Machine-learning Regression of Stellar Effective Temperatures in the Second Gaia Data Release

Yu Bai1, JiFeng Liu1, ZhongRui Bai1 Song Wang1, and DongWei Fan3

1 Key Laboratory of Optical Astronomy, National Astronomical Observatories, Chinese Academy of Sciences, 20A Datun Road, Chaoyang District, Beijing 100012, People’s Republic of China; ybai@nao.cas.cn
2 College of Astronomy and Space Sciences, University of Chinese Academy of Sciences, Beijing 100049, People’s Republic of China
3 National Astronomical Observatories, Chinese Academy of Sciences, 20A Datun Road, Chaoyang District, Beijing 100012, People’s Republic of China

Abstract

This paper reports on the application of the supervised machine-learning algorithm to the stellar effective temperature regression for the second Gaia data release, based on the combination of the stars in four spectroscopic surveys: the Large Sky Area Multi-Object Fiber Spectroscopic Telescope, Sloan Extension for Galactic Understanding and Exploration, the Apache Point Observatory Galactic Evolution Experiment, and the Radial Velocity Extension. This combination, of about four million stars, enables us to construct one of the largest training samples for the regression and further predict reliable stellar temperatures with a rms error of 191 K. This result is more precise than that given by the Gaia second data release that is based on about sixty thousands stars. After a series of data cleaning processes, the input features that feed the regressor are carefully selected from the Gaia parameters, including the colors, the 3D position, and the proper motion. These Gaia parameters are used to predict effective temperatures for 132,739,323 valid stars in the second Gaia data release. We also present a new method for blind tests and a test for external regression without additional data. The machine-learning algorithm fed with the parameters only in one catalog provides us with an effective approach to maximize the sample size for prediction, and this methodology has a wide application prospect in future studies of astrophysics.

Key words: methods: data analysis – stars: fundamental parameters – techniques: spectroscopic

Supporting material: machine-readable table

1. Introduction

The ESA space mission Gaia is performing an all-sky astrometric, photometric, and radial velocity survey at optical wavelengths (Gaia Collaboration et al. 2016). The main objective of the Gaia mission is to survey more than one billion stars in order to understand the structure, formation, and evolution of our Galaxy. The second data release (Gaia DR2; Gaia Collaboration et al. 2018) includes a total of 1.69 billion sources with G-band photometry based on 22 months of observations. Of these, 1.38 billion sources also have the integrated fluxes from the blue and red photometer (BP and RP) spectrophotometers, which span 3300–6800 Å and 6400–10500 Å, respectively.

Three broad photometric bands have been used to infer stellar effective temperatures ($T_{\text{eff}}$) for all sources brighter than $G = 17$ mag with $T_{\text{eff}}$ in the range of 3000–10,000 K (Andrae et al. 2018). A machine-learning algorithm, random forest (RF), has been applied to regress $T_{\text{eff}}$. The training data of the algorithm is a combination of five spectrum- or photometry-based catalogs with a total of 65,200 stars. A typical accuracy of the regression is 324 K which is estimated from the 50% hold-out validation, and no blind test is performed to quantify the performance of the regression and to avoid overfitting.

However, decoupling stellar temperatures and interstellar extinction is a complex problem, and more parameters than two colors is required to regress temperatures with good accuracy (Bai et al. 2019). Moreover, diversity of a sample in a parameter space has been proven to be an influential aspect, and has a strong impact on the overall performance of machine learning (Wang & Yao 2009; Wang et al. 2009). The small size of the training set in Andrae et al. (2018) could limit the diversity of the stellar sample and further cause regressed $T_{\text{eff}}$ having high systematic deviation (e.g., Pelisoli et al. 2019; Sahliholdt et al. 2019).

The availability of spectrum-based stellar parameters for large numbers is now possible thanks to the observations of large Galactic spectral surveys. Large Sky Area Multi-Object Fiber Spectroscopic Telescope (LAMOST; Luo et al. 2015) data release 5 (DR5) was available to domestic users in 2017 December, which includes over eight million observations of stars. This archive data after six years of accumulation is a treasure for various studies. One of the catalogs mounted on the archive is the A-, F-, G-, and K-type stars catalog, in which the stellar parameters, $T_{\text{eff}}$, log g, and [Fe/H] are determined by the LAMOST stellar parameter pipeline (Wu et al. 2014).

