The Astrophysical Journal, 920:33 (18pp), 2021 October 10 https://doi.org/10.3847/1538-4357/ac106f
© 2021. The American Astronomical Society. All rights reserved.

Metallicity Estimation of RR Lyrae Stars From Their *I*-Band Light Curves

István Dékány[1], Eva K. Grebel[1], and Grzegorz Pojmański[2]
[1] Astronomisches Rechen-Institut, Zentrum für Astronomie der Universität Heidelberg, Mönchhofstr. 12-14, 69120 Heidelberg, Germany
[2] Astronomical Observatory, University of Warsaw, Al. Ujazdowskie 4, 00-478 Warszawa, Poland
Received 2021 May 24; revised 2021 June 26; accepted 2021 June 29; published 2021 October 8

Abstract

We have revisited the problem of metallicity prediction of RR Lyrae stars from their near-infrared light curves in the Cousins *I* wave band. Our study is based on high-quality time series photometry and state-of-the-art high-resolution spectroscopic abundance measurements of 80 fundamental-mode (RRab) and 24 first-overtone (RRc) stars, spanning $\sim[-2.7, +0.18]$ dex and $\sim[-3, -0.5]$ dex ranges, respectively. Employing machine-learning methods, we investigated various light-curve representations and regression models to identify their optimal form for our objective. Accurate new empirical relations between the [Fe/H] iron abundance and the light-curve parameters have been obtained using Bayesian regression for both RRab and RRc stars with mean absolute prediction errors of 0.16 dex and 0.18 dex, respectively. We found that earlier *I*-band [Fe/H] estimates had a systematic positive bias of up to $\sim$0.4 dex in the metal-poor regime. Our new predictive models were deployed on large ensembles of RR Lyrae stars to obtain photometric metallicity distribution functions (MDFs) for various old stellar populations in and around the Milky Way. We find that the mode of the old bulge component's MDF is approximately -1.4 dex, in remarkable agreement with the latest spectroscopic result. Furthermore, we derive MDF modes of -1.83, -2.13, and -1.77 dex for the Large and Small Magellanic Clouds and the Sagittarius dwarf galaxy, respectively.

Unified Astronomy Thesaurus concepts: RR Lyrae variable stars (1410); Metallicity (1031); Light curves (918)

Supporting material: machine-readable tables

1. Introduction

RR Lyrae stars are keystone objects of Galactic archeology. These pulsating, low-mass, horizontal-branch stars belong to the oldest stellar populations of the universe (see Catelan & Smith 2015, for a review). Thanks to their relatively high luminosities and characteristic large-amplitude light curves caused by radial pulsation, they are straightforward to identify, making them excellent population tracers. Large time-domain photometric surveys have discovered them in vast numbers in the Milky Way and other members of the Local Group (e.g., Clementini et al. 2019; Soszyński et al. 2019; Dékány & Grebel 2020; Stringer et al. 2020, to mention only a few). They follow tight period–luminosity–metallicity relations at infrared wavelengths (e.g., Marconi et al. 2015; Muraveva et al. 2018), enabling us to employ them as high-precision standard candles and reddening estimators. Moreover, the existence of a precise relationship between their light-curve shapes and metal abundances discovered by Jurcsik & Kovács (1996) not only makes it feasible to determine their distances solely from photometric observations but also gives us simple means to obtain metallicity distributions of old stellar populations, opening the possibility to constrain their early formation histories (see Savino et al. 2020, and references therein).

In order to unlock the full potential of RR Lyrae stars as all-in-one photometric tracer objects, both precise and accurate methods for the prediction of their metallicities are required for all common photometric wave bands used by current and future surveys. The number of RR Lyrae stars with low-resolution spectroscopic metal abundance measurements based on hydrogen and calcium line widths as metallicity proxies (i.e., the "ΔS-method", see, e.g., Layden 1994) have recently proliferated (e.g., Liu et al. 2020; Muhie et al. 2021), mainly thanks to large surveys such as SEGUE (Yanny et al. 2009) and LEGUE (Deng et al. 2012). However, such observations are resource-intensive and currently unfeasible for faint stars in distant and/or highly reddened populations. In addition to being cheaper and much more straightforward to obtain, photometric RR Lyrae metallicities have also been shown to have a precision similar to [Fe/H] estimates obtained with the ΔS-method. Following up on the pioneering work by Jurcsik & Kovács (1996, hereafter JK96) who calibrated a simple linear formula for the estimation of the [Fe/H] from *V*-band light curves, several studies have extended their work to various other photometric wave bands (e.g., Smolec 2005; Nemec et al. 2013; Ngeow et al. 2016; Skowron et al. 2016; Hajdu et al. 2018; Iorio & Belokurov 2021).

A common shortcoming of all these formulae is that they were derived from either sub-optimal data sets or from one another, either via intermediate data sets or by simply using transformation formulae between photometric systems. This is due to the long-prevailing shortage of objects with both high-quality light curves and accurate and high-precision spectroscopic metallicity measurements. The various prediction models had to be trained on relatively few objects offering a limited coverage of the underlying parameter space, which in itself can induce biased predictions. Furthermore, the underlying metallicity measurements were dominantly obtained by low-resolution spectroscopy or spectro-photometry, often forming heterogeneous samples, which in turn had to be calibrated to even fewer high-resolution (HR) spectroscopic [Fe/H] abundance measurements, leading to the coexistence of various "metallicity scales." Finally, adapting a prediction model from one wave band to another by the transformation of the light-curve parameters has the additional pitfall that such transformations themselves can be metallicity-dependent (see, e.g., Skowron et al. 2016).

In the course of the past decade, HR spectroscopic analyses of RR Lyrae stars have been slowly accumulating thanks to the devotion of tremendous amounts of telescope time and the painstaking effort of several authors (see Crestani et al. 2021, and references therein). Recently, Mullen et al. (2021) capitalized on this progress to establish new prediction formulae of the metallicity from optical (V) and mid-infrared ($W1$ and $W2$) light curves based on a sizable data set of low-resolution spectroscopic [Fe/H] estimates, which, in turn, were tied to recent HR spectroscopic data. Their work revealed large systematics in earlier estimators, which were resulting from the various aforementioned problems.

In this work, we revisit the problem of metallicity estimation from Cousins I-band light curves of fundamental-mode RR Lyrae (RRab) stars, and, for the first time, attempt to establish a similar estimator for the first-overtone RR Lyrae (RRc) stars as well. This particular photometric wave band is of particular interest in contemporary research due to vast amounts of I-band photometric time series of RR Lyrae stars acquired by the Optical Gravitational Lensing Experiment (OGLE, Udalski et al. 2015). What makes OGLE unique among the large number of recent and ongoing photometric surveys is that it provides homogeneous, high-quality light curves for billions of objects over large areas of the observationally most challenging regions of the Local Group, namely the Galactic bulge and mid-plane, and the Small and Large Magellanic Clouds. These data were utilized by a plethora of studies, employing the RR Lyrae stars not only as distance indicators (e.g., Jacyszyn-Dobrzeniecka et al. 2017), but also as metallicity tracers (e.g., Skowron et al. 2016; Pietrukowicz et al. 2020).

Such studies either relied on the metallicity estimator from the early work of Smolec (2005, hereafter S05), or predictive formulae established for optical wave bands and transformed into the I-band. The S05 I-band formula was established on the basis of a very small data set available at the time, comprising metallicity estimates from low-resolution spectra and spectrophotometry on the JK96 scale. As an alternative, Skowron et al. (2016) employed the Nemec et al. (2013, hereafter N13) estimator to explore the metallicity distribution of the Magellanic Clouds using OGLE data by transforming it from the Kepler photometric band to the I-band and their method was further applied by Jacyszyn-Dobrzeniecka et al. (2017). The disadvantage of this approach is that it adds noise and potential systematics to the predictions due to the uncertainties and the metallicity-dependence of the transformations.

I-band photometric metallicities were also used as intermediate calibration data. Hajdu et al. (2018) trained a deep-learned [Fe/H] predictive model for the near-infrared K_s-band using I-band predictions of the [Fe/H] from the S05 formula. Their estimator was deployed on data from the VVV survey (Minniti et al. 2010) to explore the southern thick disk (Dékány et al. 2018). In view of the number of currently available HR spectroscopic [Fe/H] measurements, the recalibration of the I-band photometric estimation of RR Lyrae stars became timely. To this end, in order to minimize the possible sources of noise and systematics from intermediate data and calibrations, we take the strategy of relying solely on homogenized HR spectroscopic measurements.

This paper is structured as follows. First, we describe the photometric and spectroscopic data that form the basis of our analysis in Section 2. In addition, we present the parametric representations of the photometric time series that we use throughout the study, together with the procedures to obtain them. In Section 3, we search for the optimal form for the predictive models of the metallicity for RRab and RRc stars and estimate their performances using machine-learning techniques; then we obtain their final parameter distributions via Bayesian regression analysis and compare the resulting predictive models to the S05 estimator and discuss the nature and possible sources of the observed differences. In Section 4, we deploy our new [Fe/H] estimators on large data sets from various Local Group environments and derive the metallicity distribution functions (MDFs) of their old stellar populations. We also compare the resulting distributions with those obtained by the widely employed S05 formula. Finally, we summarize our findings and conclusions in Section 5.

2. Observational Data

We surveyed the literature for RR Lyrae stars with both high-resolution spectroscopic metallicity determinations and well sampled, accurate light curves in the Cousins I photometric passband. The largest single catalog of contemporary spectroscopic RR Lyrae metallicities was published recently by Crestani et al. (2021) and is based mostly on echelle spectra from the du Pont telescope at Las Campanas Observatory, complemented by data from various other instruments. According to Crestani et al. (2021), their measurements, together with the additional results by For et al. (2011), Chadid et al. (2017), and Sneden et al. (2017) form a homogeneous sample (from here on, referred to as the CFCS sample), covering $-3.1 \lesssim$ [Fe/H] $\lesssim 0.2$, i.e., virtually the full range of possible metallicities of RR Lyrae stars. We used this data set as our reference metallicity scale and complemented it with additional data sets from the literature that have a sufficient number of common objects with the CFCS sample in order to account for possible systematic offsets. The CFCS sample was thus complemented by additional metallicity measurements from Clementini et al. (1995, C95), Fernley & Barnes (1996, F96), Lambert et al. (1996, L96), Liu et al. (2013, L13), Nemec et al. (2013, N13), Govea et al. (2014, G14), Pancino et al. (2015, P15), and Andrievsky et al. (2018, A18). The resulting spectroscopic data set contains metallicities for 183 RRab and 49 RRc stars in the aforementioned metallicity range, but improving its coverage.

We searched the literature for existing I-band photometric time series of the RR Lyrae stars in the combined spectroscopic data set from various surveys and individual studies. Although the OGLE Collection of Variable Stars[3] (OCVS) includes tens of thousands of Galactic RR Lyrae stars, its footprint covers low Galactic latitudes toward the bulge and the southern mid-plane that spectroscopic studies tend to purposely avoid due to large interstellar extinction and point-source density. This mutual avoidance resulted in merely two objects in common with the spectroscopic sample. For several field RR Lyrae stars from our closer Galactic neighborhood, highly accurate I-band light curves were published by Monson et al. (2017). The All Sky Automated Survey (ASAS, Pojmanski 1997), although monitoring the sky in mostly the V-band also acquired well sampled light curves in the I-band for hundreds of field RR Lyrae stars; a large part of this sample was previously analyzed by Szczygieł et al. (2009). We found that 94 RRab

