Investigating the Cosmological Lithium Problem: The Spite Plateau in Warm Metal-Poor Dwarf Stars with Gaia DR3, GALAH DR4, and Bayesian Regression

Russell Huang[1]

Received March 22, 2026
Accepted July 15, 2026
Electronic access August 15, 2026

The cosmological lithium problem is a major discrepancy between Big Bang Nucleosynthesis predictions and stellar observations. This study addresses this problem by measuring lithium abundances in warm, metal-poor dwarf stars to investigate the Spite Plateau. In this paper, we measure the Spite Plateau in warm, metal-poor dwarf stars using GALAH DR4 and Bayesian regression. After applying quality cuts to the data to isolate warm, metal-poor dwarf stars, we analyze lithium abundance compared to metallicity using the Bayesian regression code ROXY, which also accounts for uncertainties in Spite Plateau data. This allows us to infer the slope, intercept, and scatter of Spite Plateau stars across varying metallicity cuts. Across all metallicity cuts between [Fe/H] < -1.3 and [Fe/H] < -1.7, the slopes remain consistent with zero within their uncertainties, suggesting that the data are consistent with a flat Spite Plateau. These findings suggest that even with modern surveys and Bayesian modeling, the oldest stars exhibit nearly constant lithium abundances that remain significantly lower than the values predicted by Big Bang Nucleosynthesis. Specifically, the inferred lithium abundances remain approximately 0.6 dex lower than the primordial abundance predicted by standard BBN models, which continues to be an open problem in cosmology and stellar astrophysics. The study provides new constraints on the Spite Plateau using modern surveys and advanced Bayesian methods, contributing to our understanding of early-universe nucleosynthesis.

Keywords: Lithium abundance, Spite Plateau, Big Bang Nucleosynthesis, Gaia DR3, GALAH DR4, Bayesian regression

Introduction

Big Bang Nucleosynthesis (BBN) occurred during the first few minutes after the Big Bang, when the universe was hot and dense enough for nuclear reactions to form the lightest elements, including hydrogen, helium, deuterium, and lithium[1,2]. During this period, the temperature of the early universe decreased as the universe expanded. A useful approximation for the temperature evolution is

$$T \sim 1\,\mathrm{MeV}\left(\frac{1\,\mathrm{s}}{t}\right)^{1/2} \quad (1)$$

This relation shows why BBN was limited to a short time window. At early times, the universe was too hot for nuclei to remain stable because photons were able to break them apart[3]. Later, the universe became too cool for nuclear reactions to continue efficiently. Therefore, lithium production depends more on the expansion rate, temperature, baryon density, and the nuclear reaction network active during this short period.

The expansion rate of the early universe is described by the Friedmann equation[1],

$$H^2 = \frac{8\pi G}{3}\rho \quad (2)$$

where H is the Hubble expansion rate, G is the gravitational constant, and ρ is the total energy density of the universe. During this era, when there were large amounts of radiation, relativistic particles such as photons and neutrinos dominated the energy density, causing

$$\rho_{\mathrm{rad}} \propto a^{-4} \quad (3)$$

where a is the scale factor of the universe[3]. This matters for lithium because the expansion rate controls how long nuclear reactions occur before the universe becomes too cool. If the universe expands too quickly, reactions can stop earlier, and if it expands more slowly, the reactions can continue longer. Therefore, the expansion history helps determine the predicted primordial abundance of lithium.

The baryon density is another key value in BBN calculations. Modern measurements of the cosmic microwave background, especially from the Planck mission, provide precise

[1] *Mountain View High School*

constraints on the baryon density of the universe[4]. When this baryon density is included into standard BBN models, the predicted primordial lithium abundance is approximately[1],

$$A(\mathrm{Li}) = \log_{10}\left(\frac{N_{\mathrm{Li}}}{N_{\mathrm{H}}}\right) + 12 \approx 2.7 \tag{4}$$

However, observations of old, warm, metal-poor dwarf stars typically find lithium abundances closer to[5],[6].

$$A(\mathrm{Li}) \approx 2.1 \tag{5}$$

This difference of roughly 0.5–0.6 dex, or a factor of about 2 - 3, is known as the cosmological lithium problem. The problem is especially important because deuterium and helium generally agree well with standard BBN predictions, while lithium remains significantly lower than expected[7]. Another side of this problem is the observational side, which is connected to the Spite Plateau. The Spite Plateau can be defined as the nearly constant lithium abundance observed in warm, metal-poor dwarf stars across a range of low metallicities[5,6]. These stars are useful because they are among the oldest stellar populations and formed before significant chemical enrichment of the Milky Way occurred[8]. Warm dwarf stars are also less affected by lithium depletion than cooler stars or evolved giants, making them important for tracing the amount of early lithium abundance[6,9].

One possible explanation for the lithium discrepancy is stellar depletion. In this scenario, the stars originally formed with higher lithium abundances, but processes such as atomic diffusion, gravitational settling, and turbulent mixing gradually reduced the lithium observed at the stellar surface[10,11]. Models that include diffusion and turbulent mixing typically predict lithium depletion of approximately 0.15 - 0.30 dex[12]. A depletion correction of 0.15 - 0.30 dex would raise an observed plateau value of A(Li) ≈ 2.1 dex to approximately 2.25–2.40 dex, reducing but not eliminating the gap with the BBN prediction of A(Li) ≈ 2.7 dex.

In this paper, we test whether the observed lithium abundances of warm, metal-poor dwarf stars in GALAH DR4 are consistent with a flat Spite Plateau. Gaia DR3 is used for photometric and astrometric validation of the stellar sample[13,14], while GALAH DR4 provides the values of lithium abundances, metallicities, effective temperatures, and surface gravities used in the analysis[15,16]. The objective of this study is to measure the lithium-metallicity relationship in warm, metal-poor dwarf stars and quantify the Spite Plateau using Bayesian regression.

