







Optical spectroscopic signatures of the red giant evolutionary state

Ella Xi Wang ^{1,2,3}★, Melissa Ness ², Thomas Nordlander ^{2,3,4}, Andrew R. Casey,^{3,5,6}
 Sarah Martell ^{3,7}, Marc Pinsonneault ⁸, Xiaoting Fu(符晓婷) ⁹, Dennis Stello,^{3,7,10} Claudia Reyes,²
 Marc Hon,¹¹ Madeleine McKenzie,^{2,12} Mingjie Jian(简明杰),¹ Jie Yu(余杰),² Sven Buder,^{2,3} Karin Lind,¹
 Joss Bland-Hawthorn,^{3,10} Daniel B. Zucker,^{3,13,14} Pradosh Barun Das,^{3,13,14} Richard de Grijs^{13,14,15}
 and Michael Hayden¹⁶

¹Department of Astronomy, AlbaNova University Center, Stockholm University, SE–106 91 Stockholm, Sweden

²Research School of Astronomy and Astrophysics, Australian National University, Canberra, ACT 2611, Australia

³ARC Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), Canberra, ACT 2611, Australia

⁴Theoretical Astrophysics, Department of Physics and Astronomy, Uppsala University, Box 516, 751 20 Uppsala, Sweden

⁵School of Physics & Astronomy, Monash University, Melbourne, VIC 3800, Australia

⁶Center for Computational Astrophysics, Flatiron Institute, 162 Fifth Avenue, New York, NY 10010, USA

⁷School of Physics, University of New South Wales, Sydney, NSW 2052, Australia

⁸Department of Astronomy, The Ohio State University, 140 West 18th Ave, Columbus OH 43210, USA

⁹Purple Mountain Observatory, Chinese Academy of Sciences, Nanjing 210023, People's Republic of China

¹⁰Sydney Institute for Astronomy (SIfA), School of Physics, University of Sydney, Sydney, NSW 2006, Australia

¹¹Kavli Institute for Astrophysics and Space Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

¹²Carnegie Science Observatories, 813 Santa Barbara St, Pasadena, CA 91101, USA

¹³School of Mathematical and Physical Sciences, Macquarie University, Balaclava Road, Sydney, NSW 2109, Australia

¹⁴Astrophysics and Space Technologies Research Centre, Macquarie University, Balaclava Road, Sydney, NSW 2109, Australia

¹⁵International Space Science Institute–Beijing, 1 Nanertiao, Zhongguancun, Hai Dian District, Beijing 100190, China

¹⁶Homer L. Dodge Department of Physics & Astronomy, University of Oklahoma, 440 W. Brooks St, Norman, OK 73019, USA

Accepted 2025 June 3. Received 2025 April 10; in original form 2025 February 14

ABSTRACT

Modern spectroscopic surveys output large data volumes. Theoretical models provide a means to transform the information encoded in these data to measurements of physical stellar properties. However, in detail, the models are incomplete and simplified, and prohibit interpretation of the fine details in spectra. Instead, the available data provide an opportunity to use data-driven, differential analysis techniques, as a means towards understanding spectral signatures. We deploy such an analysis to examine core helium-fusing red clump (RC) and shell hydrogen-fusing red giant branch (RGB) stars, to uncover signatures of evolutionary state imprinted in optical stellar spectra. We exploit 786 pairs of RC and RGB stars from the GALAH survey, chosen to minimize spectral differences, with evolutionary state classifications from TESS and K2 asteroseismology. We report sub-per cent residual, systematic spectral differences between the two classes of stars, and show that these residuals are significant compared to a reference sample of RC–RC and RGB–RGB pairs selected using the same criteria. First, we report systematic differences in the Swan (C_2) band and CN bands caused by stellar evolution and a difference in mass, where RGB stars at similar stellar parameters have higher masses than RC stars. Secondly, we observe systematic differences in the line-width of the H_α and H_β lines caused by a difference in microturbulence, as measured by GALAH, where we measure higher microturbulence in RC stars than RGB stars. This work demonstrates the ability of large surveys to uncover the subtle spectroscopic signatures of stellar evolution using model-free, data-driven methods.

Key words: asteroseismology – methods: data analysis – techniques: spectroscopic – surveys – stars: evolution – stars: statistics.

1 INTRODUCTION

Distinguishing between hydrogen-shell-fusing red giant branch (RGB) stars and helium-core-fusing red clump (RC) stars is important to a myriad of science cases. For example: the study of lithium rich giants and their enhancement mechanisms (Casey et al.

2019; Singh, Reddy & Kumar 2019; Martell et al. 2021; Yan et al. 2021; Zhou et al. 2022; Tayar et al. 2023), building the local cosmic distance ladder (Stanek, Zaritsky & Harris 1998; Hawkins et al. 2017), building constrained stellar samples to study the low and high α disc (Lu et al. 2022), and mass loss on the RGB (Charbonnel 2005; Howell et al. 2024).

Asteroseismology, the study of internal stellar structure through oscillations, provides an unambiguous way of distinguishing between RC and RGB stars (Montalbán et al. 2010; Bedding et al. 2011;

* E-mail: xi.wang@astro.su.se

Mosser et al. 2011, 2012). Red giants are observed to have solar-like oscillations (Frandsen et al. 2002; De Ridder et al. 2006; Barban et al. 2007). These oscillations are caused by pressure or acoustic (p) modes, which probe the convective envelope; and gravity (g) modes, which probe the core. Whilst g modes contain information on the stellar core, they usually cannot propagate to the surface of the star and are thus not observable. However, p and g modes can couple to form mixed modes (Chaplin & Miglio 2013). Mixed modes carry properties from the stellar core to the surface of the star, allowing us to distinguish between different evolutionary states (Beck et al. 2011).

Asteroseismology requires high temporal frequency radial velocity or photometric observations in order to resolve mixed modes (Hon, Stello & Yu 2018a). Although great strides have been made for asteroseismic surveys (Stello et al. 2013; Elsworth et al. 2017; Hon et al. 2018a), the number of stars for which we have asteroseismic measurements pales in comparison to the number of stars with spectroscopic data (Abdurro'uf et al. 2022; Buder et al. 2025). Using spectroscopy, we are able to measure effective temperature (T_{eff}), surface gravity ($\log(g)$), and metallicity ($[\text{Fe}/\text{H}]$) to high accuracy and precision. However, stellar evolutionary tracks at the same $[\text{Fe}/\text{H}]$ but different mass can overlap in $T_{\text{eff}}\text{-}\log(g)$ space. Although RC and RGB stars can be separated if the star's mass is known, it is difficult to infer masses for giant stars to a high accuracy and precision using only spectroscopic data.

Machine learning (ML) algorithms have been employed to expand the number of stars with an evolutionary state classification. Data-driven ML methods that leverage precision asteroseismic measurements have shown success in learning evolutionary state directly from both APOGEE and LAMOST stellar spectra (Hawkins, Ting & Walter-Rix 2018; Ting, Hawkins & Rix 2018; Casey et al. 2019; Lu et al. 2022). This is achieved through label transfer of asteroseismic parameters to spectra, allowing us to classify the evolutionary state using spectroscopic information. With these ML algorithms achieving 93–98 per cent accuracy, it is clear that there is an evolutionary state signal in stellar spectra.

Differences in the spectroscopy of RC and RGB stars have been linked to molecular features related to carbon (Masseron & Hawkins 2017; Hawkins et al. 2018; Banks et al. 2023, 2024), with the so-called deep mixing thought to be the cause of this link. At the end of the first dredge up, the convective envelope recedes to the surface, leaving behind a chemical discontinuity (the so-called μ -barrier) which prevents further mixing of material between the stellar core and surface. During evolution up the red giant branch, the hydrogen-fusing shell expands outwards and contacts this μ -barrier, effectively removing it. When this occurs, the star temporarily drops in luminosity, causing what we observe as the RGB bump. The removal of this chemical discontinuity beyond the RGB bump allows further mixing between the stellar core and surface (Sweigart & Mengel 1979; Charbonnel 1994; Charbonnel & Zahn 2007; Charbonnel & Lagarde 2010; Lagarde et al. 2012). Deep mixing has been observed on the RGB beyond the RGB bump (Gilroy 1989; Charbonnel, Brown & Wallerstein 1998; Gratton et al. 2000; Recio-Blanco & de Laverny 2007; Martell, Smith & Briley 2008a; Martell, Smith & Briley 2008b; Masseron & Gilmore 2015; Lagarde et al. 2019; Shetrone et al. 2019; Roberts et al. 2024): it depletes C, enhances N, and reduces the $^{12}\text{C}/^{13}\text{C}$ ratio, with the amount depending on the star's mass and metallicity (Lagarde et al. 2012). As a result, RC stars of similar mass and metallicity have lower $[\text{C}/\text{N}]$ and $^{12}\text{C}/^{13}\text{C}$ compared to RGB stars, causing an observable difference in carbon-related molecular features in the spectrum. Although existing studies find differences in CN, CH, and CO molecular lines (Hawkins et al.

2018; Banks et al. 2023, 2024), these studies make use of data-driven models.

There are now a number of large spectroscopic surveys with complementary classifications of evolutionary state from asteroseismic surveys for subsets of stars (Pinsonneault et al. 2014; Buder et al. 2025; Pinsonneault et al. 2025). Galactic Archaeology with HERMES (GALAH) is a spectroscopic survey, with the main mission of studying Galactic archaeology through chemical tagging (De Silva et al. 2015). In Data Release 4 (Buder et al. 2025), GALAH incorporated additional fields that observe asteroseismic targets from the K2 (Howell et al. 2014) and *TESS* (Ricker et al. 2014) surveys.

In this paper, we study the evolutionary state signal in optical stellar spectra using the GALAH, *TESS*, and K2 surveys in a data-driven but model-free way. In Section 2, we clean and crossmatch the survey data used in this work. We take pairs of RC and RGB stars in order to limit stellar parameter influence, discussed in Section 3. In Section 4, we show the optical spectroscopic signature of stellar evolutionary state, and link this signal to physical stellar parameters in Section 5. Lastly, we summarise our findings in Section 6.

2 DATA

2.1 Spectroscopic catalogue

In this work, we use measured parameters and abundances as well as observed spectra drawn from the million-star GALAH survey. GALAH is a stellar spectroscopic survey, using the 2dF fibre positioner (Lewis et al. 2002) on the 3.9 m Anglo-Australian Telescope at Siding Spring Observatory. Spectra are taken with the HERMES instrument, which has a resolution of $R = 28\,000$ over 4 CCDs, covering wavelengths: 4713–4903 Å, 5648–5873 Å, 6478–6737 Å, and 7585–7887 Å (Barden et al. 2010; Brzeski, Case & Gers 2011; Heijmans et al. 2012; Farrell et al. 2014; Sheinis et al. 2015). Data Release 4 (DR4) measured stellar parameters and abundances for $\sim 900\,000$ unique stars, these parameters along with reduced spectra are published in Buder et al. (2025).

We remove anomalous stars, flagged stars, and low S/N stars. The GALAH catalogue flags stars with issues in the data analysis. In particular, the flags `flag_sp` and `flag_mg_fe` are set to zero in order to remove stars with inaccurate stellar parameters and $[\text{Mg}/\text{Fe}]$ measurements. We follow the recommendation in Buder et al. (2025) and do not apply the flag `flag_fe_h`, as up to 34 per cent of stars with detectable Fe lines have been flagged as non-detections because of an inappropriate choice of reference spectra. Rapid rotators and lithium rich giants have anomalous spectra which are not caused by a star's evolutionary state. We remove these stars through restricting $v \sin(i) < 7 \text{ km s}^{-1}$ and $A(\text{Li}) < 1.5$, where we set `flag_a_li` < 2 in order to remove stars with poor Li abundance detections. To increase the spectral fidelity, we remove stars with S/N per pixel below 10 in CCD1, S/N below 20 in CCD2, and S/N below 30 in CCD3 and CCD4. All observed spectra have been shifted to the rest frame and linearly interpolated onto the same wavelength grid.

