

RESEARCH PAPER

The GALAH Survey: Data Release 4

S. Buder,^{1,2} Main contributors, including Janez Kos, Madeleine McKenzie, Ella Wang, Thomas Nordlander, Gregor Traven, Rok Vogrinčič, Michael Hayden, Dan Zucker,³ Joss Bland-Hawthorn, Gayandhi M. De Silva, Ken C. Freeman, Michael R. Hayden, Janez Kos, Sarah L. Martell, Geraint F. Lewis, Karin Lind, Sanjib Sharma, Jeffrey D. Simpson, Dennis Stello, Daniel B. Zucker, Tomaž Zwitter, other co-authors with significant contribution, including Kirsten Banks, Kevin Beeson, Ioana Ciucă, Chiaki Kobayashi, and The GALAH Collaboration

¹Research School of Astronomy and Astrophysics, Australian National University, Canberra, ACT 2611, Australia

²ARC Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), Australia

³To Be Decided.

Author for correspondence: S. Buder, Email: sven.buder@anu.edu.au.

(Received dd Mmm 2024; revised dd Mmm 2024; accepted dd Mmm 2024; first published online dd Mmm 2024)

Abstract

The stars in our Milky Way hold the chemical history of our Milky Way in their atmospheres as they wander through our Galaxy. Like bar codes, we can extract the chemical fingerprint of stars from high-resolution spectra. The fourth data release (DR4) of the Galactic Archaeology with HERMES (GALAH) Survey provides the chemical abundances of up to 31 elements for 917 588 stars which have exquisite orbit information from the *Gaia* satellite information.

For this data release, we utilise neural networks to simultaneously fit all stellar labels (stellar parameters and elemental abundances) for the full wavelength range. The neural networks are trained on grids of synthetic spectra that were computed with Spectroscopy Made Easy.

In a first iteration, we fit the best set of labels and in particular radial velocities for all 1 085 520 spectra (including repeat observations). These are used to properly co-add spectra (for example of spectroscopic binaries without line-splitting) for the second major iteration. This loop uses the distances inferred from astrometric data of the *Gaia* satellite and 2MASS photometry to self-consistently infer more accurate and precise surface gravities for each star.

Notable improvements of this data release are the newly available abundances of carbon and nitrogen from molecular features as well as the significantly improved precision and amount of elemental abundances, in particular for neutron-capture elements.

The combination of chemical compositions across multiple nucleosynthesis paths, orbit information, and age estimates for almost a million stars from GALAH DR4 provides a rich data set for Galactic exploration.

Keywords: Surveys; the Galaxy; methods: observational; methods: data analysis; stars: fundamental parameters; stars: abundances

1. INTRODUCTION AND WORKFLOW

1.1 Motivation

The history of our Milky Way galaxy is written in star light. By capturing and analysing the light of millions of stars that are now billions of years old, we can learn about the chemical composition of the material that was locked into their stellar atmospheres at birth. We can use stars as time capsules into the past evolution of the Milky Way. The light of stars can thus guide us to explore and map our environment, just as it guides Aboriginal and Torres Strait Islander people and their astronomers on Country since tens of thousands of years.

With this fourth data release (DR4) from the Galactic Archaeology with HERMES (GALAH), we are proudly publishing the next set of measurements of stellar chemical abundances for almost a third of the elements of the elements of the periodic table that are created by stars. The initial motivation for measuring so many elemental abundances was laid out by De Silva et al. (2015) and included the major motivation – chemical tagging – with the aim to trace back stars that were born together through their (expected) similar chemical compositions. The recent and ongoing efforts of GALAH and other surveys like the SDSS/APOGEE surveys (e.g. Abdurro'uf et al.,

2022; Kollmeier et al., 2017), LAMOST (Zhao et al., 2012), *Gaia*-ESO (Gilmore et al., 2022), or RAVE (Steinmetz et al., 2020) have taught us that the chemical evolution of our galaxy and stars is complex and we cannot (yet) recover stellar siblings due to limitations in our observations, analyses methods, and intrinsic changes to chemical composition due stellar evolution. New observations and innovations in the analysis that are presented in this data release, will allow us to make significant progress towards chemical tagging.

The unique observational setup of GALAH allows us to deliver the chemical information for a powerful and substantial set of stars: those which have exquisite orbit information from the revolutionary *Gaia* satellite (Gaia Collaboration et al., 2016) and for which we can estimate stellar ages either from empirical or theoretical models, like stellar isochrones or mass- and age-dependent changes of chemical compositions. By combining stellar ages, orbits, and chemistry, we have made major advances in our understanding of our Galaxy. In particular, the discovery of the major merger of the Milky Way with another slightly less massive galaxy between 8 and 10 Gyr ago (Belokurov et al., 2018; Helmi et al., 2018) was paradigm shifting and motivated a new rush to collect more and diverse

information about the stars in our Milky Way.

The work of this data release is motivated by expanding the previous analysis of spectra towards a more holistic workflow that can aid upcoming surveys like 4MOST (de Jong et al., 2019) and WEAVE (Dalton et al., 2014). Switching to the use of neural networks for interpolating synthetic spectra allows us to model the full wavelength range, molecular features such from C₂ and CN molecules, rather than a small selection of wavelength windows around atomic lines. It even enables us to model all stellar labels (global parameters and elemental abundances) simultaneously and derive interstellar properties from the residuals of observations and synthetic stellar spectra. Finally, we can also relax the (often wrong) assumption that we observe single stars, by superimposing synthetic spectra of two stars, and thus provide global parameters for binary systems as part of our workflow.

1.2 Workflow

The workflow of GALAH DR4 is depicted in Fig. 1. Following this workflow throughout this manuscript, we first describe the collection of data in Sec. 2, most notably the observation of HERMES spectra. We explain how we create synthetic stellar spectra to compare with the observed ones in Sec. 3. The comparison is then done in two consecutive steps. In Sec. 4, we explain how we extract stellar labels from individual observations (without non-spectroscopic information folded in), while Sec. 5 describes how we co-add repeated observations and fold in non-spectroscopic information for each star. Sec. 6 describes how we estimate main stellar labels for binary stars. We describe the post-processing of our data in Sec. 7 and validate our measurements in Sec. 8. We describe the data products of this data release in Sec. 9 and conclude this manuscript in Sec. 11.

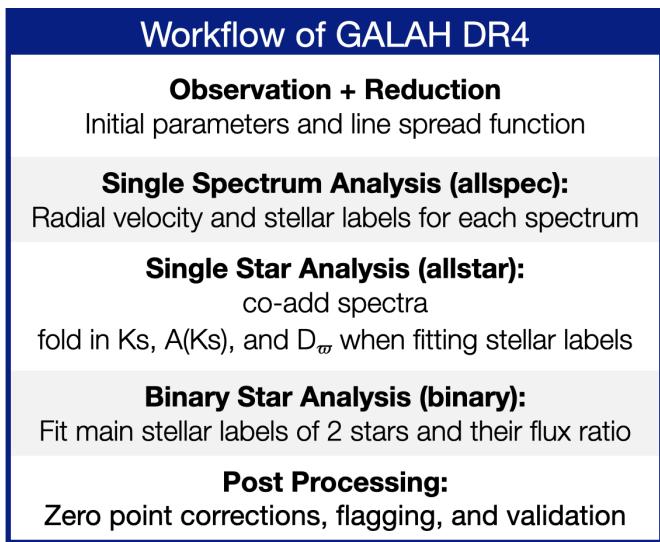


Figure 1. Workflow of GALAH DR4

2. DATA

The GALAH Survey is using the 3.9-metre Anglo-Australian Telescope at Siding Spring Observatory on Gamilaraay Country and its Two-Degree Field positioning system (2dF) to end (Lewis et al., 2002). 2dF places up to 400 fibre entrances on one of two plates which can be tumbled to allow parallel observing and fibre configuring. Light is then sent through fibres to the High Efficiency and Resolution Multi-Element Spectrograph (HERMES) spectrograph (Barden et al., 2010; Brzeski et al., 2011; Heijmans et al., 2012; Farrell et al., 2014; Sheinis et al., 2015) and dispersed into four non-contiguous wavelength bands in the optical and captured in separate rows of four separate charge-coupled devices (CCDs). The data used in this data release is primarily based on observations of stars with said setup, but also makes use of auxiliary photometric and astrometric information for the stars, where available.

In this Section, we describe which stars we have targeted and observed (Sec. 2.1) with the 2dF-HERMES setup, including the first description of the second phase of GALAH observations (GALAH Phase 2) with a sharper focus on main-sequence turnoff stars to estimate more precise ages. In Sec. 2.2, we briefly summarise the properties of the spectroscopic data and how they were reduced to one-dimensional spectra. We also point out major changes of the observations and reductions with respect to the previous third data release (Buder et al., 2021). We further elaborate on the auxiliary information that was used for the analysis in Sec. 2.3.

2.1 Target selection and observational setup

GALAH DR4 is a combination of the main GALAH survey and additional proposals to observe targets of the K2 Sharma et al. (2019) and TESS missions Sharma et al. (2018) to combine spectroscopic and asteroseismic data, as well as hand-picked and random observations of globular and open cluster members. The column `survey_name` in our catalogues denotes this origin.

2.1.1 Target selection of GALAH Phase 1 and 2

For GALAH Phase 1, we used the 2MASS photometric survey (Skrutskie et al., 2006) with its J and K_S filters as parent sample, from which we selected stars within empirical visual magnitudes

$$V_{JK} = K_S + 2(J - K_S + 0.14) + 0.382e^{(J-K_S-0.2)/0.5}. \quad (1)$$

For GALAH Phase 1, a tiling pattern (with unique `field_id` entries) with 2 deg fields of view below declination $\delta < 10$ deg were created for regions with Galactic latitude $|b| > 10$ deg to avoid crowding and strong extinction. For each tile, a selection of 400 stars within magnitudes $9 < V_{JK} < 12$ for a bright magnitude cut and $12 < V_{JK} < 14$ for the nominal magnitude cut is randomly selected from the complete parent sample of 2MASS. This typically selects 2/3 main sequence and turnoff stars and 1/3 evolved stars.

For GALAH Phase 2, a stronger focus on turn-off stars was implemented with the photometric and astrometric information of *Gaia* data release 2 as parent sample. For each

field, we therefore first allocate fibres to stars with absolute magnitude between $2 < M_G < 4$, where

$$M_G = \text{phot_g_mean_mag} + 5 \cdot \log_{10} \left(\frac{\text{parallax}}{100 \text{ mas}} \right) \quad (2)$$

with G magnitude and parallax measurements from *Gaia* DR2 (Gaia Collaboration et al., 2018; Evans et al., 2018; Lindegren et al., 2018). Remaining fibres are again filled with targets as done in Phase 1.

2.1.2 Observational setup

Table 1. Overview of stars observed for the programs included in this data release. Numbers of open and globular cluster observations were estimated after observations as described in Sec. 2.3.3. We have observed 18 globular clusters (13 with ≥ 5 stars) and 286 open clusters (94 with ≥ 5 stars).

Program	Nr. Stars	Program	Nr. Stars
galah_bright	65 188	k2_hermes	117 708
galah_main	427 699	tess_hermes	37 228
galah_faint	25 802	globular clusters	1 185
galah_phase2	125 058	open clusters	4 487
commissioning	2 625	other	20 308

We list the observations under certain programs in Table 1. Except for 2 935 spectroscopic observations with the high-resolution mode of HERMES ($R \sim 42\,000$) on 7/8/10/11/12th February 2014, all observations were made in the low-resolution mode ($R \sim 28\,000$). Exposure times depend on programs. Under sufficient conditions (no clouds and seeing below 2 arcsec), GALAH Phase 1 and TESS-HERMES observed 3x6 minutes for bright targets ($9 < V_{JK} < 12$) and 3x20 minutes for the majority of targets ($12 < V_{JK} < 14$). GALAH Phase 2 extended these times to 3x10 and 3x30 minutes, respectively, and included repeat observations of GALAH Phase 1 main targets with another 3x15 minutes. K2-HERMES observations targeted stars with $13 < V_{JK} < 15$ or even $13 < V_{JK} < 15.8$ to complement targets of the K2 Galactic Archaeology Program (Stello et al., 2015) and were observed for 2 hours, similar to most globular and open cluster stars. Worse seeing conditions or thin clouds caused triggered between one ($2 < \text{seeing} < 2.5$ arcsec) and 3 ($2.5 < \text{seeing} < 3$ arcsec) additional exposures. In addition to the science frames, fibre flat and arc observations from a ThXe were taken before or after each exposure together with bias frames at the beginning of each night.

2.2 Spectroscopic data from GALAH observations

Since the commissioning of the HERMES spectrograph in late 2013 until 22 August 2022, the GALAH collaboration and its partners have observed and successfully reduced 943 654 spectra of 827 288 stars. The spectroscopic data is collected by four separate CCDs that cover $\sim 1000 \text{ \AA}$ in the range of 4713–4903 (blue CCD or CCD1), 5648–5873 (green / CCD2), 6478 – 6737 (red / CCD3), and 7585 – 7887 \AA (infrared IR /

CCD4), as can be seen from the reduced example spectrum of the asteroid 4 Vesta in Fig. 2. The reduction process to create FITS files of reduced spectra from two-dimensional images of the cameras is employing an updated and publicly available version 6 of the already well-tested reduction pipeline (Kos et al., 2017). The extensions of these files are listed in Tab. 2 and created as follows.

Table 2. Data product 1: FITS files of reduced spectra

FITS Ext.	Description
Ext. 0	Un-normalised signal / counts
Ext. 1	Normalised signal (by reduction pipeline)
Ext. 2	Relative uncertainty of signal
Ext. 3	Subtracted sky signal / counts
Ext. 4	Applied telluric correction
Ext. 5	Scattered light / counts
Ext. 6	Cross-talk
Ext. 7	Resolution profile / FWHM

Science frames are corrected by removing biases, fixing gains of the two CCD halves, fixing bad pixels, dividing by master flats, as well as removing cosmics and scattered light. Subsequently, apertures are identified and used to extract the individual spectra.

Wavelength calibrations are performed via Chebychev polynomial functions based on the ThXe arc frames and the spectra are interpolated onto a linearly increasing wavelength grid with starting wavelength CRVAL1 and dispersion CDELT1. Both values are saved in the headers of the FITS files.

Finally, sky lines are subtracted and telluric lines removed, before a barycentric correction is applied to create the 'reduced' spectra that are saved in extension 1 of the reduction pipeline FITS files and used for the subsequent analysis.

Noise / uncertainties are saved in extension 2 and calculated from the square root of the sum of flux, sky features (extension 3), scattered light (extension 5), crosstalk (extension 6) and as well as the squared readout noise.

The wavelength dependent line spread functions (LSFs) are measured from the arc calibration frames for each spectrum and CCD by fitting modified Gaussian distributions with a boxiness parameter b and full width half maxima $fwhm$ for each wavelength point of the spectrum, that is

$$\exp(-0.693147|2 \cdot x/fwhm|^b) \quad (3)$$

The array x then includes the pixels around each wavelength step that are used to apply the convolution from higher resolution to GALAH resolution spectra. The fitted values of $fwhm$ are saved in extension 7 with b in its header.

Should we elaborate on any overall information, like SNR overview and pointing on the sky? Elaborate on how the reduction pipeline parameters `teff_x` etc. are estimated.

2.3 Auxiliary data from Gaia, 2MASS, and literature

To support our spectroscopic analysis, we make use of astrometric and photometric information from the *Gaia* satellite (Gaia

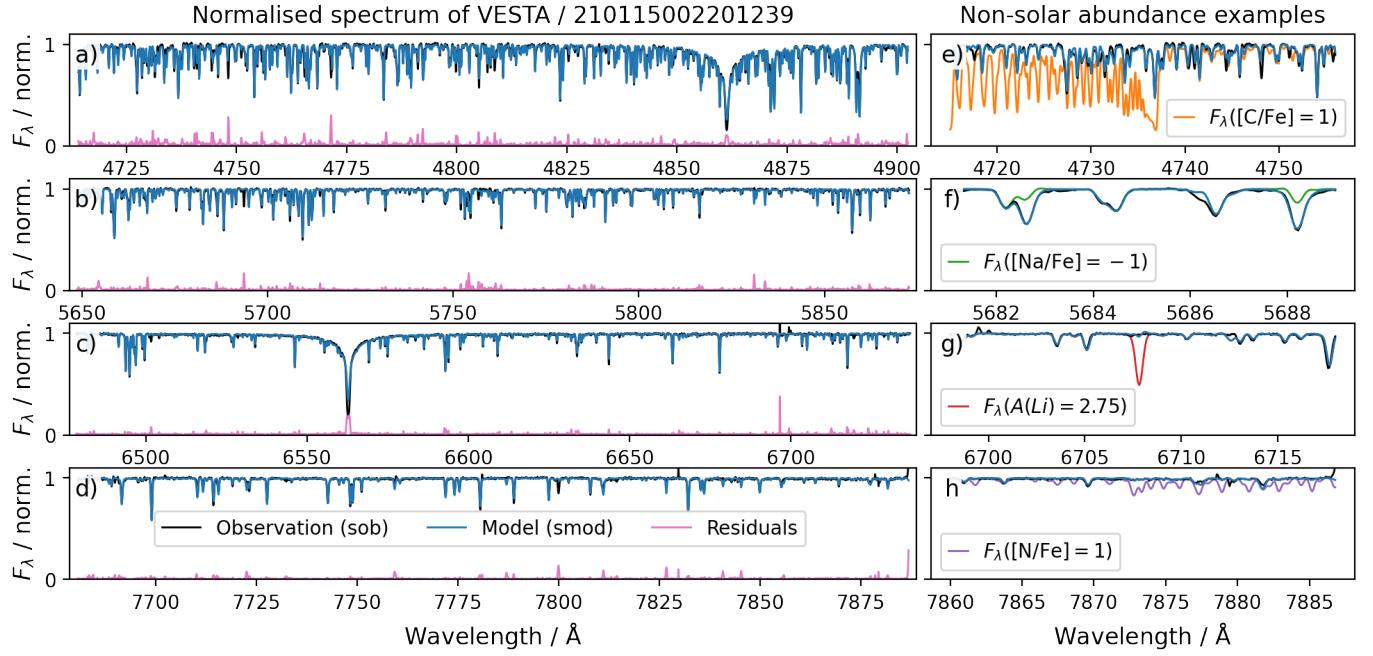


Figure 2. Comparison of normalised observed (black) and synthetic spectra (blue) of the asteroid 4 Vesta with solar composition as well as examples of synthetic spectra with non-solar abundances. Panels a-d) show the observed and best-fit synthetic spectrum as well as their absolute residual (pink) for the four wavelength channels of the HERMES spectrograph. Panel e) shows the beginning of the blue CCD 1 (left-most part of panel a) with an additional synthetic spectrum of ten-times higher C in orange, for which the C₂ molecular absorption, the Swan bands, features would be prominent. Panel f) shows the beginning of the green CCD 2 (left-most part of panel b) and exemplifies with a synthetic spectrum in green that also a ten-times lower Na abundance (for example of accreted stars) can still be detected. Panel g) shows the end of the red CCD 3 with a synthetic spectrum of primordial Li abundance of A(Li) = 2.75 in red. While this abundance could be detected, the line for the Solar value A(Li) = 1.05 is barely detectable. Panel h) shows the end of the infrared CCD 4, which would show strong molecular absorption features of the CN molecule for [N/Fe] = 1 (purple).