This study is limited by the number of extremely metal-poor stars available in GALAH DR4 and by the decreasing sample size below [Fe/H] < –1.8, which prevents a more reliable statistical analysis at lower metallicities. By measuring the slope, intercept, and intrinsic

scatter of the plateau, we compared the observed lithium abundance with previous Spite Plateau measurements and with the value predicted by standard Big Bang Nucleosynthesis. This comparison provides additional observational constraints on the cosmological lithium problem and helps evaluate whether stellar depletion processes can plausibly account for the discrepancy.

Methods

To analyze the relationship between lithium abundance and metallicity in dwarf stars, we used data from both Gaia DR3 and GALAH DR4 surveys. We created Hertzsprung-Russell and Kiel diagrams to verify our sample and the evolutionary state of the selected stars, and an A(Li) versus [Fe/H] plot to examine the lithium-metallicity relationship. We first verified that the selected stars were warm dwarf stars that satisfy the Spite Plateau. Using data from Gaia DR3, the photometry and parallaxes were used to construct a Hertzsprung–Russell diagram, which allowed us to identify stars in the main sequence and exclude giant stars. Because giant stars can experience more lithium depletion[9], excluding them from the main sample helps us ensure that the observed lithium abundances are less affected by stellar depletion which makes the main sample more useful to comparing values with primordial lithium predictions.

We also created a Kiel diagram, which plots surface gravity, $\log g$, against effective temperature, T_{eff} . In this study, the Kiel diagram is used to verify that our cuts selected warm, metal-poor dwarf stars appropriate for a Spite Plateau analysis. Since dwarf stars have higher surface gravities than giants[17], the Kiel diagram helps confirm that the selected stars satisfy the dwarf-star criteria used in the analysis. Together, the HR and Kiel diagrams verify that the final sample of stars are classified in the warm dwarf region, which is important because warm dwarf stars are believed to preserve lithium abundances closer to their initial values.

Data

This study used Gaia DR3 and the GALAH DR4 catalog because the two surveys pro-vide complementary information. Gaia DR3 provides astrometric and photometric measurements used to construct Hertzsprung–Russell and Kiel diagrams for sample validation. GALAH DR4 provides the spectroscopic quantities required for the lithium analysis, including lithium abundance, metallicity, effective temperature, surface gravity, and measurement uncertainties. Using both surveys allows the sample to be verified while also providing the spectroscopic quantities necessary for the analysis of the Spite Plateau stars with Bayesian regression.

To construct the Hertzsprung–Russell diagram, we downloaded measurements from Gaia DR3 including apparent magnitude and BP–RP color. Surface gravity and effective temperature used in the Kiel diagram were taken from the GALAH DR4 survey. From this, we were able to create graphs of Absolute Magnitude against Color Index, $\log g$ against $\log T_{eff}$, and $\log T_{eff}$ against Color Index. Using the apparent magnitudes in Gaia's measurements, the absolute magnitude for each star was calculated using the equation [18],

$$M_G = m_G + 5 - 5\log_{10}\left(\frac{1000}{\varpi}\right) \tag{6}$$

where ϖ is measured in milliarcseconds. The uncertainty in absolute magnitude was estimated through standard error propagation. Assuming the uncertainty in apparent magnitude is negligible compared to the parallax uncertainty, the propagated uncertainty in absolute magnitude is

$$\sigma_{M_G} = \frac{5}{\ln(10)}\frac{\sigma_\varpi}{\varpi} \tag{7}$$

where ϖ is the Gaia parallax and σ_ϖ is the parallax uncertainty. To improve the reliability of the Gaia sample, we required positive parallaxes and RUWE$<$1.4, which reduces the inclusion of stars with unreliable data. The parallax uncertainty and an additional parallax precision threshold were not imposed because Gaia DR3 was used primarily for sample validation and visualization rather than as a direct input to the Bayesian lithium regression, so this data would not affect the Spite Plateau slopes. To confirm that our sample consists primarily of main-sequence dwarfs, we first created the Hertzsprung–Russell diagram. The HR diagram graphs intrinsic luminosity as a function of color index and allows us to distinguish dwarfs from giants [19]. Figure 1 shows the HR diagram produced from the Gaia data.

To further confirm that we have only selected dwarf stars, we examine the Kiel diagram (Figure 2), which plots surface gravity against temperature. Dwarfs have a higher surface gravity, which confirms our data cuts [17].

The primary results are derived from GALAH DR4 measurements after additional cuts were made. The GALAH DR4 survey helped provide us with data on metallicity, effective temperature, surface gravity, and lithium abundance [15,16]. To keep our stars reliable, we made specific quality cuts such as, flag_sp = 0, flag_fe_h = 0, flag_a_li = 0, snr_px_ccd3 $>$ 30, $\log g$ $>$ 3.65, [Fe/H] $<$ –0.8, and T_{eff} $>$ 5850 K. These filters ensure that only high-quality dwarf stars are included in our data, so our predictions can be more precise.

The GALAH DR4 columns used in this analysis were a_li for lithium abundance, fe_h for metallicity, T_{eff} for effective temperature, $\log g$ for surface gravity, e_fe_h for metallicity uncertainty, and e_a_li_low and e_a_li_upp for lithium abundance uncertainties. Additional quality-control columns included flag_sp, flag_fe_h, flag_a_li, and snr_px_ccd3 [15,16].

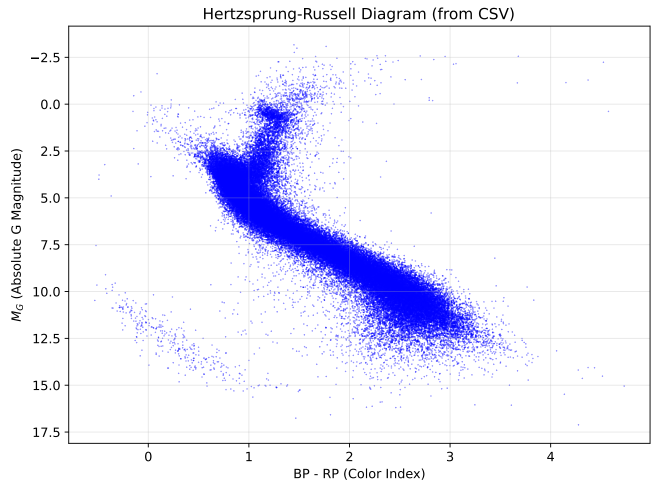


Figure. 1 Hertzsprung–Russell diagram created from the Gaia DR3 sample. The x-axis shows the BP–RP color index, while the y-axis shows absolute magnitude M_G. The main sequence is visible and shows that giant stars were removed.