[3] http://ogledb.astrouw.edu.pl/~ogle/OCVS/

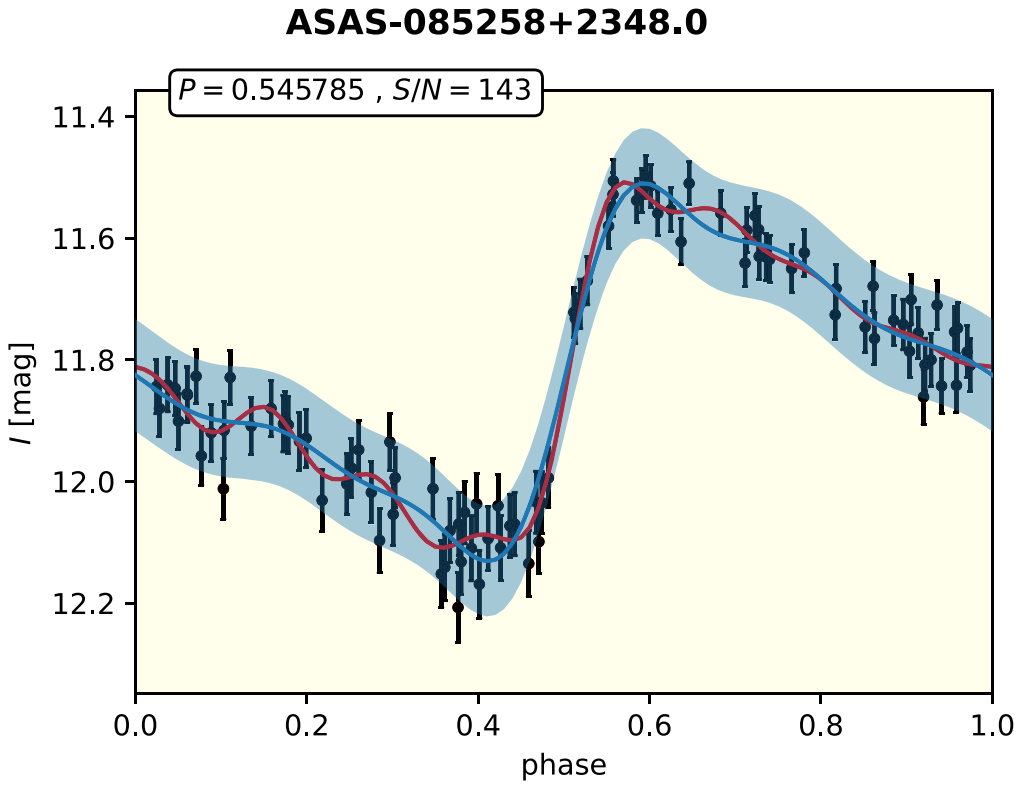

Figure 1. Phase-folded ASAS *I*-band light curve of the RRab star ASAS-085258 + 2348.0. The object's period (in days) and the S/N are shown in the inset. The individual photometric measurements and their errors are shown with black points and error bars. The red and blue curves show the best-fitting models obtained by DFF (8th order Fourier sum) and GPR, respectively. The blue shaded area shows the 95% confidence interval of the GPR model.

and 34 RRc stars from the spectroscopic data set had *I*-band light curves from the aforementioned photometric studies. Together with an additional star (RR Gem) with photometric data from Jurcsik et al. (2005), these 129 objects with combined HR spectroscopic and *I*-band time series photometric data form the basis of our study.

2.1. Light Curves and Their Representation

The optimal periods of the ASAS light curves were determined with the procedure discussed by Dékány et al. (2019) consisting of a robust, non-linear regression of a truncated Fourier series in conjunction with iterative outlier rejection and the optimal number of terms (i.e., Fourier order) determined by cross-validation. We found that the light curves from other sources already had highly accurately determined periods, therefore we directly adopted those values from the respective studies.

2.1.1. Gaussian Process Regression

Although the traditional direct Fourier fitting (DFF) provides a good generic model representation of periodic light curves, it shows considerable volatility in the case of uneven phase-sampling and/or noisy photometry, in which cases it lacks a good bias-variance trade-off at any choice of Fourier order. To obtain a more robust model representation of the light-curve shape, we employed Gaussian process regression (GPR) on the phase-folded data by the following procedure. First, we considered the $[-0.5, 1.5]$ range of the pulsation phase by repeating the measurements beyond the [0, 1] phase range in order to obtain natural periodic boundary conditions for the GPR fit. We applied GPR on the resulting data set using the product of an exponential sine-squared and a constant kernel for modeling the inherent shape of the light curves, and we added a white-kernel component to include the observational noise in the model. The optimum solutions were found by maximizing the log-marginal likelihood using the `scikit-learn` software library.

Figure 1 shows the difference between the best-fitting DFF and GPR models of an RRab star's ASAS light curve of modest quality. The GPR model can provide the complexity of a high-order Fourier sum without suffering from high variance in the absence of a large amount of high-quality data. For the latter, the difference between the results from the two methods diminishes. In addition to its robustness, another advantage of the GPR is that it yields a probabilistic model from which an arbitrary number of samples can be drawn and can then be used for computing empirical confidence intervals and error distributions of the light-curve parameters.

We transformed the fitted GPR models into the Fourier parameter space by evaluating the mean of the predictive GPR distribution over a dense equidistant phase grid and fit a 20 order Fourier sum to the resulting points. We estimated the uncertainties in the Fourier parameters by drawing 500 random samples from the GPR model, repeating the entire regression procedure for each realization and subsequently computing the standard deviations of the parameter distributions obtained in this way.

As a result of the above procedure, the light curves are represented by the parameters

$$\{P, A_i, \phi_{i1} = \phi_i - i\phi_1 \mid i = 1, 2, 3\}. \quad (1)$$

Here, P is the period, and A_i and ϕ_i are the amplitude and phase in the ith Fourier term of the following form:

$$m(t) = \sum_{i=0}^{20} A_i \sin\left(2i\pi\frac{t - t_0}{P} + \phi_i\right). \quad (2)$$

The resulting light-curve parameters and their uncertainties for our entire sample of RR Lyrae stars are listed in Table 1.

While the restriction of the above representation to the first few Fourier terms causes some potential information loss by disregarding the finest details of some RRab light curves described by higher-order parameters, in practice, such features become rather subtle for longer periods and are easily washed out by photometric noise. In addition, the use of noisy high-order parameters as regressors for metallicity prediction could easily lead to overfitting due to our modest sample size.

2.1.2. Principal Component Analysis

The traditional Fourier decomposition discussed above provides a simple and intuitive parametric representation of the RR Lyrae light curves; however, the amplitudes and phases of different orders show strong correlations, which might complicate feature selection and affect predictive performance when they are considered as descriptive variables in a regression problem such as metallicity prediction. In order to eliminate such potential effects, we performed a principal component analysis (PCA) on the Fourier parameters.

PCA is a linear transformation that decomposes a multivariate data set into a set of orthogonal components, resulting in a transformed feature space in which the individual dimensions of the data set are linearly uncorrelated and form a sequence according to the amount of explained variance, with the transformed data set showing the highest variance along the first dimension. To attain a robust transformation, the training set of the PCA should ideally fully encompass and densely populate the entire parameter space in question. We achieved this by performing PCA on the entire OGLE collection of RR Lyrae stars, i.e., the catalogs of the Galactic bulge and disk and the Large and Small Magellanic Clouds (LMC and SMC). First, all light curves in the aforementioned catalogs were fitted following the procedure in Section 2.1.1. We selected the training sets from the PCA by applying the following quality criteria to the data:

$$\text{RRab:}\quad C > 0.93;\ \mathrm{S/N} > 200;\quad N_{\text{ep.}} > 100;\ 0.1 < A_t < 1.1 \quad (3)$$

$$\text{RRc:}\quad C > 0.93;\ \mathrm{S/N} > 100;\quad N_{\text{ep.}} > 100;\ 0.1 < A_t < 0.4 \quad (4)$$

Here, C denotes the phase coverage[4], $\mathrm{S/N} = A_t\sqrt{N_{\text{ep.}}}/\sigma$, A_t is the peak-to-valley amplitude of the light-curve model, $N_{\text{ep.}}$ is the number of epochs, and σ is the residual standard deviation.

We performed PCA on the resulting ~39,000 RRab and ~21,000 RRc stars by first standardizing the features in Equation (1) to zero mean and unit variance and then performing singular value decomposition on the data matrix using the `scikit-learn` software library. Table 2 lists the ratios of the explained variance ($\hat{\sigma}_k$), singular values (s_k), and the components of the principal axes in the original feature space (i.e., right singular vectors).

[4] 1 − maximum phase lag.

As an additional, alternative representation of the light curves, we transformed the Fourier parameters into the linearly uncorrelated u_k; $k \in \{1...6\}$ feature space established by the PCA. Errors in the u_k parameters were estimated from the GPR models in the same way as for the Fourier parameters themselves (see Section 2.1.1). The u_k parameters of our RR Lyrae calibration data set, together with their uncertainties, are listed in Table 3.

2.2. Spectroscopic Data

The metallicities in our compiled spectroscopic data set are based on direct equivalent width measurements of iron absorption lines in HR spectra. For several studies, separate abundance estimates were given from both neutral (FeI) and ionized (FeII) iron lines, in which case we averaged them for each individual spectrum. We made one exception from this by using the FeII values in the case of the study by Fernley & Barnes (1996), whose FeI measurements are systematically under-abundant by ~0.1–0.15 dex, which they suspected to be the result of a bias in their temperature estimates or non-LTE effects. In case of the CFCS sample, the mean absolute difference between the FeI- and FeII-based metallicity measurements is only 0.013 dex, an order of magnitude smaller than the mean of their individual [Fe/H] error estimates provided separately for their FeI and FeII measurements. A similar, although generally smaller discrepancy between the individual uncertainties for at least one of the FeI or FeII measurements and the actual spread between them can be observed in case of the other studies as well, making the quoted [Fe/H] errors hard to interpret. We note that Lambert et al. (1996, L96), Liu et al. (2013, L13), and Andrievsky et al. (2018, A18) did not provide separate [Fe/H] measurements for FeI and FeII lines, but only combined estimates.

In addition to homogeneous [Fe/H] measurements, robust estimates of their uncertainties are also indispensable for an accurate predictive modeling due to the strong heteroskedasticity of the data. Although most of the authors provided individual error estimates, they were computed with different methods and under different assumptions. Upon close inspection of the error estimates in the various studies, we concluded that in many cases, the individual error estimates are in tension not only with the actual differences between the FeI- and FeII-based estimates but also with the standard deviations obtained from multiple measurements of the same object with the same instrument from the same authors. These are possibly due to the implicit inclusion of systematic error estimates in the quoted uncertainties. Furthermore, some authors did not provide error estimates along with their [Fe/H] measurements at all.

Due to these complications in interpreting the heterogeneous [Fe/H] uncertainties found in the literature, we opted to compute our own error estimates from the available data set itself. For this purpose, we took advantage of the fact that with only a few exceptions, the authors acquired multiple spectra for several objects using the same instrument, each yielding individual metallicity estimates. Assuming that a given author attained similar precision in the [Fe/H] determination for different objects using the same instrument, we estimated the statistical errors in the metallicities by pooled variances of the published [Fe/H] values for each instrument–author combination, using repeated [Fe/H] measurements for the same objects as individual pools. For those few cases where this approach was not possible, we adopted the uncertainties provided by the

Table 1
Names and Photometric Parameters of the RR Lyrae Stars in the Calibration Data Set

Name	Gaia DR2 source_id	ASAS ID	Type	**Blazhko**	ref.	$N_{\rm ep}$	P	$\langle I \rangle$	A_t	A_1	A_2	A_3	ϕ_{21}	ϕ_{31}	S/N
AA Aql	4224859720193721856	203815-0253.5	RRab	0	ASAS	151	0.361788	11.38	0.932	0.265(.005)	0.164(.005)	0.110(.005)	8.865(.052)	5.499(.075)	333
AA Aqr	2608819623000543744	223604-1000.9	RRab	0	ASAS	123	0.608883	12.36	0.736	0.224(.011)	0.107(.011)	0.106(.011)	8.900(.145)	5.600(.190)	121
AE Boo	1234729400256865664	144735 + 1650.7	RRc	0	M17	288	0.314887	10.26	0.234	0.120(.001)	0.009(.001)	0.008(.001)	9.641(.048)	6.597(.055)	350
AM Vir	3604450388616968576	132333-1639.9	RRab	1	ASAS	235	0.615087	10.93	0.449	0.157(.003)	0.078(.003)	0.047(.003)	9.336(.062)	6.234(.092)	257
AO Peg	1786827307055763968	212703 + 1836.0	RRab	0	ASAS	133	0.547243	12.37	0.616	0.207(.012)	0.117(.012)	0.063(.012)	8.887(.171)	5.935(.250)	98

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

Table 2
Results of the Principal Component Analysis

PCA Comp.(k)	$\hat{\sigma}_k^2$	$s_{(k)}$	P	A_1	A_2	A_3	ϕ_{21}	ϕ_{31}
				RRab				
1	0.820	438.86	0.337	−0.420	−0.417	−0.433	0.417	0.419
2	0.109	159.78	−0.683	−0.392	−0.399	−0.280	−0.288	−0.243
3	0.055	113.64	0.644	−0.143	−0.275	−0.047	−0.486	−0.501
4	0.008	43.41	−0.027	0.597	−0.077	−0.623	0.301	−0.398
5	0.004	31.69	0.055	−0.445	0.764	−0.438	−0.028	−0.154
6	0.004	30.04	−0.022	−0.309	−0.047	0.391	0.646	−0.577
				RRc				
1	0.449	239.26	−0.153	0.474	0.558	0.445	−0.253	−0.423
2	0.281	189.16	−0.603	−0.388	−0.032	−0.351	−0.506	−0.325
3	0.098	111.58	0.412	0.042	−0.118	0.114	−0.822	0.354
4	0.088	105.73	0.647	−0.083	0.082	−0.407	−0.022	−0.634
5	0.056	84.34	−0.076	0.423	0.445	−0.693	−0.018	0.369
6	0.029	61.25	0.140	−0.661	0.685	0.139	0.055	0.226

Table 3
Light-curve Parameters in the Principal Component Space of the RR Lyrae Stars in the Calibration Data Set

Name	u_1	u_2	u_3	u_4	u_5	u_6
AA Aql	−3.877(.180)	0.211(.113)	−1.525(.158)	−0.402(.155)	0.335(.174)	0.024(.114)
AA Aqr	−1.569(.410)	−0.704(.302)	0.751(.429)	−0.739(.339)	−0.530(.397)	0.198(.361)
AE Boo	−0.742(.264)	0.289(.251)	−0.153(.397)	0.265(.169)	−0.513(.260)	−0.432(.111)
AM Vir	1.680(.177)	−0.027(.123)	−0.136(.190)	0.026(.137)	0.032(.103)	0.145(.155)
AO Peg	−0.802(.554)	0.085(.320)	−0.002(.539)	−0.207(.400)	0.454(.391)	−0.772(.408)

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

authors, or in their total absence, we assumed an uncertainty of 0.1 dex.

