2022000083 and the National Key R&D Program of China through grant No. 2022YFF0504200. J.-H.Z. acknowledges support from NSFC grant No. 12103063.

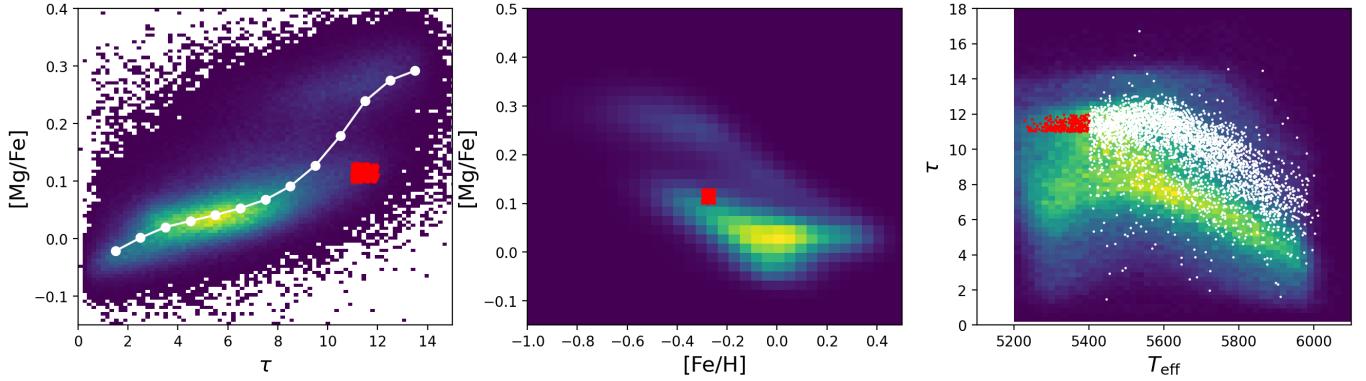
This work has made use of data products from the Guo Shou Jing Telescope (the Large Sky Area Multi-Object Fiber Spectroscopic Telescope; LAMOST). LAMOST is a National Major Scientific Project built by the Chinese Academy of Sciences. Funding for the project has been provided by the National Development and Reform Commission. LAMOST is operated and managed by the National Astronomical Observatories, Chinese Academy of Sciences.

## APPENDIX

### A. JUSTIFICATION FOR THE AGE ESTIMATION WITH OLD MG-POOR STARS

We have noticed that in the  $\tau$ -[Mg/Fe] relation plot, there are some old Mg-poor stars ( $\tau > 10$  Gyr). To validate their age estimation, we carry out a more detailed analysis on these stars. We choose low-alpha K dwarf stars located in a small box in the  $\tau$ -[Mg/Fe] and [Mg/Fe]-[Fe/H] planes, as highlighted in red dots in the left and middle panels of Figure 19 below, and plot their age as a function of temperature in the right panel. As a comparison set, we also show the result of stars with similar [Fe/H] and [Mg/Fe] but with higher  $T_{\text{eff}}$  ( $> 5400$  K). The ages of the comparison stars are plotted with white dots in the right panel of Figure 19. Their ages show a clear trend with temperature, spanning from 6 to 12 Gyr in the temperature range 5400-6000 K. Most stars with  $T_{\text{eff}} < 5600$  K have an age older than 10 Gyr. We expect that a substantial portion of the information for their age estimation is  $T_{\text{eff}}$ ,  $\log g$ , and [Fe/H], and their age estimates thus are robust. Otherwise, we would expect these stars to have a more uniform age distribution,

similar to the background rather than concentrate on 10–12 Gyr. The K dwarfs (red dots) are well on the  $\tau - T_{\text{eff}}$  trend, and we thus believe their age estimates are reasonable. In particular, we do not expect there to be an abrupt difference in the age distribution between  $T_{\text{eff}} < 5400$  K and  $T_{\text{eff}} > 5400$  K for stars in the same [Fe/H] and [Mg/Fe] bin.