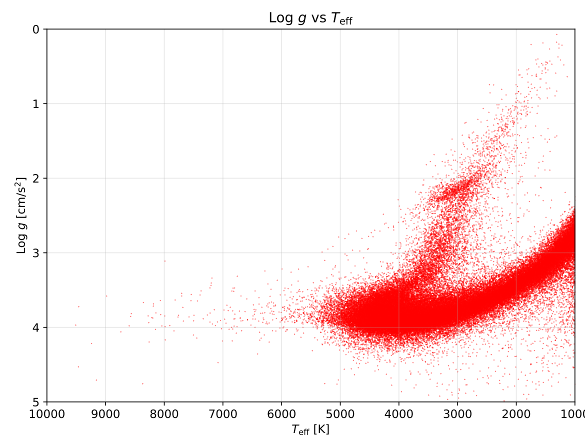


Figure. 2 Kiel diagram ($\log g$ vs T_{eff}) of the Gaia DR3 sample. The high surface gravities show that the stars align with dwarf properties.

Table 1 summarizes the amount of stars remaining after we imposed each filtering cut used to construct the final warm, metal-poor dwarf sample from GALAH DR4. Each successive cut removes stars that may add more uncertainties to the analysis. In particular, the temperature, surface gravity, and metallicity cuts isolate warm, metal-poor dwarf stars that are expected to preserve lithium abundances closer to their primordial values.

The filtered sample was then used to examine the relationship between A(Li) and [Fe/H]. This can provide additional insight into stellar evolution and the cosmological lithium prob-

Table 1 Filtering pipeline used to construct the final warm, metal-poor dwarf sample from GALAH DR4

Filter Applied	Stars Remaining
Initial GALAH DR4 sample	773,170
parallax > 0	773,170
flag_sp = 0	567,913
flag_fe_h = 0	472,579
flag_a_li = 0	236,118
snr_px_ccd3 > 30	235,810
$M_G > 1.5$	180,320
$\log g > 3.65$	153,221
$T_{\mathrm{eff}} > 5850$ K	81,789
[Fe/H] < -0.8	157

lem.

Identifying Spite Plateau stars

To isolate our sample Spite Plateau stars, we first filtered through the data for dwarf stars that met the requirements to be classified as a warm, metal-poor dwarf. These stars were selected using $\log g > 3.65$, $M_G > 1.5$, $T_{eff} > 5850$ K, and [Fe/H] < –0.8. The stars in this category are usually warm and metal-poor, which helps us analyze our data more accurately. Warm stars tend to experience less lithium depletion, meaning their current lithium abundance is closer to the primordial value than other stars[20,21]. The initial sample was first restricted to warm, metal-poor dwarf stars using [Fe/H] < –0.8. This cut defines the master sample used throughout the study. Bayesian regression was then performed on progressively lower metallicity cuts at upper limits [Fe/H] < –1.3, –1.4, –1.5, –1.6, and –1.7. These cuts were chosen to isolate increasingly metal-poor stars while maintaining sufficient sample sizes for plausible data analysis. Comparing the regression results across multiple metallicity thresholds also allows us to evaluate the reliability of the inferred slope and identify the metallicity range where GALAH DR4 can reliably constrain the Spite Plateau.

After applying all quality cuts, the final sample contained 157 warm, metal-poor dwarf stars, and the number of stars decreased rapidly at lower metallicities. Because the sample size became very small below [Fe/H] < –1.8, the primary Bayesian regression analysis was restricted to metallicity cuts between [Fe/H] < –1.3 and [Fe/H] < –1.7.

Table 2 summarizes the data at each metallicity cut where N is the total number of Spite Plateau stars that meet all the cuts, median T_{eff} (K) being the median effective temperature of all stars in the sample, median $\log g$ being the median surface gravity of the sample, median [Fe/H] being the median metallicity of the sample, and fraction excluded (%) being the percent of stars removed by the GALAH lithium quality requirement flag_a_li $\neq 0$. As the metallicity threshold decreases, the sample size rapidly declines, which contributes to larger uncertainties in the inferred regression parameters.

Table 2 Summary of the metallicity subsamples used in the Bayesian regression analysis. The median values of T_{eff} (K), $\log g$, and [Fe/H] are reported

[Fe/H] cut	N	T_{eff} (K)	$\log g$	[Fe/H]	Fraction Excluded (%)
-1.3	63	5962	4.149	-1.492	12.5
-1.4	46	5987	4.185	-1.559	13.2
-1.5	28	6004	4.153	-1.718	15.2
-1.6	19	5995	4.149	-1.790	9.5
-1.7	16	6006	4.216	-1.799	5.9

To visualize the distribution of warm dwarf stars in color–temperature space, we examine Figure 3.

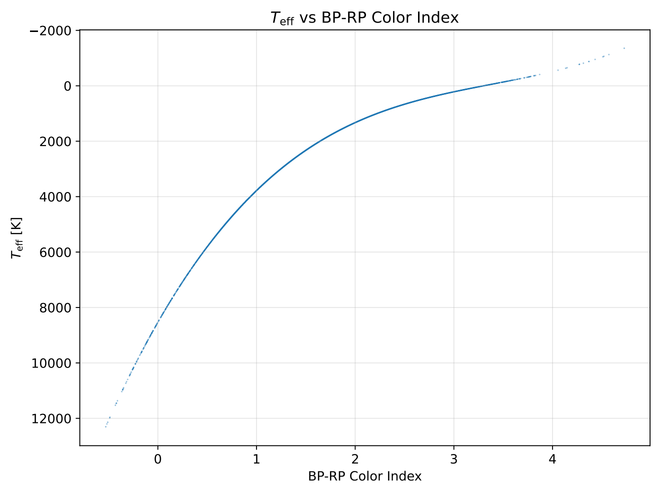


Figure. 3 BP–RP color index versus T_{eff} for the filtered dwarf sample. This figure is used to verify that the selected stars occupy the expected warm-star color and temperature range.