Another large survey is the Sloan Extension for Galactic Understanding and Exploration (SEGUE; Yanny et al. 2009). The spectra are processed through the SEGUE Stellar Parameter Pipeline (SSPP; Allende Prieto et al. 2008; Lee et al. 2008a, 2008b; Smolinski et al. 2011), which uses a number of methods to derive accurate estimates of stellar parameters, $T_{\text{eff}}$, log g, [Fe/H], [α/Fe], and [C/Fe].

Different from the two abovementioned surveys that are in the optical band, the Apache Point Observatory Galactic Evolution Experiment (APOGEE), as one of the programs in both SDSS-III and SDSS-IV, has collected high-resolution ($R \sim 22,500$), high signal-to-noise (S/N > 100) near-infrared (1.51–1.71 μm) spectra of 277,000 stars (data release 15) across the Milky Way (Majewski et al. 2017). These stars are dominated by red giants selected from the Two Micron All Sky Survey. Their stellar parameters and chemical abundances are

4 See http://dr5.lamost.org/.
estimated by the APOGEE Stellar Parameters and Chemical Abundances Pipeline (ASPCAP; García Pérez et al. 2016).
These surveys aim mainly at stars located in the northern hemisphere, while the Radial Velocity Extension (RAVE) covers the southern sky. It is designed to provide stellar parameters to complement missions that focus on obtaining radial velocities to study the motions of stars in the Milky Way’s thin and thick disk and stellar halo (Steinmetz et al. 2006). Its pipeline processes the RAVE spectra and derives estimates of $T_{\text{eff}}$, log $g$, and [Fe/H] (Kunder et al. 2017).

The large amount of spectroscopic data in these four catalogs provides us with an opportunity to apply machine-learning technology to regress $T_{\text{eff}}$ effectively. In Section 2, we present validation samples and a method of data cleaning. Various input parameters are also explored to regress temperatures in the section. We apply the regressor and present a revised version of the $T_{\text{eff}}$ catalog for Gaia DR2 in Section 3. Blind tests and external regression tests are also provided. A discussion is given in Section 4.

2. Methodology

2.1. Validation Samples

The A-, F-, G-, and K-type stars catalog of LAMOST DR5 includes the estimates of the stellar $T_{\text{eff}}$ with the application of a correlation function interpolation (Du et al. 2012) and Université de Lyon spectroscopic analysis software (Koleva et al. 2009). These two approaches are based on the distribution and morphology of absorption lines in normalized stellar spectra, independent from Galactic extinction. The temperatures are in the range of $3460 < T_{\text{eff}} < 8500$ K with the uncertainty of $\sim 110$ K (Gao et al. 2015). We extract 4,340,931 unique stars in the catalog, and cross-match them to Gaia DR2 with a radius of 2′, which yields 4,249,013 stars.

For the SEGUE survey, we adopt $T_{\text{eff}}$ estimated with the SSPP which is also based on the distribution and morphology of stellar absorption lines. The temperatures range from $4000 < T_{\text{eff}} < 9710$ K with the typical uncertainty of $\sim 180$ K. We perform a cross-match with Gaia DR2, and obtain 1,037,433 stars.

The $T_{\text{eff}}$ of the APOGEE stars is estimated by ASPCAP, which searches a multidimensional grid for the best-matching synthetic spectrum (Mészáros et al. 2013). The temperatures are in the range of $3550 < T_{\text{eff}} < 8200$ K, with a typical uncertainty of $\sim 100$ K. We cross-match these stars with Gaia DR2, and obtain 275,019 stars.

The pipeline of RAVE is based on the combination of the MATrix Inversion for Spectral SynthEsis (MATISSE; Recio-Blanco et al. 2006) algorithm and the decision tree algorithm for astrophysics (Bijaoui et al. 2012). This pipeline is valid for stars with temperatures between 4000 and 8000 K. The estimated errors in $T_{\text{eff}}$ are approximately 250 K and $\sim 100$ K for spectra with S/N $\sim 50$ (Kunder et al. 2017). The cross-match with Gaia DR2 yields 518,812 stars.

Here we only adopt the $T_{\text{eff}}$ from the spectroscopic surveys, since their stellar parameters are highly reliable (Mathur et al. 2017) compared to the photometric surveys, e.g., the Kepler Input Catalog. As a result, there are 6,080,277 Gaia matched stars in the four catalogs.

2.2. Data Cleaning

Andrae et al. (2018) applied various filters to remove bad data, and some of them are also adopted in our data cleaning processes. We remove the samples with $\varpi \leq 0$ or $\sigma_{\varpi}/\varpi > 0.2$. The samples with the high or negative relative uncertainties of the parallaxes may suffer large bias in the distance measurements (Luri et al. 2018) or could include a large fraction of non-stellar objects (Bai et al. 2018). We also exclude the samples with $\sigma(T_{\text{eff}})/T_{\text{eff}} > 0.05$ to remove inaccurate estimates.