**Figure 19.** Validation of the age estimates for old Mg-poor stars. The left panel shows the  $\tau$ –[Mg/Fe] relation for K dwarfs, from which we have selected old Mg-poor stars in the small box marked as red dots. The middle panel shows the box of [Fe/H] and [Mg/Fe] we have imposed in our selection. The right panel shows the  $\tau - T_{\text{eff}}$  relation for these old Mg-poor stars, while the white dots represent the comparison set, which is composed of stars in the same [Fe/H] and [Mg/Fe] box but with  $T_{\text{eff}} > 5400$  K. The background color-coded map indicates the full stellar sample in our catalog.

## REFERENCES

- Anders, F., Khalatyan, A., Queiroz, A. B. A., et al. 2022, *A&A*, 658, A91, doi: [10.1051/0004-6361/202142369](https://doi.org/10.1051/0004-6361/202142369)
- Angus, R., Aigrain, S., Foreman-Mackey, D., & McQuillan, A. 2015, *MNRAS*, 450, 1787, doi: [10.1093/mnras/stv423](https://doi.org/10.1093/mnras/stv423)
- Baliunas, S. L., Donahue, R. A., Soon, W. H., et al. 1995, *ApJ*, 438, 269, doi: [10.1086/175072](https://doi.org/10.1086/175072)
- Barnes, S. A. 2003, *ApJ*, 586, 464, doi: [10.1086/367639](https://doi.org/10.1086/367639)
- . 2007, *ApJ*, 669, 1167, doi: [10.1086/519295](https://doi.org/10.1086/519295)
- . 2010, *ApJ*, 722, 222, doi: [10.1088/0004-637X/722/1/222](https://doi.org/10.1088/0004-637X/722/1/222)
- Basri, G., Marcy, G. W., & Graham, J. R. 1996, *ApJ*, 458, 600, doi: [10.1086/176842](https://doi.org/10.1086/176842)
- Berger, T. A., van Saders, J. L., Huber, D., et al. 2022, *ApJ*, 936, 100, doi: [10.3847/1538-4357/ac8746](https://doi.org/10.3847/1538-4357/ac8746)
- Borucki, W. J., Koch, D., Basri, G., et al. 2010, in American Astronomical Society Meeting Abstracts, Vol. 215, American Astronomical Society Meeting Abstracts #215, 101.01
- Bouy, H., Bertin, E., Barrado, D., et al. 2015, *Astronomy & Astrophysics*, 575, A120, doi: [10.1051/0004-6361/201425505](https://doi.org/10.1051/0004-6361/201425505)
- Brown, W. R., Kilic, M., Kosakowski, A., & Gianninas, A. 2022, *ApJ*, 933, 94, doi: [10.3847/1538-4357/ac72ac](https://doi.org/10.3847/1538-4357/ac72ac)
- Brown, W. R., Lattanzi, M. G., Kenyon, S. J., & Geller, M. J. 2018, *ApJ*, 866, 39, doi: [10.3847/1538-4357/aadb8e](https://doi.org/10.3847/1538-4357/aadb8e)
- Buder, S., Sharma, S., Kos, J., et al. 2021, *MNRAS*, 506, 150, doi: [10.1093/mnras/stab1242](https://doi.org/10.1093/mnras/stab1242)
- Casamiquela, L., Reese, D. R., Lebreton, Y., et al. 2024, *A&A*, 692, A243, doi: [10.1051/0004-6361/202451677](https://doi.org/10.1051/0004-6361/202451677)
- Chen, B. Q., Liu, X. W., Yuan, H. B., et al. 2018, *MNRAS*, 476, 3278, doi: [10.1093/mnras/sty454](https://doi.org/10.1093/mnras/sty454)
- Chen, D.-C., Yang, J.-Y., Xie, J.-W., et al. 2021, *AJ*, 162, 100, doi: [10.3847/1538-3881/ac0f08](https://doi.org/10.3847/1538-3881/ac0f08)
- Chen, D.-C., Xie, J.-W., Zhou, J.-L., et al. 2023, Proceedings of the National Academy of Science, 120, e2304179120, doi: [10.1073/pnas.2304179120](https://doi.org/10.1073/pnas.2304179120)