Once dwarf stars were identified, we plotted lithium abundance as a function of metallicity to inspect the Spite Plateau structure.

Fitting the Spite Plateau

After selecting the warm, metal-poor dwarf sample, we fit the relationship between lithium abundance and metallicity using ROXY. ROXY is a Python package for fitting a function to data that contains both x and y errors[22]. ROXY implements Marginalized Normal Regression, which uses Bayesian inference to fit a function to data with errors. This is important because uncertainties in the data that aren't accounted for can lead to incorrect data and slope. The code also uses automatic

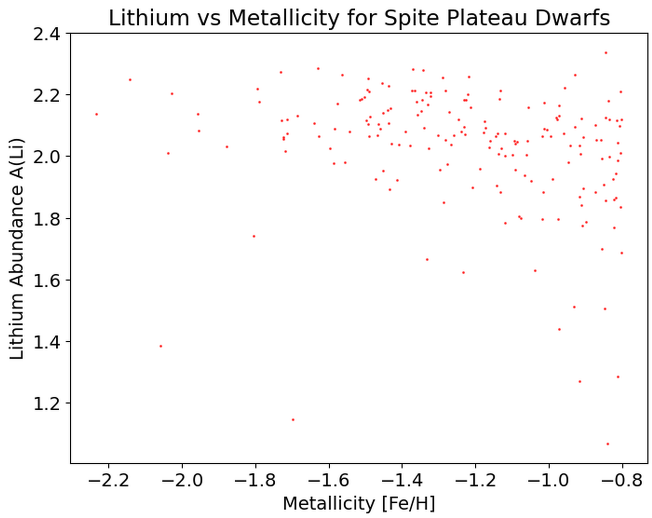


Figure. 4 Lithium abundance, A(Li), as a function of stellar metallicity, [Fe/H], for warm, metal-poor dwarf stars selected from GALAH DR4. Error bars correspond to the reported survey uncertainties.

differentiation enabled by JAX to sample the likelihood using Hamiltonian Monte Carlo to compute the derivatives required. ROXY's regression engine treats uncertain x values as latent variables. These latent x values follow the Gaussian Mixture Model, which marginalizes the latent x values to calculate the likelihood of the observed data, which helps avoid bias. ROXY accounts for measurement uncertainties in both lithium abundance and metallicity when estimating the regression parameters[22]. With the Spite Plateau stars, we then calculated the slope, intercept, and scatter values with the equation y = Ax + B, where A is the slope and B is the intercept. We used these values to create a data plot for each metallicity cut for Spite Plateau stars, and these uniform priors

were adopted for the slope ($-5 < m < 5$), intercept ($0 < b < 5$), and intrinsic scatter ($0 < \sigma < 1$). The model assumed a linear relation A(Li) = m[Fe/H] + b. For all ROXY figures, the x-axis represents stellar metallicity, [Fe/H], measured in dex relative to the Sun,

while the y-axis represents lithium abundance, A(Li), also measured in dex, and the ROXY fits were then performed on the metallicity values from GALAH. Error bars correspond to the reported GALAH DR4 uncertainties in metallicity and lithium abundance. Specifically, metallicity uncertainties were obtained from e fe h. Lithium uncertainties were computed as the average of e a li low and e a li upp. Each star retained its own individual uncertainty during the fitting process.

As a comparison model, we also performed a weighted least-squares (WLS) linear regression for each metallicity cut using the reported lithium abundance uncertainties as weights. The WLS fits produced slopes that were consistent with the ROXY posterior median slopes within their uncertainties (Table 3). This comparison suggests that the conclusion of a nearly flat Spite Plateau is not strongly dependent on the choice of regression method.

Table 3 Comparison of ROXY Bayesian regression and weighted least-squares (WLS) regression slopes. The consistency between the two methods indicates that the conclusion of a nearly flat Spite Plateau is not sensitive to the chosen regression technique

[Fe/H] cut	ROXY slope	WLS slope	ROXY intercept	WLS intercept
< -1.3	$0.024^{+0.059}_{-0.060}$	0.001 ± 0.048	$2.165^{+0.095}_{-0.093}$	2.146 ± 0.076
< -1.4	$0.004^{+0.067}_{-0.071}$	-0.042 ± 0.060	$2.130^{+0.110}_{-0.116}$	2.066 ± 0.100
< -1.5	$0.012^{+0.110}_{-0.108}$	-0.015 ± 0.087	$2.141^{+0.194}_{-0.191}$	2.116 ± 0.157
< -1.6	$-0.036^{+0.167}_{-0.169}$	0.001 ± 0.126	$2.050^{+0.305}_{-0.309}$	2.148 ± 0.236
< -1.7	$-0.115^{+0.193}_{-0.197}$	-0.009 ± 0.150	$1.897^{+0.358}_{-0.369}$	2.128 ± 0.285

The Markov Chain Monte Carlo analysis used two chains with 1000 warm-up iterations and 3000 sampling iterations. We evaluated convergence using the Gelman-Rubin statistic ($\hat{R}$), and all parameters measured values that were close to 1.0, indicating satisfactory convergence, and no divergent transitions were observed during sampling.

Figures 5a–5e show the ROXY fits for various metallicity cuts.

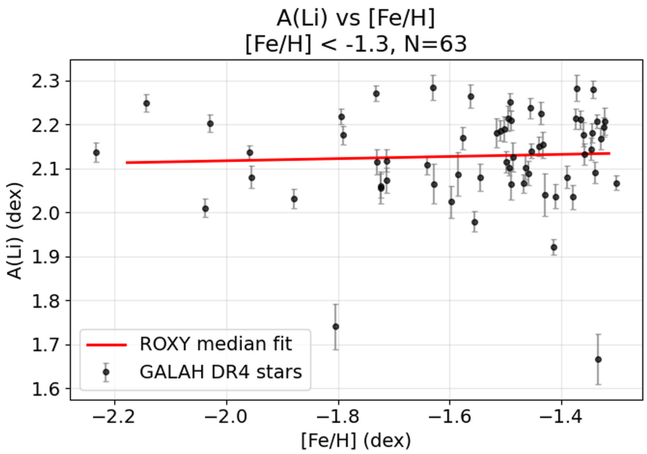
Results

In this study, the slopes were measured to be consistent with zero within uncertainties for all plots as all posterior intervals for the slope included zero. Because the slope is a continuous parameter, the probability that the slope is exactly zero is not meaningful in a Bayesian framework and therefore was not reported. Instead, we report the posterior probability that the slope is negative, $P(m < 0)$, which quantifies the evidence for a decreasing lithium abundance with metallicity. The inferred intrinsic scatter was also shown to range from approximately 0.10 to 0.12 dex.