According to Crestani et al. (2021), the CFCS sample is comprised of homogeneous metallicities, which means that they were obtained by the same method using the same input physics. However, this might not be the case for the rest of the data in our combined spectroscopic sample, therefore it was necessary to bring all the other subsamples to the CFCS scale. Assuming a simple metallicity-independent transformation between the CFCS sample and all other results from the literature, we estimated the systematic shifts between the data sets by taking the mean of the [Fe/H] differences between the individual measurements for common objects in both compared samples, weighted by the adopted uncertainties discussed earlier. The resulting shifts range from −0.26 dex to +0.16 dex, and presumably arise from the combination of systematic differences between the adopted solar metallicities as well as slightly different input physics and spectral fitting algorithms.

Table 4 lists the individual [Fe/H] values, their uncertainties, and the computed shifts (Δ) with respect to the CFCS data and the corresponding spectroscopic references for the objects that also have I-band photometric data available from the literature.

Table 4
Spectroscopic Metallicities of RR Lyrae Stars Compiled from the Literature

Name	[Fe/H]	Δ	Ref.	Instrument
AA Aql	−0.49 (.10)	0	C21	Subaru
AA Aql	−0.34 (.06)	0	C21	SALT
AA Aql	−0.32 (.04)	0	L13	Subaru
AA Aqr	−2.31 (.10)	0	C21	UVES
AE Boo	−1.62 (.17)	0	C21	HARPS

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

3. Predictive Modeling of the Metallicity

The combined photometric–spectroscopic data set discussed in Section 2 was further narrowed down by selecting objects with phase coverage of $C > 0.8$, along with $S/N > 100$ and $S/N > 70$ in the I-band light curves for RRab and RRc stars, respectively, in order to constrain the sample to stars with unbiased light-curve parameters. The remaining data were visually inspected and a few additional objects were culled where under-sampled photometry was suspected to bias their light curves. Some of the manually rejected objects had phase gaps overlapping their light-curve minima or maxima; others showed strong systematics in their light curves due to under-sampled (semi-)periodic modulation known as the Blazhko effect (see, e.g., Smolec 2016, for a review). In the case of strong amplitude/phase modulation, an incomplete coverage of the Blazhko cycle can result in a light-curve model that is significantly biased with respect to the mean light variation of the star, justifying its omission.

The resulting *development data set* comprises 379 individual spectroscopic [Fe/H] measurements of 82 RRab stars ranging from −2.6 to +0.2 dex, 22 of which have single and 60 of which have multiple independent metallicity determinations as well as 147 individual [Fe/H] values of 24 RRc stars ranging from −3 to −0.5 dex, among which 20 have multiple measurements. Importantly, there are only two RRc [Fe/H] measurements above −0.9 dex, currently posing a serious

limitation to any photometric predictive model of the metallicity for these objects. We also note that according to both the respective authors of the photometric databases that form the basis of our study, as well as our own assessment, our development set contains 24 RRab and 6 RRc stars with (adequately sampled) Blazhko modulation. They are marked in the fifth column of Table 1. For the large majority of these objects, the size of the modulation is rather small. Their possible influence on the metallicity prediction will be investigated in the following sections.

We divide the regression problem of predicting the [Fe/H] from the I-band light curves into two sub-problems: feature and model selection and final estimation of the model parameters. In the latter, given a regression-model type and an optimal set of features (i.e., descriptive variables), we estimate the final model parameters by Bayesian regression, which has two main advantages. First, it allows us to estimate credible intervals and covariances from the posterior distributions instead of mere point estimates of the regression model's parameters and second, it provides a means to properly take the heteroskedastic uncertainties in the descriptive variables into account by a probabilistic modeling of latent regressors. However, due to the costly nature of Bayesian inference, we use a more conventional, frequentist's approach for the preceding feature and model selection procedure.

We opted to use the individual metallicity measurements in the predictive modeling instead of binning the spectroscopic data into one [Fe/H] value per object in order to avoid introducing biases in the error distributions by small number statistics for objects with only a few measurements. In addition, using single, unbinned metallicity measurements provides natural sample weights for the target variable.

3.1. Feature and Model Selection

For selecting the optimal set of features, we considered both the Fourier and the PCA representations of the GPR light-curve models discussed in Section 2.1. In both scenarios, we considered linear least-squares models with either linear or second-order polynomial basis functions. We searched for the best performing regression model using the Sequential Feature Selection (SFS) algorithm implemented in the `scikit-learn` software library. The SFS is a greedy algorithm, attempting to find the global optimum of the model performance, measured by some metric, through a series of local optima. From a pool of N features, it selects a subset of M features by sequentially adding (forward-SFS) or removing (backward-SFS) the feature to/from the previous subset that leads to the next local optimum of the chosen performance metric measured by cross-validation (CV). For example, a subset of $M = 3$ features from a pool of $N = 6$ is selected in 3 iterations by forward-SFS. First, the algorithm selects one feature by evaluating the metric for all 6 possible univariate models by CV, and selects the best performing one. In the second iteration, the second feature is selected by evaluating all 5 possible bivariate models that include the feature selected in the first iteration, and choosing the best performing one. The final model is selected in the last iteration by evaluating the 4 possible multivariate models with 3 features that include the previously selected two. Likewise, backward-SFS also performs 3 iterations in the $N = 6$, $M = 3$ case, but performing more computations overall, and it may lead to different feature subsets. The best feature set is searched by running the algorithm for different values of M, evaluating each model by CV, and selecting the best performing one.

The forward-SFS algorithm was applied to both the RRab and RRc data using the mean absolute error (MAE) as the performance metric, estimated by 20 fold CV repeated 10 times, selecting up to 6 features in each case. We found that the RRab stars SV Eri and AN Ser were strong outliers in all cases, with deviations of $\sim$0.6 dex from all trial models, and were consequently removed from the development data set. Their peculiar behavior cannot be related to their photometry, since both stars have accurate, well sampled, unmodulated I-band light curves and their photometric properties in the Gaia EDR3 database do not indicate any hint of binarity or blend. The metal-poor SV Eri ([Fe/H] $= -2.22$, Crestani et al. 2021), however, is known to stand out in terms of its unusually large period increase, and is suspected to be rapidly crossing the instability strip in the final stages of its horizontal-branch evolution (Le Borgne et al. 2007). The metal-rich AN Ser on the other hand, has unusually low alpha-element abundance ([Fe/H] $= +0.05$ and [α/Fe] $= -0.2$ according to Chadid et al. 2017). AN Ser has three independent metallicity determinations, while only one [Fe/H] measurement was reported by Crestani et al. (2021) for SV Eri.

Figures 2 and 3 show the distributions of the MAE from CV for the different feature subsets selected by SFS, while Table 5 summarizes the results. For RRab stars, the two most important Fourier parameters for predicting the [Fe/H] are the period and the ϕ_{31}, while adding A_2 as a third feature consistently improves the model's performance, independently from the random realization of the CV splitting. However, adding more than three features by SFS, the predictive performance measured by MAE does not improve beyond its uncertainty. These findings are fully consistent with the earlier analysis by Smolec (2005). Using quadratic basis functions, a virtually identical performance can be achieved by using the $\{P, \phi_{31}, PA_2\}$ feature subset, with no significant further improvement by adding more features. Using the linearly uncorrelated u_i parameters of the principal component feature space leads to similar results. The combination of u_3 and u_2 in a simple linear model provides good predictive performance, which is marginally improved by adding u_5, while the rest of the features proved unimportant in predicting the metallicity. However, there is no clear benefit in using linearly uncorrelated features instead of the standard Fourier parameters and no improvement can be attained by using a second-order polynomial basis either. We concluded that for the RRab stars, based on the currently available data, a three-parameter linear model using the classical $\{P, \phi_{31}, A_2\}$ feature set is optimal for the prediction of the metallicity.

For RRc stars, there is a gradually diminishing performance increase by adding up to five Fourier parameters as features, including the amplitudes of the first three Fourier terms. However, the amplitude of the third term can become hard to accurately measure for some RRc stars in the absence of precision photometry. For this reason, and due to the modest improvement of the MAE by including A_3, we restrict our linear regression model to the $\{P, A_1, A_2, \phi_{31}\}$ feature set. In the case of including second-order polynomial features, SFS leads to a similar optimum by selecting $\{PA_2, A_2\phi_{21}, \phi_{21}^2, A_1\phi_{31}\}$ and adding more terms slightly further improves the performance. Similar results can be attained with the $\{u_4, u_6, u_1\}$ feature subset, without significant further improvement by increasing the

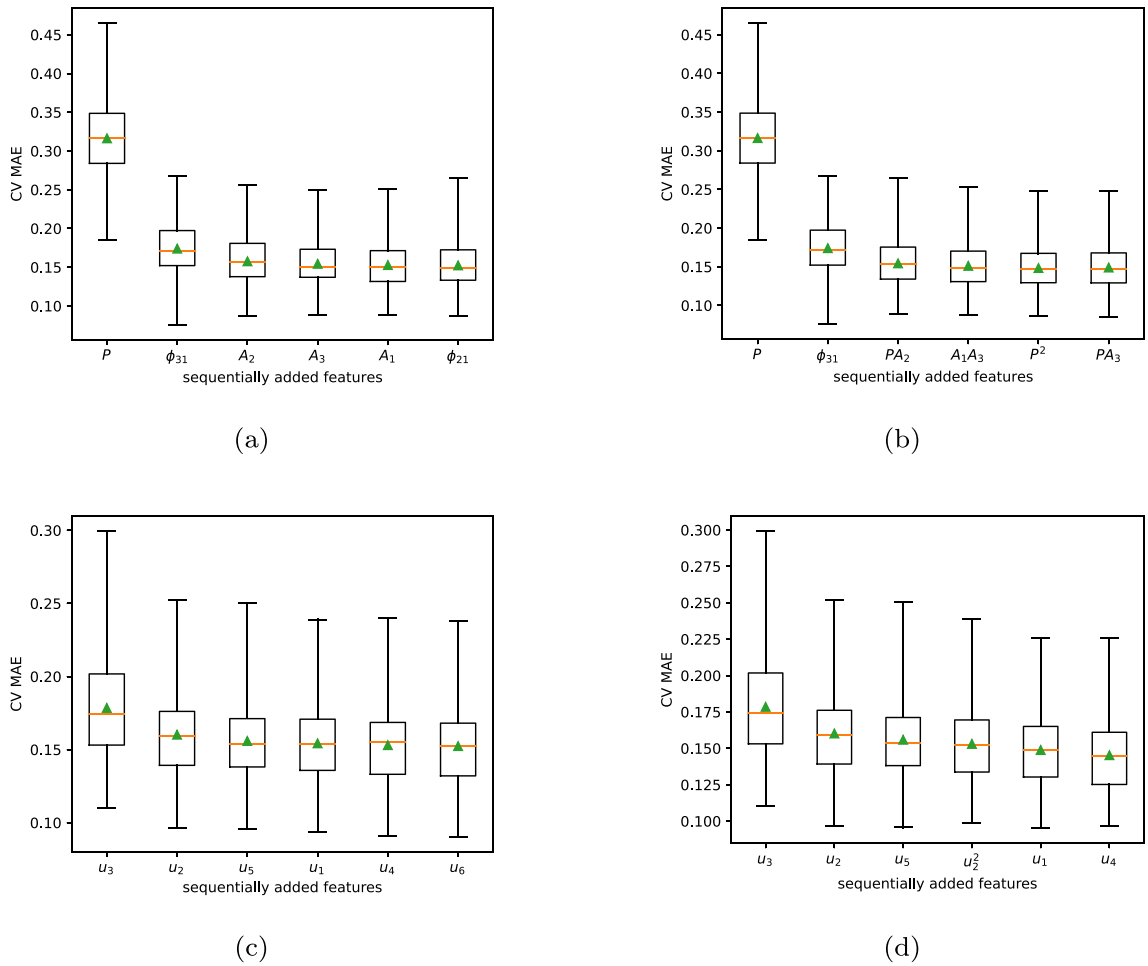

Figure 2. Boxplots of the distributions of the MAE measured by CV for different feature subsets for the RRab stars. The labels on the horizontal axes show the sequentially added features selected to be optimal by the SFS algorithm, while increasing the number of used features from 1 to 6 (left to right in each panel). The rectangular boxes encompass the interquartile ranges of the distributions with lines at the medians, with horizontal orange lines at the medians and green triangles at the means. The whiskers indicate the full ranges of the distributions. Panels (a) and (b) show the results from linear least-squares models using linear and quadratic Fourier parameters as their feature space, respectively. Panels (c) and (d) show the same but for the feature spaces transformed by PCA (see Section 2.1.2).