Collaboration et al., 2016) and 2MASS survey (Skrutskie et al., 2006) which is available for essentially all our targets. *Gaia*: 826 076 (99.9 %) of the 827 288 stars observed by GALAH have photometric measurements catalogued in *Gaia* DR3 and the 2MASS Survey. We further use the value-added catalogues, like distance estimates for field stars by (Bailer-Jones et al., 2021) as well as open and globular clusters by Cantat-Gaudin & Anders (2020) as well as Vasiliev & Baumgardt (2021) and Baumgardt & Vasiliev (2021).

2.3.1 Gaia DR3

We crossmatch our observations to the *Gaia* DR3 catalogue (Gaia Collaboration et al., 2021, 2022) via the 2MASS via the nearest neighbour crossmatches provided as part of *Gaia* DR3 (Torra et al., 2021). 820 715 (99.2 %) also have astrometric information (Lindegren et al., 2021b) and 774 915 stars (93.7 %) have radial velocity estimates (Katz et al., 2022). We apply the corrections to both photometric (Riello et al., 2021) and astrometric (Lindegren et al., 2021a) information. Where possible we prefer the photogeometric distances over the geometric distances by (Bailer-Jones et al., 2021). Where neither are available, we further try to find parallaxes from van Leeuwen (2007). The average parallax uncertainty of the GALAH stars is $\sigma_{\omega}/\omega = 1.5^{+2.6}_{-0.9}\%$. Only 2.3 % of GALAH stars have no parallax measurements or parallax measurements beyond 20% uncertainty, for which the priors by (Bailer-Jones et al., 2021)

start to dominate distance estimates.

2.3.2 2MASS, WISE, and extinction

In addition to the excellent infrared photometry for 99.9 % of our sources from the 2MASS survey (Skrutskie et al., 2006), 98.7 % of them have far-infrared measurements from the WISE mission (Cutri et al., 2014). We therefore can estimate the extinction in the K_S , via the RJCE method (Majewski et al., 2011) $A_{K_S} = 0.917 \cdot (H - W2 - 0.08)$ for most stars. We confirm this estimate by estimating the extinction in K_S via the extrapolation of the color extinction of $B - V$, that is, $A(K_S) \sim 0.36 \cdot E(B - V)$ (Cardelli et al., 1989). We revert to this value, if it is less than half the value of the RJCE estimate or if either of the H and W2 bands does not have an excellent quality flag 'A'. For negative estimates by the RJCE method and very nearby stars (< 100 pc) we null the value.

2.3.3 Open and Globular Cluster members and distances

We identify open cluster members via the membership catalogue from Cantat-Gaudin & Anders (2020) via crossmatch with the *Gaia* source_id and adjust their parallaxes and distance estimates to the average cluster values if the latter are more precise. We identify globular cluster members (with more than 70% probability) via the membership catalogue from Vasiliev & Baumgardt (2021) by crossmatching with the *Gaia* source_id. We then adjust the parallaxes and distances

for the member stars to the mean values listed by Baumgardt & Vasiliev (2021).

3. SYNTHETIC SPECTRA FOR 2DF-HERMES

The aim of our spectroscopic analysis is to estimate the best set of stellar properties (labels) that influence a stellar spectrum by minimising the difference of observed stellar spectra to synthetic ones that were created with said stellar labels. In our endeavour to push the limits even further, we are advancing our analysis to fit all 31 elemental abundances and stellar parameters across the full GALAH wavelength range simultaneously with the appropriate model spectra. In this release we implement the possibility to also create synthetic binary star spectra through the superposition of two synthetic spectra and optimise them and respective radial velocity shifts of the binary stars.

To make this computationally feasible, we follow an idea reported by Rix et al. (2016) and create flexible models for smaller parts of the parameters space from only a limited number of ab initio models (see also Ting et al., 2016). Our ab initio models are calculated with Spectroscopy Made Easy (SME Valenti & Piskunov, 1996; Piskunov & Valenti, 2017) for the whole wavelength range and all visible atomic and molecular lines for random selections of elemental abundances and stellar parameters within the range of MARCS atmospheres (Gustafsson et al., 2008) at much higher resolution and sampling than our HERMES spectra. We then select subsets of these spectra within a restricted space of the three main spectroscopic parameters T_{eff} , $\log g$, and $[\text{Fe}/\text{H}]$. This idea is comparable to selecting Solar or stellar twin spectra when analysing the Sun (Nissen, 2015, see e.g.) or differential abundance analysis of globular cluster stars (Yong et al., 2013; Monty et al., 2023, e.g.). As they cancel out several systematic issues of line data and atmospheric effect, these approaches have been highly successful (Nissen & Gustafsson, 2018). For each subset, we train a neural network that correlates stellar flux and labels (stellar parameters and abundances) similar to *The Payne* (Ting et al., 2019). With these models, we can then create model spectra with all lines over the whole wavelength range for any combination of element abundances within this restricted parameter space within less than a second (compared to minutes or hours for physics-driven syntheses).

Another reason to create smaller training sets rather than a "one-fit-all" approach was the limited flexibility of both quadratic and non-linear interpolation routines within computationally reasonable model sizes and ability to test the model accuracy over the full parameter space. Surveys like GALAH, RAVE or APOGEE aim to fit basically all types of spectra at once. This means we demand one model to predict Sun-like stars, red clump stars, metal-poor stars with almost no absorption features, cool evolved stars with strong molecular features, and hot stars with shallow and broad lines – all at once and for up to 31 elemental abundances. As one can imagine this is an impossible task and has led to numerous systematic trends in catalogues for the most extreme cases that use such model interpolations Buder et al. (2018); Casey et al. (2016); Ting et al.

(2019). In this data release, we therefore purposefully limit the spectral complexity by creating smaller models. The hot star model therefore does not need to also predict the strong molecular absorption features of a cool stars. We discuss the possible caveats and disadvantages of this approach and our particular implementation in Sec. 10.

In this section, we lay out how we create smaller bins in the parameter space from which we sample a training set (Sec. 3.2) rather than one training set for all stars. We lay out how we create the parent sample of high-resolution synthetic stellar spectra (Sec. 3.2) and train neural networks to quickly predict/interpolate new synthetic spectra (Sec. 3.3).

3.1 Stellar twin training sets rather than one-fits-all

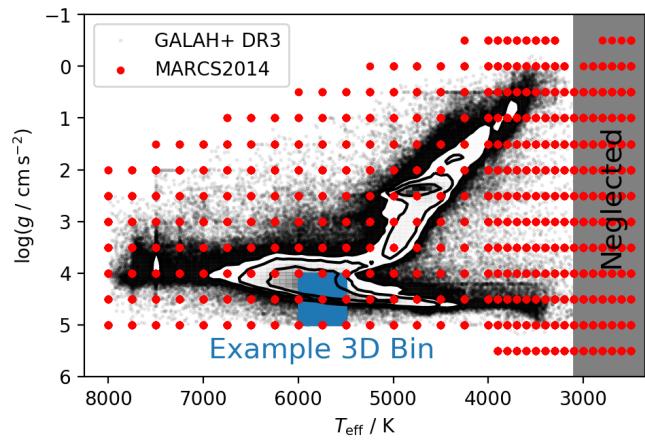


Figure 3. Coverage in T_{eff} and $\log g$ of MARCS2014 grid (red) and GALAH DR3 (black, including density contour). Shown is also an example of one of the 3D bins used to create models with *The Cannon*. MARCS grid points $T_{\text{eff}} < 3100$ K or $[\text{Fe}/\text{H}] < -3$ dex are neglected throughout GALAH DR4.

The base grid for our training set computation is the MARCS grid (Gustafsson et al., 2008), which is shown with red points in Fig. 3. Following the aforementioned idea of restricting ourselves to stellar siblings, we create multiple 3-dimensional bins in T_{eff} , $\log g$, and $[\text{Fe}/\text{H}]$ within ± 1 grid points in T_{eff} (with either ± 250 or ± 100 K), $\log g$ (± 0.5 dex), and $[\text{Fe}/\text{H}]$ (± 0.5 or ± 0.25 dex). An example box is shown for Solar siblings bin as blue box in Fig. 3.

Within these bins, we sample 280^a synthetic spectra with no rotational broadening, which are later broadened with different rotational velocities $v \sin i$ to create between 1680 and 2240 training set spectra for each bin. We explain the sampling of parameters and abundances in an exemplary way for the 3D bin centred on $T_{\text{eff}} = 5750 \pm 250$ K, $\log g = 4.5 \pm 0.5$ dex and $[\text{Fe}/\text{H}] = 0.0 \pm 0.25$ dex (see blue box in Fig. 3).

Stellar parameters (T_{eff} , $\log g$, $[\text{Fe}/\text{H}]$, v_{mic}) and elemental abundances $[\text{X}/\text{Fe}]$ of all 30 elements are randomly sampled within reasonable limits (see examples in Tab. 3) and fed into SME to create self-consistent synthetic spectra over the full wavelength range for MARCS atmospheres.

^aThis number is chosen to match the 28 CPUs of our computing nodes.

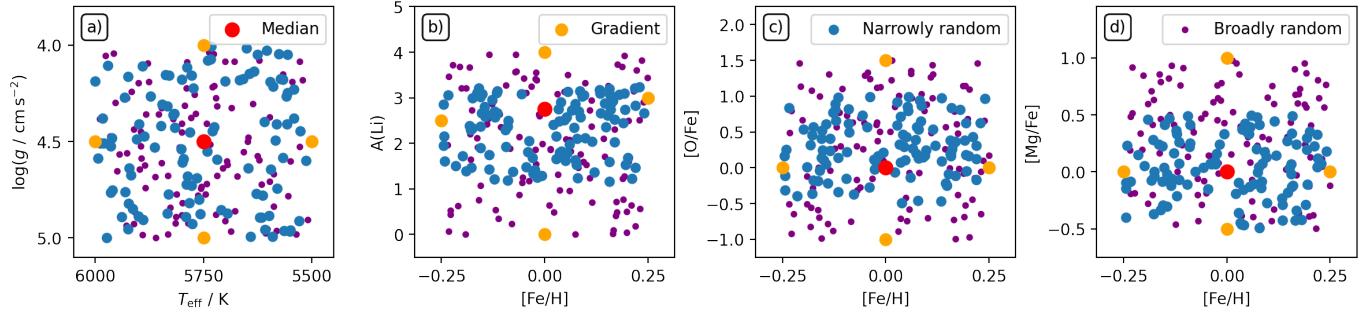


Figure 4. Coverage of stellar parameters and abundances for one of the 3D bins. Shown is the example of the solar 3D bin ($T_{\text{eff}} / \text{K} = 5750$, $\log(g / \text{cm s}^{-2}) = 4.5$, $[\text{Fe}/\text{H}] / \text{dex} = 0.0$). **Panel a)**: T_{eff} and $\log g$, **Panel b)**: $[\text{Fe}/\text{H}]$ vs. $[\text{Li}/\text{Fe}]$, **Panel c)**: $[\text{Fe}/\text{H}]$ vs. $[\text{O}/\text{Fe}]$, **Panel d)**: $[\text{Fe}/\text{H}]$ vs. $[\text{Mg}/\text{Fe}]$. While T_{eff} , $\log g$, and $[\text{Fe}/\text{H}]$ are sampled randomly within the 3D bin, the abundances are sampled both narrowly (blue) and broadly (purple) within limits as described in the text. Red point indicate the median spectrum and orange points the adjusted spectra to test the gradient change of spectra with individual label.

Table 3. Example of boundaries for the uniform sampling of synthetic spectrum labels (stellar parameters and elemental abundances) for the 3-dimensional bin of Solar siblings 5750_4.50_0.00.

Parameter	Sampling	Element	Sampling Narrow	Element	Sampling Broad
$T_{\text{eff}} / \text{K}$	5500..5750..6000	$A(\text{Li})$	1.05..2.75..3.26	$A(\text{Li})$	0.00..4.00
$\log g / \text{dex}$	4.0..4.5..5.0	C, N, O	-0.5..0.0..1.0	C, N, O	-1.0..1.5
$[\text{Fe}/\text{H}] / \text{dex}$	-0.25..0.0..0.25	Y, Ba, La, Ce, Nd	-0.5..0.0..1.0	Y, Ba, La, Ce, Nd	-1.0..1.5
$v_{\text{mic}} / \text{km s}^{-1}$	0.5, 1.5, 4.0, but see Eq. 4	$[\text{X}/\text{Fe}]$ for Mg, Si, Ti	-0.5..0.0..0.5	$[\text{X}/\text{Fe}]$ for Mg, Si, Ti	-0.5..1.0
$v \sin i / \text{km s}^{-1}$	0.0, but see Eq. 5	$[\text{X}/\text{Fe}]$ for all other elements	-0.5..0.0..0.5	$[\text{X}/\text{Fe}]$ for all other elements	-1.0..1.0

v_{mic} values are sampled uniformly between the upper and lower limits of the empirical relation from GALAH DR3 (Eqs. 4 and 5 from Buder et al., 2021) and an adjusted version of the relation by Dutra-Ferreira et al. (2016). The latter has been adjusted for $T'_{\text{eff}} = T_{\text{eff}} - 5500 \text{ K}$ as well as $\log g' = \log g - 4.0$ to return:

$$v_{\text{mic}} = \begin{cases} 1.198 + 3.16 \cdot 10^{-4} \cdot T'_{\text{eff}} - 0.253 \cdot \log g' \\ -2.86 \cdot 10^{-4} \cdot T'_{\text{eff}} \cdot \log g' + 0.165 \cdot \log g' \end{cases} \quad (4)$$

3.2 Higher-resolution synthetic spectra with SME

We create training sets from high-resolution stellar spectra for each smaller 3D bin region of the parameter space. We compute over-sampled synthetic intensity spectra at ten times higher resolution than the typical GALAH resolution with SME for seven equal-area angles of a stellar surface (see Fig. 5).

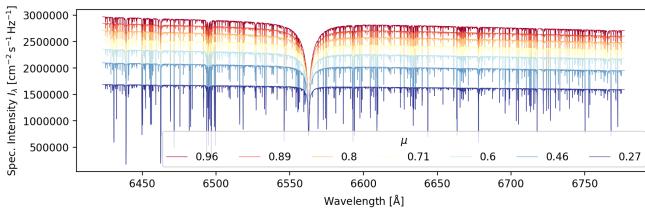


Figure 5. Example output of SME for a solar spectrum in HERMES CCD3 (around the Balmer H_{α} line). Shown are the specific intensities (sme.sint) as a function of the equal-area midpoints of each equal-area annulus μ .

For each spectrum, we first run a test on all available lines in the GALAH linelist, which is adapted from Heiter et al. (2021)

and includes small changes to correct wrong $\log g_f$ values for few lines within the GALAH wavelength range. We keep all atomic lines for the final synthesis and restrict the molecular lines to those with sme.depth above 0.001.

Spectra are computed at a resolution of $R = 300,000$ on a fine wavelength grid with 60,819 pixels spread over the extended wavelengths 4675.1–4949.9, 5624.1–5900.9, 6424.1–6775.9, and 7549.1 – 7925.9 Å. We note that these extend significantly beyond the actual GALAH wavelength range.

We use 1D MARCS atmospheres from the MARCS grid (Gustafsson et al., 2008, version 2014), and interpolate them for combinations of T_{eff} , $\log g$, and $[\text{Fe}/\text{H}]$. We use grids of non-LTE departure coefficients by Amarsi et al. (2020) for H, Li, C, N, O, Na, Mg, Al, Si, K, Ca, Mn, and Ba. For all, except C, we use grids that include background scattering.

Our synthetic grid explicitly includes C and N abundances. C was previously included in the analysis of GALAH DR3, but limited to the atomic C line. The analysis thus neglected the molecular absorption features of C_2 and CN at the beginning of CCD1 and end of CCD4, respectively. With the new self-consistent grid, we can however include these features, as they hold valuable information for both C and N, as well as several other features through the molecular equilibrium in stars (see e.g. Ting et al., 2018).

To be able to test that the flux-label correlations found by our subsequent polynomial interpolation are limited to reasonable wavelength ranges, we also calculate one spectrum that is exactly in the middle of the parameter range and additional spectra, where we increase the value of one label at a time (e.g. increase $[\text{O}/\text{Fe}]$ by 1 dex) to test the response in the synthetic

spectrum.

To save computational costs, we compute synthetic spectra with no rotational or macroturbulence broadening ($v_{\text{mac}} = v \sin i = 0 \text{ km s}^{-1}$), but save the model continuum flux (`sme.cmmod`) and the specific intensities (`sme.sint`) as a function of the equal-area midpoints of each equal-area annulus μ (see Fig. 5). We then apply the broadening of spectra due to rotation ($v \sin i$) with the flux integration code of the python-implementation PySME (Wehrhahn, 2021) of SME (Piskunov & Valenti, 2017). Depending on the expected rotational velocities (increasing with temperature) we sample a range of

$$v \sin i / \text{km s}^{-1} \in \{1.5, 3, 6, 9, 12, 18, 24^*, 36^{**}\}. \quad (5)$$

For bins with $T_{\text{eff}} \geq 5000 \text{ K}$, we also include $v \sin i = 24 \text{ km s}^{-1}$ and for those with $T_{\text{eff}} \geq 6000 \text{ K}$ we also include $v \sin i = 36 \text{ km s}^{-1}$.

3.3 Interpolating synthetic spectra with neural networks

To allow a fast interpolation with new and different stellar labels, we use the method of training descriptive models to connect stellar fluxes at given pixels from a combination of stellar labels. This method is well established in stellar spectroscopy through the successful applications of quadratic models with *The Cannon* (see e.g. Ness et al., 2015, 2016; Casey et al., 2016, 2017; Ho et al., 2017; Buder et al., 2018) as well as neural networks with *The Payne* (see e.g. Ting et al., 2019; Xiang et al., 2019, 2021). Because of the needed flexibility^b to predict synthetic spectra with 36 stellar labels for a large parameter space, we are also choosing neural networks to interpolate between our synthetic spectra in this data release.

In particular, we use the neural network architecture and training algorithms similar to *The Payne* (Ting et al., 2019) for our data. We describe the connection of stellar labels ℓ and the flux f at each wavelength pixel λ via

$$f_{\lambda} = w \cdot \text{lReLU} \left(\tilde{w}_{\lambda}^j \cdot \text{lReLU} \left(w_{\lambda i}^k \ell_k + b_{\lambda i} \right) + \tilde{b} \right) + \bar{f}_{\lambda}, \quad (6)$$

that encapsulates the so called layers of a neural network with $i = 300$ neurons and where we use the default leaky Rectified Linear Unit (lReLU)

$$\text{lReLU}(x) = \begin{cases} x & x \geq 0 \\ 0.01x & x < 0. \end{cases} \quad (7)$$

After optimising a loss function for 10^4 steps, we consider the network trained with an optimised combination of three sets of weights and biases within the minimum and maximum ranges of each label. The trained networks can then be used with new input labels to quickly create synthetic spectra for the label optimisation.

^bFor a more detailed discussion on the advantages of neural networks for predicting spectra see Ting et al. (2019).