We plot a color–color diagram in Figure 1 and select the region with number densities higher than 150 per 0.01 mag$^2$. A logarithmic function is used to fit the colors of the sample in this region. The best-fit function is $G - G_{\text{BP}} = 1.79 - \log_{10}(G_{\text{BP}} - G + 0.42) + 0.71$. We shift the function with $\pm 0.15$ mag to select the samples with good photometry. This good-quality region is marked with the black solid lines in Figure 1. The region defined by the logarithmic function shows a better consistency with the stellar locus than the cuts in Andrae et al. (2018). As a result, the training sample contains 3,810,143 stars.

The $T_{\text{eff}}$ distribution of these stars is shown in Figure 2, which is inhomogeneous. We give the impact of this on the prediction for Gaia DR2 in Section 3. The training sample is dominated by F, G, and K stars with $T_{\text{eff}} \sim 5000$-$6000$ K, different from the distribution of the training sample in Andrae et al. (2018) which concentrates in five specific temperatures.

We present the differences between the $T_{\text{eff}}$ in Gaia DR2 and the literature estimates in Figure 3. Some vertical concentrated regions are shown in the LAMOST and SSPP panels. The stars in these regions have similar temperatures in Gaia DR2, but have different estimates in the spectrum-based catalogs. This implies that the temperatures given by Gaia DR2 are probably still coupled with Galactic extinction, since the regressor was built with two colors from a small sample size. These two colors could not provide enough information to decouple the temperatures from the extinction (Davenport et al. 2014).
The working theory of the RF is that it builds an ensemble of unpruned decision trees and merges them together to obtain a regressor, but try different combinations of input parameters. One big advantage of RF is fast learning once for validation. The 10 folded cross-validation can provide an overall assessment of the regression.

The root-mean-squared error (RMSE) is adopted to stand for the performance of the regressors (Table 1). We find that the regressor of eight input parameters, $l$, $b$, $\Delta \varpi$, $\mu_\varpi$, $\mu_b$, $BP$, $G$, and $GP$, shows the best performance with the RMSE of 191 K, while the regressor that is constructed with only two colors is the worst. The one-to-one correlation of the best regression is shown in the left panel of Figure 4. We bin the regressed $T_{\text{eff}}$ with a step size of 100 K and the standard deviation ($\sigma$) are listed in Table 3.

### 3. Result

We now use the criteria below to select the samples in Gaia DR2, which yields 132,739,323 stars.

$$\frac{\Delta T_{\text{eff}}}{T_{\text{eff}}} < 0.05,$$

$$0 < \Delta \varpi / \varpi < 0.2,$$

$$G - G_{\text{BP}} \leq 1.79 \cdot \log_{10}(G_{\text{BP}} - G + 0.42) + 0.71 + 0.15,$$

$$G - G_{\text{RP}} \geq 1.79 \cdot \log_{10}(G_{\text{BP}} - G + 0.42) + 0.71 - 0.15.$$
The algorithm of RF constructed with eight input parameters is applied to regress their \( T_{\text{eff}} \), and the result is listed in Table 2.

The size of the catalog is a little smaller than that in Andrae et al. (2018), since we use more strict criteria. We compare our results with \( T_{\text{eff}} \) in \textit{Gaia} DR2 in Figure 5. The \textit{Gaia} \( T_{\text{eff}} \) is concentrated in some specific temperatures, 4000 K, 4500 K, 5000 K, 5500 K, and 6000 K. These temperatures are consisted of the peaks in the distribution of the training set of the regressor. The inhomogeneous training set yielded outputs with similar distribution (see Figure 5 and Figure 18 in Andrae et al. 2018). Our \( T_{\text{eff}} \) distribution concentrates in two much broader peaks, 5000 K and 6000 K, implying better homogeneity.

### 3.1. Blind Tests

A blind test is an effective method to measure the performance of a machine-learning classifier or regressor (Bai et al. 2019). It evaluates the prediction accuracy with data that are not in the training set, and provides validation that a regressor is working sufficiently to output reliable results.

In order to apply blind tests, we train subregressors with eight input parameters in three catalogs and use the forth catalog to test these subregressors. The LAMOST DR5 is always included in the training set, since it accounts for 87% of the stars in our training set. We omit the testing stars that are located outside the parameter spaces of the subregressors to avoid external regression. We present the results of the blind tests in Figure 6, and list the parameters of the Gaussian fit to the total residuals in Table 3.