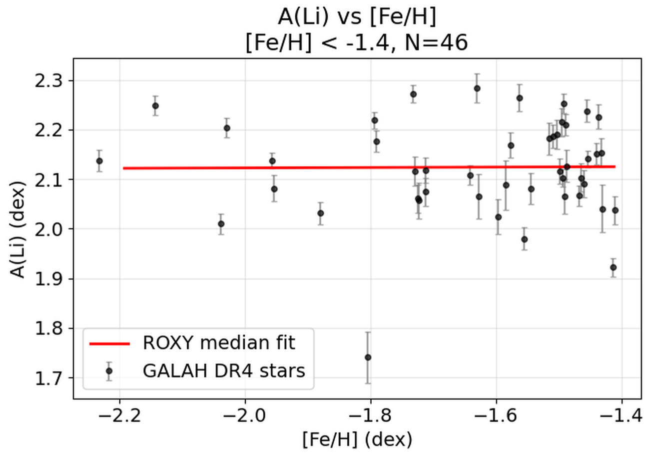
After applying all the necessary cuts, we obtained a filtered sample of metal-poor stars that fit the criteria of Spite Plateau stars. The final sample occupies the expected warm dwarf region in both the HR and Kiel diagrams, supporting the reliability of the stellar selection.

These plots of the isolated GALAH DR4 stars were then compared to the original Gaia DR3 HR and Kiel diagrams. We noted that the stars in the GALAH plots were consistent with our expected regions in the Gaia plots of warm, metal poor dwarfs which suggest that our cuts had isolated Spite Plateau dwarf stars in the warm, metal-poor category.

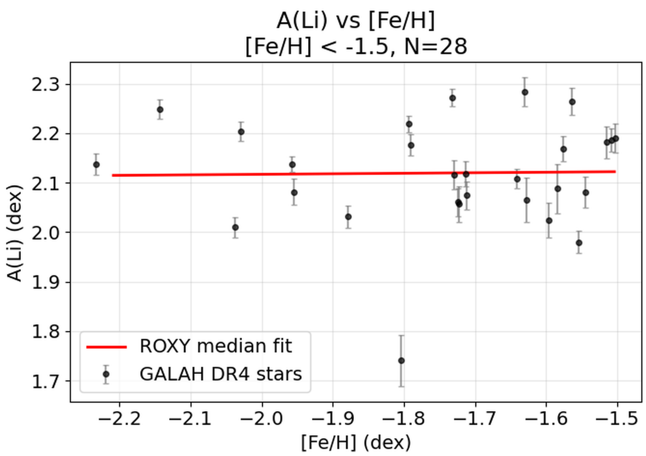
The weighted least-squares fits produced slopes that were consistent with the ROXY posterior median slopes for all metallicity cuts (Table 3). Both methods indicate that the



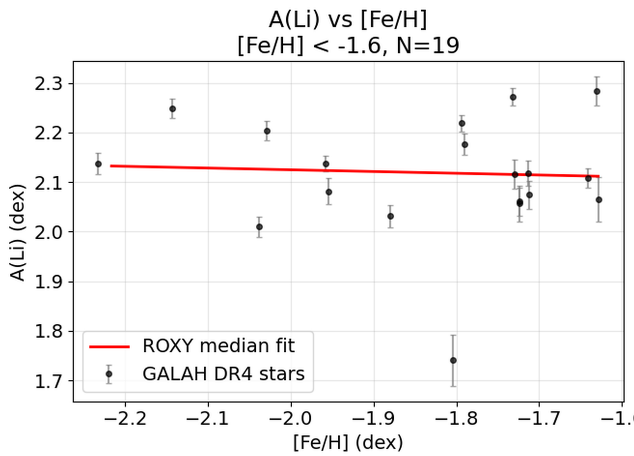
(a) ROXY Bayesian linear regression for stars with [Fe/H] < –1.3. The red line shows the median posterior regression line.



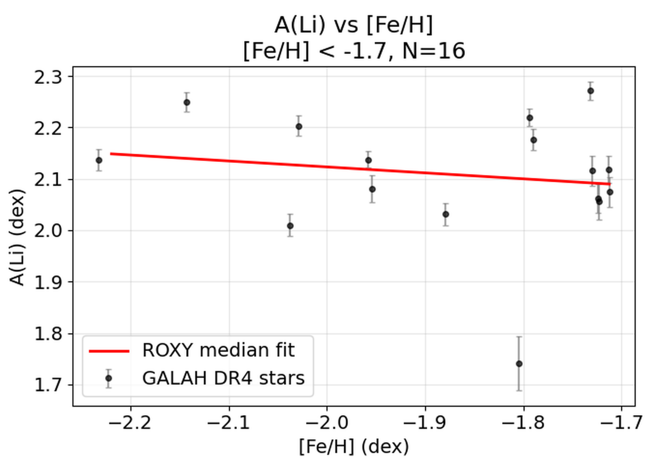
(b) Same as Figure 5a but for stars with [Fe/H] < –1.4.



(c) ROXY fit for [Fe/H] < –1.5.