- Chen, T., & Guestrin, C. 2016, in Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (ACM), doi: [10.1145/2939672.2939785](https://doi.org/10.1145/2939672.2939785)
- Choi, J., Dotter, A., Conroy, C., et al. 2016, *ApJ*, 823, 102, doi: [10.3847/0004-637X/823/2/102](https://doi.org/10.3847/0004-637X/823/2/102)
- Cui, X.-Q., Zhao, Y.-H., Chu, Y.-Q., et al. 2012, *Research in Astronomy and Astrophysics*, 12, 1197, doi: [10.1088/1674-4527/12/9/003](https://doi.org/10.1088/1674-4527/12/9/003)
- Delgado Mena, E., Moya, A., Adibekyan, V., et al. 2019, *A&A*, 624, A78, doi: [10.1051/0004-6361/201834783](https://doi.org/10.1051/0004-6361/201834783)
- Demarque, P., Woo, J.-H., Kim, Y.-C., & Yi, S. K. 2004, *ApJS*, 155, 667, doi: [10.1086/424966](https://doi.org/10.1086/424966)

- Deng, L.-C., Newberg, H. J., Liu, C., et al. 2012, Research in Astronomy and Astrophysics, 12, 735, doi: [10.1088/1674-4527/12/7/003](https://doi.org/10.1088/1674-4527/12/7/003)
- El-Badry, K., & Rix, H.-W. 2018, MNRAS, 480, 4884, doi: [10.1093/mnras/sty2186](https://doi.org/10.1093/mnras/sty2186)
- El-Badry, K., Rix, H.-W., & Heintz, T. M. 2021, MNRAS, 506, 2269, doi: [10.1093/mnras/stab323](https://doi.org/10.1093/mnras/stab323)
- Feuillet, D. K., Frankel, N., Lind, K., et al. 2019, MNRAS, 489, 1742, doi: [10.1093/mnras/stz2221](https://doi.org/10.1093/mnras/stz2221)
- Gaia Collaboration, Prusti, T., de Bruijne, J. H. J., et al. 2016, A&A, 595, A1, doi: [10.1051/0004-6361/201629272](https://doi.org/10.1051/0004-6361/201629272)
- Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2021, A&A, 649, A1, doi: [10.1051/0004-6361/202039657](https://doi.org/10.1051/0004-6361/202039657)
- Ge, J., & The ET team. 2024, in American Astronomical Society Meeting Abstracts, Vol. 243, American Astronomical Society Meeting Abstracts, 445.01
- Ge, J., Zhang, H., Zang, W., et al. 2022, ET White Paper: To Find the First Earth 2.0. <https://arxiv.org/abs/2206.06693>
- Han, H., Wang, S., Li, X., Zheng, C., & Liu, J. 2024, ApJ, 977, 138, doi: [10.3847/1538-4357/ad957a](https://doi.org/10.3847/1538-4357/ad957a)
- Hartman, J. D., Gaudi, B. S., Pinsonneault, M. H., et al. 2009, The Astrophysical Journal, 691, 342, doi: [10.1088/0004-637x/691/1/342](https://doi.org/10.1088/0004-637x/691/1/342)
- Hayden, M. R., Sharma, S., Bland-Hawthorn, J., et al. 2022, MNRAS, 517, 5325, doi: [10.1093/mnras/stac2787](https://doi.org/10.1093/mnras/stac2787)
- Haywood, M., Matteo, P. D., Lehnert, M. D., Katz, D., & Gómez, A. 2013, Astronomy & Astrophysics, 560, A109, doi: [10.1051/0004-6361/201321397](https://doi.org/10.1051/0004-6361/201321397)
- Ho, A. Y. Q., Rix, H.-W., Ness, M. K., et al. 2017, ApJ, 841, 40, doi: [10.3847/1538-4357/aa6db3](https://doi.org/10.3847/1538-4357/aa6db3)
- Huang, Y., Schönrich, R., Zhang, H., et al. 2020, ApJS, 249, 29, doi: [10.3847/1538-4365/ab994f](https://doi.org/10.3847/1538-4365/ab994f)