The blind tests show that the offsets of the total residuals are below 112 K and the standard deviations are less than 200 K. Lee et al. (2015) have applied the SSPP to LAMOST stars and compared the results to those from the RAVE and APOGEE catalogs. The offsets of \( T_{\text{eff}} \) between different pipelines are from 36 to 73 K, and the standard deviations are from 79 to 172 K. This indicates that our regressor can output the stellar temperatures at similar accuracy to the results of spectrum-based pipelines.

### 3.2. External Regression

In order to test the stars that are located outside our criteria, we adopt the subregressors trained with three catalogs and use the stars in the forth catalog to apply external regression. The stars are divided into two subclasses, located outside the quality cuts in Figure 1, and with 0.2 < \( \Delta \sigma / \sigma \) < 0.4 (Table 4).

The result is shown in Figure 7. The \( T_{\text{eff}} \) is systematically overestimated for the first testing subclass, and their RMSEs are twice as large as those of the blind tests. The photometry that feed to the subregressors is probably worse than the photometry that is located inside of the quality cuts, and the subregressors could not predict \( T_{\text{eff}} \) with good accuracy.

### Table 2

| Source ID | Regressed \( T_{\text{eff}} \) (K) |
|-----------|----------------------------------|
| 2448780173659609728 | 5128 ± 634 |
| 2448781208748235648 | 5463 ± 69 |
| 2448690965856595888 | 5984 ± 91 |
| 244869777484387072 | 4333 ± 396 |
| 2448783991887042176 | 4166 ± 65 |
| 2448690258520723712 | 5062 ± 55 |
| 2448690327240200576 | 5846 ± 89 |
| 244869811844125184 | 4328 ± 385 |
| 2448784953959717376 | 5382 ± 58 |
| 2448783991887042048 | 4888 ± 539 |

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

### Table 3

| Source ID | \( \mu \) (K) | \( \sigma \) (K) | RMSE (K) |
|-----------|---------------|----------------|---------|
| Cross-validation | -17 ± 1 | 91 ± 1 | 191 |
| SSPP     | 58 ± 2 | 87 ± 2 | 179 |
| RAVE     | -112 ± 4 | 196 ± 4 | 260 |
| APOGEE   | -28 ± 3 | 119 ± 3 | 191 |

Figure 5. Density map of \( T_{\text{eff}} \) in \textit{Gaia} DR2 vs. regressed \( T_{\text{eff}} \). Normalized histograms of the \( T_{\text{eff}} \) distributions are plotted in the top and left panels.
Figure 7. Result of external regression. Left panel: the stars that are located outside of the quality cuts in Figure 1. Right panel: the stars with $0.2 < \Delta \text{par} / \text{par} < 0.4$.

Table 4

| Catalog | First Subclass | Second Subclass |
|---------|----------------|-----------------|
| SSPP    | 666 (383)      | 74,015 (184)    |
| RAVE    | 247 (510)      | 1,245 (314)     |
| APOGEE  | 225 (406)      | 22,444 (315)    |

Note. The numbers in brackets are the RMSEs in Kelvin units.

For the second subclass, most of the stellar temperatures are also overestimated, since a large parallax relative uncertainty may refer to a complex transformation to determine a distance (Bailer-Jones et al. 2018). Such a transformation may bring noise to the subregressors and results in bad performances.

We do not test the regression with $T_{\text{eff}}$ outside of the training label range of 3700–9700 K because of the inability of RF to extrapolate. Andrae et al. (2018) fed their regressor with stars that have $T_{\text{eff}}$ outside the training interval, and those stars were assigned temperatures inside the training interval.

Therefore, it is suggested that all the criteria should be applied before regression in order to select good samples and further produce reliable $T_{\text{eff}}$.

4. Discussion

In this work, we have attempted to regress the effective temperatures for 132,739,323 stars in Gaia DR2 using a machine-learning algorithm. The regressor is trained with about four million stars in LAMOST, SSPP, RAVE, and APOGEE catalogs, one of the largest training samples ever used for machine learning in astrophysics. We have tried several combinations of input parameters, and have applied cross-validation to test the performances. The regressor with the smallest RMSE is built with $l$, $b$, $\Delta \text{par}$, $\mu_{\alpha}$, $\mu_{\delta}$, $\text{BP} - G$, and $G - \text{RP}$. The cross-validation indicates that the typical accuracy of the regression is 191 K. In order to examine the performance of the regressor, we use the majority of the training set to build three subregressors and apply the rest to the small fraction for blind testing. The testing results show similar performance to some spectrum-based pipelines. In this section we would like to discuss the processes that have not been used in other machine-learning studies.

