Photometric Metallicities of 0.8 Million KiDS Stars

Bao-Kun Sun, Bing-Qiu Chen, and Xiao-Wei Liu
South-Western Institute for Astronomy Research, Yunnan University, Kunming 650091, China; bchen@ynu.edu.cn, x.liu@ynu.edu.cn

Abstract

Accurate determinations of metallicity for large, complete stellar samples are essential for advancing various studies of the Milky Way. In this paper, we present a data-driven algorithm that leverages photometric data from the KiDS and the VIKING surveys to estimate stellar absolute magnitudes, effective temperatures, and metallicities. The algorithm is trained and validated using spectroscopic data from LAMOST, SEGUE, APOGEE, and GALAH, as well as a catalog of very metal-poor stars from the literature, and Gaia EDR3 data. This approach enables us to estimate metallicities, effective temperatures, and *g*-band absolute magnitudes for approximately 0.8 million stars in the KiDS data set. The photometric metallicity estimates exhibit an uncertainty of around 0.28 dex when compared to spectroscopic studies, within the metallicity range of -2 dex to 0.5 dex. The photometric effective temperature estimates have an uncertainty of around 149 K, while the uncertainty in the absolute magnitudes is approximately 0.36 mag. The metallicity estimates are reliable for values down to about -2 dex. This catalog represents a valuable resource for studying the structure and chemical properties of the Milky Way, offering an extensive data set for future investigations into Galactic formation and evolution.

Key words: stars: abundances – stars: distances – Galaxy: abundances

1. Introduction

Understanding the chemical composition of the Milky Way is key to unraveling its structure, formation, and evolutionary processes (e.g., Casagrande et al. 2011; Peng et al. 2013; Yan et al. 2019; Whitten et al. 2021; Martin et al. 2023; Sun et al. 2023; Hackshaw et al. 2024).

The field has been transformed by large-scale spectroscopic surveys, such as the Radial Velocity Experiment (RAVE; Steinmetz et al. 2006), Sloan Extension for Galactic Understanding and Exploration (SEGUE; Yanny et al. 2009), Large Sky Area Multi-Object Fiber Spectroscopic Telescope (LAMOST; Zhao et al. 2012; Liu et al. 2015), Apache Point Observatory Galactic Evolution Experiment (APOGEE; Majewski et al. 2017), and the Galactic Archaeology with HERMES project (GALAH; De Silva et al. 2015). These surveys have collected spectra for more than ten million stars (e.g., Luo et al. 2015; Yuan et al. 2015a), providing critical stellar parameters such as effective temperature (T_{eff}), metallicity ([Fe/H]), surface gravity ($\log g$), and other physical properties. This wealth of data has enabled significant advances in our understanding of the Milky Way's structure, chemical composition, and kinematics.

For example, using metallicity data from LAMOST, Wang et al. (2019) found that the vertical metallicity gradient is steeper in younger stars, reaches a maximum in stars of intermediate age, and becomes flatter in older populations. Rojas-Arriagada et al. (2020), employing APOGEE data, identified a trimodal metallicity distribution in the Galactic

bulge, suggesting distinct spatial distributions and kinematic structures among metal-rich, metal-intermediate, and metal-poor populations, providing important clues about their formation histories. Additionally, Guiglion et al. (2024) used a hybrid convolutional neural network (CNN) approach to reanalyze the Radial Velocity Spectrometer (RVS) sample and characterized the [Fe/H]-[α/Fe] bimodality in the Galactic disk. Further insights come from Zhang et al. (2024), who examined a large sample of giant stars using Gaia DR3 XP spectra. They found that a clear disk population emerges at [M/H] ~ -1.3 dex, with no such population present at [M/H] < -1.6 dex, suggesting two distinct halo populations in the very metal-poor regime. Similarly, Lian et al. (2023) analyzed APOGEE metallicities and discovered that the radial [Fe/H] gradients of older stellar populations are flatter than those of younger stars. Finally, Hattori (2025) estimated [M/H] and [α/M] for 48 million giant and dwarf stars in low-extinction regions using Gaia DR3 XP spectra, finding that high-[α/M] and low-[α/M] stars exhibit distinct kinematic properties, particularly among giants and low-temperature dwarfs.

Spectroscopic methods, while valuable, are resource-intensive in terms of telescope time and are often constrained by complex selection biases (e.g., Nidever et al. 2014; Wojno et al. 2017; Chen et al. 2018). Additionally, the faintness of the target stars limits the scope of exploration, particularly in the more distant regions of the Milky Way, such as the halo. In contrast, large-scale, high-precision, multi-band photometric

surveys—such as the SkyMapper Southern Survey (SMSS; Wolf et al. 2018), Javalambre/Southern Photometric Local Universe Survey (J/S-PLUS; Cenarro et al. 2019; Mendes de Oliveira et al. 2023), and Gaia (Gaia Collaboration et al. 2016, 2018, 2021)—have proven capable of reliably deriving stellar atmospheric parameters from photometry across a broad, unbiased sample of stars (e.g., Ivezić et al. 2008; Yuan et al. 2015b; Huang et al. 2022; Xu et al. 2022; Sun et al. 2023; Lu et al. 2024; Xiao et al. 2024).

Photometric methods have shown great promise in mapping the chemical properties of the Milky Way. Ivezić et al. (2008) demonstrated the ability to distinguish between the thin disk, thick disk, and halo components using photometrically derived metallicities from SDSS data. Building upon this, An & Beers (2020, 2021a, 2021b) advanced the photometric mapping of the Milky Way by integrating spatial and chemical information with Gaia proper motions. Further, Fernández-Alvar et al. (2021) identified a stellar population exhibiting the rotational velocities characteristic of the thin disk, even at metallicities as low as $[Fe/H] \sim -2$ dex, using stars from the Pristine survey (Starkenburg et al. 2017) in the direction of the Milky Way's anticenter. Huang et al. (2019, 2022, 2023) applied the stellar loci fitting technique to the photometric data from SMSS DR2 and the Stellar Abundances and Galactic Evolution Survey (SAGES; (Fan et al. 2023)) DR1. Thanks to the well-optimized narrow/medium-band u, v filters, they were able to estimate photometric metallicities down to $[Fe/H] \sim -3.5$ dex for nearly 50 million stars. Similarly, Youakim et al. (2020) mapped the metallicity distribution within the Milky Way's inner halo using Pristine Survey data, revealing a higher fraction of metal-poor stars than previously reported and estimating the likely number of metal-poor globular clusters in the halo. Chiti et al. (2021a, 2021b) explored the Milky Way's Metallicity Distribution Function (MDF), deriving metallicities for approximately 28,000 stars down to $[Fe/H] \sim -3.75$ dex using SMSS data. Additionally, Whitten et al. (2021) used S-PLUS photometry to estimate effective temperatures, metallicities, carbon abundances, and carbonicity for over 700,000 stars. These data enabled them to characterize the K-dwarf halo MDF and the frequency of Carbon-Enhanced Metal-Poor (CEMP) stars in the Milky Way. Expanding on this, Kim & Lépine (2022) compiled a catalog of 551,214 main-sequence stars in the local thick disk and halo, calibrated their photometric metallicities, and analyzed their kinematics, identifying structures consistent with known halo components. Finally, Martin et al. (2023) utilized Gaia BP/RP spectra to derive synthetic photometry of the Ca II H&K region, based on narrow-band photometry from the Pristine Survey, to illustrate how the Milky Way's structure varies with metallicity.