2.2 Asteroseismic catalogue

We adopt asteroseismic observables and the stellar evolutionary state from the *TESS* (Ricker et al. 2014) and K2 (Howell et al. 2014) surveys. The frequency of maximum acoustic power (ν_{max}) and the frequency separation between overtone modes ($\Delta\nu$) are extracted from the power spectrum using the SYD pipeline (Huber et al. 2009, 2011; Yu et al. 2018), which takes initial guess ν_{max} values from Hon, Stello & Zinn (2018b). The $\Delta\nu$ values are then

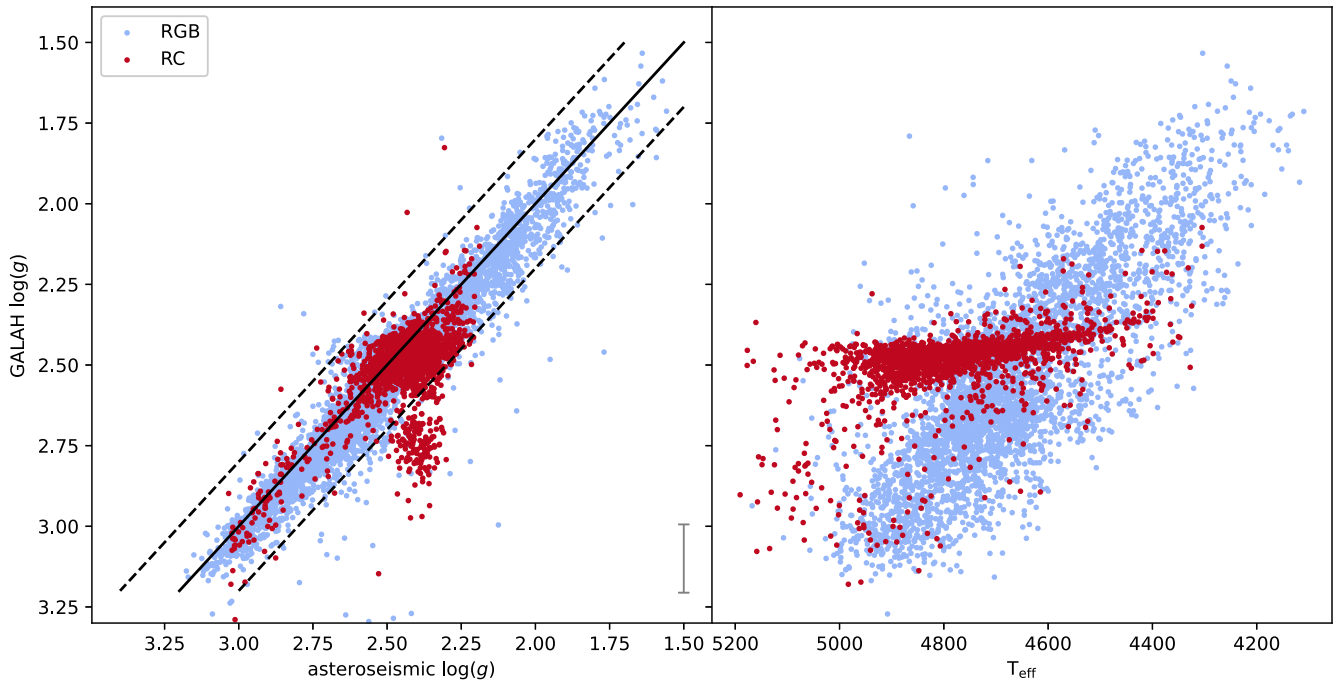


Figure 1. The left panel shows GALAH $\log(g)$ with a representative error bar compared to asteroseismic $\log(g)$. RGB stars are shown in blue whilst RC stars are shown in red. The solid black line indicates a one-to-one relationship between the two different $\log(g)$. The dashed black lines are offset by 0.2 dex from the one-to-one relationship line, and stars with $\log(g)$ difference larger than 0.2 are removed from the crossmatched catalogue. The right panel shows the T_{eff} and $\log(g)$ space that the crossmatched stars occupy after removing stars with a $\log(g)$ discrepancy.

vetted through a neural network-based classifier (Reyes et al. 2022), where we adopt the same vetting cut-off threshold of 0.5. Lastly, the evolutionary state is classified using a convolutional neural network on the folded spectrum (Hon et al. 2017; Hon et al. 2018a). Stars are classified in a continuum from 0 (a clear RGB star), to 1 (a clear RC star), with the vast majority of stars having values near 0 or 1. To avoid the ambiguous classifications near 0.5, we remove stars with classification between 0.3 to 0.7.

2.3 Crossmatched catalogue

We combine the spectroscopic data with asteroseismic data by crossmatching between stars with the smallest angular distance on sky. Using this crossmatched catalogue, we calculate asteroseismic ($\log(g)$):

$$\log(g) = \log\left(\frac{\nu_{\text{max}}}{\nu_{\text{max},\odot}}\right) + 0.5 \log\left(\frac{T_{\text{eff}}}{T_{\text{eff},\odot}}\right) + \log(g_{\odot}), \quad (1)$$

where T_{eff} is the effective temperature. We adopt Solar parameters $T_{\text{eff},\odot} = 5772$ K and $\log(g_{\odot}) = 4.438$ from Prša et al. (2016), and $\nu_{\text{max},\odot} = 3090$ μHz from Huber et al. (2011). Whilst GALAH measures $\log(g)$ through its definition ($g \propto M/R^2$) and the Stefan-Boltzmann scaling relation (Buder et al. 2025). We show GALAH $\log(g)$ compared to the asteroseismic $\log(g)$ in the left panel of Fig. 1. In general, there is good agreement between asteroseismic and spectroscopic $\log(g)$, with the exception of some RC stars with GALAH $\log(g) \approx 2.8$ and asteroseismic $\log(g) \approx 2.4$. These stars are overestimated in spectroscopic $\log(g)$ due to a mismatch of isochrones and stellar spectroscopic parameters, causing primary RC stars to be misidentified as secondary RC stars (Buder et al. 2025). GALAH measures stellar parameters iteratively and consistently, in order to retain consistency, we adopt the GALAH $\log(g)$. However, we remove stars with $\log(g)$ difference of more than 0.2 dex in

order to remove these misidentified secondary RC stars. In total we crossmatch 8163 stars; after applying the aforementioned quality cuts, 5949 stars remain, out of which 4065 are RGB stars and 1884 are RC stars. These remaining crossmatched stars are shown in the right panel of Fig. 1.

Using GALAH T_{eff} , $[\text{Fe}/\text{H}]$ and asteroseismic ν_{max} , $\Delta\nu$, we calculate mass and radius using ASFGRID (Sharma et al. 2016; Stello & Sharma 2022). ASFGRID computes mass and radius using the asteroseismic scaling relations and Stefan–Boltzmann luminosity law in addition to a corrected $\Delta\nu$ interpolated over a grid (Stello & Sharma 2022). This correction is necessary to account for the well-known fact that $\Delta\nu$ does not scale precisely with $\rho^{1/2}$ in red giant stars (see Sharma et al. 2016 for details).

3 METHOD

Stellar spectra contain information on effective temperature (T_{eff}), surface gravity [$\log(g)$], metallicity ($[\text{Fe}/\text{H}]$), and elemental abundances. As a result, it is difficult to separate RGB and RC stars using spectroscopy alone because they occupy the same T_{eff} , $\log(g)$, and $[\text{Fe}/\text{H}]$ space within uncertainties.

In order to find spectroscopic differences caused by the evolutionary state, we need to control the observed spectroscopic parameters. We achieve this with a powerful differential technique by taking pairs of RGB and RC stars with similar spectroscopic parameters – we refer to these as matched pairs. Such matched pairs of RGB and RC stars must have a difference in mass. This is because when a star evolves from the hydrogen-shell fusion state at the tip of the red giant branch to begin helium fusion, its outer envelope contracts making the star smaller and therefore hotter (i.e. as per the Stefan Boltzmann equation). If an RGB and an RC star have the same stellar parameters: T_{eff} and $\log(g)$ (which implies the stars have

the same radius), then the RGB star must be less massive than the RC star.

Whilst mass cannot be directly measured from stellar spectra, secondary effects such as changes in carbon, nitrogen, and oxygen abundances due to the first dredge up and deep mixing should be visible in the spectra of our matched pairs. At the base of the red giant branch, the convective envelope deepens into the core, dredging up material from the stellar core to the stellar surface (Iben 1964, 1967; Lambert 1981). As a consequence, CNO-processed material in the core is brought to the surface of the star, modifying the observed [C/N] and $^{12}\text{C}/^{13}\text{C}$ ratio (Lambert & Ries 1977, 1981; Smith & Lambert 1985; Halabi & Eid 2015; Masseron & Gilmore 2015; Martig et al. 2016; Ness et al. 2016). The amount of CNO-processed material in the core is a function of both stellar mass and the depth of the excursion of the convective envelope into the core. For example, lower mass stars have a lower fraction of core nitrogen abundance and a shallower excursion of the convection zone. Subsequently, lower mass stars at fixed metallicity have higher [C/N]. Deep mixing further modifies [C/N] and $^{12}\text{C}/^{13}\text{C}$ ratio in a similar direction in stars beyond the RGB bump (Lagarde et al. 2012).

3.1 Matched pairs

We define a matched pair as two stars with different evolutionary states but similar spectroscopic and observational parameters: $\delta T_{\text{eff}} < 50 \text{ K}$, $\delta \log(g) < 0.15$, $\delta [\text{Fe}/\text{H}] < 0.05 \text{ dex}$, $\delta [\text{Mg}/\text{Fe}] < 0.05 \text{ dex}$, and $\left| \frac{(S/N)_{\text{RC}}}{(S/N)_{\text{RGB}}} - 1 \right| < 0.2$ for all 4 CCDs of the HERMES instrument. These ranges are approximately based on GALAH uncertainties. These parameter criteria ensure similar spectroscopic properties and metallicity contribution from core collapse and thermonuclear supernovae (Sayeed et al. 2024). The S/N criteria ensure similar observational properties, creating RGB and RC pairs which have spectral differences driven by the remaining small parameter differences inherited from their respective evolutionary states. For each RC star, we randomly pick a RGB star that satisfies the aforementioned criteria, resulting in a total of $n = 786$ matched pairs. We note that unique RGB stars can be present in more than one matched pair, this is further discussed in Appendix A1.

For each matched pair, we calculate a difference spectrum, δf , given by:

$$\delta f_i = f_{\text{RC},i} - f_{\text{RGB},i}, \quad (2)$$

where $i \in \{1, \dots, n = 786\}$ is the index of the matched pair, f_{RC} is the normalized flux of the RC star, and f_{RGB} is the normalized flux of the RGB star. The difference spectra on the H_α region for all matched pairs is shown in the bottom panel of Fig. 2.

The median RGB spectrum, $\text{med}(f_{\text{RGB}})$, is given by

$$\text{med}(f_{\text{RGB}}) = \text{med}(f_{\text{RGB},1}, \dots, f_{\text{RGB},n}). \quad (3)$$

This median RGB spectrum is shown in the top panel of Fig. 2 as a reference for the typical spectral lines present in the H_α wavelength region. The median RC spectrum is not visibly different from the median RGB spectrum and is therefore not shown.

We want to study the difference in spectra caused by evolutionary state, not individual matched pairs, so we calculate the median difference spectrum, Δf , defined as

$$\Delta f = \text{med}(\delta f_1, \dots, \delta f_n). \quad (4)$$

The median difference spectrum for the H_α region is shown in the middle panel of Fig. 2. The standard deviation of Δf is a representation of how much Δf deviates from zero, and

is also shown in the middle panel of Fig. 2. For the H_α region, this is 0.2 per cent, which is a small difference. Therefore, the rest of the methodology investigates the significance of this difference.

3.2 Control pairs

To quantify the effect of the matched methodology on the median difference spectrum, we find two additional sets of matched stars. We take the RGB stars from the existing matched pairs, and find their matched RGB stars using the same criteria. This gives 247 RGB to RGB matched pairs. We repeat this process for RC stars, and find 650 RC to RC matched pairs. The median difference spectra for the RGB–RGB matches and RC–RC matches are shown in the middle panel of Fig. 2. These additional difference spectra are smaller compared to the RC–RGB matched set, indicating that the difference in RC and RGB spectra is not caused by our matched pair selection. Instead, some of this difference must be due to their different evolutionary states.