number of features. Using a quadratic formula, however, the mean MAE shows a gradual decrease down to ∼0.13 dex with five features. However, since such a complicated formula is highly error-prone in the case of noisy light curves and because the selection of features 3–5 is fairly sensitive to the random initialization of the CV in the SFS, we constrain our further analysis to the linear Fourier features $\{P, A_1, A_2, \phi_{31}\}$.

In order to probe whether the Blazhko effect systematically affects the feature selection, we also carried out the same analysis after the exclusion of all objects that showed signs of modulation in their light curves, resulting in modified development sets of 245 and 116 individual metallicity measurements of 56 RRab and 18 RRc stars, respectively. In spite of the significant decrease in the amount of data, this modification did not cause any significant change with respect to the original analysis, as far as the selection of linear features is concerned. In the case of quadratic features, however, SFS converges to different local optima for $M > 2$, while becoming more volatile to the random initialization of the algorithm. This, however, does not cause any significant change in the relative performances of the models and is more likely a direct effect of the reduced data set sizes (particularly for the RRc stars), rather than due to the exclusion of the Blazhko stars.

From the above analysis, we draw as a conclusion that based on the currently available data, linear models using the Fourier parameters as regressors yield satisfactory performance in predicting the metallicity for both RRab and RRc stars. The possibility of improvement by using more subtle features and/or more complicated models is not ruled out, but is not supported by the data, and should be further explored once significantly more observational data become available. We select $\{P, \phi_{31}, PA_2\}$ and $\{P, A_1, A_2, \phi_{31}\}$ as optimal feature subsets for RRab and RRc stars, respectively. In the following Section, we determine the parameters of the corresponding linear predictive models of the [Fe/H] by Bayesian regression.

3.2. Bayesian Regression of the Metallicity

In the Bayesian approach to regression problems, parameters of a predictive model are considered to be random variables and their optimal (i.e., most probable) values are inferred from posterior densities conditional on observed data. An important advantage of this approach over frequentist methods is that uncertainties in the model parameters can be expressed as credible intervals and their posterior distributions can be used for making further inferences, e.g., to analyze the propagation of error in the model's predictions. Another key benefit of

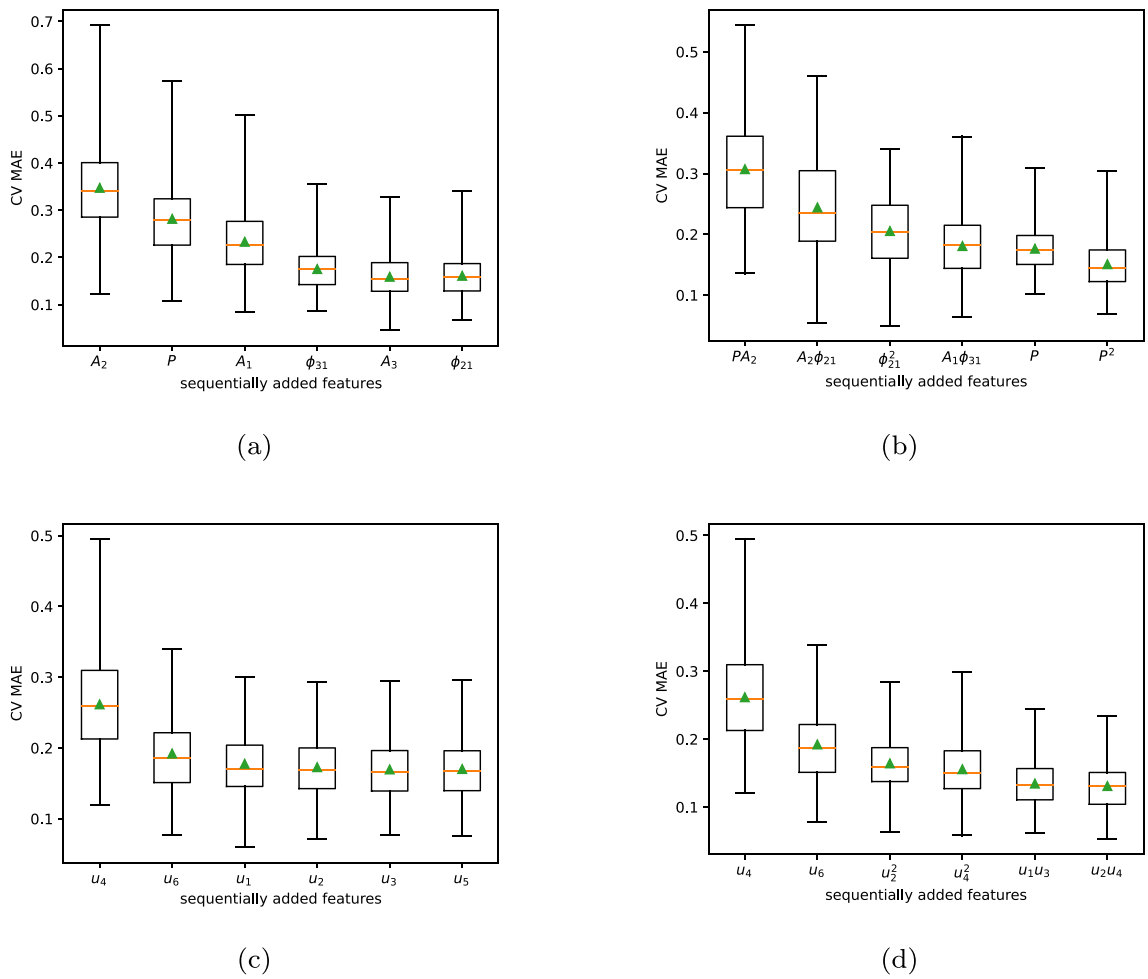

Figure 3. The same as Figure 2 but for RRc stars.

Bayesian regression lies in the possibility of taking the errors in the covariate features explicitly into account. In frequentist approaches, the descriptive variables are considered to be fixed values, which may lead to a biased regression model, e.g., in case they stem from diverse measurements or other processes that lead to heteroskedastic uncertainties. The latter problem clearly affects the regression of the metallicity to light-curve parameters since the uncertainties in the latter have a large spread as a result of the heterogeneous quality of the photometric time series acquired by different instruments under diverse conditions for objects with different apparent magnitudes.

The uncertainties in the covariates are introduced into our probabilistic model via latent regressors ξ_i:

$$\begin{aligned} y_i &= f(\boldsymbol{\xi}_i, \boldsymbol{\theta}) + \epsilon_i \\ \boldsymbol{\xi}_i &= \boldsymbol{x}_i + \boldsymbol{\nu}_i \end{aligned} \tag{5}$$

where (x_i, y_i) are the observed covariates and response data of the ith example, respectively; f is the functional form of the regression model with parameters $\boldsymbol{\theta}$; $\boldsymbol{\xi}_i$ are the latent (i.e., the actual, unobservable) covariate features, and the ϵ_i and $\boldsymbol{\nu}_i$ terms are random variables representing the stochasticity in the model. In our specific regression problem, $\boldsymbol{x}_i$ is the vector of light-curve features (i.e., the $\{P, \phi_j, A_j\}$ or $\{u_j\}$ parameters), while y_i is (are) the spectroscopic [Fe/H] measurement(s) of the ith RR Lyrae star in the development data set.

Inferences about the $\boldsymbol{\theta}$ and $\boldsymbol{\xi}_i$ random variables, conditional on (x_i, y_i), are drawn using Bayes' theorem:

$$P(\boldsymbol{\theta}, \boldsymbol{\xi}|\boldsymbol{y}, \boldsymbol{x}) \propto P(\boldsymbol{y}, \boldsymbol{x}|\boldsymbol{\xi}, \boldsymbol{\theta}) \cdot P(\boldsymbol{\theta}, \boldsymbol{\xi}|\boldsymbol{x}) \tag{6}$$

Here, the first term on the right side of the equation is the joint likelihood of the observed data given the model parameters and the latent regressors, and the second term is the prior probability distribution of the latter, given the observed covariates. Their product is proportional to the joint posterior density of $\boldsymbol{\theta}$ and $\boldsymbol{\xi}_i$. We assume that ϵ_i and $\boldsymbol{\nu}_i$ follow zero mean normal distributions with variances of the former made equal to the squared error estimates of the [Fe/H] measurements (see Section 2.2). Likewise, we use informative priors on $\boldsymbol{\xi}_i$ by setting the variances of $\boldsymbol{\nu}_i$ to be equal to the error estimates of the light-curve parameters obtained from the GPR models, as discussed in Section 2.1.1. For the elements of $\boldsymbol{\theta}$, we pose weakly informative priors assuming normal distributions with means obtained by ordinary linear least-squares regression.

In view of our conclusions from Section 3.1, we implemented the above probabilistic regression model using the `pymc3` software library for the two types of RR Lyrae stars in the

Table 5
Performances of Various Regression Models and Feature Sets for Predicting the [Fe/H] of RR Lyrae Stars

		RRab			RRc		
Model	M	⟨MAE⟩	σ(MAE)	Features	⟨MAE⟩	σ(MAE)	Features
Fourier(1)	1	0.316	0.050	$\{P\}$	0.348	0.088	$\{A_2\}$
	2	0.174	0.032	$\{P, \phi_{31}\}$	0.282	0.086	$\{A_2, P\}$
	3	0.158	0.031	$\{P, \phi_{31}, A_2\}$	0.226	0.068	$\{A_2, P, A_1\}$
	4	0.154	0.030	$\{P, \phi_{31}, A_2, A_3\}$	0.182	0.039	$\{A_2, P, A_1, \phi_{31}\}$
	5	0.153	0.030	$\{P, \phi_{31}, A_2, A_3, A_1\}$	0.170	0.041	$\{A_2, P, A_1, \phi_{31}, A_3\}$
	6	0.152	0.031	$\{P, \phi_{31}, A_2, A_3, A_1, \phi_{21}\}$	0.171	0.041	$\{A_2, P, A_1, \phi_{31}, A_3, \phi_{21}\}$
Fourier(2)	1	0.316	0.050	$\{P\}$	0.307	0.080	$\{PA_2\}$
	2	0.174	0.032	$\{P, \phi_{31}\}$	0.244	0.083	$\{PA_2, A_2\phi_{21}\}$
	3	0.154	0.030	$\{P, \phi_{31}, PA_2, \}$	0.205	0.060	$\{PA_2, A_2\phi_{21}, \phi_{21}^2\}$
	4	0.151	0.029	$\{P, \phi_{31}, PA_2, A_1A_3, \}$	0.181	0.051	$\{PA_2, A_2\phi_{21}, \phi_{21}^2, A_1\phi_{31}\}$
	5	0.148	0.029	$\{P, \phi_{31}, PA_2, A_1A_3, P^2\}$	0.176	0.037	$\{PA_2, A_2\phi_{21}, \phi_{21}^2, A_1\phi_{31}, P\}$
	6	0.149	0.029	$\{P, \phi_{31}, PA_2, A_1A_3, P^2, PA_3\}$	0.151	0.041	$\{PA_2, A_2\phi_{21}, \phi_{21}^2, A_1\phi_{31}, P, P^2\}$
PCA(1)	1	0.179	0.036	$\{u_3\}$	0.262	0.072	$\{u_4\}$
	2	0.160	0.029	$\{u_3, u_2\}$	0.192	0.053	$\{u_4, u_6\}$
	3	0.156	0.029	$\{u_3, u_2, u_5\}$	0.178	0.044	$\{u_4, u_6, u_1\}$
	4	0.154	0.027	$\{u_3, u_2, u_5, u_1\}$	0.173	0.041	$\{u_4, u_6, u_1, u_2\}$
	5	0.153	0.027	$\{u_3, u_2, u_5, u_1, u_4\}$	0.170	0.043	$\{u_4, u_6, u_1, u_2, u_3\}$
	6	0.152	0.027	$\{u_3, u_2, u_5, u_1, u_4, u_6\}$	0.170	0.043	$\{u_4, u_6, u_1, u_2, u_3, u_5\}$
PCA(2)	1	0.179	0.036	$\{u_3\}$	0.262	0.072	$\{u_4\}$
	2	0.160	0.029	$\{u_3, u_2\}$	0.192	0.053	$\{u_4, u_6\}$
	3	0.156	0.029	$\{u_3, u_2, u_5\}$	0.164	0.041	$\{u_4, u_6, u_2^2\}$
	4	0.153	0.028	$\{u_3, u_2, u_5, u_2^2\}$	0.156	0.042	$\{u_4, u_6, u_2^2, u_4^2\}$
	5	0.149	0.026	$\{u_3, u_2, u_5, u_2^2, u_1\}$	0.134	0.035	$\{u_4, u_6, u_2^2, u_4^2, u_1u_3\}$
	6	0.145	0.026	$\{u_3, u_2, u_5, u_2^2, u_1, u_4\}$	0.131	0.035	$\{u_4, u_6, u_2^2, u_4^2, u_1u_3, u_2u_4\}$