In this study, we leverage deep multi-band photometric data from the Kilo-Degree Survey (KiDS; Kuijken et al. 2019) and the VISTA Kilo-degree Infrared Galaxy (VIKING; Edge et al. 2013) survey to estimate photometric metallicities for

approximately 820,055 stars in the Milky Way. This extensive catalog allows us to explore the Milky Way's structure, formation, and evolutionary processes.

2. Data

In this study, we aim to derive photometric metallicities using multi-band photometry, trained on a foundation of spectroscopic data.

2.1. Photometric Data

The photometric data are sourced from the KiDS and VIKING surveys, which cover an area of 1350 square degrees, providing observations in both optical and near-infrared (NIR) wavelengths. The KiDS survey, conducted with the VST/OmegaCAM telescope (Capaccioli & Schipani 2011; Kuijken 2011), utilizes four optical filters (u, g, r, i) with a resolution of $0''.2$ pixel^{-1}. The r-band images are taken under optimal seeing conditions, with an average full width at half maximum (FWHM) of approximately $0''.7$. These images reach a mean limiting AB magnitude (5σ in a $2''$ aperture) of 25.02 ± 0.13 mag. The remaining bands (u, g, i) have slightly lower resolution (FWHM $< 1''.1$) and fainter limiting magnitudes of 24.23 ± 0.12 mag, 25.12 ± 0.14 mag, and 23.68 ± 0.27 mag, respectively (Kuijken et al. 2019).

The VIKING survey, using the VISTA/VIRCAM instrument (Sutherland et al. 2015), complements KiDS with five NIR bands (Z, Y, J, H, K_S). The median seeing for these images is approximately $0''.9$, with depths ranging from 21.2 mag to 23.1 mag across the passbands (Edge et al. 2013). Together, the KiDS DR4 and VIKING DR4 releases contain photometric data for over 100 million unique sources. The deep limiting magnitudes of these data sets provide an excellent sample for investigating the Milky Way's chemical properties.

For stellar distance measurements, we used Gaia EDR3 photometric and astrometric data. Rather than directly inverting Gaia parallax measurements to estimate distances, we employed the distance catalog from Bailer-Jones et al. (2021), which provides distance estimates for nearly 1.47 billion stars. This catalog is based on Gaia parallaxes and includes a weak prior that accounts for the spatial distribution of stars in the Milky Way.

Since the KiDS and VIKING surveys primarily target regions at high Galactic latitudes ($|b| > 22°$), the observed stars are minimally affected by extinction. The magnitudes and colors presented in this study have been corrected for reddening using $E(B - V)$ values from Schlegel et al. (1998), with extinction coefficients for the KiDS and VIKING passbands adopted from Kuijken et al. (2019).

2.2. Spectroscopic Data

To derive photometric T_{eff} and metallicities from the multi-band photometry of the KiDS and VIKING surveys, we rely on

training samples with robust spectroscopic metallicity measurements. Our primary spectroscopic data set is drawn from the LAMOST Galactic survey. The LAMOST telescope, a quasi-meridian reflecting Schmidt instrument, is capable of simultaneously obtaining spectra for up to 4000 celestial objects within a 5° diameter field of view. The spectra cover a wavelength range from 3700 to 9100 Å, with a resolution of $R \approx 1800$. The LAMOST DR9 release includes over 11 million spectra, comprising 242,569 galaxies, 10,907,516 stars, and 76,167 quasars. Stellar atmospheric parameters, such as metallicity ([Fe/H]), are determined using the LAMOST Stellar Parameter Pipeline (LASP; Wu et al. 2014), providing reliable metallicity estimates for stars with [Fe/H] values above -2.5 dex.

To ensure broader coverage across a range of stellar types and effective temperatures (3300 K $\leqslant T_{\rm eff} \leqslant$ 8800 K), we also incorporate data from other spectroscopic surveys, including APOGEE, SEGUE, and GALAH DR2. However, combining multiple spectroscopic data sets introduces challenges due to systematic differences in stellar parameter estimations, which can arise from variations in instrumentation and data processing methods across the surveys. To mitigate these discrepancies, we cross-match stars from LAMOST, GALAH, and SEGUE with those in APOGEE. This allows us to apply zero-point corrections for metallicity and effective temperature, harmonizing the metallicity scales between LAMOST and the other data sets.

For the characterization of extremely metal-poor stars—a regime where many spectroscopic surveys face limitations—we augment our data set with stars from specialized studies targeting this population (Jacobson et al. 2015; Li et al. 2018; Da Costa et al. 2019; Aguado et al. 2019). This sample includes 10,224 stars with metallicities predominantly in the range -3.5 to -2 dex, all with [Fe/H] < -2 dex. These extremely metal-poor stars enhance the precision of our metallicity estimates at the lowest end of the metallicity distribution.

2.3. Training Data Set

We constructed two distinct training sets to derive metallicities, $T_{\rm eff}$, and distances for stars. The first set, referred to as the "KS" sample, consists of objects common to both the KiDS data set and the spectroscopic samples described earlier. The second set, named the "KG" sample, includes stars that are cross-matched between the KiDS data set and Gaia EDR3.

The selection criteria for both the KS and KG samples are as follows:

1. Stars must be detected in all nine KiDS and VIKING filters, with photometric uncertainties below 0.08 mag.
2. The star classification probability, as indicated by the KiDS image classification ("class_star"), must be at least 0.9.

Table 1
Parameter Ranges for Stars in the KS and KG Training Samples

Parameter	Range
$T_{\rm eff}$	[3311–8846] K
$\log(g)$	[0.20–5.53] dex
[Fe/H]	[-3.92–0.44] dex
$M_{\rm g}$	[2.26–16.22] mag
$(g - r)_0$	[-0.28–1.12] mag

3. Reddening values, sourced from Schlegel et al. (1998), must satisfy $E(B - V) < 0.1$ mag.

Additionally, for the KS sample (which includes spectroscopic data), we apply the following constraints:

1. The spectral signal-to-noise ratio (S/N) must exceed 30.
2. The uncertainty in effective temperature ($\delta T_{\rm eff}$) should be less than 150 K.
3. Errors in surface gravity ($\delta \log(g)$) must not exceed 0.2 dex.
4. Metallicity errors (δ[Fe/H]) should be below 0.2 dex.

For the KG sample, which relies on Gaia data, the additional selection criteria are:

1. Relative errors in Gaia parallaxes ($\delta\varpi/\varpi$) must be less than 0.1.
2. The Renormalized Unit Weight Error (RUWE) must be 1.4 or lower.

For the metal-poor star sample, we relaxed the spectroscopic data quality requirements to ensure sufficient representation of metal-poor stars in the final sample. After applying these selection criteria, the KS sample contains 11,318 stars, while the KG sample includes 136,708 stars. Table 1 summarizes the parameter ranges for stars within both training sets.