Despite selecting matched pairs that have similar stellar parameters, there is still an expected overall non-zero difference in stellar parameters between matched pairs. Table 1 shows the median and mean difference in stellar parameters for all three matched sets. This non-zero mean difference further implies that the distribution in stellar parameter difference is non-uniform, as shown in Fig. 3. Despite the non-zero and non-uniform distribution of the difference in stellar parameters, it cannot fully account for the signal in the RC–RGB median difference spectrum that we see. The distribution, median difference, and mean difference in stellar parameters are very similar between the three matched sets. Therefore, we expect the impact on the median difference spectrum to be similar. However, the RC–RGB matched median difference spectrum shows a larger signal than the RC–RC and RGB–RGB matched median difference spectrum, implying that some of the RC–RGB median difference spectrum is caused by evolutionary state parameters that are not included in our matched criteria.

Furthermore, we use PYSME (Wehrhahn, Piskunov & Ryabchikova 2023) to calculate synthetic spectra at 10 K apart, and show that the median difference spectrum cannot be fully explained by the non-zero δT_{eff} reported in Table 1. The PYSME difference spectrum is shown in the middle panel of Fig. 2. However, we caution that spectral synthesis of the H_α line in giant stars differs from the observed spectra by up to 20 per cent – a well known issue (Bergemann et al. 2016) – therefore, the flux difference that a 10 K T_{eff} perturbation creates could be different in reality. We further explore the impact of selection effects in the matched methodology on the median difference spectrum in Appendix A.

3.3 Flux uncertainties

The median difference spectrum is more significant than its standard error and uncertainty. The standard error of the median difference spectra, $\text{std}_{\Delta f}$, is approximated by

$$\text{std}_{\Delta f} = \frac{\text{std}(\delta f_1, \dots, \delta f_n)}{\sqrt{n}}. \quad (5)$$

The uncertainty of the median difference spectrum assuming a Gaussian distribution, $\sigma_{\Delta f}$, is approximated by

$$\sigma_{\Delta f} = \left(\sum_{i=1}^n \frac{1}{\sigma_{\text{RGB},i}^2 + \sigma_{\text{RC},i}^2} \right)^{-\frac{1}{2}}, \quad (6)$$

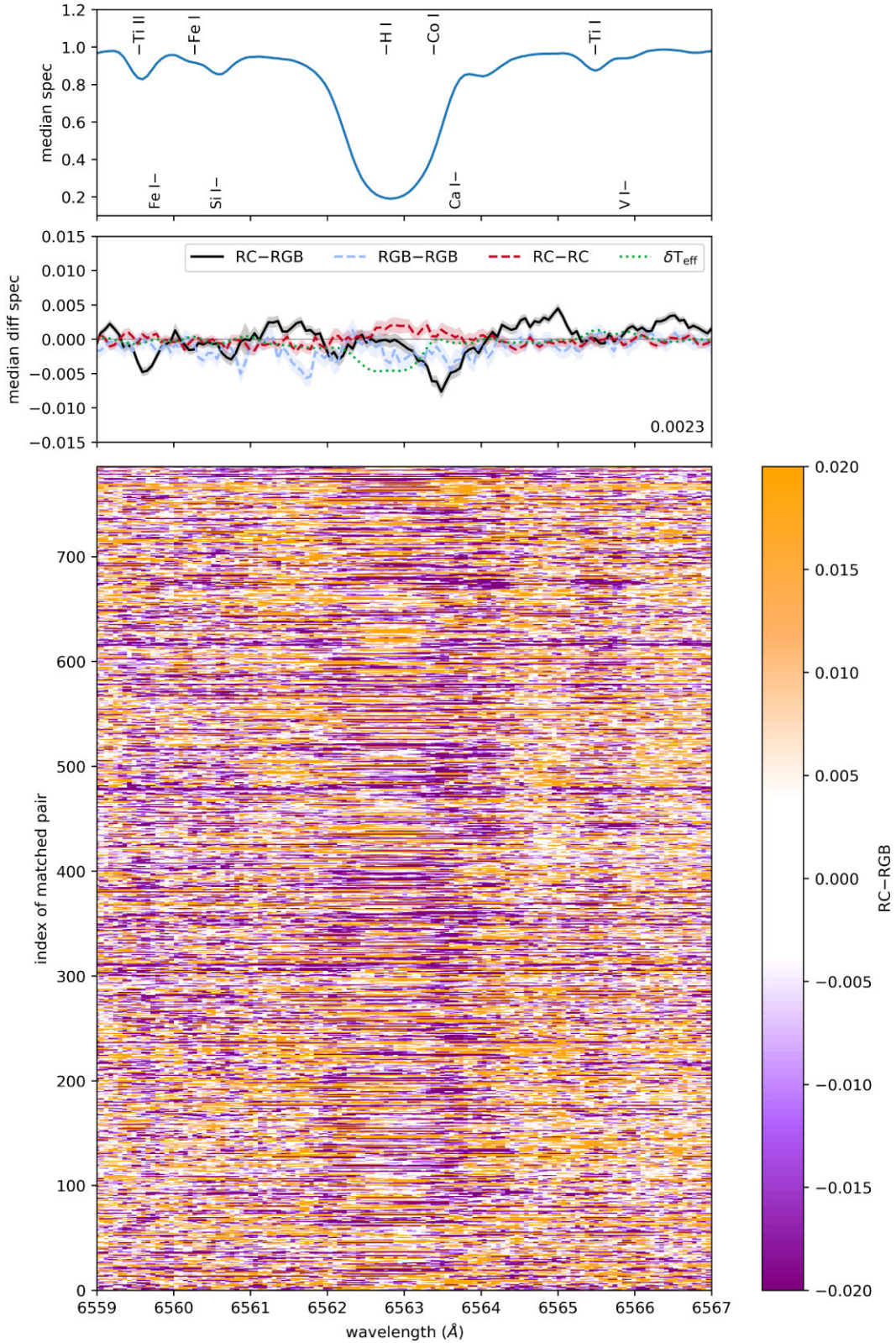
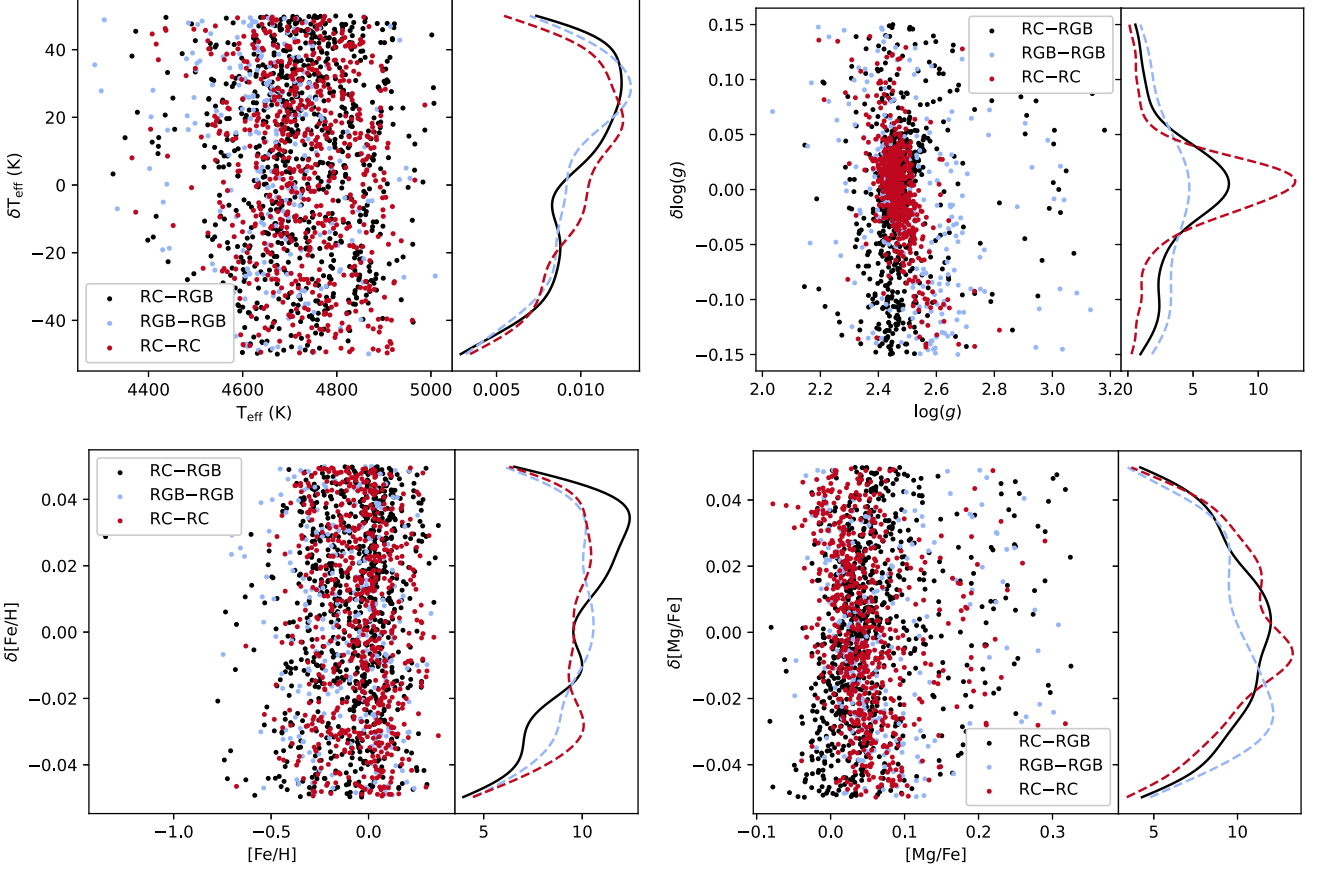


Figure 2. The top panel is the median spectrum for RGB stars with labelled atomic spectral features, to show the spectral features in this wavelength region. The middle panel shows the median difference spectra: RC–RC stars (solid black), RGB–RC stars (dashed blue), and RC–RC stars (dashed red). The standard error on the difference spectra (equation 5) is shown for all 3 difference spectra in the corresponding colours. In order to show the deviation of the RC–RC median difference spectrum from zero, we plot a thin grey line at zero, and report the standard deviation for the H_{α} wavelength region on the bottom right. The difference spectrum for a change of 10 K in T_{eff} is calculated through PYSME and shown in dotted green. The bottom panel shows the difference spectra for all RC–RC matched pairs, with a truncated colour bar.

Table 1. The median and mean difference in stellar parameters for the full crossmatched data set and the matched sets. These differences are of a similar magnitude, and so the impact of stellar parameter on the median difference spectra is similar between the matched sets.

Parameter	Units	Full		RC–RGB		RGB–RGB		RC–RC	
		Median	Mean	Median	Mean	Median	Mean	Median	Mean
T_{eff}	K	87.04	101.31	10.87	7.17	12.12	7.23	7.96	4.89
$\log(g) / 10^{-2}$	$\log(\text{cm s}^{-2})$	-11	-4.7	-0.19	-0.75	-1.4	-1.4	0.32	-0.0047
$[\text{Fe}/\text{H}] / 10^{-2}$	dex	15	13	0.92	0.58	0.32	0.26	0.25	0.22
$[\text{Mg}/\text{Fe}] / 10^{-3}$	dex	-52	-53	-0.25	-0.35	-5.6	-3.1	0.38	0.94

**Figure 3.** The correlation between RC stellar parameters and the stellar parameter difference, shown in the left panel of each subfigure; and the distribution of stellar parameter difference modelled using a kernel density estimator, shown in the right panel of each subfigure. Each subfigure shows a different stellar parameter: T_{eff} and $\log(g)$ in the top row, $[\text{Fe}/\text{H}]$ and $[\text{Mg}/\text{Fe}]$ in the bottom row. All three matched sets are shown: RC–RGB (solid black), RGB–RGB (dashed blue), and RC–RC (dashed red).

where σ_{RGB} is the flux error of the RGB star and σ_{RC} is the flux error of the RC star. See Appendix B for a derivation of equation (6). $\sigma_{\Delta f}$ is smaller than $\text{std}_{\Delta f}$ by a factor of 1.5 in the H_{α} region; therefore, we show the larger error, $\text{std}_{\Delta f}$, from equation (5) in the middle panel of Fig. 2 as shaded regions for all three matched sets. Despite the sub-per cent magnitude of the median difference spectrum, the RC–RGB median difference spectrum is significant when compared to the uncertainty estimates.