following forms:

$$\text{RRab:}\quad [\text{Fe/H}] = \theta_P P + \theta_{\phi_{31}}\phi_{31} + \theta_{A_2}A_2 + \theta_0 \tag{7}$$

$$\text{RRc:}\quad [\text{Fe/H}] = \theta_P P + \theta_{\phi_{31}}\phi_{31} + \theta_{A_1}A_1 + \theta_{A_2}A_2 + \theta_0 \tag{8}$$

Posterior distributions were computed numerically via Hamiltonian Markov chain Monte Carlo simulations using the No-U-Turn Sampler (NUTS) of Hoffman & Gelman (2014). We performed the sampling in multiple parallel Markov chains and assessed convergence numerically through the $\hat{R}$ statistic and visually using trace and posterior density plots using the `ArviZ` software package.

Figures 4 and 5 show marginal and paired kernel density estimates of the resulting posterior distributions for the components of $\boldsymbol{\theta}$ for RRab and RRc stars, respectively. Credible intervals were determined by computing highest density intervals (HDIs) of the marginal posteriors for 90% and 95% probabilities. Table 6 summarizes the statistics of the posterior distributions of our [Fe/H] regression models' parameters, while the covariance and Pearson correlation matrices between these parameters are shown in Tables 7 and 8, respectively.

Figure 6 shows the predicted [Fe/H] values against their HR spectroscopic equivalents for the RR Lyrae stars in our development set using the parameter distributions' means from Table 6. The residual standard deviations of the fits are 0.20 dex for RRab and 0.21 dex for RRc stars. The scatter of the RRab predictions is slightly higher than that of the S05 formula (0.14) but the latter was trained on a tiny data set of 28 objects with single spectroscopic measurements, which alone can result in a downward bias of the dispersion.

We investigated how the inclusion of Blazhko-modulated stars in the development set might influence our predictive models. In Figure 6, data corresponding to stars with the Blazhko effect are highlighted with red crosses. No systematic deviation or increased scatter of these data with respect to the rest of the samples can be seen. We also performed Bayesian regression on the RRab development set after excluding the data of all objects that show the Blazhko effect in their I light curves. In the right panel of Figure 6, we compare the [Fe/H] values predicted by the modified model versus those by the original one for the *entire* development set with the Blazhko stars highlighted. There is no significant difference between the predictions of the two models; thus we conclude that the photometric metallicity prediction of RRab stars is robust against the presence of the Blazhko effect in the light curves, in case the modulation cycle is well sampled. Finally, we see no point in performing a separate test for RRc stars since the excessively reduced sample size of only 18 non-modulated objects would have a large negative effect on the resulting model's performance.

Concerning the predictive model's application on Blazhko RR Lyrae stars, we emphasize that in the case of large phase and/or amplitude modulation, a partial photometric coverage of the Blazhko cycle can obviously lead to rather biased light-curve parameters with respect to those of the mean light variation of the same star and may consequently result in a biased [Fe/H] prediction. To demonstrate the above effect, Figure 7 shows the light curve of a very strongly modulated

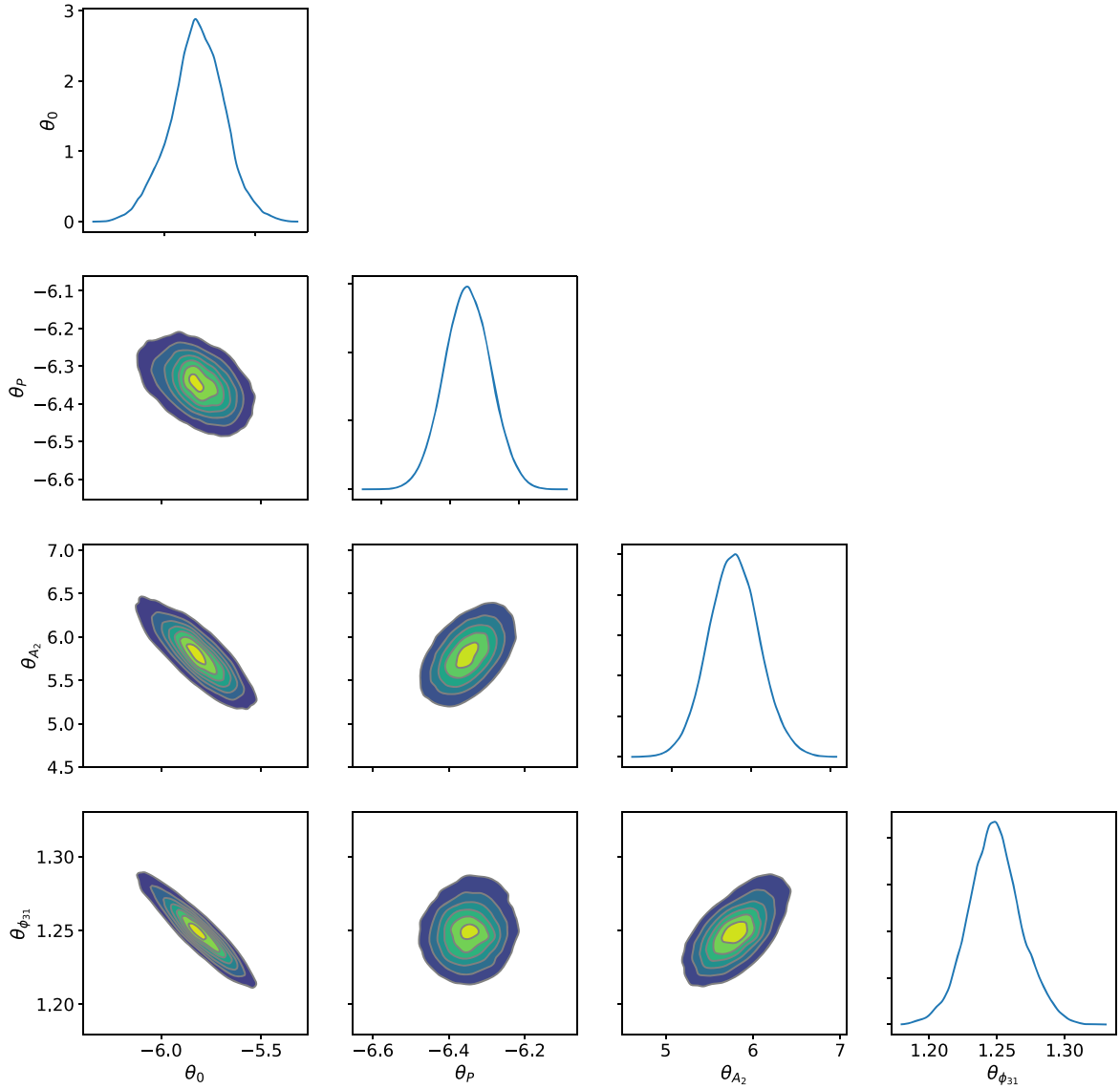

Figure 4. Posterior distributions of the parameters of the metallicity prediction model for RRab stars. Kernel density estimates of marginal posterior MCMC samples are shown for single parameters and all possible parameter pairs.

bulge RRab star of a rather rare morphological type with a very long modulation cycle of 233.65 days (see Skarka et al. 2020). Observations with a limited time baseline would severely under-sample the light curves of similar objects, which we illustrate with three ∼11.5 day long light-curve segments at different modulation phases, marked with different colors in Figure 7. The light-curve solutions of these various segments yield biased [Fe/H] predictions with respect to the value derived from the mean light-curve model based on all data, with errors of up to 0.4 dex. This caveat should be considered when photometrically estimating the [Fe/H] for individual objects or small samples.

3.3. Comparison with Earlier [Fe/H]-estimators

To visualize the differences between our Bayesian predictive model and other *I*-band photometric [Fe/H] formulae from the literature, we confronted the metallicity predictions from these with our development data set of RRab stars in Figure 8. From left to right, we plotted our predictions against those from the S05 formula, as well as the JK96 and N13 formulae transformed into the *I*-band, following the method of Skowron et al. (2016). We refer to these formulae as S16J and S16N, respectively. The predictions from the S05 and S16J formulae are on the Jurcsik & Kovács (1996) metallicity scale, which was established on the basis of ΔS measurements of field RR Lyrae stars (e.g., Layden 1994; Suntzeff et al. 1994), while the S16N formula was directly tied to spectroscopic [Fe/H] measurements of 26 RRab stars. We can observe a significant upward bias of the [Fe/H] values computed from the S05 and JK96 formulae with respect to our predictions. The deviations from our model's output show a dependence on the [Fe/H], with a systematically increasing positive bias toward lower metallicities, amounting to as much as ∼0.5 dex for [Fe/H] $\lesssim -2$. While the J16N predictions are in a better agreement at low metallicities, they too show a positive bias with respect to our values over the entire [Fe/H] range, in addition to a large scatter.

Similar discrepancies were found by Mullen et al. (2021) when comparing [Fe/H] predictions from their *V*-band regression model (trained on CFCS metallicities) to those