4. SINGLE SPECTRUM ANALYSIS (ALLSPEC)

As outlined in Sec. 1, the workflow of GALAH DR4 includes a first analysis step of all observed spectra without taking non-spectroscopic information for the optimisation. This allows to identify shifts of radial velocities between spectroscopic observations of the same stars^c and a more accurate co-adding of spectra for the *allstar* analysis (see Sec. 5). Another motivation for this step is to get a first estimate of stellar labels without potentially biased photo- and astrometric, for example for binary stars.

The optimisation of stellar labels is thus aiming to minimise the absolute difference between synthetic and observed spectrum for all unmasked pixels, weighted by their uncertainty. Starting from a set of initial labels (Sec. 4.1), we create high-resolution synthetic spectra and downgrade them to the observed resolution and wavelength grid of each spectrum. In an important improvement of our analysis, we perform an on-the-fly re-normalisation of the observed spectrum, which allows a more accurate comparison of synthetic and observed spectra (Sec. 4.2) and thus a more accurate stellar label optimisation (see Sec. 4.3).

4.1 Initial stellar labels

Initial values of all stellar labels are needed for creating a first synthetic spectrum. For v_{rad} , T_{eff} , $\log g$, and $v \sin i$ we use a combination of sources. Where possible, we use the previous estimates from GALAH DR3 (Buder et al., 2021) and otherwise those from the GALAH DR4 reduction pipeline (Sec. 2.2). Because of the limited accuracy of the latter for cool stars with $T_{\text{eff}} < 4000 \text{ K}$ as well as the hottest stars with $T_{\text{eff}} > 6500 \text{ K}$, we perform a consistency check with photometric information from *Gaia* DR3 (Gaia Collaboration et al., 2021), and 2MASS (Skrutskie et al., 2006). For most of the aforementioned cool and hot stars, we therefore prefer the parameters from the *Gaia* DR3 photometric pipeline GSP-Phot (Andrae et al., 2022; Fouesneau et al., 2022).

In selected cases, we further adjust the starting parameters towards reasonable limits, for example for hot stars which are likely to be young and close to Solar [Fe/H]. Furthermore, we recalculate initial v_{mic} based on Eq. 4 and limits rotational broadening values to $3 \leq v \sin i \leq 10 \text{ km s}^{-1}$ for stars below $T_{\text{eff}} = 5500 \text{ K}$ and $3 \leq v \sin i \leq 20 \text{ km s}^{-1}$ for hotter stars. The explicit choices of starting values for T_{eff} , $\log g$, [Fe/H], v_{mic} , and $v \sin i$ are described in our online repository and are depicted in Fig. 28.

Based on the value of [Fe/H] we apply an offset to the $-$ -elements O, Mg, Si, Ca, and Ti within our wavelength range. The initial value is 0.4 for $[\text{Fe}/\text{H}] < -1$, 0.0 for $[\text{Fe}/\text{H}] > 0$, and $-0.4 \cdot [\text{Fe}/\text{H}]$ for $-1 \leq [\text{Fe}/\text{H}] \leq 0$. All other abundances are initialised at $[\text{Fe}/\text{H}] = 0$.

^cWhile repeat observations were only done for quality assurance in GALAH Phase 1, we have started to revisit a large amount of the fields for another 2x15 minutes to reach the updated needed exposure time for main GALAH fields of 1.5 hours.

4.2 Comparison of synthetic spectra to observations

The major aim of our spectroscopic analysis is to predict the best set of stellar labels by minimising the uncertainty-weighted difference of observed and synthetic spectra. In this section, we describe several important steps to enable the pixel-level comparison of the higher resolution, oversampled synthetic spectra created with the neural networks from Sec. 3.3 and the observational data at actually measured resolution and sampling (presented in Sec. 2.2).

4.2.1 Downgrading of synthetic spectra to observed resolution

Because dedicated line-spread-function measurements are available for every spectrum (see Sec. 2.2), we use this information to downgrade our synthetic spectra to the measured resolution of each observations. We then interpolate the over-sampled synthetic spectrum onto the observed wavelength.

4.2.2 On-the-fly re-normalisation of observed spectrum

Measurements of the GALAH flux and flux uncertainty are reported in counts by the reduction pipeline. To compare with our synthetic spectra, which are normalised to the continuum, we fit an outlier-robust polynomial function to the ratio of observed and synthetic spectrum and re-normalise our observed spectra and their uncertainties via this normalisation function.

This specific approach is similar to the internal routine of SME (Piskunov & Valenti, 2017) and has the important advantage, that no continuum points have to be defined. This is advantageous, because we try to cover the full parameter range of FGKM stars for which positions of continuum points – corresponding to 1 on a (pseudo-)continuum-normalised spectrum – differ significantly or for which continuum points may not even be present (as is the case for M stars).

We make two additional adjustments to the reduced spectra, which come in the form of counts and uncertainty per wavelength, f_λ and $\sigma_{f,\lambda}$.

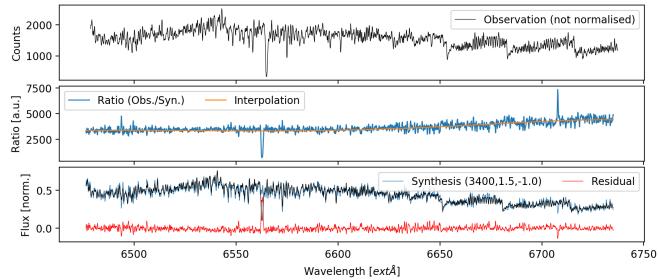


Figure 6. Example of normalisation for GALAH DR4 for a model spectrum that is selected during the label optimisation. Panel (a): Observed spectrum (counts) of star 2MASS XYZ. Panel (b): Ratio (blue) of observed spectrum and model spectrum as well as Chebychev polynomial fit (orange). Panel (c): Normalised observed spectrum (black) compared to the model spectrum (blue). Residuals (red) can then be used as input for the likelihood function.

As we compare the observation to model spectra, we do not have to restrict ourselves to an a priori normalisation, but can take into account the residual information on the continuum

in parts of the spectra. For each model spectrum that we compare to, we therefore perform a normalisation by fitting a Chebyshev polynomial with outlier clipping to the ratio of model and observation (see Fig. 6). This allows us to both overcome previous shortcomings of the synthetic analysis in GALAH+ DR3 (Buder et al., 2021), which had to be restricted to small wavelength segments and assumed a linear relation for those. Our new approach allows us to properly assess the structure of deep and steep molecular features for cool stars, which dominate spectra of cool stars and carry significant information on T_{eff} as well as v_{rad} .

4.3 Stellar label optimisation

In up to four major loops, we optimise the radial velocities and all other stellar labels and report a) their values, b) their co-variances, c) the best fit synthetic and re-normalised spectra along with their uncertainties and masks that indicate which pixels were used in the final optimisation.

Starting from the initial values, a first synthetic spectrum is computed and compared with the observation in order to assess the initial radial velocity. This is done by applying the `SCIPY.SIGNAL.FIND_PEAKS` algorithm on the residuals^d of non-shifted observed and synthetic spectra, when the latter is shifted by $v_{\text{rad}} = -1000..(2)..1000 \text{ km s}^{-1}$ (see Fig. 7a). If no peak is found, the initial v_{rad} value is used hereafter. If more than one peak is found (see Fig. 7 with *Gaia* DR3 agreeing with the systemic radial velocity), the two strongest peaks are reported and trigger a stellar label estimation assuming binary star (Sec. 6). For the purpose of the single star analysis, a narrower search is conducted around the highest peak with a v_{rad} shift of $-20.00..(0.04)..20.00 \text{ km s}^{-1}$ around said peak by fitting a Gaussian function to the inverse of the residuals that were normalised with the smallest residual values (see Fig. 7c). The mean of this fit and its uncertainty are reported by the pipeline.

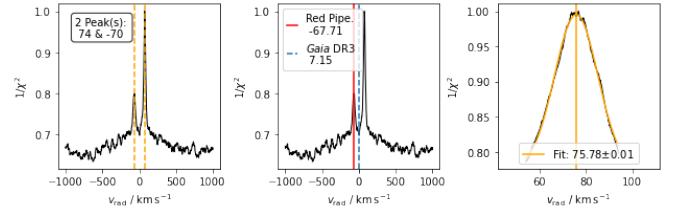


Figure 7. Output of the radial velocity fitting module. Panel a) shows the initial broad search on a v_{rad} array of $-1000..(2)..1000 \text{ km s}^{-1}$. In the case of 2MASS J06084657-7815235, two peaks (yellow, dashed) are visible for this line-splitting spectroscopic binary. Panel b) shows the same plot, but overlaid with the GALAH DR4 reduction pipeline (red) and *Gaia* DR3 (blue, dashed) estimates. Panel c) shows the narrow window of $-20.00..(0.04)..20.00 \text{ km s}^{-1}$ around the highest peak and the Gaussian fit (yellow) to it.

The centerpiece of our optimisation is the `SCIPY.OPTIMISE` routine `CURVE_FIT` module (Virtanen et al., 2020), which we call with counts and uncertainties (our absolute sigmas) as input for a placeholder function that self-consistently re-normalises

^dWe use the inverse of the residuals that were normalised with the smallest residual values.

the observed spectrum. We estimate the labels via the least squares optimisation within less than 10^4 iterations and a desired relative error (`xto1`) below 0.0001.

For each optimisation loop, a new, best-fit 3D bin and neural network is identified via a grid search in the T_{eff} , $\log g$, and [Fe/H] dimensions with `SKLEARN.cKDTree`. If the stellar labels that are being fitted have changed (for example if an element is deemed not detectable for the new 3D bin), the label and its value is either deleted or initialised with $[\text{X}/\text{Fe}] = 0$.

While the optimisation has not converged (the final parameters T_{eff} , $\log g$, and [Fe/H] are not within the current 3D bin), the optimisation is repeated, starting with the previous best-fit parameters as starting guesses.

4.3.1 Which labels are optimised?

As part of the synthetic grid computations, we have perturbed each label individually to our chosen maximum and minimum ranges (see Sec. 3.1). This allows us to also judge which stellar labels to fit for each given star. We choose to fit a stellar label, if either of the two cases applies to said label for the GALAH wavelength range when neglecting the cores of the Balmer lines: Does the spectrum between minimum and maximum label value at any pixel change more than a certain threshold (0.07 of the normalised spectrum)? Does the spectrum between minimum and maximum label value change by more than 0.005 of the normalised spectrum for at least 25% of the spectrum? While the first case is constructed for atomic lines, such as Li I 6708, the second case is addressing in particular molecular lines like the C₂ and CN lines. For spectra that are missing the infrared arm (CCD4) we exclude the otherwise possibly fit labels for N, O, L, and Rb.

Initial tests of the pipeline have revealed that in cases where the initial parameter guesses are deviating significantly from the final ones, several elemental abundance estimates were shifted towards their boundaries, leading to a masking of their elemental abundances lines by the masking module (Sec. 4.3.2) at the beginning of each optimisation loop. To minimise this effect, we therefore shift the interim abundance values towards the narrow label boundaries. In practise, we this limit the initial and interim abundances to 1.05..3.26 for A(Li), $[\text{X}/\text{Fe}] = -0.5..1.0$ for C, N, O, Y, Ba, La, Ce, and Nd, and $[\text{X}/\text{Fe}] = -0.5..0.5$ for all other elements before optimising them again. For warm and hot stars ($T_{\text{eff}} > 6000$ K), this effect was seen to effect multiple abundances, such that we needed to implement a zeroing of all abundances (except Lithium) for stars above 6000 K, which would on average be expected to be young and have a solar-like composition.

4.3.2 Masking of unreliable wavelength regions

Not all pixels of the observed or synthetic spectra might prove useful for estimating reliable stellar labels. Observations can include bad pixels/patterns and incorrect corrections (for example of telluric or sky lines). Flux predictions of synthetic spectra are only as good as the input physics (limited for example for specific lines via uncertain oscillator strengths).

To minimise the influence of inaccurate synthetic pixel predictions, we have compared a 2dF-HERMES observation of the asteroid 4 Vesta and a high-quality Solar spectrum by Hinkle et al. (2000) with the flux that would be predicted through our pipeline for a star with Solar labels ($T_{\text{eff}} = 5772$ K, $\log g = 4.438$ dex, $[\text{Fe}/\text{H}] = 0.00$ dex, $v_{\text{mic}} = 1.06 \text{ km s}^{-1}$, $v \sin i = 1.6 \text{ km s}^{-1}$, $v_{\text{mac}} = 4.2 \text{ km s}^{-1}$ (Prša et al., 2016; Jofré et al., 2017), and $[\text{X}/\text{Fe}] = 0.00$ dex for the default Solar abundance pattern for MARCS by Grevesse et al. (2007)).

We have identified all lines that showed differences of the normalised flux of more than 0.1, lines where either a synthetic line or an observed one was completely missing, or lines that were significantly misaligned. Examples of this list of masks are shown in Fig. 8. To avoid the influence of bad spectrum regions with an observational origin, we mask pixels where the synthetic and re-normalised observed spectrum differ by more than 5σ or a flux of 0.3 (0.4 before the initial optimisation). To avoid the masking of lines that are vital for our spectroscopic analysis, we have created a list with segments of such lines that is mainly based on the previous element masks from GALAH DR3 (Buder et al., 2021). The final mask of pixels to use for the optimisation then includes all vital line regions and those wavelengths that are not unreliable in the synthesis show a too strong disagreement between observation and synthesis.

In addition to this default masking, we exclude pixels for each major iteration, for which the flux of observation and synthetic differ by more than 5σ and 0.3 of the normalised flux and by more than 1 for the vital line regions.

We further indirectly take into account the currently less constrained molecular data for cool stars in optical spectra, in particular towards the blue (e.g. Rains et al., 2021). For presumably cool stars (with initial $T_{\text{eff}} < 4100$ K), we therefore double the observational uncertainty of the blue arm.

5. SINGLE STAR ANALYSIS (ALLSTAR)

After the `allspec` module (Sec. 4) was used to estimate spectroscopic labels for all spectra, we use the `allstar` module to co-add spectra and analyse one spectrum per star while taking into account photometric and astrometric information to constrain the surface gravities. This approach was already very successfully applied for GALAH DR3 (Buder et al., 2021) with *Gaia* DR2 distances (Bailer-Jones et al., 2018) to overcome spectroscopic degeneracies. For the co-adding, we test if the radial velocity estimates of individual exposures agree within 2σ . Below this threshold, we apply no radial velocity correction and fit a global radial velocity. Above this threshold (which is useful for spectroscopic binaries of type 1 without line splitting as shown in Fig. 9), we apply a radial velocity correction before the co-adding.

To speed up computation, we use the mean results of the `allspec` analyses as initial stellar labels. All other methodology of the comparison of synthetic spectra to observations (Sec. 4.2 and label optimisation (Sec. 4.3) apply also to this module, with the exception of the optimisation of $\log g$. Contrary to the `allspec` approach, we do not fit $\log g$ in this module, but estimate the logarithmic surface gravity, $\log g$, as a combination of its

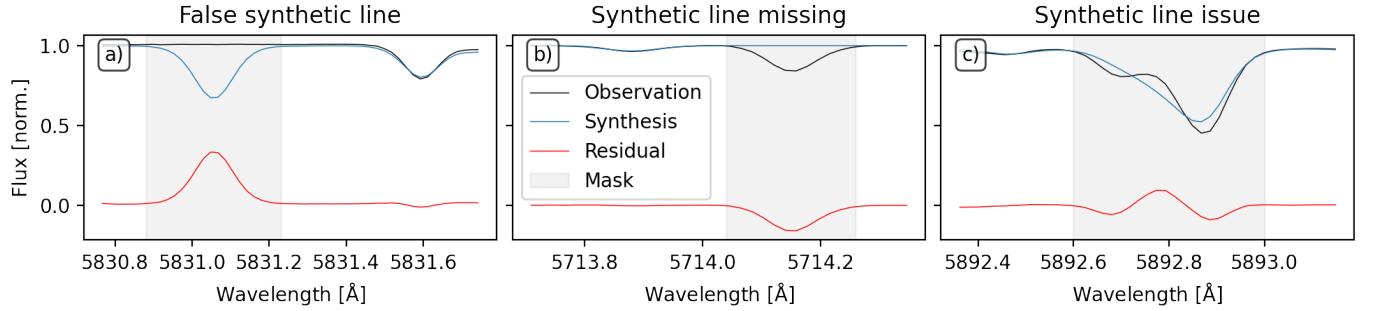


Figure 8. Examples of masks applied to unreliable pixels for the spectrum fitting, that is the minimisation of residuals (red) between observation (black) and synthesis (blue). **Panel a)** showing a strong synthetic line, where no line is observed in the Sun. **Panel b)** showing an observed line without any line being synthesised. **Panel c)** showing significant disagreement between the two observed lines and the synthesis.

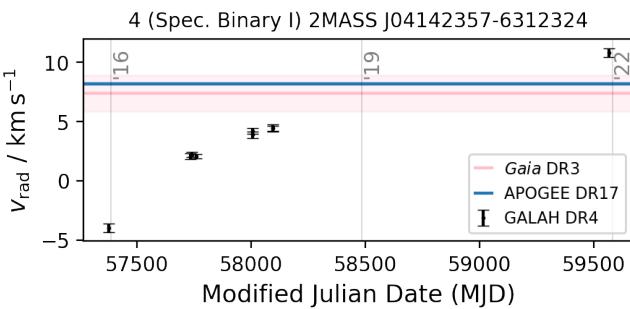


Figure 9. Example of the radial velocity evolution over time for a spectroscopic binary (SB) of type I (no line split).

definition ($g \propto \frac{M}{R^2}$) and the Stefan-Boltzmann law relative to the Solar values:

$$\log g = \log g_{\odot} + \log \frac{M}{M_{\odot}} + 4 \log \frac{T_{\text{eff}}}{T_{\text{eff},\odot}} - \log \frac{L_{\text{bol}}}{L_{\text{bol},\odot}} \quad (8)$$

While we can use our spectroscopically determined T_{eff} , the other values have to be estimated through models or non-spectroscopic information. The logarithmic bolometric luminosity, L_{bol} , can be estimated from the bolometric magnitude, such that $\log \frac{L_{\text{bol}}}{L_{\text{bol},\odot}} = -0.4 \cdot (M_{\text{bol}} - M_{\text{bol},\odot})$. The bolometric magnitude can be estimated from any given apparent magnitude, if we correct the latter by the distance modulus, bolometric correction, and extinction. Because for essentially all stars of GALAH DR4, we have outstanding infrared magnitudes available which suffer less from (uncertain) extinction corrections, we continue to chose K_S as the magnitude to estimate our bolometric magnitudes and luminosities via

$$M_{\text{bol}} = K_S - 5 \cdot \log \frac{D_{\infty}}{10} + BC(K_S) - A(K_S). \quad (9)$$

While the values for K_S , D_{∞} , and $A(K_S)$ are readily available (see Sec. 2.3), we need to estimate the bolometric correction from tabulated values using the routines provided by Casagrande & Vandenberg (2018):

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

We choose to assume an extinction value of $E(B-V) = 0$ mag for this particular interpolation and post-correct the value by $A(K_S)$ our actual extinctions can exceed the tabulated values of $E(B-V) = 0.72$ mag by Casagrande & Vandenberg (2018) and our available $E(B-V)$ values from Schlegel et al. (1998) are upper limits that can overestimate the actual extinction of our observed targets.