The metallicities in the KS sample predominantly range from -0.5 to -1.5 dex, while the absolute magnitudes in the KG sample ($M_{\rm g}$) are primarily between 5 and 10 mag. This distribution introduces an imbalance in the sample, which could affect the accuracy of regression models due to insufficient coverage of certain regions in parameter space. To mitigate this, we applied a grid-based averaging technique within a three-dimensional space defined by $T_{\rm eff}$, $\log(g)$, and [Fe/H] for the KS sample. We computed the average metallicity and $T_{\rm eff}$ at each grid. The grid dimensions were set to 150 K in effective temperature, 0.05 dex in surface gravity, and 0.02 dex in metallicity. This process generated 9312 synthetic stars, including 8248 dwarfs and 1064 giants.

For the KG sample, we divided the parameter space into two dimensions: $M_{\rm g}$ and $(g - r)_0$, with grid steps of 0.01 mag for both parameters. This resulted in 26,996 synthetic stars, with 348 giants and 26,648 dwarfs. The relatively low number of giants is due to the long exposure times of the KiDS survey and

the larger relative errors in Gaia parallaxes for distant, faint stars. These gridded populations of 9001 stars (KS sample) and 26,996 stars (KG sample) constitute our final training data set.

3. Methods

This section describes the methodology employed in this study. First, we distinguish between giant and dwarf stars using a Random Forest Classifier (RFC; Breiman 2001). Once the stars are classified, we apply two separate Random Forest Regression (RFR) models—one for dwarfs and one for giants—to estimate photometric metallicities and T_{eff}. Additionally, we train another pair of RFR models to predict absolute magnitudes for each stellar class. The Random Forest (RF) algorithm (Breiman 2001; Hastie et al. 2009) is an ensemble learning method that aggregates the predictions of numerous decision trees to arrive at a final estimate. Each tree in the forest is built on a randomly selected subset of features and is trained on a bootstrapped sample of the data, ensuring diversity among the trees. The final RF prediction is obtained by averaging the outputs from all trees in the ensemble. Importantly, the algorithm also calculates feature importance ($Imp(x)$) by assessing how much each feature contributes to the decision-making process across all trees. These importance scores are then averaged to provide an overall ranking of feature significance.

In this study, all RF models are implemented using the *scikit-learn* package (Pedregosa et al. 2011) in Python. To construct the RF models, we randomly split the training data, using 70% for model training and reserving the remaining 30% for testing and evaluating predictive performance. In addition to the RF algorithm, we experimented with other machine learning (ML) techniques, including Support Vector Machines (SVM), Multi-layer Perceptron (MLP), and Deep Learning (DL) approaches. After comparing the performance of these methods, Random Forest was found to be the most effective for accurately determining stellar parameters.

3.1. Dwarf-Giant Classification

In this subsection, we describe the classification algorithm used to distinguish between dwarf and giant stars. We employ a Random Forest Classifier (RFC) that processes intrinsic color indices: $(u - g)_0$, $(g - r)_0$, $(r - i)_0$, $(i - Z)_0$, $(Z - Y)_0$, $(Y - J)_0$, $(J - H)_0$, $(H - K_s)_0$, along with nine de-reddened magnitudes. We made an empirical classification of giants and dwarfs on the Kiel diagram following the approach of previous works (e.g., Thomas et al. 2019; Guo et al. 2021). Figure 1 shows a Kiel diagram for all stars in the KS training sample. Stars within the red polygon are classified as giants, while those outside are considered dwarfs. The RFC provides the probability of a star being a giant, P_{giant}, or a dwarf, P_{dwarf}, where P_{dwarf} is complementary to P_{giant}, i.e., $P_{dwarf} = 1 - P_{giant}$.

To evaluate the performance of our classification, we calculate the completeness and purity for both the dwarf and
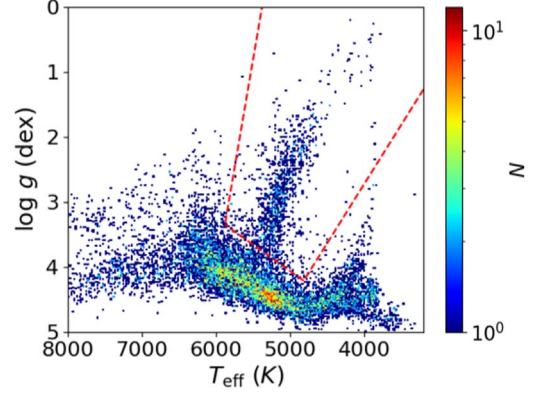
Figure 1. The Kiel diagram for the KS sample used in the dwarf-giant classification. The red polygon encloses stars classified as giants, while those outside are classified as dwarfs.

Table 2
Completeness and Purity of Dwarf and Giant Classifications in the Test Set

Class	Fraction of Sample	Completeness	Purity
Dwarf	0.96	0.93	0.99
Giant	0.04	0.47	0.31

Note. The second column shows the proportion of giants and dwarfs in the sample, the third column lists the completeness of the classifications, and the fourth column provides the purity.

giant classes within the test set, as predicted by the trained RFC. These metrics are presented in Table 2. The depth of the KiDS survey, which reaches beyond 25 mag in the g-band, results in a sample dominated by dwarf stars. The RFC proves highly effective at classifying dwarfs, achieving a completeness of 93% with less than 1% contamination by giant stars. For the less frequent giants, the RFC successfully classifies nearly half, with a purity of around 31%.

3.2. Estimating Metallicities, Effective Temperature, and Absolute Magnitudes

In this step, the algorithm estimates metallicity ([Fe/H]), effective temperature (T_{eff}), and g-band absolute magnitude (M_g) for both giant and dwarf stars. We use a RFR to predict these parameters, with eight intrinsic colors serving as input features. Once trained, the RFR models are applied to the test samples to predict [Fe/H], T_{eff}, and M_g, and the results are compared to the reference values, as shown in Figure 2.