4 RESULTS

The median difference spectrum for RC–RGB matched stars is non-zero and larger than the flux error and control matched sets for many wavelengths within GALAH spectra. If additional parameters of stellar evolutionary state not included in our criteria (i.e. mass) do

not affect spectra, then we expect the median difference spectrum to be zero, as the matched pairs were selected to have similar stellar parameters. Therefore, the spectroscopic signature of evolutionary state is where the median difference spectrum is non-zero. We find that the RC–RGB matched median difference spectrum deviates from zero where C_2 and CN bands are located, and in the line-width near the inner wings of the H_{α} and H_{β} lines.

4.1 C_2 and CN

We find that the spectroscopic signature of evolutionary state correlates with C_2 and CN molecular band heads. Fig. 4 shows the Swan bands in CCD1 and CN bands in CCD4 captured by GALAH. We extract a stellar line list at $T_{\text{eff}} = 4750 \text{ K}$, $\log(g) = 2.5$, microturbulence $v_{\text{mic}} = 1.46 \text{ km s}^{-1}$ from VALD3 (Ryabchikova et al. 2015),

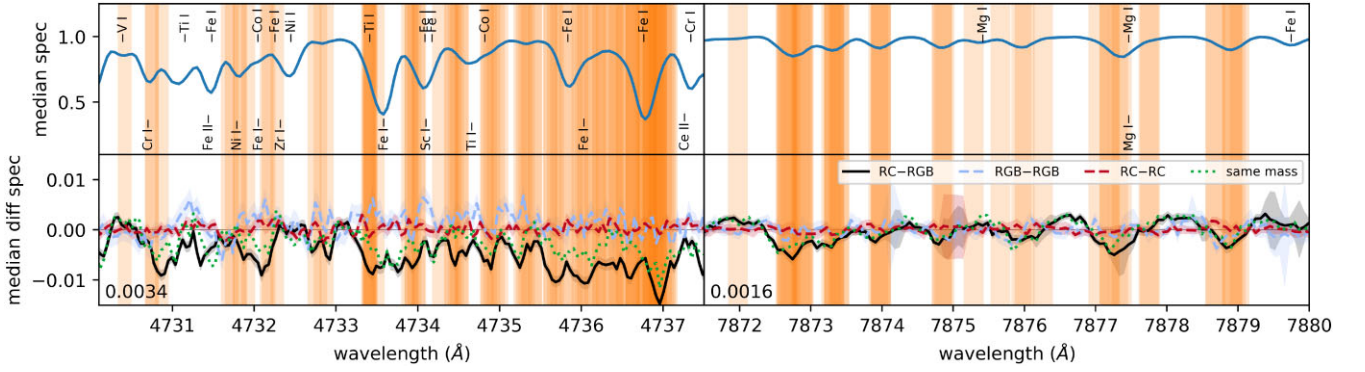


Figure 4. The median difference spectrum is correlated with the Swan bands (left) and the CN bands (right). For each subfigure, the top panel shows the median RGB spectrum with labelled atomic spectral features, as a reference for the spectral features of this wavelength region. The bottom panel shows the median difference spectrum for RC-RGB stars (solid black); the uncertainty in the median difference spectrum given by equation (6) (shaded grey); RGB-RGB matched pairs (dashed blue) and RC-RC matched pairs (dashed red). C₂ and CN molecular lines are shown in both panels as orange shaded rectangles, where darker orange indicates overlapping C₂ and CN lines. The median difference spectrum for the mass constrained matched set is shown in the bottom panel (dotted green) of each subfigure and is discussed in Section 5.

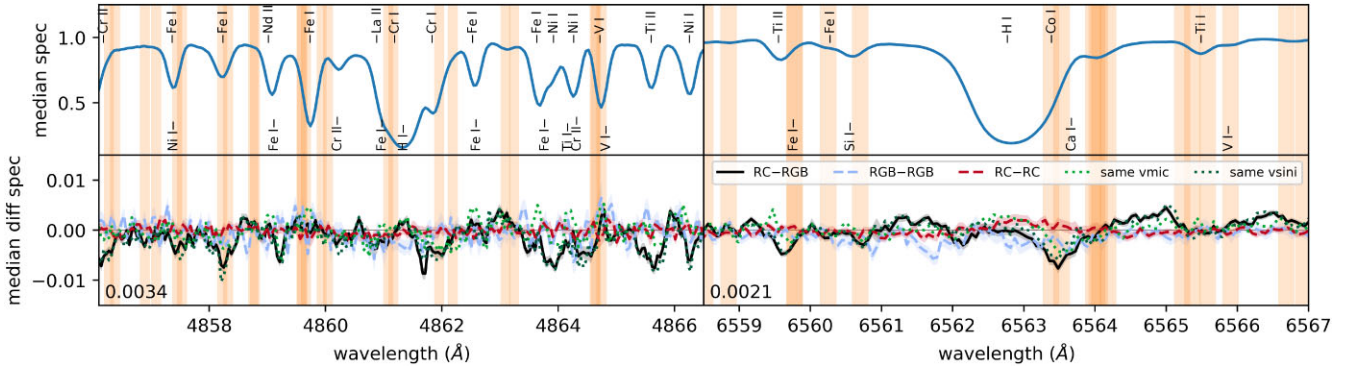


Figure 5. The line-width of the H_α (right) and H_β (left) line is different between RC and RGB stars of similar stellar parameters. Labels follow Fig. 4. The median difference spectrum for matched sets matched to agree in v_{mic} and $v \sin(i)$ respectively, in addition to T_{eff} , $\log(g)$, $[\text{Fe}/\text{H}]$, and $[\text{Mg}/\text{Fe}]$, are shown in the bottom panel (dotted green lines) of each subfigure and are discussed in Section 5.

and show C₂ and CN molecular lines (Obbarius & Kock 1982; Lawler et al. 2001; Pickering, Thorne & Perez 2001) with VALD3 central depth parameter larger than 0.005 shaded in orange. The width of the shaded region (w) is indicative of the instrumental broadening of the molecular line, and is defined as $w = \lambda_l/R$, where λ_l is the wavelength of the molecular line, and R is the resolution of the spectra. The median difference spectrum deviates from zero at the same wavelengths as the C₂ and CN molecular band heads. The median difference spectrum is negative, implying that the C₂ and CN features are stronger in RGB stars than RC stars. We note that despite C₂ and CN molecular lines being present at many other wavelengths, the correlation with the median difference spectrum is only clear in the band heads. Band heads are where many lines from the same molecular species are present; therefore, this correlation only being clear in the band heads implies that every molecular line carries a small bit of evolutionary state signal which only becomes visible when many of these lines are close in wavelength.

4.2 Line-width

We see a difference in the line-width of the H_α and H_β lines in the spectrum, shown in Fig. 5. The median difference spectrum is close to zero in the core of the line, but negative in the wings of the line, implying that lines in RC stars are broader than in RGB

stars. Whilst we only show H_α and H_β in Fig. 5, the line-width difference can be seen from other lines in the spectrum. However, it becomes difficult to distinguish this difference in line-width when the lines are heavily blended. The H_α median difference spectrum is visibly asymmetric, likely caused by both a difference in the line-width and two CN molecular lines situated in the red wing of the H_α line at $\lambda \approx 6563.4$ Å. Whilst the smaller negative median difference spectrum in the blue wing of the H_α line is caused only by a difference in the line-width.

4.3 Formation depth

The median difference spectrum does not appear to be driven by the formation depth. We define the formation depth per wavelength as the geometric depth at which the majority of photons for a wavelength originate. We approximate the geometric depth using the optical depth (τ_λ). Specifically, we calculate τ_{5000} where $\tau_\lambda = 1$ for every wavelength using Korg (Wheeler et al. 2023), and adopt $\log_{10}(\tau_{5000})$ as our proxy for formation depth. We compare the formation depth to the median RGB spectrum and median difference spectrum in Fig. 6. We find that the formation depth is higher where the normalized flux is closer to zero, e.g. at $\lambda = 4736.8$ Å, as expected. However, we do not see a correlation with the amplitude of the median difference spectrum, indicating that the origin of the spectroscopic signal of

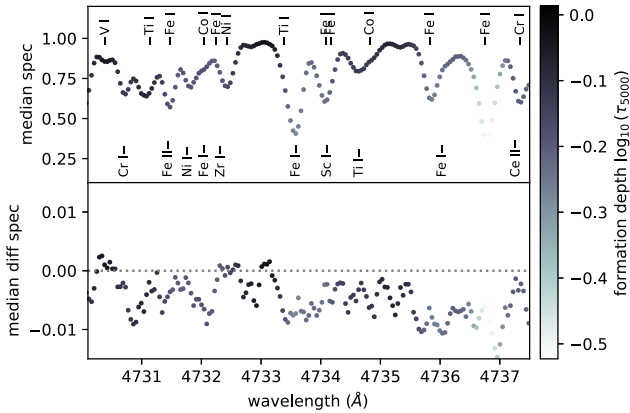


Figure 6. The top panel shows the median RGB spectrum with labelled atomic spectral features coloured by formation depth. Where normalized flux is closer to 1, the formation depth is geometrically higher, and vice versa. The bottom panel shows the median difference spectrum coloured by formation depth. There is no correlation between the formation depth and the median difference spectrum.

evolutionary state is not related to the geometric height of the atmosphere.

5 DISCUSSION

We find spectroscopic signatures of evolutionary state in the C_2 and CN bands, and in the wings of the H_α and H_β lines. In this section, we relate these results to physical parameters.

5.1 C_2 and CN

For a pair of RGB and RC stars to have similar spectroscopic stellar parameters, they must have different masses. We find that the average mass of our RGB matched stars is $1.39 M_\odot$ with standard deviation $0.35 M_\odot$, while for our RC matched stars it is $1.14 M_\odot$ with standard deviation $0.33 M_\odot$. To further investigate the impact of mass on our median difference spectrum, we make an additional matched set, using the same matched criteria as previously, except with an extra mass restriction of $\delta M < 0.3 M_\odot$. The median difference spectrum of this mass restricted matched set is shown in the bottom panel of Fig. 4. The amplitude of the mass restricted median difference spectrum is smaller than the amplitude of the median difference spectrum in the C_2 and CN bands, indicating that part of the evolutionary state signal in these regions is due to a difference in mass.

The strength of the C_2 and CN lines depends on the molecular equilibrium and isotopic ratio of relevant species. The CNO abundances in giant stars have been impacted by the first dredge up, and RC CNO abundances have been further modified through deep mixing. Similarly, C isotopic ratios are modified through both the first dredge up and deep mixing. Furthermore, the amount that these abundances and isotopic ratios change by depends on mass and metallicity. In Table 2 we report CNO abundances and isotopic ratios at the average mass and $\log(g)$ of our RGB and RC stars, predicted using stellar evolution models incorporating thermohaline instability and rotation-induced mixing (Lagarde et al. 2012).

To investigate the impact of the first dredge up and deep mixing on the median difference spectrum, we compute synthetic spectra using the 1D local thermodynamic equilibrium spectrum synthesis

Table 2. The predicted CNO abundance and C isotopic ratios for stellar evolution models representing the mean mass of RGB and RC stars in our RC–RGB matched pairs. We adopt stellar evolution models from Lagarde et al. (2012) at $\log(g) = 2.5$ and Solar metallicity ($Z = 0.014$), these models incorporate thermohaline instability and rotation-induced mixing. These values are linearly interpolated from grid models.