Because of the appearance of $\log g$ in Eq. 10, we iterate the calculation of $BC(K_S)$ and subsequently $\log g$ up to four times or until the latter value changes less than 0.02 dex between iterations. Similarly, we need to estimate the stellar masses (and ages as a byproduct) from tabulated values, that is,

$$\mathcal{M}, \tau = f(T_{\text{eff}}, \log g, [\text{Fe}/\text{H}], L_{\text{bol},\odot}) \quad (11)$$

For this on-the-fly estimate of masses and ages we use an earlier version of the ELLI code by Lin et al. (2018) for a likelihood-weighted estimate with default uncertainties of 100 K, 0.25 dex, 0.2 dex, and an average uncertainty of $L_{\text{bol},\odot}$ from propagated uncertainties of Eq. 9.

We interpolate over the default tables of PARSEC+COLIBRI isochrones (Bressan et al., 2012; Marigo et al., 2017) from that cover the logarithmic ages of $\log(\tau / \text{Gyr}) = 8.00..(0.01)..10.18$ by default and metallicities $[\text{M}/\text{H}] = -2.75..(0.25)..-0.75$ as well as $[\text{M}/\text{H}] = -0.6..(0.1)..0.7$. We exclude hot stars above 10 000 K as well as extremely evolved stars ($\log g > 6$ dex or $J - K_S > 2$ mag) as they fall far outside our spectroscopic pipeline range. We convert between the theoretical $[\text{M}/\text{H}]$ and our measured $[\text{Fe}/\text{H}]$ as well as $[\alpha/\text{Fe}]$ via the correlation by Salaris & Cassisi (2006), $[\text{M}/\text{H}] = [\text{Fe}/\text{H}] + \log(10^{[\alpha/\text{Fe}]} \cdot 0.694 + 0.306)$. For open clusters with age estimates below 1 Gyr as well as unevolved stars that are more luminous than expected from the oldest cool main sequence isochrone with matching $[\text{M}/\text{H}]$, we sample $\log(\tau / \text{Gyr}) = 6.19..(0.01)..10.18$. For globular cluster stars identified in the crossmatch with Baumgardt & Vasiliev (2021), we limit the isochrones to a minimum age of 4.5 Gyr.

6. BINARY STAR ANALYSIS (BINARY)

This module is motivated by the extensive study of GALAH binary star spectra by Traven et al. (2020) and our ability to

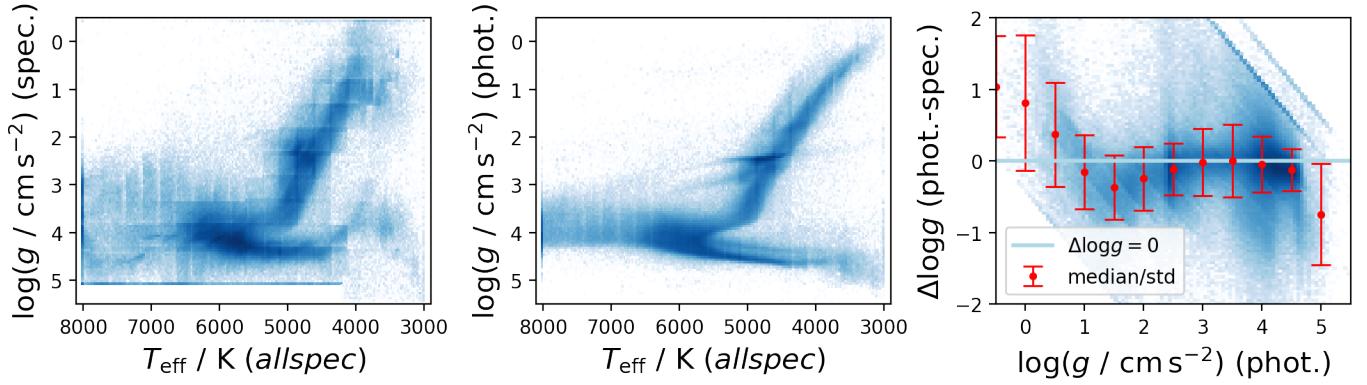


Figure 10. Comparison of spectroscopic and photometric $\log g$ estimates of the *allspec* analysis Panel a) shows the distribution of spectroscopic $\log g$ and T_{eff} from the *allspec* module. Panel b) shows the distribution of the same T_{eff} and photometric $\log g$. Panel c) shows the difference of photometric $\log g$ and spectroscopic $\log g$ as a function of photometric $\log g$. Red error bars indicate the 1σ percentiles of this difference in 0.5 dex bins.

model the full spectrum via neural networks. It is activated, if the post-processing of a spectrum for binary signatures in the residual spectra or differences to *Gaia* DR3 radial velocities (Sec. 7.1.1) warrant a closer inspection of binarity.

The module is still in development, but already now delivers at least useful radial velocity estimates `rv_comp_1` and `rv_comp_2` as well as an overall χ^2 of a binary spectrum, as shown in Fig. 11.

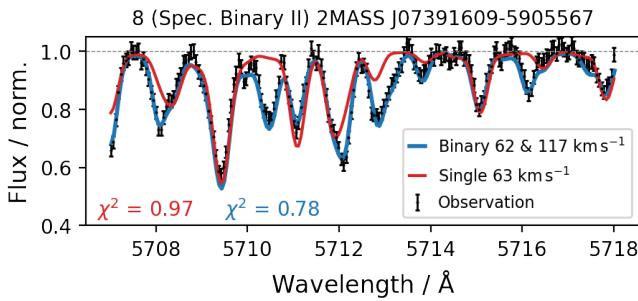


Figure 11. Examples of the spectrum of a line-splitting binary (SB type II) which is better fit with our binary fitting algorithm.

In the current implementation, we superimpose two synthetic spectra with a relative flux contribution factor f_{contr} between 0 and 1. In addition to fitting this flux contribution ratio, we fit $T_{\text{eff},s}$, $\log g_s$, $v_{\text{mic},s}$, $v_{\sin i,s}$ for synthetic spectra of two stars s with the same [Fe/H] to the observed spectrum.

Because we know that the stars most likely share not only the same [Fe/H], but also age τ , another implementation could be to fit a system age and [Fe/H] and two different masses M_s . In this case, $T_{\text{eff},s}$ and $\log g_s$ would be interpolated from the PARSEC isochrones for a given age, mass, and [Fe/H] combination. This could be tricky though because of the turn-off region which may cause local χ^2 minimum.

7. POST-PROCESSING

After the *allspec* (Sec. 4), *allstar* and (Sec. 5) modules have been run for a night, a post-processing routine is used to estimate additional parameters from the residuals of the spectra

(Sec. 7.1), estimate and validate accuracy and precision uncertainties (Sec. 7.2) and perform quality assurance tests on a global scale (`flag_sp`, see Sec. 7.3) as well as for the individual abundances of elements X (`flag_x_fe`, see Sec. 7.4).

7.1 Analysis of spectral residuals

7.1.1 Binary signatures to trigger the binary module

The residual spectrum of our best fitting single star analysis can help us to identify a second flux contributor to the observed spectrum. In our case, there are two points in the analysis, where we can identify such an influence. Firstly, the residuals are visible in the χ^2 distribution as a function of radial velocity shifts (see Fig. 7). While a single star would only show one peak (saved in `rv_comp_1`), a binary system like 2MASS J06084657-7815235 shows a second peak (-70 km s⁻¹ in addition to 74 km s⁻¹) that is saved in the `rv_comp_2`. Secondly, we perform an automatic search for re-occurring residuals as a function of radial velocity for a few selected lines. We chose the combination of strong lines in the spectra (Balmer lines, Fe lines at 4890 and 4891 Å, Ni at 6644 Å) as well as those with the largest expected wavelength shift in the infrared detector (O triplet at 7772 – 7775 Å as well as Mg at 7692 Å). If we find several peaks with a reasonably similar radial velocity, the likely $X \in 16, 50, 84^{\text{th}}$ percentiles of this radial velocities are saved in `sb2_rv_X`^e

Because radial velocities from the *Gaia* radial velocity spectrometer (Katz et al., 2022) are reported in *Gaia* DR3 for 94% (774 914) of the stars observed for GALAH DR4, we can also compare the radial velocity estimates. For 6% (50 577) of our stars, we find difference with respect to *Gaia* DR3 beyond 10 km s⁻¹. For these as well as stars for which we estimate unrealistic v_{mic} and $v \sin i$ below 0 km s⁻¹ in *allspec* or $v \sin i$ above 10 km s⁻¹ or identify a secondary component in `rv_comp_2`

^eShould we better run CCF or the Gaussian fitting to residuals in RV space, as we do during the RV estimation?

^f*allspec* was run without boundary conditions for global parameters and thus also resulted in negative velocities, which are later flagged. *allstar*, however, was run with v_{mic} and $v \sin i$ forced to be above 0 km s⁻¹.

and `sb2_rv_X`, we trigger the run of the *binary* module (see Sec. 6).

7.1.2 Post-correction of $\log g$ for *allspec* results

While we estimate logarithmic surface gravities $\log g$ solely from spectra in the *allspec* results, we also perform a post-processing estimate where we employ the methodology of Sec. 5 while fixing all other stellar parameters. While planned as an initial test, it clearly confirmed the previous conclusions of GALAH DR1–DR3 that the spectroscopic information in HERMES spectra to estimate $\log g$ is not sufficient for the majority of the parameter space. We show the spectroscopic $\log g$ in Fig. 10a and the photometric $\log g$ and their difference in Figs. 10b and c, respectively.

We see an overall good agreement of both $\log g$ estimates for stars between $4250 < T_{\text{eff}} < 6500$ K. Hotter stars show a strong dispersion of spectroscopic $\log g$ due to limited information from fewer and shallower lines. Cooler stars show a significant trend towards much lower $\log g$ for main sequence stars and much higher $\log g$ for cool evolved stars up to an order of $\Delta \log g$ of 1 dex. This trend was previously seen in GALAH DR2 (Buder et al., 2018) and is believed to be caused by the onset of molecular absorption features which suppress the continuum for almost the entire HERMES wavelength range (see for example Fig. 6), thus introducing several degeneracies^g. In addition, we can notice a significantly lower precision of the spectroscopic $\log g$ in comparison to the excellent precision of photometric $\log g$, for example of the red clump stars.

At closer inspection, we notice several trends in Fig. 10a. Most notably, we see nodding patterns along the T_{eff} and $\log g$ grids where the *allspec* module switches between different neural network models. Our investigation of these nodding effects is addressed in Sec. 10. In comparison to Fig. 10b, where a clear same mass binary sequence is showing just above the cool main sequence, we do not see such a sequence in Fig. 10a. The difference between spectroscopic and photometric $\log g$ will therefore be useful to identify photometric binaries at least for the high quality spectra (where $\log g$ precisions are below the single to binary system offset of $\Delta \log g = 0.3$ dec), as discussed in Sec. 7.3.

7.1.3 Interstellar absorption

Because we can create synthetic stellar spectra for the full wavelength range, we can now also trace interstellar absorption in the residuals of observed spectra. By default, we try to calculate the equivalent width via Gaussian fits to the three diffuse interstellar bands (5780.59, 5797.19, 6613.66 Å) with central wavelengths identified by Vogrinčič et al. (2023) as well as for interstellar K (7698.9643 Å), see Fig. 12. We report the equivalent widths `eq_x`, standard deviations `sigma_x` and radial velocities `rv_x` for x in `k_is` for interstellar K and x in `DIB_5780`, `DIB_5797`, and `DIB_6613`.

^gGet back to this and include the private communication with Charlie Conroy on the $\log g$ trends for cool stars.

7.1.4 Emission estimates for the Balmer lines

The difference between synthetic and observed Balmer line absorption holds valuable information both on emission stars as well as the known inaccuracy of the synthetic Balmer lines and possibly even information on unresolved binary systems (Sayeed et al., in prep.) as well as masses for evolved stars (Bergemann et al., 2016). We therefore perform a trapezoidal integration around the Balmer line at 4861.3230 and 6562.7970 Å whose values we report in `ew_h_beta` and `ew_h_alpha`. By default we integrate in a window of ± 0.75 and 1.25 Å for H_{β} and H_{α} , respectively, and increase this window to 5 Å if the average observed, normalised flux within ± 0.5 Å of the Balmer line core exceeds 1. An example of such a star is shown in Fig. 13, for which we measure a residual EW of 1.09/AA.

7.2 Uncertainty estimation and validation

The uncertainties that we report for our spectroscopic data analysis are based on the covariance matrix of the fitting process. While we report the covariance matrix for each spectrum (see Fig. 23 for the covariance matrices of VESTA and Arcturus), we restrict ourselves to the standard deviations of each feature (square root values of the diagonal variance entries).

Table 4. List of accuracy and representative precision uncertainties for stellar parameters of GALAH DR4. Accuracy values are estimated from comparisons with literature references (see Fig. 16). Precision values are extracted from the median covariance uncertainties for stars with $SNR = 50 \pm 10$ in CCD2 (see Fig. 21).

Parameter / Unit	Accuracy	Precision ($SNR = 50$)
$T_{\text{eff}} / \text{K}$	66	23 ± 5
$\log(g / \text{cm s}^{-2})$	0.042	-
$[\text{Fe}/\text{H}] / \text{dex}$	0.064	0.025 ± 0.004
$v_{\text{mic}} / \text{km s}^{-1}$	0.28	0.05 ± 0.03
$v \sin i / \text{km s}^{-1}$	1.4	0.5 ± 0.2
$v_{\text{rad}} / \text{km s}^{-1}$	0.15	0.17 ± 0.02

7.2.1 Accuracy estimation and validation

Estimating the accuracy of spectroscopic measurements has always been a complicated endeavour, because there are no universal benchmark sets for all parameters. Subsequently, we there describe the numerous comparisons that we have performed for both stellar parameters (T_{eff} , $\log g$, $[\text{Fe}/\text{H}]$, v_{mic} , $v \sin i$, and v_{rad}) as well as the elemental abundance measurements. Due to the limited coverage of benchmark literature, we continue to apply a universal accuracy estimate for the stellar parameters and no accuracy estimate for the elemental abundances (as for GALAH DR3 Buder et al., 2021).

Fig. 16 for stellar parameters and Fig. 19 for abundances.

- *Gaia* FGK Benchmark stars (Jofré et al., 2018) Fig. 16a–c
- IRFM (Casagrande et al., 2020)
- K2/TESS Zinn et al. (2020) Fig. 16b
- APOGEE DR17 Fig. 16a–f
- *Gaia* DR3 for v_{rad} Fig. 16f

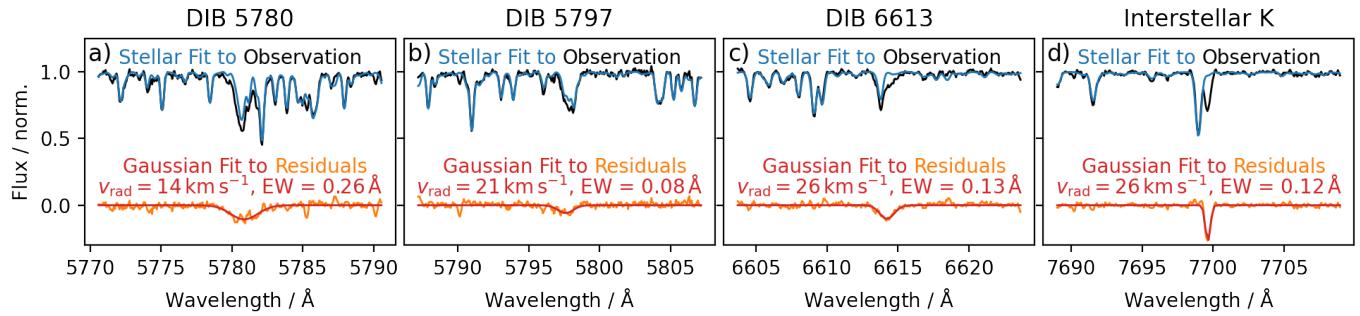


Figure 12. Example of interstellar diffuse band (DIB) and interstellar K measurements for 2MASS J06453479-0102137 with $E(B-V) = 0.84$ mag from Schlegel et al. (1998).. Shown are the observation (black) and stellar fit (blue) as well as a Gaussian fit (red) to the residual (orange), resulting in an estimate of the equivalent width (EW) as well as radial velocity.

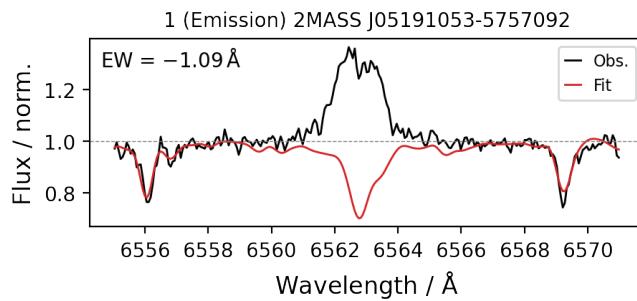


Figure 13. Example of a stars with clear emission in the Balmer lines (here H_β) and raised major quality flag flag_sp for emission.

- Globular Clusters for low [Fe/H] accuracy Fig. 20
 - 47 Tuc Carretta et al. (2009c,a)
 - NGC 6362
 - M4
 - NGC 362 Monty et al. (2023)?
 - NGC 288 Monty et al. (2023)?
 - NGC 1851
 - M2 NGC 7089, with 3 populations in [Fe/H] according to Yong et al. (2014). 21331790-0048198 was even observed by Yong et al. (2014) as NR 79 / HI-104 with [Fe/H] = -1.69 . We get -1.37 ± 0.26 , but at SNR around 10.
 - Omega Cen Johnson & Pilachowski (2010)
 - NGC 6397 Carretta et al. (2009c,a)
 - M30 Carretta et al. (2009c,a)
- Abundance Zeropoints
 - VESTA
 - APOGEE DR17
 - Solar Twins
 - M 67, Ruprecht 147 and a few others cover stars across several evolutionary stages. Check them roughly (could also be problematic because of atomic diffusion etc.)

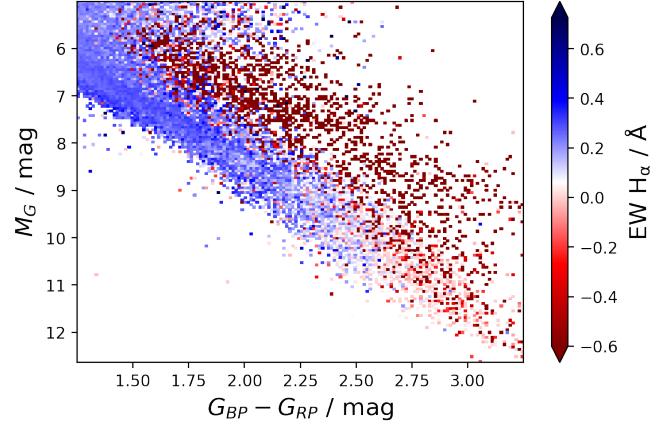


Figure 14. Distribution of residual equivalent width between observation and synthesis in the Gaia DR3 color-absolute magnitude diagram ($G_{\text{BP}} - G_{\text{RP}}$ vs. M_G) for cool dwarfs. Balmer lines in emission result in negative color values, as is the case for most young stars that are more luminous than the cool main sequence. Should we point out the interesting sequence of binary MS around $M_G \sim 7$ mag?