- Kim, Y. C., Yi, S. K., & Demarque, P. 2001, in American Astronomical Society Meeting Abstracts, Vol. 198, American Astronomical Society Meeting Abstracts #198, 43.01
- Li, T., Li, Y., Bi, S., et al. 2022, ApJ, 927, 167, doi: [10.3847/1538-4357/ac4fbf](https://doi.org/10.3847/1538-4357/ac4fbf)
- Liu, C., Deng, L.-C., Carlin, J. L., et al. 2014a, ApJ, 790, 110, doi: [10.1088/0004-637X/790/2/110](https://doi.org/10.1088/0004-637X/790/2/110)
- Liu, C., Xu, Y., Wan, J.-C., et al. 2017, Research in Astronomy and Astrophysics, 17, 096, doi: [10.1088/1674-4527/17/9/96](https://doi.org/10.1088/1674-4527/17/9/96)
- Liu, X. W., Yuan, H. B., Huo, Z. Y., et al. 2014b, in Setting the scene for Gaia and LAMOST, ed. S. Feltzing, G. Zhao, N. A. Walton, & P. Whitelock, Vol. 298, 310–321, doi: [10.1017/S1743921313006510](https://doi.org/10.1017/S1743921313006510)
- Lu, Y., Minchev, I., Buck, T., et al. 2022, arXiv e-prints, arXiv:2212.04515, doi: [10.48550/arXiv.2212.04515](https://doi.org/10.48550/arXiv.2212.04515)
- Mamajek, E. E., & Hillenbrand, L. A. 2008, ApJ, 687, 1264, doi: [10.1086/591785](https://doi.org/10.1086/591785)
- Martig, M., Rix, H.-W., Silva Aguirre, V., et al. 2015, MNRAS, 451, 2230, doi: [10.1093/mnras/stv1071](https://doi.org/10.1093/mnras/stv1071)
- Martig, M., Fouesneau, M., Rix, H.-W., et al. 2016, MNRAS, 456, 3655, doi: [10.1093/mnras/stv2830](https://doi.org/10.1093/mnras/stv2830)
- Meibom, S., Barnes, S. A., Platais, I., et al. 2015, Nature, 517, 589, doi: [10.1038/nature14118](https://doi.org/10.1038/nature14118)
- Miglio, A., Chiappini, C., Mackereth, J. T., et al. 2021, A&A, 645, A85, doi: [10.1051/0004-6361/202038307](https://doi.org/10.1051/0004-6361/202038307)
- Mints, A., Hekker, S., & Minchev, I. 2019, A&A, 629, A127, doi: [10.1051/0004-6361/201935864](https://doi.org/10.1051/0004-6361/201935864)
- Nataf, D. M., Schlaufman, K. C., Reggiani, H., & Hahn, I. 2024a, arXiv e-prints, arXiv:2407.18307, doi: [10.48550/arXiv.2407.18307](https://doi.org/10.48550/arXiv.2407.18307)
- . 2024b, ApJ, 976, 87, doi: [10.3847/1538-4357/ad7c4e](https://doi.org/10.3847/1538-4357/ad7c4e)
- Ness, M., Hogg, D. W., Rix, H. W., et al. 2016, ApJ, 823, 114, doi: [10.3847/0004-637X/823/2/114](https://doi.org/10.3847/0004-637X/823/2/114)
- Niu, Z., Yuan, H., & Liu, J. 2023, ApJ, 950, 104, doi: [10.3847/1538-4357/accf8b](https://doi.org/10.3847/1538-4357/accf8b)
- Pace, G., & Pasquini, L. 2004, Astronomy & Astrophysics, 426, 1021, doi: [10.1051/0004-6361:20040568](https://doi.org/10.1051/0004-6361:20040568)
- Perryman, M. A. C., Brown, A. G. A., Lebreton, Y., et al. 1998, A&A, 331, 81, doi: [10.48550/arXiv.astro-ph/9707253](https://doi.org/10.48550/arXiv.astro-ph/9707253)
- Ratcliffe, B., Minchev, I., Cesutti, G., et al. 2024, MNRAS, 528, 3464, doi: [10.1093/mnras/stae226](https://doi.org/10.1093/mnras/stae226)
- Richer, H. B., Fahlman, G. G., Rosvick, J., & Ibata, R. 1998, ApJL, 504, L91, doi: [10.1086/311586](https://doi.org/10.1086/311586)