	Units	RGB	RC
M	M_\odot	1.39	1.14
A(C)	dex	8.286	8.283
A(N)	dex	8.187	8.193
A(O)	dex	8.710	8.714
$^{12}\text{C} : ^{13}\text{C}$	–	95 : 5	89 : 11

code MOOG (Snedden 1973). We adopt the average T_{eff} , $\log(g)$, $[\text{Fe}/\text{H}]$, and $[\text{Mg}/\text{Fe}]$ of the RC–RGB matched set, along with CNO abundances and isotopic ratios from Lagarde et al. (2012) as listed in Table 2, and apply solar abundances for all other elements. We run the 2017 version of MOOG¹ with the PYMOOGI GUI (Adamow 2017) using ATLAS9 model atmospheres interpolated using PyKMOD² and a line list generated by linemake (Placco et al. 2021). This line list adopts C_2 data from Ram et al. (2014), CN data from Sneden et al. (2014) and CH data from Masseron et al. (2014). We show two MOOG synthetic spectra in the top panel of Fig. 7: RGB abundances and isotopic ratio, and RC abundances and isotopic ratio. The difference between the MOOG RC spectrum and RGB spectrum then corresponds to our observed RC–RGB median difference spectrum and is shown in the bottom panel of Fig. 7. In both the Swan and CN bands, the synthetic difference spectrum is positive, despite the observed negative median difference spectrum. The magnitude of the MOOG difference spectrum is roughly the same as the observed median difference spectrum in the Swan band. However, in the CN band, the magnitude of the MOOG difference spectrum is smaller compared to the observed median difference spectrum.

To determine whether the median difference spectrum depends on the change in CNO abundances or C isotopic ratio, we compute two additional synthetic spectra with MOOG: RGB abundances with RC isotopic ratio, and RC abundances with RGB isotopic ratio. The synthetic difference spectrum with RC abundances but different isotopic ratios, and the difference spectrum with the RC isotopic ratio but different abundances, are shown in the bottom panel of Fig. 7. In both wavelength regions, the MOOG RC–RGB difference spectrum is approximately the sum of both the change in abundance and the change in isotopic ratio. In the Swan bands, the change in isotopic ratio is larger than the change in abundance; whilst in the CN bands wavelength region, the change in isotopic ratio is smaller than the change in abundance.

We caution that these synthetic spectra were computed using abundances from stellar evolution models (Lagarde et al. 2012), and do not match GALAH observations, which measure abundances using radiative transfer codes (Buder et al. 2025). Notably, the synthetic spectra differ from the observed spectra by up to 20 per cent at many wavelengths, as shown in the top panel of Fig. 7. The difference spectra could change; if the synthetic spectra are fitted to the observed normalised flux instead. However, synthetic spectra are not accurate to the sub-per cent level, and will introduce model errors into such fits.

¹<https://github.com/alexji/moog17scat>

²<https://github.com/kolecki4/PyKMOD>

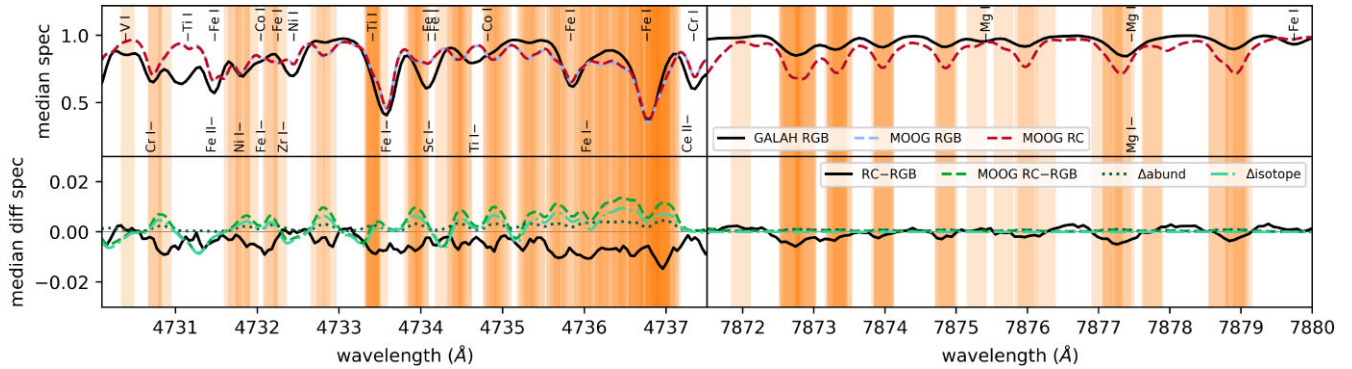


Figure 7. The top panel shows the normalized flux for GALAH observed median RGB spectrum (black) with labelled atomic spectral features, MOOG synthetic spectrum calculated with RGB abundances and isotopic ratio (blue), and RC abundances and isotopic ratio (red) based on stellar evolution models (Lagarde et al. 2012). Bottom panel shows difference spectra: observed RC–RGB matched pairs (black), MOOG synthetic RC–RGB using Lagarde et al. (2012) abundances and isotopic ratios (dashed green), MOOG synthetic different abundances same isotopic ratios (dotted dark green), and MOOG synthetic same abundances different isotopic ratios (dash dotted light green).

5.2 Line-width

Line-width is controlled by stellar rotation, $v \sin(i)$, and velocity fields within the stellar atmosphere, which are approximated by v_{mac} and v_{mic} within 1D model atmospheres. GALAH measures v_{mic} and $v \sin(i)$, where $v \sin(i)$ represents the line broadening from both v_{mac} and stellar rotation. In order to determine the parameter driving the line broadening observed, we make additional matched sets. The $v \sin(i)$ constrained matched set incorporates an extra constraint of $\delta(v \sin(i)) < 0.8 \text{ km s}^{-1}$, whilst the v_{mic} constrained matched set incorporates an extra constraint of $\delta v_{\text{mic}} < 0.07 \text{ km s}^{-1}$. We find that the v_{mic} constrained matched set has a median difference spectrum closer to zero whilst the $v \sin(i)$ constrained matched set does not (shown in bottom panel of Fig. 5), implying that the difference in line-width we observe is due to v_{mic} and not $v \sin(i)$. Furthermore, the median difference spectrum is mostly negative for both H_{α} and H_{β} , and the integrated median difference spectrum is: -0.23 \AA , -0.12 \AA , -0.08 \AA , -0.05 \AA , respectively for the 4 CCDs of HERMES. $v \sin(i)$ conserves the line strength whilst v_{mic} does not, further implying that the difference in line-width is caused by v_{mic} and not $v \sin(i)$. In addition, we find that v_{mic} improves random forest classifier accuracy (see Appendix C), whilst $v \sin(i)$ does not.

Our RGB matched stars have the same $\log(g)$ but are more massive than RC matched stars, as a result, these RGB matched stars must also have a larger radius. We incorporate an extra constraint of $\delta R < 0.6 R_{\odot}$ and compute the median difference spectrum. We find that the median difference spectrum for this radius restricted matched set is similar to the RC–RGB matched set, indicating that radius is not the cause of the observed difference in line-width.

v_{mic} parametrizes small scale velocity motions in the stellar photosphere. We measure an average $v_{\text{mic}} = 1.42 \text{ km s}^{-1}$ with $0.00014 \text{ km s}^{-1}$ standard error in the mean and $\sigma_{v_{\text{mic}}} = 0.11 \text{ km s}^{-1}$ in our RGB matched stars, and an average $v_{\text{mic}} = 1.49 \text{ km s}^{-1}$ with $0.00012 \text{ km s}^{-1}$ standard error in the mean and $\sigma_{v_{\text{mic}}} = 0.098 \text{ km s}^{-1}$ in our RC matched stars. As previously discussed, our RC matched stars are lower mass than the corresponding RGB matched stars. Sphericity effects related to this mass difference cause the RC star’s temperature to be lower by up to 13 K in the line forming regions, as shown in Fig. 8. This lower temperature could cause RC stars to have weaker H_{α} wings than RGB stars. However, density and photospheric radius also change with mass, so the overall impact of mass on line strength is unclear. Although 3D hydrodynamic model atmospheres simulate convection from first principles and do not use

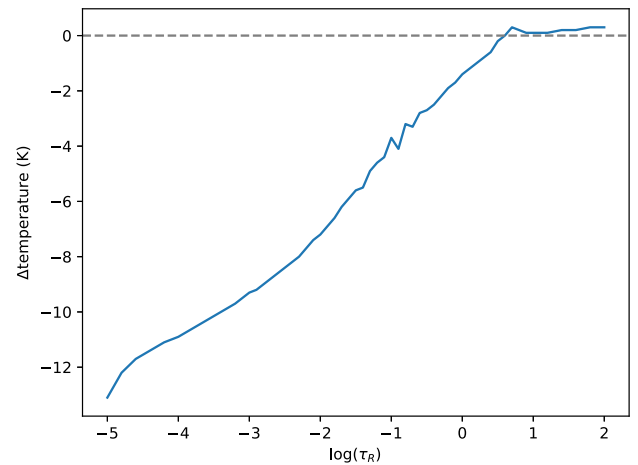


Figure 8. Temperature difference at different Rosseland optical depth for spherically symmetric stellar atmosphere models with $M = 1 M_{\odot}$ compared to $M = 2 M_{\odot}$. Model atmospheres from the MARCS grid (Gustafsson et al. 2008), located at $T_{\text{eff}} = 5000 \text{ K}$, $\log(g) = 2$, $[\text{Fe}/\text{H}] = 0$, $v_{\text{mic}} = 1 \text{ km s}^{-1}$.

the v_{mic} parameter, RC and RGB stars are modelled in the same way. Therefore, there is currently no known difference in the velocity field of RC stars compared to RGB stars. It has been shown that changes in CNO abundances can affect the atmospheric structure (Gallagher et al. 2017; Zhou et al. 2023), however, no detailed studies have been conducted on the impact of CNO abundances on the velocity field in 3D stellar atmospheres.

The core of the H_{α} and H_{β} lines form in the chromosphere and are linked to stellar activity and age, where younger stars are more active. The average age of our RGB stars are 6.08 Gyr with standard deviation 2.23 Gyr, and 6.38 Gyr with standard deviation 2.04 Gyr for RC stars, estimated through isochrone fitting (see Buder et al. 2025 for details). Whilst we do not expect the majority of our stars to be active, there may be some younger RGB stars which have a more active chromosphere; studies have shown that young stars appear to exhibit elevated v_{mic} values (e.g. Baratella et al. 2020). However, this is the opposite of what we observe in our matched set. We note that the majority of our matched RC stars are primary RC stars; using a $2 M_{\odot}$ cut-off (Girardi 1999), only 16 pairs contain a secondary RC

stars. Therefore, the differences that we see between RC and RGB stars are mainly driven by primary RC stars.

In addition to v_{mic} , mass also impacts the median difference spectrum in the red wing of the H_{α} line. Bergemann et al. (2016) showed that the line-width of the H_{α} line is broader for lower mass stars. We find a similar result, where our RC stars have broader H_{α} and H_{β} lines along with lower mass compared to our RGB stars. This mass dependence is likely due to the CN molecular lines present in the red wing of the H_{α} line, where the strength of the CN molecular lines is related to mass (see Section 5.1). In particular, the median difference spectrum in the red wing goes closer to zero when mass is restricted, reducing the asymmetry.

6 CONCLUSIONS

In this work, we present a model-free analysis of optical spectroscopic evolutionary state signals, and find sub-percent amplitude changes at all wavelengths in the spectra caused by evolutionary state. Despite the small signal, we show that this is significant compared to control sets and uncertainty estimates. For RC and RGB stars at similar stellar parameters, we find stronger C_2 and CN features in RGB stars compared to RC stars, caused by a difference in stellar evolution and mass; and we find that the H_{α} and H_{β} lines are broader in RC stars than RGB stars, caused by a difference in microturbulence, as measured by GALAH. We find that rotational velocity, radius, and formation depth do not impact spectroscopic evolutionary state signals. This analysis is demonstrative of the utility of large surveys to undertake a data-driven and model-independent population analysis, revealing the subtle differences between RC and RGB stars in the line forming regions of the stellar atmosphere. Future studies classifying stellar evolutionary states for spectroscopic surveys can use these results to inform which parameters and wavelengths contain evolutionary state information. Given the small amplitudes of these residual differences, analysis of these in individual stars would require $S/N > 200$, at resolution ≈ 28000 .

Whilst we identify some of the sub-percent amplitude changes and link them to evolutionary state, there are still many wavelength regions which we did not study. Future work could look at these other wavelengths and find additional parameters that these amplitude differences depend on, and extend this model-free technique to infrared spectra where there are more carbon-related molecular lines.