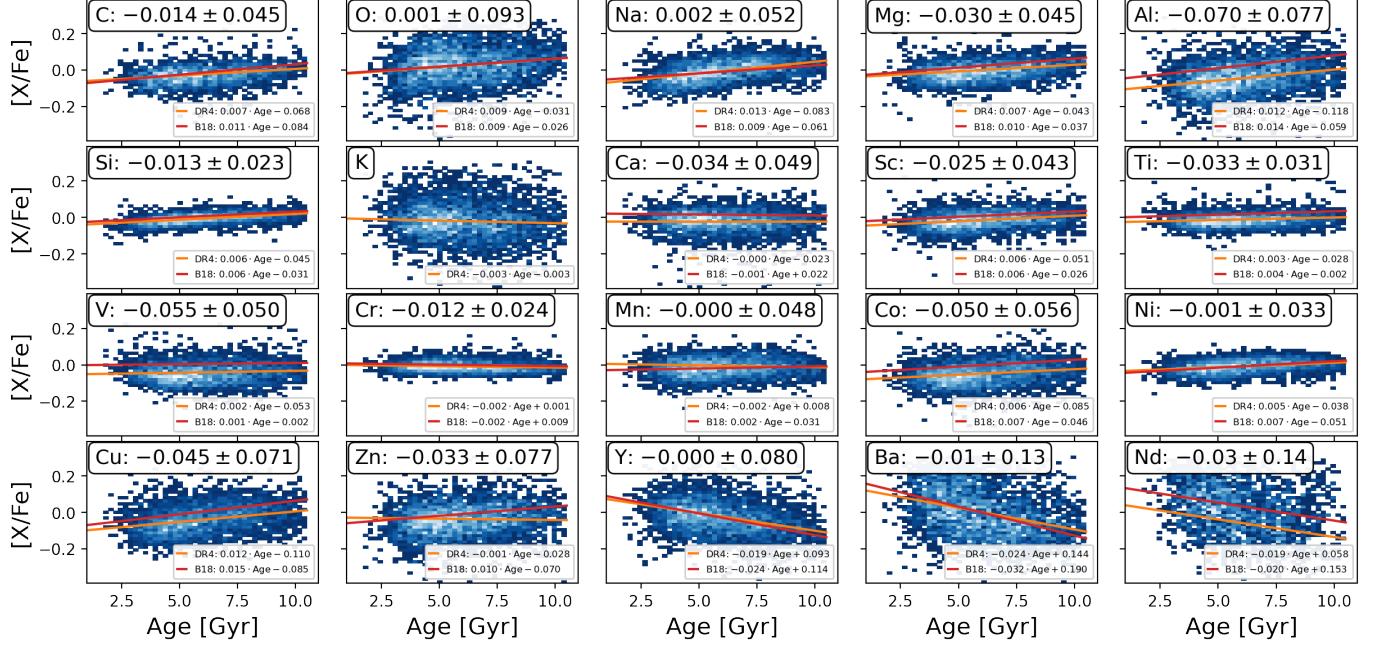


Figure 15. Chemical abundances $[X/Fe]$ of Solar twin stars as a function of on-the-fly computed stellar age. We overplot linear fits to our age-abundance relations for Solar twins in orange as well as the literature values from ? in red.

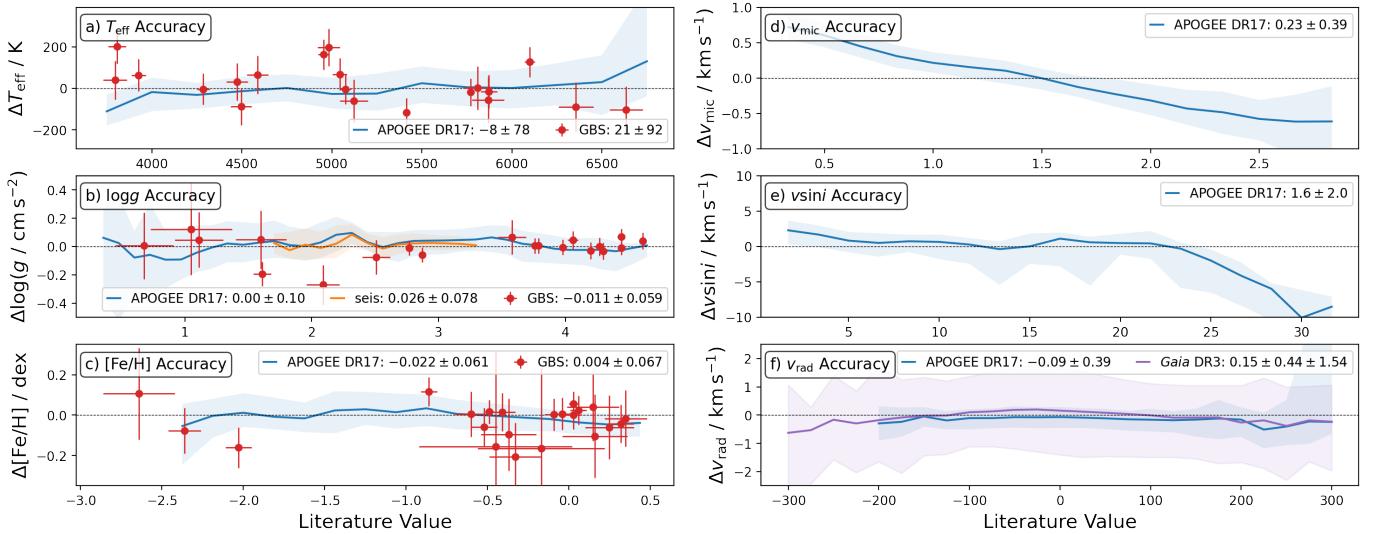


Figure 16. Accuracy of the main stellar parameters T_{eff} , $\log g$, $[\text{Fe}/\text{H}]$, v_{mic} , $v \sin i$, and v_{rad} for GALAH DR4.. Each panel shows the comparison to literature (DR4 - literature). Comparisons are performed for the Gaia FGK Benchmark stars (red), APOGEE DR17 (blue), $\log g$ inferred from asteroseismic measurements (orange) and Gaia DR3 radial velocities (purple).

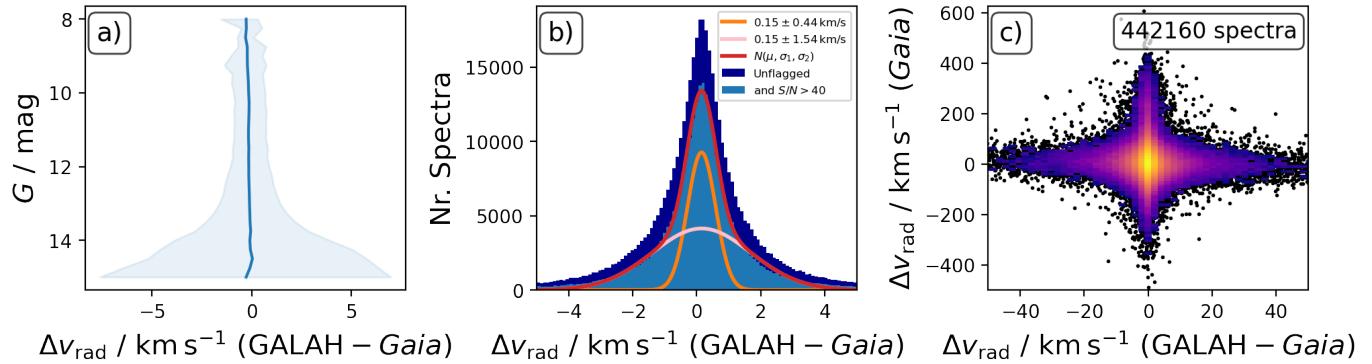


Figure 17. Comparison of radial velocities between GALAH DR4 and Gaia DR3. Panel a) shows the difference of radial velocities as function of Gaia's G photometry. Panel b) shows a histogram of the difference with two Gaussian distributions (with same mean) fitted to them to estimate a more robust, that is binary independent, radial velocity difference. Panel c) shows the difference of radial velocities as function of radial velocity, showing the systematic scatter introduced by binaries.

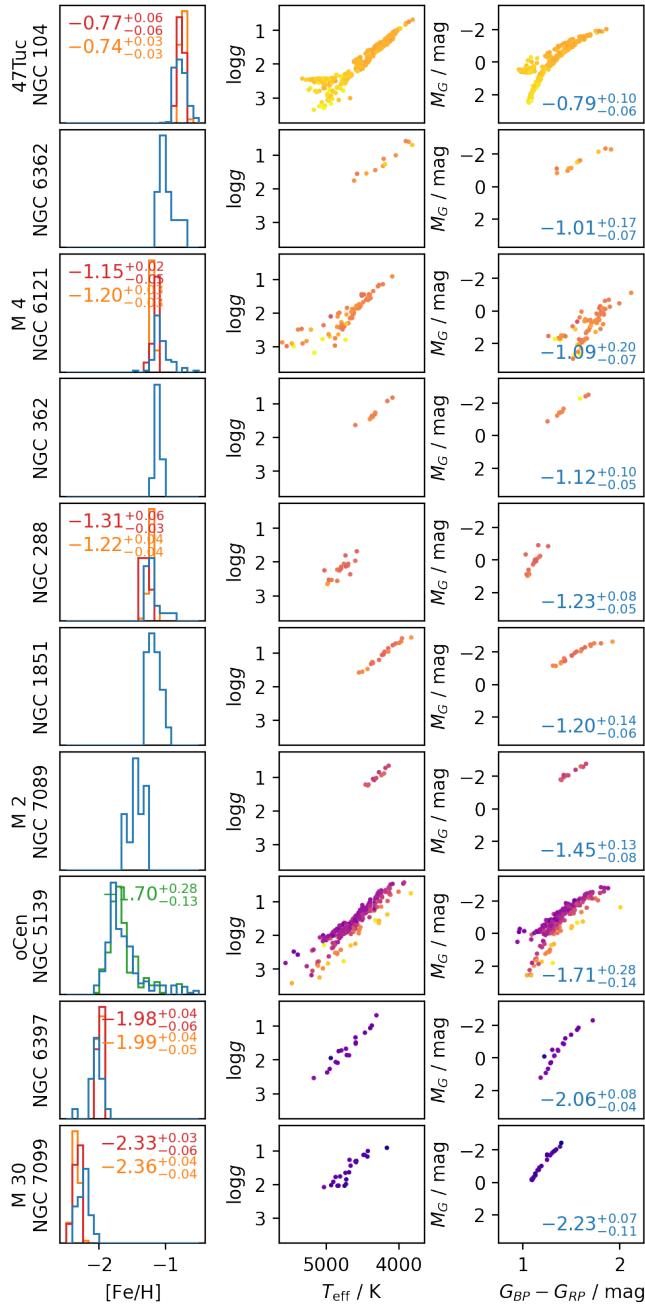


Figure 18. Comparison of iron abundances (16th, 50th and 84th percentiles) and overview of spectroscopic and photometric properties of globular cluster stars in GALAH DR4. Left panels show histograms of iron abundances from GALAH DR4 (blue) as well as literature estimates for the globular clusters from Giraffe (orange) and UVES (red) observations by (Carretta et al., 2009c,b) as well as observations from Johnson & Pilachowski (2010). Middle panels show the spectroscopic T_{eff} - $\log g$ diagrams colored by iron abundance $[\text{Fe}/\text{H}]$. Right panels show the Gaia DR3 color absolute magnitude diagrams.

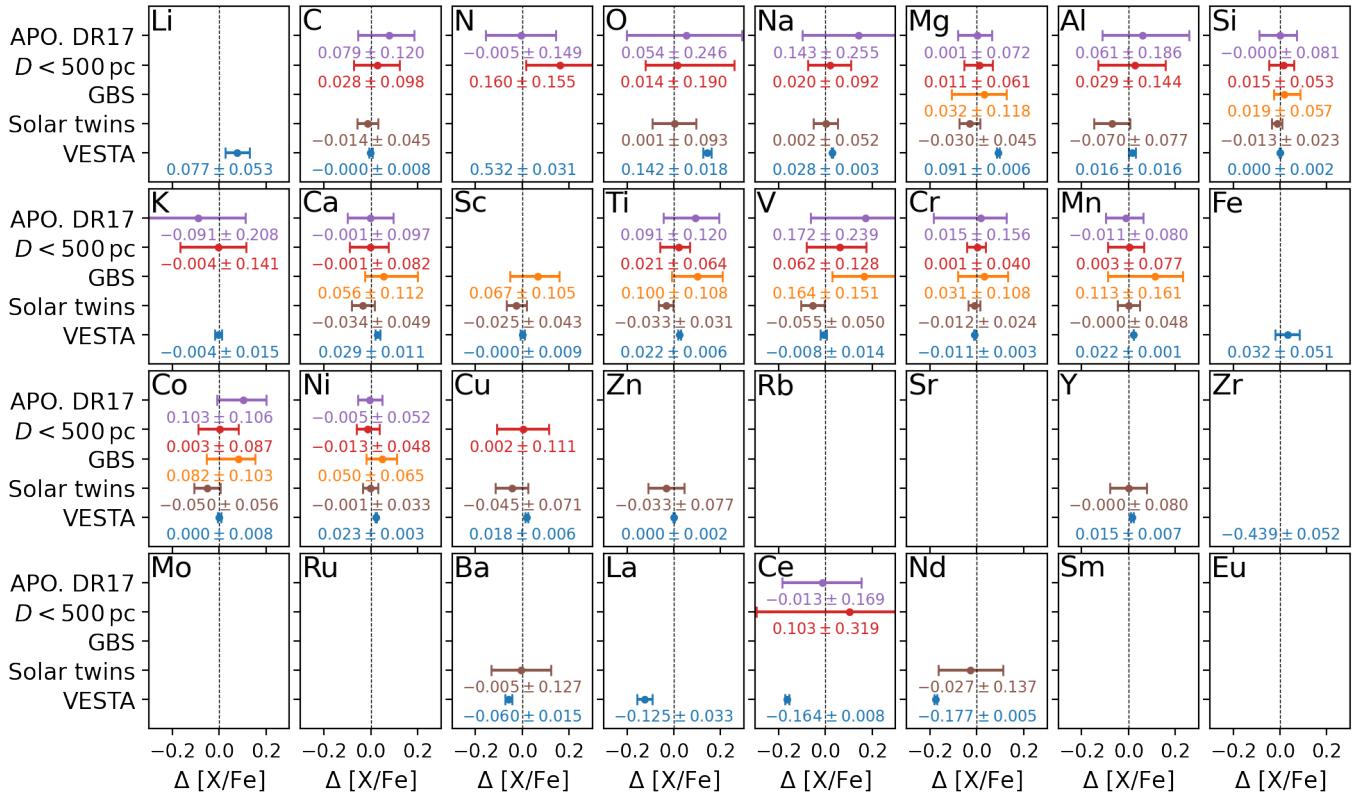


Figure 19. Zeropoint estimates of elemental abundances for GALAH DR4.. Each panel shows the comparison to literature (DR4 - literature) for VESTA (blue), Gaia FKG Benchmark Stars (orange), Stars with $|[Fe/H]| \leq 0.1$ closer than $D_\omega < 0.5$ kpc (red), as well as the overlap with APOGEE DR17 (purple).

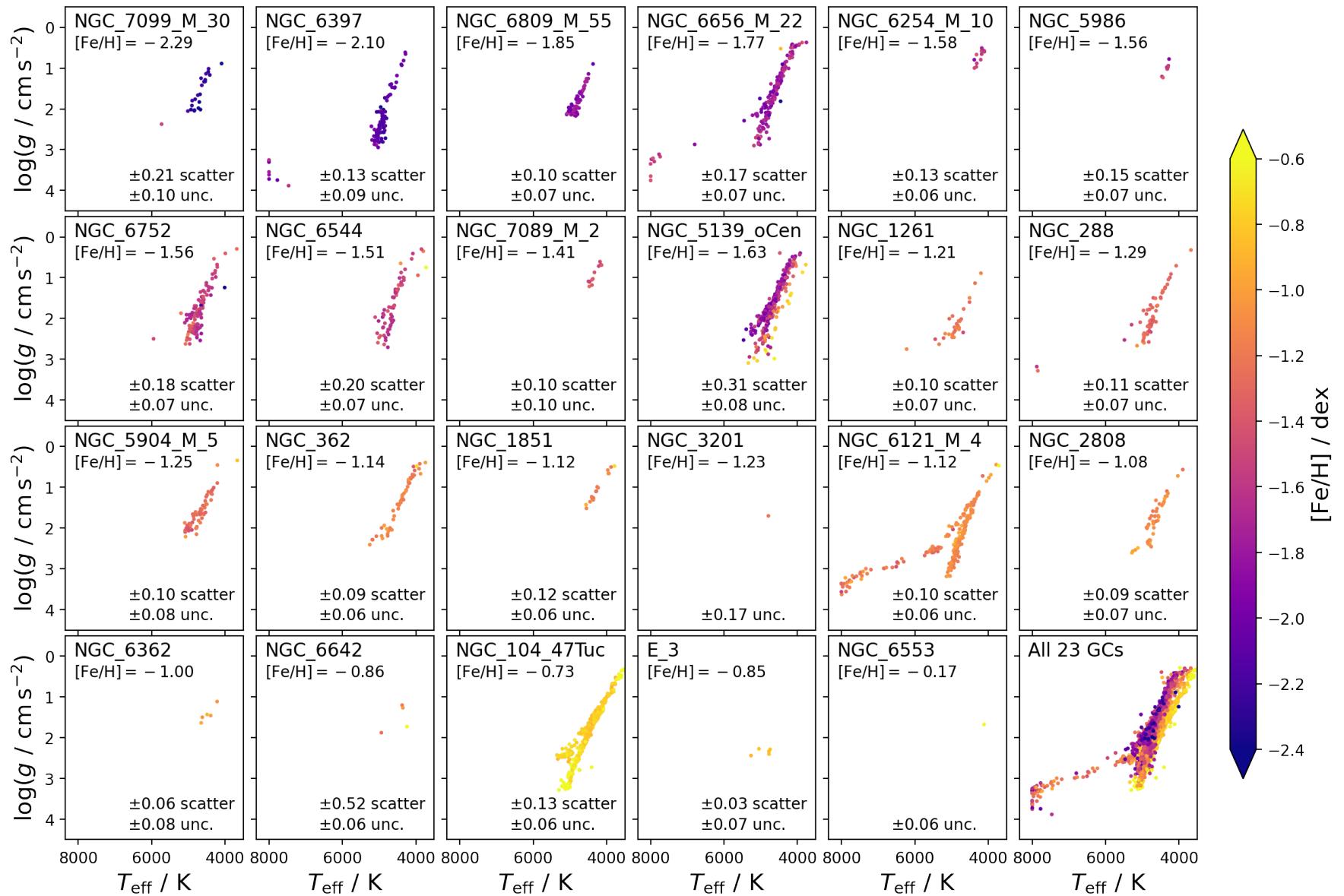


Figure 20. Collage of Globular Clusters in the T_{eff} - $\log g$ space, colored by stellar abundance $[\text{Fe}/\text{H}]$. Multiple cluster, most pronounced NGC 5139 (!Cen) show a spread in $[\text{Fe}/\text{H}]$, whereas others are consistent with no spread within the uncertainties when comparing scatter with median uncertainties in the lower right of each panel.