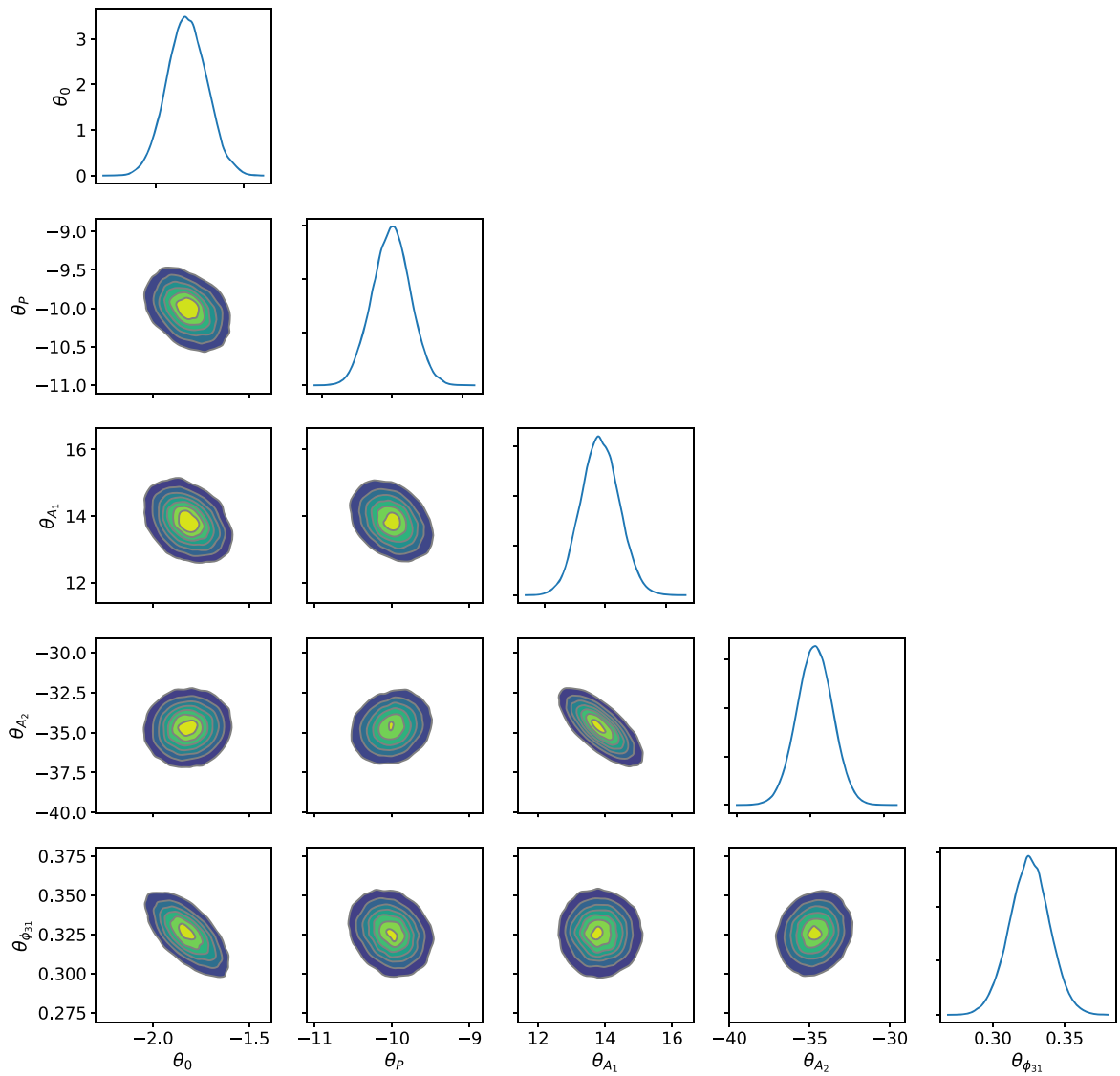

Figure 5. The same as Figure 4 but for RRc stars.

Table 6
Posterior Distribution Statistics of the [Fe/H] Regression Models' Parameters

Type	Parameter	Mean	σ[a]	HDI 2.5%	HDI 5%	HDI 95%	HDI 97.5%
RRab	θ_0	−5.819	0.149	−6.124	−6.079	−5.582	−5.524
	θ_P	−6.350	0.067	−6.484	−6.461	−6.240	−6.220
	θ_{A_2}	5.785	0.320	5.158	5.253	6.308	6.423
	$\theta_{\phi_{31}}$	1.248	0.020	1.211	1.216	1.282	1.290
RRc	θ_0	−1.821	0.113	−2.049	−2.008	−1.636	−1.604
	θ_P	−10.014	0.269	−10.546	−10.469	−9.578	−9.493
	θ_{A_1}	13.835	0.623	12.645	12.799	14.838	15.071
	θ_{A_2}	−34.704	1.211	−37.090	−36.704	−32.721	−32.340
	$\theta_{\phi_{31}}$	0.325	0.014	0.298	0.302	0.348	0.353

Note.
[a] Standard deviation of the marginal posterior.

obtained from the original JK96 and N13 formulae. To directly investigate the source of the deviating trend between our predictions and the S05 and S16J ones, we cross-matched our HR spectroscopic sample to their calibration data sets. Figure 9 shows individual HR spectroscopic metallicity measurements for the 19 stars in common with the S05 calibration data set plotted against their low-resolution counterparts on the JK96 scale. We can observe a similar positive bias of the S05 calibration data with respect to HR spectroscopy. The same [Fe/H] measurements (with only one exception) were also part

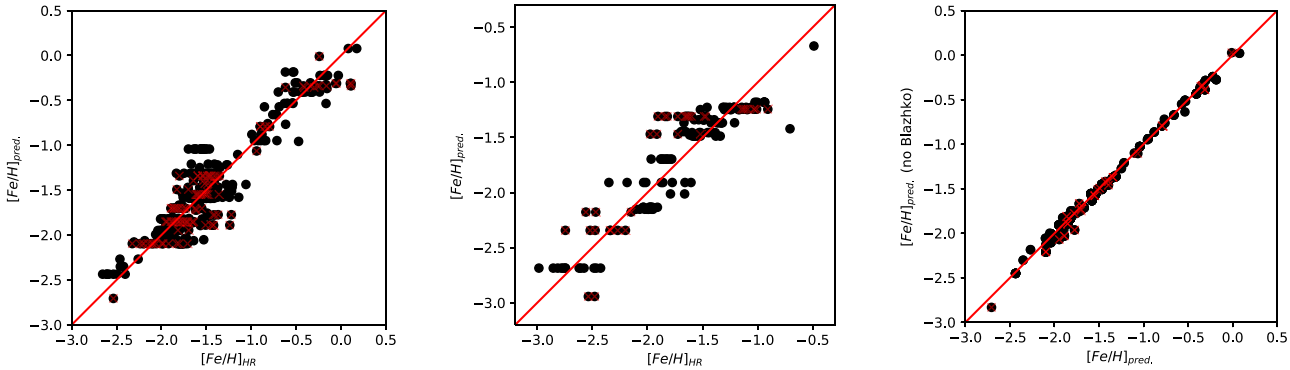

Figure 6. Predicted vs. true metallicities for the RRab (left) and RRc (right) stars in the development set. Black points denote the entire development set; red crosses mark data of only those stars that show the Blazhko effect. Note the different scales of the figures. Right: comparison of RRab metallicity predictions by our Bayesian regression models trained on data sets including/excluding Blazhko-modulated stars.

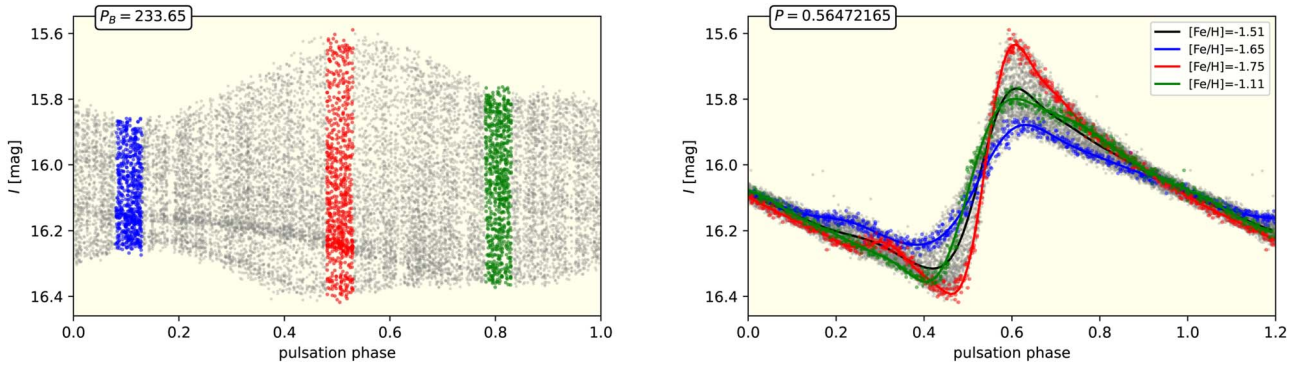

Figure 7. OGLE-IV *I*-band photometric time series of the bulge RRab star OGLE-BLG-RRLYR-06473 as a function of the modulation phase (left) and the pulsation phase (right). Gray points denote all data; the blue, red, and green points show different segments of the data in equally long modulation phase intervals of 0.05. The black and colored curves show GPR fits to the entire light curve and its various segments plotted with the corresponding colors, respectively. The inset of the right panel shows [Fe/H] values derived from the Fourier representations of the various GPR fits using our Bayesian predictive model. The left and right figure headers show the modulation period (Skarka et al. 2020) and the pulsation period in days, respectively.

Table 7
Covariance and Pearson Correlation Matrices of the RRab [Fe/H] Prediction Model's Parameters Computed from Their Joint Posterior Distribution

	Covariance Matrix				Pearson Correlation Matrix			
	θ_0	θ_P	θ_{A_2}	$\theta_{\phi_{31}}$	θ_0	θ_P	θ_{A_2}	$\theta_{\phi_{31}}$
θ_0	0.0222	−0.0042	−0.0394	−0.0027	1.0000	−0.4227	−0.8240	−0.9287
θ_P	−0.0042	0.0045	0.0105	0.0001	−0.4227	1.0000	0.4858	0.0952
θ_{A_2}	−0.0394	0.0105	0.1026	0.0039	−0.8240	0.4858	1.0000	0.6246
$\theta_{\phi_{31}}$	−0.0027	0.0001	0.0039	0.0004	−0.9287	0.0952	0.6246	1.0000

Table 8
The Same as Table 7, but for RRc Stars

	Covariance Matrix					Pearson Correlation Matrix				
	θ_0	θ_P	θ_{A_1}	θ_{A_2}	$\theta_{\phi_{21}}$	θ_0	θ_P	θ_{A_1}	θ_{A_2}	$\theta_{\phi_{21}}$
θ_0	0.0129	−0.0109	−0.0221	0.0063	−0.0010	1.0000	−0.3556	−0.3123	0.0461	−0.6558
θ_P	−0.0109	0.0725	−0.0551	0.0574	−0.0007	−0.3556	1.0000	−0.3281	0.1759	−0.1989
θ_{A_1}	−0.0221	−0.0551	0.3881	−0.5421	−0.0003	−0.3123	−0.3281	1.0000	−0.7186	−0.0355
θ_{A_2}	0.0063	0.0574	−0.5421	1.4663	0.0027	0.0461	0.1759	−0.7186	1.0000	0.1588
$\theta_{\phi_{31}}$	−0.0010	−0.0007	−0.0003	0.0027	0.0002	−0.6558	−0.1989	−0.0355	0.1588	1.0000

of the data set used by JK96 for the calibration of their *V*-band [Fe/H] prediction formula. In Figure 8 we also compare the [Fe/H] values from their sample with more recent HR spectroscopic measurements (all of these are from the CFCS data set). The same trend between the two is apparent, leading us to the conclusion that the systematically lower photometric metallicity estimates from our work with respect to the S05 and S16J formulae stem mainly from a linear bias in the JK96 scale.

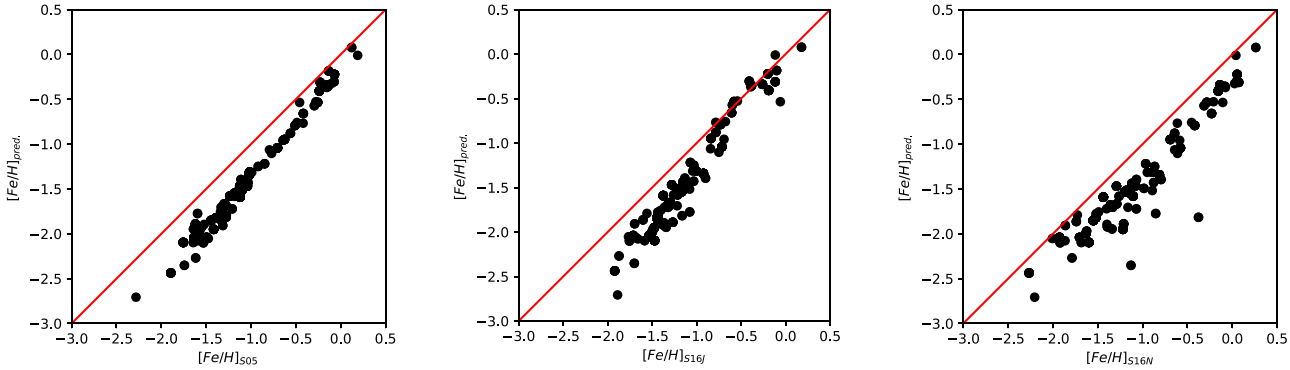

Figure 8. Metallicities of the RRab stars in the development set predicted by our regression model, in comparison with their equivalents obtained by the S05 three-parameter formula (left), the JK96 formula (middle), and the N13 formula (right). The latter two were transformed into the *I*-band following the method of Skowron et al. (2016).