ACKNOWLEDGEMENTS

We thank Adam Wheeler for providing formation depths for Section 4.3. We thank Alexander Soen for various discussions regarding equation (6) and helping formulate Appendix B. This work has made use of the VALD data base, operated at Uppsala University, the Institute of Astronomy RAS in Moscow, and the University of Vienna. EW and KL acknowledge funds from the European Research Council under the European Union's Horizon 2020 research and innovation program (grant agreement 852977). KL and TN acknowledge support from the Knut and Alice Wallenberg Foundation. DS is supported by the Australian Research Council (DP190100666). XF acknowledges National Natural Science Foundation of China (NSFC) no. 12203100, the China Manned Space Project with No. CMS-CSST-2021-A08. PBD is supported by the Australian Government International Training Program (iRTP) Scholarship.

DATA AVAILABILITY

The GALAH DR4 spectroscopic and catalogue data are available on <https://www.galah-survey.org/dr4/overview/>. Other data available upon request.

REFERENCES

- Abdurro'uf et al., 2022, *ApJS*, 259, 35
- Adamow M. M., 2017, in American Astronomical Society Meeting Abstracts #230. p.216.07
- Banks K. A. et al., 2023, *MNRAS*, 523, L80
- Banks K. A. et al., 2024, *MNRAS*, 529, 3912
- Baratella M. et al., 2020, *A&A*, 634, A34
- Barban C. et al., 2007, *A&A*, 468, 1033
- Barden S. C. et al., 2010, in McLean I. S., Ramsay S. K., Takami H., eds, Proc. SPIE Conf. Ser. Vol. 7735, Ground-based and Airborne Instrumentation for Astronomy III. SPIE, Bellingham, p. 773509
- Beck P. G. et al., 2011, *Science*, 332, 205
- Bedding T. R. et al., 2011, *Nature*, 471, 608
- Bergemann M. et al., 2016, *A&A*, 594, A120
- Brzeski J., Case S., Gers L., 2011, in Hatheway A. E., ed., Proc. SPIE Conf. Ser. Vol. 8125, Optomechanics 2011: Innovations and Solutions. SPIE, Bellingham, p. 812504
- Buder S. et al., 2025, *PASA* 42 e051
- Casey A. R. et al., 2019, *ApJ*, 880, 125
- Chaplin W. J., Miglio A., 2013, *ARA&A*, 51, 353
- Charbonnel C., 1994, *A&A*, 282, 811
- Charbonnel C., 2005, in Barnes T. G., III, Bash F. N., eds, ASP Conf. Ser. Vol. 336, Cosmic Abundances as Records of Stellar Evolution and Nucleosynthesis. Astron. Soc. Pac., San Francisco, p. 119
- Charbonnel C., Lagarde N., 2010, *A&A*, 522, A10
- Charbonnel C., Zahn J. P., 2007, *A&A*, 467, L15
- Charbonnel C., Brown J. A., Wallerstein G., 1998, *A&A*, 332, 204
- De Ridder J., Barban C., Carrier F., Mazumdar A., Eggenberger P., Aerts C., Deruyter S., Vanautgaerden J., 2006, *A&A*, 448, 689
- De Silva G. M. et al., 2015, *MNRAS*, 449, 2604
- Elsworth Y., Hekker S., Basu S., Davies G. R., 2017, *MNRAS*, 466, 3344
- Farrell T. J., Birchall M. N., Heald R. W., Shortridge K., Vuong M. V., Sheinis A. I., 2014, in Chiozzi G., Radziwill N. M., eds, Proc. SPIE Conf. Ser. Vol. 9152, Software and Cyberinfrastructure for Astronomy III. SPIE, Bellingham, p. 915223
- Frandsen S. et al., 2002, *A&A*, 394, L5
- Gallagher A. J., Caffau E., Bonifacio P., Ludwig H. G., Steffen M., Homeier D., Plez B., 2017, *A&A*, 598, L10
- Gilroy K. K., 1989, *ApJ*, 347, 835
- Girardi L., 1999, *MNRAS*, 308, 818
- Gratton R. G., Sneden C., Carretta E., Bragaglia A., 2000, *A&A*, 354, 169
- Gustafsson B., Edvardsson B., Eriksson K., Jørgensen U. G., Nordlund Å., Plez B., 2008, *A&A*, 486, 951
- Haddouchi M., Berrado A., 2019, in 2019 1st International Conference on Smart Systems and Data Science (ICSSD). p. 1
- Halabi G. M., Eid M. E., 2015, *MNRAS*, 451, 2957
- Hawkins K., Leistedt B., Bovy J., Hogg D. W., 2017, *MNRAS*, 471, 722
- Hawkins K., Ting Y.-S., Walter-Rix H., 2018, *ApJ*, 853, 20
- Heijmans J. et al., 2012, in McLean I. S., Ramsay S. K., Takami H., eds, Proc. SPIE Conf. Ser. Vol. 8446, Ground-based and Airborne Instrumentation for Astronomy IV. SPIE, Bellingham, p. 84460W
- Hon M., Stello D., Yu J., 2017, *MNRAS*, 469, 4578
- Hon M., Stello D., Yu J., 2018a, *MNRAS*, 476, 3233
- Hon M., Stello D., Zinn J. C., 2018b, *ApJ*, 859, 64
- Howell S. B. et al., 2014, *PASP*, 126, 398
- Howell M., Campbell S. W., Stello D., De Silva G. M., 2024, *MNRAS*, 527, 7974
- Huber D., Stello D., Bedding T. R., Chaplin W. J., Arentoft T., Quirion P. O., Kjeldsen H., 2009, *Commun. Asteroseismology*, 160, 74
- Huber D. et al., 2011, *ApJ*, 743, 143

Iben I., Jr, 1964, *ApJ*, 140, 1631
 Iben I., Jr, 1967, *ARA&A*, 5, 571
 Lagarde N., Decressin T., Charbonnel C., Eggenberger P., Ekström S., Palacios A., 2012, *A&A*, 543, A108
 Lagarde N. et al., 2019, *A&A*, 621, A24
 Lambert D. L., 1981, in Iben I., Jr, Renzini A., eds, Springer Dordrecht Astrophysics and Space Science Library Vol. 88, Physical Processes in Red Giants. p. 115
 Lambert D. L., Ries L. M., 1977, *ApJ*, 217, 508
 Lambert D. L., Ries L. M., 1981, *ApJ*, 248, 228
 Lawler J. E., Wickliffe M. E., den Hartog E. A., Sneden C., 2001, *ApJ*, 563, 1075
 Lewis I. J. et al., 2002, *MNRAS*, 333, 279
 Lu Y. L., Ness M. K., Buck T., Zinn J. C., Johnston K. V., 2022, *MNRAS*, 512, 2890
 Martell S. L., Smith G. H., Briley M. M., 2008a, *PASP*, 120, 7
 Martell S. L., Smith G. H., Briley M. M., 2008b, *AJ*, 136, 2522
 Martell S. L. et al., 2021, *MNRAS*, 505, 5340
 Martig M. et al., 2016, *MNRAS*, 456, 3655
 Masseron T., Gilmore G., 2015, *MNRAS*, 453, 1855
 Masseron T., Hawkins K., 2017, *A&A*, 597, L3
 Masseron T. et al., 2014, *A&A*, 571, A47
 Montalbán J., Miglio A., Noels A., Scuflaire R., Ventura P., 2010, *ApJ*, 721, L182
 Mosser B. et al., 2011, *A&A*, 532, A86
 Mosser B. et al., 2012, *A&A*, 540, A143
 Ness M., Hogg D. W., Rix H. W., Martig M., Pinsonneault M. H., Ho A. Y. Q., 2016, *ApJ*, 823, 114
 Obbair H. U., Kock M., 1982, *J. Phys. B At. Mol. Phys.*, 15, 527
 Pedregosa F. et al., 2011, *J. Mach. Learn. Res.*, 12, 2825
 Pickering J. C., Thorne A. P., Perez R., 2001, *ApJS*, 132, 403
 Pinsonneault M. H. et al., 2014, *ApJS*, 215, 19
 Pinsonneault M. H. et al., 2025, *ApJS*, 276, 69
 Placco V. M., Sneden C., Roederer I. U., Lawler J. E., Den Hartog E. A., Hejazi N., Maas Z., Bernath P., 2021, *Res. Notes Am. Astron. Soc.*, 5, 92
 Prša A. et al., 2016, *AJ*, 152, 41
 Ram R. S., Brooke J. S. A., Bernath P. F., Sneden C., Lucatello S., 2014, *ApJS*, 211, 5
 Recio-Blanco A., de Laverny P., 2007, *A&A*, 461, L13
 Reyes C., Stello D., Hon M., Zinn J. C., 2022, *MNRAS*, 511, 5578
 Ricker G. R. et al., 2014, in Oschmann Jacobus M. J., Clampin M., Fazio G. G., MacEwen H. A., eds, Proc. SPIE Conf. Ser. Vol. 9143, Space Telescopes and Instrumentation 2014: Optical, Infrared, and Millimeter Wave. SPIE, Bellingham, p. 914320
 Roberts J. D. et al., 2024, *MNRAS*, 530, 149
 Ryabchikova T., Piskunov N., Kurucz R. L., Stempels H. C., Heiter U., Pakhomov Y., Barklem P. S., 2015, *Phys. Scr.*, 90, 054005
 Sayeed M. et al., 2024, *ApJ*, 964, 42
 Sharma S., Stello D., Bland-Hawthorn J., Huber D., Bedding T. R., 2016, *ApJ*, 822, 15
 Sheinis A. et al., 2015, *J. Astron. Telesc. Instrum. Syst.*, 1, 035002
 Shetrone M. et al., 2019, *ApJ*, 872, 137
 Singh R., Reddy B. E., Kumar Y. B., 2019, *MNRAS*, 482, 3822
 Smith V. V., Lambert D. L., 1985, *ApJ*, 294, 326
 Sneden C. A., 1973, PhD thesis, University of Texas, Austin
 Sneden C., Lucatello S., Ram R. S., Brooke J. S. A., Bernath P., 2014, *ApJS*, 214, 26
 Stanek K. Z., Zaritsky D., Harris J., 1998, *ApJ*, 500, L141
 Stello D., Sharma S., 2022, *Res. Notes Am. Astron. Soc.*, 6, 168
 Stello D. et al., 2013, *ApJ*, 765, L41
 Sweigart A. V., Mengel J. G., 1979, *ApJ*, 229, 624
 Tayar J., Carlberg J. K., Aguilera-Gómez C., Sayeed M., 2023, *AJ*, 166, 60
 Ting Y.-S., Hawkins K., Rix H.-W., 2018, *ApJ*, 858, L7
 Wehrhahn A., Piskunov N., Ryabchikova T., 2023, *A&A*, 671, A171
 Wheeler A. J., Abruzzo M. W., Casey A. R., Ness M. K., 2023, *AJ*, 165, 68
 Yan H.-L. et al., 2021, *Nat. Astron.*, 5, 86

Yu J., Huber D., Bedding T. R., Stello D., Hon M., Murphy S. J., Khanna S., 2018, *ApJS*, 236, 42
 Zhou Y., Wang C., Yan H., Huang Y., Zhang B., Ting Y.-S., Zhang H., Shi J., 2022, *ApJ*, 931, 136
 Zhou Y., Amarsi A. M., Aguirre Børsen-Koch V., Karlslose K. G., Collet R., Nordlander T., 2023, *A&A*, 677, A98

APPENDIX A: IMPACT ON MEDIAN DIFFERENCE SPECTRUM

The amplitude of the median difference spectrum is small but significant, as shown through comparison to controlled matched sets of RGB–RGB and RC–RC stars (see Section 3). Given the small amplitude, minor changes in the RC–RGB matched set could cause visible changes in the median difference spectrum. In this section, we investigate the impact of RC–RGB matched set definition on the median difference spectrum.

A1 Duplicate RGB stars

There are repeated RGB stars in the RC–RGB matched set. For each RC star, we find a corresponding RGB star; however, the RGB stars are spread out in $\log(g)$ whilst RC stars are concentrated at $\log(g) \approx 2.5$, as shown in Fig. A1. As a result, the RGB matched stars are not unique.

We create a new matched set, where pairs containing a duplicate RGB star are removed, resulting in 394 matched pairs. The median difference spectrum for this no duplicate matched set is shown in Fig. A2. The median difference spectrum is similar between the RC–RGB matched set and the no duplicate matched set, indicating that these duplicate RGB stars do not impact the spectroscopic signal of evolutionary states.