- Riello, M., De Angeli, F., Evans, D. W., et al. 2021, A&A, 649, A3, doi: [10.1051/0004-6361/202039587](https://doi.org/10.1051/0004-6361/202039587)
- Salaris, M., Chieffi, A., & Straniero, O. 1993, ApJ, 414, 580, doi: [10.1086/173105](https://doi.org/10.1086/173105)
- Sanders, J. L., & Das, P. 2018, MNRAS, 481, 4093, doi: [10.1093/mnras/sty2490](https://doi.org/10.1093/mnras/sty2490)
- Serenelli, A., Johnson, J., Huber, D., et al. 2017, ApJS, 233, 23, doi: [10.3847/1538-4365/aa97df](https://doi.org/10.3847/1538-4365/aa97df)
- Sestito, P., & Randich, S. 2005, A&A, 442, 615, doi: [10.1051/0004-6361:20053482](https://doi.org/10.1051/0004-6361:20053482)
- Sharma, S., Hayden, M. R., Bland-Hawthorn, J., et al. 2022, MNRAS, 510, 734, doi: [10.1093/mnras/stab3341](https://doi.org/10.1093/mnras/stab3341)
- Silva Aguirre, V., Lund, M. N., Antia, H. M., et al. 2017, ApJ, 835, 173, doi: [10.3847/1538-4357/835/2/173](https://doi.org/10.3847/1538-4357/835/2/173)
- Silva Aguirre, V., Bojsen-Hansen, M., Slumstrup, D., et al. 2018, MNRAS, 475, 5487, doi: [10.1093/mnras/sty150](https://doi.org/10.1093/mnras/sty150)
- Skrutskie, M. F., Cutri, R. M., Stiening, R., et al. 2006, AJ, 131, 1163, doi: [10.1086/498708](https://doi.org/10.1086/498708)
- Skúladóttir, Á., Hansen, C. J., Salvadori, S., & Choplin, A. 2019, A&A, 631, A171, doi: [10.1051/0004-6361/201936125](https://doi.org/10.1051/0004-6361/201936125)

- Skumanich, A. 1972, ApJ, 171, 565, doi: [10.1086/151310](https://doi.org/10.1086/151310)
- Smiljanic, R., Romano, D., Bragaglia, A., et al. 2016, A&A, 589, A115, doi: [10.1051/0004-6361/201528014](https://doi.org/10.1051/0004-6361/201528014)
- Soderblom, D. R. 2010, Annual Review of Astronomy and Astrophysics, 48, 581, doi: [10.1146/annurev-astro-081309-130806](https://doi.org/10.1146/annurev-astro-081309-130806)
- Sun, T., Chen, X., Bi, S., et al. 2023, MNRAS, 523, 1199, doi: [10.1093/mnras/stad1499](https://doi.org/10.1093/mnras/stad1499)
- Ting, Y.-S. 2024, arXiv e-prints, arXiv:2412.05806, doi: [10.48550/arXiv.2412.05806](https://doi.org/10.48550/arXiv.2412.05806)
- Tripathi, B., Nandy, D., & Banerjee, S. 2021, MNRAS, 506, L50, doi: [10.1093/mnrasl/slab035](https://doi.org/10.1093/mnrasl/slab035)
- van Saders, J. L., Ceillier, T., Metcalfe, T. S., et al. 2016, Nature, 529, 181, doi: [10.1038/nature16168](https://doi.org/10.1038/nature16168)
- van Saders, J. L., Pinsonneault, M. H., García, R. A., & Ceillier, T. 2015, in European Physical Journal Web of Conferences, Vol. 101, European Physical Journal Web of Conferences, 05006, doi: [10.1051/epjconf/201510105006](https://doi.org/10.1051/epjconf/201510105006)
- Vickers, J. J., Shen, J., & Li, Z.-Y. 2021, ApJ, 922, 189, doi: [10.3847/1538-4357/ac27a9](https://doi.org/10.3847/1538-4357/ac27a9)
- von Hippel, T., & Gilmore, G. 2000, AJ, 120, 1384, doi: [10.1086/301528](https://doi.org/10.1086/301528)