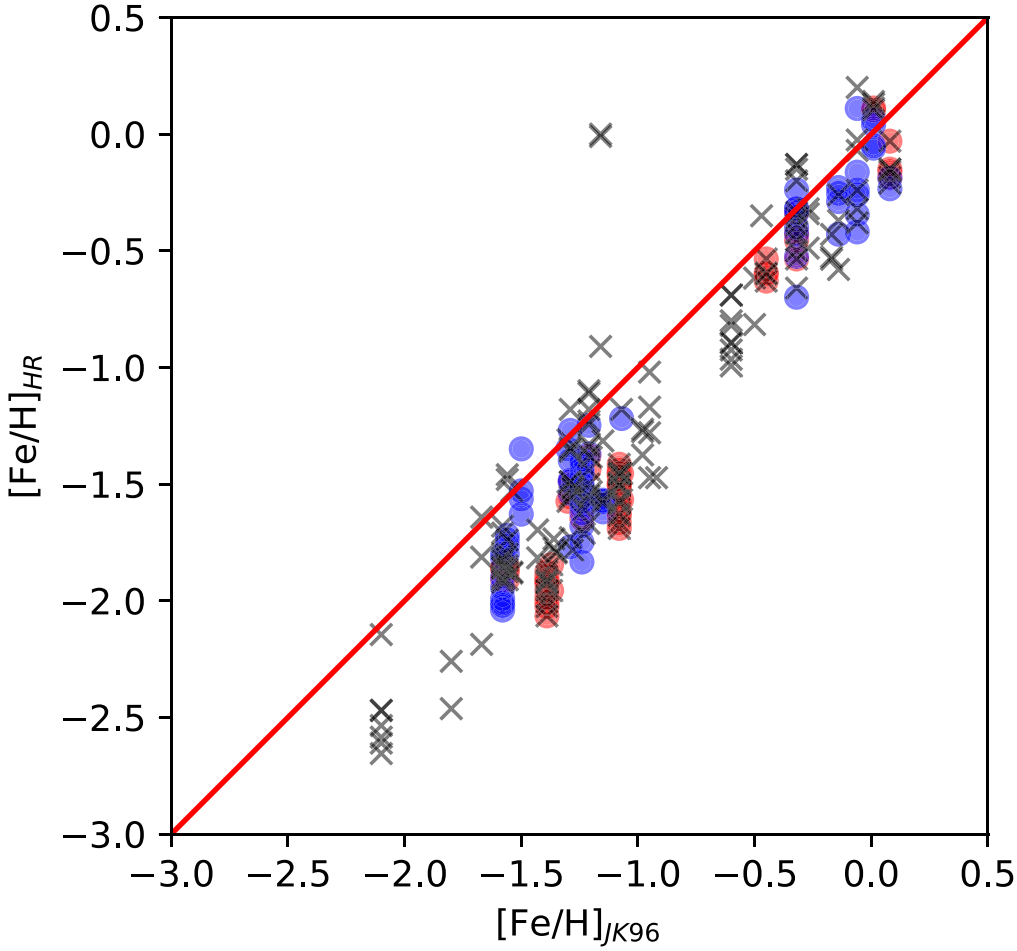

Figure 9. Individual [Fe/H] measurements from our HR spectroscopic data set plotted against their counterparts on the JK96 scale obtained from low-resolution spectra. Circles: objects used for calibrating the *I*-band S05 formula, which have HR spectroscopic [Fe/H] measurements either from the CFCS data set (red) or from other sources (blue, after being shifted to the CFCS scale). Black crosses: objects in the "basic data set" of JK96.

The latter may arise from the very modest amount of HR spectroscopic RR Lyrae [Fe/H] determinations at that time to which they tied the ΔS measurements. Apart from this, a further cause of the discrepancy between the photometric metallicities may be biased regression parameters due to the small underlying data sets deficient of very metal-poor stars particularly in case of the N13 and S05 regressions. The former includes non-linear terms that apparently do not generalize well, the latter is constrained to $[\mathrm{Fe/H}] \gtrsim -1.6$.

4. Photometric Metallicity Distributions

The systematic, metallicity-dependent biases in S05 and S16 estimators have various implications. Perhaps the most important one is their effect on the photometric MDFs derived for various stellar populations in and around the Milky Way, while capitalizing on the vast amount of accurate *I*-band time series data from the OGLE surveys of the bulge, southern disk, and the Magellanic Clouds. Most recently, Pietrukowicz et al. (2020) exploited the entire OGLE RR Lyrae catalog to study

the MDFs of the old stellar populations of the Galaxy, the LMC, SMC, and the Sagittarius dwarf galaxy, and drew inferences on their formation histories therefrom. To assess the effect of the systematic difference between the various earlier formulae and our new regression model on our view of the MDFs of these populations, we deployed all of them on the *I*-band light curves of RR Lyrae stars in the OCVS and computed the MDFs of the same cosmic environments. We used the same simple positional and brightness criteria in selecting the pertinent stellar populations as the ones that were applied by Pietrukowicz et al. (2020), which are sufficient to keep sample contamination on the level of a few percent, and aid direct comparability and reproducibility. With the additional quality criteria of $\{C > 0.9$ and $\mathrm{S/N} > 100$ and $N_{\mathrm{ep}} > 100\}$, we obtained RR Lyrae samples for the inner and outer bulge, bulge-disk transition area, southern disk, Sagittarius dwarf spheroidal galaxy, LMC, SMC, and Galactic halo.

We computed the light-curve features of all OCVS RR Lyrae stars in the above samples by applying the same light-curve fitting procedures as to the development data set, as discussed in 2.1, and used them as input of the predictive models. The resulting photometric MDFs of the various stellar populations are shown in Figure 10. The most obvious difference between the RRab MDFs from the earlier formulae and this work is the significant offset of their modes by 0.3–0.4 dex. As opposed to earlier estimates of ∼–1 dex for the bulge MDF's mode, in fact it is located at −1.37 dex and −1.38 dex for the inner and outer bulge samples, respectively. These latter values show a remarkable agreement with the results of Savino et al. (2020), who determined a median of −1.39 dex for their bulge RR Lyrae MDF obtained from medium-resolution BRAVA-RR spectra (Kunder et al. 2016). The shape of their bulge MDF also shows phenomenological consistency with our results (*cf.* their Figure 3). The bulge MDF's mode seems to shift slightly toward lower metallicity with increasing Galactocentric distance, reaching −1.42 for the transition region sample, albeit this is based on an order of magnitude smaller sample than the bulge values.

As for the shape of the MDFs, the ones derived by the S05 formula show the closest resemblance to our new solutions. The inner bulge distribution shows weak excesses around the main mode, similarly to the S05 version of the MDF, while the S16 formulae predict a much stronger excess at the metal-rich side. Their relative sizes of the distribution's tails increase toward the outer bulge, also apparent (although shifted) in the S05 and S16 MDFs, but they show significantly less spread in the S05 and S16 versions. The latter can be attributed to the metallicity-varying upward bias of the older formula having a shrinking effect on the MDFs' tails. This effect is counter-balanced in the S16 formula by the increased noise due to the applied wave band transformation and the non-linear terms in the case of S16N.

The metal-rich excess evolves into a secondary mode of the MDF as we increase the sample's Galactocentric angular distance along the southern disk. At the same time, the main mode becomes even more metal-poor, converging to the mode of the halo MDF at approximately −1.7 dex. We note that the halo MDF shown in Figure 10 is based on less than a thousand RRab stars toward the Magellanic Cloud footprint of the OGLE survey which were separated from endemic LMC and SMC objects based on their apparent magnitudes following the criteria of Pietrukowicz et al. (2020). The mode of the halo MDF is in good agreement with the approximately −1.6 dex median metallicity of the halo population derived by Crestani et al. (2021) from a sample of 211 RR Lyrae stars with HR spectroscopy. Furthermore, our photometric halo MDF has a strong tail toward extremely low metallicities, which is seen greatly diminished in the MDFs based on the S05 and S16 estimators.

The Sagittarius dwarf spheroidal galaxy shows a Gaussian-like [Fe/H] distribution and it closely matches the mode of the halo's MDF. The [Fe/H] distribution of the LMC RRab stars is also close to normal, with its mode at −1.83 dex. The SMC is significantly more metal-poor than the LMC with a [Fe/H] mode at −2.13 dex, and its MDF is slightly skewed toward low metallicities. The derived modes are significantly more metal-poor than earlier photometric Fourier-decomposition-based RR Lyrae metallicity estimates (Haschke et al. 2012; Pietrukowicz et al. 2020).

Figure 10 also shows MDFs derived from the RRc stars using our predictive model discussed in Section 3. This model has serious limitations with its training data being largely constrained to $[\mathrm{Fe/H}] \lesssim -1$. In addition, the RRc sample sizes are significantly smaller than those for RRab stars in all of the eight studied environments. In spite of this, the modes of the resulting RRc MDFs show good agreement with those of the RRab stars, with less than 0.1 dex offsets between the two. However, subtler features of the RRab MDFs, are generally poorly recovered by the RRc stars, especially at higher metallicities, such as the secondary mode of the disk MDF.

4.1. On the Effect of the He-abundance

Finally, we would like to emphasize certain important limitations of the general applicability of the RR Lyrae stars' photometric metallicity estimation. While our predictive model captures the dependency of the light-curve parameters on the bulk heavy-element content, there are other physical variables that can significantly influence the light-curve shapes of RR Lyrae stars, but which remained latent in the current formulation of the [Fe/H] regression problem and which currently cannot be taken into account due to the lack of data. For example, helium abundance is arguably such a latent physical quantity, which is hard to measure directly but known to vary across different Galactic environments and can have a significant spread even within monometallic globular clusters (for a review, see Bastian & Lardo 2018).

Our predictive model cannot be expected to generalize well to data outside the (not precisely known, but assumed to be close to canonical) He-abundance range of its training set of field RR Lyrae stars and deploying it on such data poses the risk of significantly biased [Fe/H] predictions. We illustrate this on the extreme example of the metal-rich globular cluster NGC 6441, which hosts RR Lyrae stars with anomalously long periods (Pritzl et al. 2001), interpreted as a result of a very large helium abundance of $Y \gtrsim 0.35$ (Caloi & D'Antona 2007). We analyzed the OGLE *I*-band photometry of its known RRab stars (see Pritzl et al. 2001; Corwin et al. 2006) following the procedure in Section 2.1, applied the same quality criteria as earlier, and omitted suspected blends, which resulted in a sample of 14 objects with very high-quality light curves. Applying our predictive model on this data set results in a predicted mean metallicity of $\langle[\mathrm{Fe/H}]\rangle \approx -1.8$, with individual estimates ranging from −1.5 to −2.3 dex. These values are