(d) ROXY Bayesian fit for [Fe/H] < –1.6



(e) ROXY fit for [Fe/H] < -1.7

Figure. 5

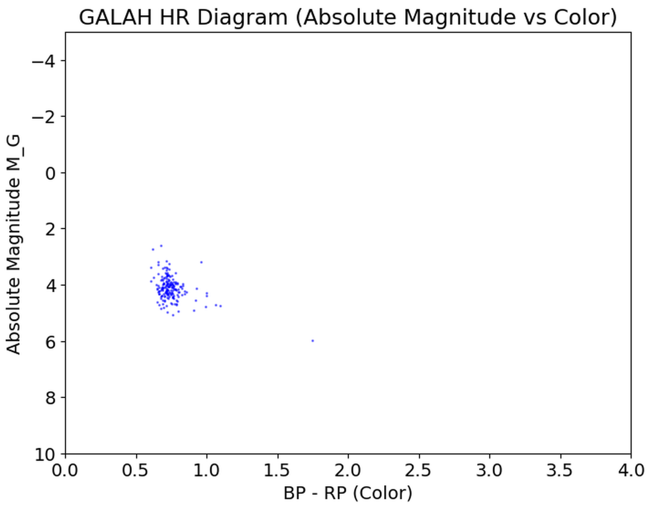


Figure. 6 Hertzsprung–Russell diagram of the final warm, metal-poor dwarf sample selected from GALAH DR4. Unlike Figure 1, this figure shows only the stars used in the Spite Plateau analysis. The tight concentration along the main sequence confirms that the adopted quality cuts successfully isolated dwarf stars.

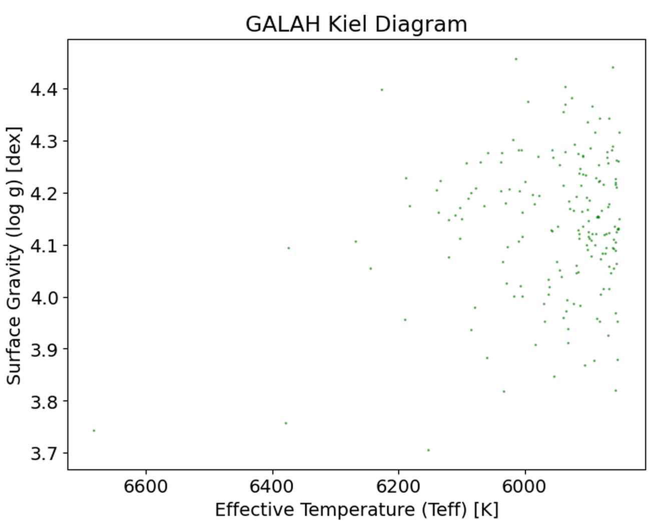


Figure. 7 Kiel diagram of the final warm, metal-poor dwarf sample used in this study. The stars occupy a narrow range of high surface gravities and temperatures, consistent with warm dwarf stars that are expected to preserve near-primordial lithium abundances.

inferred lithium-metallicity relationship is approximately flat over the metallicity range sampled by GALAH DR4.

The posterior intervals for all metallicity cuts include zero, indicating that the data do not strongly favor either a positive or negative trend. This suggests that the lithium abundance in warm, metal-poor dwarf stars does not show a significant dependence on metallicity which

Table 4 Bayesian regression results for each metallicity cut. $P(m < 0)$ represents the posterior probability that the slope is negative

[Fe/H] cut	N	Slope m	A(Li) Intercept b	Scatter σ_{int}	$P(m<0)$
< -1.3	63	$0.024^{+0.059}_{-0.060}$	$2.165^{+0.095}_{-0.093}$	$0.102^{+0.010}_{-0.010}$	0.344
< -1.4	46	$0.004^{+0.067}_{-0.071}$	$2.130^{+0.110}_{-0.116}$	$0.096^{+0.013}_{-0.010}$	0.474
< -1.5	28	$0.012^{+0.110}_{-0.108}$	$2.141^{+0.194}_{-0.191}$	$0.107^{+0.019}_{-0.015}$	0.454
< -1.6	19	$-0.036^{+0.167}_{-0.169}$	$2.050^{+0.305}_{-0.309}$	$0.119^{+0.026}_{-0.020}$	0.583
< -1.7	16	$-0.115^{+0.193}_{-0.197}$	$1.897^{+0.358}_{-0.369}$	$0.122^{+0.033}_{-0.024}$	0.721

is consistent with the Spite Plateau. Although the posterior probability of a negative slope increases at lower metallicities, the decreasing sample size can increase the uncertainties.

Discussion

In this study, we used GALAH DR4 data to perform a Bayesian regression analysis of lithium abundance as a function of metallicity in warm, metal-poor dwarf stars. By applying quality cuts to isolate warm, metal-poor dwarfs and analyzing their lithium abundances through the Bayesian regression tool ROXY, we inferred the slope, intercept, and scatter of A(Li) as a function of metallicity for several metallicity cuts. The results show that the slope remains consistent with zero for all cuts between [Fe/H] $<$ –1.3 and [Fe/H] $<$ –1.7, indicating that the lithium abundance does not show a significant dependence on metallicity. These results are consistent with a flat Spite Plateau over the metallicity range sampled by GALAH DR4. A major limitation of this study is the small number of extremely metal-poor stars available in GALAH DR4 after applying quality and stellar parameter cuts. Although these cuts improve the reliability of the sample, they reduce the number of reliable stars below [Fe/H] $<$ –1.8 to fewer than ten. Therefore, this study is more of a measurement of the Spite Plateau over the metallicity range sampled by GALAH DR4 rather than a measurement of dwarfs at ultra-low metallicities. Future work incorporating surveys that specifically target extremely metal-poor stars would allow the plateau to be investigated at lower metallicities, as Ryan et al. and Aoki et al. were able to measure abundance values at a much lower metallicity [23,24].