The systematic deviations between the predicted and true values of [Fe/H], T_{eff}, and M_g are minimal, with mean offsets of 0.00 dex for [Fe/H], -2 K for $T_{ef,f}$, and -0.01 mag for M_g. The standard deviations of the residuals are 0.34 dex for [Fe/H], 149 K for T_{eff}, and 0.36 mag for M_g. The dispersion of the [Fe/H] residuals is 0.28 dex for stars with metallicity in the
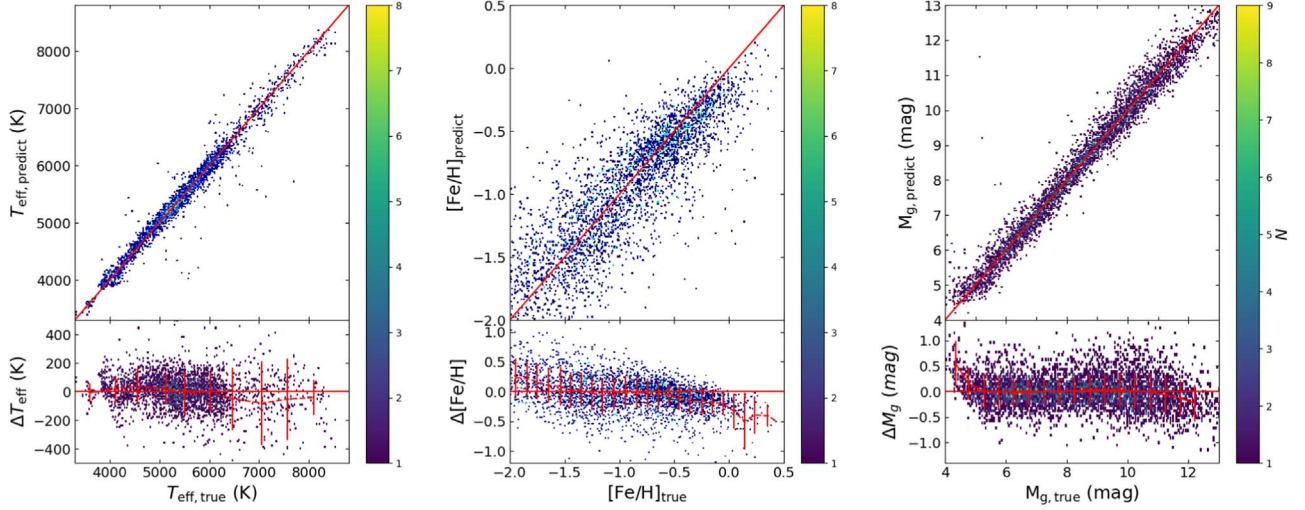
Figure 2. Comparisons between true and RFR-predicted $T_{\rm eff}$ (left panel), [Fe/H] (middle panel), and M_g (right panel) for the test samples. The top panels show predicted values vs. true values, with the red lines indicating the one-to-one correspondence. The bottom panels show the residuals (predicted minus true values), with red lines and error bars indicating the mean differences and standard deviations within individual bins.

range $-2 <$ [Fe/H] < 0.5 dex. Across most of the metallicity range, there are no significant trends in the [Fe/H] residuals, except for the most metal-poor stars ([Fe/H] < -2 dex). For these stars, the algorithm tends to overestimate the metallicity compared to spectroscopic measurements, suggesting reduced reliability for stars with [Fe/H] < -2 dex. The broad u-band filter used in the KiDS survey, similar to that in SDSS, is less sensitive to changes in metallicity for stars with [Fe/H] below -2 dex, which may contribute to this reduced accuracy. The residuals of the effective temperature ($T_{\rm eff}$) show no significant trends across the range of 3500 K–8800 K.

For absolute magnitudes, M_g, most stars in the training set fall within a range of 5–12 mag. The residuals show no significant trends within this range, though some deviations are observed at the extremes ($M_g < 4$ mag and $M_g > 12$ mag). These outliers could be due to smaller sample sizes at the boundaries of the data set and the inherent limitations of machine learning algorithms when extrapolating beyond the core of the training distribution.

3.3. Estimating Uncertainties in the Training Sample

RF models inherently exhibit variability due to the randomization involved in bootstrap sampling and feature selection. Typically, a single partition of the data set into training and test sets is used to train the RF models. In this study, however, we improve the robustness of our results by dividing the training sample into seven distinct training and test set pairs, each generated using a different random seed, while maintaining a consistent partition ratio of 70% training and 30% testing. For each of the seven splits, we train four independent RFR models, each initialized with a different

random state. This approach results in a total of 28 models. The variation in random states ensures that each decision tree evaluates a different subset of features at each node, enhancing the diversity of the ensemble. These models are then used to predict metallicities ([Fe/H]), effective temperature ($T_{\rm eff}$), and absolute magnitudes (M_g) for the stars in our sample. The final predicted values are the averages across the ensemble of 28 models, and the standard deviations of these predictions are taken as a measure of the internal uncertainty of our method.

However, this internal uncertainty does not account for the full predictive error, as it excludes uncertainties from the input measurements. To capture the total uncertainty, we follow the approach of Nepal et al. (2023) and Guiglion et al. (2024). Specifically, we calculate the dispersion of the differences between the true labels and those predicted by the RF models as a function of the true labels in the training sample. This dispersion provides an estimate of the precision relative to the training input, referred to as the "systematic" uncertainty.

In Figure 3, we present both types of uncertainties—internal and systematic—for the [Fe/H], $T_{\rm eff}$, and M_g predictions. The systematic errors are measured by the spread of predictions across the 28 RFR models. For [Fe/H], the systematic errors are 0.24 dex for giants, 0.20 dex for dwarfs, and 0.21 dex overall. For $T_{\rm eff}$, the systematic errors are 90 K for giants, 127 K for dwarfs, and 124 K overall. For M_g, the systematic errors are 0.30 mag for giants, 0.22 mag for dwarfs, and 0.22 mag overall. The fitting errors, which are derived from the dispersion between the true and predicted values in the training data, are also calculated. For [Fe/H], the fitting errors are 0.14 dex for giants, 0.12 dex for dwarfs, and 0.12 dex overall. For $T_{\rm eff}$, the fitting errors are 46 K for giants, 54 K for dwarfs,
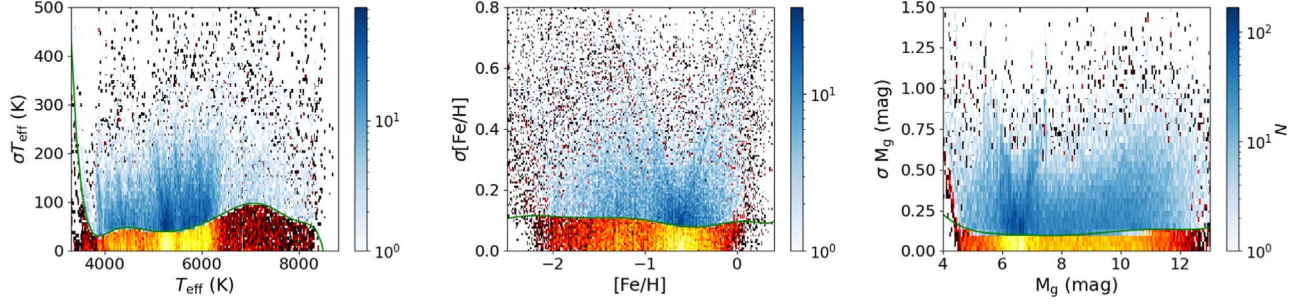
Figure 3. Uncertainty estimates for $T_{\rm eff}$ (left panel), [Fe/H] (middle panel), and M_g (right panel) as a function of the predicted values for the training samples. The red colormap represents the internal uncertainty, computed from the spread of predictions across the 28 RFR models. The green line shows the running dispersion derived from the training labels, representing systematic errors. The blue colormap depicts the 2D histogram of the total uncertainty, calculated as the quadratic sum of internal and systematic errors, providing a comprehensive uncertainty estimate.