- Wang, J., Cao, Z., Huang, Y., & Yuan, H. 2023, Research in Astronomy and Astrophysics, 23, 025020, doi: [10.1088/1674-4527/acaef6](https://doi.org/10.1088/1674-4527/acaef6)
- Wu, Y., Xiang, M., Chen, Y., et al. 2021, MNRAS, 501, 4917, doi: [10.1093/mnras/staa3949](https://doi.org/10.1093/mnras/staa3949)
- Wu, Y., Xiang, M., Zhao, G., et al. 2023, MNRAS, 520, 1913, doi: [10.1093/mnras/stad223](https://doi.org/10.1093/mnras/stad223)
- Wu, Y., Xiang, M., Bi, S., et al. 2018, MNRAS, 475, 3633, doi: [10.1093/mnras/stx3296](https://doi.org/10.1093/mnras/stx3296)
- Wu, Y., Xiang, M., Zhao, G., et al. 2019, MNRAS, 484, 5315, doi: [10.1093/mnras/stz256](https://doi.org/10.1093/mnras/stz256)
- Wu, Y.-Q., Xiang, M.-S., Zhang, X.-F., et al. 2017, Research in Astronomy and Astrophysics, 17, 5, doi: [10.1088/1674-4527/17/1/5](https://doi.org/10.1088/1674-4527/17/1/5)
- Wu, Z.-Y., Zhou, X., Ma, J., & Du, C.-H. 2009, Monthly Notices of the Royal Astronomical Society, 399, 2146, doi: [10.1111/j.1365-2966.2009.15416.x](https://doi.org/10.1111/j.1365-2966.2009.15416.x)
- Xiang, M., & Rix, H.-W. 2022, Nature, 603, 599, doi: [10.1038/s41586-022-04496-5](https://doi.org/10.1038/s41586-022-04496-5)
- Xiang, M., Rix, H.-W., Ting, Y.-S., et al. 2021, ApJS, 253, 22, doi: [10.3847/1538-4365/abd6ba](https://doi.org/10.3847/1538-4365/abd6ba)
- Xiang, M., Liu, X., Shi, J., et al. 2017, ApJS, 232, 2, doi: [10.3847/1538-4365/aa80e4](https://doi.org/10.3847/1538-4365/aa80e4)
- Xiang, M., Liu, X., Shi, J., et al. 2017, The Astrophysical Journal Supplement Series, 232, 2, doi: [10.3847/1538-4365/aa80e4](https://doi.org/10.3847/1538-4365/aa80e4)
- Xiang, M., Ting, Y.-S., Rix, H.-W., et al. 2019, ApJS, 245, 34, doi: [10.3847/1538-4365/ab5364](https://doi.org/10.3847/1538-4365/ab5364)
- Xiang, M.-S., Liu, X.-W., Yuan, H.-B., et al. 2015, Research in Astronomy and Astrophysics, 15, 1209, doi: [10.1088/1674-4527/15/8/009](https://doi.org/10.1088/1674-4527/15/8/009)
- Yan, H., Li, H., Wang, S., et al. 2022, The Innovation, 3, 100224, doi: [10.1016/j.xinn.2022.100224](https://doi.org/10.1016/j.xinn.2022.100224)
- Ye, L., Bi, S., Zhang, J., et al. 2024, ApJS, 271, 19, doi: [10.3847/1538-4365/ad1eee](https://doi.org/10.3847/1538-4365/ad1eee)
- Yuan, H., Liu, X., Xiang, M., et al. 2015, ApJ, 799, 135, doi: [10.1088/0004-637X/799/2/135](https://doi.org/10.1088/0004-637X/799/2/135)
- Zhang, J., Xiang, M., Yu, J., et al. 2024, ApJS, 272, 40, doi: [10.3847/1538-4365/ad41b6](https://doi.org/10.3847/1538-4365/ad41b6)
- Zhang, M., Xiang, M., Zhang, H.-W., et al. 2021, ApJ, 922, 145, doi: [10.3847/1538-4357/ac22a5](https://doi.org/10.3847/1538-4357/ac22a5)
- Zhao, G., Zhao, Y.-H., Chu, Y.-Q., Jing, Y.-P., & Deng, L.-C. 2012, Research in Astronomy and Astrophysics, 12, 723, doi: [10.1088/1674-4527/12/7/002](https://doi.org/10.1088/1674-4527/12/7/002)
- Zhao, J. K., Oswalt, T. D., Rudkin, M., Zhao, G., & Chen, Y. Q. 2011, The Astronomical Journal, 141, 107, doi: [10.1088/0004-6256/141/4/107](https://doi.org/10.1088/0004-6256/141/4/107)