The lithium abundance values also aligned with measurements in previous literature. François and Monique Spite calculated an abundance value of $\approx$ 2.1 dex [5], later in 2006, Bonifacio et al. calculated a similar abundance value of $\approx$ 2.1 dex [9], and in 2010, Sbordone et al. calculated abundance values from multiple stars which ranged from $\approx$ 1.8 - 2.2 dex [6]. The results table shows a calculated lithium value of $\approx$ 1.90 - 2.17 dex, which is consistent with the results from previous studies. When compared with Big Bang Nucleosynthesis predictions, our measured lithium abundances remain significantly lower than expected. Standard BBN models constrained by Planck 2018 cosmological parameters predict a primordial abundance of approximately A(Li) $\approx$ 2.72 dex[25], whereas our inferred plateau values range from approximately 1.90 to 2.17. This corresponds to a discrepancy of roughly 0.55–0.82 dex, which reinforces the cosmological lithium problem[26] even when using modern surveys and Bayesian regression. The similarity between our results and other studies may suggest that the discrepancy is unlikely to be a result of a particular survey but instead a persistent feature in metal-poor stellar populations. One possible explanation is stellar depletion through processes such as atomic diffusion and turbulent mixing, although these mechanisms have not yet fully resolved the discrepancy[27].

A significant outcome of this analysis is the relatively small intrinsic scatter ($\sigma_{int} \approx 0.10$ dex) inferred by ROXY. This is consistent with previous observations that warm, metal-poor dwarf stars exhibit relatively uniform lithium abundances across a range of metallicities[28].

Overall, this work shows that even with high-quality data from modern surveys and advanced Bayesian modeling, the Spite Plateau continues to be a significant challenge. Its low lithium abundance is difficult to align with the predictions of Big Bang Nucleosynthesis, continuing to make the lithium problem a significant problem in cosmology and stellar astrophysics.

Acknowledgments

This work made use of the Fourth Data Release of the GALAH Survey (Buder et al. 2021). The GALAH Survey is based on data acquired through the Australian Astronomical Observatory, under programs: A/2013B/13 (The GALAH pilot survey); A/2014A/25, A/2015A/19, A2017A/18 (The GALAH survey phase 1); A2018A/18 (Open clusters with HERMES); A2019A/1 (Hierarchical star formation in Ori OB1); A2019A/15, A/2020B/23, R/2022B/5, R/2023A/4, R2023B/5 (The GALAH survey phase 2); A/2015B/19, A/2016A/22, A/2016B/10, A/2017B/16, A/2018B/15 (The HERMES-TESS program); A/2015A/3, A/2015B/1, A/2015B/19, A/2016A/22, A/2016B/12, A/2017A/14, A/2020B/14 (The HERMES K2-follow-up pro-gram); R/2022B/02 and A/2023A/09 (Combining asteroseismology and spectroscopy in K2); A/2023A/8 (Resolving the chemical fingerprints of Milky Way mergers); and A/2023B/4 (s-process variations in southern globular clusters). We acknowledge the traditional owners of the land on which the AAT stands, the Gamilaraay people, and pay our respects to elders past and present. This paper includes data that has been provided by AAO Data Central (datacentral.org.au). I would also like to thank my mentor Dr. Tariq Yasin from the University of Oxford for his guidance throughout this project.

References

1 R. H. Cyburt and B. D. Fields and K. A. Olive and T.-H. Yeh, Big bang nucleosynthesis: Present status. Reviews of Modern Physics. Vol. 88, 2016., https://doi.org/10.1103/RevModPhys.88.015004.

2 C. Pitrou and A. Coc and J.-P. Uzan and E. Vangioni, Precision big bang nucleosynthesis with improved helium-4 predictions. Physics Reports. Vol. 754, pp. 1–66, 2018., https://doi.org/10.1016/j.physrep.2018.04.005.

3 G. Steigman, Primordial nucleosynthesis in the precision cosmology era. Annual Review of Nuclear and Particle Science. Vol. 57, pp. 463–491, 2007., https://doi.org/10.1146/annurev.nucl.56.080805.140437.

4 Planck Collaboration and N. Aghanim and Y. Akrami and M. Ashdown and J. Aumont and C. Baccigalupi and et al., Planck 2018 results. VI. Cosmological parameters. Astronomy & Astrophysics. Vol. 641, A6, 2020., https://doi.org/10.1051/0004-6361/201833910.

5 F. Spite and M. Spite, Abundance of lithium in unevolved halo stars and old disk stars: Interpretation and consequences. Astronomy & Astrophysics. Vol. 115, pp. 357–366, 1982.

6 L. Sbordone and P. Bonifacio and E. Caffau and H.-G. Ludwig and N. T. Behara and J. I. González Hernández and M. Steffen and R. Cayrel and B. Freytag and C. Van't Veer and P. Molaro and B. Plez and T. Sivarani and M. Spite and F. Spite and T. C. Beers and N. Christlieb and P. François and V. Hill, The metal-poor end of the Spite plateau: I. Stellar parameters, metallicities, and lithium abundances. Astronomy & Astrophysics. Vol. 522, A26, 2010., https://doi.org/10.1051/0004-6361/200913282.

7 B. D. Fields, The Primordial Lithium Problem. Annual Review of Nuclear and Particle Science. Vol. 61, pp. 47–68, 2011., https://doi.org/10.1146/annurev-nucl-102010-130445.

8 A. Frebel and J. E. Norris, Near-field cosmology with extremely metal-poor stars. Annual Review of Astronomy and Astrophysics. Vol. 53, pp. 631–688, 2015., https://doi.org/10.1146/annurev-astro-082214-122423.

9 P. Bonifacio and P. Molaro and T. Sivarani and R. Cayrel and M. Spite and F. Spite and B. Plez and J. Andersen and B. Barbuy and T. C. Beers and E. Depagne and V. Hill and P. François and B. Nordström and F. Primas, First stars VII: Lithium in extremely metal poor dwarfs. Astronomy & Astrophysics. Vol. 462, pp. 851–864, 2006., https://doi.org/10.1051/0004-6361:20064834.

10 O. Richard and G. Michaud and J. Richer, Implications of WMAP Observations on Li Abundance and Stellar Evolution Models. The Astrophysical Journal. Vol. 619, pp. 538–548, 2005., https://doi.org/10.1086/426470.

11 A. J. Korn and F. Grundahl and O. Richard and P. S. Barklem and L. Mashonkina and R. Collet and B. Gustafsson and N. Piskunov, Atomic diffusion and lithium abundances in metal-poor stars. The Astrophysical Journal. Vol. 671, pp. 402–419, 2007., https://doi.org/10.1086/523098.