Table 3
Field Description for the Final Sample

Column	Label	Description	Unit
1	R.A.	R.A. (J2000) from KiDS DR4	degrees
2	decl.	decl. (J2000) from KiDS DR4	degrees
3	l	Galactic longitude	degrees
4	b	Galactic latitude	degrees
5–8	$u/g/r/i$	Photometric magnitudes from KiDS DR4	mag
9–13	$z/y/J/H/K_{\rm s}$	Photometric magnitudes from VIKING DR4	mag
14–17	$e\_u/e\_g/e\_r/e\_i$	Photometric uncertainties from KiDS DR4	mag
18–22	$e\_Z/e\_Y/e\_J/e\_H/e\_K_{\rm s}$	Photometric uncertainties from VIKING DR4	mag
23	$E(B-V)$	Reddening values	mag
24	$p\_{\rm giant}$	Classification as giant (1) or dwarf (0)	…
25	Teff	Photometric effective temperature	K
26	e_Teff	Uncertainty in effective temperature estimates	K
27	feh	Photometric metallicity	dex
28	e_feh	Uncertainty in metallicity estimates	dex
29	$M_{\rm g}$	Absolute magnitude in g-band	mag
30	$e\_M_{\rm g}$	Uncertainty in $M_{\rm g}$ estimates	mag
31	d	Distance	kpc
32-34	$X/Y/Z$	Galactic Cartesian coordinates	kpc
35	R	Galactocentric distance	kpc

and 53 K overall. For M_g, the fitting errors are 0.17 mag for giants, 0.13 mag for dwarfs, and 0.13 mag overall. We combine these systematic and fitting uncertainties in quadrature to compute the final uncertainties for our predictions. As a result, the typical uncertainty in [Fe/H] is found to be 0.24 dex, the typical uncertainty in $T_{\rm eff}$ is approximately 134 K and the typical uncertainty in M_g is approximately 0.26 mag.

4. Results and Discussion

4.1. The Catalog

We compiled a sample of 820,055 stars from the KiDS and VIKING surveys to estimate their metallicities, effective temperatures, and g-band absolute magnitudes. The selection criteria were as follows:

1. The stars have detections in all nine KiDS and VIKING filters with photometric uncertainties below 0.1 mag.
2. They are classified as stars, with the KiDS image classification parameter, "class_star" > 0.9.

We applied the trained RFC and RFR models to the selected stars to estimate their [Fe/H], $T_{\rm eff}$, and M_g values. We have identified 814,383 dwarfs and 5672 giants in the final sample. The complete catalog, including our estimated parameters and associated uncertainties, is available online at https://nadc.china-vo.org/res/r101527/. Table 3 provides a description of the catalog contents. For each star, we list its coordinates, photometric magnitudes, and corresponding errors in each filter, along with reddening values from Schlegel et al. (1998). Additionally, we have included our classification of the star as a giant (1) or dwarf (0), along with our estimates for its metallicity, effective
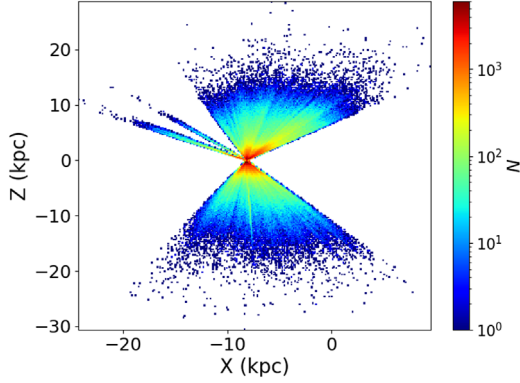
Figure 4. Spatial number-density distributions for stars in the final sample. The panel shows the *X-Z* plane. The color bar at right indicates the number densities. The Sun is positioned at $(X, Y, Z) = (-8.122, 0.0, 0.0)$ kpc.

temperature, *g*-band absolute magnitudes, and the associated uncertainties. The three-dimensional (3D) positions of each star in Galactic Cartesian coordinates (X, Y, Z) are also provided. Here, the coordinate system is centered on the Galactic center: X points away from the Sun, Y aligns with the direction of Galactic rotation, and Z extends toward the North Galactic Pole.

Figure 4 illustrates the spatial distribution of our sample stars in the *X-Z* plane. The Sun is located at $(X, Y, Z) = (-8.122, 0.0, 0.0)$ kpc (Abuter et al. 2018). As shown, the majority of stars lie within 30 kpc of the Sun, with X coordinates extending from about -20 kpc to 5 kpc and Z coordinates -30 kpc to 30 kpc. As previously discussed, our sample predominantly consists of dwarf stars, which are more numerous in the KiDS survey due to their fainter absolute magnitudes. Although dwarfs are detectable at relatively shallow distances, their distance estimates are accurate. However, thanks to the deep magnitude range covered by KiDS, we can probe substantial distances. For the small fraction of giant stars in our sample, the detection limit extends beyond 40 kpc.

4.2. Validation of Our Metallicity Effective Temperature and Distance Estimates

In this section, we validate the accuracy of our photometrically derived metallicities, effective temperature and distances by comparing them with independent measurements from previous studies, including Lin et al. (2022), Huang et al. (2022), Martin et al. (2023), Vallenari et al. (2023), Lu et al. (2024), and Zhang et al. (2024).

Zhang et al. (2024) employed the DD-PAYNE method on the low resolution spectra from Dark Energy Spectroscopic Instrument (DESI) Early Data Release (EDR; DESI Collaboration et al. 2016, 2024), providing abundances for 12 elements and stellar atmospheric parameters. They obtained a precision of 0.05 dex for metallicity and an uncertainty of 20 K for effective temperature. We have identified 21,687 stars common

to both data sets, with 143 classified as giants and 21,544 as dwarfs. Assuming the classifications by Zhang et al. (2024) are flawless, our classification achieves a purity of 93% for dwarfs and 67% for giants. The left panel of Figure 6 shows a comparison between our results and their metallicities. Overall, our metallicities are in good agreement with those from Zhang et al. (2024), with a negligible systematic difference of $\mu = -0.02$ dex and a standard deviation of $\sigma = 0.29$ dex. In the left panel of Figure 6, we observe an anomalous region where some stars have DESI metallicities between -1.1 and -0.7 dex, while our estimates range from -0.4 to 0 dex. Upon further investigation, these stars are primarily cool giants with $T_{\mathrm{eff}} \approx 4000$ K. It is likely that the DD-PAYNE method used by Zhang et al. (2024) underestimates the metallicities for these cooler giants. The left panel of Figure 5 compares our T_{eff} estimates with those from Zhang et al. (2024). Overall, the results are consistent, with a mean difference of $\mu = 80$ K and a standard deviation of $\sigma = 139$ K. However, a larger systematic offset is observed for cooler stars ($T_{\mathrm{eff}} < 4300$ K). Cross-matching with the spectroscopic data used in this study reveals an offset of 89 K in this regime, suggesting that the Zhang et al. (2024) results may have limitations in accurately measuring T_{eff} for cooler stars.

Martin et al. (2023) calculated metallicities by utilizing synthetic photometries obtained from the Gaia BP/RP spectra, in conjunction with the observed photometries of the narrow-band filters from the Pristine Survey. Metallicity estimates were obtained using the stellar loci method. They released two photometric metallicity catalogs: the Pristine-Gaia synthetic catalog covering most of the sky, and the higher-quality Pristine DR1 catalog, which is limited to the Pristine footprint. We compare our results with their Pristine DR1 catalog, which includes 27,565 giants and 276 dwarfs common to both data sets. The middle panel of Figure 6 shows excellent agreement between our metallicities and those from Martin et al. (2023), with a negligible median difference of $\mu = -0.02$ dex and a small standard deviation of $\sigma = 0.28$ dex.