7.2.2 Precision estimation and validation

- Fig. 21 and 22
- Create Table with overall scatter values for individual elements based on:
- Covariance uncertainties
- Repeat observations
- Scatter in open clusters.
- Wide binary scatter

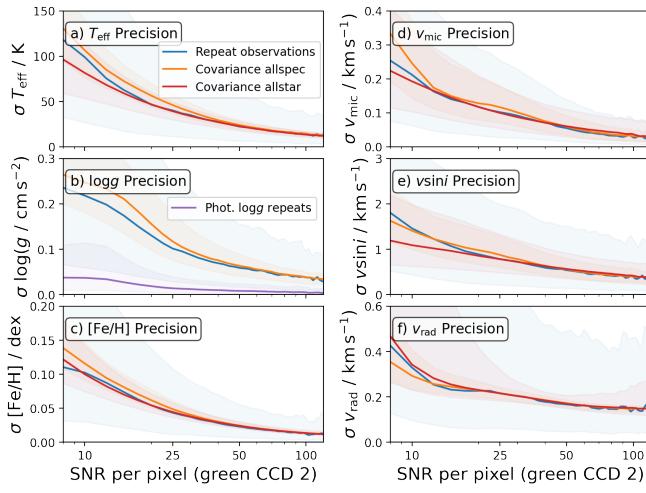


Figure 21. Precision monitoring of stellar parameters as a function of SNR for the green CCD2 across for GALAH DR4.. Each panel shows the behavior for bins of size 10 for the scatter of repeat observations of the *allspec* runs (blue), covariance uncertainties of *allspec* (orange) and *allstar* (red) setups as well scatter of photometric $\log g$ from repeat observations (purple).

Giribaldi Titans [Fe/H] stars

7.2.3 Covariances

We also calculate and report the covariances of all fitted labels, see Fig. 23 for VESTA and Arcturus.

7.3 Stellar parameter flags `flag_sp`

We have implemented a number of post-processing routines for quality assurance. Here, we describe the major quality flags of global parameters and overall fitting performance. If a spectrum does not pass one of these tests, a unique bit flag will be raised and added to `flag_sp`.

The description of the implemented bits/flags for `flag_sp` and how often they were raised is listed in Tab. 5 and distributions in the Kiel diagram (T_{eff} and $\log g$) are shown for each raised bit in Fig. ?? for the *allstar* catalogue.

Describe how the vsini and vmic flags are quite informative about binarity. Look at the binary main sequence for that!

SB2 warning currently often activated for luminous RGB stars. Is that maybe caused by detectable extinction? Inspect more examples and figure out what is going on!

Why are the chi2 stars mainly on the RGB? We use a strict cutoff value rather than an SNR dependent one. Maybe implement an SNR dependent on (GBS stars are also all flagged because of this issue)?

Fig. 13 and Fig. 9

Fig. ??

Table 5. List of major quality flag `flag_sp` listing the bit, description and how often the flag was raised for the *allstar* and *allspec* routines. Notes: Multiple bits can be raised for each of the 943 654 spectra spectra of the of 827 288 stars. (*) No $\log g$ could be estimated for 15,319 spectra in the post-processing due to missing astrometric information.

Raised Bit	Flag	Description	<i>allstar</i>	<i>allspec</i>
	0	No flag	166764	633322*
0	1	Emission	1576	6490
1	2	CCD missing	221	48787
2	4	SB1	9292	0
3	8	SB2	6132	31404
4	16	$\chi^2 > 3\sigma$	11552	58862
5	32	$v \sin i$ warning	17325	118341
6	64	v_{mic} warning	15495	81621
7	128	Triple Binary warning	0	0
8	256	T_{eff} warning	0	0
9	512	$\log g$ warning	0	0
10	1024	[Fe/H] warning	0	0
11	2048	S/N low	17323	89881
12	4096	not converged	0	18042
13	8192	Model extrapolated	5316	56742
14	16384	No Results	2835	6599

7.4 Elemental abundance flags `flag_x_fe`

Table 6. List of elemental abundance quality flags `flag_fe_h` for [Fe/H] or `flag_x_fe` for element X.

Raised Bit	Flag	Description
	0	detection
0	1	upper limit
1	2	no measurement available
2	4	no convergence
3	8	measurement above limit
4	16	measurement below limit
5	32	measurement issue of CNO
6	64	measurement of Li, Ca, or Ba

7.5 Abundance detection or upper limit

To assess if the abundance estimates are a true detection or an upper limit for element X, we produce synthetic spectra with the same stellar labels, but decrease the abundance [X/Fe] to the lower limit of the neural network (or decrease the [Fe/H] value by -0.5 for Fe). The residuals in units of sigma between the best-fit spectrum and spectrum with lowest possible [X/Fe] or lowered [Fe/H] then allow us to identify a detection (with maximum differences beyond 3 sigma) or upper limits, for which we raise the flag `FLAG_X_FE` by 1. We note again that our initial test of overall detectability (Sec. 4.3.1) allowed us identify

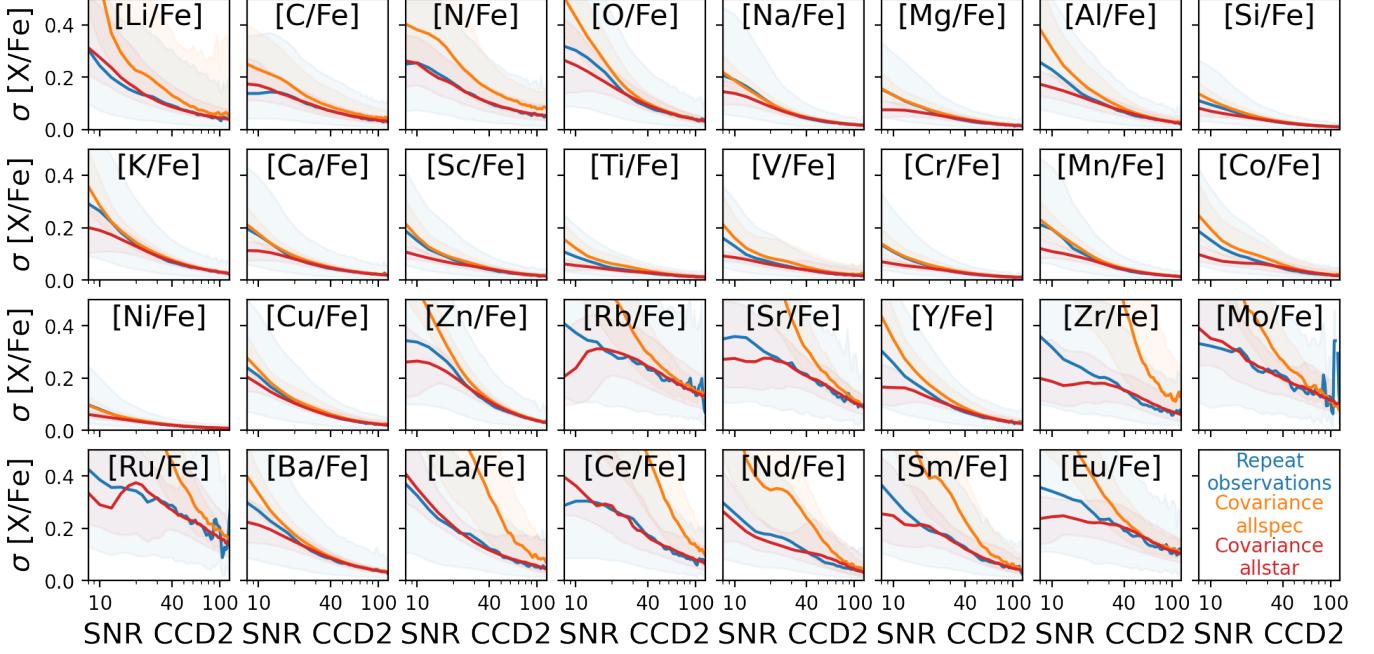


Figure 22. Precision monitoring of elemental abundances as a function of SNR for the green CCD2 across for GALAH DR4.. Each panel shows the behavior for bins of size 10 for the scatter of repeat observations of the allspec runs (blue) as well as covariance uncertainties of allspec (orange) and allstar (red) setups. Missing legend

elements for which not even an upper limit was expected, raising flag `FLAG_X_FE` by 2.

8. VALIDATION

TESS CVZ asteroseismic info (?) and tables from Hon_2022_AJ_164_fits

8.1 Automatic Post-Processing Performance

In this section, we try to estimate how reliable the automatic flags of GALAH DR4 are.

Spectra have been classified by eye

- Spectroscopic Binary Type 1 (SB1) classification, Fig. 9
- Spectroscopic Binary Type 2 (SB2) classification, Fig. 11
- Equivalent width estimates of three diffuse interstellar bands (5780.59, 5797.19, 6613.66 Å) with central wavelengths identified by Vogrinčič et al. (2023) as well as for interstellar K (7698.9643 Å), see Fig. 12
- Emission classification and measurement, Fig. 13
- Upper Limit estimation

8.2 Comparison with GALAH DR3 and APOGEE DR17

Fig. 24

8.3 False positive rates for spectroscopic binaries

Binaries from Bashi et al. : 2304.00043 (?)

Thanks to the *Gaia* satellite the radial velocity estimates provided for bright stars from the *Gaia* radial velocity spectrometer (Katz et al., 2022), we have radial velocity estimates for 94% (774 914) of the stars observed for GALAH DR4.

Fig. 25

Binarity. How accurate is the flagging. Compare ratios of true/false, false-positive detections aided by Traven et al. (2020). Get in contact with Alex Wallace and Andy Casey regarding their binarity identification from BP/RP spectra.

8.4 Variable stars in GALAH DR4

We have crossmatched the sources from GALAH DR4 with those from *Gaia* DR3's variability catalogues (?).

Put list of vari catalogues and their respective publications in here, e.g. *Gaia* DR3 binary star parameters from `gaiadr3.nss_non_linn.fits` (?).

1 White dwarf crossmatched, indeed the one (2MASS J05005185–0930549) that was found in GALAH data by ?.

8.5 Minimal Science Verification

Fig. 3 from <https://arxiv.org/abs/2203.05463>: logAge vs. [C/N]

9. DATA RELEASE PRODUCTS

9.1 Data release catalogues

1. `galah_dr4_allspec.fits`: analysis for each spectrum (incl. RV) assuming single star
2. `galah_dr4_allstar.fits`: analysis for each star based on co-added spectra for each star and using non-spectroscopic information to constrain $\log g$
3. `galah_dr4_binary.fits`: analysis for those spectra that are suspected line-splitting spectroscopic binaries (SB2) assuming two sources for spectrum with same [Fe/H] but different RV

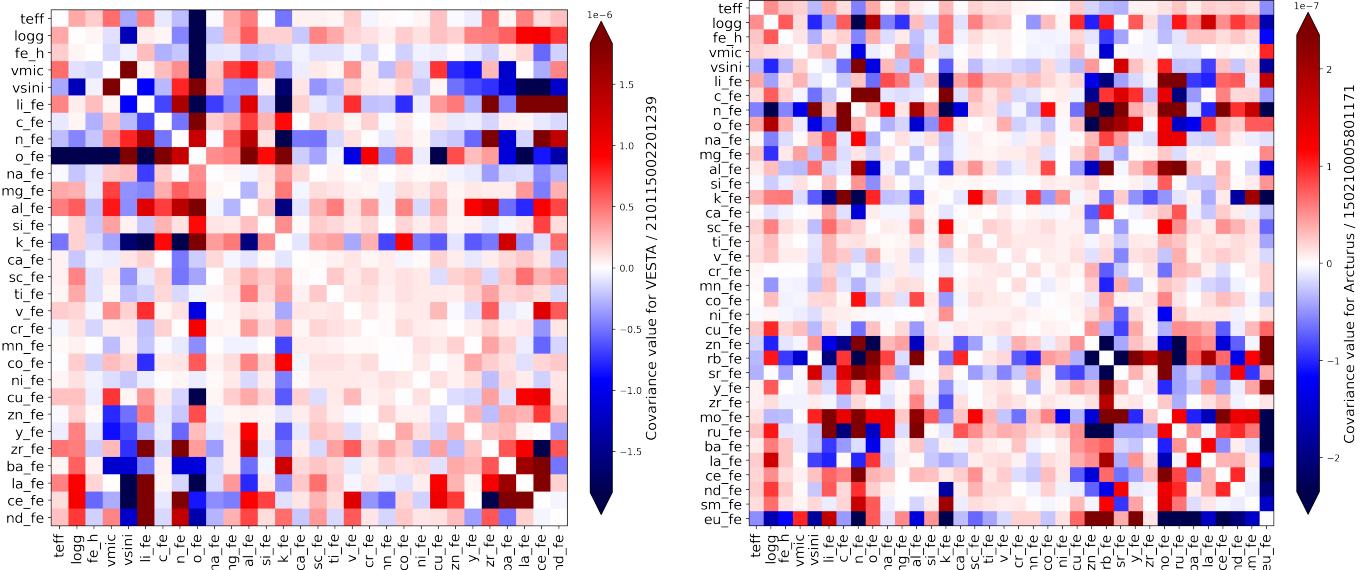


Figure 23. Covariance matrices for VESTA (panel a) and Arcturus (panel b).

Table 7. Overlap of GALAH DR4 with Gaia DR3 variability catalogues. Classifications are taken from the `gaiadr3.vari_classifier_result` table and described in ?. Abbreviations: ACV (α^2 Canum Venaticorum), MCP and CP ((magnetic) chemically peculiar), ROAM and ROAP (rapidly oscillating Am- and Ap-type), SXARI (SX Arietis), BE (B-type emission line), GCAS (γ Cassiopeiae), SDOR (SDoradus), WR (Wolf-Rayet), DSCT (δ Scuti), GDOR (γ Doradus), SXPHE (SX Phoenicis), SOLAR_LIKE (solar-like variable, including BY Draconis type, rotating spotted star, and flaring stars).

Classification	DR4 allspec	DR4 allstar
ACV CP MCP ROAM ROAP SXARI	330	285
ACYG (α Cygni-type)	45	34
BCEP (β Cephei)	16	15
BE GCAS SDOR WR	67	54
CEP (Cepheid)	133	103
CV (cataclysmic variable)	3	3
DSCT GDOR SXPHE	16342	14477
ECL (eclipsing binary)	4610	4074
ELL (ellipsoidal variable)	95	11
EP (star with exoplanet transits)	19	14
LPV (long-period variable)	7351	6247
RCB (R Coronae Borealis variable)	4	1
RR (RR Lyrae type)	561	401
RS (RS Canum Venaticorum variable)	3950	3355
S (short-timescale object)	40	33
SDB (subdwarf B star)	6	6
SOLAR_LIKE	20027	17256
SPB (slowly pulsating B-type variable)	31	24
SYST (symbiotic variable star)	3	3
WD (variable white dwarf)	1	1
YSO (young stellar object)	1534	1096
Total	55168	47493

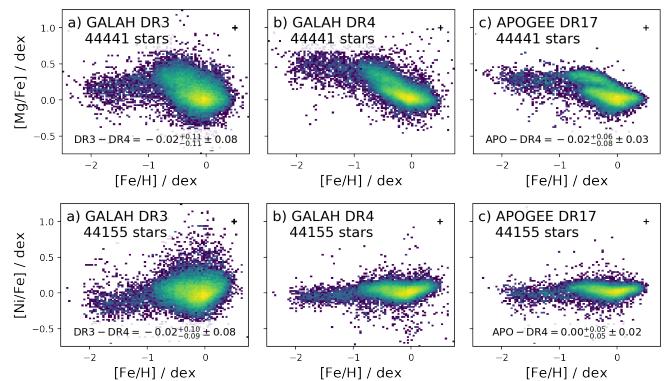


Figure 24. Comparison of stars with measurements for all of GALAH DR3 (left panel), GALAH DR4 (middle panel) as well as APOGEE DR17 (right panel) for [Mg/Fe] (top row), [Ni/Fe] (middle row).

Table 8. Table schema of the GALAH DR4 main catalogues.

Keyword	Type	Description	Section
sobject_id	int64	GALAH identifier	
tmass_id	str	2MASS identifier	
gaiadr3_source_id	int64	Gaia DR3 source_id	
survey_name	str	HERMES-2dF Survey/Program/Pointing (other if not available)	
field_id	int32	GALAH Field ID (-1 if not available)	
setup	str	Analysis setup: single/binary/coadds	
mjd	float32	Modified Julian Date	
ra	float64	propagated from Gaia DR3	
dec	float64	propagated from Gaia DR3	
best_spec4star	bool	Best spectrum for given star (identifiable via 2MASS ID)	
flag_sp	int32	Major spectroscopic quality bitmask flag	
flag_sp_fit	int32	Major spectroscopic fitting quality bitmask flag	
opt_loop	int32	Nr of optimisation loops used for fitting	
flag_red	int32	Quality bitmask flag of reduction pipeline	
snr_px_ccdx	float32	Average signal-to-noise ratio (per pixel) of CCDX	
chi2_sp	float32	Chi2 value of spectroscopic fitting	
px_used_perc	int64	Percentage of spectrum used for spectroscopic fit	
model_name	str	Neural network model used for creating synthetic spectra	
closest_model	str	Neural network model closest for Tefflogg/[Fe/H] combination	
comp_time	float32	Computation time spent on spectrum	
flux_contr	float32	Flux contribution of main source	
e_flux_contr	float32	Uncertainty flux_contr	
rv_comp_nr	int64	Number of peaks in RV cross-correlation function (CCF)	
rv_comp_x	float32	Radial velocity of primary or secondary sources	
e_rv_comp_x	float32	Uncertainty of rv_comp_x	
rv_comp_x_h	float32	Height of rv_comp_x in CCF	
rv_comp_x_p	float32	Prominence of rv_comp_x in CCF	
rv_gaia_dr3	float32	Radial velocity in Gaia DR3	
e_rv_gaia_dr3	float32	Uncertainty of rv_gaia_dr3	
v_bary_eff	float64	Barycentric velocity applied to reduced spectra	
teff	float32	Spectroscopic effective temperature (used for fitting)	
e_teff	float32	Uncertainty teff	
logg	float64	Surface gravity adjusted via parallax information	
e_logg	float32	Uncertainty logg_plx	
fe_h	float32	Abundance of Fe and all other elements not fitted in GALAH (Fe: 1D-NLTE)	
e_fe_h	float32	Uncertainty fe_h	
flag_fe_h	int32	Quality flag fe_h	
vmic	float32	Microturbulence velocity (fitted)	
e_vmic	float32	Uncertainty vmic	
vsini	float32	Broadening velocity (fitted sme.vsini with sme.vmac=0)	
e_vsini	float32	Uncertainty of vsini	
x_comp_2	float32	Parameter value x of potential secondary source	
e_x_comp_2	float32	Uncertainty of x_comp_2	
x_fe	float32	Elemental abundance for [X/Fe]	
e_x_fe	float32	Uncertainty x_fe	
flag_x_fe	int32	Quality bitmask flag of x_fe	

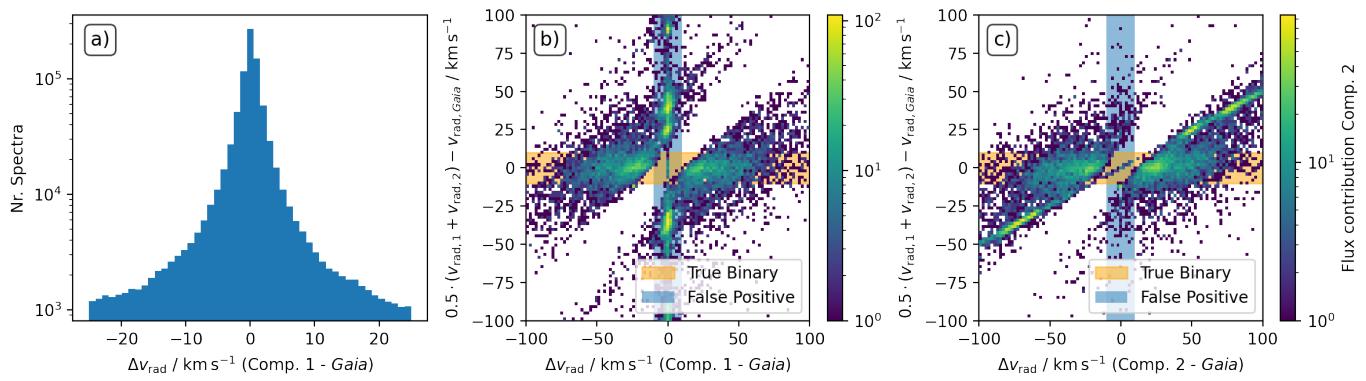


Figure 25. Validation of C2 and CN molecular line fitting as part of GALAH DR4 for the !Cen member 2MASS J13283993-4726329.