4.1. Feature Selection

In machine-learning technology, feature selection is a process of selecting features in the data that are most useful or most relevant for the problem. The problem in this paper is regressing $T_{\text{eff}}$ with parameters in Gaia DR2. We adopt the RMSE to indicate the relevance of the problem for the different subsets of the input parameters.

One of the most popular parameters is a stellar color, since a stellar temperature could be roughly described by a color. However, this description suffers from temperature-extinction coupling. When we try to use Gaia colors or magnitudes to regress $T_{\text{eff}}$, the performance is bad. This implies that the color parameters are relevant to our problem, but the problem could not be fully described by these colors. The additional input parameters are required to provide information about the Galactic interstellar extinction.

Many works have been done to draw the 3D dusty map of the Milky Way (e.g., Green et al. 2018). The extinction value is a function of the stellar location. When we add $\alpha$, $\delta$, and parallax to the parameter subsets, the performance becomes better. The RMSE is slightly smaller for the regressor with the $l$ and $b$ inputs than the $\alpha$ and $\delta$ inputs, probably due to the transformation between equatorial and galactic coordinates. The algorithm needs to find this potential transformation when building the regressor with $\alpha$ and $\delta$, which may add additional noise and result in a larger RMSE. When we use $l$ and $b$ instead of $\alpha$ and $\delta$ to build the regressor, $l$ and $b$ become the most two important parameters (Figure 8). It implies that the information on Galactic extinction plays an important role in the $T_{\text{eff}}$ regression.

The proper motion can also improve the performance of the regressor, and its importance is higher than those of the Gaia colors (Figure 8), implying that its more relevant than colors in our $T_{\text{eff}}$ regressing process. The proper motion could provide assistant information on stellar distance statistically, based on the fact that the systematic errors in distance would result in the correlations between the measured $U$, $V$, and $W$ velocity components (Schönrich et al. 2012; Wang et al. 2016). This implies that when we add the proper motion to the parameter subsets, the parallax could give more information about the reliability of the stellar distance.

4.2. Blind Tests for Subregressors

The training set for the regressor is dominated by the stars in the LAMOST catalog, over 87%, and the other three catalogs comprise $\sim$4% of SSPP, $\sim$5% of RAVE, and $\sim$4% of APOGEE. We build three subregressors with a combination of three catalogs that are $\sim$95% of the training set in order to apply blind tests and further to avoid potential overfitting. Each one of the subregressors could be used to predict the $T_{\text{eff}}$ for the stars in Gaia DR2, while we use all four catalogs to train the final regressor in order to maximize the performance.

It has been proven that the performance of the regressor can be increased by adding more data to the training set.
than noise into the training set, and using four catalogs rather than three could raise the performance.

On the one hand, the systematic error of the regression is mainly from the biases among different input catalogs (Figure 2). Such an error does not decrease when we add more data to build the regressor. The residuals in Figure 6 show that the biases are constrained to $\mu < 112$ and $\sigma < 200$ for the four spectrum-based catalogs. On the other hand, performing a regression on the combination of four catalogs can be just as biased as performing the regression on three of them. In these cases, it can be reasonable to use an averaging scheme, when there is enough samples in every bin of the grid. However, the training stars are dominated by F, G, and K stars, and the sample sizes of high and low mass are not enough to smooth the fluctuation in the bins. We would take advantage of the averaging scheme to train a regressor more effectively, with the help of Gaia DR3 (next year) and LAMOST DR6 plus an early version of DR7 (more than ten million spectra in this summer).

Therefore, it is reasonable that we use the majority of the training set to build the subregressor and apply the rest to the small fraction for the blind test. This process can be applied in other machine-learning regression when there is not additional data for a blind test.

This work was supported by the National Program on Key Research and Development Project (grant No. 2016YFA0400804) and the National Natural Science Foundation of China (NSFC) through grants NSFC-11603038/11425313/11403056. This work presents results from the European Space Agency (ESA) space mission Gaia. Gaia data are being processed by the Gaia Data Processing and Analysis Consortium (DPAC). Funding for the DPAC is provided by national institutions, in particular the institutions participating in the Gaia Multilateral Agreement (MLA). The Gaia mission website is https://www.cosmos.esa.int/gaia. The Gaia archive website is https://archives.esac.esa.int/gaia.