12 A. J. Korn, How stars in globular clusters reveal the depletion of the Spite plateau of lithium. arXiv, 2021., https://arxiv.org/abs/2111.00913.

13 Gaia Collaboration and A. Vallenari and A. G. A. Brown and T. Prusti and J. H. J. de Bruijne and F. Arenou and C. Babusiaux and et al., Gaia Data Release 3: Summary of the content and survey properties. Astronomy & Astrophysics. Vol. 674, A1, 2023., https://doi.org/10.1051/0004-6361/202243940.

14 Gaia Collaboration, T. Prusti, J. H. J. de Bruijne, A. G. A. Brown, A. Vallenari, C. Babusiaux, C. A. L. Bailer-Jones, U. Bastian, M. Biermann, D. W. Evans, L. Eyer, F. Jansen, C. Jordi, S. A. Klioner, U. Lammers, L. Lindegren, X. Luri, M. Mignard, D. J. Milligan, C. Panem, V. Poinsignon, D. Pourbaix, S. Randich, G. Sarri, P. Sartoretti, H. I. Siddiqui, C. Soubiran, V. Valette, F. van Leeuwen, N. A. Walton et al.,

The Gaia Mission. Astronomy & Astrophysics. Vol. 595, A1, 2016., https://doi.org/10.1051/0004-6361/201629272.
15 S. Buder and J. Kos and E. X. Wang and M. McKenzie and M. Howell and S. L. Martell and M. R. Hayden and D. B. Zucker and T. Nordlander and B. T. Montet and G. Traven and J. Bland-Hawthorn and G. M. De Silva and K. C. Freeman and G. F. Lewis and K. Lind and S. Sharma and J. D. Simpson and D. Stello and T. Zwitter and The GALAH Collaboration, The GALAH Survey: Data Release 4. arXiv, 2025., https://arxiv.org/abs/2409.19858.
16 G. M. De Silva and K. C. Freeman and J. Bland-Hawthorn and S. Martell and E. W. de Boer and M. Asplund and S. Keller and S. Sharma and D. B. Zucker and T. Zwitter and et al., The GALAH survey: Scientific motivation. Monthly Notices of the Royal Astronomical Society. Vol. 449, pp. 2604–2617, 2015., https://doi.org/10.1093/mnras/stv327.
17 X. Gao and K. Lind and A. M. Amarsi and S. Buder and J. Bland-Hawthorn and S. W. Campbell and M. Asplund and A. R. Casey and G. M. De Silva and K. C. Freeman and et al., The GALAH survey: A new constraint on cosmological lithium and Galactic lithium evolution from warm dwarf stars. Monthly Notices of the Royal Astronomical Society: Letters. Vol. 497, pp. L30–L34, 2020., https://doi.org/10.1093/mnrasl/slaa109.
18 F. Anders and A. Khalatyan and A. B. A. Queiroz and C. Chiappini and J. Ardèvol and L. Casamiquela and F. Figueras and Ó. Jiménez-Arranz and C. Jordi and et al., Photo-astrometric distances, extinctions, and astrophysical parameters for Gaia EDR3 stars brighter than G = 18.5. Astronomy & Astrophysics. Vol. 658, A91, 2022., https://doi.org/10.1051/0004-6361/202142369.
19 M. Asplund and D. L. Lambert and P. E. Nissen and F. Primas and V. V. Smith, Lithium isotopic abundances in metal-poor halo stars. The Astrophysical Journal. Vol. 644, pp. 229–259, 2006., https://doi.org/10.1086/503538.
20 J. Meléndez and I. Ramírez, Reappraising the Spite Lithium Plateau: Extremely thin and marginally consistent with WMAP data. The Astrophysical Journal. Vol. 615, pp. L33–L36, 2004., https://doi.org/10.1086/425962.
21 J. Meléndez and N. G. Shchukina and I. Ramírez and I. Vasiljeva, The temperature scale of metal-poor dwarfs: lithium and oxygen abundances. Proceedings of the International Astronomical Union. Vol. 1, pp. 265–266, 2005., https://doi.org/10.1017/S1743921305005752.
22 D. J. Bartlett and H. Desmond, Marginalised normal regression: unbiased curve fitting in the presence of x-errors. Open Journal of Astrophysics. Vol. 6, 2023., https://doi.org/10.21105/astro.2309.00948.
23 J. E. Norris and D. Yong and A. Frebel and S. G. Ryan, A critique of the Spite Plateau and the astration of primordial lithium. Monthly Notices of the Royal Astronomical Society. Vol. 522, pp. 1358–1376, 2023., https://doi.org/10.1093/mnras/stad936.
24 W. Aoki and P. S. Barklem and T. C. Beers and N. Christlieb and S. Inoue and A. E. G. Pérez and J. E. Norris and D. Carollo, Lithium abundances of extremely metal-poor turnoff stars. The Astrophysical Journal. Vol. 698, pp. 1803–1812, 2009., https://doi.org/10.1088/0004-637X/698/2/1803.
25 B. D. Fields and K. A. Olive and T.-H. Yeh and C. Young, Big-bang nucleosynthesis after Planck. Journal of Cosmology and Astroparticle Physics. Vol. 2020, pp. 010–010, 2020., https://doi.org/10.1088/1475-7516/2020/03/010.
26 O. D. Miranda, The cosmological lithium problem. Astronomy & Astrophysics. Vol. 701, A164, 2025., https://doi.org/10.1051/0004-6361/202554482.
27 X. Fu and A. Bressan and P. Molaro and P. Marigo, Lithium evolution in metal-poor stars: From pre-main sequence to the Spite plateau. Monthly Notices of the Royal Astronomical Society. Vol. 452, pp. 3256–3265, 2015., https://doi.org/10.1093/mnras/stv1384.
28 S. G. Ryan and J. E. Norris and T. C. Beers, The Spite Lithium Plateau: Ultrathin but postprimordial. The Astrophysical Journal. Vol. 523, pp. 654–677, 1999., https://doi.org/10.1086/307769.

Supplementary Information

The online version contains supplementary material available at https://nhsjs.com/?p=45072