Lu et al. (2024) combined photometric data from GALEX GR6+7 AIS and Gaia. They established relationships between stellar loci, metallicity, and absolute magnitude (M_G) to estimate stellar metallicities. We found 4004 common sources between our catalog and theirs, with 20 stars consistently classified as giants and 3825 as dwarfs. Assuming that their classifications are correct, our classification shows a purity of 87% for dwarfs and 96% for giants. The right panel of Figure 6 shows that our metallicities are consistent with those from Lu et al. (2024), with only a small median difference of $\mu = -0.08$ dex and a standard deviation of $\sigma = 0.31$ dex.

Lin et al. (2022) used a Bayesian isochrone-fitting method to derive distances, extinctions, and stellar parameters (T_{eff}, $\log g$, [Fe/H]) for nearly 18 million stars from SMSS DR3. Their analysis incorporated 14 photometric bands from four surveys (SMSS, Gaia, 2MASS, and WISE). For the comparison, we
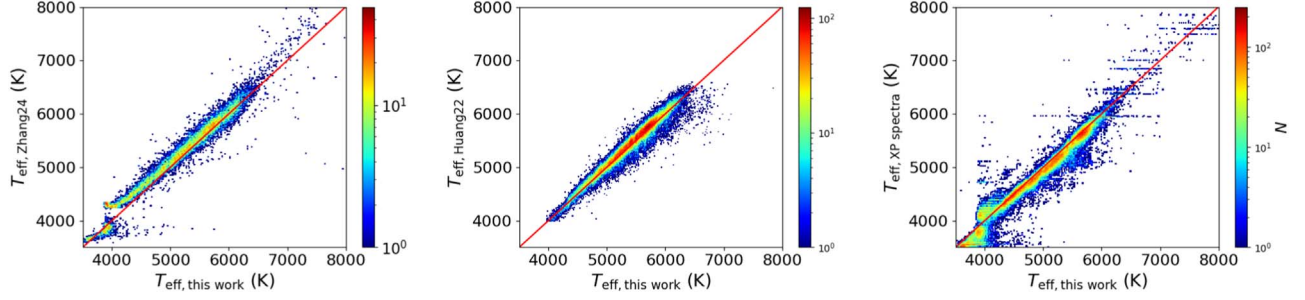
Figure 5. Comparison of our effective temperature estimates with those from Zhang et al. (2024, left panel), Huang et al. (2022, middle panel), and Vallenari et al. (2023, right panel). The color scales represent source density, and the red lines indicate equality between the data sets in each panel for reference.
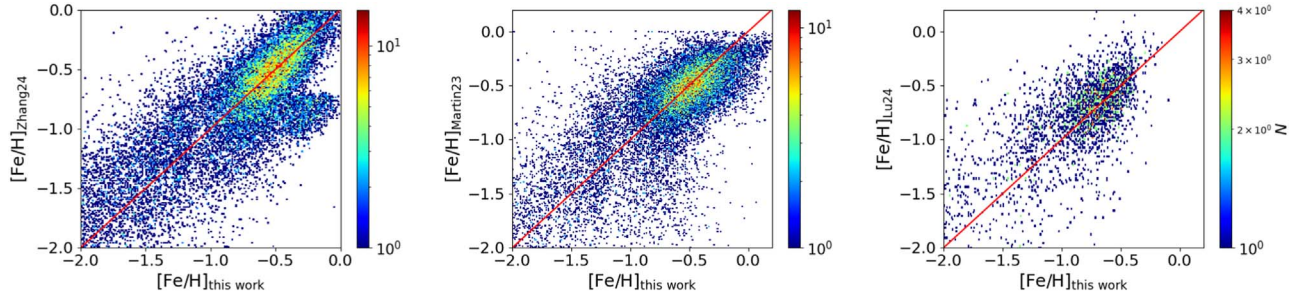

Figure 6. Comparison of our metallicity estimates with those from Zhang et al. (2024, left panel), Martin et al. (2023, middle panel), and Lu et al. (2024, right panel). The color scales represent the density of sources. The red lines indicate equality between the two data sets in each panel to guide the eye.
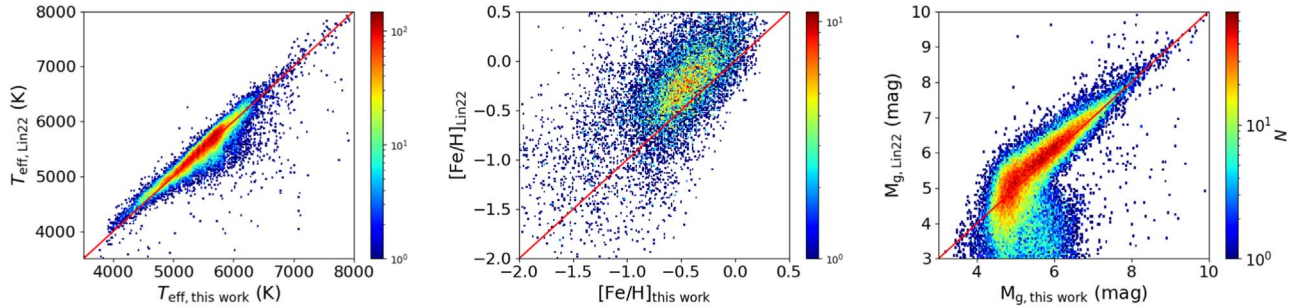
Figure 7. Comparisons of our findings with those of Lin et al. (2022) in terms of $T_{\rm eff}$ (left panel), [Fe/H] (middle panel), and M_g (right panel). The color scales indicate the density of sources, and the red lines represent equality between the two data sets.

selected stars from both catalogs with reliable measurements. In our KiDS sample, we required metallicity uncertainties of $\sigma\_$ [Fe/H] < 0.2 and $\sigma\_T_{\rm eff}$ < 150 K. From the Lin et al. (2022) catalog, we selected stars with metallicity uncertainties ERR_FEH < 0.15, distance uncertainties ERR_DIST_A-DOP/DIST_APOD < 0.1, and clean photometry (band_flags set to "00000000000000"). This yields 13,796 common stars for the metallicity and $T_{\rm eff}$ comparison, and 57,893 stars for the distance comparison. The comparison between our results and those from Lin et al. (2022) is shown in Figure 7. For $T_{\rm eff}$, the agreement is excellent, with a negligible mean difference of $\mu = 3$ K and a standard deviation of $\sigma = 150$ K. For metallicity,

however, we find a significant systematic difference of $\mu = -0.24$ dex. Given the strong agreement between our metallicities and those of Zhang et al. (2024), Martin et al. (2023), and Lu et al. (2024), we suspect that Lin et al. (2022) may have overestimated the metallicities of their stars. For absolute magnitudes (M_g), we find good agreement with Lin et al. (2022), especially for stars with $M_g > 4$ mag, where the systematic difference is only $\mu = -0.04$ mag and the dispersion is $\sigma = 0.49$ mag.