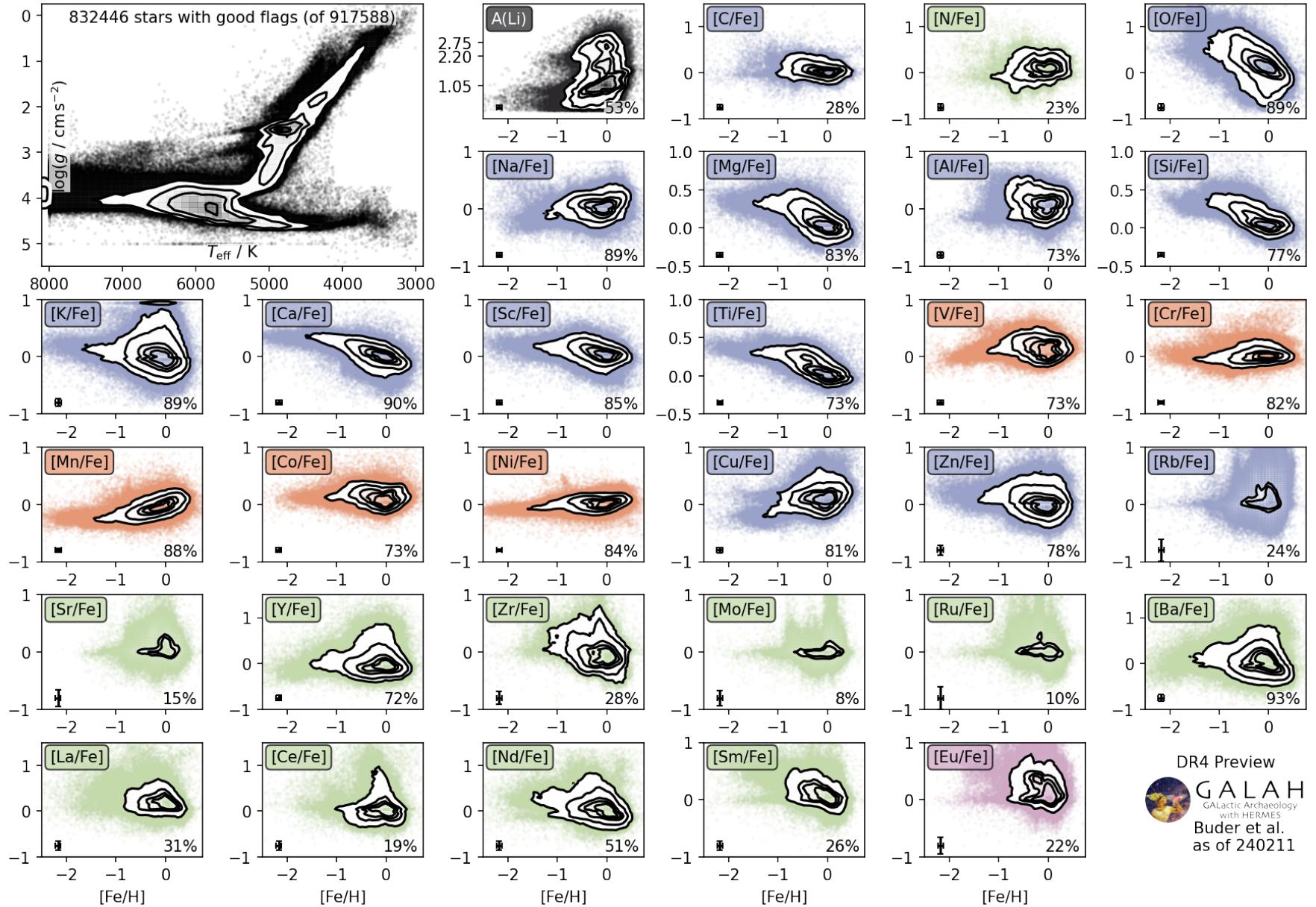


Figure 26. Overview of stellar parameters and elemental abundances for the *allstar* estimates of GALAH DR4. The top left panel shows the density distribution of stars in the Kiel diagram of T_{eff} and $\log g$. All other panels show the logarithmic elemental abundances (for elements indicated in the top left of the panel) as a function of the logarithmic iron abundances $[\text{Fe}/\text{H}]$. Elements are colored by different nucleosynthetic channels (black for big bang nucleosynthesis, blue for core-collapse supernovae, red for supernovae type Ia, green for asymptotic giant branch star contributions and pink for the rapid neutron capture process with contributions from merging neutron stars) following the color schema from Kobayashi et al. (2020).

9.2 Data products for each spectrum

1. 210115002201239_single_ft_comparison.pdf (see Fig. 29)
Maybe actually use OmegaCen star 140314005201392:
cool, metal-poor, strong CNO features and a good spectrum
to explain why continuum points may not always work for a pipeline.
2. 210115002201239_single_fit_covariances.npz
3. 210115002201239_single_fit_results.fits
4. 210115002201239_single_fit_rv.png (see Fig. 7)
5. 210115002201239_single_fit_spectrum.fits

For easier downloading, we also provide these files in tarfiles per night. Furthermore, we have interpolated all normalised allstar spectra onto a common wavelength that includes the largest possible wavelength coverage of 99.9% of radial velocity corrected spectra (neglecting pixels below 7680Å for the infrared channel).

9.3 Value-Added-Catalogs

9.3.1 Dynamics

Fig. ??

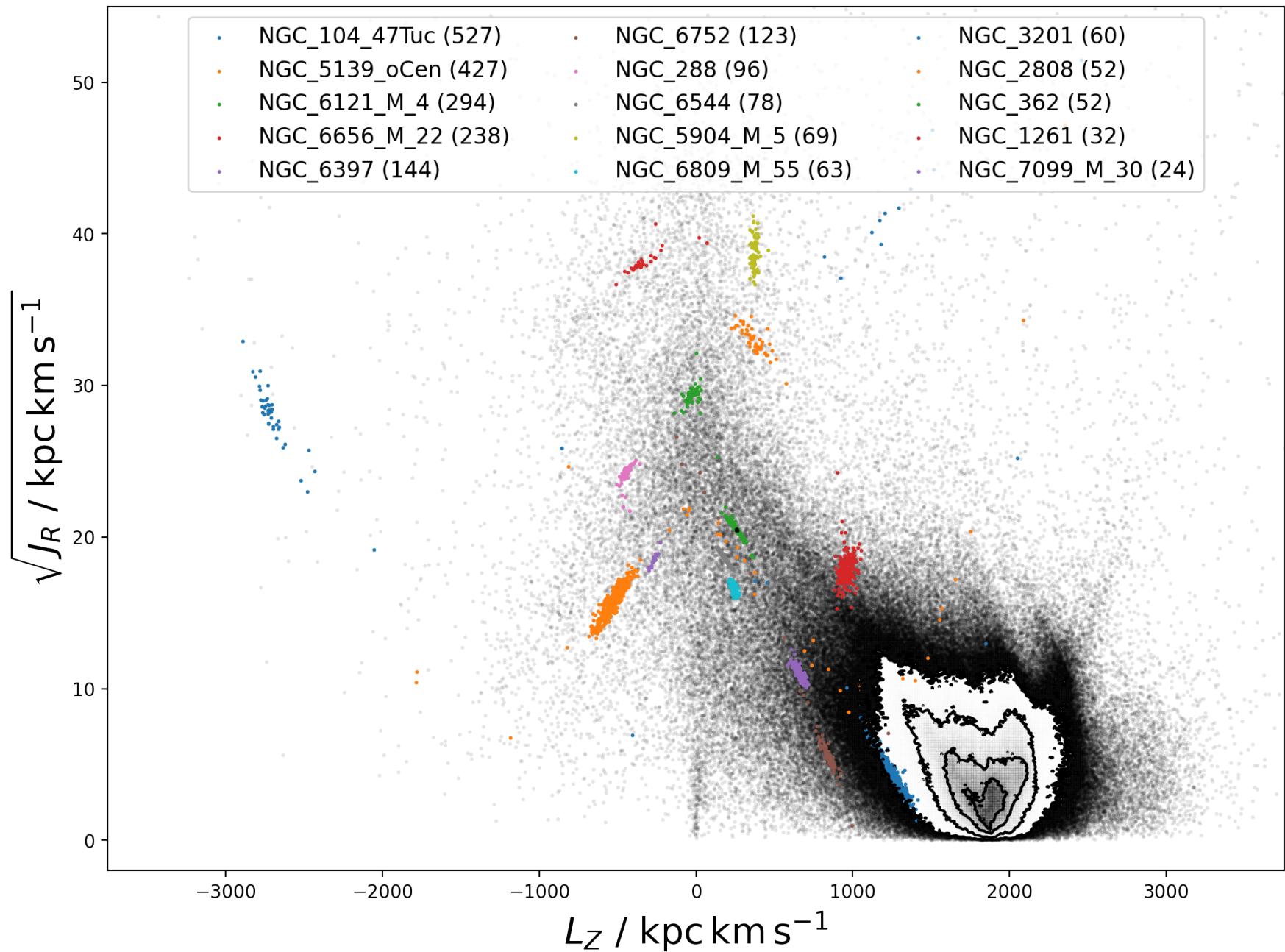


Figure 27. Overview of the dynamic content of GALAH DR4 (black) as well as selected globular clusters (colored) in action space of angular momentum L_Z versus radial action J_R . Cluster members were selected as those with more than 70 percent membership probability according to Vasiliev & Baumgardt (2021).

9.3.2 Ages

10. CAVEATS AND FUTURE IMPROVEMENTS

10.1 Actual Bugs

`flag_fe_h` was computed similar to the elemental abundances, that is by comparing the best-fit spectrum with a spectrum with the lowest grid value. In the case of [Fe/H], however, this is incorrect. It can lead to a comparison of a spectrum with -0.74 to one with -0.75, which can easily be too similar within the SNR range. Wrongfully affecting 34% of stars and making the flag useless.

10.2 Rerunning?

neglecting ccds with flux uncertainty below 0 for more than 5% of spectra. test cases were 200514003201389 and 200514002701209. Is there a way to identify these stars more systematically? These were found because their [Mg/Fe] maxed out without any reason or `flag_sp` raised. Is it only 200514?

10.3 Masses

Should a prior be used for the mass estimates? While this is maybe not too important for the majority of stars, we noticed potentially significant differences in the area around the primary and even more so around the more massive secondary red clump, where different masses will lead to different $\log g$ estimates. One could in principle include a prior on mass based on the initial mass function, as for example done by Sharma et al. (2018) with the exponential form of the initial mass function by ?. Such choices should be carefully tested in the future as they can significantly influence the estimated stellar parameters, most notably masses and ages.

10.4 Limitations of our implemented multiple neural networks

10.4.1 Consistency between models and nodding in the 3D-domain

10.4.2 Flexibility of our neural network implementation for extreme abundances

While this approach has proved to be powerful for all elements across their abundance ranges, we have noticed sinusoidal shapes for weak Li lines (see also Wang et al., 2020). This is likely caused by the large dynamical range that has to be covered by the neural network of $0 < A(\text{Li}) < 4$. Elaborate on this! Ella's EW estimation code will limit the impact of this caveat!

10.4.3 Coverage of chemical compositions

Before the neural networks are computed, it should actually be tested, what the abundance zeropoints are. In the case of several elements, like Na, Al, Eu, they were significant on the order of 0.2–0.3dex. In that case, stars with high abundances of 0.7–0.8 dex are not be sufficiently covered, although expected (e.g. in old stars and especially globular clusters). → rerun neural networks for these cases?

10.5 Which analysis setup should be used?

10.5.1 To use or not to use non-spectroscopic information?

An implementation of non-spectroscopic information, as done in our `allstar` module, has the advantage of overcome spectroscopic degeneracies (as proven for the limited information on $\log g$ in the HERMES wavelength range) as well as improving accuracy and precision also for the lowest quality spectra (because $\log g$ is no longer solely dependent on the spectrum information).

However, this approach is only useful if the non-spectroscopic information not biased (as it would be for astrometric and photometric information in the case of unresolved binarity). In this case, the `binary` module should be favoured and we have tried to implement a sophisticated decision algorithm to decide if a spectrum is including contributions from a second source or just a fast rotating star.

10.6 Binary or fast rotator?

Example of a warm main sequence turn-off star that could be a binary, but shows no clear line-splitting.

10.7 Unreliable wavelength solutions

For each CCD, the reduction has found the most suitable wavelength solution, linking pixels with actual wavelengths, based on the ThXe arc lines. In GALAH DR3 (Buder et al., 2021), we have found several issues for spectra, where not enough ThXe lines could be used to constrain the wavelength solution. Improvements have been made for the new reduction version to improve the number of useful ThXe lines. Two additional pieces of information are, however, unused by the reduction pipeline, namely (i) the telluric lines that are present throughout GALAH spectra and (ii) the absorption lines of the stellar spectra. Both hold valuable information, as their position (in rest wavelength) is known very well. This has affected mainly CCD3 with Li and Eu. One could estimate wavelength solutions from arc lines + telluric lines + absorption features.

Note that I have also monitored the CRVAL1 and CDELT1 estimates with quite a few surprising outliers^h: 7.9% for CCD3 have CDELT1 significantly outside (mainly above) the expected values. CCD4 shows 2.2 and 2.1%, respectively, with outliers in CRVAL1 and CDELT1.

Sanity check: Are all `e_fe_h` actually above the accuracy value for `fe_h`?

10.8 Chemical composition differences of model atmospheres and spectra

For several of our synthetic spectra, the chosen chemical composition differs significantly from the scaled-Solar pattern of the MARCS model atmospheres. We note that APOGEE DR17 (Abdurro'uf et al., 2022) have managed to compute MARCS that match their chemical compositions.

^hvalidation/figures/galah_dr4_crval_cdelt_histograms.png

10.9 Red clump stars

Some red clump stars have too high $\log g$, because the isochrones were selected imperfect. Possible fixes include selecting better isochrones or including [C/N] when adjusting $\log g$? could help to put the RCs closer to 2.5dex

11. CONCLUSIONS

C, N etc. also in high-res spectra: (?) <https://ui.adsabs.harvard.edu/abs/arXiv:2308.10118>

Future work: holistic reduction (??)

Acknowledgements

We acknowledge the traditional owners of the land on which the AAT and ANU stand, the Gamilaraay, the Ngunnawal and Ngambri people. We pay our respects to elders past, present, and emerging and are proud to continue their tradition of surveying the night sky in the Southern hemisphere. This work was supported by the Australian Research Council Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), through project number CE170100013.

Facilities

AAT with 2df-HERMES at Siding Spring Observatory: AAT observations for this data release were performed under programs 2013B/13, 2014A/25, 2015A/3, 2015A/19, 2015B/1, 2015B/19, 2016A/22, 2016B/10, 2016B/12, 2017A/14, 2017A/18, 2017B/16, 2018A/18, 2018B/15, 2019A/1, 2019A/15, and 2020/B23. This paper includes data that has been provided by AAO Data Central (datacentral.ao.org.au). **Gaia:** This work has made use of data from the European Space Agency (ESA) mission *Gaia* (<http://www.cosmos.esa.int/gaia>), processed by the *Gaia* Data Processing and Analysis Consortium (DPAC, <http://www.cosmos.esa.int/web/gaia/dpac/consortium>). Funding for the DPAC has been provided by national institutions, in particular the institutions participating in the *Gaia* Multi-lateral Agreement. **Other facilities:** This publication makes use of data products from the Two Micron All Sky Survey (Skrutskie et al., 2006) and the CDS VizieR catalogue access tool (Ochsenbein et al., 2000).