A2 Stellar parameters

There is a correlation between stellar parameter and stellar parameter difference, as shown in Fig. 3. We fit a linear line to the RC–RGB matched set to show the gradient of this trend in Fig. A3. To study the impact of this correlation and distribution in stellar parameter difference on the median difference spectrum, we create a new matched set per stellar parameter, where the stellar parameter

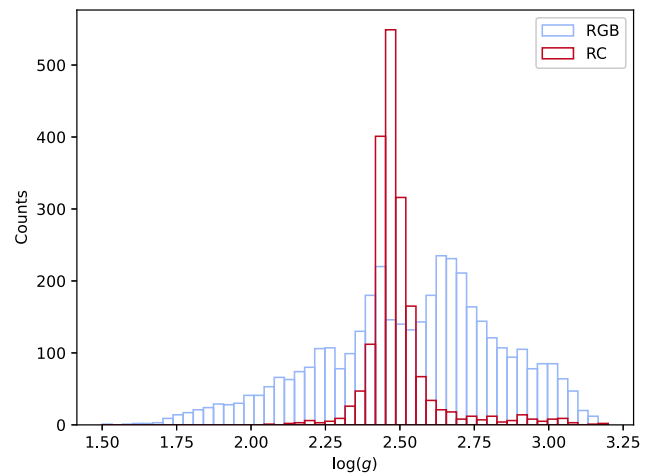


Figure A1. $\log(g)$ vs counts for the crossmatched catalogue, where RGB stars are shown in blue and RC stars are shown in red. The RC stars are more clustered in $\log(g)$ whilst RGB stars are more spread out.

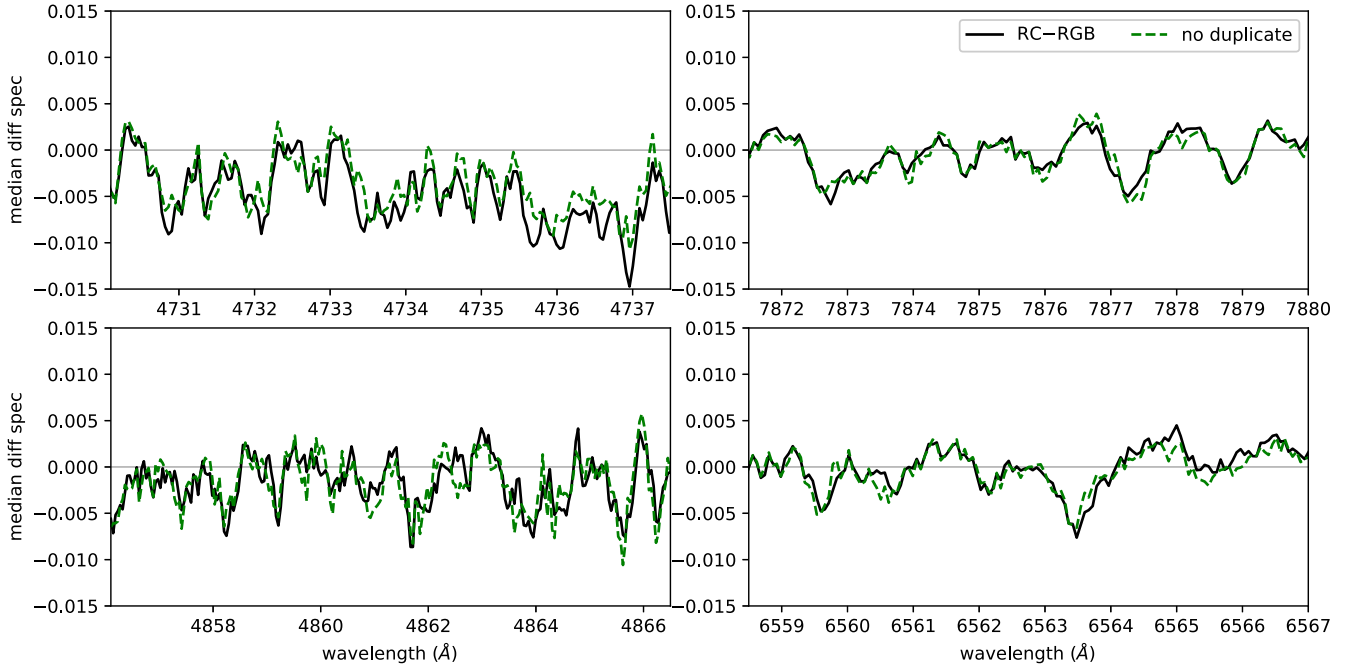


Figure A2. The median difference spectrum for the full RC–RGB matched set with duplicates (black solid) and a no duplicate matched set (green dashed). The difference between the median different spectra is minimal, indicating that duplicate RGB stars do not affect the median difference spectrum.

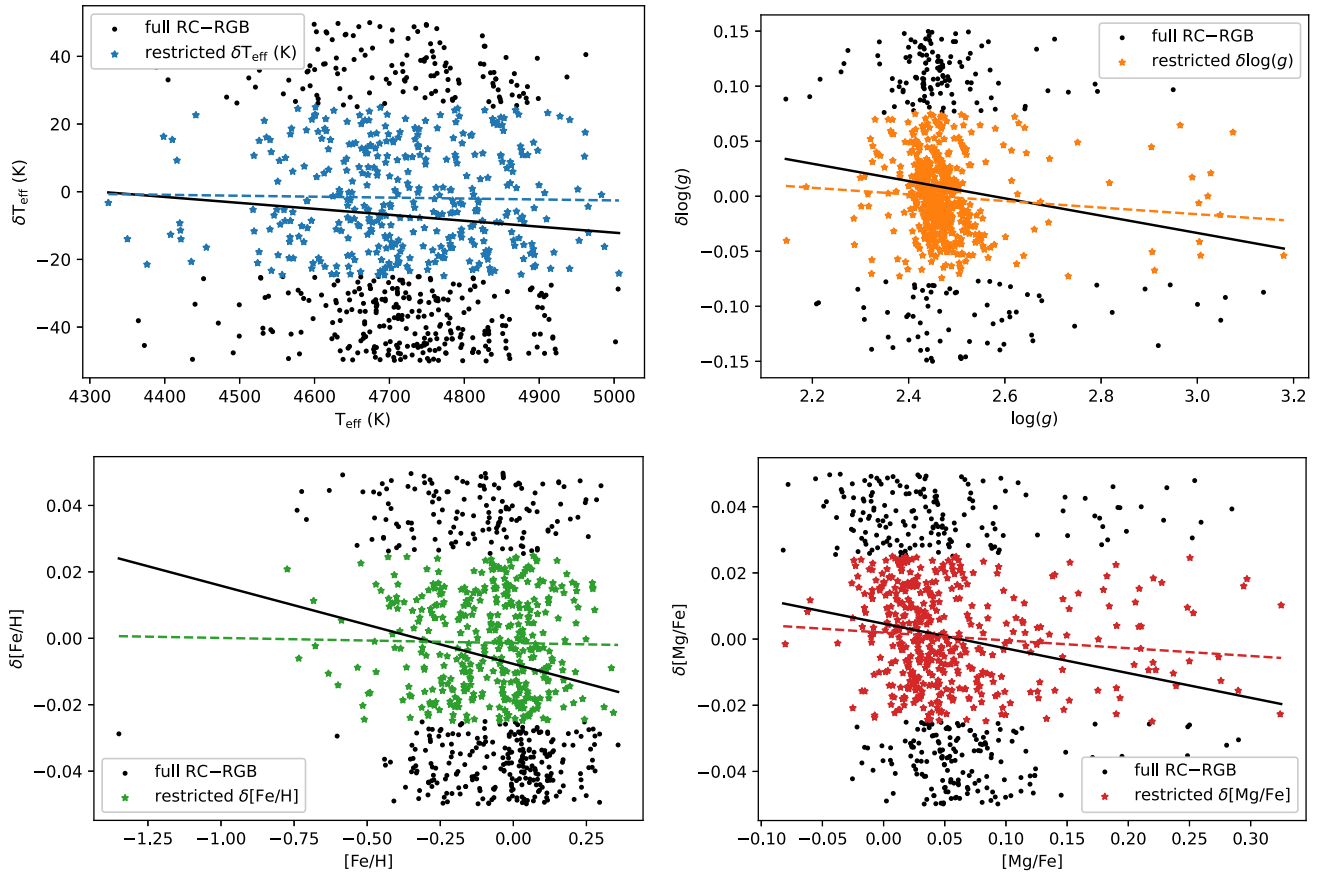


Figure A3. The correlation between RC stellar parameters and the stellar parameter difference, where each panel shows a different stellar parameter. The full RC–RGB matched set is shown in black whilst the restricted matched sets are shown in different colours as well as panels: restricted T_{eff} is shown in the top left (blue stars), restricted $\log(g)$ is shown in the top right (orange stars), restricted $[\text{Fe}/\text{H}]$ is shown in the bottom left (green stars), and restricted $[\text{Mg}/\text{Fe}]$ is shown in the bottom right (red stars). The solid line is a linear fit for each parameter and matched set, showing that the gradient in the restricted sets are reduced compared to the full set.

Table A1. The mean difference in stellar parameters for the full RC–RGB matched set and the restricted matched sets. The mean difference for the restricted matched set is reduced compared to the full RC–RGB matched set for the corresponding stellar parameter, whilst the other mean stellar parameter differences of the non-restricted stellar parameters remains at a similar magnitude, limiting its impact on the median difference spectrum.

Parameter	Units	RC–RGB	δT_{eff}	$\delta \log(g)$	$\delta[\text{Fe}/\text{H}]$	$\delta[\text{Mg}/\text{Fe}]$
T_{eff}	K	7.17	1.79	6.56	4.87	7.25
$\log(g) / 10^{-3}$	$\log(\text{cm s}^{-2})$	-7.5	-8	0.59	-7.1	-7.9
$[\text{Fe}/\text{H}] / 10^{-3}$	dex	5.8	3.5	5.8	1.3	5.5
$[\text{Mg}/\text{Fe}] / 10^{-4}$	dex	-3.5	-1.3	-0.71	-3.3	-6.5

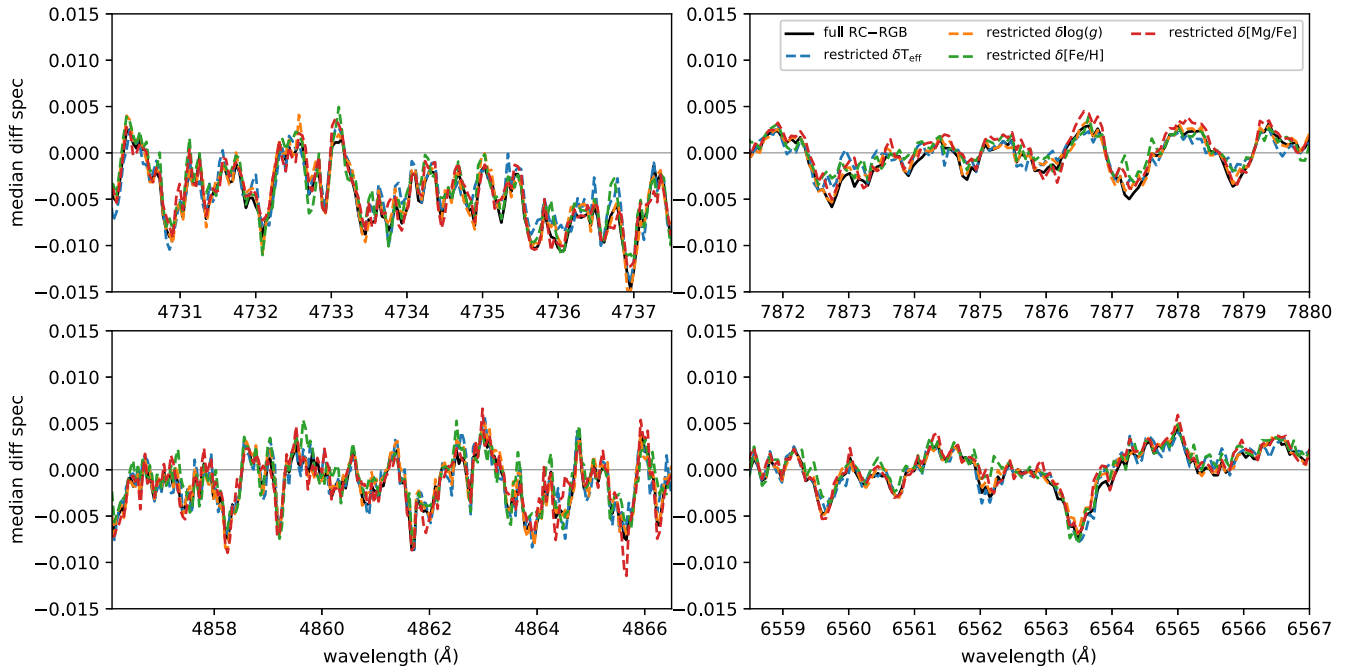


Figure A4. The median difference spectra: RC–RGB (solid black), restricted δT_{eff} (dashed blue), restricted $\delta \log(g)$ (dashed orange), restricted $\delta[\text{Fe}/\text{H}]$ (dashed green), and restricted $\delta[\text{Mg}/\text{Fe}]$ (dashed red). The difference between the full RC–RGB median difference spectrum and the restricted median difference spectrum is minimal for the wavelength regions shown, where we show the same wavelength regions as in Section 4. These are the same restricted matched sets as in Fig. A3.

difference restriction is smaller by a factor of 2 compared to the RC–RGB matched set. The correlation between stellar parameter and stellar parameter difference of these new matched sets are shown in Fig. A3. We fit a linear line to the new matched sets, showing that the gradient is reduced compared to the full RC–RGB matched set. The mean stellar parameter difference for the new matched sets are reported in Table A1, showing that our new matched sets have a reduced mean stellar parameter difference for the parameter that is controlled, whilst the mean difference between the other parameters do not change significantly.