Huang et al. (2022) derived $T_{\rm eff}$ by calibrating photometric data from SMSS DR2 and Gaia EDR3 using metallicity-dependent color–$T_{\rm eff}$ relations. Among the 118,549 common
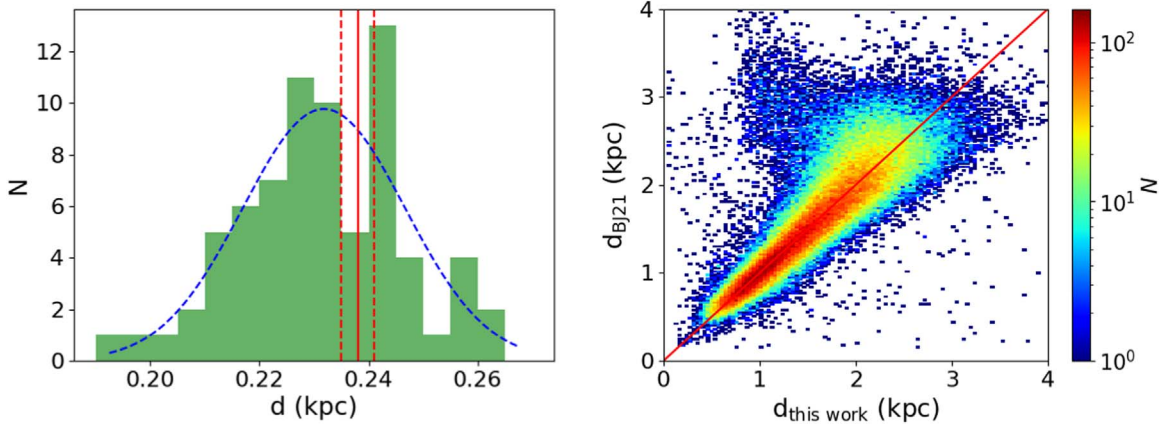
Figure 8. Left panel: Histogram of our distance estimates for the member stars of the open cluster Blanco_1, with the blue dashed line showing a Gaussian fit. The vertical red solid and dashed lines indicate the distance and uncertainty of Blanco_1 from Cantat-Gaudin et al. (2020). Right panel: Comparison of our distance estimates with those from Bailer-Jones et al. (2021). The red line denotes equality between the two data sets.

sources between our catalog and theirs, we restricted the comparison to stars with errors smaller than 100 K for reliability. The middle panel of Figure 5 shows that our T_{eff} values agree well with Huang et al. (2022), with a mean difference of $\mu = -24$ K and a dispersion of $\sigma = 88$ K. The observed small negative offset is primarily attributed to a systematic difference of -68 K between the T_{eff} values reported by Huang et al. (2022) and the spectroscopic data used in this study.

Vallenari et al. (2023) derived effective temperatures (T_{eff}) using Gaia XP spectra through the GSP-Phot module, which estimates T_{eff} by fitting observed XP spectra to stellar models. The right panel of Figure 5 compares our T_{eff} estimates with those from Gaia XP spectra. For reliable comparisons, we selected 98,720 common sources from Vallenari et al. (2023) with well-constrained temperatures (B_T_{eff}_xa $-$ b_T_{eff}_x $<$ 20 K). Our results show a mean difference of $\mu = 89$ K and a dispersion of $\sigma = 202$ K. While our T_{eff} estimates are generally consistent with those of Vallenari et al. (2023), we observe a noticeable systematic offset. Additionally, discrepancies are more pronounced at lower temperatures ($T_{\mathrm{eff}} \sim 4000$ K). A comparison between the effective temperatures from Vallenari et al. (2023) and spectroscopic data sets, such as those from LAMOST, reveals similar offsets and inconsistencies at the low-temperature end.

To assess the accuracy of the distances in our sample, we compared our results with two other independent data sets: the open cluster distances from Cantat-Gaudin et al. (2020) and distances derived from Gaia parallaxes by Bailer-Jones et al. (2021). Due to the fact that the KiDS survey primarily covers high Galactic latitude regions and masks globular clusters (Kuijken et al. 2019), we identified only one open cluster suitable for comparison: Blanco_1. Our sample contains 85 members of this cluster. As shown in the left panel of Figure 8,

the literature distance for Blanco_1 is 0.237 ± 0.003 kpc (Cantat-Gaudin et al. 2020), while our KiDS-derived mean distance is 0.225 kpc, with a dispersion of 0.027 kpc. The close agreement between our results and the literature values confirms the reliability of our distance estimates for these stars.

We also compared our distance estimates with those provided by Bailer-Jones et al. (2021), which are based on Gaia parallax measurements. To ensure the quality of the comparison, we selected sources with reliable Gaia parallaxes, defined as having fractional parallax errors of $0.1 < \delta\varpi/\varpi < 0.2$. This selection resulted in 99,526 common sources, of which 385 are classified as giants and 99,141 as dwarfs. The comparison is shown in the right panel of Figure 8. Our distances are in good agreement with those from Bailer-Jones et al. (2021), with a mean difference of 0.048 kpc and a dispersion of 0.38 kpc.

5. Conclusion

Using spectroscopic training data sets, we developed RFC and RFR models to estimate the metallicity ([Fe/H]), effective temperature (T_{eff}), and absolute magnitude (M_{g}) of stars, based on photometric data from KiDS DR4 and VIKING DR4. The input features for the RFC model, which classifies stars as giants or dwarfs, include intrinsic colors and de-reddened magnitudes, while the output is the classification of each star as either a giant or a dwarf. The RFR models use intrinsic colors as input features to predict T_{eff}, [Fe/H], and M_{g} for each star. Our models demonstrate strong predictive performance. For T_{eff}, the systematic difference compared to spectroscopic measurements is -2 K with a scatter of 149 K. For metallicity ([Fe/H]), the systematic difference is 0.00 dex with a scatter of 0.28 dex. For absolute magnitudes (M_{g}), we find a systematic difference of -0.01 mag and a dispersion of 0.36 mag in the

test sample. These results indicate the reliability of our models in deriving accurate stellar parameters from photometric data.

Using our models, we derived metallicity, T_{eff}, and M_g values for 820,055 stars in the KiDS and VIKING data sets, including 814,383 dwarfs and 5672 giants. When comparing our T_{eff} and metallicity estimates with those from spectroscopic data and other studies, we find small systematic offsets. For T_{eff}, the offset ranges between 0 and 100 K, with a dispersion of 100–200 K. For metallicity, the systematic offset is between 0 and 0.08 dex, with a dispersion of 0.27–0.37 dex. These comparisons highlight the consistency of our results with previous studies and spectroscopic measurements. The distances derived in our final sample also show good agreement with values from the literature. Furthermore, our algorithm will be applied to larger stellar samples obtained from upcoming surveys such as the Multi-channel Photometric Survey Telescope (Mephisto) and the Chinese Space Station Telescope (CSST).