Software

The research for this publication was coded in PYTHON (version 3.7.4) and included its packages ASTROPY (v. 3.2.2; Astropy Collaboration et al., 2013, 2018), ASTROQUERY (v. 0.4; Ginsburg et al., 2019), CORNER (v. 2.0.1; Foreman-Mackey, 2016), GALPY (version 1.6.0; Bovy, 2015), IPYTHON (v. 7.8.0; Pérez & Granger, 2007), MATPLOTLIB (v. 3.1.3; Hunter, 2007), NUMPY (v. 1.17.2; Walt et al., 2011), SCIPY (version 1.3.1; Virtanen et al., 2020), SKLEARN (v. 0.21.3; Pedregosa et al., 2011). We further made use of TOPCAT (version 4.7; Taylor, 2005);

References

- Abdurro'uf, Accetta, K., Aerts, C., et al. 2022, ApJS, 259, 35
 Amarsi, A. M., Lind, K., Osorio, Y., et al. 2020, A&A, 642, A62
 Andrae, R., Fouesneau, M., Sordo, R., et al. 2022, arXiv e-prints, arXiv:2206.06138
 Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33
 Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, AJ, 156, 123
 Bailer-Jones, C. A. L., Rybizki, J., Fouesneau, M., Demleitner, M., & Andrae, R. 2021, AJ, 161, 147
 Bailer-Jones, C. A. L., Rybizki, J., Fouesneau, M., Mantelet, G., & Andrae, R. 2018, AJ, 156, 58
 Barden, S. C., Jones, D. J., Barnes, S. I., et al. 2010, in Proc. SPIE, Vol. 7735, Ground-based and Airborne Instrumentation for Astronomy III, 773509
 Bashi, D., Mazeh, T., & Faigler, S. 2023, MNRAS, 522, 1184
 Baumgardt, H., & Vasiliev, E. 2021, MNRAS, 505, 5957
 Bedell, M., Bean, J. L., Meléndez, J., et al. 2018, ApJ, 865, 68
 Belokurov, V., Erkal, D., Evans, N. W., Koposov, S. E., & Deason, A. J. 2018, MNRAS, 478, 611
 Bergemann, M., Serenelli, A., Schönrich, R., et al. 2016, A&A, 594, A120
 Bovy, J. 2015, ApJS, 216, 29
 Bressan, A., Marigo, P., Girardi, L., et al. 2012, MNRAS, 427, 127
 Brzeski, J., Case, S., & Gers, L. 2011, Proc. SPIE, 8125, 812504
 Buder, S., Asplund, M., Duong, L., et al. 2018, MNRAS, 478, 4513
 Buder, S., Sharma, S., Kos, J., et al. 2021, MNRAS, 506, 150
 Cantat-Gaudin, T., & Anders, F. 2020, A&A, 633, A99
 Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, ApJ, 345, 245
 Carretta, E., Bragaglia, A., Gratton, R., D’Orazi, V., & Lucatello, S. 2009a, A&A, 508, 695
 Carretta, E., Bragaglia, A., Gratton, R., & Lucatello, S. 2009b, A&A, 505, 139
 Carretta, E., Bragaglia, A., Gratton, R. G., et al. 2009c, A&A, 505, 117
 Casagrande, L., & VandenBerg, D. A. 2018, MNRAS, 475, 5023
 Casagrande, L., Lin, J., Rains, A. D., et al. 2020, arXiv e-prints, arXiv:2011.02517
 Casey, A. R., Hogg, D. W., Ness, M., et al. 2016, ArXiv e-prints, arXiv:1603.03040
 Casey, A. R., Hawkins, K., Hogg, D. W., et al. 2017, ApJ, 840, 59
 Chabrier, G. 2001, ApJ, 554, 1274
 Cutri, R. M., et al. 2014, VizieR Online Data Catalog, 2328
 Dalton, G., Trager, S., Abrams, D. C., et al. 2014, in Proc. SPIE, Vol. 9147, Ground-based and Airborne Instrumentation for Astronomy V, 91470L
 de Jong, R. S., Agertz, O., Berbel, A. A., et al. 2019, The Messenger, 175, 3
 De Silva, G. M., Freeman, K. C., Bland-Hawthorn, J., et al. 2015, MNRAS, 449, 2604
 Dutra-Ferreira, L., Pasquini, L., Smiljanic, R., Porto de Mello, G. F., & Steffen, M. 2016, A&A, 585, A75
 Evans, D. W., Riello, M., De Angeli, F., et al. 2018, ArXiv e-prints, arXiv:1804.09368
 Farrell, T. J., Birchall, M. N., Heald, R. W., et al. 2014, in Proc. SPIE, Vol. 9152, Software and Cyberinfrastructure for Astronomy III, 915223
 Foreman-Mackey, D. 2016, The Journal of Open Source Software, 1, 24
 Fouesneau, M., Frémant, Y., Andrae, R., et al. 2022, arXiv e-prints, arXiv:2206.05992
 Gaia Collaboration, Prusti, T., de Bruijne, J. H. J., et al. 2016, A&A, 595, A1
 Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2018, A&A, 616, A1
 —. 2021, A&A, 649, A1
 Gaia Collaboration, Vallenari, A., Brown, A. G. A., et al. 2022, arXiv e-prints, arXiv:2208.00211
 Gilmore, G., Randich, S., Worley, C. C., et al. 2022, A&A, 666, A120
 Ginsburg, A., Sipőcz, B. M., Brasseur, C. E., et al. 2019, AJ, 157, 98
 Grevesse, N., Asplund, M., & Sauval, A. J. 2007, Space Sci. Rev., 130, 105
 Gustafsson, B., Edvardsson, B., Eriksson, K., et al. 2008, A&A, 486, 951
 Halbwachs, J.-L., Pourbaix, D., Arenou, F., et al. 2023, A&A, 674, A9
 Heijmans, J., Asplund, M., Barden, S., et al. 2012, in Proc. SPIE, Vol. 8446, Ground-based and Airborne Instrumentation for Astronomy IV, 84460W

- Heiter, U., Lind, K., Bergemann, M., et al. 2021, A&A, 645, A106
- Helmi, A., Babusiaux, C., Koppelman, H. H., et al. 2018, Nature, 563, 85
- Hinkle, K., Wallace, L., Valenti, J., & Harmer, D., eds. 2000, Visible and Near Infrared Atlas of the Arcturus Spectrum, 3727–9300 Å (Astron. Soc. Pac.)
- Ho, A. Y. Q., Ness, M. K., Hogg, D. W., et al. 2017, ApJ, 836, 5
- Hon, M., Kuszlewicz, J. S., Huber, D., Stello, D., & Reyes, C. 2022, AJ, 164, 135
- Hunter, J. D. 2007, Comput Sci Eng, 9, 90
- Jofré, P., Heiter, U., Tucci Maia, M., et al. 2018, Research Notes of the American Astronomical Society, 2, 152
- Jofré, P., Heiter, U., Worley, C. C., et al. 2017, A&A, 601, A38
- Johnson, C. I., & Pilachowski, C. A. 2010, ApJ, 722, 1373
- Katz, D., Sartoretti, P., Guerrier, A., et al. 2022, arXiv e-prints, arXiv:2206.05902
- Kawka, A., Simpson, J. D., Vennes, S., et al. 2020, MNRAS, 495, L129
- Kobayashi, C., Karakas, A. I., & Lugaro, M. 2020, ApJ, 900, 179
- Kollmeier, J. A., Zasowski, G., Rix, H.-W., et al. 2017, arXiv e-prints, arXiv:1711.03234
- Kos, J., Lin, J., Zwitter, T., et al. 2017, MNRAS, 464, 1259
- Kos, J., Bland-Hawthorn, J., Betters, C. H., et al. 2018, MNRAS, 480, 5475
- Lewis, I. J., Cannon, R. D., Taylor, K., et al. 2002, MNRAS, 333, 279
- Lin, J., Dotter, A., Ting, Y.-S., & Asplund, M. 2018, MNRAS, doi:10.1093/mnras/sty709
- Lindegren, L., Hernández, J., Bombrun, A., et al. 2018, A&A, 616, A2
- Lindegren, L., Bastian, U., Biermann, M., et al. 2021a, A&A, 649, A4
- Lindegren, L., Klioner, S. A., Hernández, J., et al. 2021b, A&A, 649, A2
- Majewski, S. R., Zasowski, G., & Nidever, D. L. 2011, ApJ, 739, 25
- Marigo, P., Girardi, L., Bressan, A., et al. 2017, ApJ, 835, 77
- Monty, S., Yong, D., Marino, A. F., et al. 2023, MNRAS, 518, 965
- Ness, M., Hogg, D. W., Rix, H.-W., Ho, A. Y. Q., & Zasowski, G. 2015, ApJ, 808, 16
- Ness, M., Hogg, D. W., Rix, H.-W., et al. 2016, ApJ, 823, 114
- Nissen, P. E. 2015, A&A, 579, A52
- Nissen, P. E., & Gustafsson, B. 2018, A&A Rev., 26, 6
- Ochsenbein, F., Bauer, P., & Marcout, J. 2000, A&AS, 143, 23
- Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2011, J Mach Learn Res, 12, 2825
- Pérez, F., & Granger, B. E. 2007, Comput Sci Eng, 9, 21
- Piskunov, N., & Valenti, J. A. 2017, A&A, 597, A16
- Prša, A., Harmanec, P., Torres, G., et al. 2016, AJ, 152, 41
- Rains, A. D., Žerjal, M., Ireland, M. J., et al. 2021, MNRAS, 504, 5788
- Riello, M., De Angeli, F., Evans, D. W., et al. 2021, A&A, 649, A3
- Rimoldini, L., Holl, B., Gavras, P., et al. 2023, A&A, 674, A14
- Rix, H.-W., Ting, Y.-S., Conroy, C., & Hogg, D. W. 2016, ApJ, 826, L25
- Salaris, M., & Cassisi, S. 2006, Evolution of Stars and Stellar Populations (J. Wiley)
- Saydjari, A. K., Zucker, C., Peek, J. E. G., & Finkbeiner, D. P. 2022, arXiv e-prints, arXiv:2212.03879
- Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, ApJ, 500, 525
- Sharma, S., Stello, D., Budar, S., et al. 2018, MNRAS, 473, 2004
- Sharma, S., Stello, D., Bland-Hawthorn, J., et al. 2019, MNRAS, 490, 5335
- Sheinis, A., Anguiano, B., Asplund, M., et al. 2015, J. Astron. Telesc. Instrum. Syst., 1, 035002
- Skrutskie, M. F., Cutri, R. M., Stiening, R., et al. 2006, AJ, 131, 1163
- Steinmetz, M., Matijević, G., Enke, H., et al. 2020, AJ, 160, 82
- Stello, D., Huber, D., Sharma, S., et al. 2015, ApJ, 809, L3
- Taylor, M. B. 2005, in ASPC, Vol. 347, Astronomical Data Analysis Software and Systems XIV, ed. P. Shopbell, M. Britton, & R. Ebert, 29
- Ting, Y.-S., Conroy, C., & Rix, H.-W. 2016, ApJ, 826, 83
- Ting, Y.-S., Conroy, C., Rix, H.-W., & Asplund, M. 2018, ApJ, 860, 159
- Ting, Y.-S., Conroy, C., Rix, H.-W., & Cargile, P. 2019, ApJ, 879, 69
- Torra, F., Castañeda, J., Fabricius, C., et al. 2021, A&A, 649, A10
- Traven, G., Feltzing, S., Merle, T., et al. 2020, A&A, 638, A145
- Valenti, J. A., & Piskunov, N. 1996, A&AS, 118, 595
- van Leeuwen, F. 2007, A&A, 474, 653
- Vasiliev, E., & Baumgardt, H. 2021, MNRAS, 505, 5978
- Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, Nature Methods, 17, 261
- Vogrinčič, R., Kos, J., Zwitter, T., et al. 2023, arXiv e-prints, arXiv:2303.14016
- Walt, S. v. d., Colbert, S. C., & Varoquaux, G. 2011, Comput Sci Eng, 13, 22
- Wang, E. X., Nordlander, T., Asplund, M., et al. 2020, MNRAS, arXiv:2010.15248
- Wehrhahn, A. 2021, in The 20.5th Cambridge Workshop on Cool Stars, Stellar Systems, and the Sun (CS20.5), Cambridge Workshop on Cool Stars, Stellar Systems, and the Sun, 1
- Xiang, M., Ting, Y.-S., Rix, H.-W., et al. 2019, ApJS, 245, 34
- Xiang, M., Rix, H.-W., Ting, Y.-S., et al. 2021, arXiv e-prints, arXiv:2108.02878
- Yong, D., Meléndez, J., Grundahl, F., et al. 2013, MNRAS, 434, 3542
- Yong, D., Roederer, I. U., Grundahl, F., et al. 2014, MNRAS, 441, 3396
- Zhao, G., Zhao, Y.-H., Chu, Y.-Q., Jing, Y.-P., & Deng, L.-C. 2012, Research in Astronomy and Astrophysics, 12, 723
- Zinn, J. C., Stello, D., Elsworth, Y., et al. 2020, ApJS, 251, 23

Appendix 1. Initial parameters

Appendix 2. Example Fit Comparison

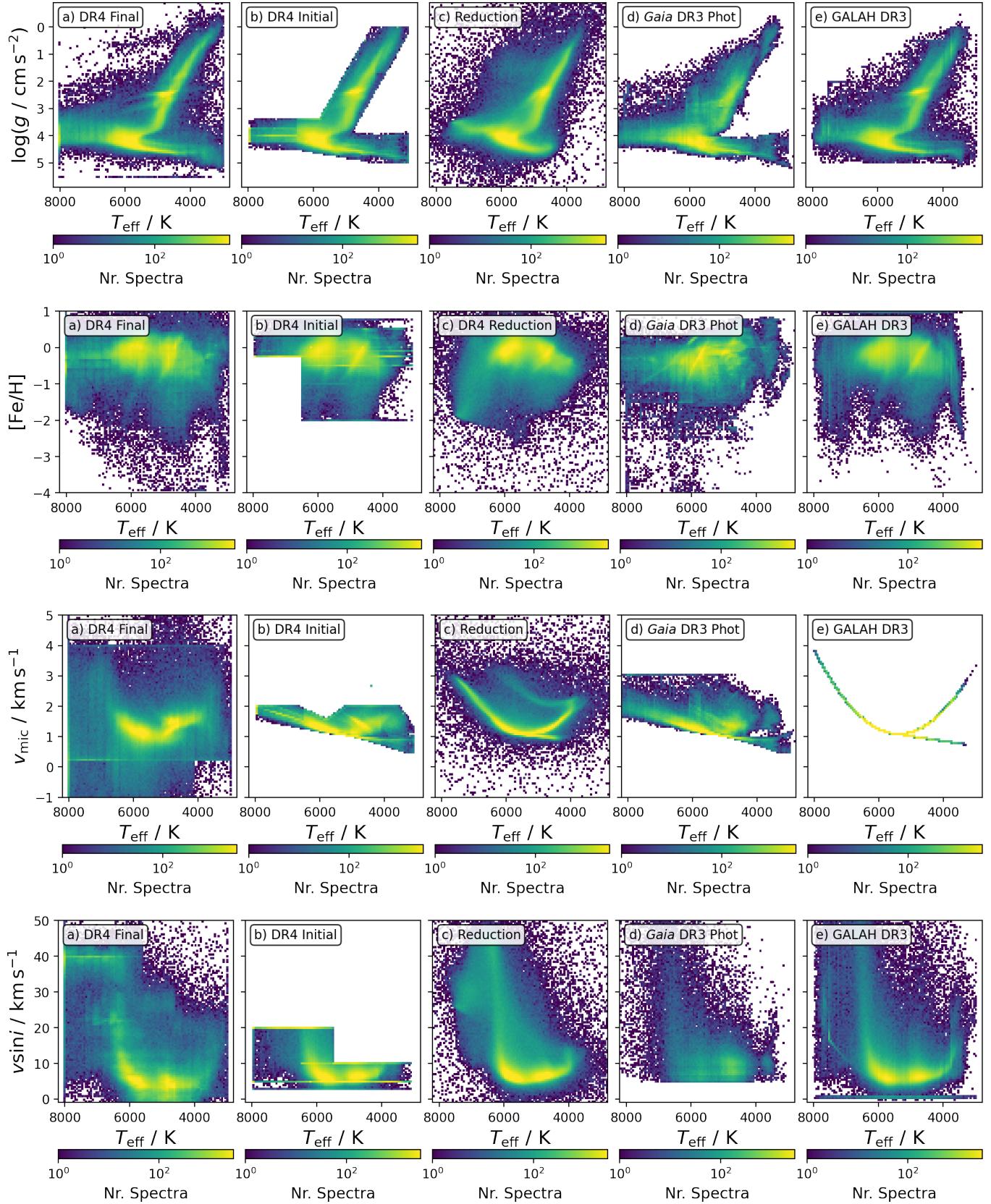
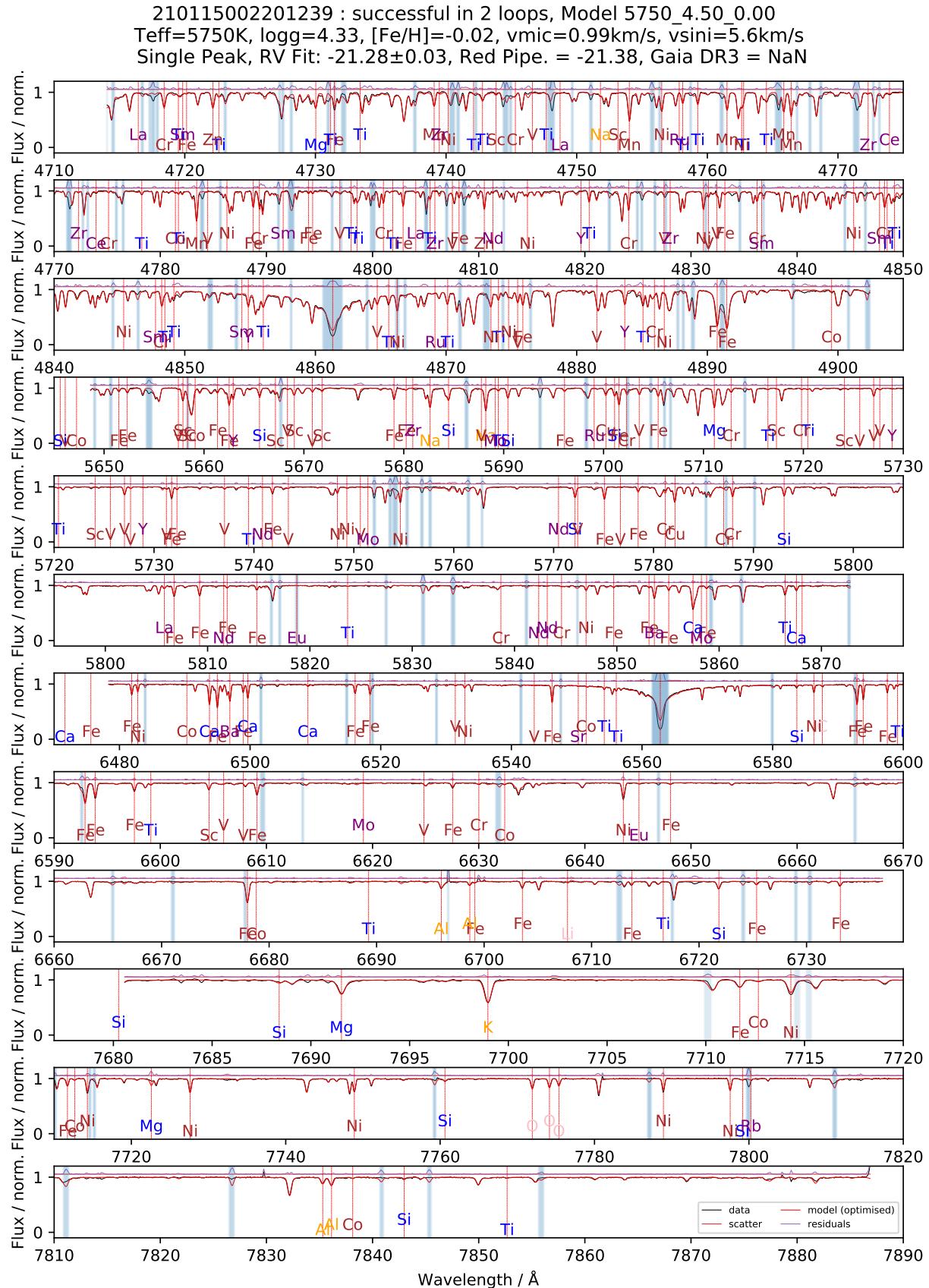


Figure 28. Comparison of initial parameters used for GALAH DR4 (first column) with the GALAH DR4 reduction pipeline (second column), Gaia DR3 (third column with v_{mic} based on the adjusted formula from Dutra-Ferreira et al. (2016)), and GALAH DR3 (fourth column).

Figure 29. Example output of the *allspec* fitting routine for VESTA / 210115002201239.