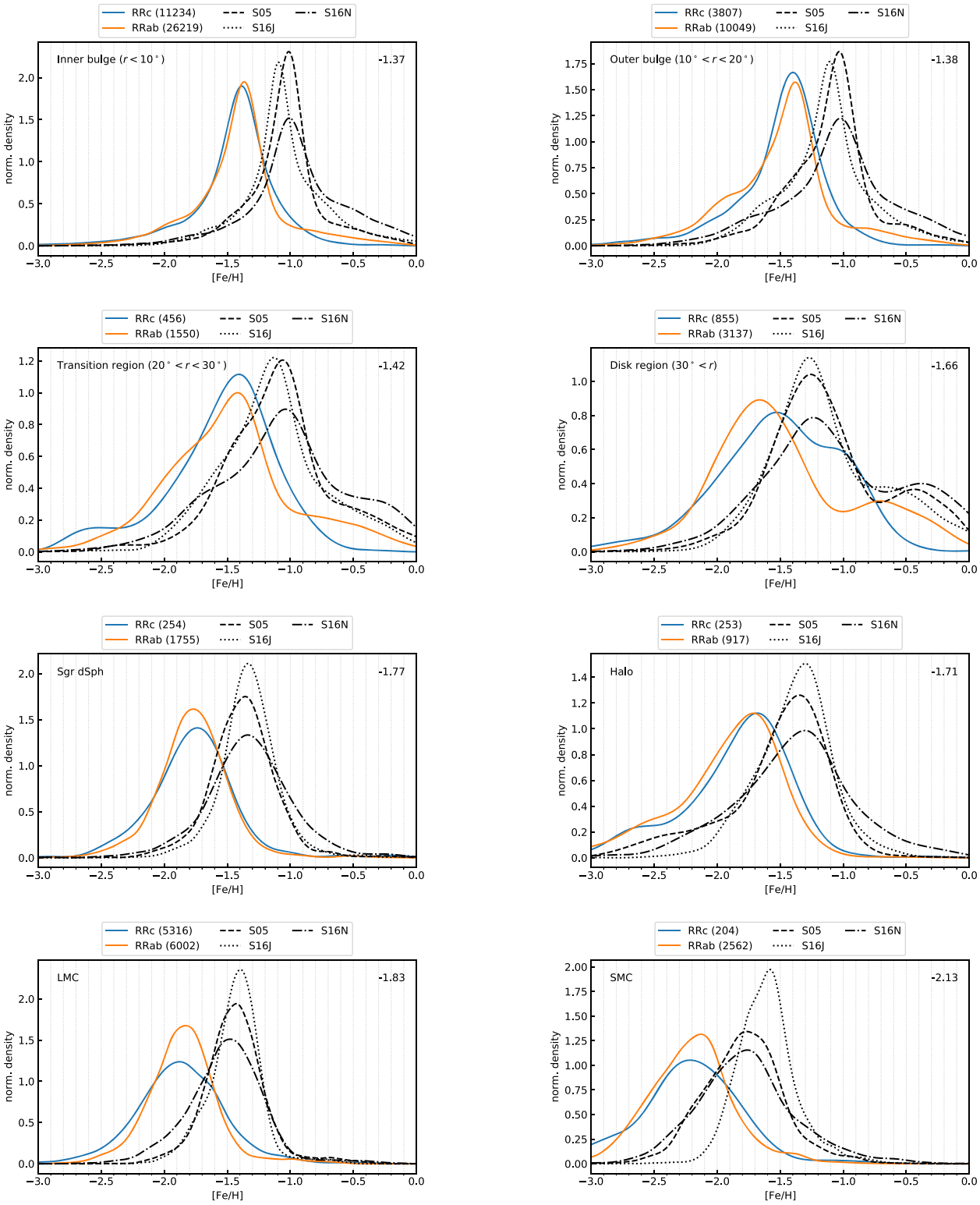

Figure 10. Photometric MDFs of RR Lyrae stars from the OCVS in different Local Group environments shown in the upper left corners of each panel. The curves denote kernel density estimates using Gaussian kernels and bandwidths according to Scott's rule. The modes estimated from the RRab (orange) density curves obtained from our predictive model are shown in the upper right corners of the individual panels. Orange and blue curves show the results computed using the [Fe/H] predictive models of this work for RRab and RRc stars, respectively. The dashed, dotted, and dashed–dotted black curves show the MDFs obtained by using the S05 three-parameter formula, the S16J, and the S16N formula, respectively. The number of stars used for computing the density estimates are shown in parentheses in each panel's legend. The angular Galactocentric distance (r) ranges of the relevant samples are also shown.

extremely biased with respect to the cluster's mean spectroscopic metallicity of $[\mathrm{Fe/H}] \approx -0.4$ measured by Clementini et al. (2005). Although this is probably one of the most exaggerated cases for the effect of He-abundance variations, it underlines our caveat to the applicability of the photometric metallicity estimator.

5. Discussion and Conclusions

In this work, we have provided new empirical relations for predicting the metallicities of RR Lyrae stars from Cousins *I*-band light curves. Our results are based on a homogeneous set of [Fe/H] abundances derived from HR spectroscopic measurements compiled from the contemporary literature, and high-quality photometric time series.

We used robust fitting techniques to derive unbiased parametric representations of the light curves. Using a standard machine-learning approach, we concluded that with the available data, the best trade-off between [Fe/H] predictive performance and model complexity for RRab stars can be obtained by linear models using the pulsation period and the ϕ_{31} and A_2 parameters, while the optimal results for RRc stars can be achieved with the $\{P, A_1, A_2, \phi_{31}\}$ feature set. We used cross-validation to measure the mean absolute error in the predictions to be 0.16 and 0.18 dex for the RRab and RRc stars, respectively, which are on par with the typical error level of low-resolution spectroscopic [Fe/H] measurements (e.g., Layden 1994). The final results were obtained by following a full Bayesian approach, yielding robust estimates of the optimal regression model parameters as well as their correlations and credible intervals. Being trained on a sizable HR spectroscopic data set, our predictive model yields accurate photometric metallicity estimates over a wide range (from solar to $[\mathrm{Fe/H}] \approx -2.5$). At the same time, our formula for the [Fe/H] estimation of RRc stars requires the important caveat that it is unconstrained for $[\mathrm{Fe/H}] \gtrsim -1$ due to the limited range and small size of its training data set.

We deployed our predictive models on a large number RR Lyrae light curves from the OGLE survey to obtain photometric metallicity distributions for various stellar populations in the Local Group. We found that despite its limitations, our predictive model for RRc stars recovers the distributions' modes computed from RRab data with an error of less than 0.1 dex. We would like to emphasize that targeted *I*-band time series observations of only about a dozen metal-rich field RRc stars with existing HR spectroscopic abundance measurements would already result in an enormous improvement of the [Fe/H] prediction formula.

Comparisons with individual [Fe/H] estimates and MDFs obtained from the previously employed *I*-band [Fe/H] prediction formulae of Smolec (2005) and Skowron et al. (2016) show that the latter have metallicity-dependent positive biases. In case of the S05 and the *I*-band transformed JK96 formulae, we attribute these to a similar bias in the underlying Jurcsik & Kovács (1996) metallicity scale. We found that all previously used formulae overestimate the modes of the MDFs of the Galactic bulge, disk, and the Magellanic Clouds by 0.35–0.4 dex and cause distortions in their shapes. This prevalent bias in the RR Lyrae photometric metallicities affects their use as standard candles as well by making the objects appear closer. Considering the state-of-the-art theoretical period–luminosity–metallicity relations of Marconi et al. (2015), a systematic error of the aforementioned size in the [Fe/H] corresponds to an offset of 0.06 magnitudes in the distance modulus. This translates to a bias of roughly 200 pc at the distance of the bulge, but amounts to an offset of almost 1.4 kpc at the distance of the LMC and errors of similar size may propagate into the distances of halo objects as well.

Having unbiased photometric metallicity estimates of RR Lyrae stars in various common photometric passbands is a key ingredient to their employment as distant indicators and old population tracers. By relating their accurate metallicity distribution through population synthesis models (see Savino et al. 2020) could give stringent constraints on the formation epoch and early evolution of various stellar environments. Furthermore, in tandem with widely used intermediate-age population tracers such as red clump giants, they can provide important ingredients to the inference of chemical enrichment histories.

The authors are grateful to the anonymous referee for constructive suggestions that helped to improve this study. The authors thank Dorota Skowron and Zdeněk Prudil for fruitful and enlightening discussions. I.D. and E.K.G. were supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—Project-ID 138713538—SFB 881 ("The Milky Way System", subproject A03).

Software: numpy (Harris et al. 2020), scikit-learn (Pedregosa et al. 2011), pymc3 (Salvatier et al. 2016), ArviZ (Kumar et al. 2019).

ORCID iDs

István Dékány https://orcid.org/0000-0001-7696-8331
Eva K. Grebel https://orcid.org/0000-0002-1891-3794
Grzegorz Pojmański https://orcid.org/0000-0002-6495-0676

References

Andrievsky, S., Wallerstein, G., Korotin, S., et al. 2018, PASP, 130, 024201
Bastian, N., & Lardo, C. 2018, ARA&A, 56, 83
Caloi, V., & D'Antona, F. 2007, A&A, 463, 949
Catelan, M., & Smith, H. A. 2015, Pulsating Stars (Weinheim: Wiley-VCH), 2015
Chadid, M., Sneden, C., & Preston, G. W. 2017, ApJ, 835, 187
Clementini, G., Carretta, E., Gratton, R., et al. 1995, AJ, 110, 2319
Clementini, G., Gratton, R. G., Bragaglia, A., et al. 2005, ApJL, 630, L145
Clementini, G., Ripepi, V., Molinaro, R., et al. 2019, A&A, 622, A60
Corwin, T. M., Sumerel, A. N., Pritzl, B. J., et al. 2006, AJ, 132, 1014
Crestani, J., Fabrizio, M., Braga, V. F., et al. 2021, ApJ, 908, 20
Dékány, I., & Grebel, E. K. 2020, ApJ, 898, 46
Dékány, I., Hajdu, G., Grebel, E. K., et al. 2018, ApJ, 857, 54
Dékány, I., Hajdu, G., Grebel, E. K., & Catelan, M. 2019, ApJ, 883, 58
Deng, L.-C., Newberg, H. J., Liu, C., et al. 2012, RAA, 12, 735
Fernley, J., & Barnes, T. G. 1996, A&A, 312, 957
For, B.-Q., Sneden, C., & WPreston, G. 2011, ApJS, 197, 29
Govea, J., Gomez, T., Preston, G. W., & Sneden, C. 2014, ApJ, 782, 59
Hajdu, G., Dékány, I., Catelan, M., Grebel, E. K., & Jurcsik, J. 2018, ApJ, 857, 55
Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, Natur, 585, 357
Haschke, R., Grebel, E. K., Duffau, S., & Jin, S. 2012, AJ, 143, 48
Hoffman, M. D., & Gelman, A. 2014, JMLR, 15, 1593
Iorio, G., & Belokurov, V. 2021, MNRAS, 502, 5686
Jacyszyn-Dobrzeniecka, A. M., Skowron, D. M., Mróz, P., et al. 2017, AcA, 67, 1
Jurcsik, J., & Kovács, G. 1996, A&A, 312, 111
Jurcsik, J., Sódor, A., Váradi, M., et al. 2005, A&A, 430, 1049
Kumar, R., Carroll, C., Hartikainen, A., & Martin, O. 2019, JOSS, 4, 1143
Kunder, A., Rich, R. M., Koch, A., et al. 2016, ApJL, 821, L25
Lambert, D. L., Heath, J. E., Lemke, M., & Drake, J. 1996, ApJS, 103, 183
Layden, A. C. 1994, AJ, 108, 1016
Le Borgne, J. F., Paschke, A., Vandenbroere, J., et al. 2007, A&A, 476, 307
Liu, G.-C., Huang, Y., Zhang, H.-W., et al. 2020, ApJS, 247, 68
Liu, S., Zhao, G., Chen, Y.-Q., Takeda, Y., & Honda, S. 2013, RAA, 13, 1307
Marconi, M., Coppola, G., Bono, G., et al. 2015, ApJ, 808, 50
Minniti, D., Lucas, P. W., Emerson, J. P., et al. 2010, NewA, 15, 433
Monson, A. J., Beaton, R. L., Scowcroft, V., et al. 2017, AJ, 153, 96

Muhie, T. D., Dambis, A. K., Berdnikov, L. N., Kniazev, A. Y., & Grebel, E. K. 2021, MNRAS, 502, 4074
Mullen, J. P., Marengo, M., Martínez-Vázquez, C. E., et al. 2021, ApJ, 912, 144
Muraveva, T., Delgado, H. E., Clementini, G., Sarro, L. M., & Garofalo, A. 2018, MNRAS, 481, 1195
Nemec, J. M., Cohen, J. G., Ripepi, V., et al. 2013, ApJ, 773, 181
Ngeow, C.-C., Yu, P.-C., Bellm, E., et al. 2016, ApJS, 227, 30
Pancino, E., Britavskiy, N., Romano, D., et al. 2015, MNRAS, 447, 2404
Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2011, JMLR, 12, 2825
Pietrukowicz, P., Udalski, A., Soszyński, I., et al. 2020, AcA, 70, 121
Pojmanski, G. 1997, AcA, 47, 467
Pritzl, B. J., Smith, H. A., Catelan, M., & Sweigart, A. V. 2001, AJ, 122, 2600
Salvatier, J., Wiecki, T. V., & Fonnesbeck, C. 2016, PeerJ Comp. Sci., 2, e55
Savino, A., Koch, A., Prudil, Z., Kunder, A., & Smolec, R. 2020, A&A, 641, A96
Skarka, M., Prudil, Z., & Jurcsik, J. 2020, MNRAS, 494, 1237
Skowron, D. M., Soszyński, I., Udalski, A., et al. 2016, AcA, 66, 269
Smolec, R. 2005, AcA, 55, 59
Smolec, R. 2016, in 37th Meeting of the Polish Astronomical Society, ed. A. Różańska & M. Bejger (Warszawa: Polish Astronomical Society), 22
Sneden, C., Preston, G. W., Chadid, M., & Adamów, M. 2017, ApJ, 848, 68
Soszyński, I., Udalski, A., Wrona, M., et al. 2019, AcA, 69, 321
Stringer, K. M., Drlica-Wagner, A., Macri, L., et al. 2020, ApJ, 911, 109
Suntzeff, N. B., Kraft, R. P., & Kinman, T. D. 1994, ApJS, 93, 271
Szczygieł, D. M., Pojmański, G., & Pilecki, B. 2009, AcA, 59, 137
Udalski, A., Szymański, M. K., & Szymański, G. 2015, AcA, 65, 1
Yanny, B., Rockosi, C., Newberg, H. J., et al. 2009, AJ, 137, 4377