References

Abuter, R., Amorim, A., Anugu, N., et al. 2018, A&A, 615, 10
Aguado, D. S., Youakim, K., González Hernández, J. I., et al. 2019, MNRAS, 490, 2241
An, D., & Beers, T. C. 2020, ApJ, 897, 39
An, D., & Beers, T. C. 2021a, ApJ, 907, 101
An, D., & Beers, T. C. 2021b, ApJ, 918, 74
Bailer-Jones, C., Rybizki, J., Fouesneau, M., Demleitner, M., & Andrae, R. 2021, AJ, 161, 147
Breiman, L. 2001, Machine learning, 45, 5
Buder, S., Sharma, S., Kos, J., et al. 2021, MNRAS, 506, 150
Cantat-Gaudin, T., Anders, F., Castro-Ginard, A., et al. 2020, A&A, 640, A1
Capaccioli, M., & Schipani, P. 2011, Msngr, 146, 2
Casagrande, L., Schönrich, R., Asplund, M., et al. 2011, A&A, 530, A138
Cenarro, A. e., Moles, M., Cristóbal-Hornillos, D., et al. 2019, A&A, 622, A176
Chen, B., Liu, X., Yuan, H., et al. 2018, MNRAS, 476, 3278
Chiti, A., Frebel, A., Mardini, M. K., et al. 2021a, ApJS, 254, 31
Chiti, A., Mardini, M. K., Frebel, A., & Daniel, T. 2021b, ApJL, 911, L23
DESI Collaboration, Adame, A., Aguilar, J., et al. 2024, AJ, 168, 33
DESI Collaboration, Aghamousa, A., Aguilar, J., et al. 2016, arXiv:1611.00036
Da Costa, G. S., Bessell, M. S., Mackey, A., et al. 2019, MNRAS, 489, 5900
De Silva, G. M., Freeman, K. C., Bland-Hawthorn, J., et al. 2015, MNRAS, 449, 2604
Edge, A., Sutherland, W., Kuijken, K., et al. 2013, Msngr, 154, 32
Fan, Z., Zhao, G., Wang, W., et al. 2023, ApJS, 268, 9
Fernández-Alvar, E., Kordopatis, G., Hill, V., et al. 2021, MNRAS, 508, 1509
Gaia Collaboration, Brown, A., Vallenari, A., et al. 2018, A&A, 616, A1
Gaia Collaboration, Brown, A. G., Vallenari, A., et al. 2016, A&A, 595, A2
Gaia Collaboration, Brown, A. G., Vallenari, A., et al. 2021, A&A, 649, A1
Guiglion, G., Nepal, S., Chiappini, C., et al. 2024, A&A, 682, A9
Guo, H.-L., Chen, B.-Q., Yuan, H.-B., et al. 2021, ApJ, 906, 47
Hackshaw, Z., Hawkins, K., Filion, C., et al. 2024, ApJ, 977, 143

Hastie, T., Tibshirani, R., Friedman, J. H., & Friedman, J. H. 2009, The Elements of Statistical Learning: Data Mining, Inference, and Prediction, 2 (Berlin: Springer)

Hattori, K. 2025, ApJ, 980, 90

Huang, Y., Beers, T. C., Wolf, C., et al. 2022, ApJ, 925, 164

Huang, Y., Beers, T. C., Yuan, H., et al. 2023, ApJ, 957, 65

Huang, Y., Chen, B.-Q., Yuan, H.-B., et al. 2019, ApJS, 243, 7

Ivezić, Ž., Sesar, B., Jurić, M., et al. 2008, ApJ, 684, 287

Jacobson, H. R., Keller, S., Frebel, A., et al. 2015, ApJ, 807, 171

Kim, B., & Lépine, S. 2022, MNRAS, 510, 4308

Kuijken, K., Heymans, C., Dvornik, A., et al. 2019, A&A, 625, A2

Kuijken, K., et al. 2011, Msngr, 146, 3

Li, H., Tan, K., & Zhao, G. 2018, ApJS, 238, 16

Lian, J., Bergemann, M., Pillepich, A., Zasowski, G., & Lane, R. R. 2023, NatAs, 7, 951

Lin, J., Casagrande, L., & Asplund, M. 2022, MNRAS, 510, 433

Liu, X.-W., Zhao, G., & Hou, J.-L. 2015, RAA, 15, 1089

Lu, X., Yuan, H., Xu, S., et al. 2024, ApJS, 271, 26

Luo, A.-L., Zhao, Y.-H., Zhao, G., et al. 2015, RAA, 15, 1095

Majewski, S. R., Schiavon, R. P., Frinchaboy, P. M., et al. 2017, AJ, 154, 94

Martin, N. F., Starkenburg, E., Yuan, Z., et al. 2023, A&A, 692, A115

Mendes de Oliveira, C. E., Ribeiro, T., Schoenell, W., et al. 2019, MNRAS, 489, 241

Nepal, S., Guiglion, G., De Jong, R., et al. 2023, A&A, 671, A61

Nidever, D. L., Bovy, J., Bird, J. C., et al. 2014, ApJ, 796, 38

Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2011, JLMR, 12, 2825

Peng, X., Du, C., Wu, Z., Ma, J., & Zhou, X. 2013, MNRAS, 434, 3165

Rojas-Arriagada, A., Zasowski, G., Schultheis, M., et al. 2020, MNRAS, 499, 1037

Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, ApJ, 500, 525

Starkenburg, E., Martin, N., Youakim, K., et al. 2017, MNRAS, 471, 2587

Steinmetz, M., Zwitter, T., Siebert, A., et al. 2006, AJ, 132, 1645

Sun, M., Chen, B., Guo, H., et al. 2023, AJ, 166, 126

Sutherland, W., Emerson, J., Dalton, G., et al. 2015, A&A, 575, A25

Thomas, G. F., Annau, N., McConnachie, A., et al. 2019, ApJ, 886, 23

Vallenari, A., Brown, A. G., Prusti, T., et al. 2023, A&A, 674, A1

Wang, C., Liu, X., Xiang, M., et al. 2019, MNRAS, 482, 2189

Whitten, D. D., Placco, V. M., Beers, T. C., et al. 2021, ApJ, 912, 147

Wojno, J., Kordopatis, G., Piffl, T., et al. 2017, MNRAS, 468, 3368

Wolf, C., Onken, C. A., Luvaul, L. C., et al. 2018, PASA, 35, e010

Wu, Y., Du, B., Luo, A., Zhao, Y., & Yuan, H. 2014, in Proc. the Int. Astronomical Union, 10 (Cambridge: Cambridge Univ. Press)

Xiao, K., Huang, B., Huang, Y., et al. 2024, ApJL, 968, L24

Xu, S., Yuan, H., Niu, Z., et al. 2022, ApJS, 258, 44

Yan, Y., Du, C., Liu, S., et al. 2019, ApJ, 880, 36

Yanny, B., Rockosi, C., Newberg, H. J., et al. 2009, AJ, 137, 4377

Youakim, K., Starkenburg, E., Martin, N. F., et al. 2020, MNRAS, 492, 4986

Yuan, H., Liu, X., Xiang, M., Huang, Y., & Chen, B. 2015b, ApJ, 803, 13

Yuan, H.-B., Liu, X.-W., Huo, Z.-Y., et al. 2015a, MNRAS, 448, 855

Zhang, H., Ardern-Arentsen, A., & Belokurov, V. 2024, MNRAS, 533, 889

Zhang, M., Xiang, M., Ting, Y.-S., et al. 2024, ApJS, 273, 19

Zhao, G., Zhao, Y.-H., Chu, Y.-Q., Jing, Y.-P., & Deng, L.-C. 2012, RAA, 12, 723