The Guoshoujing Telescope (the Large Sky Area Multi-Object Fiber Spectroscopic Telescope, LAMOST) is a National Major Scientific Project which is built by the Chinese Academy of Sciences, funded by the National Development and Reform Commission, and operated and managed by the National Astronomical Observatories, Chinese Academy of Sciences.

Funding for the Sloan Digital Sky Survey IV has been provided by the Alfred P. Sloan Foundation, the U.S. Department of Energy Office of Science, and the Participating Institutions. SDSS-IV acknowledges support and resources from the Center for High-Performance Computing at the University of Utah. The SDSS website is http://www.sdss.org/.

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

**References**

Allende Prieto, C., Sivarani, T., Beers, T. C., et al. 2008, AJ, 136, 2070

Andrae, R., Fouesneau, M., Creevey, O., et al. 2018, A&A, 616, A8

Bai, Y., Liu, J.-F., & Wang, S. 2018, RAA, 18, 118

Bai, Y., Liu, J.-F., Wang, S., & Yang, F. 2019, AJ, 157, 9

Bailer-Jones, C. A. L., Andrae, R., Arcay, B., et al. 2013, A&A, 559, A74

Bailer-Jones, C. A. L., Rybizki, J., Fouesneau, M., Mantelet, G., & Andrae, R. 2018, AJ, 156, 58

Banko, M., & Brill, E. 2001, in Proc. 39th Annual Meeting of the Association for Computational Linguistics, ed. B. L. Webber (Stroudsburg, PA: ACL), 26

Bijaoui, A., Recio-Blanco, A., de Laverny, P., & Ordenovic, C. 2012, Stat. Methodol., 9, 55

Bretzmann, L. 2001, Mach. Learn., 45, 5

Davenport, J. R. A., Ivezic, Z., Becker, A. C., et al. 2014, MNRAS, 440, 3430

Du, B., Luo, A., Zhang, J., Wu, Y., & Wang, F. 2012, Proc. SPIE, 8451, 845137

Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2018, A&A, 616, A1

Gaia Collaboration, Prusti, T., de Bruijne, J. H. J., et al. 2016, A&A, 595, A1

Gao, H., Zhang, H.-W., Xiang, M.-S., et al. 2015, RAA, 15, 2204

García Pérez, A. E., Allende Prieto, C., Holtzman, J. A., et al. 2016, AJ, 151, 144

Green, G. M., Schlafly, E. F., Finkbeiner, D., et al. 2018, MNRAS, 478, 651

Koleva, M., Prugniel, P., Bouchard, A., & Wu, Y. 2009, A&A, 501, 1269

Kunder, A., Kordopatis, G., Steinmetz, M., et al. 2017, AJ, 153, 75

Lee, Y. S., Beers, T. C., Carlín, J. L., et al. 2015, AJ, 150, 187

Lee, Y. S., Beers, T. C., Sivarani, T., et al. 2008a, AJ, 136, 2022

Lee, Y. S., Beers, T. C., Sivarani, T., et al. 2008b, AJ, 136, 2050

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

Luri, X., Brown, A. G. A., Sarro, L. M., et al. 2018, A&A, 616, A9

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

Mathur, S., Huber, D., Batalha, N. M., et al. 2017, ApJS, 229, 30

Mészáros, S., Holtzman, J., García Pérez, A. E., et al. 2013, AJ, 146, 133

Pelisoli, I., Bell, K. J., Kepler, S. O., & Koester, D. 2019, MNRAS, 482, 3831

Plášzy, J., & Tikk, D. 2006, in Proc. Third ACM Conf. on Recommender Systems, ed. L. Bergman (New York: ACM), 93

Recio-Blanco, A., Bijaoui, A., de Laverny, P. 2006, MNRAS, 370, 141

Sahliholt, C. L., Felting, S., Lindegren, L., & Church, R. P. 2019, MNRAS, 482, 895

Schönrich, R., Binney, J., & Asplund, M. 2012, MNRAS, 420, 1281

Smolinski, J. P., Lee, Y. S., Beers, T. C., et al. 2011, AJ, 141, 89

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

Wang, S., Tang, K., & Yao, X. 2009, in Proc. Int. Joint Conf. Neural Netw. (Piscataway, NJ: IEEE), 1796

Wang, S., & Yao, X. 2009, in Proc. IEEE Symp. Comput. Intell. Data Mining, ed. Y. Saygin (Piscataway, NJ: IEEE), 324

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