We show the median difference spectrum for each difference restricted set and the original RC–RGB matched set in Fig. A4. Whilst the median difference spectrum is closer to zero at some wavelengths (e.g. the $[\text{Fe}/\text{H}]$ restricted matched set at 7875.9 \AA), the impact of lowering the mean stellar parameter difference and correlation between stellar parameter and stellar parameter difference is minimal.

A3 Misclassifications

Whilst we adopt asteroseismic classifications, they still could contain misclassifications. In order to study the effect of misclassifications on our median difference spectrum, we create a new matched set with some misclassifications injected into the data. For this misclassified matched set, we take 10 percent of our RC–RGB matched pairs and find a new pair, except this time with the same evolutionary state. We set our misclassification percentage to 10 percent based on the random forest accuracy from Appendix C. Fig. A5 shows the median difference spectrum for this misclassified matched set, which is similar to the RC–RGB matched set, indicating that misclassifications do not significantly affect our results.

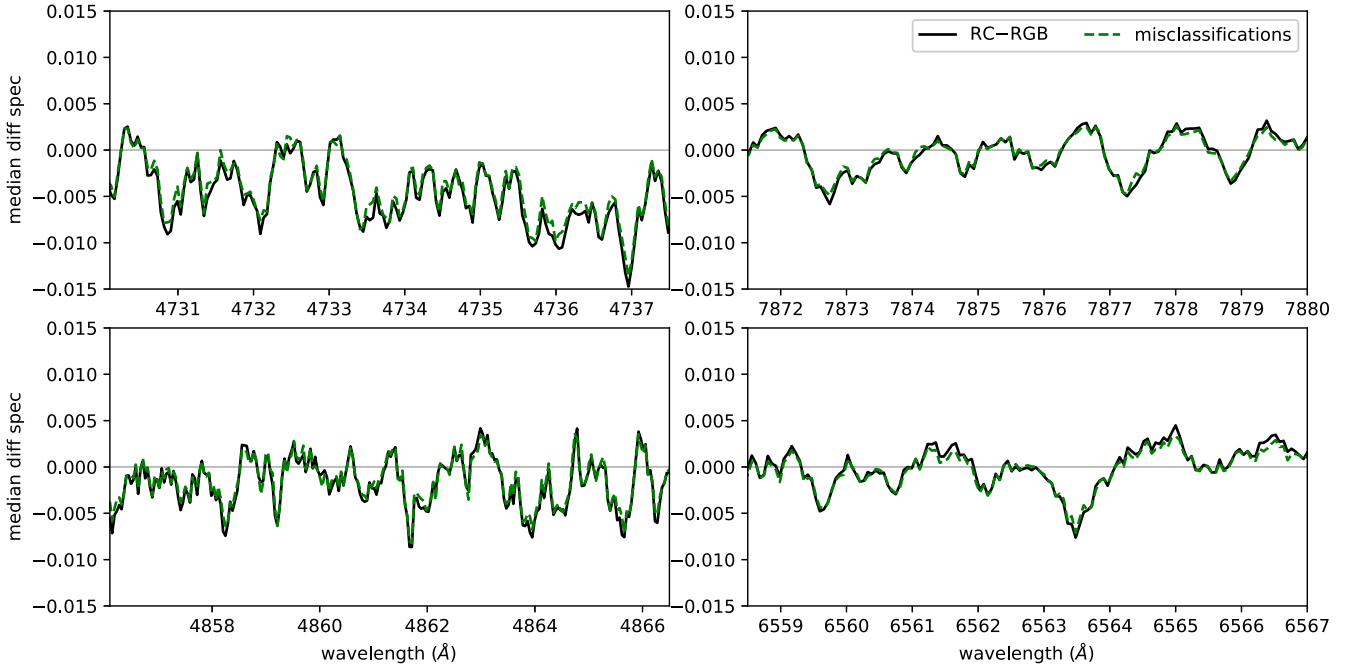


Figure A5. The median difference spectrum for the RC–RGB matched set (black solid) and a set of matched stars with 10 per cent misclassifications (green dashed). The difference between the median different spectra is minimal, indicating that 10 per cent misclassified RGB stars do not affect the median difference spectrum.

APPENDIX B: UNCERTAINTY OF DIFFERENCE SPECTRUM

Each spectrum of GALAH has flux error at every wavelength. We are interested in quantifying the error in the median difference spectrum due to this flux error. In the following, we derive equation (6).

Let f_i be the spectrum for the i th star, then for a total of n stars, $i \in 1, \dots, n$. We stack the spectra by taking the median: $f_m = \text{med}(f_1, \dots, f_n)$. Assuming the spectra are independent, the errors can then be combined in terms of their S/N by summation in quadrature:

$$[S/N](f_m) = \sqrt{\sum_{i=1}^n [S/N]^2(f_i)}, \quad (\text{B1})$$

Note that the flux error for spectrum f_i is the inverse of the S/N,

$$\sigma_i = \frac{1}{[S/N](f_i)}. \quad (\text{B2})$$

Then the error of the median spectrum is given by the combination of equation (B1) and (B2):

$$\sigma_m = \left(\sum_{i=1}^n \frac{1}{\sigma_i^2} \right)^{-\frac{1}{2}}. \quad (\text{B3})$$

In this work, we take the median of the difference spectra, $\delta f_i = f_{\text{RC},i} - f_{\text{RGB},i}$. The error in the difference spectrum is given by the additivity of variance:

$$\sigma_i^2 = \sigma_{\text{RC},i}^2 + \sigma_{\text{RGB},i}^2. \quad (\text{B4})$$

therefore, by combining equation (B3) and equation (B4) we reach equation (6).

APPENDIX C: RANDOM FOREST

In an effort to expand the number of stars with asteroseismic classifications, machine learning classification algorithms are often employed to transfer the evolutionary state label from asteroseismic data sets to spectroscopic data sets (Hawkins et al. 2018; Ting et al. 2018; Casey et al. 2019; Lu et al. 2022). Random forests are often employed for this purpose due to their speed of training and high accuracy. In this appendix section, we apply random forests to our crossmatched catalogue training on different combinations of both stellar parameters and spectra, to identify which features contribute to the evolutionary state classification.

For these tests we, utilize our full crossmatched catalogue which contains a total of 5949 stars, where we adopt a 80–20 train-test cross validation split for evaluation of our random forest performance. We use `RandomForestClassifier` from the `SCIKIT-LEARN` library (Pedregosa et al. 2011), and optimise hyperparameters using `GridSearchCV` over hyperparameter ranges:

- (i) Criterion utilized: { gini, entropy };
- (ii) Max depth of a single tree: { 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15};
- (iii) Number of estimators of the forest: { 10, 100, 500, 1000, 5000, 10000, 20000};
- (iv) Maximum samples utilized in training a single tree: { 0.3, 0.5, 0.7, 0.9}.

We find that when training on stellar parameters it is difficult to improve the accuracy beyond 91 per cent, which is achieved through training on T_{eff} , $\log(g)$, $[\text{Fe}/\text{H}]$; likely due to the low accuracy of additional features such as v_{mic} , $[\text{C}/\text{N}]$ on a star-by-star basis. When learning solely on spectra, the random forest finds evolutionary state information in every wavelength of spectra, indicating that there is evolutionary state information in spectra beyond just the C_2 and CN molecular lines.

Table C1. Different combinations of features used to train the random forest and the accuracy achieved by the trained random forest.

T_{eff}	$\log(g)$	[Fe/H]	$v \sin(i)$	[C/N]	v_{mic}	Accuracy
✓						70 per cent
	✓					81 per cent
✓	✓					85 per cent
✓	✓	✓				91 per cent
✓	✓	✓	✓			91 per cent
✓	✓	✓		✓		91 per cent
✓	✓	✓			✓	92 per cent
✓	✓	✓		✓	✓	92 per cent

C1 Stellar parameters

We train random forests on different combinations of stellar parameters to classify the evolutionary state. The combination of stellar parameters trained on and the accuracy achieved are listed in Table C1. The stellar parameters T_{eff} , $\log(g)$, and [Fe/H] can provide a classification accuracy of 91 per cent for our crossmatched catalogue. Out of these three stellar parameters, $\log(g)$ contains the most information, with classification accuracy reaching 81 per cent when trained on $\log(g)$ alone. It may seem surprising that the accuracy when training on $\log(g)$ alone is relatively high, but this is expected, as RC stars are concentrated at $\log(g) = 2.5$. The inclusion of T_{eff} and [Fe/H] to classification using $\log(g)$ both provide an increase of ~ 5 per cent in accuracy. Classifying on v_{mic} and the stellar parameters provides an increase of 1 per cent to the accuracy, whilst classifying on [C/N] or $v \sin(i)$ offers no improvement to classification accuracy. This result is expected as, although we observe v_{mic} , [C/N], and $v \sin(i)$ differences for the full population of RC and RGB stars, the accuracy and precision for individual stars are not enough to classify evolutionary state accurately on a star-by-star basis. In general, the accuracy improves as more features are

used to train the random forest. However, this improvement reaches a limit around 91 per cent, where it becomes difficult to improve the accuracy further.

We find 78 per cent–94 per cent of incorrectly classified stars are common between random forests trained on different combinations of features. That is, if a star is incorrectly classified through a particular combination of features, it is also likely to be incorrectly classified through a different combination of features. Therefore, these accuracy percentages are highly dependent on the data set. It is possible that these stars have an incorrect evolutionary state label from our asteroseismic catalogue; however, we do not see a correlation between incorrect classification by the random forest and the evolutionary state classification or $\Delta\nu$ vetting.

C2 Stellar spectra

There is evolutionary state information in every wavelength in spectra. We trained two hyperparameter optimized random forests on the Swan bands. One random forest used ~ 350 wavelengths with C_2 molecular features, whilst the other random forest used ~ 350 wavelengths without C_2 molecular features. Both random forests achieved an accuracy of ~ 84 per cent, indicating that there is evolutionary state information outside of the C_2 molecular features. In Sections 4 and 5, we find that RC stars have broader H_α and H_β than RGB stars, linked to the v_{mic} parameter. Line-width information is present at every wavelength, therefore, it is possible that the random forest is classifying based on differences in line-width present in every line in the spectrum. Despite these results, identifying the exact features that the random forest uses to make its predictions requires further careful study due to the interpretability issues associated with large random forest models (Haddouchi & Berrado 2019).

This paper has been typeset from a $\text{\TeX}/\text{\LaTeX}$ file prepared by the author.